Change detection in multisensor SAR images using bivariate gamma distributions

Florent Chatelain[†], Jean-Yves Tourneret[†] and Jordi Inglada[*]

[†]IRIT/ENSEEIHT/TéSA, 2 rue Camichel, BP 7122, 31071 Toulouse cedex 7, France

[*]CNES, 18 avenue E. Belin, BPI 1219, 31401 Toulouse cedex 9, France

florent.chatelain@enseeiht.fr, jean-yves.tourneret@enseeiht.fr, jordi.inglada@cnes.fr

Abstract—This paper studies a family of distributions constructed from multivariate gamma distributions to model the statistical properties of multisensor synthetic aperture radar (SAR) images. These distributions referred to as multisensor multivariate gamma distributions (MuMGDs) are potentially interesting for detecting changes in SAR images acquired by different sensors having different numbers of looks. The first part of the paper compares different estimators for the parameters of MuMGDs. These estimators are based on the maximum likelihood principle, the method of inference function for margins and the method of moments. The second part of the paper studies change detection algorithms based on the estimated correlation coefficient of MuMGDs. Simulation results conducted on synthetic and real data illustrate the performance of these change detectors.

Index Terms—Multivariate gamma distributions, correlation coefficient, maximum likelihood, change detection.

EDICS Category: GEO-RADR

I. INTRODUCTION

Combining information acquired from multiple sensors has become very popular in many signal and image processing applications. In the case of earth observation applications, there are two reasons for that. The first one is that the fusion of the data produced by different types of sensors provides a complementarity which overcomes the limitations of a specific kind of sensor. The other reason is that, often, in operational applications, the user does not have the possibility to choose the data to work with and has to use the available archive images or the first acquisition available after an event of interest. This is particularly true for monitoring applications where image registration and change detection approaches have to be implemented on different types of data [1], [2].

Both image registration and change detection techniques consists of comparing two images I, the reference, and J, the secondary image, acquired over the same landscape – scene – at two different dates. Usually, the reference image is obtained from an archive and the acquisition of the secondary image is scheduled after an abrupt change, like a natural disaster. In the case of the change detection, the goal is producing an indicator of change for each pixel of the region of interest. This indicator of change is the result of applying locally a similarity measure to the two images. This similarity measure is usually

chosen as the correlation coefficient or other statistical feature in order to deal with noisy data.

The estimation of the similarity measure is performed locally for each pixel position. Since a statistical estimation has to be performed, and only one realization of the random variable is available, the images are supposed to be locally stationary and the ergodicity assumption allows to make estimates using several neighbor pixels. This neighborhood is the so-called estimation window. In order for the stationarity assumption to hold, this estimation window has to be small. On the other hand, robust statistical estimates need a high number of samples. Therefore, the key point of the estimation of the similarity measure is to perform high quality estimates with a small number of samples. One way to do so is to introduce a priori knowledge about the image statistics.

In the case of power radar images, it is well known that the pixels are marginally distributed according to gamma distributions [3]. Therefore, multivariate gamma distributions (having univariate gamma margins) seem good candidates for the robust estimation of the correlation coefficient between radar images. When multi-date power radar images are acquired from different sensors, the numbers of looks associated with the different images can be different. As the number of looks is the shape parameter of the gamma distribution, this leads to study multivariate gamma distributions whose margins have different shape parameters.

A family of multivariate gamma distributions has been recently defined by S. Bar Lev and P. Bernardoff [4], [5]. These distributions are defined from an appropriate moment generating function. Their margins are distributed according to univariate gamma distributions having the same shape parameter. They have recently shown interesting properties for registration and change detection in SAR images acquired by the same sensor (i.e. for images having the same number of looks) [6], [7]. This paper studies a new family of multivariate distributions whose margins are univariate gamma distributions with different shape parameters referred to as multisensor multivariate gamma distributions (MuMGDs). The application of MuBGDs to change detection in SAR images is also investigated.

This paper is organized as follows. Section 2 recalls important results on monosensor multivariate gamma distributions (MoMGDs). Section 3 defines the family of MuMGDs considered for change detection in multisensor SAR images. Section 4 studies the maximum likelihood estimator (MLE), the infer-

ence function for margins (IFM) estimator and the estimator of moments for the unknown parameters of MuMGDs. Section 5 presents some simulation results illustrating the performance of MuMGDs for parameter estimation and change detection on synthetic and real SAR images. Conclusions and perspectives are finally reported in Section 6.

II. MONOSENSOR MULTIVARIATE GAMMA DISTRIBUTIONS

A. Definition

A random vector $\boldsymbol{X} = (X_1, ..., X_d)^T$ is distributed according to an MoMGD on $\mathbb{R}_+^d$ with shape parameter q and scale parameter P if its moment generating function, or Laplace transform, is defined as [5]:

$$\psi_{q,P}(\boldsymbol{z}) = E\left(e^{-\sum_{i=1}^{d} X_i z_i}\right) = [P(\boldsymbol{z})]^{-q}, \qquad (1)$$

where $\boldsymbol{z} = (z_1, ..., z_d)$, $q \geq 0$ and $P(\boldsymbol{z})$ is a so-called affine polynomial[1]. The Laplace transform of X_i is obtained by setting $z_j = 0$ for $j \neq i$ in (1). This shows that X_i is distributed according to a univariate gamma distribution with shape parameter q and scale parameter p_i, denoted as $X_i \sim \mathcal{G}(q, p_i)$. Thus, all margins of $\boldsymbol{X}$ are univariate gamma distributions with the same shape parameter q.

A monosensor bivariate gamma distribution (MoBGD) corresponds to the particular case $d = 2$ and is defined by its affine polynomial

$$P(\boldsymbol{z}) = 1 + p_1 z_1 + p_2 z_2 + p_{12} z_1 z_2, \qquad (2)$$

with the following conditions

$$p_1 > 0, \ p_2 > 0, \ p_{12} > 0, \ p_1 p_2 - p_{12} \geq 0. \qquad (3)$$

It is important to note that the conditions (3) ensure that (2) is the Laplace transform of a probability distribution defined on $[0, \infty[^2$. However, in the general case ($d > 2$), determining necessary and sufficient conditions on P and q such that (1) is the Laplace transform of a probability distribution defined on $[0, \infty[^d$ is a difficult problem (see [5] for more details). The main properties of MoBGDs have been studied in [6]. Some important results required for the present paper are recalled below.

B. Moments

The moments of an MoBGD can be obtained by differentiating the Laplace transform (2). For instance, the mean and variance of X_i (denoted as $E[X_i]$ and $\text{var}(X_i)$ respectively) can be expressed as follows

$$\mathsf{E}[X_i] = qp_i, \ \text{var}(X_i) = qp_i^2, \qquad (4)$$

for $i = 1, 2$. Similarly, the covariance $\text{cov}(X_1, X_2)$ and correlation coefficient $r(X_1, X_2)$ of an MoBGD are:

$$\text{cov}(X_1, X_2) = \mathsf{E}[X_1 X_2] - \mathsf{E}[X_1]\mathsf{E}[X_2] = q(p_1 p_2 - p_{12}),$$

$$r(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{var}(X_1)}\sqrt{\text{var}(X_2)}} = \frac{p_1 p_2 - p_{12}}{p_1 p_2}. \qquad (5)$$

[1]A polynomial $P(\boldsymbol{z})$ where $\boldsymbol{z} = (z_1, \ldots, z_d)$ is *affine* if the one variable polynomial $z_j \mapsto P(\boldsymbol{z})$ can be written $Az_j + B$ (for any $j = 1, \ldots, d$), where A and B are polynomials with respect to the z_i's with $i \neq j$.

It is important to note that when $\text{cov}(X_1, X_2) = 0$ (or equivalently $p_{12} = p_1 p_2$) the Laplace transform of $\boldsymbol{X}$ can be factorized as follows:

$$\psi_{q,P}(z_1, z_2) = [1 + p_1 z_1 + p_2 z_2 + p_1 p_2 z_1 z_2]^q = [1 + p_1 z_1]^q [1 + p_2 z_2]^q$$

where the two factors in the right hand side are the Laplace transforms of X_1 and X_2. As a consequence, the random variables X_1 and X_2 of an MoBGD are independent if and only if they are uncorrelated (as in the Gaussian case).

C. Probability density function (pdf)

The pdf of an MoBGD can be expressed as follows (see [8, p. 436] for a similar result)

$$f_{2D}(\boldsymbol{x}) = \exp\left(-\frac{p_2 x_1 + p_1 x_2}{p_{12}}\right) \frac{x_1^{q-1} x_2^{q-1}}{p_{12}^q \Gamma(q)} f_q(cx_1 x_2) \mathbb{I}_{\mathbb{R}_+^2}(\boldsymbol{x}), \qquad (6)$$

where $\mathbb{I}_{\mathbb{R}_+^2}(\boldsymbol{x})$ is the indicator function on $[0, \infty[^2$ ($\mathbb{I}_{\mathbb{R}_+^2}(\boldsymbol{x}) = 1$ if $x_1 > 0, x_2 > 0$ and $\mathbb{I}_{\mathbb{R}_+^2}(\boldsymbol{x}) = 0$ otherwise), $c = (p_1 p_2 - p_{12})/p_{12}^2$ and $f_q(z)$ is related to the confluent hypergeometric function [8, p. 462] defined by

$$f_q(z) = \sum_{k=0}^{\infty} \frac{z^k}{k! \Gamma(q+k)}.$$

III. MULTISENSOR GAMMA DISTRIBUTIONS

A. Definition

A random vector $\boldsymbol{Y} = (Y_1, ..., Y_d)^T$ is distributed according to a multisensor multivariate gamma distribution (MuMGD) with scale parameter P and shape parameter $\boldsymbol{q} = (q_1, \ldots, q_d)$, denoted as $\boldsymbol{X} \sim \mathcal{G}(\boldsymbol{q}, P)$, if it can be constructed as follows:

$$\begin{aligned} Y_1 &= X_1, \\ Y_i &= X_i + Z_i, \qquad 2 \leq i \leq d. \end{aligned} \qquad (7)$$

where

- $\boldsymbol{X} = (X_1, ..., X_d)^T$ is a random vector distributed according to an MoMGD on $\mathbb{R}_+^d$ with shape parameter q_1 and scale parameter P, i.e. $\boldsymbol{X} \sim \mathcal{G}(q_1, P)$,
- $Z_1, ..., Z_d$ are independent random variables distributed according to univariate gamma distributions (with the convention $Z_i = 0$ when $q_i - q_1 = 0$) $Z_i \sim \mathcal{G}(q_i - q_1, p_i)$ with $q_i \geq q_1$.
- The vector $\boldsymbol{Z} = (Z_2, ..., Z_d)^T$ is independent on $\boldsymbol{X}$.

By using the independence property between $\boldsymbol{X}$ and $\boldsymbol{Z}$, the Laplace transform of $\boldsymbol{Y}$ can be written:

$$\begin{aligned} \psi_{\mathcal{G}(\boldsymbol{q}, P)}(\boldsymbol{z}) &= E\left(e^{-\sum_{i=1}^{d} Y_i z_i}\right) \\ &= [P(\boldsymbol{z})]^{-q_1} \prod_{i=1}^{d} (1 - p_i z_i)^{-(q_i - q_1)}. \end{aligned} \qquad (8)$$

By setting $z_j = 0$ for $j \neq i$ in (9), we observe that the random variable Y_i is distributed according to a univariate gamma distribution with scale parameter p_i and shape parameter q_i, i.e. $Y_i \sim \mathcal{G}(q_i, p_i)$. Thus, all margins of $\boldsymbol{Y}$ have different shape parameters in the general case. Note that the definition above assumes that the first univariate margin Y_1 has a shape parameter q_1 smaller than all other shape parameters $q_i, i \geq 2$

without loss of generality. Note also that an MuMGD reduces to an MoMGD for $q_i = q_1, \forall i \geq 2$.

A multisensor bivariate gamma distribution (MuBGD) corresponds to the particular case $d = 2$ and is defined by its Laplace transform

$$\psi(\boldsymbol{z}) = \left(1 + \sum_{i=1}^{2} p_i z_i + p_{12} z_1 z_2\right)^{-q_1} (1 - p_2 z_2)^{-(q_2 - q_1)}, \quad (9)$$

with the following conditions:

$$p_1 > 0, \ p_2 > 0, \ p_{12} > 0, \ p_1 p_2 - p_{12} \geq 0 \ \text{ and } q_2 \geq q_1. \quad (10)$$

In the bi-dimensional case, the conditions (10) ensure that (9) is the Laplace transform of a probability distribution defined on $[0, \infty[^2$.

B. MuBGD pdf

According to (7), a vector $\boldsymbol{Y} = (Y_1, Y_2)^T$ distributed according to an MuBGD (i.e. $\boldsymbol{Y} \sim \mathcal{G}(\boldsymbol{q}, P)$) is constructed from a random vector $\boldsymbol{X} = (X_1, X_2)^T$ distributed according to an MoBGD whose pdf is denoted as $f_{\boldsymbol{X}}(\boldsymbol{x})$ and a random variable $Z \sim \mathcal{G}(q_2 - q_1, p_2)$ independent on $\boldsymbol{X}$ with pdf $f_Z(z)$. By using the independence assumption between $\boldsymbol{X}$ and Z, the density of $\boldsymbol{Y}$ can be expressed as

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \int f_{\boldsymbol{X}}(y_1, s) f_Z(y_2 - s) ds. \quad (11)$$

Straightforward computations leads to the following expression:

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \left(\frac{p_1 p_2}{p_{12}}\right)^{q_1} \frac{y_1^{q_1-1} y_2^{q_2-1}}{p_1^{q_1} p_2^{q_2}} \frac{e^{-\left(\frac{p_2}{p_{12}} y_1 + \frac{p_1}{p_{12}} y_2\right)}}{\Gamma(q_2) \Gamma(q_1)} \times \\ \Phi_3\left(q_2 - q_1; q_2; c\frac{p_{12}}{p_2} y_2, c y_1 y_2\right), \quad (12)$$

where $c = (p_1 p_2 - p_{12})/p_{12}^2$ and where Φ_3 is the so-called Horn function. The Horn function is one of the twenty convergent confluent hypergeometric series of order two, defined as [9]:

$$\Phi_3(a; b; x, y) = \sum_{m,n=0}^{\infty} \frac{(a)_m}{(b)_{m+n} m! n!} x^m y^n, \quad (13)$$

where $(a)_m$ is the Pochhammer symbol such that $(a)_0 = 1$ and $(a)_{k+1} = (a + k)(a)_k$ for any positive integer k. It is interesting to note that the relation $f_q(c y_1 y_2) = \Phi_3\left(0; q; c\frac{p_{12}}{p_2} y_2, c y_1 y_2\right) / \Gamma(q)$ allows one to show that the MuBGD pdf defined in (13) reduces to the MoBGD pdf (6) for $q_1 = q_2 = q$.

C. MuBGD moments

The moments of $\boldsymbol{Y}$ can clearly be obtained from the moments of $\boldsymbol{X}$ and $\boldsymbol{Z}$. This section concentrates on MuBGDs defined by $\boldsymbol{Y} = (X_1, X_2 + Z)^T$, where $\boldsymbol{X} = (X_1, X_2)^T$ is an MoBGD with mean (m_1, m_2), correlation coefficient r' and shape parameter q_1, and Z is a univariate gamma distribution with mean m_2 and shape parameter $q_2 - q_1$. Using

the independence property between $\boldsymbol{X}$ and $\boldsymbol{Z}$, the following results can be obtained:

$$E\left[Y_1^m Y_2^n\right] = \sum_{i=1}^{n} \binom{n}{i} E\left[X_1^m X_2^i\right] E\left[Z^{n-i}\right], \\ = \sum_{i=1}^{n} \binom{n}{i} m_2^{n-i} \frac{(q_2 - q_1)_{n-i}}{(q_2 - q_1)^{n-i}} E\left[X_1^m X_2^i\right], \quad (14)$$

for all $(m, n) \in \mathbb{N}^2$. The moments of an MoBGD were derived in [6]:

$$E[X_1^{n_1} X_2^{n_2}] = m_1^{n_1} m_2^{n_2} \frac{(q_1)_{n_1}}{q_1^{n_1}} \frac{(q_1)_{n_2}}{q_1^{n_2}} \times \\ \sum_{k=0}^{\min(n_1, n_2)} \frac{(-n_1)_k (-n_2)_k}{(q_1)_k} \frac{(r')^k}{k!}, \quad (15)$$

for all $(n_1, n_2) \in \mathbb{N}^2$. Expressions (14) and (15) can be used to derive analytical expressions of MuMGD moments. For instance, the first and second order moments can be written as

$$E[Y_i] = q_i p_i, \ \text{var}(Y_i) = q_i p_i^2, \quad i = 1, 2 \\ \text{cov}(Y_1, Y_2) = \text{cov}(X_1, X_2) = q_1 (p_1 p_2 - p_{12}), \\ r(Y_1, Y_2) = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{var}(Y_1)} \sqrt{\text{var}(Y_1)}} = \frac{q_1}{\sqrt{q_1 q_2}} \frac{p_1 p_2 - p_{12}}{p_1 p_2}.$$

It is interesting to note that the conditions (3) ensure that the correlation coefficient satisfy the constraint $0 \leq r(Y_1, Y_2) \leq \sqrt{q_1/q_2}$. In other words, the normalized correlation coefficient defined by

$$r'(Y_1, Y_2) = \sqrt{\frac{q_2}{q_1}} \ r(Y_1, Y_2) = \frac{p_1 p_2 - p_{12}}{p_1 p_2},$$

is such that $0 \leq r'(Y_1, Y_2) \leq 1$. As explained in II-B, the random variables X_1 and X_2 are independent if and only if $p_1 p_2 - p_{12} = 0$. Since Z_2 is independent from X_1 and X_2, a necessary and sufficient condition for the margins of an MuBGD Y_1 and Y_2 to be independent is $r'(Y_1, Y_2) = 0$. Note finally that for known values of the shape parameters q_1 and q_2, an MuBGD is fully characterized by the parameter vector $\boldsymbol{\theta} = (E[Y_1], E[Y_2], r'(Y_1, Y_2))$, since $\boldsymbol{\theta}$ and (p_1, p_2, p_{12}) are related by a one-to-one transformation.

IV. PARAMETER ESTIMATION FOR MuBGDs

This section studies different methods for estimating the parameters of MuBGDs.[2] The following notations are used in the rest of the paper

$$m_1 = E[Y_1], \ m_2 = E[Y_2], \ r' = r(Y_1, Y_2) \sqrt{\frac{q_2}{q_1}},$$

inducing $\boldsymbol{\theta} = (m_1, m_2, r')^T$. Note that the parameters p_1, p_2 and p_{12} can be expressed as functions of $\boldsymbol{\theta}$ as follows $p_1 = \frac{m_1}{q_1}, p_2 = \frac{m_2}{q_2}$ and $p_{12} = \frac{m_1 m_2}{q_1 q_2}(1 - r')$. Note also that the parameters q_1 and q_2 are assumed to be known in this paper, as in most practical applications. In the case where q_1 and q_2

---

[2]The results proposed here could be used to estimate the parameters of MuMGDs by using the concept of composite likelihood. The interested reader is invited to consult [10], [11] and references therein for more details.

are unknown, these parameters should be included in $\boldsymbol{\theta}$ and estimated jointly with m_1, m_2 and r'[3].

A. Maximum Likelihood (ML) Method

1) Principles: The ML method can be applied to $\boldsymbol{Y}$ since a closed-form expression of its pdf is available. After removing the terms which do not depend on $\boldsymbol{\theta}$, the log-likelihood function of $\boldsymbol{Y}$ can be written

$$l(\boldsymbol{Y}; \boldsymbol{\theta}) = - nq_1 \log (1 - r') - nq_1 \log m_1 - nq_2 \log m_2$$
$$- n\frac{q_1}{m_1(1-r')}\overline{Y}_1 - n\frac{q_2}{m_2(1-r')}\overline{Y}_2 \qquad (16)$$
$$+ \sum_{i=1}^{n} \log \Phi_3 \left(q_2 - q_1; q_2; dY_2^i, cY_1^iY_2^i\right),$$

where $d = \frac{r'q_2}{m_2(1-r')}$, $\overline{Y}_1 = \frac{1}{n}\sum_{i=1}^{n}Y_1^i$, $\overline{Y}_2 = \frac{1}{n}\sum_{i=1}^{n}Y_2^i$ are the sample means of Y_1 and Y_2 and c defined previously can be expressed as function of $\boldsymbol{\theta}$ using the relation $c = \frac{r'q_1q_2}{m_1m_2(1-r')^2}$. By differentiating the log-likelihood with respect to (wrt) $\boldsymbol{\theta}$, the MLE of $\boldsymbol{\theta}$ is obtained as a solution of:

$$\boldsymbol{u}(\boldsymbol{Y}; \boldsymbol{\theta}) = \left(\frac{\partial l(\boldsymbol{Y}; \boldsymbol{\theta})}{\partial m_1}, \frac{\partial l(\boldsymbol{Y}; \boldsymbol{\theta})}{\partial m_2}, \frac{\partial l(\boldsymbol{Y}; \boldsymbol{\theta})}{\partial r'}\right)^T = \boldsymbol{0}^T,$$

where $\boldsymbol{u}(\boldsymbol{Y}; \boldsymbol{\theta}) = \dfrac{\partial l(\boldsymbol{Y}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ is the so-called score function, or equivalently by solving

$$- nq_1 + \frac{nq_1}{m_1(1-r')}\overline{Y}_1 - r'\Delta_2 = 0, \qquad (17)$$

$$- nq_2 + \frac{nq_2}{m_2(1-r')}\overline{Y}_2 - r'(\Delta_1 + \Delta_2) = 0, \qquad (18)$$

$$nq_1 - \sum_{i=1}^{2}\frac{nq_i}{m_i(1-r')}\overline{Y}_i + \Delta_1 + (1+r')\Delta_2 = 0. \qquad (19)$$

with

$$\Delta_1 = \frac{q_2 - q_1}{m_2(1-r')}\sum_{i=1}^{n}Y_2^i \frac{\Phi_3(q_2 - q_1 + 1; q_2 + 1; dY_2^i; cY_1^iY_2^i)}{\Phi_3(q_2 - q_1; q_2; dY_2^i; cY_1^iY_2^i)},$$

$$\Delta_2 = \frac{q_1(1-r')^{-2}}{m_1m_2}\sum_{i=1}^{n}Y_1^iY_2^i \frac{\Phi_3(q_2 - q_1 + 1; q_2 + 1; dY_2^i; cY_1^iY_2^i)}{\Phi_3(q_2 - q_1; q_2; dY_2^i; cY_1^iY_2^i)}.$$

The MLE of m_2 can be obtained by summing (17), (18), (19) and replacing the value of $\Delta_1 + \Delta_2$ in (18):

$$\widehat{m}_{2\text{ML}} = \overline{Y}_2. \qquad (20)$$

The MLEs of m_1 and r' are obtained by replacing m_2 by $\widehat{m}_{2\text{ML}}$ in (16) and by maximizing the resulting log-likelihood $l(\boldsymbol{Y}; (m_1, \widehat{m}_{2\text{ML}}, r'))$ wrt m_1 and r'. This last maximization is achieved by using a constrained ($m_1 > 0$ and $r' \in [0,1]$) quasi-Newton method, since an analytical expression of the log-likelihood gradient is available[4]. Some elements regarding

---

[3]The interested reader is invited to consult [12] for a related example where the shape parameter of a mono sensor multivariate gamma distribution ($q_1 = q_2 = q$) was estimated from mixed Poisson data. This section addresses the problem of estimating the unknown parameter vector $\boldsymbol{\theta}$ from n vectors $\boldsymbol{Y} = (\boldsymbol{Y}^1, \ldots, \boldsymbol{Y}^n)$, where $\boldsymbol{Y}^i = (Y_1^i, Y_2^i)$ is distributed according to an MuBGD with parameter vector $\boldsymbol{\theta}$

[4]The negative log-likelihood function has a unique minimum with repect to r' in all pratical cases. The reador is invited to consult [13] for discussions and simulations results.

the numerical evaluation of the Horn Function are detailed in appendices I and II. It is important to note that the MLE of m_1 differs from $\overline{Y}_1$ in the general case[5]. Finally, the MLE of the correlation coefficient r is deduced by functional invariance as

$$\widehat{r}_{\text{ML}} = \sqrt{\frac{q_1}{q_2}}\, \widehat{r}'_{\text{ML}}.$$

2) Performance: The properties of the ML estimator $\widehat{m}_{2\text{ML}}$ can be easily derived from the properties of the univariate gamma distribution $\mathcal{G}(q_2, p_2)$. This estimator is obviously unbiased, convergent and efficient. However, the performance of $\widehat{m}_{1\text{ML}}$ and $\widehat{r}_{\text{ML}}$ are more difficult to study. Of course, the MLE is known to be asymptotically unbiased and asymptotically efficient, under mild regularity conditions. Thus, the mean square error (MSE) of the estimates can be approximated for large data records by the Cramer-Rao lower bound (CRLB). For unbiased estimators, the CRLB is obtained by inverting the following Fisher information matrix I:

$$I(\boldsymbol{\theta}) = -E\left[\frac{\partial \boldsymbol{u}(\boldsymbol{Y}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right].$$

Thus, the computation of I requires to determine the negative expectations of second-order derivatives of $l(\boldsymbol{Y}; \boldsymbol{\theta})$ wrt m_1, m_2 and r in (16). Closed-form expressions for the elements of I are difficult to obtain because of the term $\log \Phi_3$. In such situation, it is very usual to approximate the expectations by using Monte Carlo methods. This will provide interesting approximations of the ML MSEs (see simulation results of Section V).

B. Inference function for margins (IFM)

1) Principles: IFM is a two-stage estimation method whose main ideas can be found for instance in [14, Chapter 10] and are summarized below in the context of MuBGDs:

- estimate the unknown parameters m_1 and m_2 from the marginal distributions of Y_1 and Y_2. This estimation is conducted by maximizing the marginal likelihoods $l(Y_1; m_1)$ and $l(Y_2; m_2)$ wrt m_1 and m_2 respectively,
- estimate the parameter r' by maximizing the joint likelihood $l(\boldsymbol{Y}; \widehat{m}_{1\text{ML}}, \widehat{m}_{2\text{ML}}, r')$ wrt r'. Note that the parameters m_1 and m_2 have been replaced in the joint likelihood by their estimates resulting from the first stage of IFM.

The IFM procedure is often computationally simpler than the ML method which estimates all the parameters simultaneously from the joint likelihood. Indeed, a numerical optimization with several parameters is much more time-consuming compared with several optimizations with fewer parameters. The marginal distributions of an MuBGD are univariate gamma distributions with shape parameters q_i and means m_i, for $i = \{1, 2\}$. Thus, the IFM estimators of m_1, m_2, r' are obtained as a solution of:

$$\boldsymbol{g}(\boldsymbol{Y}; \boldsymbol{\theta}) = \left(\frac{\partial l_1(\boldsymbol{Y_1}; m_1)}{\partial m_1}, \frac{\partial l_2(\boldsymbol{Y_2}; m_2)}{\partial m_2}, \frac{\partial l(\boldsymbol{Y}; \boldsymbol{\theta})}{\partial r'}\right)^T = \boldsymbol{0}^T$$

---

[5]There is no closed-form expression for the MLE of m_1 contrarily to m_2. Indeed, there is some kind of dissymmetry between Y_1 and Y_2 inherent to the proposed model (7). This dissymmetry will disappear in the method based on the inference for margins studied in section B.

where l_i is the marginal log-likelihood function associated to the univariate random variable Y_i, for $i = \{1, 2\}$, and l is the joint log-likelihood defined in (16). The IFM estimators of m_1 and m_2 are classically obtained from the properties of the univariate gamma distribution:

$$\widehat{m_1}_{\text{IFM}} = \overline{Y}_1, \qquad \widehat{m_2}_{\text{IFM}} = \overline{Y}_2. \tag{21}$$

The IFM estimator of r' is obtained by replacing m_1 and m_2 by $\overline{Y}_1$ and $\overline{Y}_2$ in (16) and by minimizing the resulting log-likelihood $l(\boldsymbol{Y}; \overline{Y}_1, \overline{Y}_2, r')$ wrt r'. This last minimization is achieved by using a constrained quasi-Newton method (with the constraint $r' \in [0, 1]$), since an analytical expression of the log-likelihood gradient is available.

Note that the ML method presented before requires to optimize the log-likelihood $l(\boldsymbol{Y}; (m_1, \widehat{m_2}_{\text{ML}}, r')$ wrt m_1 and r' whereas the IFM method only requires to optimize $l(\boldsymbol{Y}; \overline{Y}_1, \overline{Y}_2, r')$ wrt a single variable r'. The optimization procedure is therefore much less time-consuming for IFM than for the ML method. Note also that the estimator of m_2 is the same for the ML and IFM methods. Finally, it is interesting to point out that the joint likelihood is the product of univariate gamma pdfs when $r' = 0$. As a consequence, the ML and IFM estimators are the same when $r' = 0$.

2) Performance: Asymptotic properties of the IFM estimator can be derived from the set of inferences functions $\boldsymbol{g}(\boldsymbol{Y}; \boldsymbol{\theta})$ under the usual regularity conditions for the MLE (the interested reader is invited to consult [14] for more details). In particular, the IFM estimator of $\boldsymbol{\theta}$ denoted as $\widehat{\boldsymbol{\theta}}_{\text{IFM}}$ is such that $\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{\text{IFM}} - \boldsymbol{\theta}\right)$ converges in distribution to the normal distribution $\mathcal{N}(\boldsymbol{0}, V)$, where the asymptotic covariance matrix V is the inverse Godambe information matrix defined as:

$$V = D_g^{-1} M_g D_g^{-T} \tag{22}$$

where

$$D_g = \mathsf{E}\left[\partial \boldsymbol{g}(\boldsymbol{Y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta}\right], \qquad M_g = \mathsf{E}\left[\boldsymbol{g}(\boldsymbol{Y}; \boldsymbol{\theta}) \boldsymbol{g}^T(\boldsymbol{Y}; \boldsymbol{\theta})\right].$$

Straightforward computations yield the following expressions for matrices D_g and M_g [15]:

$$D_g = \begin{pmatrix} \mathcal{J}_{11} & 0 & 0 \\ 0 & \mathcal{J}_{22} & 0 \\ \mathcal{I}_{13} & \mathcal{I}_{23} & \mathcal{I}_{33} \end{pmatrix}, \quad M_g = \begin{pmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} & 0 \\ \mathcal{J}_{12} & \mathcal{J}_{22} & 0 \\ 0 & 0 & \mathcal{I}_{33} \end{pmatrix}$$

where

- $\mathcal{I}_{ij}$ are the entries of the Fisher information matrix, $I = (\mathcal{I}_{ij})_{1 \le i,j \le 3}$,
- $\mathcal{J}_{11}$ and $\mathcal{J}_{22}$ are the Fisher information associated with the margins Y_1 and Y_2 respectively,
- $\mathcal{J}_{12} = \mathsf{E}\left[g_1(\boldsymbol{Y}; \boldsymbol{\theta}) g_2(\boldsymbol{Y}; \boldsymbol{\theta})\right] = \mathsf{E}\left[\frac{\partial l_1(\boldsymbol{Y_1}; m_1)}{\partial m_1} \frac{\partial l_2(\boldsymbol{Y_2}; m_2)}{\partial m_2}\right]$.

The terms $\mathcal{J}_{11}, \mathcal{J}_{22}, \mathcal{J}_{12}$ associated to MuBGDs are easily derived by considering the univariate log-likelihoods $l_1(\boldsymbol{Y_1}; m_1)$ and $l_2(\boldsymbol{Y_2}; m_2)$:

$$\mathcal{J}_{11} = \frac{q_1}{m_1^2}, \qquad \mathcal{J}_{22} = \frac{q_2}{m_2^2}, \qquad \mathcal{J}_{12} = \frac{q_1}{m_1 m_2} r'.$$

As explained in IV-A.2, the Fisher information entries I_{ij} do not have closed-form expressions. Consequently, these terms have been computed by using numerical integration (Simpson quadrature). Note that this method allows one to control the approximation error.

C. Method of Moments

The estimators of (m_1, m_2, r) derived in this paper will be compared to the standard estimators based on the method of moments

$$\widehat{m}_{1\text{Mo}} = \overline{X}_1, \quad \widehat{m}_{2\text{Mo}} = \overline{X}_2, \tag{23}$$

$$\widehat{r}_{\text{Mo}} = \frac{\sum_{i=1}^n (X_1^i - \overline{X}_1)(X_2^i - \overline{X}_2)}{\sqrt{\sum_{i=1}^n (X_1^i - \overline{X}_1)^2} \sqrt{\sum_{i=1}^n (X_2^i - \overline{X}_2)^2}}. \tag{24}$$

The asymptotic performance of the estimator $\hat{\boldsymbol{\theta}}_{\text{Mo}} = (\widehat{m}_{1\text{Mo}}, \widehat{m}_{2\text{Mo}}, \widehat{r}_{\text{Mo}})^T$ can be derived by imitating the results of [16] derived in the context of time series analysis. More precisely, the moment estimator of $\boldsymbol{\theta}$ can be rewritten as

$$\widehat{\boldsymbol{\theta}}_{\text{Mo}} = g\left(\boldsymbol{s}_n\right) = \left(s_n^1, s_n^2, \frac{s_n^5 - s_n^1 s_n^2}{\sqrt{s_n^3 - (s_n^1)^2}\sqrt{s_n^4 - (s_n^2)^2}}\right)^T,$$

where $\boldsymbol{s}_n = \left(s_n^1, \ldots, s_n^5\right)^T = \left[\frac{1}{n}\sum_{i=1}^n Y_1^i, \frac{1}{n}\sum_{i=1}^n Y_2^i, \frac{1}{n}\sum_{i=1}^n (Y_1^i)^2, \frac{1}{n}\sum_{i=1}^n (Y_2^i)^2, \frac{1}{n}\sum_{i=1}^n Y_1^i Y_2^i\right]^T$ contains the appropriate first and second order empirical moments of $\boldsymbol{Y} = (Y_1, Y_2)^T$. By denoting as $\Sigma(\boldsymbol{\theta}) = n \operatorname{cov}(\boldsymbol{s}_n)$ the covariance matrix of the vector $\sqrt{n}\boldsymbol{s}_n$ and $G(\boldsymbol{\theta})$ the jacobian of the function g defined above, it can be shown that the asymptotic covariance matrix of $\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{\text{Mo}} - \boldsymbol{\theta}\right)$ is $G(\boldsymbol{\theta})\Sigma(\boldsymbol{\theta})G(\boldsymbol{\theta})^T$ [16]. The determination of the covariance matrix $\Sigma(\boldsymbol{\theta})$ requires to know appropriate theoretical moments of $\boldsymbol{Y} = (Y_1, Y_2)^T$ (up to the fourth order). These moments can be determined by using the results of section III-C. The reader is invited to consult [6] for more details regarding the asymptotic performance of the moment estimator $\hat{\boldsymbol{\theta}}_{\text{Mo}}$ for MuBGDs.

V. SIMULATION RESULTS

Many simulations have been conducted to validate the previous theoretical results. This section presents some experiments obtained with a vector $\boldsymbol{Y} = (Y_1, Y_2)^T$ distributed according to an MuBGD whose Laplace transform is (9).

A. Generation of synthetic data

According to the definition given in Section III-A, a vector $\boldsymbol{Y}$ distributed according to an MuBGD can be generated by adding a random variable Z distributed according to a univariate gamma distribution to a random vector $\boldsymbol{X}$ distributed according to an MoBGD. The generation of a vector $\boldsymbol{X}$ whose Laplace transform is (1) has been described in [6] and is summarized below:

- simulate $2q$ independent multivariate Gaussian vectors of $\mathbb{R}^2$ denoted as $Z^1, \ldots, Z^{2q}$ with means $(0, 0)$ and the 2×2 covariance matrix $C = (c_{i,j})_{1 \le i,j \le 2}$ with $c_{i,j} = r^{\frac{|i-j|}{2}}$,
- compute the kth component of $\boldsymbol{X} = (X_1, X_2)$ as $X_k = \frac{m_k}{2q}\sum_{1 \le i \le 2q}(Z_k^i)^2$, Z_k^i being the kth component of Z^i.

It is interesting to note that the generation of a random vector distributed according to a multivariate gamma distribution is straightforward here since $2q$ is an integer (this assumption is not a problem in practical applications since q is the number of looks of the SAR image). However, if $2q$ wouldn't be an integer, the generation of the random vector $\boldsymbol{X}$ could be achieved by using an accept-reject procedure such as the one detailed in [17, p. 51].

B. Estimation performance

1) ML method and method of moments: The first simulations compare the performance of the estimators based on the method of moments and the ML method as a function of the sample size n. Note that the possible values of n correspond to the numbers of pixels of squared windows of size $(2p+1) \times (2p+1)$, where $p \in \mathbb{N}$. These values are appropriate to the change detection problem. The number of Monte Carlo runs is 10000 for all figures presented in this section. The other parameters for this example are $m_1 = 100$, $m_2 = 100$, $q_1 = 1$ (number of looks of the first image) and $q_2 = 2$ (number of looks of the second image). Figures 1(a), 1(b) and 1(c) show the MSEs of the estimated normalized correlation coefficient for different values of r' ($r' = 0.2$, $r' = 0.5$ and $r' = 0.8$). The losange curves correspond to the
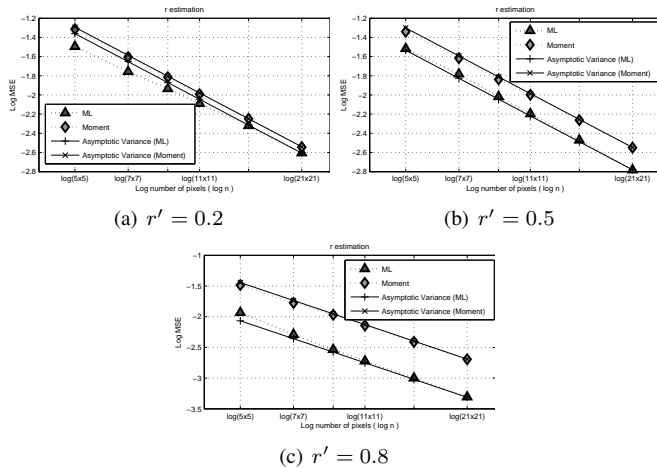

Fig. 1. log MSEs versus log n for parameter r ($q_1 = 1$, $q_2 = 2$, $m_1 = 100$ and $m_2 = 100$).

estimator of moments whereas the triangle curves correspond to the MLE. This figure shows the interest of the ML method, which is much more efficient for this problem than the method of moments, particularly for large values of the correlation coefficient r'. Note that the theoretical asymptotic MSEs of both estimators are also depicted (continuous lines). They are clearly in good agreement with the estimated MSEs, even for small values of n. Finally, these figures show that "reliable" estimates of r' can be obtained for values of n larger than 9×9, i.e. even for relatively small window sizes.

Figures 2(a) and 2(b) compare the MSEs of the estimated mean m_1 obtained for the ML method and the method of moments for two values of r' ($r' = 0.8$ and $r' = 0.9$). Both estimators perform very similarly for this parameter, even if the difference is slightly more noticeable for larger values of
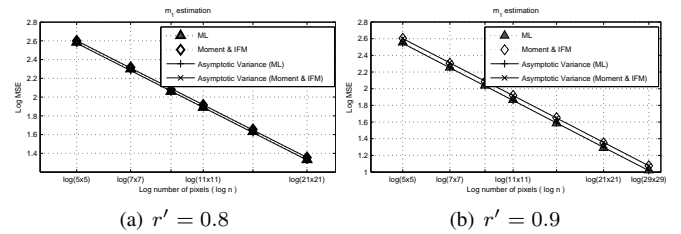

Fig. 2. log MSEs versus log n for parameter m_1 ($q_1 = 1$, $q_2 = 2$, $m_1 = 100$ and $m_2 = 100$).

r'. Note that the estimators of m_2 obtained for the ML and moment methods are the same. Thus, the corresponding MSEs have not been presented here for brevity.

2) ML and IFM: This section compares the performance of the ML and IFM estimators for the parameters r and m_1. Figure 3 first shows the asymptotic performance of both estimators by depicting the ratio of their asymptotic variances, referred to as asymptotic ratio efficiency (ARE), as a function of r. This figure shows that the ML and IFM estimators of the correlation coefficient r have very similar asymptotic variances when r' is not too close from 1. This result is
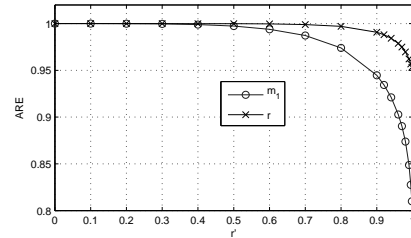

Fig. 3. Asymptotic Ratio Efficiency (ARE) ($q_1 = 1$, $q_2 = 5$, $m_1 = 1$ and $m_2 = 1$).

confirmed in Fig. 4 which shows the MSEs of the estimated correlation coefficient obtained with the ML and IFM methods for different values of the sample size n (the parameters for this simulation are $q_1 = 1$, $q_2 = 2$, $m_1 = 100$, $m_2 = 100$ and $r' = 0.9$). Figure 3 also shows that the asymptotic
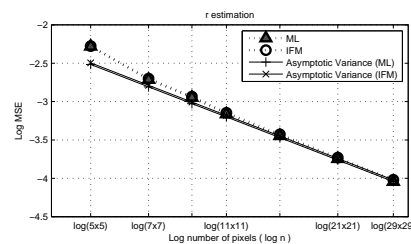

Fig. 4. log MSEs versus log n for parameter r ($r' = 0.9$, $q_1 = 1$, $q_2 = 2$, $m_1 = 100$ and $m_2 = 100$).

performance of the ML and IFM estimators for parameter m_1 differ significantly when r' approaches 1. However, this is not a major problem since the change detection algorithms proposed in this paper will be based on r only (see next section). Based on these results, the IFM method will be preferred to the ML method since it involves much smaller computational cost.

C. Detection performance

This section considers synthetic vectors $\boldsymbol{x} = (x_1, x_2)^T$ (coming from 762×292 synthetic images) distributed according to MuBGDs with $r = 0.3$ and $r = 0.7$, modeling the presence and absence of changes, respectively. The correlation coefficient r of each bivariate vector $\boldsymbol{x}^{(i,j)} = (x_1^{(i,j)}, x_2^{(i,j)})^T$ (for $1 \leq i \leq 762$, $1 \leq j \leq 292$) is estimated locally from pixels belonging to windows of size $n = (2p+1) \times (2p+1)$ centered around the pixel of coordinates (i, j) in the two analyzed images. The change detection problem in multisensor SAR images is addressed by using the following decision rule:

$$\begin{aligned} &\text{Decide } H_0 \text{ (absence of change) if } \widehat{r} > \lambda, \\ &\text{Decide } H_1 \text{ (presence of change) if } \widehat{r} \leq \lambda, \end{aligned} \quad (25)$$

where λ is a threshold depending on the probability of false alarm (PFA) and $\widehat{r}$ is an estimator of the correlation coefficient (obtained from the method of moments or the IFM method). The performance of the change detection strategy (25) can be defined by the two following probabilities [18, p. 34]

$$\begin{aligned} P_{\mathrm{D}} &= P\left[\text{accepting } H_1 \,|H_1 \text{ is true}\right] = P\left[\widehat{r} < \lambda \,|H_1 \text{ is true}\right], \\ P_{\mathrm{FA}} &= P\left[\text{accepting } H_1 \,|H_0 \text{ is true}\right] = P\left[\widehat{r} < \lambda \,|H_0 \text{ is true}\right]. \end{aligned}$$

Thus, a pair $(P_{\mathrm{FA}}, P_{\mathrm{D}})$ can be defined for each value of λ. The curves representing P_{D} as a function of P_{FA} are called receiver operating characteristics (ROCs) and are classically used to assess detection performance [18, p. 38].

The ROCs for the change detection problem (25) are depicted on Figs 5(a), 5(b) and 5(c) for three representative values of (q_1, q_2) and two window sizes (9×9) and (21×21). The IFM estimator clearly outperforms the moment estimator

(a) $(q_1, q_2) = (1, 2)$
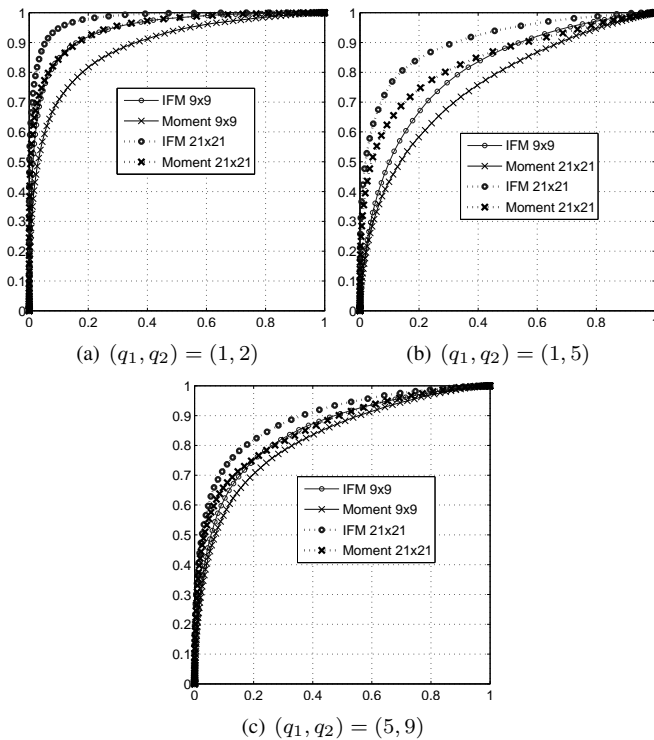
(b) $(q_1, q_2) = (1, 5)$

(c) $(q_1, q_2) = (5, 9)$

Fig. 5. ROCs for synthetic data.

for these examples. Figures 5(a) and 5(b) also show that

the detection performance seems to decrease when $q_2 - q_1$ increases, i. e. when the difference between the numbers of looks of the two images increases. In order to confirm this observation, we have derived theoretical ROCs by using the asymptotic Gaussian distribution for the estimated correlation coefficient (see section IV, B. 2). In this case, by denoting r_0 and r_1 the true values of the correlation coefficient under hypotheses H_0 and H_1, the following results can be obtained:

$$\begin{aligned} P_{\mathrm{D}} &= P\left[\widehat{r} < \lambda \,|H_1 \text{ is true}\right] = P\left[\widehat{r} < \lambda \,|\widehat{r} \sim \mathcal{N}\left(r_1, \sigma_1^2\right)\right], \\ P_{\mathrm{FA}} &= P\left[\widehat{r} < \lambda \,|H_0 \text{ is true}\right] = P\left[\widehat{r} < \lambda \,|\widehat{r} \sim \mathcal{N}\left(r_0, \sigma_0^2\right)\right], \end{aligned}$$

where σ_0^2 and σ_1^2 are the asymptotic variances of the estimated correlation coefficient $\widehat{r}$ under hypotheses H_0 and H_1 (calculated from the inverse Godambe information matrix defined in (22)) . By denoting as $\Phi(x)$ the cumulative distribution function of the Gaussian distribution $\mathcal{N}(0, 1)$, the following result is then classically obtained

$$P_{\mathrm{D}} = \Phi\left[\frac{r_0 - r_1}{\sigma_1} + \frac{\sigma_0}{\sigma_1}\Phi^{-1}(P_{\mathrm{FA}})\right]. \quad (26)$$

This result provides theoretical asymptotic expressions for the ROCs associated to the detection problem (25) and allow us to analyze detection performance as functions of the MuBGD parameters. For instance, Fig. 6 shows P_{D} as functions of q_1 and q_2 for a given probability of false alarm $P_{\mathrm{FA}} = 0.3$. This figure clearly confirms that the detection performance is
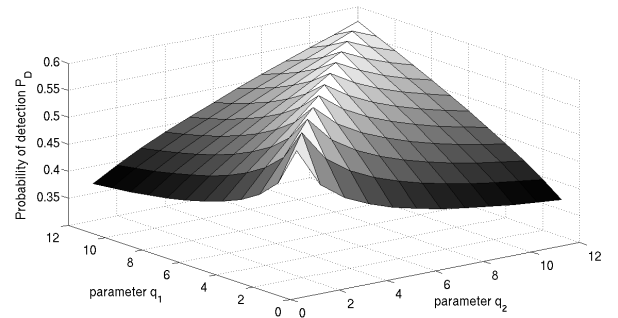

Fig. 6. P_{D} versus shape parameters q_1 and q_2 ($P_{\mathrm{FA}} = 0.3$, $n = 1$).

a decreasing function of $q_2 - q_1$.

D. Change detection in real images

This section first considers images acquired at different dates around Gloucester (England) before and during a flood (on Sept. 9, 2000 and Oct. 21, 2000 respectively). The 1-look images as well as a mask indicating the pixels affected by the flood are depicted on Figs 7(a), 7(b) and 7(c). The reference map 7(c) was obtained by photo-interpreters – who used the same SAR images we are using – and a reference map built from Landsat and SPOT data acquired one day after the radar image. The original 1-look images have been transformed into images with larger numbers of looks by replacing each pixel by the average of pixels belonging to a given neighborhood. This section compares the performance of the following change detectors
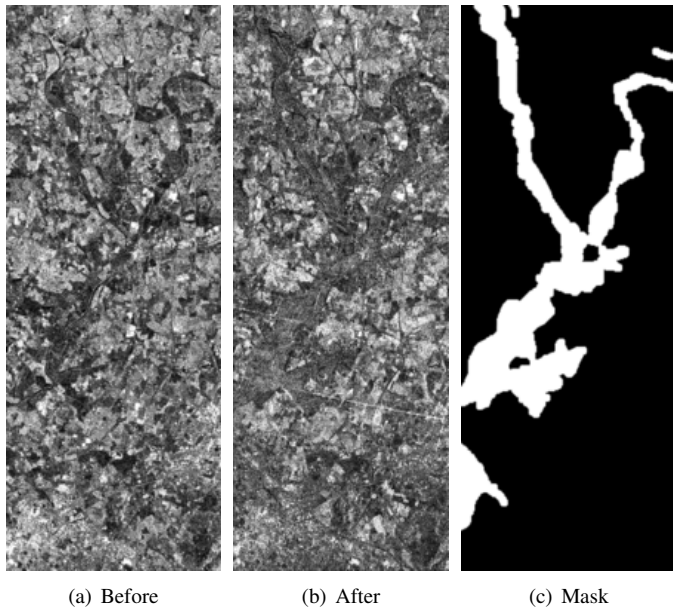
(a) Before   (b) After   (c) Mask

Fig. 7. Radarsat images of Gloucester before and after flood.

(a) Window size $n = 9 \times 9$ (b) Window size $n = 15 \times 15$

(c) Window size $n = 21 \times 21$

Fig. 8. ROCs for Gloucester images ($q_1 = 1$, $q_2 = 5$).

- the ratio edge detector which has been intensively used for SAR images [19], [20]. This detector mitigates the effects of the multiplicative speckle noise by computing the ratio of averages of pixel values belonging to neighborhoods of the pixels under consideration.
- the correlation change detector, where $\widehat{r}$ in (25) has been estimated with the moment estimator (referred to as "*Correlation Moment*"),
- the correlation change detector, where $\widehat{r}$ in (25) has been estimated with the IFM method for BGDs (referred to as "*Correlation IFM*").

The ROCs for this change detection problem are shown on Figs 8(a), 8(b) and 8(c) for different window sizes ($n = 9 \times 9$, $n = 15 \times 15$ and $n = 21 \times 21$). The numbers of looks for the two images are $q_1 = 1$ and $q_2 = 5$. The correlation IFM detector clearly provides the best results.

The second set of experiments is related to a couple of Radarsat images acquired before and after the eruption of the Nyiragongo volcano which occurred in January 2002. The Radarsat images are depicted on figures 9(a) (before eruption) and 9(b) (after eruption). Note that some changes due to the eruption can be clearly seen on the landing track for example. Figure 9(c) indicates the pixels of the image which have been affected by the eruption (white pixels). The ROCs for this change detection problem are shown on Figs 10(a), 10(b) and 10(c) for different window sizes ($n = 9 \times 9$, $n = 15 \times 15$ and $n = 21 \times 21$). The numbers of looks for the two images are $q_1 = 3$ and $q_2 = 5$. The correlation IFM detector provides better performance than the conventional correlation moment detector in all cases. The ratio edge detector also shows interesting detection performance for this example because the volcano eruption has produced significant changes in the pixel intensities. Note however that the proposed correlation IFM detector gives better performance for large PFAs. Even if these large PFA values are usually considered as a bad
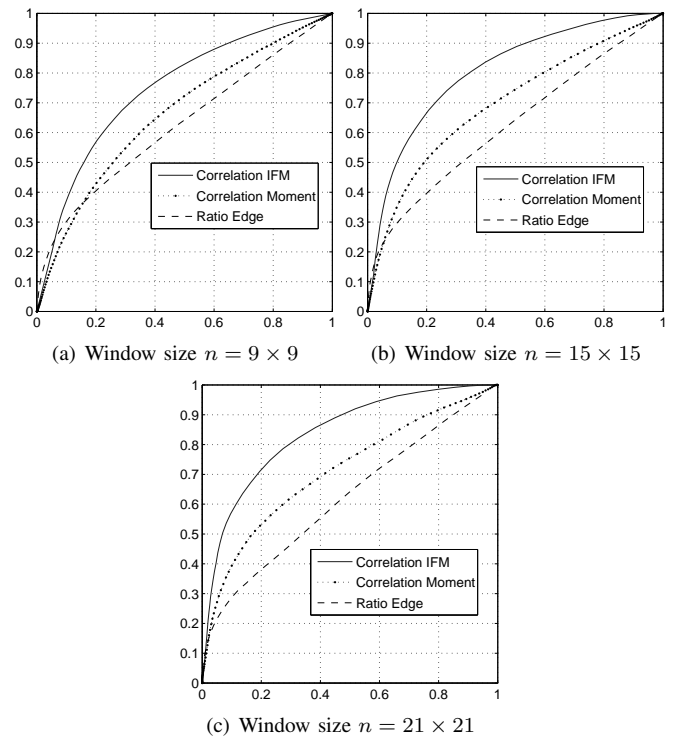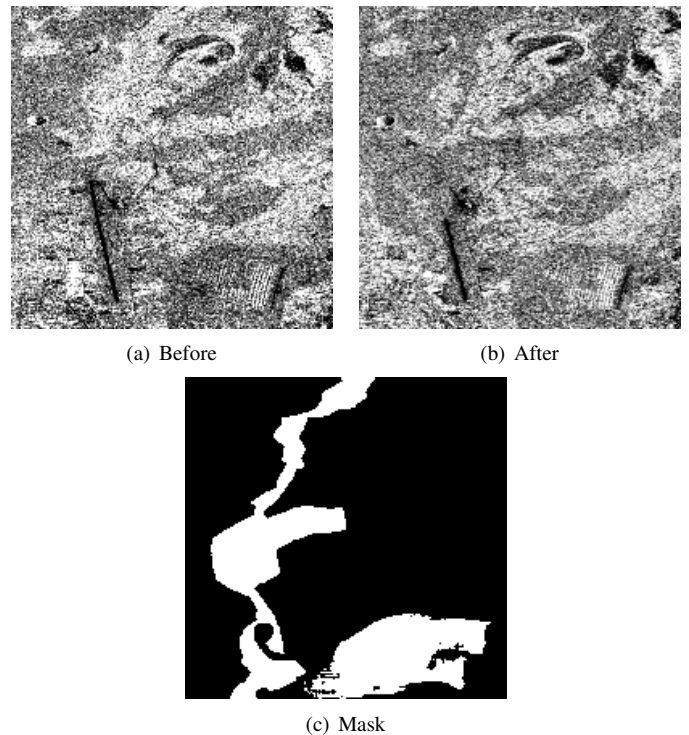

(a) Before   (b) After

(c) Mask

Fig. 9. Radarsat images of Nyiragongo before and after eruption.

result in classical detection problems, the reader has to bear in mind that when working with images, simple post-processing strategies can dramatically improve the change detection performance. Indeed, when looking at detection maps, two types of false alarms can be observed: isolated pixels and boundary pixels. For the first type of errors, a simple median filter a
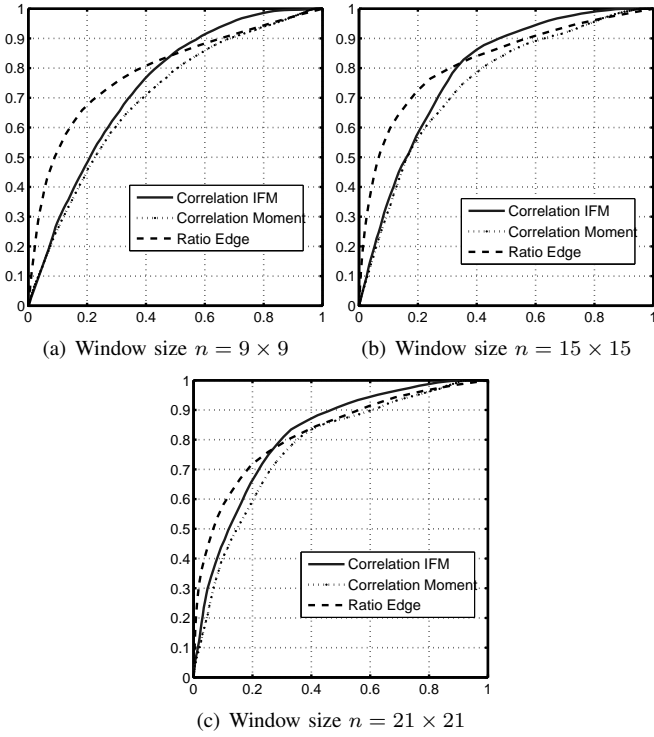
Fig. 10. ROCs for Nyiragongo images ($q_1 = 3$, $q_2 = 6$).

morphological opening, gives very good results. The second type of false alarm is due to the spatial extent of the estimation windows, which over-detect at the output boundaries of the change areas. This is not a main drawback in terms of change map production, since the change areas remain the same and only the spatial resolution of the map is affected.

VI. Conclusions

This paper studied a new family of multivariate gamma based distributions for multisensor SAR images referred to as MuMGDs. Estimation algorithms based on the ML method, the IFM principle and the methods of moments were studied to estimate the parameters of these distributions. In particular, the estimated correlation coefficient of MuMGDs showed interesting properties for detecting changes in radar images with different numbers of looks.

Being able to handle images with different numbers of looks is very useful, not only when the images have been acquired by different sensors, but also when both sensors have the same theoretical number of looks. Indeed, change detection algorithms require precise image co-registration which is usually achieved by image interpolation. Image interpolation and other image pre-processing steps modify locally the equivalent number of looks of the images. Therefore, even if the images have been acquired by the same sensor in the same imaging mode, differences in the equivalent number of looks can be observed. The algorithms presented in this paper could be used for detecting changes in this kind of images. Of course, in the case where the equivalent number of looks has to be estimated locally, an assessment of the influence of the estimation errors in the final MuMGD parameter estimation should be addressed. This point is currently under investigation.

Appendix I
Numerical evaluation of the Horn function Φ_3

Some series representation in terms of special functions are useful to compute hypergeometric series of order two [21]. For the Horn function Φ_3 defined in (13), the following expansion is particularly useful:

$$\Phi_3(a; b; x, y) = \sum_{n=0}^{\infty} \frac{y^n}{(b)_n n!} \, _1F_1[a, b + n, x],$$

where $_1F_1$ is the confluent hypergeometric series of order one, i. e. $_1F_1[a, b, x] = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n n!} x^n$. This confluent hypergeometric series $_1F_1[a, b, x]$ can be expressed as follows [22]:

$$\begin{aligned} _1F_1[a, b, x] = \frac{\Gamma(b)}{\Gamma(a)} e^x x^{a-b} \sum_{i \geq 0} \frac{(b-a)_i (1-a)_i}{i! x^i} \times \\ F_\gamma(x; i + b - a), \end{aligned} \quad (27)$$

where $F_\gamma(x; \nu)$ is the cumulative distribution function of a univariate gamma distribution with shape parameter ν and scale parameter 1. Note that the summation in (27) is finite since $a \geq 1$ is an integer. This yields the following expression of Φ_3:

$$\begin{aligned} \Phi_3(a; b; x, y) = \frac{\Gamma(b)}{\Gamma(a)} e^x x^{a-b} \sum_{n=0}^{\infty} \frac{(y/x)^n}{n!} \times \\ \sum_{i \geq 0} \frac{(b+n-a)_i (1-a)_i}{i! x^i} F_\gamma(x; i + b + n - a), \end{aligned} \quad (28)$$

where the last summation ($i \geq 0$) is finite. Equation (28) provides a numerically stable way of evaluating $\Phi_3(a; b; x, y)$ for large values of x and y. When (x, y) is close to $(0, 0)$, the definition of Φ_3 in (13) will be preferred.

Appendix II
Derivatives of the Horn function Φ_3

From the series representation of the function Φ_3 defined in (13), the following results can be obtained:

$$\begin{aligned} \frac{\partial}{\partial x} \Phi_3(a; b; x, y) &= \sum_{m \geq 1, n \geq 0} \frac{(a)_m}{(b)_{m+n} (m-1)! n!} x^{m-1} y^n, \\ &= \frac{\Gamma(a+1)}{\Gamma(a)} \frac{\Gamma(b)}{\Gamma(b+1)} \sum_{m, n \geq 0} \frac{(a+1)_m}{(b+1)_{m+n} m! n!} x^m y^n, \\ &= \frac{a}{b} \Phi_3(a+1; b+1; x, y). \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial y} \Phi_3(a; b; x, y) &= \sum_{m \geq 0, n \geq 1} \frac{(a)_m}{(b)_{m+n} m! (n-1)!} x^m y^{n-1}, \\ &= \frac{\Gamma(b)}{\Gamma(b+1)} \sum_{m, n \geq 0} \frac{(a)_m}{(b+1)_{m+n} m! n!} x^m y^n, \\ &= \frac{1}{b} \Phi_3(a; b+1; x, y). \end{aligned}$$

REFERENCES

[1] J. Inglada, "Similarity measures for multisensor remote sensing images," in *Proc. IEEE IGARSS-02*, Toronto, Canada, June 2002, pp. 104–106.

[2] J. Inglada and A. Giros, "On the possibility of automatic multi-sensor image registration," *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 10, pp. 2104–2120, Oct. 2004.

[3] T. F. Bush and F. T. Ulaby, "Fading characteristics of panchromatic radar backscatter from selected agricultural targets," *IEEE Trans. Geosci. Remote Sensing*, vol. 13, no. 4, pp. 149–157, 1975.

[4] S. B. Lev, D. Bshouty, P. Enis, G. Letac, I. L. Lu, and D. Richards, "The diagonal natural exponential families and their classification," *J. of Theoret. Probab.*, vol. 7, no. 4, pp. 883–928, Oct. 1994.

[5] P. Bernardoff, "Which multivariate Gamma distributions are infinitely divisible?" *Bernoulli*, vol. 12, no. 1, pp. 169–189, 2006.

[6] F. Chatelain, J.-Y. Tourneret, A. Ferrari, and J. Inglada, "Bivariate gamma distributions for image registration and change detection," *IEEE Trans. Image Processing*, 2006, to appear.

[7] F. Chatelain, J.-Y. Tourneret, J. Inglada, and A. Ferrari, "Parameter estimation for multivariate gamma distributions. Application to image registration." in *Proc. EUSIPCO-06*, Florence, Italy, Sept. 2006.

[8] S. Kotz, N. Balakrishnan, and N. L. Johnson, *Continuous Multivariate Distributions*, 2nd ed.   New York: Wiley, 2000, vol. 1.

[9] F. O. A. Erdlyi, W. Magnus and F. Tricomi, *Higher Transcendental Functions*.   New York: Krieger, 1981, vol. 1.

[10] B. G. Lindsay, "Composite likelihood methods," *Contempory Mathematics*, vol. 50, pp. 221–239, 1988.

[11] C. C. Heyde, *Quasi-Likelihood and its application. A general approach to optimal parameter estimation.*   New York: Springer, 1997.

[12] F. Chatelain and J.-Y. Tourneret, "Composite likelihood estimation for multivariate mixed Poisson distributions," in *Proc. IEEE-SP Workshop Stat. Signal Processing*, Bordeaux, France, July 2005, pp. 49–54.

[13] ——, "Estimating the correlation coefficient of bivariate gamma distributions using the maximum likelihood principle and the inference functions for margins," University of Toulouse (IRIT/ENSEEIHT), Tech. Rep., May 2007, available at http://florent.chatelain.free.fr/Publi/ChatelainTechReport07.pdf.

[14] H. Joe, *Multivariate Models and Dependence Concepts*.   London: Chapman & Hall, May 1997, vol. 73.

[15] ——, "Asymptotic efficiency of the two-stage estimation method for copula-based models," *J. Multivar. Anal.*, vol. 94, no. 2, pp. 401–419, 2005.

[16] B. Porat and B. Friedlander, "Performance analysis of parameter estimation algorithms based on high-order moments," *International Journal of adaptive control and signal processing*, vol. 3, pp. 191–229, 1989.

[17] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer, 2004.

[18] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: Part I*. New York: Wiley, 1968.

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

[20] E. J. M. Rignot and J. J. van Zyl, "Change Detection Techniques for ERS-1 SAR Data," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, no. 4, pp. 896–906, 1993.

[21] F. D. Colavecchia and G. Gasaneo, "f1: a code to compute Appell's F_1 hypergeometric function," *Computer Physics Communications*, vol. 157, pp. 32–38, feb 2004.

[22] K. E. Muller, "Computing the confluent hypergeometric function, $M(a, b, x)$," *Numerische Mathematik*, vol. 90, no. 1, pp. 179–196, Nov. 2001.