Designing SAR Images Change-Point Estimation Strategies using an MSE Lower Bound

Ammar Mian [1,2] Lucien Bacharach [3] Guillaume Ginolhac [2] Alexandre Renaux [3] M. Nabil El Korso [4] Jean-Philippe Ovarlez [1,5]

[1]CentraleSupélec SONDRA [2]LISTIC, Université Savoie Mont-Blanc [3]L2S, Université Paris-Saclay [4]LEME, Université Paris-Nanterre [5]DEMR ONERA, Université Paris-Saclay

I. Some theory about hybrid estimation and MSE Lower Bound

Hybrid Estimation problem:

Suppose we have N i.i.d observations $\{\mathbf{x}_k \in \mathbb{C}^p | 1 \leq k \leq N\}$ of a random vector : $\mathbf{x} \sim f_{\mathbf{x};\boldsymbol{\theta}}(\mathbf{x};\boldsymbol{\theta})$.
We denote by $\hat{\boldsymbol{\theta}}(\mathbf{x})$ an estimator of the **hybrid** parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ of dimension M. Hybrid means that:

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_d, \boldsymbol{\theta}_r]^{\mathrm{T}}$$

Deterministic Random

MSE lower bound:

We want to have the lower bound on the Mean Square Error (MSE) of the estimator with regards to the true value. The MSE is defined as:

$$\mathbf{MSE}(\hat{\boldsymbol{\theta}}) = \mathbb{E}_{\mathbf{x},\boldsymbol{\theta}_d,\boldsymbol{\theta}_r} \left\{ (\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta})^{\mathrm{T}} \right\}, \quad (1)$$

A lower-bound can be obtained using the covariance equality:

$$\mathbf{MSE}(\hat{\boldsymbol{\theta}}) \succeq \mathbf{V}\mathbf{P}^{-1}\mathbf{V}^{\mathrm{T}}, \quad (2)$$

where $\mathbf{V}$ is a $M \times M$ matrix whose elements are given by

$$(\mathbf{V})_{k,l} = \mathbb{E}_{\mathbf{x},\boldsymbol{\theta}_d,\boldsymbol{\theta}_r} \left\{ \left((\hat{\boldsymbol{\theta}}(\mathbf{x}))_k - (\boldsymbol{\theta})_k \right) \Psi_l(\mathbf{x},\boldsymbol{\theta}) \right\},$$

$\mathbf{P}$ is a $M \times M$ matrix whose elements are given by

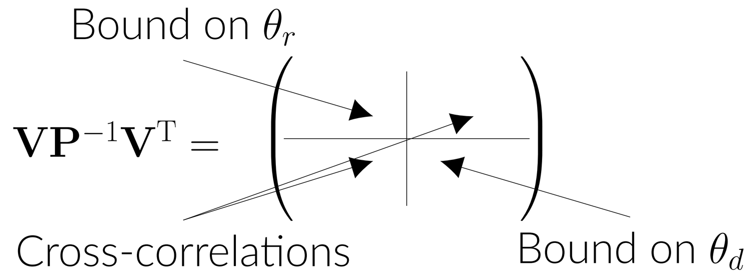
$$(\mathbf{P})_{k,l} = \mathbb{E}_{\mathbf{x},\boldsymbol{\theta}_d,\boldsymbol{\theta}_r} \{\Psi_k(\mathbf{x},\boldsymbol{\theta})\Psi_l(\mathbf{x},\boldsymbol{\theta})\}$$

and $\{\Psi_k(\mathbf{x},\boldsymbol{\theta}) \,|\, k \in [\![1, M]\!]\}$ are real-valued functions.

Example of the Cramer-Rao Bound (CRB):

$\mathbf{MSE}(\hat{\boldsymbol{\theta}}) \succeq \mathbf{F}^{-1}$, where $\mathbf{F} = -\mathbb{E}_{\mathbf{x},\boldsymbol{\theta}_r} \left\{ \frac{\partial^2 f_{\mathbf{x};\boldsymbol{\theta}}}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathrm{T}}} \right\}$

Problem 1: What if hybrid estimation problem?
Problem 2: What if $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ where $\boldsymbol{\Theta}$ is discrete ($\partial\boldsymbol{\theta}$)?
→ Use an **hybrid lower-bound**:



In this paper, we consider the hybrid **Cramer-Rao** and **Weiss-Weinstein** lower bound, which we will denote **HCRWWB**.

II. Change-point estimation in SAR image time series analysis

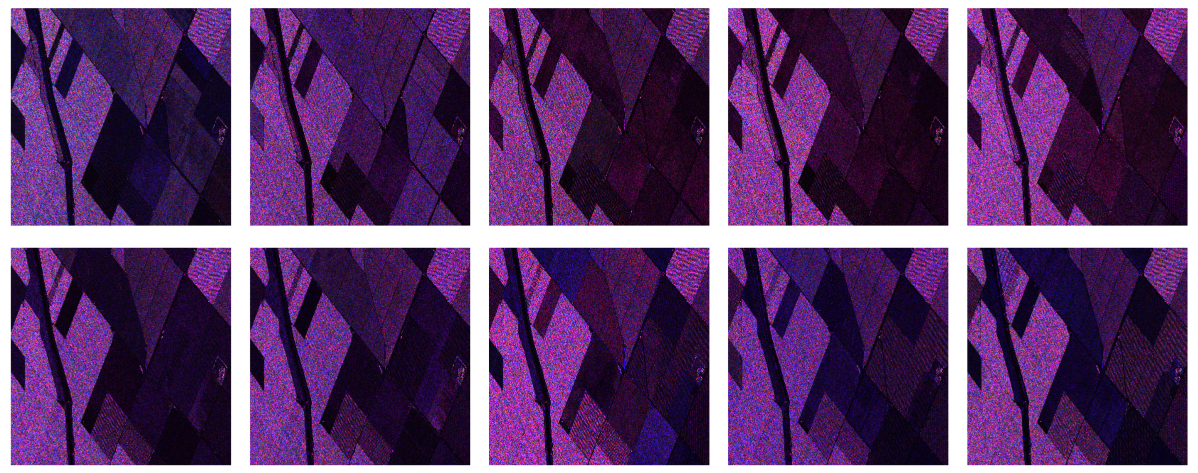


Figure 1. Example of UAVSAR/JPL image time-series between 2008 and 2018 over California

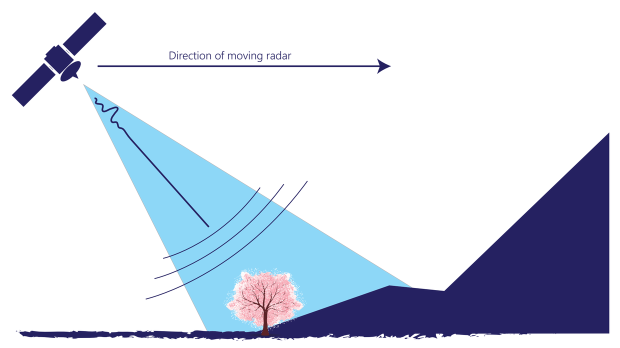


Figure 2. SAR acquistion principle

SAR images are useful to monitor changes for **large areas** (several km) over a **long time-period** (several years).

Data model: On a local window, the Sample Covariance Matrix (SCM) is assumed to be Wishart distributed:

$$f_{\mathbf{X}_t;\boldsymbol{\Sigma}}(\mathbf{X}_t;\boldsymbol{\Sigma}) = \frac{|\mathbf{X}_t|^{N-p}}{\Gamma_p(N)\,|\boldsymbol{\Sigma}|^N} \operatorname{etr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{X}_t\right), \quad (3)$$

where, $\Gamma_p(N) = \pi^{p(p-1)/2} \prod_{j=1}^{p} \Gamma(N-j+1)$, $\Gamma(.)$ is the Gamma function and $\operatorname{etr}(.)$ is the exponential trace function.

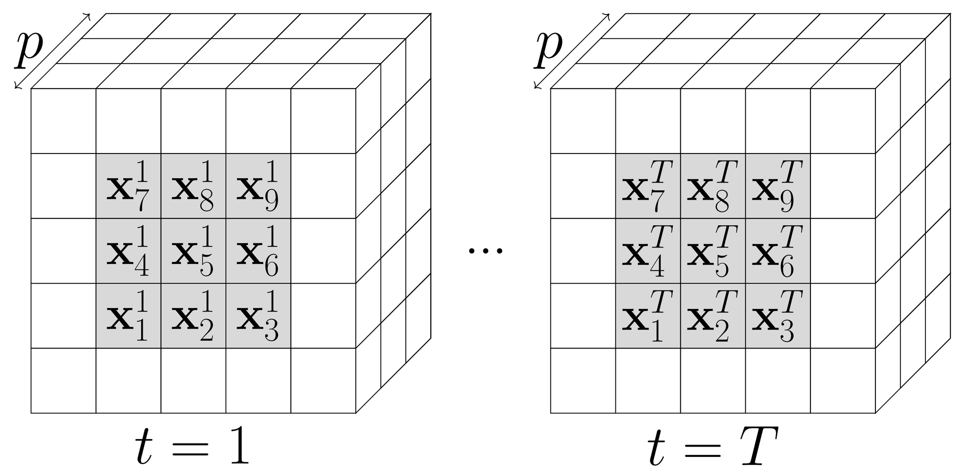


Figure 3. Local data selection for computing SCM

Single change point estimation problem:

For a single pixel, we consider the following model:

$$\begin{cases} \mathbf{X}_t \sim \mathbb{C}\mathcal{W}(p, N, \boldsymbol{\Sigma}_0) \text{ for } t = 1, \dots, t_C \\ \mathbf{X}_t \sim \mathbb{C}\mathcal{W}(p, N, \boldsymbol{\Sigma}_1) \text{ for } t = t_C + 1, \dots, T \end{cases}. \quad (4)$$

The objective is to esimate $\boldsymbol{\Sigma}_0$, $\boldsymbol{\Sigma}_1$ and t_C called the **change-point**. We assume a uniform prior this parameter.

Tuning problem:

The performance of any estimator of t_C depends on the set of parameters (p, N) (dimension of each pixel, number of samples in the local windows) which can be controlled. How can we tune the methods to have a desired performance ?
→ Since the MSE of the estimators is untractable, we focus on obtaining a **lower-bound** !

We consider a CRB on the **deterministic unknown** covariances $\boldsymbol{\Sigma}_0$, $\boldsymbol{\Sigma}_1$ and the bayesian Weiss Weinstein Bound (WWB) on the **random unknown** change-point t_C.

III. Statement of the result

In the context of change-point estimation, the right-hand side of the inequality at eq. (2) can be obtained by using the semi closed-form expression provided in [1]:

$$\mathbf{V} = \begin{bmatrix} -\mathbf{I}_{2p^2} & \mathbf{0}_{2p^2,1} \\ \mathbf{0}_{1,2p^2} & v_{22} \end{bmatrix} \text{ and } \mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^{\mathrm{T}} & P_{22} \end{bmatrix}, \quad (5)$$

where the block-matrices are defined as follows:

- $\mathbf{P}_{11} = T/2 \operatorname{diag}(\mathbf{F}(\boldsymbol{\Sigma}_0), \mathbf{F}(\boldsymbol{\Sigma}_1))$, where $\mathbf{F}(\boldsymbol{\Sigma}_0)$ (resp. $\mathbf{F}(\boldsymbol{\Sigma}_1)$) is the Fisher information matrix with regards to $\boldsymbol{\Sigma}_0$ (resp. $\boldsymbol{\Sigma}_1$).
- $P_{22} =$ $u(h)\left(\rho^{|h|}(\epsilon_h(2s)) + \rho^{|h|}(\epsilon_h(2s-1))\right) - 2u(2h)\rho^{|h|}(\epsilon_h(s))$, where:

$$u(h) \triangleq \begin{cases} (T-1-|h|)/(T-1) & \text{if } |h| < T-1 \\ 0 & \text{otherwise} \end{cases},$$

$$\epsilon_h(s) = \begin{cases} s & \text{if } h > 0 \\ 1-s & \text{if } h < 0 \end{cases} \text{ and}$$

$$\rho(s) \triangleq \int_{\mathbb{S}_{\mathbb{H}}^p} f^s_{\mathbf{X}_t;\boldsymbol{\Sigma}_0}(\mathbf{X}_t;\boldsymbol{\Sigma}_0) f^{1-s}_{\mathbf{X}_t;\boldsymbol{\Sigma}_1}(\mathbf{X}_t;\boldsymbol{\Sigma}_1)\,\mathrm{d}\mathbf{X}_t. \quad (6)$$

- $v_{22} = hu(h)\rho^{|h|}(\epsilon_h(s))$.

[1] L. Bacharach, M. N. E. Korso, A. Renaux and J. Tourneret, "A Hybrid Lower Bound for Parameter Estimation of Signals With Multiple Change-Points," in IEEE Transactions on Signal Processing

- $\mathbf{P}_{12} = \left[\mathbf{p}^{\mathrm{T}}, \mathbf{q}^{\mathrm{T}}\right]^{\mathrm{T}}$, where the elements of vectors $\mathbf{p}$ and $\mathbf{q}$ are given by:

$$(\mathbf{p})_\ell = -hu(h)\rho^{|h|-1}(\epsilon_h(s))\,\phi_{\sigma_0,\ell}(\epsilon_h(s)),$$
$$(\mathbf{q})_\ell = hu(h)\rho^{|h|-1}(\epsilon_h(s))\,\phi_{\sigma_1,\ell}(\epsilon_h(s)),$$

and given $j \in \{0,1\}$, $\ell \in [\![1, p^2]\!]$, $s \in]0,1[$:

$$\phi_{\sigma_j,\ell}(s) \triangleq \int_{\mathbb{S}_{\mathbb{H}}^p} \left.\frac{\partial \ln f_{\mathbf{X}_t;\boldsymbol{\Sigma}}(\mathbf{X}_t;\boldsymbol{\Sigma})}{\partial\left([\operatorname{vech}(\boldsymbol{\Sigma})]_{\mathbb{CR}}\right)_\ell}\right|_{\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_j} \times f^s_{\mathbf{X}_t;\boldsymbol{\Sigma}_0}(\mathbf{X}_t;\boldsymbol{\Sigma}_0)\, f^{1-s}_{\mathbf{X}_t;\boldsymbol{\Sigma}_1}(\mathbf{X}_t;\boldsymbol{\Sigma}_1)\,\mathrm{d}\mathbf{X}_t. \quad (7)$$

Derivations lead to:

- $\rho(s) = \dfrac{\left|s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right|^{-N}}{|\boldsymbol{\Sigma}_0|^{sN}\,|\boldsymbol{\Sigma}_1|^{(1-s)N}}$.
- $\mathbf{F}(\boldsymbol{\Sigma}) = f_{\mathbb{CR}}\left(N\mathbf{D}_p^{\mathrm{T}}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1})\mathbf{D}_p\right)$, where $\mathbf{D}_p$ is the duplication matrix defined for any matrix $\mathbf{X} \in \mathbb{C}^{p\times p}$, by

$$\mathbf{D}_p \operatorname{vech}(\mathbf{X}) = \operatorname{vec}(\mathbf{X}).$$

- The different terms of $\phi_{\sigma_j,\ell}(s)$ for $\ell \in [\![1, p^2]\!]$, $j \in \{0,1\}$ are given by $\phi_{\sigma_j,\ell}(s) = \left([\operatorname{vech}(\boldsymbol{\Phi}_j(s))]_{\mathbb{CR}}\right)_\ell$, where $\boldsymbol{\Phi}_j(s)$ is a $p \times p$ matrix given by:

$$\boldsymbol{\Phi}_j(s) = N\rho(s)\boldsymbol{\Sigma}_j^{-1}\left(s\boldsymbol{\Sigma}_0^{-1} + (1-s)\boldsymbol{\Sigma}_1^{-1}\right)^{-1}\boldsymbol{\Sigma}_j^{-1} - N\rho(s)\boldsymbol{\Sigma}_j^{-1}.$$

IV. Simulation results

Estimators:

Two estimators of the change-point have been considered:

- The Maximum A Posteriori (MAP) estimator which has the knowledge of the covariance matrices before and after the change:

$$\hat{t}_C = \underset{t_C \in [\![1, T-1]\!]}{\operatorname{argmax}} \; f_{\mathbf{x},t_C}(\mathbf{x}, t_C). \quad (8)$$

- The following Maximum A Posteriori/Maximum Likelihood estimator:

$$\hat{t}_C = \underset{t_C \in [\![1, T-1]\!]}{\operatorname{argmax}} \; f_{\mathbf{x},t_C;\hat{\boldsymbol{\sigma}}}(\mathbf{x}, t_C; \hat{\boldsymbol{\sigma}}), \quad (9)$$

where $\hat{\boldsymbol{\sigma}} = \left[[\operatorname{vech}(\hat{\boldsymbol{\Sigma}}_0)]_{\mathbb{CR}}^{\mathrm{T}}, [\operatorname{vec}(\hat{\boldsymbol{\Sigma}}_1)]_{\mathbb{CR}}^{\mathrm{T}}\right]^{\mathrm{T}}$ with:

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{1}{t_C N}\sum_{t=1}^{t_C} \mathbf{X}_t \text{ and } \hat{\boldsymbol{\Sigma}}_1 = \frac{1}{(T - t_C)N}\sum_{t=t_C+1}^{T} \mathbf{X}_t.$$

Tuning example:

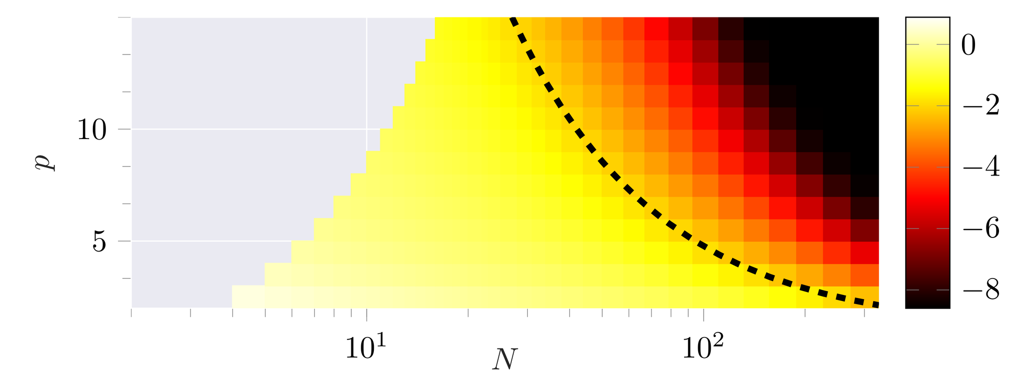


Figure 4. Evolution of $\log_{10}\sqrt{(\mathbf{HCRWWB})_{M,M}}$ for several parameters p and N, $T = 100$, $\alpha_0 = 0.1$ and $\alpha_1 = 0.3$. The dashed line corresponds to the region where $\sqrt{(\mathbf{HCRWWB})_{M,M}} = 10^{-2}$.

Validation of the bound:

In order to validate the bound derived in this paper, Wishart time series subjected to a change-point as described in eq. (4) have been generated. t_C is generated using a uniform random prior and the covariance matrices have been chosen as Toeplitz matrices of the form: $(\boldsymbol{\Sigma}_{k=0,1})_{i,j} = \alpha_k^{|i-j|}$.

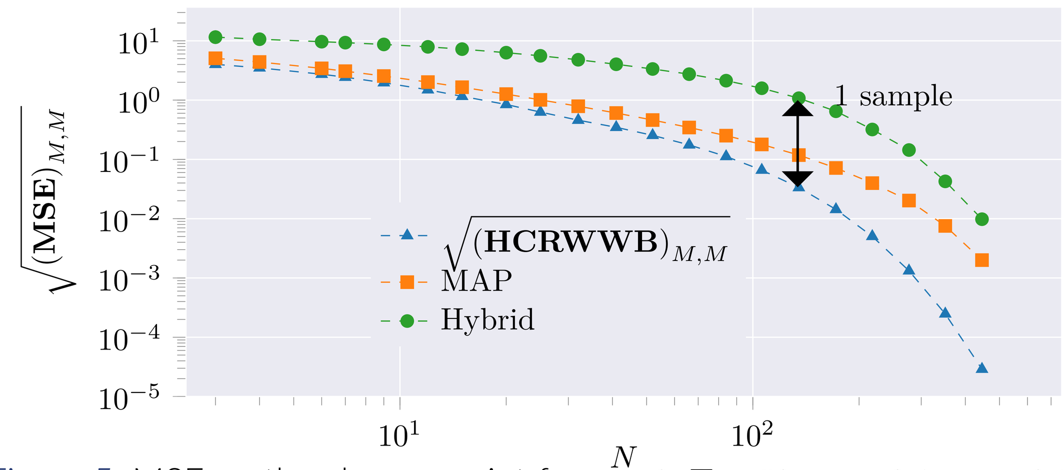


Figure 5. MSE on the change-point for $p = 3$, $T = 50$, $\alpha_0 = 0.1$, $\alpha_1 = 0.3$. The estimators curves have been computed with 10^6 Monte Carlo trials.

Table 1. Time-consumption in seconds.

N	HCRWWB	MAP	Hybrid estimator
10	0.17	305	310
10^2	0.17	510	568
10^3	0.17	1462	1476