

Supervised High Resolution Dual Polarization SAR Image Classification by Finite Mixtures and Copulas

Vladimir Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia

▶ **To cite this version:**

Vladimir Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia. Supervised High Resolution Dual Polarization SAR Image Classification by Finite Mixtures and Copulas. IEEE Journal of Selected Topics in Signal Processing, 2011, 10.1109/JSTSP.2010.2103925 . inria-00562326

HAL Id: inria-00562326

https://inria.hal.science/inria-00562326

Submitted on 3 Feb 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

Supervised High Resolution Dual Polarization SAR Image Classification by Finite Mixtures and Copulas

Vladimir A. Krylov, Gabriele Moser, *Member, IEEE,* Sebastiano B. Serpico, *Fellow, IEEE,* and Josiane Zerubia, *Fellow, IEEE*

***Abstract*—In this paper a novel supervised classification approach is proposed for high resolution dual polarization (dual-pol) amplitude satellite synthetic aperture radar (SAR) images. A novel probability density function (pdf) model of the dual-pol SAR data is developed that combines finite mixture modeling for marginal probability density functions estimation and copulas for multivariate distribution modeling. The finite mixture modeling is performed via a recently proposed SAR-specific dictionary-based stochastic expectation maximization approach to SAR amplitude pdf estimation. For modeling the joint distribution of dual-pol data the statistical concept of copulas is employed, and a novel copula-selection dictionary-based method is proposed. In order to take into account the contextual information, the developed joint pdf model is combined with a Markov random field approach for Bayesian image classification. The accuracy of the developed dual-pol supervised classification approach is validated and compared with benchmark approaches on two high resolution dual-pol TerraSAR-X scenes, acquired during an epidemiological study. A corresponding single-channel version of the classification algorithm is also developed and validated on a single polarization COSMO-SkyMed scene.**

***Index Terms*—Polarimetric synthetic aperture radar, supervised classification, probability density function (pdf), dictionary-based pdf estimation, Markov random field, copula.**

I. Introduction

IN MODERN remote sensing, the use of synthetic aperture radar (SAR) represents an important source of information for Earth observation. Recent improvements have enabled modern satellite SAR missions, such as COSMO-SkyMed and TerraSAR-X, to acquire high resolution (HR) data (up to metric resolution) with a very short revisit time (e.g. 12 hours for COSMO-SkyMed). In addition, SAR is robust with respect to lack of illumination and atmospheric conditions. Together, these factors explain the rapidly growing interest in SAR imagery for various applications, such as flood/fire monitoring, urban mapping and epidemiological surveillance. Classification is one of the most important image processing tasks applied to remote sensing imagery. Classification maps can either be directly used in applications or they can serve as an input to further SAR processing problems, e.g., for change detection.

Contemporary satellite SAR missions are capable of providing polarimetric imagery, which provide a more complete description of landcover scattering behavior than single-channel SAR data [1], [2]. The potential for improved classification accuracy with data in several polarizations, compared to single-channel data, explains the special interest to polarimetric SAR image classification. Furthermore, several current satellite SAR systems, e.g., TerraSAR-X, COSMO-SkyMed, support, at least, dual-pol acquisition mode. In this paper we investigate the dual polarization (dual-pol) SAR imagery scenario, as well as single polarization (single-pol) SAR, as a special case.

A wide variety of methods have been developed earlier for the classification of polarimetric SAR data [2]. We list some recent methods based on the employed methodological approach: maximum likelihood [3]–[7], neural networks [8], [9], support vector machines [10], fuzzy methods [11], [12], stochastic complexity [13], spectral graph partitioning [14], wavelet texture models [15] and other approaches [16], [17]. In this paper we develop a classification method based on the maximum likelihood approach. As such, this method specifies a probability density function (pdf) describing the statistics of polarimetric SAR data. Previously, several models have been proposed for this purpose: the classical Wishart distribution [3], [18], the K-distribution [4], [19] for textured areas, the K-Wishart distribution [6] designed to improve the distinguishability of non-Gaussian regions, the G-distribution [5], [20] for extremely heterogeneous areas, the Ali-Mikhail-Haq copula-based model [21] combined with the Sinclair matrix representation, and the KummerU distribution [7] for Fisher distributed texture. These models were developed for the multilook complex-valued SAR image statistics. In this paper, we study SAR classification using only the amplitude data and not the complex-valued data. This is an important data typology because several image products provided by novel high resolution satellite SAR systems are geocoded ellipsoid-corrected amplitude (intensity) images, and because several earlier coarser resolution sensors (e.g., ERS) primarily used this modality.

The classification technique developed in this paper combines the Markov random field (MRF) approach to Bayesian image classification with the dictionary-based stochastic expectation maximization (DSEM) amplitude pdf estimator.

This research was a collaborative effort between the Institut National de Recherche en Informatique et en Automatique (INRIA) Sophia Antipolis - Méditerranée Center, France, and the Dept. of Biophysical and Electronic Engineering (DIBE) of the University of Genoa, Italy. It was carried out with the partial financial support of the INRIA and the French Space Agency (CNES).

V.A. Krylov is with the Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, 119991, Moscow, Russia, and also with the EPI Ariana, INRIA/I3S, F-06902, Sophia Antipolis Cedex, France (e-mail: vkrylov@cs.msu.ru).

G. Moser, S.B. Serpico are with DIBE, University of Genoa, I-16145, Genoa, Italy (e-mail: gabriele.moser@unige.it, sebastiano.serpico@unige.it).

J. Zerubia is with the EPI Ariana, INRIA/I3S, F-06902, Sophia Antipolis Cedex, France (e-mail: Josiane.Zerubia@inria.fr).

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

MRFs represent a general family of probabilistic image models that provide a convenient and consistent way to characterize context dependent data [22]. The latter has been developed for the purpose of modeling the amplitude distributions of SAR images [23], [24]. Contrary to various pdf models [1], [25]–[30], DSEM uses a mixture of several distinct SAR-specific pdfs to accurately model the amplitude statistics. The benefit of using the finite mixture DSEM approach is twofold: First, DSEM is an efficient and automatic tool for SAR amplitude pdf estimation, capable of providing estimates of higher accuracy than single parametric pdf models, and, second, the underlying mixture assumption in DSEM enables accurate characterization of the inhomogeneous classes of interest, i.e., classes which contain several different landcover subclasses. This is particularly important for HR imagery. The resulting DSEM-MRF technique is a simple and efficient tool for single-channel SAR classification. In order to support dual-pol SAR data, copula theory is used for modeling the joint class-conditional distributions of the dual-pol channels, resulting in a Copula-DSEM-MRF approach (CoDSEM-MRF). The employed joint distribution modeling tool, copulas [31], is a rapidly developing statistical tool that was designed for constructing joint distributions from marginals with a wide variety of allowable dependence structures. For every class the choice of an optimal copula from a dictionary of copulas is performed by a dedicated criterion. The concept of copulas is relatively new in image processing, and has just emerged in remote sensing methods [21], [32], [33]. The approach suggested in [21] can also be used for dual-pol SAR and it is compared with the developed in this paper technique in Section IV. The proposed CoDSEM-MRF HR dual-pol SAR classification technique is based on three flexible statistical modeling concepts, i.e. copulas, finite mixtures and MRFs, and constitutes an efficient and robust approach with respect to possibly sophisticated classes of interest.

The contributions of this paper are twofold: First, we develop a novel single-channel HR SAR classification approach that is based on SAR-specific DSEM probability density function estimation technique and a contextual MRF model, and, second, we use a novel parametric statistical modeling approach, which is based on a dictionary of copulas, to develop a dual-pol HR SAR classification method. This method outperforms benchmark approaches, such as the parametric 2D Nakagami-Gamma [18] model combined with MRF, and the nonparametric "K-nearest neighbors" [34] method combined with MRF.

This paper is organized as follows. In Section II, we present an overview of the developed DSEM-MRF and CoDSEM-MRF approaches to HR SAR classification. In Section III, the description of methodological components of the designed approach, i.e., the DSEM approach to amplitude pdf estimation and bivariate copulas, are presented. Section IV reports the experiments of the developed SAR classification approach on HR dual-pol TerraSAR-X and single-pol COSMO-SkyMed images. Conclusions are drawn in Section V.

II. Method overview

In this section, we present an overview of the developed CoDSEM-MRF approach to dual-pol SAR classification. We treat single-channel SAR classification (DSEM-MRF) as a special case of SAR classification. The first two steps of the algorithm (the DSEM and Copula steps) are presented in more detail in the next section.

A. Dual-pol case

In this subsection, we present the key steps of the proposed CoDSEM-MRF classification algorithm for the case of dual-pol SAR (2 polarization channels). We study supervised classification with M classes of interest. Thus we assume training pixels for these classes to be available.

DSEM step. In the first step, the marginal pdfs of polarization channels are estimated separately by applying DSEM to the training pixels available for the considered classes of interest. For each m-th class and d-th channel ($m = 1, \ldots, M$, $d = 1, 2$), the mixture DSEM pdf estimators $p_{dm}(y_d|\omega_m)$ and the corresponding cumulative distribution functions (CDFs) $F_{dm}(y_d|\omega_m)$ are as follows:

$$\begin{aligned} p_{dm}(y_d|\omega_m) &= \sum_{i=1}^{K_{dm}} P_{dmi} p_{dmi}(y_d), \\ F_{dm}(y_d|\omega_m) &= \sum_{i=1}^{K_{dm}} P_{dmi} F_{dmi}(y_d), \end{aligned} \tag{1}$$

where ω_m is the event of the observation belonging to the m-th class, y_1 and y_2 are the amplitudes from the two polarization channels and K_{dm} is the number of components in the mixture. F_{dmi} and p_{dmi} represent the i-th mixture component in the CDF and pdf domains, respectively, and P_{dmi} is the related mixture proportion; p_{dmi} is automatically drawn by DSEM from a dictionary of several SAR-specific parametric pdfs, $i = 1, \ldots, K_{dm}$. There are two reasons why DSEM was used instead of a single parametric pdf model. First, DSEM is an efficient and automatic tool for SAR amplitude pdf estimation, and it is capable of providing estimates of higher accuracy compared to single parametric pdf models [23], [24]. Specifically, DSEM was experimentally found to accurately estimate the statistics of a HR satellite SAR image [24]. Second, the underlying mixture assumption in DSEM enables accurate characterization of the inhomogeneous classes of interest, i.e., the classes that contain several different landcover subclasses. This is a very important property when dealing with HR imagery, since the corresponding image statistics are usually strongly mixed due to the high level of spatial detail appreciable at high resolution. However, even in the case of homogenous classes, DSEM can be viewed as a tool for choosing the best single pdf model from the set of pdfs in the DSEM dictionary.

Copula step. The goal of this step is to merge the marginal pdfs, which correspond to polarization channels estimated on DSEM-step, into a joint bivariate pdf describing the joint amplitude distribution of the dual-pol SAR image. To this end, the joint pdfs $p_m(y_1, y_2|\omega_m)$ for classes $m = 1, \ldots, M$ are

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

modeled via copulas from the marginal distributions in Eq. (1). The choice of copulas as a tool for joint distribution estimation is attractive because of the simplicity in its analytical formulations, and because a wide variety of dependence structures can be modeled. Thus, joint pdfs are constructed as:

$$p_m(y_1, y_2|\omega_m) = p_{1m}(y_1|\omega_m)p_{2m}(y_2|\omega_m) \times \frac{\partial^2 C_m^*}{\partial y_1 \partial y_2}(F_{1m}(y_1|\omega_m), F_{2m}(y_2|\omega_m)), \quad (2)$$

where C_m^* is a specific copula for the m-th class and is automatically picked by the algorithm from a dictionary of considered copulas ($m = 1, \ldots, M$; details are found in Sec. III).

MRF step. In order to take into consideration the contextual information disregarded by the pixel-wise Copula-DSEM technique, and to gain robustness against the inherent noise-like phenomenon of SAR known as speckle, we adopted a contextual approach based on an MRF model. Following the classical definitions of MRFs (see, e.g., [35], [36]) on the two dimensional lattice S of N observations $\mathbf{y} = \{y_1, \ldots, y_N\}$ and class labels $\mathbf{x} = \{x_1, \ldots, x_N\}$, $x_i \in \{1, \ldots, M\}$, we introduced an isotropic second-order neighborhood system C with cliques of size 2. The Hammersley-Clifford theorem [36] allows for the presentation of the joint probability distribution of an MRF as a Gibbs distribution:

$$P(\mathbf{x}) = Z^{-1} \exp(-H(\mathbf{x}|\beta)),$$

where $Z = \sum_{\mathbf{z}} \exp(-H(\mathbf{z}|\beta))$ is a normalizing constant, β is a positive parameter and $H(\mathbf{x}|\beta)$ is the MRF energy function of the class labels. More specifically $H(\mathbf{x}|\beta)$ takes the following form:

$$H(\mathbf{x}|\beta) = \sum_{\{s,s'\}\in C} \left[-\beta\, \delta_{x_s = x_{s'}}\right],$$

where δ is the Kronecker delta:

$$\delta_{x_s = x_{s'}} = \begin{cases} 1, & \text{if } x_s = x_{s'} \\ 0, & \text{otherwise.} \end{cases}$$

Image classification poses a problem of recovering the unobserved data, i.e., class labels. In the case of hidden MRFs, the unobserved data $\mathbf{x}$ are modeled by an MRF and the observed data $\mathbf{y}$, i.e., SAR amplitudes, are assumed to be conditionally independent given $\mathbf{x}$ [35]:

$$p(\mathbf{y}_i|\mathbf{y}_{S\setminus\{i\}}, x_i) = p(\mathbf{y}_i|x_i), \qquad \forall i \in S,$$

where $\mathbf{y}_i$ is the pair (y_1, y_2) at pixel i, $\mathbf{y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$, and $S\setminus\{i\}$ represents all the pixels of S except i. Thus, given the conditional pdfs in Eq. (2), we model the full data $(\mathbf{y}, \mathbf{x})$ as a hidden MRF with the energy function:

$$U(\mathbf{x}|\mathbf{y}, \beta) = \sum_{i\in S} U_i(x_i|\mathbf{y}_i, \mathbf{x}_{S\setminus\{i\}}, \beta), \quad (3)$$

where for $m = 1, \ldots, M$:

$$U_i(x_i = \omega_m|\mathbf{y}_i, \mathbf{x}_{S\setminus\{i\}}, \beta) = -\log p_m(\mathbf{y}_i|\omega_m) - \beta \sum_{s:\{i,s\}\in C} \delta_{x_i = x_s}. \quad (4)$$

The energy function in Eq. (3) has a single parameter β that must be estimated. Conveniently, all of the parameter estimation involved in Eq. (1) is incorporated into DSEM. Thus, the resulting energy function (1)-(3) also has only one parameter. In order to estimate β, we used a simulated annealing procedure [37] with a pseudo-likelihood function PL [36] of the following form:

$$\log PL(\mathbf{x}|\beta) = \log \left[\prod_{s\in S} P(x_s|\mathbf{x}_{S\setminus\{s\}}, \beta)\right],$$

where:

$$P(x_s|\mathbf{x}_{S\setminus\{s\}}, \beta) = \frac{\exp(-U(x_s|\mathbf{x}_{S\setminus\{s\}}, \beta))}{\sum_{z_s\in X_S} \exp(-U(z_s|\mathbf{x}_{S\setminus\{s\}}, \beta))},$$

and $X_s = \{\omega_1, \ldots, \omega_M\}$.

The implemented simulated annealing procedure generated a sequence $\{\beta_t\}$ of estimates, and employed the normal proposal distribution $N(\beta_t, 1)$ [38] together with an exponentially decreasing cooling schedule $T_t = 0.95 \cdot T_{t-1}$. Considering the convergence in distribution of the related Metropolis algorithm, the final estimate β^* was set by averaging the estimates on the last n iterations. The estimation of β was performed on an ML pre-classification of the image, which associated every pixel with the highest probability-density label according to the Copula-DSEM models.

Energy minimization step. This step involves the minimization of the energy in Eq. (3). For this optimization problem the iterative deterministic modified Metropolis dynamics (MMD) [40] algorithm was adopted. This is a compromise between the deterministic "iterated conditional modes" algorithm (ICM) [36], which is a fast local minimization technique, that is strongly dependent on the initial configuration, and simulated annealing (SA) [37], which is a very computationally intensive global minimization approach. MMD is computationally feasible and provides reasonable results in real classification problems [40]. MMD is set with a cooling schedule (like SA) and proceeds as follows:

1. sample a random initial label configuration $\mathbf{x}^0$, define an initial temperature T^0 and parameters $\alpha \in (0, 1)$, $n_1 \in \mathbb{N}$, $\tau \in (0, 1)$, $\gamma \in \mathbb{R}^+$, initialize $k = 0$ and $T_0 = T^0$;
2. set $i = 0$;
3. using uniform distribution pick up a label configuration $\boldsymbol{\eta}$ which differs exactly in one element from $\mathbf{x}^k$;
4. compute $\Delta = U(\boldsymbol{\eta}) - U(\mathbf{x}^k)$ and accept $\boldsymbol{\eta}$ according to the rule:
$$\mathbf{x}^{k+1} = \begin{cases} \boldsymbol{\eta}, & \text{if } \Delta \leqslant 0 \\ \boldsymbol{\eta}, & \text{if } \Delta > 0 \text{ and } \ln(\alpha) \leqslant -\frac{\Delta}{T_k} \\ \mathbf{x}^k, & \text{otherwise} \end{cases};$$
5. calculate $\Delta U_i = |U(\mathbf{x}^{k+1}) - U(\mathbf{x}^k)|$;
6. if $i \leqslant n_1$, set $i = i + 1$ and goto Step 3;
7. calculate $\Delta U = \sum_{i=1}^{n_1} \Delta U_i$;
8. if $\Delta U / U(\mathbf{x}^{k+1}) > \gamma$ decrease the temperature $T_{k+1} = \tau T_k$, increase $k = k+1$, and goto Step 2; stop otherwise.

During the early stages of this iterative procedure, the behavior of MMD is close to that of SA, and thus it provides much better exploratory properties as compared to ICM.

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

TABLE I
PDFS AND MOLC EQUATIONS FOR THE PARAMETRIC FAMILIES INCLUDED IN THE CONSIDERED DICTIONARY $\mathcal{D}_M$: LOG-NORMAL, WEIBULL, NAKAGAMI AND GENERALIZED GAMMA (GΓD) DISTRIBUTIONS. HERE $\Gamma(\cdot)$ IS THE GAMMA FUNCTION, $\Psi(\cdot)$ THE DIGAMMA FUNCTION AND $\Psi(\nu,\cdot)$ THE ν^{TH} ORDER POLYGAMMA FUNCTION [39]

Family	Probability density function	MoLC equations
Log-normal	$f_1(r\|m,\sigma)=\frac{1}{\sigma r\sqrt{2\pi}}\exp\left[-\frac{(\ln r-m)^2}{2\sigma^2}\right],\ r>0$	$\kappa_1=m$ $\kappa_2=\sigma^2$
Weibull	$f_2(r\|\mu,\eta)=\frac{\eta}{\mu^\eta}r^{\eta-1}\exp\left[-\left(\frac{r}{\mu}\right)^\eta\right],\ r\geqslant 0$	$\kappa_1=\ln\mu+\Psi(1)\eta^{-1}$ $\kappa_2=\Psi(1,1)\eta^{-2}$
Nakagami	$f_3(r\|\lambda,L)=\frac{2}{\Gamma(L)}(\lambda L)^L r^{2L-1}\exp\left(-\lambda L r^2\right),\ r\geqslant 0$	$2\kappa_1=\Psi(L)-\ln\lambda L$ $4\kappa_2=\Psi(1,L)$
GΓD	$f_4(r\|\kappa,\sigma,\nu)=\frac{\nu}{\sigma\Gamma(\kappa)}\left(\frac{r}{\sigma}\right)^{\kappa\nu-1}\exp\left\{-\left(\frac{r}{\sigma}\right)^\nu\right\},\ r\geqslant 0$	$\kappa_1=\Psi(\kappa)/\nu+\ln\sigma$ $\kappa_j=\Psi(j-1,\kappa)/\nu^j,\ j=2,3$

B. Single-pol case

Here we consider single-pol classification as a particular case of the developed algorithm. In this case the target joint pdf is univariate. Thus there is no need for the **Copula-step**, because the **DSEM-step** estimates provide the goal pdfs and can be directly plugged into Eq. (3) in the **MRF-step**. In the single-pol case the algorithm (DSEM-MRF) represents the straightforward application of the DSEM technique to MRF-based SAR image classification.

III. METHODOLOGICAL COMPONENTS

A. Dictionary-based Stochastic Expectation Maximization

In this subsection, we present the outline of DSEM. This approach was initially developed in [23] for medium-resolution SAR and then enhanced and validated on high resolution SAR in [24].

To take into account the heterogeneity, when several distinct land-cover typologies are present in the same SAR image, a finite mixture model [41] for the distribution of grey levels is assumed. Specifically, DSEM is applied to estimate marginal class-conditional statistics. Separately focusing on each class ω_m and each channel y_d, $m=1,\ldots,M$, $d=1,2$, dropping the subscripts m and d for ease of notation, and denoting y_d simply as r, we assume that the d-th component of the training samples of ω_m are independent and identically distributed random variables, drawn from a mixture pdf with K components:

$$p(r)=\sum_{i=1}^{K}P_i p_i(r),\quad r\geqslant 0, \tag{5}$$

where $p_i(\cdot)$ is the i-th mixture component and $\{P_i\}$ is a set of mixing proportions, i.e., $\sum_{i=1}^{K}P_i=1$ and $0\leqslant P_i\leqslant 1$ for $i=1,\ldots,K$. Each component $p_i(\cdot)$ is modeled by resorting to a finite dictionary $\mathcal{D}_M=\{f_1,\ldots,f_4\}$ (see Table I) of four SAR specific distinct parametric pdfs $f_j(r|\theta_j)$, parameterized by $\theta_j\in A_j$, $j=1,\ldots,4$ [24]. Assuming that channel r is quantized on the levels $\{0,1,\ldots,Z-1\}$, we also denote by $h(z)$, $z=0,\ldots,Z-1$, the histogram of r, when restricted to the training samples of the considered class.

The calculation of CDFs is needed to merge the marginal DSEM pdf estimates of polarization channels into joint pdfs via copulas (see Eq. (2)). Therefore, the pdfs in the dictionary have been chosen so that, for each of them, there is either an analytical closed-form expression (f_1,f_2), or a simple numerical approximation procedure (f_3,f_4) for the related CDF. For the rest of the models considered in [24], where the dictionary includes also Fisher [27], $\mathcal{K}$-root [25], generalized Gaussian-Rayleigh [29] and symmetric-α-stable generalized Rayleigh [28] pdf models, there are no closed-form CDFs and numerical approximation is computationally intensive. Furthermore, the experiments in [24] demonstrate that very accurate pdf estimates can be obtained with this dictionary of four models. Therefore, the dictionary $\mathcal{D}_M$ considered in this paper has been restricted to the four models in Table I.

As discussed in [23], considering the variety of estimation approaches for finite mixtures, the appropriate choice for this particular estimation problem is the iterative stochastic expectation maximization (SEM) scheme [42]. Instead of adopting ML estimates as the classical SEM scheme suggests [42], DSEM employs the Method of Log-Cumulants (MoLC) [27], [43] for component parameter estimation, which has been demonstrated to be a feasible and effective estimation tool for all the pdfs in $\mathcal{D}_M$ [23], [24], [27], [30]. The MoLC equations relate the unknown pdf parameters with κ_1, κ_2 and κ_3:

$$\begin{cases}\kappa_1=E\{\ln r\}\\ \kappa_j=E\{(\ln r-\kappa_1)^j\}\end{cases},\quad j=2,3,$$

which are the 1st, 2nd, 3rd order logarithmic cumulants, respectively. The MoLC equations have a single solution for any observed values of log-cumulants for all pdfs in $\mathcal{D}_M$ (see Table I). MoLC equations for f_1 and f_2 allow analytical solutions. For f_3 and f_4, the strict monotonicity of the involved functions allows solutions to be reached by numerical procedures. As in [24], we implemented a K-estimation procedure that consists of initializing SEM with $K^0=K_{max}$, and then allowing components to be eliminated from the mixture during the DSEM iterative process, once their priors become too small, thus decreasing K.

Thus, each t-th iteration of DSEM goes as follows:

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

TABLE II
CONSIDERED DICTIONARY $\mathcal{D}_C$ OF COPULAS: CLAYTON, ALI-MIKHAIL-HAQ, GUMBEL, FRANK, A12, A14, FARLIE-GUMBEL-MORGENSTERN (FGM), MARCHAL-OLKIN, GAUSSIAN AND STUDENT-T, EACH DEFINED BY THE FUNCTION $C_c(u,v|\theta)$ $(c=1,\ldots,8)$, ALONG WITH THE $\theta(\tau)$ DEPENDENCIES AND τ-INTERVALS. HERE $\phi^{-1}(t)$ AND $t_\nu^{-1}(t)$ DENOTE THE QUANTILE FUNCTIONS OF A STANDARD UNIVARIATE NORMAL AND A STANDARD UNIVARIATE t_ν DISTRIBUTIONS RESPECTIVELY.

Copula	$C_c(u,v\|\theta)$	$\theta(\tau)$ dependence	τ interval
Clayton	$(u^{-\theta}+v^{-\theta}-1)^{-1/\theta}$	$\theta=\frac{2\tau}{1-\tau}$	$\tau\in(0,1]$
Ali-Mikhail-Haq	$\frac{uv}{1-\theta(1-u)(1-v)}$	$\tau=\frac{3\theta-2}{3\theta}-\frac{2}{3}\left(1-\frac{1}{\theta}\right)^2\log(1-\theta)$	$\tau\in[-0.1817,0.3333]$
Gumbel	$\exp\left(-\left[(-\log(u))^\theta+(-\log(v))^\theta\right]^{1/\theta}\right)$	$\theta=\frac{1}{1-\tau}$	$\tau\in[0,1]$
Frank	$-\frac{1}{\theta}\log\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$\tau=1-\frac{4}{\theta^2}\int_0^\theta\frac{t}{e^{-t}-1}dt$	$\tau\in[-1,0)\cup(0,1]$
A12	$\left(1+\left[(u^{-1}-1)^\theta+(v^{-1}-1)^\theta\right]^{1/\theta}\right)^{-1}$	$\theta=\frac{2}{3-3\tau}$	$\tau\in[0.3334,1]$
A14	$\left(1+\left[(u^{-1/\theta}-1)^\theta+(v^{-1/\theta}-1)^\theta\right]^{1/\theta}\right)^{-\theta}$	$\theta=\frac{1+\tau}{2-2\tau}$	$\tau\in[0.3334,1]$
FGM	$uv(1+\theta(1-u)(1-v))$	$\theta=\frac{9}{2}\tau$	$\tau\in[-0.2222,0.2222]$
Marchal-Olkin	$\min\left(u^{1-\theta}v,uv^{1-\theta}\right)$	$\theta=\frac{2\tau}{\tau+1}$	$\tau\in[0,1]$
Gaussian	$\int_{-\infty}^{\phi^{-1}(u)}\int_{-\infty}^{\phi^{-1}(v)}\frac{1}{2\pi\sqrt{1-\theta^2}}\exp\left(\frac{2\theta xy-x^2-y^2}{2(1-\theta^2)}\right)dxdy$	$\theta=\sin\left(\frac{\pi}{2}\tau\right)$	$\tau\in(-1,1)$
Student-t	$\int_{-\infty}^{t_\nu^{-1}(u)}\int_{-\infty}^{t_\nu^{-1}(v)}\frac{1}{2\pi\sqrt{1-\theta^2}}\left\{1+\frac{x^2-2\theta xy+y^2}{\nu(1-\theta^2)}\right\}^{-(\nu+2)/2}dxdy$	$\theta=\sin\left(\frac{\pi}{2}\tau\right)$	$\tau\in(-1,1)$

- **E-step:** compute, for each greylevel z and i-th component, the posterior probability estimates corresponding to the current pdf estimates, i.e. $z=0,\ldots,Z-1$:

$$\tau_i^t(z)=\frac{P_i^t p_i^t(z)}{\sum_{j=1}^{K^t}P_j^t p_j^t(z)},\quad i=1,\ldots,K^t;$$

- **S-step:** sample a component label $s^t(z)\in\{1,\ldots,K^t\}$ of each greylevel z according to the current estimated posterior probability distribution $\{\tau_i^t(z):i=1,\ldots,K^t\}$, $z=0,\ldots,Z-1$;
- **MoLC-step:** for the i-th mixture component, compute the following histogram-based estimates of the mixture proportions and the first three log-cumulants:

$$P_i^{t+1}=\frac{\sum_{z\in Q_{it}}h(z)}{\sum_{z=0}^{Z-1}h(z)},\quad \kappa_{1i}^t=\frac{\sum_{z\in Q_{it}}h(z)\ln z}{\sum_{z\in Q_{it}}h(z)},$$
$$\kappa_{bi}^t=\frac{\sum_{z\in Q_{it}}h(z)(\ln z-\kappa_{1i}^t)^b}{\sum_{z\in Q_{it}}h(z)},\quad i=1,\ldots,K^t,$$

  where $b=2,3$ and $Q_{it}=\{z:s^t(z)=i\}$ is the set of grey levels assigned to the i-th component by the S-step; then, solve the corresponding MoLC equations (see Table I) for each parametric family $f_j(\cdot|\theta_j)$ $(\theta_j\in A_j)$ in the dictionary, thus computing the resulting MoLC estimate θ_{ij}^t, $j=1,\ldots,4$;
- **K-step:** for each $i=1,\ldots,K^t$, if P_i^{t+1} is below a given threshold, eliminate the i-th component and update K^{t+1};
- **Model Selection-step:** for the i-th mixture component, compute the log-likelihood of each estimated pdf $f_j(\cdot|\theta_{ij}^t)$ according to the data assigned to the i-th component:

$$L_{ij}^t=\sum_{z\in Q_{it}}h(z)\ln f_j(z|\theta_{ij}^t),\quad i=1,\ldots,K^{t+1},$$

  and define $p_i^{t+1}(\cdot)$ as the estimated pdf $f_j(\cdot|\theta_{ij}^t)$ yielding the highest value of L_{ij}^t, $j=1,\ldots,4$.

B. Copulas

In this section, we present the employed dictionary of copulas and the novel copula selection procedure. An overview of bivariate copula concepts is presented in Appendix A.

We used copulas to model the joint distributions of a dual-pol SAR image, given estimates of the related marginal distributions [31]. As mentioned earlier, many parametric models have been proposed for marginal statistics of SAR amplitudes or intensities, but only a few models are available for the joint distribution of several SAR amplitudes, e.g. the Nakagami-Gamma model developed in [18]. In order to overcome this limitation, we combine the marginal pdfs provided by DSEM by means of copulas.

A bivariate copula is a 2-variate distribution defined on $[0,1]^2$ such that marginal distributions are uniform on $[0,1]$. In this paper, we consider several parametric families of one-parameter copulas. By taking advantage of the connection of copulas with the Kendall's tau ranking coefficient τ (see Appendix A), we can obtain a closed-form relationship between τ and the copula parameter θ for each parametric copula. Given a sample-rank-statistics estimate of τ (Appendix A), an estimate of θ can be obtained by inverting the corresponding equation.

In order to fully exploit the modeling potential of copulas, we consider a dictionary $\mathcal{D}_C=\{C_1,\ldots,C_{10}\}$ of 10 uni-parametric copulas $C_c(u,v|\theta)$ $(0\le u,v\le 1;c=1,\ldots,10)$: 6 Archimedean (Clayton, Ali-Mikhail-Haq, Gumbel, Frank, A12, A14) [31], a copula with a quadratic section (Farlie-Gumbel-Morgenstern) [31], 2 elliptical (Gaussian and Student-

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

t[1]) [44] and a non-Archimedean copula with simultaneous presence of an absolutely continuous and a singular component (Marchal-Olkin) [31]. Here the names for A12 and A14 originate from their positions in the list of Archimedean copulas in [31]. This choice of copulas is capable of modeling a wide variety of dependence structures, and covers most copula applications [45]. The considered dictionary of copulas is summarized in Table II (further information can be found in [31], [44], [45]).

The choice for every class $m = 1, \ldots, M$ of a specific copula C_m^* from the dictionary (see Eq. (2)) consists of the following two steps. First, given an estimator $\hat{\tau}$ of τ (computed with the training samples of the m-th class), one should decide whether a specific copula is appropriate for modeling the dependence with a corresponding value of Kendall's tau rank correlation. Indeed, some copulas are specific to marginals with a low level of dependence, others deal with strongly dependent marginals, and still others are capable of modeling all levels of dependency. In other words, for each class, the list of copulas is limited to those which are capable of accurately modeling the specific empirically estimated value $\hat{\tau}$ (see Table II). Thus, we first discard the copulas for which the current sample estimate $\hat{\tau}$ is outside the related τ-relevance interval; for the remaining copulas, we derive an estimate $\hat{\theta}$ of the related parameter by the above-mentioned Kendall's tau method.

Second, for each class, we choose the copula with the highest p-value in a Pearson chi-square test-of-fitness (PCS) [46]. In general, PCS tests the null hypothesis that the frequency distribution of certain events observed in a sample is consistent with a particular theoretical distribution. The chi-square statistic is constructed as follows:

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{6}$$

where O_i and E_i are the observed and the hypothetical frequencies, respectively, and n the number of outcomes. PCS is one of the statistical tests whose results are a chi-square distribution [46], i.e. $X^2 \sim \chi^2_{n-r-1}$, where r is the number of reductions of degrees of freedom (typically, the number of parameters for parametric CDFs). The reference to χ^2 distribution allows to calculate p-values for several different null hypothesis.

In our case, the null hypothesis in PCS is that the sample frequencies of the pairs of the form $(F_{1m}(y_1|\omega_m), F_{2m}(y_2|\omega_m))$, as (y_1, y_2) varies in the set of training samples of the m-th class, are consistent with the theoretical frequencies (probabilities) predicted by the parametric copula C_c ($m = 1, \ldots, M; c = 1, \ldots, 10$). This is correct because if $\mathbf{x}$ is distributed with CDF $F(\cdot)$, and $\mathbf{x}_1, \cdots, \mathbf{x}_N$ are independent observations of $\mathbf{x}$, then $F(\mathbf{x}_i)$, $i = 1, \cdots, N$, are independent $[0, 1]$-uniformly distributed random variables [46].

[1]Unlike all other considered copulas, the Student-t copula depends on two parameters: the linear correlation coefficient θ and the number of degrees of freedom ν. To avoid a cumbersome $\hat{\nu}$ ML estimation [44] we employ the following approach: we consider separately $\nu = 3k$, with k from 1 to 9 due to the fact that as ν grows large ($\nu \geqslant 30$) Student-t copula becomes indistinguishable from a Gaussian copula [44]. In other words, instead of a single biparametric Student-t copulas we consider here nine uniparametric Student-t copulas with fixed values of ν.

IV. Experiments

A. Datasets for experiments

The two versions of the algorithm, CoDSEM-MRF and DSEM-MRF, were tested on the following HR SAR datasets:

- Dual-pol HH/VV **TerraSAR-X**, Stripmap (6.5 m ground resolution), geocorrected, 2.66-look image acquired over Sanchagang, China, in the framework of an epidemiology monitoring application (©Infoterra GmbH). We present experiments on the 1000×1200 subimage TSX1 (see Fig. 1) and on the 750×750 subimage TSX2 (see Fig. 2).
- Single-pol HH, 1-look **COSMO-SkyMed** (CSK®), Stripmap (2.5 m ground resolution), acquired over Piemonte, Italy (©ASI). We present experiments on the 700×1000 subimage CSK1 (see Fig. 3).

B. Experimental settings

The experiments involved the classification of humid regions into $M = 3$ classes: "water", "wet", and "dry soil".

The following experimental settings were used. In the **DSEM-step**, given the relative homogeneity of the target M classes on the experiment datasets, an average value of the initial number of components $K_0 = 3$ was selected, thus assuming mixtures with just a few components for the classes. The threshold for component elimination on the K-step was 0.005 (as discussed in [24], the choice of this value only marginally affects the results). In the **Copula-step**, the following settings in the PCS test were used. The $[0, 1] \times [0, 1]$ square was divided into 25 equal squares $\mathcal{S}_i$ by 4 horizontal and 4 vertical lines parallel to the axes (i.e., $n = 25$). For every class $m = 1, \ldots, M$ and every cluster $\mathcal{S}_i$, $i = 1, \ldots, n$:

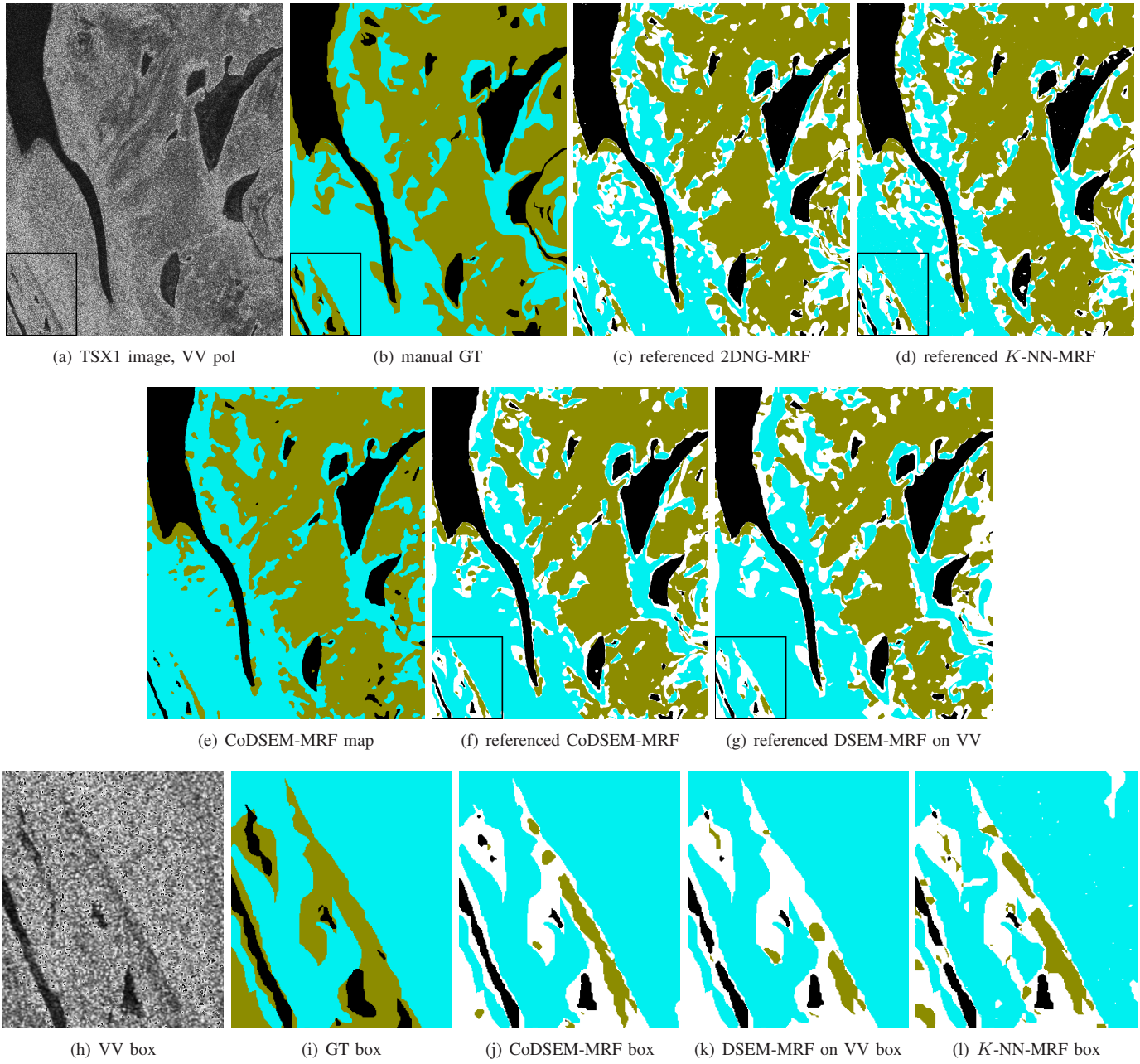
$$O_{im} = \sum I_{\mathcal{S}_i} \left[F_{1m}(y_1|\omega_m), F_{2m}(y_2|\omega_m) \right],$$

where I is the indicator function and the summation is taken over all the training pixels available for class m. For each copula $c = 1, \ldots, 10$ in $\mathcal{D}_C$, the value of E_{im} was calculated by integrating the related pdf over the square $\mathcal{S}_i$, thus retrieving the probability that a pair of $[0, 1]$-uniform random variables, whose joint distribution is defined by the c-th copula, belong to $\mathcal{S}_i$ ($i = 1, \ldots, n$). In the **Energy minimization step**, the revisit scheme on Step 3 of MMD was implemented raster-wise and the following parameter values were used : $T^0 = 5.0$, $\alpha = 0.3$, $\tau = 0.97$, $\gamma = 10^{-4}$ and n_1 was set equal to the size of the image in pixels. Compared to ICM, MMD generated significantly better results (above 5% of accuracy gain), starting at ML initialization (maximizing Eq. (2) at each pixel). SA obtained slightly better results (roughly a 1% accuracy gain) compared to MMD. However, SA generated a drastic increase in computational complexity (around 200 iterations for MMD and over 1500 for SA).

For every dataset, the proposed algorithm was trained on a small 250×250 sub-image endowed with a manually annotated non-exhaustive ground truth (GT), that did not overlap with the test areas.

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.



(a) TSX1 image, VV pol (b) manual GT (c) referenced 2DNG-MRF (d) referenced K-NN-MRF

(e) CoDSEM-MRF map (f) referenced CoDSEM-MRF (g) referenced DSEM-MRF on VV

(h) VV box (i) GT box (j) CoDSEM-MRF box (k) DSEM-MRF on VV box (l) K-NN-MRF box

Fig. 1. (a) TSX1 image (1000×1200) in VV polarization and (h) its subimage (240×270). (b) and (i) manually created ground truth (GT), (e) CoDSEM-MRF classification map (water ■, wet soil ■, dry soil ■). Classification maps referenced to the GT: (c) 2DNG-MRF, (d) and (l) K-NN-MRF, (f) and (j) CoDSEM-MRF, (g) and (k) DSEM-MRF on the VV pol image (correctly classified water ■, wet soil ■, dry soil ■, misclassification of all types □).

C. Experimental results

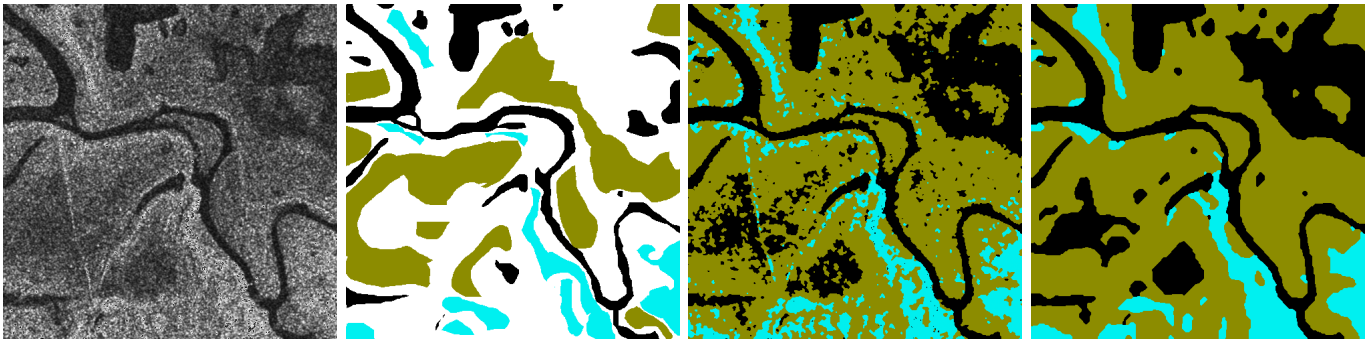
Table III reports the accuracy obtained on the test samples for each class ("class accuracy"), the resulting average accuracy (i.e., the arithmetic mean of the class accuracies) and overall accuracy (i.e., the percentage of correctly classified test samples, irrespective of their classes). The first test image TSX1 (see Fig. 1) covers an area of a river delta in Sanchagang, China. For this experiment, an exhaustive ground truth was manually created, and the classification results were referenced to it in order to calculate the accuracies and demonstrate the results visually. The automatically selected copulas were: Gumbel for the "water", and Frank for the "wet soil" and "dry soil" classes. The β estimation provided $\beta^* = 1.408$ and the overall classification accuracy achieved by CoDSEM-MRF was 84.5%. The resulting CoDSEM-MRF classification map is shown in Fig. 1(e) and referenced to the GT in Fig. 1(f).

In order to appreciate the classification gain from dual-pol data (HH/VV), compared to single-channel data, we performed an experiment of DSEM-MRF classification on the VV channel, which reported an overall accuracy of 80.6%. On average, the overall accuracy gain from adding the second polarization channel on the considered datasets was $3 \div 7\%$.

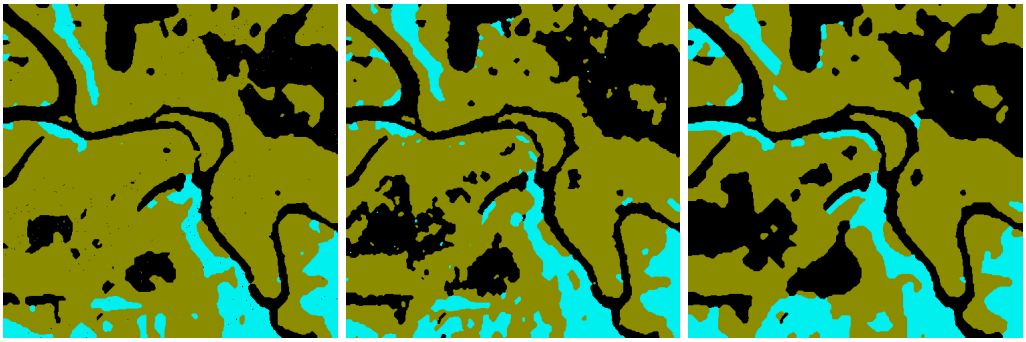
For the sake of comparison, we also provide classifica-

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.



(a) TSX2 image, HH pol (b) nonexhaustive GT map (c) CoDSEM-MRF map, $\beta = 0.5$ (d) CoDSEM-MRF map, $\beta^* = 1.324$



(e) K-NN-MRF map (f) 2DNG-MRF map (g) DSEM-MRF on HH pol map

Fig. 2. (a) TSX2 image (750×750) in HH polarization, (b) nonexhaustive ground truth (GT) map (water ■, wet soil ■, dry soil ■, outside GT □) and classification maps (water ■, wet soil ■, dry soil ■): CoDSEM-MRF with (c) manually set $\beta = 0.5$ and (d) automatically estimated $\beta^* = 1.324$, (e) K-NN-MRF, (f) 2DNG-MRF and (g) DSEM-MRF on HH pol.

tion maps obtained by the benchmark 2D Nakagami-Gamma (2DNG) model [18] (Fig. 1(c)) and K-nearest neighbors (K-NN) method [34] (Fig. 1(d)). In order to perform a fair comparison among contextual methods, the above models were combined with the contextual MRF approach. The same MRF parameter estimation and energy minimization procedures as in CoDSEM-MRF were considered in these benchmark experiments as well. For the 2DNG model, the equivalent number of looks was set to $L = 2.66$, and for the K-NN model, $K^* = 40$ was estimated by cross-validation [34]. The overall accuracies of these two benchmark approaches were 81.5% for 2DNG-MRF and 81.9% for K-NN-MRF on the dual-pol HH/VV TSX1 image. These accuracies were lower than the overall classification accuracy achieved by CoDSEM-MRF. K-NN-MRF had higher accuracy for the "wet soil" class, however this was achieved at the expense of accuracy for the "dry soil." A visual comparison of the zoomed area in Fig. 1(h)-(l) confirms these comments.

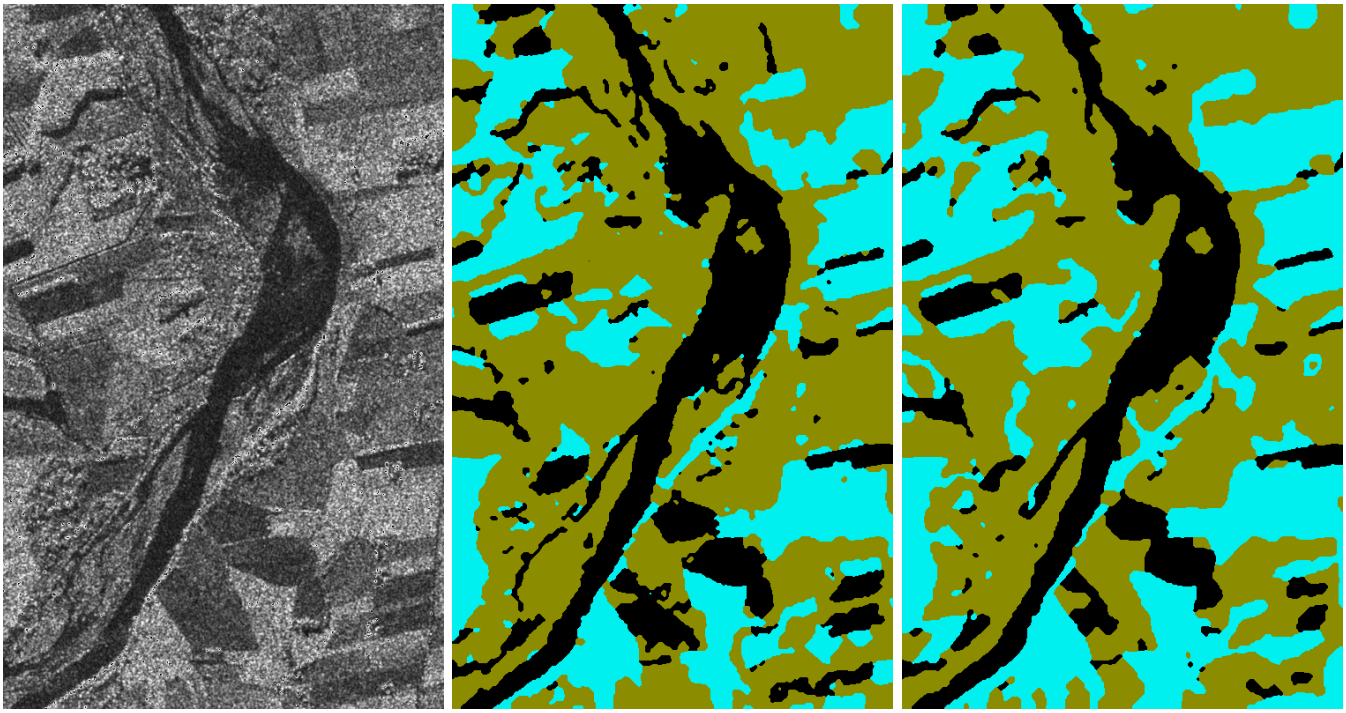
TABLE III
CLASSIFICATION ACCURACIES ON THE CONSIDERED TEST IMAGES: CLASS ACCURACIES, AVERAGE ACCURACIES, AND OVERALL ACCURACIES.

Image	Method	Water	Wet soil	Dry soil	Average	Overall
TSX1	CoDSEM-MRF	90.02%	82.56%	84.80%	85.79%	**84.55%**
	2DNG-MRF	89.33%	82.12%	76.51%	82.78%	**81.49%**
	K-NN-MRF	87.15%	87.81%	71.45%	82.14%	**81.95%**
	DSEM-MRF VV	90.00%	69.93%	91.28%	83.74%	**80.61%**
TSX2	CoDSEM-MRF, β^*	92.48%	94.59%	85.16%	90.49%	**92.41%**
	CoDSEM-MRF, β	91.52%	89.31%	77.48%	86.11%	**87.98%**
	2DNG-MRF	92.59%	89.33%	86.01%	89.31%	**89.74%**
	K-NN-MRF	90.21%	98.56%	78.91%	89.23%	**92.61%**
	DSEM-MRF HH	93.93%	83.50%	90.27%	90.23%	**87.11%**
CSK1	DSEM-MRF	95.57%	92.22%	92.45%	93.41%	**93.14%**
	K-root-MRF	89.64%	87.78%	96.83%	91.53%	**90.94%**

The second test image TSX2 (see Fig. 2) was taken from the same dataset as TSX1. It employed the same learning image and, thus, the same set of copulas were selected. The overall classification accuracy obtained by CoDSEM-MRF was 92.4% with $\beta^* = 1.324$. This accuracy is higher than that reported for the TSX1 image, however, this increase is mostly due to non-exhaustive test map employed in this experiment, which did not contain much of the class transition regions where misclassified pixels may be visually spotted. The single channel DSEM-MRF classification on HH-pol image(Fig. 2(g)) provided the overall accuracy of 87.1%, which is about 5% less than that of CoDSEM-MRF. Again, we provide comparisons with benchmark 2DNG-MRF (Fig. 2(f)) and K-NN-MRF (Fig. 2(e)) classification approaches. Like in the TSX1 image test, here CoDSEM-MRF outperformed appreciably 2DNG-MRF (89.7% of overall accuracy). K-NN-MRF with 92.6%, however, reported about the same level of overall accuracy as CoDSEM-MRF. Here we make the same observation as for TSX1 image: K-NN-MRF performed better on "wet soil" and far worse on "dry soil". We notice also

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.



(a) CSK1 image (b) DSEM-MRF map (c) $\mathcal{K}$-root-MRF map

Fig. 3. (a) CSK1 image (500×800) in HH polarization, and classification results (water ■, wet soil ■, dry soil ■): (b) DSEM-MRF and (c) $\mathcal{K}$-root-MRF classification maps.

that the average accuracy of K-NN-MRF is inferior to that of CoDSEM-MRF on both TSX1 and TSX2. Moreover, from the methodological point of view, CoDSEM is preferable to K-NN as it provides an explicit description of a statistical model for the data, whereas the latter operates as a "black box".

A slight oversmoothing effect can be noticed in the results obtained by the developed method, see Fig. 2(d). Therefore, for the TSX2 image, we also show the classification map obtained by manually setting the MRF parameter $\beta = 0.5$ (Fig. 2(c)). One can see that the the spatial details are more precise at the expense of a noisier segmentation. The classification accuracies achieved with $\beta = 0.5$ (see Table III) are inferior to those obtained in the $\beta^* = 1.324$ case. This suggests that, at least for this dataset, a stronger regularization is overally more preferable.

The third experiment was conducted on a single-pol CSK1 image (see Fig. 3(a)). It also employed a non-exhaustive GT. The overall classification accuracy of DSEM-MRF (Fig. 3(b)) was 93.1% with $\beta^* = 1.566$. We compare this with the $\mathcal{K}$-root-MRF contextual classification approach that is based on a $\mathcal{K}$-root model [25] for each class-conditional statistics. The $\mathcal{K}$ distribution is a well-known model for a possibly textured SAR multilook single-channel intensity and $\mathcal{K}$-root is the corresponding amplitude parametric pdf. Overall classification accuracy reported in this experiment (Fig. 3(c)) was equal to 90.9%. The improved performance of DSEM-MRF is due to more accurate pdf estimates by DSEM compared to $\mathcal{K}$-root model.

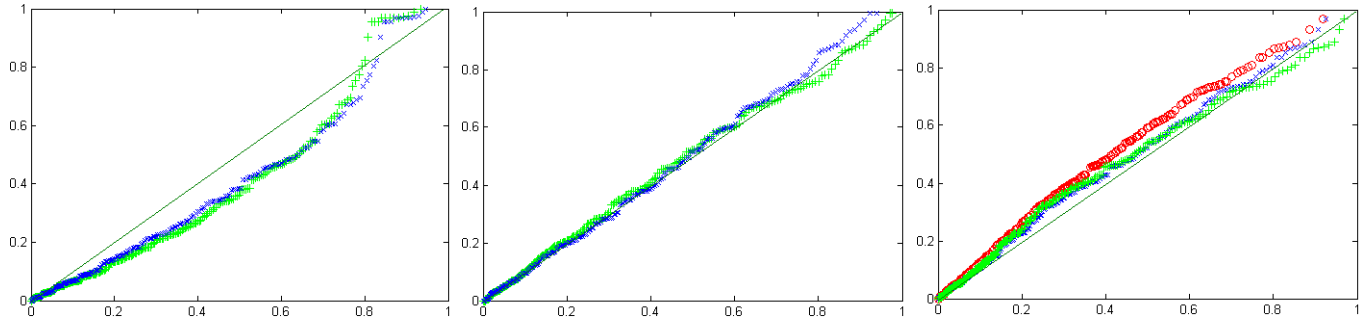
The restriction of the DSEM dictionary $\mathcal{D}_M$ to four models compared to eight constituting the dictionary in [24] (see Sec. III(a)) resulted in a negligible decrease in modeling accuracy. In other words, this restriction affected the classification accuracy marginally, yet allowed a significant gain in calculation speed.

In addition, we want to demonstrate the accuracy of the Copula-DSEM joint pdf model directly. To do so we present the K-plots [47] of the sample HH-VV dependence and the dependence estimated by copulas (see Fig. 4) on the learning image for TSX1. K-plots is a tool for graphical goodness-of-fit presentations and its definition is recalled in Appendix B. The demonstrated copulas were automatically selected from the dictionary $\mathcal{D}_M$ and one can see a good agreement between observations and the model.

Let us now briefly compare the developed CoDSEM-MRF approach with the previously proposed copula-based method [21]. Firstly, the approach [21] is based on the use of the Ali-Mikhail-Haq (AMH) copula [31], which can model dependencies corresponding to Kendall's correlation coefficient $\tau \in [-0.1817, 0.3333]$. On the employed TSX datasets we have observed empirical values $\hat{\tau}_{\text{water}} \approx 0.36, \hat{\tau}_{\text{wet}} \approx 0.39, \hat{\tau}_{\text{dry}} \approx 0.27$. Therefore, the AMH copula can be used only for the dry soil class and its K-plot is presented in Fig. 4(c). However, even in this case, the goodness-of-fit provided by the automatically selected Frank copula is better. Thus, the use of dictionary-based copula selection approach is more accurate and constitutes a more flexible model with a wider range of applicability. Secondly, the use of the DSEM finite mixture estimation approach for marginal pdf estimation provides higher accuracy estimates than singular pdf models the use of which was suggested in [21], especially for HR

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.



(a) Water: Gumbel copula, $\hat{\tau} = 0.358887$ (b) Wet soil: Frank copula, $\hat{\tau} = 0.393471$ (c) Dry soil: Frank and AMH copulas, $\hat{\tau} = 0.265$

Fig. 4. K-plots display the graphical goodness-of-fit for copula models on the learning stage for TSX1 image. For the considered three classes the sample HH-VV dependences (+) are fitted by automatically selected copulas (×). The K-plot for the Ali-Mikhail-Haq (AMH) copula (o) is also presented for the dry soil class (c). The diagonal (–) represents the independency scenario.

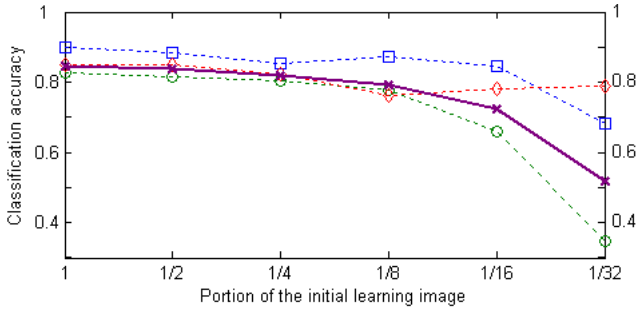


Fig. 5. The sensitivity of CoDSEM-MRF classification accuracy to the size of the learning image. The class accuracies (water □, wet soil ○, dry soil ◇) and overall accuracies (solid line, **x**) are reported for TSX1 image. The initial learning image is 250×250 pixels.

SAR images [24].

It is well known, that the acquisition of the training sets required for the supervised classification is a very costly procedure. Therefore, the classification models are designed to be as robust as possible with respect to small learning sets. To evaluate this characteristic of the proposed algorithm, we present an experimental study (see Fig. 5) of the classification accuracy as a function of size of the employed training set. We start with the learning image of 250×250 pixels, and then iteratively reduce its size by a half at each step, i. e. by discarding 50% randomly selected training pixels for each thematic class, untill $1/32$ of the initial learning image is present. At each step we evaluate the class and overall accuracies of classification on TSX1. To make the accuracy estimates more consistent, the whole process was repeated three times and the averaged results are presented in Fig. 5. We observe that the algorithm behaves robustly till $1/8$ (90×90 pixels) of the initial learning image is kept. When very strong subsampling rates ($\geqslant 16$) are applied, the wet area classification rate drops significantly. This is due to the fact that the histogram of wet soil lies "between" those of water and dry soil. Thus, when the training set becomes unrepresentatively small, a lot of wet soil pixels are misclassified as dry soil. This result suggests the capability of the developed algorithm to perform fairly good and consistently on small training sets, i.e. up to 100×100 training images for the considered dataset.

The most time consuming stages of the algorithm are the **MRF** and **Energy minimization** steps. Given the fairly low value $K_0 = 3$, the **DSEM** step was very fast. The **Copula** and **MRF** steps do not involve any intensive computation procedures. The experiments were conducted on a Core 2 Duo 1.83GHz, 1Gb RAM, WinXP system. With the number of iterations equal to 200, the β-estimation on a roughly 1000×1000 image took around 80 seconds. The average of 200 iterations required for convergence on MMD (**Energy minimization** step) took about 100 seconds. Thus about 200 seconds were required for the complete classification process of a 1000×1000 dual-pol image with three classes.

V. Conclusions

Contemporary satellite SAR missions, such as TerraSAR-X and COSMO-SkyMed, are capable of providing high resolution imagery. In this paper, a novel supervised classification algorithm for HR single channel and dual polarization satellite SAR imagery has been developed. It combines the Markov random field approach to Bayesian image classification and a finite mixture technique for probability density function estimation. Finite mixture modeling was done via dictionary-based stochastic expectation maximization for amplitude pdf estimation, which provides a high level of estimation accuracy for the algorithm [24]. These two concepts allowed us to formulate a classification algorithm for the case of single-channel SAR imagery. The second contribution of this paper is the introduction of a novel statistical approach based on a dictionary of copulas for the problem of dual polarization SAR image classification. This approach enables a new level of flexibility in modeling the joint distribution from marginal single-channel distributions, which results in higher classification accuracy. Together with Markov random fields and finite mixture approaches, these three statistical concepts ensure high flexibility and applicability of the developed method to dual-pol HR SAR image classification.

The proposed classification algorithm is supervised and semiautomatic (a few DSEM and MMD parameters have to be specified). The accuracy of the proposed algorithm was validated for water/wet soil/dry soil classification on two high resolution satellite SAR images (a dual-pol TerraSAR-X image

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

and a single-pol COSMO-SkyMed image). The experiments demonstrated a high level of accuracy on the experiment datasets and outperformed the considered parametric and non-parametric contextual benchmark algorithms.

Finally, we would like to point out several directions of further development of this work. First of all, a very interesting topic would be the generalization of the developed copula-based approach to multi-polarized imagery and its application to quad-pol TerraSAR-X imagery. The theory of copulas allows an extension of bivariate theory to multivariate cases, thus preserving the same copula types and parameter estimation techniques. However, contrary to the symmetric copulas approach considered here, which implied equal "contributions" of all marginal channels in the joint distribution, non-symmetric copulas would have to be considered for multi-polarization scenario. Another interesting direction of work is to explore the use of more efficient optimization approaches. A very promising candidate to replace MMD is an appropriate graph cut approach [48]. Such approaches yield very good approximations in the MAP segmentation problem and are known to be fast. The third direction of development lies in the specialization of this model to urban area classification, which is an important and relevant application of SAR classification. To this end, the contextual MRF model would also need to incorporate geometrical information.

APPENDIX A
BIVARIATE COPULA THEORY

A *bivariate copula* is a function $C : [0,1]^2 \to [0,1]$, which satisfies the following properties:

1. both marginals are uniformly distributed on $[0,1]$;
2. for every u,v in $[0,1]$: $C(u,0) = C(0,v) = 0$, and $C(u,1) = u$, $C(1,v) = v$;
3. a 2-increasing property: $\forall u_1 \leqslant u_2$, $v_1 \leqslant v_2$ in $[0,1]$: $C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geqslant 0$.

The importance of copulas in statistics is explained by Sklar's Theorem [31], which states the existence of a copula C, that models the joint distribution function H of arbitrary random variables X and Y with CDFs F and G:

$$H(x,y) = C(F(x), G(y)), \tag{7}$$

for all x, y in $\mathbb{R}$. If, in addition, F and G are continuous, then C is unique. Thus, copulas link joint distribution functions to their one-dimensional marginals.

Given absolutely continuous random variables with pdfs $f(x)$ and $g(y)$ and corresponding CDFs $F(x)$ and $G(y)$, the pdf of the joint distribution $h(x,y)$ corresponding to (7) is given by:

$$h(x,y) = f(x)g(y)\frac{\partial^2 C}{\partial x \partial y}(F(x), G(y)), \tag{8}$$

where $\frac{\partial^2 C}{\partial x \partial y}(x,y)$ is pdf corresponding to the copula $C(x,y)$.

An important family of copulas are Archimedean copulas, which have a simple analytical form and yet provide a wide variety of modeled dependence structures. An *Archimedean* copula is a bivariate copula C, defined as:

$$C(u_1,u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)),$$

where the *generator function* $\phi(u)$ is any function satisfying the following properties:

1. $\phi(u)$ is continuous on $[0,1]$;
2. $\phi(u)$ is decreasing, $\phi(1) = 0$;
3. $\phi(u)$ is convex.

A common way of copula estimation is by using its connection with Kendall's tau, which is a ranking correlation coefficient [31]. *Kendall's tau* is a concordance-discordance measure between two independent realizations (X,Y) and $(\hat{X},\hat{Y})$ from the same CDF $H(x,y)$:

$$\tau = \text{Prob}\{(X-\hat{X})(Y-\hat{Y}) > 0\} - \text{Prob}\{(X-\hat{X})(Y-\hat{Y}) < 0\}.$$

Given realizations (x_l, y_l), $l = 1,\ldots,N$, the empirical estimator of Kendall's tau is given by:

$$\hat{\tau} = \frac{4}{N(N-1)} \sum_{i \neq j} I[x_i \leqslant x_j] I[y_i \leqslant y_j] - 1. \tag{9}$$

By integrating in the definition of τ over the distribution of $(\hat{X},\hat{Y})$, we get the general relationship between Kendall's τ and the copula C associated with $H(x,y)$, expressed by the following Lebesgue-Stieltjes integral:

$$\tau + 1 = 4 \int_0^1 \int_0^1 C(u,v) dC(u,v), \tag{10}$$

In the specific case of Archimedean copulas, the relationship is expressed in terms of generator function $\phi(t)$:

$$\tau = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt. \tag{11}$$

By plugging the empirical estimate $\hat{\tau}$ in place of τ in Eq. (10) or (11), we get parameter estimates $\hat{\theta}$ (see Table II).

APPENDIX B
GRAPHICAL GOODNESS-OF-FIT BY K-PLOTS

A K-plot (from Kendall-plot) [47] is a rank based graphical tool developed for visualization of dependence structure between two random variables X and Y. This technique is based on plotting the pairs $(W_{i:N}, H_{(i)})$ for $i = 1,\ldots,N$:

- $H_{(i)}$ are defined as $H_{(1)} \leqslant H_{(2)} \leqslant \cdots \leqslant H_{(N)}$, i.e. the order statistics of quantities
$$H_i = \frac{1}{N-1} \#\{j \neq i : x_i \leqslant x_j,\ y_i \leqslant y_j\}.$$
- $W_{i:N}$ represents the expectation of the ith order statistic in a random sample of size N from the distribution $K_0(\omega) = \text{Prob}\{H(X,Y) \leqslant \omega\}$ of the H_i under the null hypothesis of independency. It writes [47]:
$$W_{i:N} = N \binom{N-1}{i-1} \times \int_0^1 \omega \{K_0(\omega)\}^{i-1} \{1 - K_0(\omega)\}^{N-i} dK_0(\omega), \tag{12}$$
where $K_0(\omega) = \omega - \omega \log \omega$.

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

In order to compare this sample dependence representation with the one provided by a given copula we need to find the form of $K_\theta(\omega) = \text{Prob}\{C_\theta(U,V) \leqslant \omega\}$ specific to this copula, where $U \sim F(X), V \sim G(Y)$. It can be found as [49]

$$K_\theta(\omega) = \omega - \frac{\phi_\theta(\omega)}{\phi'_\theta(\omega)}, \quad \omega \in (0,1),$$

for an Archimedean copula with generator $\phi_\theta(\cdot)$. Replacing then $K_0(\omega)$ by $K_\theta(\omega)$ in (12) we can also draw a plot $(\widehat{W}_{i:N}, H_{(i)})$ that corresponds to a given copula and where θ is estimated via Kendall's tau (see Appendix A). Finally, the generators for Gumbel and Frank copulas are given by [31]:

$$\phi_{\text{Gumbel}}(\omega) = (-\ln \omega)^\theta, \quad \phi_{\text{Frank}}(\omega) = -\ln \frac{e^{-\theta\omega} - 1}{e^{-\theta} - 1}.$$

ACKNOWLEDGMENT

The authors would like to thank the Italian Space Agency for providing the COSMO-SkyMed (CSK®) image of Piemonte (COSMO-SkyMed Product - ©ASI - Agenzia Spaziale Italiana - 2008. All Rights Reserved). The TerraSAR-X image of Sanchagang was taken from *http://www.infoterra.de/*, where it is available for free testing (©Infoterra GmbH, 2008). The authors would also like to thank the anonymous reviewers for their helpful and constructive comments.

REFERENCES

[1] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, 2nd ed. NC, USA: SciTech, Raleigh, 2004.

[2] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. New York: CRC Press, 2009.

[3] J.-S. Lee, M. R. Grunes, and R. Kwok, "Classification of multilook polarimetric SAR imagery based on complex Wishart distribution," *Int. J. Remote Sensing*, vol. 15, no. 11, pp. 2299–2311, 1994.

[4] J.-M. Beaulieu and R. Touzi, "Segmentation of textured polarimetric SAR scenes by likelihood approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2063–2072, 2004.

[5] A. C. Frery, C. C. Freitas, and A. H. Correia, "Classifying multifrequency fully polarimetric imagery with multiple sources of statistical evidence and contextual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3098–3109, 2007.

[6] A. P. Doulgeris, S. N. Anfinsen, and T. Eltoft, "Classification with a non-Gaussian model for PolSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2999–3009, 2008.

[7] L. Bombrun, G. Vasile, M. Gay, and F. Totir, "Hierarchical segmentation of polarimetric SAR images using heterogeneous clutter models," *IEEE Trans. Geosci. Remote Sens.*, 2010, (to appear). DOI:10.1109/TGRS.2010.2060730.

[8] Y. Ito and S. Omatu, "Polarimetric SAR data classification using competitive neural networks," *Int. J. Remote Sensing*, vol. 19, no. 14, pp. 2665–2684, 1998.

[9] K. S. Chen, W. P. Huang, D. H. Tsay, and F. Amar, "Classification of multifrequency polarimetric SAR imagery using a dynamic learning neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 3, pp. 814–820, 1996.

[10] C. Lardeux, P.-L. Frison, C. Tison, J.-C. Souyris, B. Stoll, B. Fruneau, and J.-P. Rudant, "Support vector machine for multifrequency SAR polarimetric data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4143–4152, 2009.

[11] C. T. Chen, K. S. Chen, and J.-S. Lee, "The use of fully polarimetric information for the fuzzy neural classification of SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 2089–2100, 2003.

[12] P. R. Kersten, J.-S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 519–527, 2005.

[13] J. Morio, F. Goudail, X. Dupuis, P. Dubois-Fernandez, and P. Refregier, "Polarimetric and interferometric SAR image partition into statistically homogeneous regions based on the minimization of the stochastic complexity," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3599–3609, 2007.

[14] K. Ersahin, I. Cumming, and R. Ward, "Segmentation and classification of polarimetric SAR data using spectral graph partitioning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 164–174, 2010.

[15] G. D. De Grandi, J. S. Lee, and D. L. Schuler, "Target detection and texture segmentation in polarimetric SAR images using a wavelet frame: Theoretical aspects," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3437–3453, 2007.

[16] 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, 1997.

[17] J.-S. Lee, M. R. Grunes, E. Pottier, and L. Ferro-Famil, "Unsupervised terrain classification preserving polarimetric scattering characteristics," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 4, pp. 722–731, 2004.

[18] J.-S. Lee, K. W. Hoppel, S. A. Mango, and A. R. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1017–1028, 1994.

[19] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, and L. M. Novak, "K-distribution and polarimetric terrain radar clutter," *J. Electromagn. Waves Applicat.*, vol. 3, pp. 747–768, 1989.

[20] C. C. Freitas, A. C. Frery, and A. H. Correia, "The polarimetric G distribution for SAR data analysis," *Environmetrics*, vol. 16, no. 1, pp. 13–31, 2005.

[21] G. Mercier and P.-L. Frison, "Statistical characterization of the Sinclair matrix: Application to polarimetric image segmentation," in *Proceedings of IGARSS*, Cape Town, South Africa, 2009, pp. III–717–III–720.

[22] Z. Kato, J. Zerubia, and M. Berthod, "Unsupervised parallel image classification using Markovian models," *Pattern Recognit.*, vol. 32, no. 4, pp. 591–604, 1999.

[23] G. Moser, S. Serpico, and J. Zerubia, "Dictionary-based Stochastic Expectation Maximization for SAR amplitude probability density function estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 1, pp. 188–199, 2006.

[24] V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "Enhanced dictionary-based SAR amplitude distribution estimation and its validation with very high-resolution data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 148–152, 2011.

[25] E. Jakeman and P. N. Pusey, "A model for non-Rayleigh sea echo," *IEEE Trans. Antennas Propagat.*, vol. 24, pp. 806–814, 1976.

[26] Y. Delignon, A. Marzouki, and W. Pieczynski, "Estimation of generalized mixtures and its application in image segmentation," *IEEE Trans. Image Process.*, vol. 6, no. 10, pp. 1364–1375, 1997.

[27] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre, "A new statistical model for Markovian classification of urban areas in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2046–2057, 2004.

[28] E. E. Kuruoglu and J. Zerubia, "Modelling SAR images with a generalization of the Rayleigh distribution," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 527–533, 2004.

[29] G. Moser, J. Zerubia, and S. B. Serpico, "SAR amplitude probability density function estimation based on a generalized Gaussian model," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1429–1442, 2006.

[30] H.-C. Li, W. Hong, and Y.-R. Wu, "Generalized Gamma distribution with MoLC estimation for statistical modeling of SAR images," in *Proceedings of APSAR*, Huangshan, China, 2007, pp. 525–528.

[31] R. B. Nelsen, *An Introduction to Copulas*, 2nd ed. New-York: Springer, 2006.

[32] G. Mercier, S. Derrode, W. Pieczynski, J.-M. Nicolas, A. Joannic-Chardin, and J. Inglada, "Copula-based stochastic kernels for abrupt change detection," in *Proceedings of IGARSS*, Denver, USA, 2006, pp. 204–207.

[33] G. Mercier, G. Moser, and S. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, 2008.

[34] C. M. Bishop, *Pattern Recognition and Machine Learning*. New-York: Springer, 2006.

[35] V. Krylov and J. Zerubia, "High resolution SAR image classification," INRIA, Research Report 7108, 2009. [Online]. Available: http://hal.archives-ouvertes.fr/docs/00/44/81/40/PDF/RR-7108.pdf

[36] J. Besag, "On the statistical analysis of dirty pictures," *J. Royal Stat. Soc. B*, vol. 48, pp. 259–302, 1986.

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

[37] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 6, pp. 721–741, 1984.

[38] W. Hastings, "Monte Carlo sampling method using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[39] M. Abramowitz and I. Stegun, Eds., *Handbook of Mathematical Functions*. New York: Dover, 1964.

[40] Z. Kato, J. Zerubia, and M. Berthod, "Satellite image classification using a modified Metropolis dynamics," in *Proceedings of ICASSP*, 1992, pp. 573–576.

[41] M. A. F. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, 2002.

[42] G. Celeux, D. Chauveau, and J. Diebolt, "Stochastic versions of the EM algorithm: an experimental study in the mixture case," *J. Statist. Comp. Sim.*, vol. 55, no. 4, pp. 287–314, 1996.

[43] J.-M. Nicolas, "Introduction aux statistiques de deuxième espèce: applications des logs-moments et des logs-cumulants à l'analyse des lois d'images radar," *Traitement du Signal (in french)*, vol. 19, no. 3, pp. 139–167, 2002.

[44] G. Frahm, M. Junker, and A. Szimayer, "Elliptical copulas: Applicability and limitations," *Stat. Prob. Lett.*, vol. 63, pp. 275–1286, 2003.

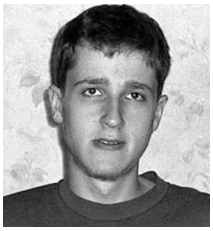
[45] D. Huard, G. Évin, and A.-C. Favre, "Bayesian copula selection," *Comput. Statist. Data Anal.*, vol. 51, no. 2, pp. 809–822, 2006.

[46] E. Lehmann and J. Romano, *Testing statistical hypotheses*, 3rd ed. New York: Springer, 2005.

[47] C. Genest and J.-C. Boies, "Detecting dependence with Kendall plots," *Amer. Statistician*, vol. 57, no. 4, pp. 275–284, 2003.

[48] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222 –1239, 2001.

[49] C. Genest and L.-P. Rivest, "Statistical inference procedures for bivariate Archimedean copulas," *J. Am. Stat. Assoc.*, vol. 88, pp. 1034–1043, 1993.



Vladimir A. Krylov received the specialist degree in applied mathematics from the Lomonosov Moscow State University, Moscow, Russia, in 2007, where he is currently working towards the PhD degree in mathematics at the Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics. In 2009 and 2010 he has spent several months with Ariana research group, Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France, as a visiting PhD student.

His main research interest lies in the field of statistical image analysis and processing. In particular, his work concerns statistical modeling, classification and multitemporal change detection of remote sensing imagery.



Gabriele Moser (S'03–M'05) received the "laurea" (M.S.) degree (*summa cum laude*) in telecommunications engineering and the Ph.D. degree in space sciences and engineering from the University of Genoa, Genoa, Italy, in 2001 and 2005, respectively. Since 2001, he has cooperated with the Signal Processing and Telecommunications Research Group (SP&T) of the Department of Biophysical and Electronic Engineering (DIBE), University of Genoa, in the field of remote-sensing image analysis. From January to March 2004, he was a visiting student at the Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France, working with the Ariana research group on the problem of SAR data modeling. He is currently Assistant Professor of Telecommunications at the University of Genoa.

His research activity is focused on image-processing and pattern-recognition methodologies for remote-sensing data interpretation. In particular, his current research interests include SAR data analysis, multitemporal image classification, hyperspectral image analysis, contextual classification, and geo/biophysical parameter estimation. He has been a reviewer for several international journals. Since 2008, he has been an Associate Editor of the IEEE Geoscience and Remote Sensing Letters.



Sebastiano B. Serpico (M'87-SM'00-F'08) received the Laurea degree in electronic engineering and the Doctorate from the University of Genoa, Italy, in 1982 and 1989, respectively. Full Professor of telecommunications at the Faculty of Engineering of the University of Genoa since 2004, he is currently the Head of the Signal Processing and Telecommunications laboratory of the Department of Biophysical and Electronic Engineering with the University of Genoa.

Dr. Serpico is the Chairman of the Institute of Advanced Studies in Information and Communication Technologies (ISICT). He has been the project manager of numerous national and international projects and an evaluator of project proposals for various programmes of the European Union. He is author (or coauthor) of more than 200 scientific articles published in journals and conference proceedings. He is an Associate Editor of the international journal IEEE Transactions on Geoscience and Remote Sensing (TGRS). He was a guest editor of the Special Issues of TGRS on "Analysis of hyperspectral image data" (July 2001) and "Advances in techniques for analysis of remotely sensed data" (March 2005). From 1998 to 2002 he was the chairman of a SPIE/EUROPTO series of conferences on Signal and Image Processing for Remote Sensing. In 2003 he was one of the chairmen of the IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data (NASA Goddard Visitor Center, Greenbelt MD, USA). He held seminars and invited talks at the leading universities and research centers of various countries, such as France, Germany, UK, USA, Japan, Norway, and Iceland.

His current research interests include pattern recognition for remote sensing images and for biomedical images.



Josiane Zerubia (S'78–M'81-SM'99-F'03) received the M.Sc. degree from the Department of Electrical Engineering at Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble, Grenoble, France, in 1981, and the Dr.Eng. degree, the Ph.D. degree, and the "Habilitation," all from the University of Nice, Sophia-Antipolis, France, in 1986, 1988, and 1994, respectively. She has been a permanent research scientist at l'Institut National de Recherche en Informatique et en Automatique (INRIA) since 1989, and director of research since July 1995. She was head of the PASTIS remote sensing laboratory (INRIA Sophia-Antipolis) from mid-1995 to 1997. Since January 1998, she has been head of the Ariana research group (INRIA/CNRS/University of Nice), which also works on remote sensing. She has been professor at SUPAERO (ISAE) in Toulouse since 1999. Before that, she was with the Signal and Image Processing Institute of the University of Southern California (USC) in Los-Angeles as a postdoc. She also worked as a researcher for the LASSY (University of Nice/CNRS) from 1984 to 1988 and in the Research Laboratory of Hewlett Packard in France and in Palo-Alto (CA) from 1982 to 1984.

Dr. Zerubia is a member of the IEEE IMDSP and IEEE BISP Technical Committees (SP Society). She was associate editor of IEEE Trans. on IP from 1998 to 2002; area editor of IEEE Trans. on IP from 2003 to 2006; guest co-editor of a special issue of IEEE Trans. on PAMI in 2003; and member-at-large of the Board of Governors of the IEEE SP Society from 2002 to 2004. She has also been a member of the editorial board of the French Society for Photogrammetry and Remote Sensing (SFPT) since 1998, of the International Journal of Computer Vision since 2004, and of the Foundation and Trends in Signal Processing since 2007. She has been associate editor of the on-line resource: Earthzine (IEEE CEO and GEOSS). She was co-chair of two workshops on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR'01, Sophia Antipolis, France, and EMMCVPR'03, Lisbon, Portugal); co-chair of a workshop on Image Processing and Related Mathematical Fields (IPRM'02, Moscow, Russia); chair of a workshop on Photogrammetry and Remote Sensing for Urban Areas, Marne La Vallee, France, 2003; and co-chair of the special sessions at IEEE ICASSP 2006 (Toulouse, France) and at IEEE ISBI 2008 (Paris, France). She is a member of the organizing committees of IEEE ICIP 2011 (Brussels, Belgium) and IEEE ICIP 2014 (Paris, France).

Her current research interests are in image processing using probabilistic models and variational methods. She also works on parameter estimation and optimization techniques.

Copyright (c) 2010 IEEE. Personal use is permitted. For any other purposes, Permission must be obtained from the IEEE by emailing pubs-permissions@ieee.org.