Bayesian off-line detection of multiple change-points corrupted by multiplicative noise: application to SAR image edge detection

Jean-Yves Tourneret[a,*], Michel Doisy[a], Marc Lavielle[b]

[a] *IRIT/ENSEEIHT/TéSA, 2 rue Camichel, BP 7122, 31071 Toulouse Cedex 7, France*
[b] *Université Paris Sud, 91405 Orsay Cedex, France*

Abstract

This paper addresses the problem of Bayesian off-line change-point detection in synthetic aperture radar images. The minimum mean square error and maximum a posteriori estimators of the changepoint positions are studied. Both estimators cannot be implemented because of optimization or integration problems. A practical implementation using Markov chain Monte Carlo methods is proposed. This implementation requires a priori knowledge of the so-called hyperparameters. A hyperparameter estimation procedure is proposed that alleviates the requirement of knowing the values of the hyperparameters. Simulation results on synthetic signals and synthetic aperture radar images are presented.

Keywords: SAR image segmentation; Change-point detection; MCMC methods

1. Introduction

Increasing interest is being shown in many signal processing applications for change-point estimation and detection. These applications include segmentation, fault detection or monitoring (for an overview see [2] and references therein). Of course, the problems of estimating and detecting change-points have received much attention in the signal processing and statistical literature. For example, the generalized likelihood ratio (GLR) detector proposed in [47] has shown interesting properties for these problems [1,2]. Some shortcomings of the GLR detector were eliminated by considering the change-point amplitudes as nuisance parameters and by using marginalization [22]. Marginalization is a common procedure in Bayesian inference, which has also been widely considered for change-point estimation and detection. The segmentation of non-stationary signals which can be represented by autoregressive processes in independent segments is studied in [14]. The off-line segmentation of signals using a large sample approximation of the MAP criterion is studied in [12]. The derivation is general in the sense that it is valid for signals that can be parametrized by linear or nonlinear functions embedded in additive possibly non-Gaussian and colored noise. However, instead of pursuing the exact MAP solution, a MAP approximation based on asymptotic Bayesian theory is studied. The

* Corresponding author. Tel.: +33-5-61-58-84-77.
E-mail addresses: jean-yves.tourneret@tesa.prd.fr (J.-Y. Tourneret), doisy@len7.enseeiht.fr (M. Doisy), marc.lavielle@math.u-psud.fr (M. Lavielle).

intractability of the a posteriori distributions for the unknown change-point parameters has led to some interesting approaches based on Markov Chain Monte Carlo (MCMC) methods [19]. For example, the identification of multiple change-points in linearly modelled data was investigated in [33]. In this study, the parameter space is partitioned in three subspaces: linear coefficients, noise parameters and change-points which are sampled using the Gibbs sampler. A Gibbs sampling approach to Bayesian inference for single change-point problems and its extension to multiple change-points were also presented by Carlin et al. [6] and Stephens [36]. A new parametrization of the change-point model and an associated MCMC algorithm were recently studied in [7].

Most of the previous studies have been carried out to detect changes in signals contaminated by additive noise, i.e., to detect mean shifts in the observed time series. Consequently, the proposed algorithms cannot be used when signals are corrupted by non-additive interferences. Some authors have solved this problem by considering more general change-point models including random level-shift models [27,28] or multiplicative noise models [15,34,38,39]. This paper addresses the problem of change-point detection in multiplicative noise models. The problem has received much attention for edge detection in synthetic aperture radar (SAR) images. Indeed, because of the multiplicative speckle noise, most standard edge detectors such as gradient-based detectors or Bayesian detectors [13,14,37] perform poorly when applied to SAR images. Touzi et al. [43] studied a statistical and geometrical edge detector for SAR images. This detector denoted the ROA detector was based on a ratio of averages (ROA). It was shown in [43] to outperform the gradient detector, the Sobel detector and the Frost et al. detector [16]. A new edge detector for SAR images was recently studied in [15]. This detector denoted the ROEWA detector performed a line-by-line and column-by-column change-point detection, by computing the ratio of exponentially weighted averages (ROEWA) on opposite sides of the central pixel in the horizontal and vertical directions. These averages were computed by filtering the image intensity by the infinite symmetric exponential filter. The use of the infinite symmetric exponential filter was motivated by the fact that it yields the best unbiased linear reflectivity estimator (which minimizes the mean square error between the real image reflectivity and the filtered noisy image reflectivity). The ROEWAs in the horizontal and vertical directions were then combined to yield an Edge Strength Map (ESM). A high pixel value in the ESM indicates the presence of an edge at this position. Finally, local maxima were extracted and attributed to edges using the watershed algorithm [46]. Fjortoft et al. showed that the ROEWA detector outperforms the ROA detector in a multiedge context [15].

The main contribution of this paper is to adapt the Bayesian detector proposed by Lavielle in [25] to the edge detection problem in SAR images. An interesting hyperparameter estimation procedure is also studied. The noisy SAR images are modelled as piecewise constant fields corrupted by multiplicative speckle noise. The edge detection is performed off-line, line-by-line and column-by-column as in [13,15,37]. Note that the on-line approaches studied by Basseville [2] are not appropriate for SAR image segmentation, since the whole image is available. Moreover, as emphasized in [37], in image segmentation, *a retrospective scheme is more attractive as it reflects the global rather than local aspects of the edge detection problem*. In a Bayesian framework, the unknown change-point parameters are estimated using their a posteriori distribution via the minimum mean square error (MMSE) or marginal maximum a posteriori (MMAP) estimators. These estimators are optimal in the sense that they minimize an appropriate cost function [45, p. 55]. It is interesting to note that change-point parameters could also be estimated using the maximum likelihood (ML) method [41]. However, the resulting maximum likelihood estimator (MLE) has serious limitations, especially when multiple abrupt changes occur. For instance, the MLE is sensitive to over-parametrization [33, p. 9]. Unfortunately, the implementation of the MMSE and MMAP estimators is difficult in a multiedge situation. MCMC methods are then used to simulate the posterior distribution of the change-point positions and to compute the estimates.

The SAR image model is described briefly in Section 2. Section 3 studies the change-point instant MMSE and MAP estimators. The simulation of a posteriori change-point instant distributions using MCMC methods is discussed in Section 4. The problem of hyperparameter estimation is discussed

in Section 5. Simulation results and conclusions are reported in Sections 6 and 7.

2. Signal model

The *complex* SAR image (also referred to as *parent process* [24]) is computed after the SAR system receives the coherent sum of reflected monochromatic microwaves. The magnitude and magnitude-squared of this complex image are denoted *amplitude* SAR image and *intensity* SAR image, respectively. SAR amplitude images can be obtained in several ways (1) by averaging L amplitude images, (2) by averaging L intensity images and then taking the square root, or (3) coherently averaging complex images to obtain an intensity image and then taking the square root [23]. This paper studies edge detection in SAR images produced via the second approach.

2.1. Speckle statistics for a single pixel

The complex SAR image at position (x, y) denoted $z(x, y)$ (or z for brevity) is usually modelled by a complex zero-mean circular Gaussian variable (considering the very large number of image cells in the radar field of view and invoking the central limit theorem [4] [10, p. 215] with probability density function (pdf):

$$f(z) = \frac{1}{2\pi\sigma^2(x,y)}\, e^{-|z|^2/2\sigma^2(x,y)}, \quad z(x,y) \in \mathbb{C}. \qquad (1)$$

The *intensity* image at position (x, y) is defined as $w(x, y) = |z(x, y)|^2$. A standard change of variables shows that $w(x, y)$ is exponentially distributed with mean $E[w(x, y)] = 2\sigma^2(x, y)$ [21]. Consequently, the intensity image at position (x, y) can be written as the product of the terrain reflectivity at position (x, y) denoted $m(x, y) = 2\sigma^2(x, y)$ and a random variable $\tilde{w}(x, y)$ (exponentially distributed with parameter $E[\tilde{w}(x, y)] = 1$) which is independent of $m(x, y)$, i.e., $w(x, y) = m(x, y)\tilde{w}(x, y)$. To reduce the speckle variance, several independent images denoted "looks" are usually averaged. When L images are averaged, the resulting intensity at a pixel can be written as $v(x, y) = 1/L \sum_{j=1}^{L} w_j(x, y)$ (the index j denotes the jth look) whose pdf is [29, p. 95]:

$$f(v) = \left(\frac{L}{m}\right)^L \frac{v^{L-1}}{\Gamma(L)} \exp\left(-\frac{Lv}{m}\right), \quad v \geqslant 0, \qquad (2)$$

where $\Gamma(t) = \int_0^{+\infty} u^{t-1} \mathrm{e}^{-u} \mathrm{d}u$ is the standard Gamma function and L (the number of looks) is assumed to be known in the rest of the paper. It is interesting to note that the variable $v(x, y)$ can be written as a product of the real image intensity at position (x, y), that is $m(x, y)$ and a random variable $b(x, y) = 1/L \sum_{j=1}^{L} \tilde{w}_j(x, y)$ which is independent of $m(x, y)$. This property explains the terminology "multiplicative noise" for $b(x, y)$ and assumes (as in [15,16]; [29, p. 95]) (1) that the transfer function of the SAR system does not vary significantly over the bandwidth of interest, (2) that the additive measurement noise can be neglected. The statistical properties of $b(x, y)$ can be easily derived from Eq. (2): the variable $b(x, y)$ is distributed according to a Gamma distribution with parameters L and L [31, p. 381]. The mean and variance of this distribution are $E[b(x, y)] = 1$ and $\mathrm{Var}[b(x, y)] = 1/L$, which shows the speckle noise reduction due to the averaging of L independent images.

2.2. Speckle statistics for a line of the SAR image intensity

Denote N as the number of pixels in a line of the SAR image intensity and T as the sampling period. Using the single pixel statistics (described in Section 2.1), a line of the SAR image intensity can be modelled as follows:

$$v_n = b_n m_n, \quad n = 1, \ldots, N, \qquad (3)$$

where $b_n = b(nT), m_n = m(nT), v_n = v(nT)$ are the multiplicative speckle noise, the uncorrupted and corrupted line of the SAR image intensity respectively. The properties of b_n and m_n can be defined as follows:

- the autocovariance function of the speckle may decrease very rapidly. In this case, the speckle noise sequence b_n can be approximated by an *independent identically distributed* (iid) sequence of random variables with Gamma distribution whose parameters are L and L [15]; [29, p. 99]; [44, p. 1914],
- The uncorrupted line of the SAR image intensity m_n can be modelled by K steps, when K fields with different reflectivities are considered. This model referred to as the Cartoon Model [29, p. 197] is a

good approximation for important scene types such as agricultural fields. Denote l_{i-1} (with $l_0 = 0$ and $l_K = N$) as the sample point after which there is the ith sudden change in the signal ($i = 1,\ldots,K$). In the following, the integers l_{i-1} will be referred to as change locations and the corresponding actual change locations are $t_{i-1} = l_{i-1}T + \tau$, with $0 < \tau < T$. The uncorrupted line of the SAR image can then be defined by

$$m_n = A_i, \quad n \in]l_{i-1}, l_i], \qquad i = 1,\ldots,K, \quad (4)$$

where $A_i > 0$ is the ith step amplitude.

The line-by-line edge detection problem consists of estimating the change-point locations l_i for $i \in \{1,\ldots,K-1\}$ from the observed data v_n. This edge detection problem is crucial in image segmentation. Once the change-point locations have been estimated, the line of the SAR image can be recovered by estimating the change-point amplitudes A_i. Consequently, the edge detection problem can also be used to recover the radar reflectivity (ideal image without speckle) (this problem is currently referred to as speckle filtering).

3. Bayesian inference

3.1. Likelihood

The likelihood function of the observed data $v = (v_1,\ldots,v_N)^t$ (where t denotes transposition), conditioned upon the change-point locations $l = (l_1,\ldots,l_{K-1})^t$ and amplitudes $A = (A_1,\ldots,A_K)^t$ is defined by

$$\begin{aligned} f(v|K,A,l) &= \prod_{k=1}^{K} \prod_{i=l_{k-1}+1}^{l_k} \left(\frac{L}{A_k}\right)^L \\ &\quad \times \frac{v_i^{L-1}}{\Gamma(L)} \exp\left(-\frac{Lv_i}{A_k}\right) \\ &\propto \prod_{k=1}^{K} \frac{1}{A_k^{Ln_k}} \exp\left(-\sum_{k=1}^{K} \frac{LS_k}{A_k}\right) \end{aligned} \quad (5)$$

with $n_k = l_k - l_{k-1}$ and $S_k = \sum_{i=l_{k-1}+1}^{l_k} v_i$. The problem can be reparameterized by introducing binary variables defined by

$$\begin{aligned} r_j &= 1 \quad \text{if there is a} \\ &\qquad \text{change-point at pixel } j, \quad j = 1,\ldots,N-1. \\ r_j &= 0 \quad \text{otherwise,} \end{aligned} \quad (6)$$

Conventionally, $r_N = 1$ such that the number of step changes equals the number of steps denoted as $K(r) = \sum_{j=1}^{N} r_j$ with $r = (r_1,\ldots,r_{N-1})^t$. The likelihood function of v can then be rewritten as

$$f(v|\theta) \propto \exp\left(-L \sum_{k=1}^{K(r)} \left\{ \frac{S_k(r)}{A_k} + n_k(r) \log A_k \right\}\right), \quad (7)$$

where "$\propto$" means "proportional to", $n_k(r) = l_k(r) - l_{k-1}(r)$, $\theta = (r, A)$ and $S_k(r) = \sum_{i=l_{k-1}(r)+1}^{l_k(r)} v_i$.

3.2. Parameter priors

The choice of priors in Bayesian inference is important and has received much attention in the literature [5, p. 183]; [3, p. 264]; [35]. This study uses the following priors for the change-point detection problem:

- Independent *Bernoulli priors* are chosen for the change-point locations:

$$f(r) = \lambda^{K(r)-1}(1-\lambda)^{N-K(r)}, \quad \text{where } r \in \{0,1\}^{N-1}. \quad (8)$$

  The parameter $\lambda \in]0,1[$ is the Bernoulli parameter which represents the a priori probability of having a change-point at a given position.
- Independent *inverted-gamma* (*IG*) *priors* (denoted as $A_i \sim \mathrm{IG}(\alpha,\gamma)$) are chosen for the step amplitudes:

$$f(A|r) = \prod_{i=1}^{K(r)} \frac{\gamma^\alpha}{\Gamma(\alpha) A_i^{\alpha+1}} \exp\left(-\frac{\gamma}{A_i}\right) \times I_{[0,+\infty[}(A_i), \quad (9)$$

  where $\alpha > 0$ and $\gamma > 0$ are two constants, and $I_{[0,+\infty[}(.)$ is an indicator function ($I_{[0,+\infty[}(t) = 1$ if $t \in [0,+\infty[$ and $I_{[0,+\infty[}(t) = 0$ if $t \in]-\infty,0[$). Suitable choices of parameters α and γ allow to

incorporate either very vague or more specific prior information about the step amplitude (see discussions in [20]). The motivation for choosing the IG prior (whose main properties can be found in [3, p. 119] or [20]) is that the IG belongs to the conjugate family of priors for A with respect to the likelihood $f(v|\theta)$ [3, p. 265]. In other words, $f(A|r)$ has the same "structure" as $f(v|\theta)$, when $f(v|\theta)$ is viewed as a function of A. This yields analytically tractable integration of $f(\theta|v)$ with respect to A_i, i.e., allows marginalization.

3.3. MMSE and MAP estimators

Using Bayes' theorem, we can express the parameter posterior pdf as

$$f(\theta|v) \propto f(v|\theta)f(\theta) = f(v|A,r)f(A|r)f(r), \quad (10)$$

where $f(v|\theta)$ has been defined in Eq. (7) and $f(\theta) = f(A|r)f(r)$ is the *a priori* distribution for $\theta = (r, A)$. Edge detection only requires the estimation of the change-point vector r. Consequently, the so called "nuisance parameters" A_i can be eliminated by integrating out A_i from the posterior pdf (10). Some straightforward computations allow to obtain the marginal a posteriori pdf of r:

$$f(r|v) = C(v,L)\lambda^{K(r)-1}(1-\lambda)^{N-K(r)}\frac{\gamma^{\alpha K(r)}}{\Gamma(\alpha)^{K(r)}} \times \prod_{k=1}^{K(r)} \frac{\Gamma(\alpha + Ln_k(r))}{(\gamma + LS_k(r))^{Ln_k(r)+\alpha}} \quad (11)$$

with $C(v,L) = (L^L/\Gamma(L))^N \prod_{i=1}^N v_i^{L-1}$. Equivalently, the marginal pdf of r can be written as $f(r|v) \propto \exp(-U(r|v))$ where

$$U(r|v) = \beta K(r) + \sum_{k=1}^{K(r)} \log \frac{(\gamma + LS_k(r))^{Ln_k(r)+\alpha}}{\Gamma(\alpha + Ln_k(r))} \quad (12)$$

is referred to as the energy function and $\beta = \log[(1-\lambda)/\lambda] - \log[\gamma^\alpha/\Gamma(\alpha)]$. Note that the parameter β is a decreasing function of λ. Consequently, the smaller β, the higher the a priori probability of a change and the fewer the omissions. On the other hand, the bigger β, the fewer the false alarms. The parameter β controls the resolution level of the segmentation: changes with small amplitudes will be detected for small values of β.

The unknown parameter vector r can be estimated from the posterior distribution $f(r|v)$ by minimizing the mean of an appropriate cost function [45, p. 55]. Standard Bayesian estimators are the (marginal) minimum mean square estimator (MMSE) and (marginal) maximum a posteriori (MAP) estimators:

- The MMSE estimator of r, which minimizes the quadratic cost function, is defined by

  $$\hat{r} = E[r|v].$$

  Here, since $r \in \{0,1\}^N$, the MMSE estimator of r yields the change-point a posteriori probabilities, which will be useful to define an SAR image edge strength map (see Section 7),
- The MAP estimator of r, which minimizes the 0-1 cost function, is defined by

  $$\hat{r} = \underset{r}{\operatorname{argmax}}\, f(r|v) = \underset{r}{\operatorname{argmin}}\, U(r|v). \quad (13)$$

Unfortunately, a closed-form expression of the MMSE and MAP estimators of r cannot be obtained.

4. MCMC methods

The previous section showed that a closed-form expression for MMAP or MMSE estimators of the change-point positions cannot be obtained. Indeed, both estimators suffer from optimization or integration problems. Numerical techniques based on MCMC methods can then be explored to solve these problems. The detection of change-points corrupted by additive noise using MCMC methods has received much attention in the literature [6–8,21,26,30,33,35,36]. However, to our knowledge no analysis has been provided for multiplicative noise models. The major contribution of this section is to study two simple MCMC methods for estimating change-points corrupted by multiplicative noise:

(1) The MMSE estimator of r is obtained by constructing a homogeneous Markov chain using the Metropolis-Hastings (MH) algorithm with the invariant distribution $f(r|v)$ defined in (11). The mean of (11) is then estimated by the time average of the last Markov chain output samples, which converges to the MMSE estimator according to the ergodic theorem for Markov chains.

(2) The MAP estimator of r is determined by using a Simulated Annealing (SA) algorithm. The SA algorithm defines a non-homogeneous Markov chain which converges under appropriate conditions to the minimum of the energy function $U(r|v)$, which is here the MAP estimator.

The Markov chain transition kernels are the same for both strategies. The introduction of a decreasing temperature schedule in the SA algorithm (which modifies the acceptance probability) is the only difference between the MMSE and MAP estimator algorithms. This section is organized as follows: the Markov chain jumps are described in parts (a), (b), (c) and the acceptance probabilities for the MMSE and MAP estimators are detailed in parts (1) and (2). The Markov chain state space and current state are denoted by $\Omega = \{0,1\}^{N-1}$ and $\Theta^n = (\Theta_i^n)_{i=1,\ldots,N-1} \in \Omega$, respectively. The Markov chain moves are defined as follows:

(a) The candidate $z^{n+1} \in \Omega$ is drawn independently of the current location Θ^n yielding *the independence sampler* [19] defined by $q(z^{n+1}|\Theta^n) = q(z^{n+1})$, where q is an instrumental distribution. For our experiment, q is a Bernoulli distribution with parameter λ. The parameter λ is the *a priori* probability of having a change-point at a fixed pixel, and it adjusts the mean value of the number of change-points. For example, when $N = 250$ and $\lambda = 0.02$, the mean number of change-points is $E[K(r)] = N\lambda = 5$. In this procedure, the candidate z^{n+1} is selected using the classical acceptance probability. The independence sampler allows to move rapidly to distant parts of the state space. However, the global acceptance probability for this sampler is very low for long datasets. Consequently, such sampling scheme is only used during the first iterations, and it does not increase the computational cost of the algorithm.

(b) Local changes are made following the so-called *one-variable-at-a-time* MH algorithm. This variable-at-a-time step was suggested for instance in [8] or [19, p. 10] to increase the convergence speed. More precisely, a random permutation of $\{1,\ldots,N-1\}$ is uniformly drawn. According to this permutation, each component is flipped from 0 to 1 or from 1 to 0. The move is then accepted with the usual acceptance probability. This move visits each site randomly and all sites are visited in each scan. It belongs to the class of random scan Gibbs samplers. Note that the use of random scan Gibbs samplers where the successive components are chosen at random, either independently or in a multinomial fashion (which amounts to select a random permutation), has been suggested by many authors (see [32, p. 45] or [19, p. 15]). Note also that this second step requires $N-1$ acceptance procedures.

(c) Change-points are moved in the neighborhood of their current location. This move is particularly interesting for the MAP estimator, because it is difficult to escape from a position close to a real change-point. In this move, an actual change-point is randomly selected and a neighborhood of this instant is defined. The change-point instant is finally moved in its neighborhood and accepted or not according to the acceptance probability. Such move is very important since it avoids trapping in a change-point neighborhood.

Each kernel is used in turn and the resulting *hybrid strategy* is called a *cycle*. The resulting cycle kernel is clearly irreducible and aperiodic (see [40] for details). The acceptance probabilities for the MMSE and MAP algorithms are defined as follows:

(1) In the MMSE algorithm (summarized in Appendix A), the chain is constructed to simulate the target distribution $f(r|v)$ defined in (11). At each step of the cycle, the acceptance probability is

$$\alpha(\Theta^n, z^{n+1}) = \min\left\{1, \frac{f(z^{n+1}|v)q(\Theta^n|z^{n+1})}{f(\Theta^n|v)q(z^{n+1}|\Theta^n)}\right\}, \quad (14)$$

where q is the transition probability associated with the jumps described in (a), (b) and (c), Θ^n and z^{n+1} are the Markov chain current state and candidate respectively. Equivalently, if Rand is the outcome of a uniform drawing on $[0,1]$:

$$\begin{cases} \Theta^{n+1} = z^{n+1} & \text{if Rand} < \frac{f(z^{n+1}|v)q(\Theta^n|z^{n+1})}{f(\Theta^n|v)q(z^{n+1}|\Theta^n)}, \\ \Theta^{n+1} = \Theta^n & \text{otherwise.} \end{cases} \quad (15)$$

In procedures (a), (b) and (c) the instrumental distribution q is symmetric such that $q(\Theta^n|z^{n+1}) = q(z^{n+1}|\Theta^n)$. Since $f(r|v) \propto \exp(-U(r|v))$, the acceptance in the MH algorithm defined in (15) reduces to

$$\begin{cases} \Theta^{n+1} = z^{n+1} & \text{if } U(\Theta^n|v) - U(z^{n+1}|v) > \eta, \\ \Theta^{n+1} = \Theta^n & \text{otherwise,} \end{cases} \quad (16)$$

Table 1
Acceptance probabilities for the different moves

Acceptance probabilities	Move a	Move b	Move c
Minimum	0.00022	0.727	0.365
Maximum	0.0022	0.824	0.442
Mean	0.0011	0.777	0.401

where $\eta = \ln \text{Rand}$. After a sufficiently long *burn-in*, the MMSE estimator of the change-point positions is determined by computing the time average of the Markov chain output samples.

(2) In the MAP algorithm (summarized in Appendix B), the schedule for lowering the temperature is defined by $T_k = 0.99T_{k-1}$, where T_0 is greater than a numerical constant Δ depending on the energy function $U(r|v)$ [17]. This temperature decrease is made at each step of the independence sampler (move a), at each permutation draw σ of the one-variable-at-a-time sampler (move b) and at each permutation draw τ of move c (i.e., three temperature decreases per cycle). The acceptance procedure is defined by

$$\begin{cases} \Theta^{n+1} = z^{n+1} & \text{if } U(\Theta^n|v) - U(z^{n+1}|v) > \eta T_k, \\ \Theta^{n+1} = \Theta^n & \text{otherwise,} \end{cases} \tag{17}$$

where T_k is the current temperature.

Remarks.

- Table 1 shows the minimum, maximum and average acceptance probabilities for each move, computed from 50 Monte Carlo runs (each Monte Carlo run corresponds to 50000 cycles). The independent sampler (move a) allows to move rapidly to distant parts of the state space. The table indicates that the global acceptance probability for the move a is very low. This explains why the move a is used only during the first burn-in iterations, in the proposed edge detection procedure.
- According to (12), the total energy U is a sum of local potentials. Consequently, a local perturbation of the configuration Θ^n (as in moves b and c) affects few terms of this sum, which ensures fast computations of the energy variations ΔU.
- Convergence results for the proposed MCMC method can be found in standard textbooks or many papers such as [9,17–19,26,32,40,48].

5. Hyperparameter estimation

The implementation of an MCMC algorithm as described above assumes that the set of hyperparameters of the model is known. These hyperparameters are the parameters α and γ of the IG distribution for the change-point amplitudes A_k, and the prior proportion of changes λ. Here, we propose to estimate these hyperparameters in a maximum likelihood framework, by using ideas developed in [11]. Let $\phi = (\alpha, \gamma, \lambda)$. The algorithm simply consists of inserting a step for updating ϕ in the MCMC algorithm used for simulating the a posteriori distribution of r:

- Choose an initial guess $\phi^{(0)}$ and an initial configuration of change-point instants $r^{(0)}$,
- At step j
  - Perform one iteration of MCMC using the current value of the parameters $\phi^{(j-1)}$ to simulate $r^{(j)}$ from $r^{(j-1)}$,
  - Compute the maximum likelihood estimate $T(r^{(j)})$ of ϕ by maximizing the joint distribution of $(v, r^{(j)})$ and update $\phi^{(j)}$

$$\phi^{(j)} = \phi^{(j-1)} + \eta_j (T(r^{(j)}) - \phi^{(j-1)}), \tag{18}$$

where the stepsize sequence (η_j) decreases to 0.

Remarks. (1) A decreasing sequence (η_j) is chosen in order to obtain a pointwise convergence of the sequence $\phi^{(j)}$ to a value $\phi^\star$, that will be used later for estimating the change-points locations. A satisfactory schedule consists of setting $\eta_j = 1$ during some iterations, which ensures a fast convergence to a neighborhood of $\phi^\star$. Then, η_j decreases as $1/j$.

(2) This algorithm is a slight modification of the Stochastic Approximation version of the EM algorithm (SAEM algorithm) proposed by Delyon [11]. The convergence of the SAEM algorithm to a local maximum of the observed likelihood can be proved under appropriate conditions. Some conditions are not satisfied here since (1) the joint pdf of (v, r) does not belong to the exponential family and (2) r is not sampled from its exact full distribution, but by using an MCMC procedure. A theoretical analysis of the algorithm is beyond the scope of the paper. However, we have noticed that the algorithm always converges to a value which is close to the unknown true value.

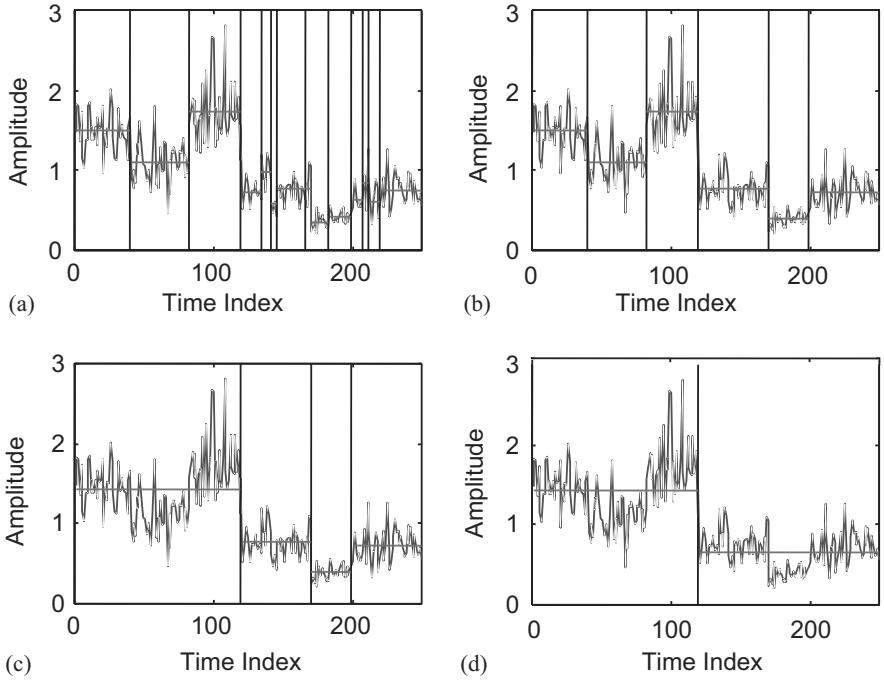

Fig. 1. Signal reconstruction using the MAP algorithm for different resolutions: (a) $\beta = -1.4$; (b) $\beta = 0.68$; (c) $\beta = 20$; and (d) $\beta = 50$.

(3) The maximum likelihood estimate of (α, γ), that maximises the joint distribution of $(v, r^{(j)})$, cannot be computed in a closed form. However, a Newton–Raphson algorithm which converges in few iterations can be used. At step j, the initial guess for this algorithm is the current value $(\alpha^{(j-1)}, \gamma^{(j-1)})$. On the other hand, it is easy to see that the maximum likelihood estimate of λ is the empirical mean $(K(r)-1)/(N-1)$.

6. Simulation results

6.1. Effect of the hyperparameter λ (or equivalently of β)

Consider a synthetic signal subject to multiple change-points with parameters A=(1.5, 1.1, 1.6, 0.8, 0.4, 0.7), $N = 250$, and $l = (0, 40, 80, 120, 170, 200, 250)$. A Markov chain with invariant distribution $f(r|v)$ is simulated on Ω. Once the MAP estimates of the change-point instants have been determined, the different signal amplitudes are estimated by maximizing $f(A|v, r)$, which allows for signal reconstruction. Straightforward computations yield

$$\hat{A}_k = \frac{S_k(r)}{n_k(r)} \frac{1 + \gamma/LS_k(r)}{1 + (\alpha+1)/Ln_k(r)}. \tag{19}$$

Note that $S_k(r)/n_k(r)$ is the standard estimated mean of v on the kth segment $]l_{k-1}, l_k]$. Eq. (19) shows that $\hat{A}_k$ approximately equals $S_k(r)/n_k(r)$ when $\gamma/LS_k(r) \ll 1$ and $(\alpha+1)/[Ln_k(r)] \ll 1$. Such conditions are satisfied when the changes do not happen too frequently, i.e. for instance for agricultural areas where the fields are big compared to the sensor resolution (pixel size). Fig. 1 shows the signal reconstructions obtained for different values of β: (a) $\beta = -1.4$, (b) $\beta = 0.68$, (c) $\beta = 20$, (d) $\beta = 50$. These figures clearly show that β controls the resolution level of the segmentation: for $\beta = 50$ (low resolution), only the largest change-point is detected, whereas for $\beta = -1.4$ (high resolution), additional change-points are detected. The effect of δ on the change-point MMSE estimator is very similar to the effect of β on the MAP estimator (see for instance [42]).

6.2. Hyperparameter estimation

Next, we study the performance of the hyperparameter estimation algorithm on the previous synthetic signal. A Markov chain with invariant distribution $f(r|v)$ is simulated on Ω and hyperparameters are estimated as described in Section 5. Fig. 2 shows the mean of the

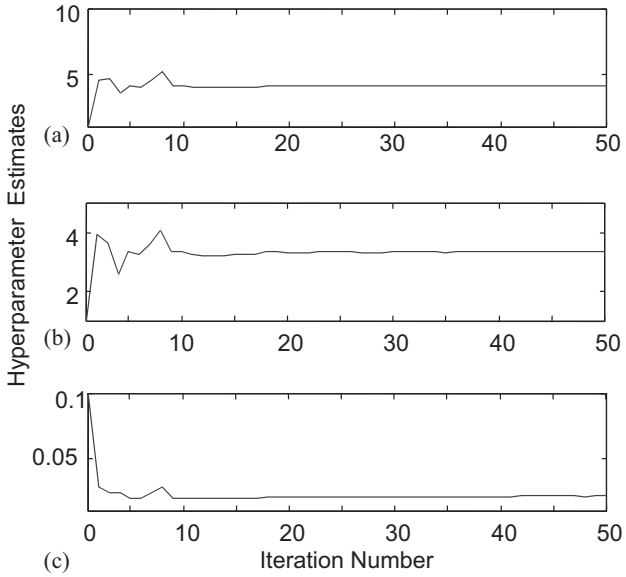

Fig. 2. Hyperparameter estimates: (a) parameter α; (b) parameter γ; and (c) parameter λ.

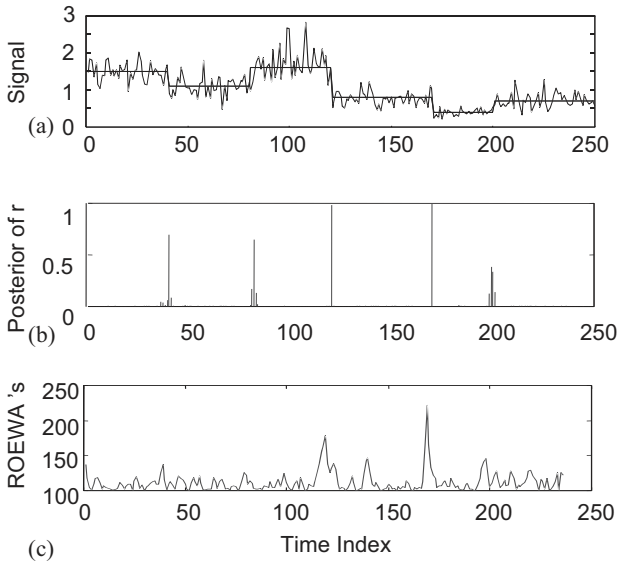

Fig. 3. Noisy signal and reflectivity (a), Changepoint a posteriori probabilities (b), and ROEWAs (c).

hyperparameter estimates computed from 50 Monte Carlo runs. The algorithm clearly converges in few iterations.

6.3. Convergence of the sampler

The vector $\bar{\Theta}=(1/N_{\mathrm{MC}})\sum_{n=N_{\mathrm{bi}}+1}^{N_{\mathrm{bi}}+N_{\mathrm{MC}}}\Theta^n$ computed for $N_{\mathrm{MC}}=10^5$ and $N_{\mathrm{bi}}=50$ is plotted in Fig. 3b. This vector estimates the a posteriori probability of an abrupt change at each lag. A change-point with large amplitude ($l=40,80,120,170$) yields a large a posteriori probability at the corresponding lag whereas there is some kind of ambiguity for a change-point with small amplitude ($l=200$). In this latter case, the algorithm switches between several lags in a neighborhood of $l=200$, because of the smoother transition in the observed signal. The a posteriori change-point probabilities shown in Fig. 3b do not significantly vary from one MCMC simulation to another, due to the high number of cycles. However, a so large number of cycles cannot be used in practical applications, where images have to be processed. Consequently, the appropriate number of required cycles has to be determined by studying the convergence of the MCMC sampler. Many convergence diagnostics can be found in the literature (see [32] and references therein). However, as specified in [20], these diagnostics are not completely

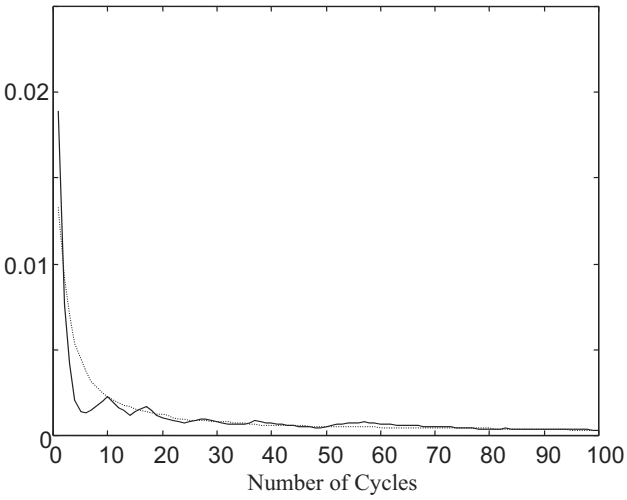

Fig. 4. MSEs between the "true" and estimated a posteriori change-point probabilities as a function of N_{MC} (for a fixed burn-in $N_{\mathrm{bi}}=50$) (solid line). Averaged MSEs computed from 50 Monte Carlo runs (dotted line).

reliable. Here, we examine the change-point probabilities obtained for different values of the hyperparameter initial values. Fig. 4 shows the MSE between the "true" (computed from 10^5 cycles) and estimated a posteriori change-point probabilities as a function of the number of cycles N_{MC} for a fixed burn-in $N_{\mathrm{bi}}=50$. The averaged MSE between the "true" and estimated

a posteriori change-point probabilities computed from 50 Monte Carlo runs is also depicted. This figure clearly shows that $N_{MC} = 50$ and $N_{bi} = 50$ are sufficient to ensure convergence of the sampler. Consequently, simulations on synthetic and real images have been conducted $N_{MC} = 50$ and $N_{bi} = 50$.

6.4. Comparison with the ROEWA edge detector

In SAR image segmentation, postprocessing algorithms such as morphological closing [43] or the watershed algorithm [46] are often used to remove false edge-points and extract closed skeleton boundaries. In these situations, the MMSE detector should be preferred to the MAP detector. Indeed, the MMSE detector yields change-point a posteriori probabilities which can be viewed as an ESM. Such ESM is well suited to postprocessing algorithms such as the watershed algorithm. The MAP estimator performs very differently since it provides the most likely a posteriori change-point configuration given the data. The ESM corresponding to the MAP estimator has on each line $K(r)$ components equal to 1 and $N - K(r)$ components equal to 0. Such ESM is not appropriate for the watershed algorithm (see comments in [15] for more details).

The MMSE change-point detector and the ROEWA detector are then compared on the same synthetic signal (which represents a line of a SAR image). Fig. 3c shows the results of the ROEWA detector obtained with the infinite symmetric exponential filter. This figure has to be compared with Fig. 3b, which shows the results obtained with the MMSE detector. For this particular synthetic signal, the MMSE detector yields better results than the ROEWA detector. Additional simulations have been conducted and they confirm that the MMSE detector provides better accuracy in change-point detection than the ROEWA detector. However, it is important to note that the ROEWA detector has lower computational cost than the proposed detector.

Simulations are then presented for a synthetic 4 look scene represented in Fig. 5. We follow the strategy used in [15] for the ROEWA detector: in order to compute the horizontal edge strength component, the image $v(x, y)$ has first to be smoothed column by column using the 1-D ROEWA filter, which yields $f(y) * v(x, y)$. The filtered image is then processed line by line to compute the exponentially weighted averages

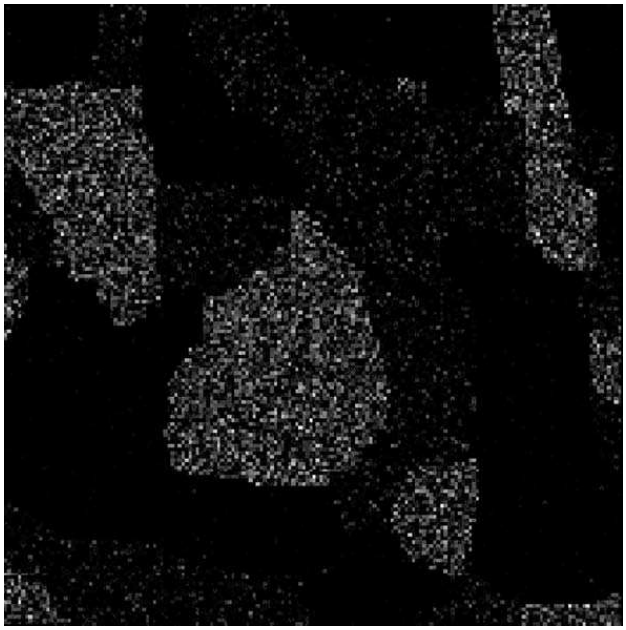

Fig. 5. Noisy 4-look 180 × 180 SAR image.

$$\hat{\mu}_{X_1}(x, y) = f_1(x) * \{f(y) * v(x, y)\},$$

$$\hat{\mu}_{X_2}(x, y) = f_2(x) * \{f(y) * v(x, y)\},$$

where f_1 and f_2 are the causal and anticausal filters associated to the ROEWA impulse response f. The averages $\hat{\mu}_{X_1}(x, y)$ and $\hat{\mu}_{X_2}(x, y)$ are then used to form the normalized ratio $r^X_{\max}(x, y)$. The vertical edge strength component $r^Y_{\max}(x, y)$ is computed similarly after computing the averages

$$\hat{\mu}_{Y_1}(x, y) = f_1(y) * \{f(x) * v(x, y)\}$$

$$\hat{\mu}_{Y_2}(x, y) = f_2(y) * \{f(x) * v(x, y)\}$$

The horizontal and vertical edge strength components are finally combined to form the 2-D ESM

$$r^{2-D}_{\max}(x, y) = \sqrt{[r^X_{\max}(x, y)]^2 + [r^Y_{\max}(x, y)]^2}.$$

For the MMSE detector, the image $v(x, y)$ is first smoothed column by column using the 1-D ROEWA filter. The MMSE detection strategy is then applied line by line on the filtered image (the change-point a posteriori probabilities are estimated for each line),

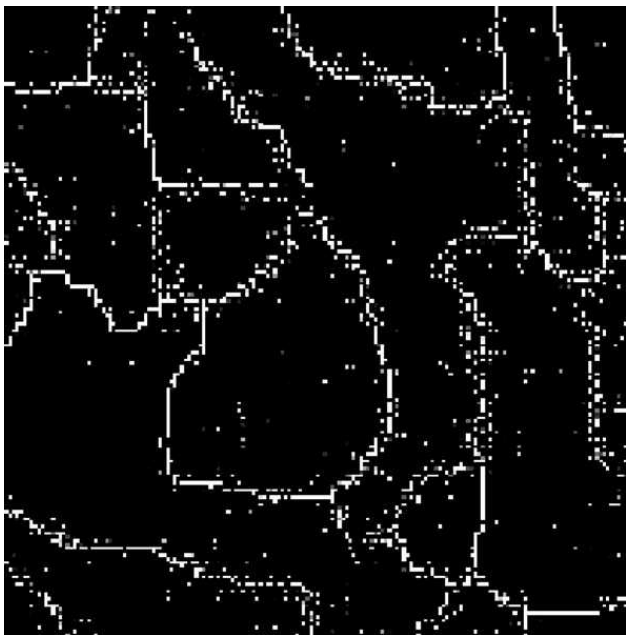

Fig. 6. SAR image edge strength map (MMSE detector).

Fig. 7. SAR image edge strength map (ROEWA detector).

which yields the horizontal ESM. The vertical ESM is computed similarly by (1) line smoothing using the 1-D ROEWA filter and (2) column by column change detection using the MMSE detector (the change-point a posteriori probabilities are estimated for each column). The 2-D ESM is then obtained by computing the magnitude of the horizontal and vertical ESM components [15]. In this ESM, a high pixel value indicates the presence of an edge. Figs. 6 and 7 show the 2-D ESMs obtained for the MMSE and ROEWA detectors. The MMSE detector seems to yield better results than the ROEWA detector for this synthetic image.

6.5. *Posterior distribution of K*

This section addresses the important question of the choice of K (number of change-points). The algorithm studied in this paper draws vectors r^i (for $i=1,\ldots,N_{\mathrm{MC}}$) distributed according to the joint distribution $f(r|v)$ defined in (11). For each vector r^i, the number of change-points is $K(r^i)=\sum_{j=1}^{N} r_j^i$. As a consequence, the posterior distribution of K can be easily estimated from the vectors r^i. Fig. 8 shows the 150th line of the synthetic image depicted in Fig. 5 and the estimated posterior distribution $\hat{f}(r|v)$ obtained with 50 burn-in cycles and $N_{\mathrm{MC}} = 100$. As can be seen,

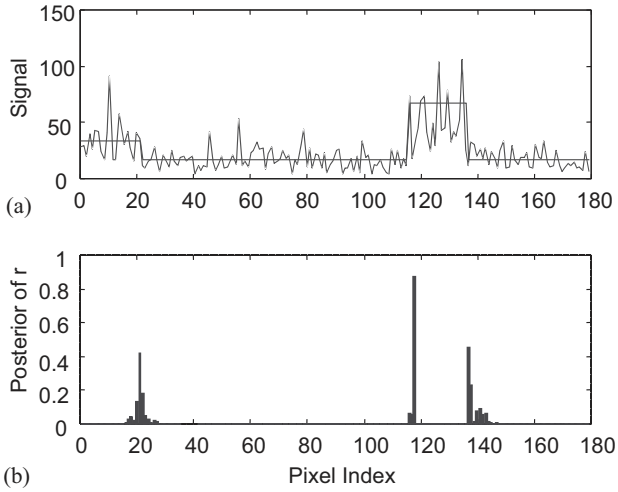

Fig. 8. Line 150 of the synthetic SAR image and the estimated posterior distribution $\hat{f}(r|v)$.

the change-point locations can be easily estimated by thresholding $\hat{f}(r|v)$. Fig. 9 shows the estimated posterior distribution of $K(r)$ obtained by computing the histogram of $K(r^i)$, for $i=51,\ldots,100$. The histogram has a maximum value for $K(r)=4$, which is in good agreement with the actual number of change-points (indeed, we have assumed that there is a change at $N=180$).

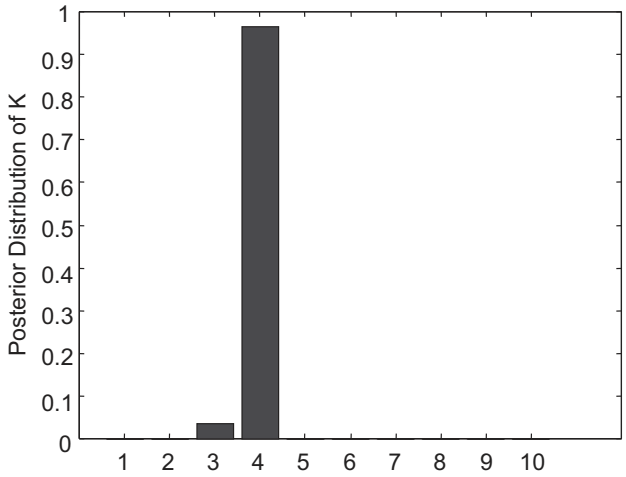

Fig. 9. Estimated posterior distribution of K (number of change-points) for the signal of Fig. 8.

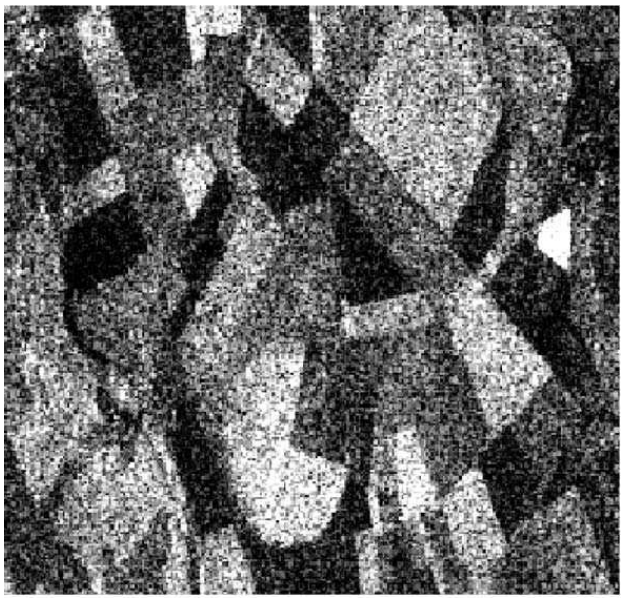

Fig. 10. 3-look 256×256 SAR image of an agricultural scene near Bourges, France. © Copyright ESA-ERS1 data-1993-Distribution SPOT IMAGE.

6.6. *Postprocessing*

By thresholding the 2-D ESMs, pixels belonging to edges are obtained with a certain probability of false alarm. However, in practical applications, ESM thresholding has to be combined with morphological closing or/and the watershed algorithm to obtain closed skeleton boundaries (see for instance [15]). Figs. 10–12 show a 3-look real SAR image of an agricultural scene near Bourges in France and the ESMs of this image using the raw MMSE and ROEWA detectors. The results obtained after postprocessing (ESM+watershed algorithm) are shown in Figs. 13 and 14. The MMSE detector combined with postprocessing performs well on this real image.

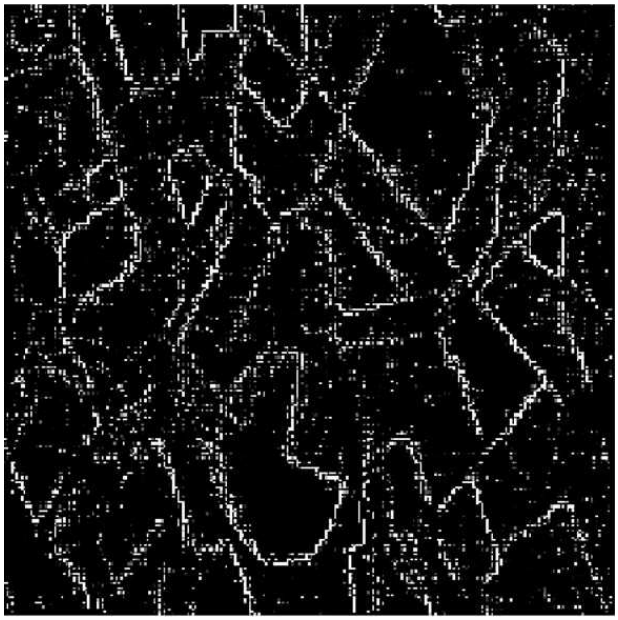

Fig. 11. ESMs of the SAR image represented in Fig. 10 using the MMSE detector.

Fig. 12. ESMs of the SAR image represented in Fig. 10 using the ROEWA detector.

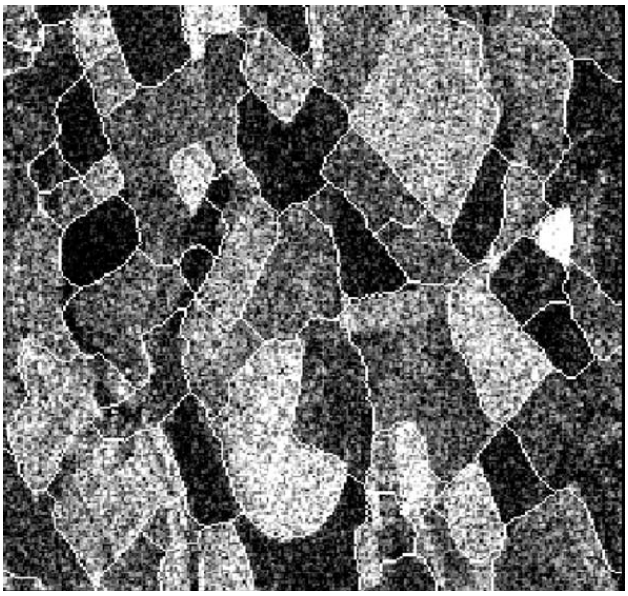

Fig. 13. SAR image segmentation using the MMSE detector and postprocessing.

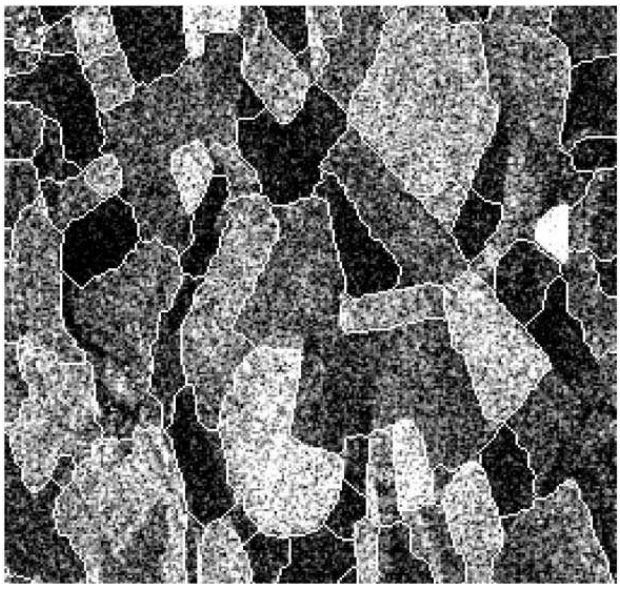

Fig. 14. SAR image segmentation using the ROEWA detector and postprocessing.

7. Conclusions

This paper studies Bayesian off-line change-point detectors based on the MMSE and MAP principles, for SAR image segmentation. The MMSE and MAP detectors were implemented by MCMC methods. Appropriate jumps ensured fast convergence for the Markov Chain. The resolution level in the segmentation was shown to depend (for both detectors) on the *a priori* knowledge of the hyperparameters. This *a priori* knowledge can be replaced by the hyperparameter estimation procedure proposed in Section 5.

It is important to note that the MMSE and MAP detectors were implemented very similarly. However, these detectors do not provide the same information: the MMSE detector estimates *the a posteriori probability of having a change-point* at each pixel whereas the MAP detector *determines the change-point locations which maximize an appropriate a posteriori change-point location distribution*. We feel that the MMSE detector is more attractive than the MAP detector for image segmentation. Indeed, the MMSE detector provided an ESM which can be combined with powerful image postprocessing algorithms including morphological closing or watershed algorithm. Moreover, the MMSE detector simulates the joint a posteriori distribution of change-point locations. Many statistical properties of the change-points (such as the probability to have change-points in a given interval or the posterior distribution of the number of change-points) can be estimated using this joint a posteriori distribution.

The change-point location MMSE detector belongs to the class of edge detectors. It was shown to yield slightly better results than the ROEWA detector (which is one of the most powerful edge detectors for SAR images) in terms of ESM. However, the MMSE detector has a higher computational cost than the ROEWA detector. Future work includes (1) comparison on postprocessed images using appropriate image quality measures, (2) comparison in terms of image quality and execution time with respect to Region Merging and Region Fitting algorithms and (3) incorporation of the row-to-row or column-to-column dependencies by using appropriate masks. In this last extension, potential edges are detected in the horizontal or vertical directions and the edge candidates are selected using the MAP criterion (see [14] for more details).

Acknowledgements

The authors thank the anonymous referees for their constructive comments and suggestions. They

are greatly indebted to R. Fjortoft, C. Lemaréchal and P. Marthon for their help regarding the implementation of the ROEWA detector. The authors would also like to acknowledge Spot Image for providing the real image © Copyright ESA-ERS1 data-1993.

Appendix A. Summary of the MCMC algorithm for MMSE estimation

1. Initialization. Sample $N-1$ i.i.d. Bernoulli variables $\Theta^0(i) \sim B(\lambda)$, $i = 1,\ldots,N-1$ (where $B(\lambda)$ is a Bernoulli distribution with parameter $\lambda = (1+e^{\beta})^{-1}$,

2. for $j = 1,\ldots,N_{\text{bi}} + N_{\text{MC}}$

- perform an independent MH step, i.e.,

  (a) sample a candidate $z^j = (z_1^j,\ldots,z_{N-1}^j)$ such that z_i^j are $N-1$ i.i.d. $B(\lambda)$ variables,
  (b) sample Rand $\sim U[0,1]$ (where $U[0,1]$ is the uniform distribution on $[0,1]$),
  (c) set $\begin{cases} \Theta^j = z^j & \text{if } U(\Theta^{j-1}|x) - U(z^j|x) > \eta = \ln \text{Rand} \\ \Theta^j = \Theta^{j-1} & \text{otherwise,} \end{cases}$

- perform a one-variable-at-a-time MH step, i.e.,

  (a) draw uniformly a permutation σ on the set $\{1,\ldots,N-1\}$,
  (b) for $k = 1,\ldots,N-1$
  for $l = 1,\ldots,N-1$
  set $\tilde{\Theta}^j(l) = \begin{cases} \Theta^j(l) & \text{if } l \neq \sigma(k), \\ 1-\Theta^j(l) & \text{if } l = \sigma(k), \end{cases}$
  end
  sample $\text{Rand}_k \sim U[0,1]$
  set $\begin{cases} \Theta^j \leftarrow \tilde{\Theta}^j & \text{if } \ln \text{Rand}_k < -U(\tilde{\Theta}^j|x) + U(\Theta^j|x), \\ \Theta^j \leftarrow \Theta^j & \text{otherwise,} \end{cases}$

  end

- perform a change-point move, i.e.,

  (a) draw uniformly a permutation τ on the set $\{1,\ldots,K(r)\}$ ($K(r)$ being the change-point number),
  (b) for $k = 1,\ldots,K(r)$
  set I_k to the position of the kth change-point,
  set $\tilde{\Theta}^j(I_{\tau(k)}) = 0$ (the change-point at position $\tau(k)$ is deleted),
  draw uniformly a number $I_{\xi(k)}$ on the set
  $S_{k,\varepsilon} = \{I_{\tau(k)-\varepsilon},\ldots,I_{\tau(k)-1},I_{\tau(k)+1},\ldots,I_{\tau(k)+\varepsilon}\}$ (which defines a neighborhood of $I_{\tau(k)}$)
  where ε is for instance the mean number of change-points i.e. $\varepsilon = E[K(r)] = N\lambda$
  set $\tilde{\Theta}^j(I_{\xi(k)}) = 1$
  sample $\text{Rand}_k \sim U[0,1]$
  set $\begin{cases} \Theta^j \leftarrow \tilde{\Theta}^j & \text{if } \ln \text{Rand}_k < -U(\tilde{\Theta}^j|x) + U(\Theta^j|x), \\ \Theta^j \leftarrow \Theta^j & \text{otherwise,} \end{cases}$
  end

3. compute

$$\bar{\Theta} = \frac{1}{N_{\text{MC}}} \sum_{n=N_{\text{bi}}+1}^{N_{\text{bi}}+N_{\text{MC}}} \Theta^n$$

Appendix B. Summary of the MCMC Algorithm for MAP Estimation

1. Initialization.

- Sample $N-1$ i.i.d. Bernoulli variables $\Theta^0(i) \sim B(\lambda)$, $i = 1, \ldots, N-1$ (where $B(\lambda)$ is a Bernoulli distribution with parameter $\lambda = (1+e^{\beta})^{-1}$,
- set $T_0 \geqslant \Delta$

2. for $j = 1, \ldots, N_{\text{bi}} + N_{\text{MC}}$

- set $T_j = 0.99 T_{j-1}$ and perform an independent SA step i.e.,

  (a) sample a candidate $z^j = (z_1^j, \ldots, z_{N-1}^j)$ such that z_i^j are $N-1$ i.i.d. $B(\lambda)$ variables,
  (b) sample Rand $\sim U[0,1]$ (where $U[0,1]$ is the uniform distribution on $[0,1]$),
  (c) set $\begin{cases} \Theta^j = z^j & \text{if } \ln \text{Rand} < \frac{1}{T_j}\{-U(z^j|x) + U(\Theta^{j-1}|x)\}, \\ \Theta^j = \Theta^{j-1} & \text{otherwise,} \end{cases}$

- set $T_j \leftarrow 0.99 T_j$ and perform a one-variable-at a time MH step i.e.,

  (a) draw uniformly a permutation σ on the set $\{1, \ldots, N-1\}$,
  (b) for $k = 1, \ldots, N-1$

$$\text{set } \tilde{\Theta}^j(l) = \begin{cases} \Theta^j(l) & \text{if } l \neq \sigma(k), \\ 1 - \Theta^j(l) & \text{if } l = \sigma(k), \end{cases}$$

sample $\text{Rand}_k \sim U[0,1]$

$$\text{set } \begin{cases} \Theta^j \leftarrow \tilde{\Theta}^j & \text{if } \ln \text{Rand}_k < \frac{1}{T_j}\{-U(\tilde{\Theta}^j|x) + U(\Theta^j|x)\}, \\ \Theta^j \leftarrow \Theta^j & \text{otherwise,} \end{cases}$$

  end

- set $T_j \leftarrow 0.99 T_j$ and perform a change-point move (the change-point $I_{\tau(k)}$ is moved to $I_{\tau(k)}$) i.e.

  (a) draw uniformly a permutation τ on the set $\{1, \ldots, K(r)\}$ ($K(r)$ being the change-point number),
  (b) for $k = 1, \ldots, K(r)$
  set I_k the position of the kth change-point,
  set $\tilde{\Theta}^j(I_{\tau(k)}) = 0$ (the change-point at position $I_{\tau(k)}$ is deleted),
  draw uniformly a number $I_{\xi(k)}$ on the set
  $S_{k,\varepsilon} = \{I_{\tau(k)-\varepsilon}, \ldots, I_{\tau(k)-1}, I_{\tau(k)+1}, \ldots, I_{\tau(k)+\varepsilon}\}$ (which defines a neighborhood of $I_{\tau(k)}$)
  where ε is for instance the mean number of change-points i.e. $\varepsilon = E[K(r)] = N\lambda$
  set $\tilde{\Theta}^j(I_{\xi(k)}) = 1$ (a change-point at position $I_{\xi(k)}$ is created),
  sample $\text{Rand}_k \sim U[0,1]$

$$\text{set } \begin{cases} \Theta^j \leftarrow \tilde{\Theta}^j & \text{if } \ln \text{Rand}_k < \frac{1}{T_j}\{-U(\tilde{\Theta}^j|x) + U(\Theta^j|x)\}, \\ \Theta^j \leftarrow \Theta^j & \text{otherwise,} \end{cases}$$

(3) set

$$\hat{\Theta} = \Theta^{N_{\text{MC}}}$$

References

[1] R. Andre-Obrecht, A new statistical approach for automatic segmentation of continuous speech signals, IEEE Trans. Acoust. Speech Signal Process. 36 (1988) 29–40.

[2] M. Basseville, I.V. Nikiforov, Detection of Abrupt Changes: Theory and Application, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[3] J.M. Bernardo, A.F.M. Smith, Wiley, New York, 1994.

[4] A.C. Bovik, On detecting edges in speckle imagery, IEEE Trans. Acoust. Speech Signal Process. 36 (October 1988) 1618–1627.

[5] G.L. Bretthorst, Bayesian Spectrum Analysis and Parameter Estimation, Springer, New York, 1988.

[6] B.P. Carlin, A.E. Gelfand, A.F.M. Smith, Hierarchical Bayesian analysis of change-point problems, Appl. Stat. 41 (1) (1992) 389–405.

[7] S. Chib, Estimation and comparison of multiple change-point models, J. Econometrics 86 (1998) 221–241.

[8] S. Chib, E. Greenberg, Understanding the Metropolis-Hastings algorithm, Am. Statistician 49 (1995) 327–335.

[9] E. Cinlar, Introduction to Stochastic Processes, Prentice-Hall, Englewood Cliffs, NJ, 1975.

[10] J.C. Curlander, R.N. McDonough, Synthetic Aperture Radar, Wiley Interscience, New York, 1991.

[11] B. Delyon, M. Lavielle, E. Moulines, Convergence of a stochastic approximation version of the EM algorithm, Ann. Stat. 27 (1) (1999) 94–128.

[12] P.M. Djuric, A MAP solution to off-line segmentation of signals, in: Proceedings of IEEE ICASSP'94, Adelaide, Australia, 1994, pp. 505–508.

[13] P.M. Djuric, J.K. Fwu, On the detection of edges in vector images, IEEE Trans. Image Proc. 6 (11) (November 1997) 1595–1601.

[14] P.M. Djuric, S.M. Kay, D. Faye Boudreaux-Bartels, Segmentation of nonstationary signals, in: Proceedings of IEEE ICASSP'92, San Francisco, CA, USA, 1992, pp. 161–164.

[15] R. Fjørtoft, A. Lopès, P. Marthon, E. Cubero-Castan, An optimal multiedge detector for SAR image segmentation, IEEE Trans. Geosci. Remote Sensing 36 (3) (May 1998) 793–802.

[16] V.S. Frost, J.A. Stiles, K.S. Shanmugan, J.C. Holtzman, A model for radar images and its application to adaptive digital filtering of multiplicative noise, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-4 (March 1982) 157–166.

[17] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. PAMI- 6 (November 1984) 721–741.

[18] C. Geyer, Practical Markov chain Monte Carlo, Stat. Sci. 7 (1992) 473–482.

[19] W.R. Gilks, S. Richardson, D.J. Spiegelhalter, Introducing Markov chain Monte Carlo, in: W.R. Gilks, S. Richardson, D.J. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, Chapman & Hall, London, UK, 1996, pp. 1–19.

[20] S.J. Godsill, P.J.W. Rayner, Statistical reconstruction and analysis of autoregressive signals in impulsive noise using the Gibbs sampler, IEEE Trans. Speech Audio Proc. 6 (July 1998) 352–372.

[21] P.J. Green, Reversible jump MCMC computation and Bayesian model determination, Biometrika 82 (1995) 711–732.

[22] F. Gustafsson, The marginalized likelihood ratio test for detecting abrupt changes, IEEE Trans. Automat. Control 41 (1996) 66–77.

[23] E. Hervet, R. Fjørtoft, P. Marthon, A. Lopès, Comparison of wavelet-based and statistical speckle filters, in: Proceedings of the European Symposium on Remote Sensing, SAR Image Analysis, Modelling, and Techniques III, Vol. SPIE 3497, Barcelona, Spain, 21–25 September 1998.

[24] P.A. Kelly, H. Derin, K. Hartt, Adaptive segmentation of speckled images using a hierarchical random field model, IEEE Trans. Acoust. Speech Signal Process. 36 (10) (October 1988) 1628–1641.

[25] M. Lavielle, Detection of multiple changes in a sequence of random variables, Stoch. Process Appl. 83 (1) (September 1999) 79–102.

[26] M. Lavielle, Optimal segmentation of random processes, IEEE Trans. Signal Process. 46 (5) (May 1998) 1365–1373.

[27] R.E. McCulloch, R.S. Tsay, Bayesian inference and prediction for mean and variance shifts in autoregressive time series, J. Am. Stat. Assoc. 88 (1993) 968–978.

[28] R.E. McCulloch, R.S. Tsay, Bayesian analysis of autoregressive time series via the Gibbs sampler, J. Time Ser. Anal. 15 (1994) 235–250.

[29] C.J. Oliver, S. Quegan, Understanding Synthetic Aperture Radar Images, Artech House, Norwood, MA, 1998.

[30] D.B. Phillips, A.F.M. Smith, Bayesian model comparison via jump diffusion, in: W.R. Gilks, S. Richardson, D.J. Spiegelhalter (Eds.), Markov Chain Monte Carlo in Practice, Chapman & Hall, London, UK, 1996, pp. 215–239.

[31] C.P. Robert, The Bayesian Choice: A Decision-Theoretic Motivation, Springer, New York, 1994.

[32] C.P. Robert, Discretization and MCMC Convergence Assessment, in: Lectures Notes in Statistics, Vol. 135, Springer, Berlin, 1998.

[33] J.O. Ruanaidh, W.J. Fitzgerald, Numerical Bayesian Methods Applied to Signal Processing, Statistics and Computing, Springer, Berlin, 1996.

[34] B.M. Sadler, A. Swami, Analysis of multiscale products for step detection and estimation, IEEE Trans. Inform Theory 45 (3) (April 1999) 1043–1051.

[35] J.A. Stark, W.J. Fitzgerald, S.B. Hladky, Multiple-order Markov chain Monte Carlo methods with application to a change-point model, Technical report CUED/F-INFENG-TR 302, Department of Engineering, Cambridge University, Cambridge, UK, September 1997.

[36] D.A. Stephens, Bayesian retrospective multiple-change-point identification, Appl. Stat. 43 (1) (1994) 159–178.

[37] D.A. Stephens, A.F.M. Smith, Bayesian edge detection in images via change-point methods, in: W. Hardle, L. Simar (Eds.), Computer intensive methods in statistics, Physica Verlag, Wurzburg, 1993, pp. 1–29.

[38] A. Swami, Multiplicative noise models: parameter estimation using cumulants, Signal Processing 36 (1994) 355–373.

[39] A. Swami, B. Sadler, Step-change localization in additive and multiplicative noise via multiscale products, in: 32nd Asilomar Conference on Signal Systems and Computers, Pacific Grove, CA, November 1998.
[40] L. Tierney, Markov chains for exploring posterior distributions, Ann. Stat. 22 (4) (1994) 1701–1728.
[41] J.Y. Tourneret, M. Chabert, Off-line detection and estimation of abrupt changes corrupted by multiplicative colored Gaussian noise, Proceedings of ICASSP'97, Munich, April 1997.
[42] J.Y. Tourneret, M. Coulon, M. Doisy, Least squares estimation of multiple abrupt changes contaminated by multiplicative noise using MCMC, Proceedings of HOS'99, Caesarea, Israel, June 1999.
[43] R. Touzi, A. Lopes, P. Bousquet, A statistical and geometrical edge detector for SAR images, IEEE Trans. Geosci. Remote Sensing 26 (6) (November 1988) 764–773.
[44] F.T. Ulaby, R.K. Moore, A.K. Kung, Microwave Remote Sensing, Vol. 3, Artech House, Dedham, MA, 1986.
[45] H.L. Van Trees, Detection, Estimation and Modulation Theory, Part I, Wiley, New York, 1968.
[46] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Trans. Pattern Anal. Mach. Intell. PAMI 13 (May 1991) 583–598.
[47] A.W. Willsky, H.L. Jones, A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems, IEEE Trans. Automat. Control 21 (February 1976) 108–112.
[48] G. Winkler, Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction, Springer, Berlin, 1995.