Signal models for changes in Polarimetric SAR data

Armando Marino, *Member, IEEE*, Matteo Nannini.

Abstract—
SAR polarimetry (PolSAR) can improve change detection in terms of detection capabilities. In this work, we are proposing to extend the idea of target decomposition to changes affecting partial targets. This will allow the separation of polarimetric dependent changes, providing extra information that can be used to better understand the processes affecting the targets.

Three models for changes are proposed and compared. The methodologies are based on Lagrangian optimisations of distinct operators built using quadratic forms for a power ratio and a power difference. The optimisations can be accomplished by diagonalisations of specific matrices derived from polarimetric covariance matrices. These are therefore spectral decompositions of an appropriate matrix which we define as change matrix.

The theoretical validity of the models is assessed using Monte Carlo simulations. Additionally, we perform real data validation exploiting L-band quad-polarimetric data from the E-SAR (DLR) SARTOM 2006 campaign and ALOS PALSAR (JAXA) acquisitions in Morecombe Bay (UK).

We observed that the two algorithms based on power difference allow to decompose the change into the minimal set of scattering mechanisms that have been added or removed from the scene. The two algorithms differ on the initial assumption on the change. On the other hand the ratio operator provides a better detection performance although the eigenvalues do not correspond to meaningful scattering mechanisms. A combination of the three methodologies can therefore improve detection and classification of changes.

Keywords— **Synthetic Aperture Radar, Polarimetry, Change Detection, Signal Models, Decompositions.**

I. Introduction

Change detection is a valuable topic in SAR remote sensing. Polarimetric SAR (PolSAR) can improve the results of single polarimetric algorithms [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. The polarization of the transmitted and received waves can be exploited to extract more information about the observed targets. This information enhances classification. In this work, we want to propose decompositions to help classify the change.

It is always possible to perform approaches where the two set of PolSAR acquisitions are analysed separately. However due to the complexity of the multi-dimensional SAR data, it is generally hard and not streighforward to identify the scattering mechanisms (SM) which have suffered the change. Specifically, from an algebraic point of view, there is no methodology to identify the minimal unique set of scattering mechanisms (SM) that have changed. Often we can just measure the increase or decrease of a set of targets (surface, dihedral [11], [12]) or parameters (entropy, α [13]), but not point at the exact scattering mechanism that produced such changes. In this work, we want to propose an operation that is able to synthesise the change and reveal in a straightforward way the scattering mechanism (SM) which produced the maximum change.

In the following there is a brief introduction to PolSAR tools that we will use. A single target is defined as a deterministic target which does not change its polarimetric behavior in time/space. Therefore, it can be represented by a single scattering matrix $[S]$ or equivalently a single scattering vector [14], [15]:

$$\underline{k}_L = [HH, HV, VH, VV]^T , \quad (1)$$

where H and V stands for linear horizontal and vertical and the repeated letter indicates the transmitter-receiver sequence. The previous is obtained employing the Lexicographic basis set. In the case of a reciprocal medium and monostatic sensor, $HV = VH$ and $\underline{k}$ is three-dimensional complex (i.e. $\underline{k}_P \in \mathbb{C}^3$) [14]. Another largely used basis set to convert $[S]$ into a scattering vector is the Pauli basis. In the reciprocal case, this is $\underline{k}_P = 1/\sqrt{(2)}\,[HH + VV, HH - VV, 2 * HV]^T$. A Scattering Mechanisms (SM) $\underline{\omega}$ is an ideal target and is defined as a normalized scattering vector.

The targets observed by a SAR system are often distributed over an area larger than the resolution cell and composed by different objects. For this reason, each pixel of the image containing such distributed targets shows a specific polarimetric behaviour. Such targets take the name of *partial targets* and they are generally described exploiting the second order statistics [14], [15]. In this context, a target covariance matrix can be estimated as $[C] = \langle \underline{k}\ \underline{k}^{*T} \rangle$, where $\langle . \rangle$ is the finite averaging operator. In case that the Single Look Complex (SLC) pixel can be modelled by a Complex Gaussian, the second order statistics are necessary and sufficient to completely characterize a partial target (and the covariance matrix is Wishart distributed). In case of the Pauli basis, the covariance matrix is often named *Coherency* matrix.

An important concept exploited in this work is the quadratic form of a covariance matrix

$$P_T = \underline{\omega}^{*T}[C]\underline{\omega}. \quad (2)$$

The latter represent the power backscattered by the scattering mechanism represented by ω. This is provided the scattering mechanism is present. Otherwise it just represents the power on that projection vector ω.

The novelties of this paper are the following: a) developping and validating for the first time (in a peer reviewed publication) the change detector based on optimisation of DIFFERENCE of power [16], b) provide a comparison using simulated and real data of three methodologies based on optimisations of quadratic forms for change detection including [17], [18], [19].

II. Signal models for change detection with polarimetric data

The starting point of this work is the consideration that the power of a SM can be estimated using a quadratic form of the covariance matrix $[C]$ and the appropriate projection vector $\underline{\omega}$. In the following we will present analyses using a general covariance matrix, however it is important to keep in mind that if we want to interpret the results in sight of the Cloude-Pottier decomposition [20], the scattering vector needs to be in Pauli basis.

Armando Marino is with The University of Stirling, Natural Sciences, Stirling, UK (e-mail: armando.marino@stir.ac.uk). Matteo Nannini is with Microwaves and Radar Institute, German Aerospace Agency, DLR, Wessling, Germany

The analyses in this paper are based on the following two key assumptions:
i) We refer to "polarimetric scattering mechanism" (SM). These are idealised targets that can be represented by a single and unique normalised scattering matrix (these can be different from interferometric scattering mechanisms). We will also omit the word "polarimetric" unless this can produce confusion.
ii) If there is a change of SM between two acquisitions, this will reveal itself in a change of the matrix $[C]$. Unfortunately, PolSAR signal decomposition is ill posed and we can only monitor three independent targets (unless a physical model is used). If changes in the scene are such that the covariance matrix remains identical, then it is not possible to detect such changes using PolSAR.

The three change models proposed and compared here are based on spectral signal decompositions of what we define as change matrices. Therefore, the output of these three algorithms are sets of eigenvalues and eigenvectors. It is important to keep in mind, that a change matrix does not necessarily represent a unique partial target, but a set of targets or transformations. In the following, we sometime refer to these optimisations as "change decompositions", however these are more similar to signal decompositions and therefore should not be confused with traditional "target decompositions" such as [13], [11], [12]. These concepts will be described in more details and validated in the reminder of this paper.

The first step in interpreting the new decompositions is to build a change framework based on detection hypotheses [21]:
- The Null hypothesis H_0 states that there is no modification of SM
- The Alternative hypothesis H_1 states that there is a change of SM.

Since we are interested in identifying optimal basis for change matrices, we can apply optimisations techniques:

$$\underline{\omega}^r_{max} = \underset{\underline{\omega}\in\Omega}{Argmax}\left[F([C_1],[C_2],\underline{\omega})\right], \tag{3}$$

where $F[.]$ is a function of covariance matrices for the two acquisitions ($[C_1]$ and $[C_2]$) and the projection vector $\underline{\omega}$. Ω is a unitary complex sphere, therefore $\underline{\omega}\in\mathbb{C}^3, ||\underline{\omega}|| = 1$.

Although many different functions can be designed, in the following we focus on two main expressions for $F[.]$, since these are derived from the two most straightforward operations in change detection: the ratio and the difference.

Another final note is on the fact that all the detectors tested in this paper are not scale invariant. This is because, in general applications, the overall intensity (or SPAN) of the acquisitions provides useful information. Therefore, changes in span will affect the detectors. However, if we want to use these detectors in applications where the overall intensity is introducing errors (e.g. if there are calibration errors), then each of the covariance matrices could be normalised by their Trace before using the matrices in the detectors.

A. Multiplicative model

A change detector based on the optimisation of the power ratio was already proposed by the authors for PolSAR data [22], [23], [24].

$$F([C_1,C_2,\underline{\omega}]) = M = \rho_{21} = \frac{\underline{\omega}^{*T}C_2\underline{\omega}}{\underline{\omega}^{*T}C_1\underline{\omega}}. \tag{4}$$

$$\underline{\omega}^r_{max} = \underset{\underline{\omega}\in\mathbb{C}^3}{Argmax}\left[\frac{\underline{\omega}^{*T}C_2\underline{\omega}}{\underline{\omega}^{*T}C_1\underline{\omega}}\right]. \tag{5}$$

Additionally, the optimisation of this functional is not new to the community, since it was previously proposed few decades ago by Swartz and Novak with the Polarimetric Match Filter (PMF) [25], [17]. The PMF aims at optimising the contrast between a target $[C_2]$ and clutter $[C_1]$ by selecting the projection vector that maximises the ratio of the quadratic forms. It also has to be said that this optimisation was previously known to the pattern recognition community and it sits in the context of Linear Discriminant Analysis [26].

In our approach, the optimization can be implemented using a constrained Lagrangian methodology which returns a diagonalization:

$$\begin{aligned} L &= \underline{\omega}^{*T}[C_2]\underline{\omega} - \lambda(\underline{\omega}^{*T}[C_1]\underline{\omega} - C), \\ \frac{\partial L}{\partial \underline{\omega}^{*T}} &= [C_2]\underline{\omega} - \lambda[C_1]\underline{\omega} = 0 \\ [C_1]^{-1}[C_2]\underline{\omega} &= \lambda\underline{\omega}. \end{aligned} \tag{6}$$

To conclude, $[C_1]^{-1}[C_2] = [U_r]^{*T}[\Sigma_r][U_r]$, where $[\Sigma_r] = diag(\lambda_{1r},\lambda_{2r},\lambda_{3r})$ with $\lambda_{1r},\lambda_{2r},\lambda_{3r}\in\mathbb{R}^+$ and the columns of $[U_r]$ are the eigenvectors. The maximum eigenvalue λ_{1r} represents the value of the maximum ratio. Interestingly optimising the ratio:

$$\rho_{12} = \frac{\underline{\omega}^{*T}[C_1]\underline{\omega}}{\underline{\omega}^{*T}[C_2]\underline{\omega}}, \tag{7}$$

ends up with the diagonalization of the matrix: $[C_2]^{-1}[C_1]$.

Therefore, the projection vector that suffers the maximum change $\underline{\omega}_{max}$ is:

$$\underline{\omega}_{max} = Argmax\left[\underset{\underline{\omega}\in\Omega}{Opt}\rho_{12}, \underset{\underline{\omega}\in\Omega}{Opt}\rho_{21}\right]. \tag{8}$$

It is interesting to note that $([C_1]^{-1}[C_2])^{-1} = [C_2]^{-1}[C_1]$, therefore they will have the same eigenvectors, but inverted eigenvalues (elevating a matrix does not change the eigenvectors) [27]. We can define the change matrix to diagonalise as $[C_m] = [C_1]^{-1}[C_2]$. In [22] we showed how it can be proved that $[C_m]$ can be diagolinised and it has real positive eigenvalues. Additionally, the matrix is not Normal and therefore its eigenvectors are not bound to be orthogonal.

$$\underline{\omega}_{max} = Argmax\left[\underset{\underline{\omega}\in\Omega}{max}(\rho_{12}), 1/\underset{\underline{\omega}\in\Omega}{min}(\rho_{12})\right]. \tag{9}$$

The eigenvalues are included in the interval between zero and infinity. A change reducing power will appear close to zero, while a change increasing power will appear as a large number. With the goal of improving the visualization of the results, the eigenvalues can be inverted when they are smaller than one. In other words:

$$\begin{cases} \hat{\rho}_{max} = \rho_{max} & \text{if } \rho_{max} \geq 1, \\ \hat{\rho}_{max} = \frac{1}{\rho_{max}} & \text{if } \rho_{max} < 1. \end{cases} \tag{10}$$

In the following, for the sake of brevity, we will refer to this change decomposition as RATIO.

B. Additive model

Another model considers the difference of quadratic forms for the function $F[]$. Preliminary results were already presented at conferences [28], [16], [29] and recently a modification of this optimisation was proposed in [19].

B.1 Difference matrix

The function we want to optimise now is

$$F([C_1, C_2, \underline{\omega}]) = D = \underline{\omega}^{*T} C_2 \underline{\omega} - \underline{\omega}^{*T} C_1 \underline{\omega}. \quad (11)$$

D can be re-written as $D = \underline{\omega}^{*T} [C_2 - C_1] \underline{\omega}$. We can define the change matrix C_c as $C_c = C_2 - C_1$. C_c is a Normal matrix, since it is the difference of two Normal matrices. This means that the diagonal elements of C_c are real and the upper triangular part is the complex conjugate of the lower triangular part. However, there is a difference with ordinary coherency matrices which are also Hermitian. C_c is not bound to be positive semi-definite, because the diagonal element can be negative. This means that its quadratic forms (using a generic $\underline{\omega}$) or its Trace can be negative. In other words, $P = \underline{\omega}^{*T}[C_c]\underline{\omega} \in \mathbf{R}$ can be negative for some $\underline{\omega}$ and $Trace([C_c]) \in \mathbf{R}$ can be negative.

Not being bound to be positive semi-definite could be a desirable feature for the change matrix since it allows to retrieve if the change in the partial target is increasing or reducing the power of a SM.

To find the maximum and minimum $\underline{\omega}$ we can apply again the Lagrangian optimisation for the quadratic form $\underline{\omega}^{*T}[C_c]\underline{\omega}$. That is:

$$\underline{\omega}_{max} = \underset{\underline{\omega} \in \Omega}{Argmax} \left[\underline{\omega}^{*T} [C_c] \underline{\omega} \right]. \quad (12)$$

By constraining $\underline{\omega}$ to be unitary we can obtain the Lagrangian as:

$$L = \underline{\omega}^{*T}[C_c]\underline{\omega} - \lambda(\underline{\omega}^{*T}\underline{\omega} + C). \quad (13)$$

After differentiating over $d\underline{\omega}^{*T}$ and setting the derivative equal to zero we obtain the equation:

$$C_c \underline{\omega} = \lambda \underline{\omega}. \quad (14)$$

The Lagrangian multipliers can be calculated by performing a diagonalisation of the matrix $[C_c]$. Since $[C_c]$ is Normal, the eigenvalues will exist and be real, but not necessarily positive. This is because a change can increase or decrease the resulting power of a scattering mechanism. Additionally the eigenvectors are orthogonal, which allow this decomposition to capture the minimal change (please see next section for more details).

We will refer to this decomposition as DIFFERENCE, or DIFF in short.

B.2 Constrained difference matrix

In [19] a constraint of the difference matrix is proposed. The rational is to allow the final change matrix to regain the property of being positive semi-definite. In this way, the matrix will represent a single physically feasible partial target. Clearly, one covariance matrix can only represent one partial target and therefore we need to define in which time direction the change has happened. In other words, we need to define if one partial target has been added to or removed from the scene.

The optimisation is again based on a diagonalisation of a matrix $[C_p]$, where:

$$\begin{cases} [C_p] = [C_2] - r_p[C_1] : \text{ if } r_p > r_m \\ [C_p] = [C_1] - r_m[C_2] : \text{ if } r_p < r_m \end{cases} \quad (15)$$

where,

$$\begin{cases} r_p = \max \text{ eig}([C_1]^{-1}[C_2]) \\ r_m = \dfrac{1}{\max \text{ eig}([C_1]^{-1}[C_2])} \end{cases} \quad (16)$$

The optimisation initially checks for the bigger r_p and r_m value, and then produces a covariance matrix representing a partial target that has been either added or removed. The constraint ensures that the resulting matrix is positive semi-definite [19].

We will refer to this decomposition as PARTIAL DIFFERENCE, or ParDIFF in short. The word "partial" wants to lead the attention on the fact that the constraint produces a change matrix that can represent a partial target (if the assumptions for the change are met).

III. Theoretical evaluation of signal models

The three previous decompositions are focused at very different typologies of changes, therefore the eigenvectors have very different interpretations. To interpret these, we need to analyse the assumptions of the different physical models.

A. Additive models: Unconstrained optimisation

In this model, the change is caused by an increase or reduction of power for a scattering mechanism. The detection hypotheses for the addition case are in the following (the subtraction can be achieved by flipping the role of 1 and 2):

$$\begin{cases} H_0 : [C_2] = [C_1] \\ H_1 : [C_2] = [C_1] + [C_c] \end{cases} \quad (17)$$

The Null hypothesis sees no change, while the Alternative hypothesis converts the first acquisition into the second by adding the partial targets enclosed in $[C_c]$.

We name this signal model as *Additive*, since the change is additive. This signal model covers all the circumstances where a target has been added or removed from the scene. For instance, in the agricultural context, we may have that the first scene only has scattering from the ground $[C_{1g}]$, while the second may have only horizontal dihedrals $[C_{2d}]$ due to plants. H_1 is true, with matrices $[C_c] = [C_{2d}] - [C_{1g}]$. The matrix $[C_c]$ will encapsulate that a dihedral scattering mechanism was added and a surface scattering mechanism was removed.

This decomposition is data driven and when the number of scattering mechanisms increases and surpasses the number of

observables (the 3 polarisation channels), the analysis becomes more complex and extra physical models need to be used to interpret the targets encapsulated in $[C_c]$ (as it is for the model based decompositions that can have more than 3 components) [14].

The decomposition DIFF tries to provide an answer to the Additive hypothesis testing.

In the circumstance that the additive model is a good fit to reality these results can be physically interpreted. The largest eigenvalue (if positive) represents the power of the largest scattering mechanism that has been added to the scene, while the minimum (if negative) is the power of the scattering mechanism that has been maximally removed from the scene. The maximum eigenvector represents the SM that has been added. The minimum eigenvector represents the SM that has been removed.

An interesting point is that the set of eigenvectors is orthogonal, because the matrix $[C_c]$ is Normal. This means that the eigenvectors also represent the minimal change that one could apply additively to transform the first partial target into the second partial target. It is minimal because the added and subtracted components are uncorrelated and therefore they have no projection on one another. If they were not orthogonal one could always add the same partial target to both $[C_a]$ and $[C_s]$, and the result would not change (this would also have made the decomposition not unique). Orthogonality therefore is an important property for this decomposition.

It is also important to stress that the eigenvectors provide complementary information from just knowing the covariance matrices before $[C_1]$ and after $[C_2]$. It is decomposing the change into its orthogonal (and therefore minimal and unique) components.

This theoretical result will be tested in the following with Monte Carlo simulations and real data.

B. Additive models: Constrained optimisation

In this model, the change is caused by either an increase or reduction of power of a partial target. The detection hypotheses for the addition case are:

$$\begin{cases} H_0 : [C_2] = [C_1] \\ H_1 : [C_2] = r[C_1] + [C_p] \end{cases} \tag{18}$$

The Null hypothesis sees no change, while the Alternative hypothesis converts a scaled first acquisition into the second by adding one partial target represented by $[C_p]$.

The change is additive although it is constrained by a scaling of the covariance matrix of the acquisition that is reducing. This signal model covers the circumstances where a unique partial target has been added or removed from the scene. For instance, in the agricultural context, we may have that the first scene only has ground component $[C_{1g}]$, while the second may have a combination of ground and horizontal dihedrals $r[C_{1g}] + [C_{2d}]$ due to plants. The scaling is needed to make sure that the change matrix $[C_p]$ can be interpreted as a unique feasible partial target.

$[C_p]$ is forced to be rank deficient (rank 2) by the constraint. It therefore only represents the 2 main spectral components of the added or removed target [19].

Since it is focused on a unique $[C_p]$ matrix, this decomposition is only valid when one unique partial target is either added or removed. The decomposition is resilient in dealing with situations when we have changes in both directions (part of a target is removed and a new target added), because r can be small to allow for reductions of one of the partial targets. However, if a partial target is completely removed $[C_1]$ the constraint will result in a very small r value. In fact, it can be proven that the only way to avoid negative eigenvalues for $[C_p]$ when we remove $[C_1]$ orthogonal to $[C_2]$ is by having $r_p = r_m = 0$. In general, the constraint will result in a $[C_p]$ that converges toward $[C_1]$ or $[C_2]$ when the assumption of a unique change is not met.

One may think to run two separate optimisations and obtain two matrices $[C_{p1}]$ and $[C_{p2}]$. This will not ensure that one of the resulting matrices may end up not being positive semi-definite (it was discouraged in [19]). Additionally, this will not remove the convergence effect. However, if we only focus on the dominant eigenvector of the two optimisations, the algorithm should still obtain estimations of the dominant change, providing the scattering mechanisms that suffer the maximum addition or subtraction.

To conclude, in the circumstance that the additive model is a good fit and a unique partial target is added or removed, the decomposition assumptions are met. The largest eigenvalue represents the power of the largest scattering mechanism that has been either added or removed. The maximum eigenvector represents the SM that has been added or removed.

This theoretical result is tested and validated in the following with Monte Carlo simulations and real data.

C. Multiplicative model

In this model, we consider that a change in the partial target is not due to adding or removing scattering mechanisms, but rather a transformation of the partial target itself. The detection hypotheses are:

$$\begin{cases} H_0 : [C_2] = [C_1] \\ H_1 : [C_2] = [A][C_1] \end{cases} \tag{19}$$

The change that transforms the second partial target into the first partial target is modelled as a multiplication by a transformation matrix $[A]$. We can derive that

$$[A] = [C_2][C_1]^{-1} \tag{20}$$

We can see that:

$$[A]^{-1} = ([C_2][C_1]^{-1})^{-1} = [C_1]^{-1}[C_2] = [C_m]. \tag{21}$$

The matrix $[C_1]^{-1}[C_2] = [C_m]$ is the change matrix used in the spectral decomposition. This means that the eigenvectors of $[A]$ are the same as $[C_m]$ and they represent transformations.

When the multiplicative model is a good fit to reality these results can be physically interpreted. The maximum eigenvector represents the direction in which the ellipsoid of the first partial target has been maximally stretched out to transform into the ellipsoid of the second partial target. On the other hand the minimum eigenvector is the direction in which the ellipsoid has been squashed (compressed) to become the second ellipsoid. The eigenvalues present the ratio between the ellipsoid surface (i.e. the power) in the directions where the stretching happened.

Please note, it is not possible to interpret these eigenvectors as scattering mechanisms since they are directions of stretching for the partial targets. It is therefore more similar to a signal decomposition than a classic target decomposition and the optimal $\underline{\omega}$ are not SM.

This theoretical result will be tested in the following using Monte Carlo simulations and real data.

D. When to use the different decompositions

(i) Additive unconstrained: If partial targets have been added or removed than DIFF will allow to classify the change. Examples of such scenarios include several environmental changes such as deforestation (trees are removed), coastal erosion (the coast line has been replaced by water), some agricultural crops (ground is substituted by small stems, larger plants and ground again). This is also useful for artificial targets (a vehicle appearing or disappearing) and urban (building built or demolished). It is important to note that this model also fits situations when the "visible" change is not clear. For example in agriculture when the plants get dry and the ground underneath get visible. Or when different moisture levels change the amount of surface scattering.

(ii) Additive constrained: If we have a priori knowledge that a unique partial target has been added or removed then ParDIFF will allow to classify the change in a better way. Examples include deforestation of sparse forests (we need to see the ground when the forest is present) and changes on crops. Practically we still need to have portions of both partial targets $[C_1]$ and $[C_2]$ in both images to avoid convergence of the change matrix into the two original partial targets $[C_1]$ and $[C_2]$. Using a modification of ParDIFF that applies two optimisations instead of one, we can provide results similar to the unconstrained additive model, however it still suffers from convergence of the change matrix.

(iii) Multiplicative: This decomposition fits when the same target is modified and the parts that appear as being added or removed do not constitute proper interpretable polarimetric targets. For instance, if we rotate a car of 20 degrees, it is very likely to show different scattering mechanisms, but it is hard to argue that these are added to the scene, since the real point target producing them is the same. In this case a multiplicative model is more appropriate, although retrieving information from the transformation may still be not straightforward and it may need the development of physical models.

Concerning detection capability, we expect the larger dynamic range of the RATIO to produce better point detection results. For instance, if we consider the Signal to Clutter Ratio (SCR) as an indicator of separability, then RATIO optimises this and therefore provides the highest SCR achievable. However, SCR is not the only parameter that impact detection performance and the standard deviation of target and clutter are also very important.

In the following, we are performing Monte Carlo simulations and use real data to test these theoretical considerations.

IV. Test with Simulated data

A. Monte Carlo Simulations

In this simulation, we assume fully developed speckle with scattering vectors that can be modelled using 3 dimensional zero-mean complex Gaussian. This assumption makes interpretation of the results easier, but interestingly it can also cover classification results when the texture does not change the structure of the covariance matrix.

The generation of the random variables is mostly following the processing steps proposed in [15]:

(1) A Monte Carlo method is used to generate N realizations of scattering vectors drawn by a 3D-Complex Circular Gaussian distribution. This is performed twice (one for each acquisition). Therefore, we generate two sets of "white random vectors": $\underline{k}_1^w(i)$ and $\underline{k}_2^w(i)$, where $i = 1, ..., N$.

(2) The white random vectors are coloured using two asymptotic covariance matrices $[\dot{C}_1]$ and $[\dot{C}_2]$ (representing the partial targets observed in the two acquisitions): $[\dot{C}_1]^{\frac{1}{2}}\underline{k}_1^w(i) = \underline{k}_1(i)$ and $[\dot{C}_2]^{\frac{1}{2}}\underline{k}_2^w(i) = \underline{k}_2(i)$

(3) The simulated covariance matrices are calculated by averaging N realizations of the outer product of simulated vectors: $[C_1] = \langle \underline{k}_1 \underline{k}_1^{*T} \rangle_N$ and $[C_2] = \langle \underline{k}_2 \underline{k}_2^{*T} \rangle_N$, where $\langle . \rangle_N$ is the finite average of N realizations. Since the simulation provides random variables that are close to be independent, the value of ENL can be approximated by N.

(4) Points 1 to 3 are repeated K times to provide more samples for evaluating statistics (e.g. they represent K experiments). In other words, a set of K covariance matrices is produced: $[C_1(k)]$ and $[C_2(k)]$, with $k = 1, ..., K$. Each of these realizations is slightly different due to speckle.

(5) For each pair of the K experiments (e.g. $[C_1(k)]$ and $[C_2(k)]$) we evaluate the diagonalisations: $([C_2(k)] - [C_1(k)])\underline{\omega} = \lambda\underline{\omega}$ and $[C_1(k)]^{-1}[C_2(k)]\underline{\omega} = \lambda\underline{\omega}$..

To summarize, K samples are generated averaging N realizations and for each sample the change optimisations are applied. We therefore have two K sets of eigenvalues and eigenvectors, one set for each change detector. A block diagram of the simulation procedure is presented in Figure 1.

B. Classifying changes: Eigenvector analysis

In this section we test the capability of the three optimisations to classify a polarimetric change. In all the following simulations we use an Equivalent Number of Look ENL equal to 50. This may correspond to an averaging window of around 11×11 for most of the current sensors. Since the purity of polarisation of the partial target is very important, we perform simulations where we modify the entropy of the considered targets. The α values provided in the following are refereed to the dominant component of the partial target using Pauli basis. We use this because all the optimisations focus on the dominant changes (or the first two dominant in case of ParDIFF) and α allows a clear and synthetic representation of the target characteristics.

In this section we will be showing figures where the horizontal axis is the Realisation (or repeat of the experiment), while the vertical axis is the dominant α angle that has been retrieved by the algorithm. The captions give information about the actual α values that are simulated in each of the plots.

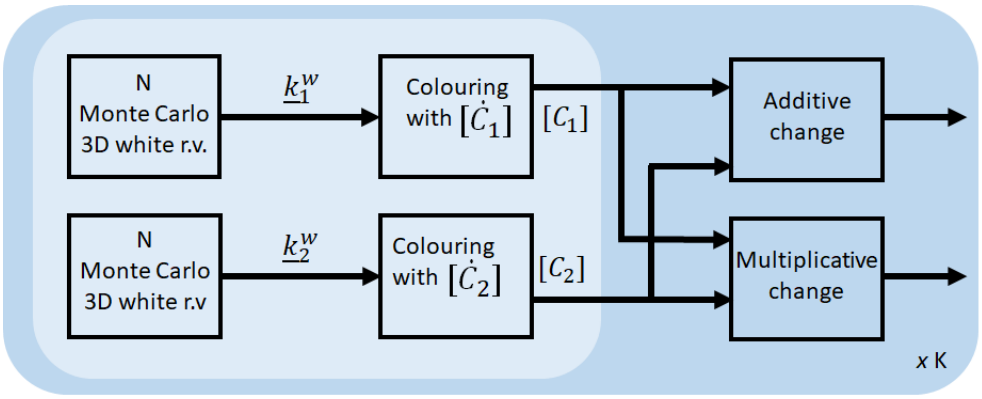


Fig. 1. Block diagram for the Monte Carlo simulations. r.v.: random variables. $\times K$ (bottom right of shaded areas) indicates the number of times the operation is repeated.

B.1 Adding one target

In the first test, the second partial target is obtained by adding 1 partial target to the first one. This follow the hypothesis H_1 : $C_2 = C_1 + C_a$, where C_a is a unique partial target. This physical change meets the assumptions for both DIFF and ParDIFF. We want to assess the retrieval of the dominant SM of C_a.

Polarised targets: Figure 2 shows a simulation where both targets C_1 and C_a are polarised (almost single targets) with an entropy of 0.1. The dominant α of the first target C_1 is zero (a surface) and the dominant α of C_a is swept from zero to 90 degree (from surface to dihedral). A good retrieval will therefore estimate an α that goes from 0 to 90 along a line throughout the different realisations. The maximum eigenvector will estimate the dominant α of C_a and C_p, since this is the SM that has been maximally added to the scene. In the plot, this is shown as a red line. We can observe that DIFF and ParDIFF are able to identify the polarimetric characteristic of the added target, where the red line follows a ramp in the plot. On the other hand, the RATIO is not able to retrieve the correct α. Looking at the estimation errors it seems both difference optimisations do not show any evident pattern and provide decent estimations with errors that are overall smaller for the ParDIFF with a RMS of 3.9 degrees compared to 4.7 for DIFF. Please note, ParDIFF minimum eigenvector is not shown because the matrix C_p is rank deficient. Additionally the DIFF minimum eigenvector (the target removed) is unstable (blue line), because nothing has been "removed" and the third eigenvalue is practically zero.

Partially polarised target: Figure 3 shows a more challenging scenario when the entropy of $[C_1]$ and $[C_a]$ is 0.5. Please note that in this simulation we are not adding a single target, but a proper partial target (formed by 3 single targets). DIFF and ParDIFF can only evaluate a portion of the total number of SM as discussed previously. This is why we are showing and concentrating only on the dominant change. As for the previous case, DIFF and ParDIFF can retrieve the added α, while RATIO provides completely unrelated values. Looking at the RMS error this has grown, where now DIFF can obtain a better result with an average 9.2 degree of error while ParDIFF has 12.8 degree.

Depolarised target: The final test in Figure 4 shows the same simulation when the entropy of C_1 and C_a is increased to 0.99. There is very little polarimetric information left and therefore retrieving any SM is very challenging. We can conclude that almost completely depolarised targets do not allow to retrieve α changes when the equivalent number of looks is 50. Things improve when averaging much more (e.g. $N = 200$), but these simulations are excluded for the sake of brevity. The RMS errors are now 27.8 degree for DIFF and 28.5 degree for ParDIFF.

The preliminary analysis shows that if the physical change is the addition of a unique partial target, DIFF and ParDIFF are able to retrieve the dominant SM that has been added to the scene, provided the entropy is not too high.

B.2 Adding and removing targets

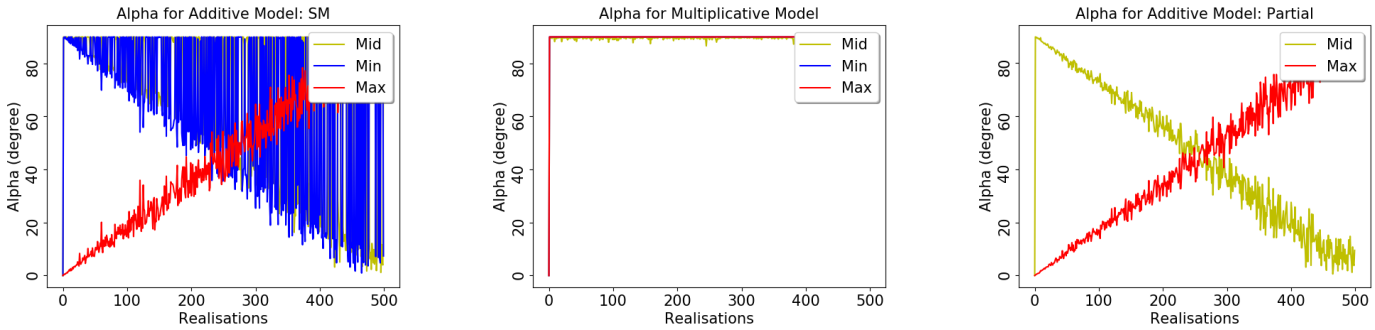
In this more challenging scenario we want to test if we can retrieve changes to two (and not one) partial targets. Specifically we are adding a partial target AND removing another partial target. Note, this is the most general additive change and this assumption is not met by ParDIFF, which is only modelling the change to a unique partial target.

Please note in this experiment we are adding AND subtracting, considering both time directions (e.g. a dense forest, which becomes bare ground, the forest is subtracted and the ground is added). If we add (subtract) more partial targets in the same time direction the superposition of effects will make it appears as a unique final partial target. That is $H_1 : [C_2] = [C_1] + [C_{a1}] + [C_{a2}] \leftrightarrow [C_2] = [C_1] + [C_a]$, where $[C_a] = [C_{a1}] + [C_{a2}]$. In this case, ParDIFF will meet its assumptions and clearly, the only way to unpack the partial targets of origin $[C_{a1}]$ and $[C_{a2}]$ is by using scattering models.

The alternative hypothesis is now equal to $H_1 : [C_2] = [C_1] + [C_a] - [C_s]$. The targets we want to retrieve now are two, the added $[C_a]$ and the subtracted $[C_s]$ which we will analyse using the dominant α parameters for the maximum and minimum eigenvectors respectively.

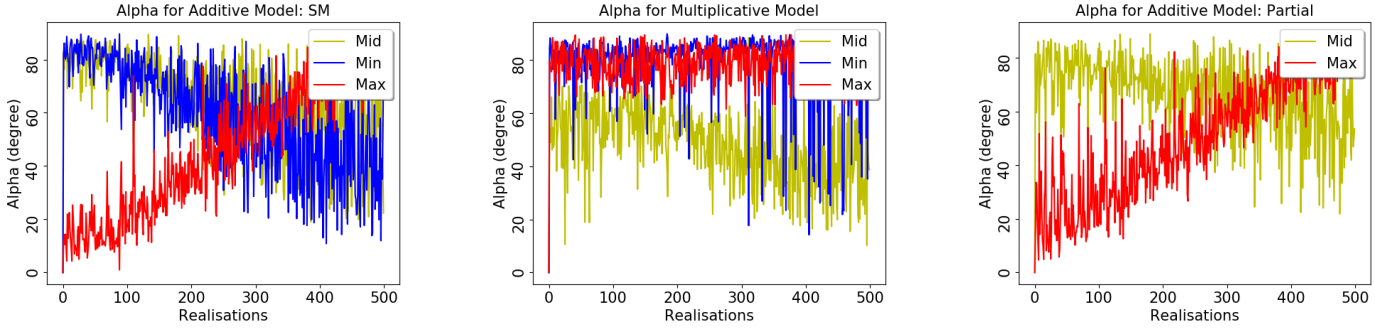
In the first test we select two orthogonal targets, where the dominant α of C_1 is zero, the α of C_a is 90 degree (a dihedral) and the α of C_s is 0. In this way, we have that $C_s = C_1$ and $C_a = C_2$. In practice, we are substituting a surface with a dihedral, both of equal power.

Polarised targets: Figure 5 shows a plot where the entropy of all the targets is 0.1 (they are almost single targets). Interestingly, DIFF can retrieve both the targets added $\alpha = 90$ (red line) and removed $\alpha = 0$ (blue line). The RATIO cannot retrieve



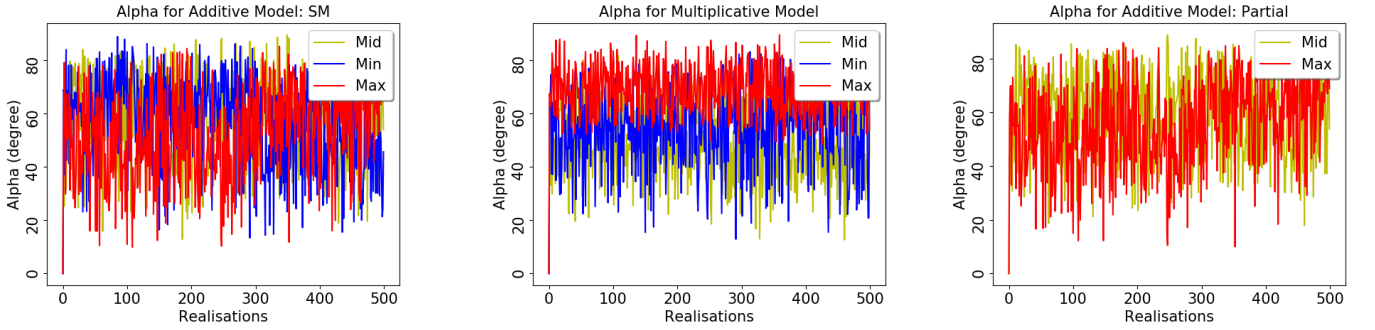
(a) DIFF (b) RATIO (c) ParDIFF

Fig. 2. Adding one target. Simulation of retrieved α values when the (a) DIFF, (b) RATIO and (c) ParDIFF optimisations. The dominant α of the target that has been added is swept from 0 to 90 degree. Entropy of targets = 0.1



(a) DIFF (b) RATIO (c) ParDIFF

Fig. 3. Adding one target. Simulation of retrieved α values when the (a) DIFF, (b) RATIO and (c) ParDIFF optimisations. The dominant α of the target that has been added is swept from 0 to 90 degree. Entropy of targets = 0.5



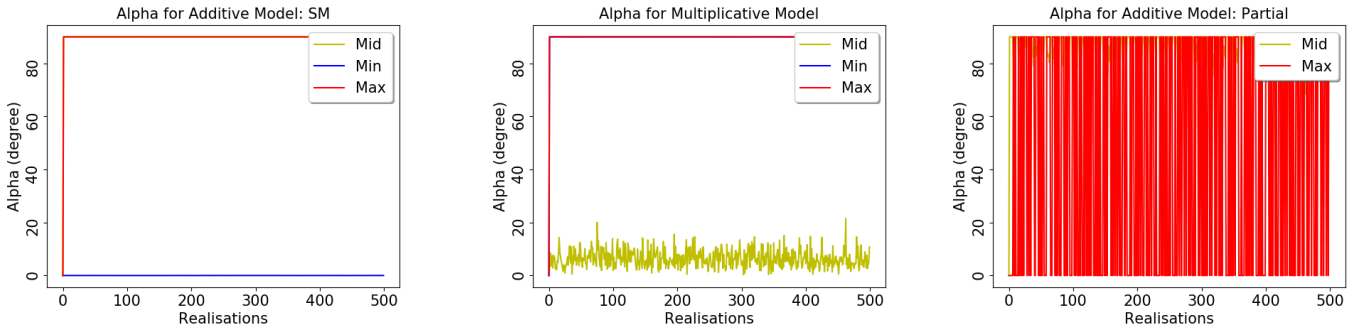
(a) DIFF (b) RATIO (c)

Fig. 4. Adding one target. Simulation of retrieved α values when the (a) DIFF, (b) RATIO and (c) ParDIFF optimisations. The dominant α of the target that has been added is swept from 0 to 90 degree. Entropy of targets = 0.99

the α for the minimum, but the maximum is correct. The ParDIFF presents a very peculiar behaviour where the maximum eigenvector jumps randomly between the target that has been added and the one that has been removed. The algorithm has not enough information to evaluate if a target has been added or if it has been removed and selects r_m or r_p based on consideration on total power [19], which in this case is simulated as being in average the same. However, if we fix the time direction then the estimations become comparable to DIFF with RMS around 0.005 degrees. The plots with ParDiff applied fixing the time direction are not shown for brevity.

Partially polarised targets: The same test was done when the entropy of all the targets is 0.5. For the sake of brevity the figure is not shown, but the plots of α are very similar (just with an increased spreading). Again, DIFF retrieves both the dominant added target $\alpha = 90$ and the removed one $\alpha = 0$. The quality of the retrieval is still very good with estimation error of 2.8 degree (for DIFF). This shows that DIFF is able to cope with situations when one target is added and one removed. Again ParDIFF suffers from the same ambiguity where the algorithm does not know which direction to select and only detects "half" of the change. By fixing the time direction and running the algorithm twice we can retrieve results just as good as the DIFF.

Depolarised targets: The Figure when the entropy of all the



(a) DIFF (b) RATIO (c)

Fig. 5. Adding and removing targets. Simulation of retrieved α values when the (a) DIFF, (b) RATIO and (c) ParDIFF optimisations. The added α is 90 degree, the removed α is 0 degree. Entropy of target equal to 0.1.

targets is 0.99 is not shown for sake of brevity. Not surprisingly the retrievals are very poor. However it is still possible to see some pattern in the DIFF retrieval where we have a mean error of 24.1 degrees. Interesting these results are much better than the one obtained when only one target was added. The reason for this is possibly that the change now is more marked than before and even a tiny amount of polarimetric information is magnified and more visible.

B.3 Adding and removing targets: non orthogonal

In the final test, we consider non orthogonal targets. This means that the subtraction is not removing completely C_1 and the addition is not adding completely C_2. In particular we have a surface which becomes a dipole. The dominant α for C_1 is 0, while the dominant α for C_2 is 45 degree. Clearly, mathematically there is no unambiguous combination of C_a and C_s that could have produced that change. We could consider an infinite number of different combinations that could have lead to such result. What the optimisation is able to tell us is the minimal addition C_a and subtraction C_s that would have transformed C_1 into C_2. This by definition leads to dominant scattering mechanisms in C_a and C_s that do not have projections on each other (i.e. they are orthogonal).

If in nature the change is obtained in a different way, we simply don't have enough information in the SAR data alone to disentangle this. However, we can always consider a scattering model that is able to start from the minimal changes and project these onto some target basis that we believe are responsible for the physical scattering. This can be done for instance in agriculture [30].

Polarised targets: Figure 6 shows the simulation using non-orthogonal targets with entropy equal to 0.1. The α of the added target is swept between 0 and 90 degrees, while the removed target has always an alpha equal to 0. We observe that the maximum and minimum of DIFF are not equal to the targets before and after. They are equal to the minimal change that would lead the first target to become the second. These two SM will modify the surface and make it into the second target. Interestingly, if we look at the envelop of the red signal (top and bottom of red area), ParDIFF seems to be able to cope with this estimation, provided we fix the direction of time. This is because the operation that is making $[C_s]$ positive semi-definite is recognising the assumptions are not met and estimating values of r_m and r_p very close to zero (around 10^{-7} in this experiment). This mean that $[C_s] \approx [C_1]$ or $[C_s] \approx [C_2]$. The change matrix converged to the partial targets in the two scenes. Clearly the same result could have been obtained doing a diagonalisation of $[C_1]$ or $[C_2]$ in the first place, but ParDIFF has the advantage to test the validity of adding a unique target and converges into the analysis of the separate images only when its own assumptions are not met.

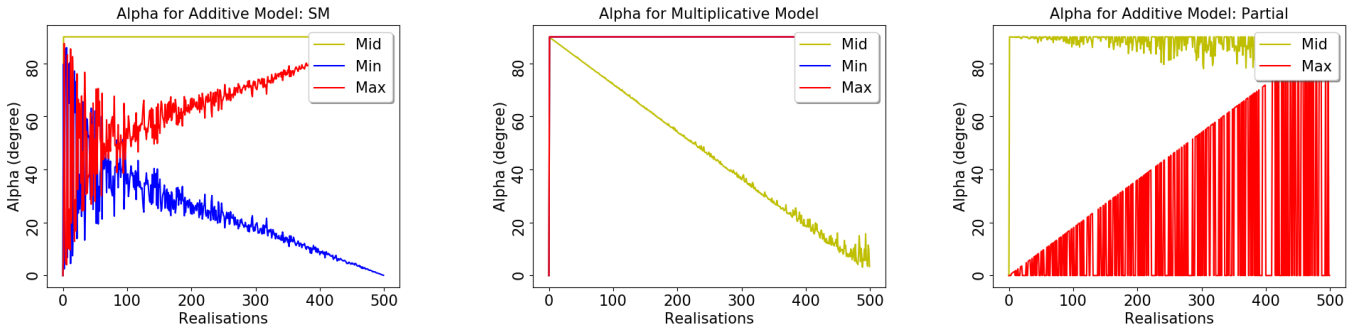
Partially polarised targets: Figure 7 shows the simulation using non-orthogonal targets with entropy equal to 0.5. As for the previous case, DIFF provides the α of minimum and maximum change, while ParDIFF is able to recognise the model for a sole partial target is not true and returns the eigenvectors of the two matrices separately. Now the values of r_m and r_p are around 0.15. The differences between the estimation errors of DIFF and ParDIFF (provided we apply ParDIFF twice) start getting much smaller, because the change matrix needs to converge less toward the original targets.

Depolarised targets: similarly to previous cases, none of the algorithms, can estimate the scattering mechanisms and the results between DIFF and ParDIFF are very similar. The plots are not showed here for the sake of brevity.

B.4 Summary of Classification

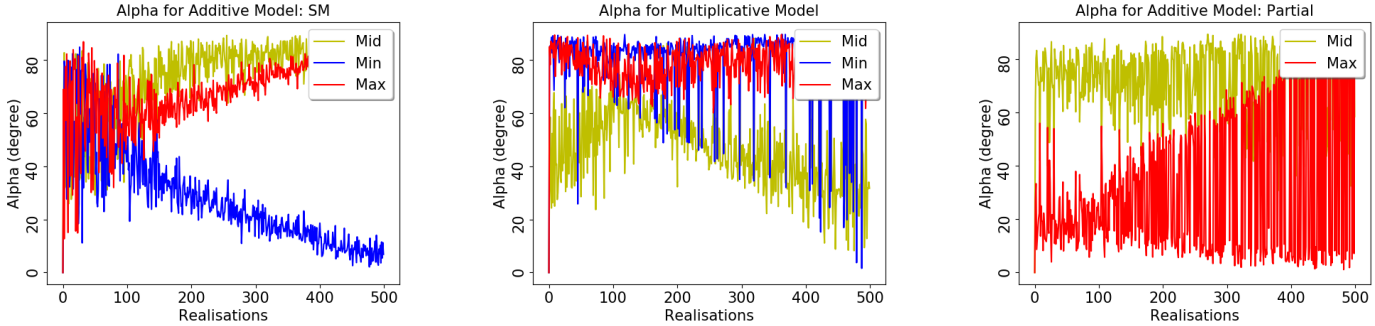
From the simulation we could produce the following conclusions:

- Additive models: adding OR removing
  - DIFF can retrieve the 2 dominant SM that have been added and/or removed.
  - ParDIFF can retrieve 1 dominant SM either added or removed. If repeated twice disabling the automatic selection of the time direction, than it can retrieve the 2 dominant SM.
- Additive models: adding AND removing
  - DIFF retrieves the minimal changes to SM to go from $[C_1]$ to $[C_2]$. This information is complementary from knowing $[C_1]$ and $[C_2]$.
  - ParDIFF recognises its assumptions are not met and retrieves the two partial targets of origin or end (due to the change matrix converging toward the original partial targets).
- RATIO produces eigenvectors that cannot be interpreted as SM.



(a) DIFF (b) RATIO (c)

Fig. 6. Adding and removing targets. Simulation of retrieved α values when the (a) DIFF, (b) RATIO and (c) ParDIFF optimisations. The added α is swept between 0 and 90 degree, the removed α is 0 degree. Entropy of target equal to 0.1



(a) DIFF (b) RATIO (c) ParDIFF

Fig. 7. Adding and removing targets. Simulation of retrieved α values when the (a) DIFF, (b) RATIO and (c) ParDIFF optimisations. The added α is swept between 0 and 90 degree, the removed α is 0 degree. Entropy of target equal to 0.5

- When the entropy approaches 1 all the algorithms fail.

C. Detection: Eigenvalue analysis

In this second simulation we want to test the detection capability of the three change optimisations. To do this we will produce Receiving Operating Characteristics curves for different situations and Signal to Clutter Ratio (SCR) values. These are plots showing Probability of Detection P_d against Probability of False Alarm P_f. The probability are calculated simulating realisations of targets (H_1) and clutter (H_0).

C.1 Adding one target

The SCR is calculated considering the trace of C_a as our signal and the trace of C_1 as our clutter. This is because we want to detect an added target, embedded in the target of the first acquisition. For instance a $SCR = 0.5$ means that the trace of C_a is half the one of C_1. We simulated 10,000 instances of the first target and 10,000 instances of the second target, to calculate P_d and P_f.

In these simulations as benchmark we considered the Whisart change detector based on the likelihood ratio [18]. This was developped by Conradsen et al [18] and it is sometime refereed to as the Conradsen change detector. In the following figures we will show ROC curves, where we express the probability of false alarm P_f with a base 10 logarithmic scale. In each of these plots, the color of the line identifies the algorithm, while the style (solid, dashed, dotted) identifies the value of Signal to Clutter Ratio (SCR).

Polarised targets: Figure 8 shows the ROC curves when $[C_1]$ and $[C_a]$ have an entropy equal to 0.1 (note $[C_2] = [C_1] + [C_a]$). The solid line is for $SCR = 0.5$, dashed line is for $SCR = 0.1$ and the dotted line (appearing as thicker in the plot) is for $SCR = 0.05$. The black line (the lowest performance possible) represents the random assignment (flipping a coin and deciding if it is target or clutter). Please note the scale for P_f is in log_{10}.

Interestingly, the performances of RATIO and Wishart are top of the scale for $SCR = 0.5$ and for $SCR = 0.1$. They provide perfect detection in our simulations (i.e. $P_d = 1$ and $P_f < 10^{-4}$) and they cannot be seen because they overlap on each other at the top of the ROC plot (only visible a blue line, the green line is underneath). Please note with 10,000 simulations we can only estimate $P_f < 10^{-4}$. Since this is a good limit for change detection we didn't simulate more samples which would have lead to a slow down of the computation. The DIFF is providing lower detection performance, which we may have expected considering the ratios have larger dynamics. Finally ParDIFF is providing the worst performance. This can be justified by the fact that to produce a positive semi-definite change matrix we reduce the difference (r_s and r_m are generally smaller than 1). The power of the test is therefore reduced while this is not affecting false alarms.

Also note that when the SCR is very low (the added target is 5% of the initial target) the detection performance is very low for all the detectors resembling the random assignment.

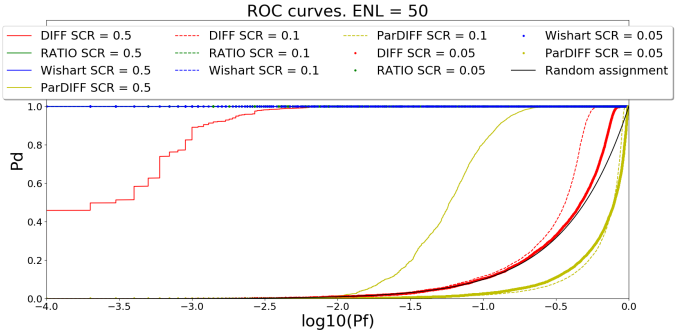


Fig. 8. ROC curves, one target added. Simulation for different SCR when the first target is a surface and the target added has a dominant $\alpha = 90$ degree. Entropy of targets equal to 0.1

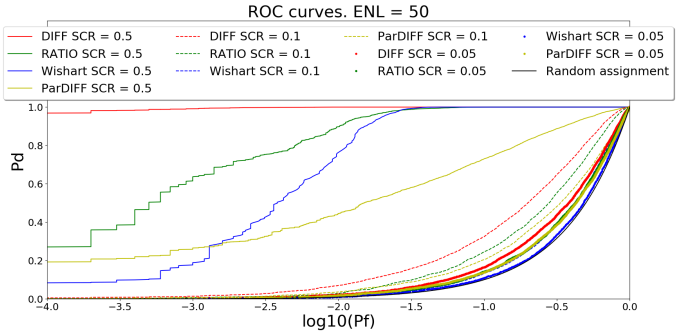


Fig. 10. ROC curves, one target added. Simulation for different SCR when the first target is a surface and the target added has a dominant $\alpha = 90$ degree. Entropy of targets equal to 0.99.

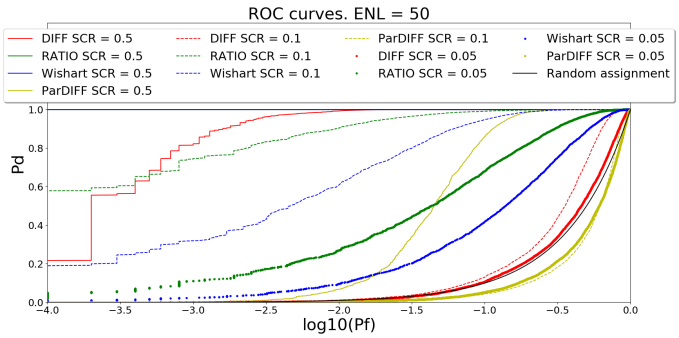


Fig. 9. ROC curves, one target added. Simulation for different SCR when the first target is a surface and the target added has a dominant $\alpha = 90$ degree. Entropy of targets equal to 0.5.

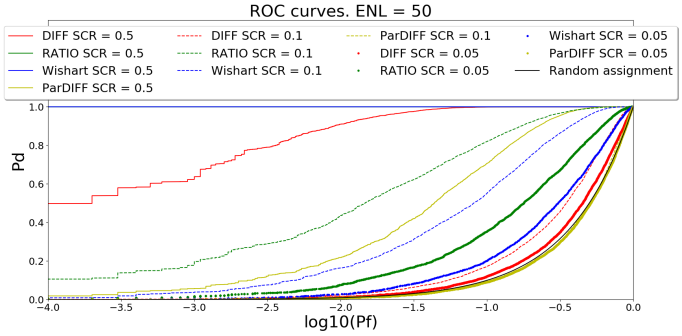


Fig. 11. ROC curves, one target added. Simulation for different SCR when the first target is a surface and the target added has a dominant $\alpha = 45$ degree. Entropy of targets equal to 0.5.

Partially polarised targets: Figure 9 shows the ROC curves when C_1 and C_a have an entropy equal to 0.5. Again the RATIO and Wishart provide the best results and performance degrades when the SCR reduces (green and blue lines overlap again). For $SCR = 0.5$, we have that RATIO and Wishart obtain the perfect performance, with DIFF again inferior in detection performance. For $SCR = 0.1$, RATIO and Wishart have a degraded performance and we can clearly see that RATIO performs better than Wishart. For $SCR = 0.05$ no algorithm is able to provide decent performance.

Depolarised targets: Figure 10 shows the ROC curves when C_1 and C_a have an entropy equal to 0.99. A surprising thing happens here when the SCR is high enough ($SCR = 0.5$). Once polarimetry becomes less important because the depolarisation is really high, DIFF performs much better than other algorithms. When it comes to detection, DIFF sees more separability when the polarimetric information is less important and the discrimination is done purely on the total intensity of the partial targets. This is a result that we will see in real data as well. However, the good detection performance of DIFF can only be achieved when the SCR is large enough. If $SCR = 0.1$, no detector can provide a decent performance.

To summarise these first results, when we observe single targets (e.g. vehicles, small buildings), then we will have larger discrimination if we use RATIO or Wishart. On the other hand, when dealing with high entropy systems, we would get better detection performance by using DIFF.

Finally in Figure 11 we can observe the ROC curves when we add a dominant dipole $\alpha = 45$. The entropy is 0.5. We can see that the performance reduces especially for RATIO and Wishart with lower SCR, but the ranking of the detectors remains the same.

C.2 Adding and removing targets

In this section, we want to simulate the general situation when we add and remove targets. In the following simulations we always start with an initial target $[C_1]$ equal to a surface and modify this into a second target $[C_2]$ with a variable dominant α (choosing between 90, 45 and 22 degree).

Polarised targets: when the entropy is 0.1 we can obtain perfect detection for all the algorithms. The ROC curve is not shown for sake of brevity.

Partially polarised targets: The situation is identical to Figure 12 and not showed here for the sake of brevity. An entropy of 0.5 still provides perfect detection (i.e. $P_d = 1$ and $P_f < 10^{-4}$). Targets with a difference of 22 degree cannot be discriminated using 50 ENL.

Depolarised targets: If we simulate very depolarised targets with entropy 0.99 (Figure 13) we discover that all the detectors perform very bad. The problem is that there is very small polarimetric difference with entropy 0.99 and the SCR of the two targets is identical. Therefore DIFF cannot separate the targets based on the overall power as it was doing before for depolarised targets with different Trace. To have separability when there is no polarimetric information we need a difference in overall intensity.

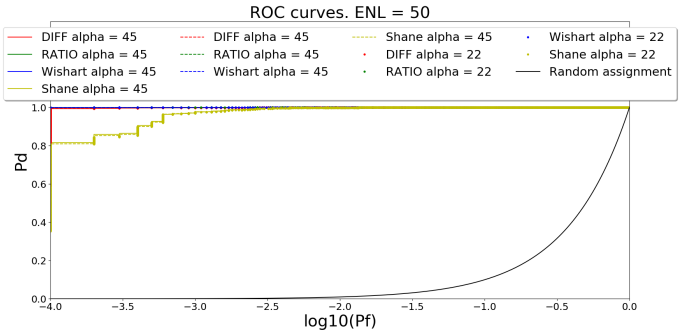


Fig. 12. ROC curves, targets added and removed. Simulation for different α of the second target (90, 45 and 22 degree). Entropy of targets equal to 0.5.

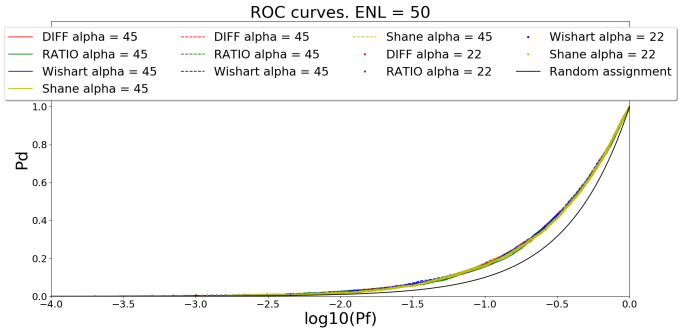


Fig. 13. ROC curves, one target added. Simulation for different SCR when first target is a surface and the α added is 45 degree. Entropy of target 0.99.

V. Tests with real data

A. Data Presentation: SARTOM E-SAR (DLR)

The first dataset was acquired by the E-SAR DLR system and it is L-band quad-polarimatric. The data collection called SARTOM [31] was performed in 2006 and it was mainly aimed at target detection using tomography and polarimetry. Interestingly, several targets were deployed in open field and under vegetation. Additionally, some of the targets were moved during the acquisitions. Figure 14 shows the RGB Pauli images of the test area for two separated acquisitions. The spatial baseline is zero (in average) and the temporal one is four days. Several targets among two trihedral corner reflectors (CR) in open field and one CR in the forest were removed during the four days.

B. Change models on SARTOM

The three optimisations were applied to the SARTOM data and Figure 15 shows the images of the eigenvalues.

All the targets with validated changes can be easily detected in the maximum eigenvalues of the RATIO (some in the maximum and some in the minimum). This includes the corner reflectors under canopy cover. Interestingly, using only the backscattering the latter cannot be detected. Moreover, some further point targets have a relatively large RATIO. Note that the northernmost of the corner reflectors was positioned on the ground without pedestal. This was removed at the end of the first acquisition and then replaced on the ground four days after. It was not used for calibration purposes, but it only functioned as a target for detection. We can still detect the change, because its different positioning made it look different in the two images (looking at the Pauli RGB images reveals that the color around the corner reflector looks different in the two images).

Targets are also visible in the DIFF and ParDIFF eigenvalues, however the forest clutter is very high and an automatic change detector may struggle separating the point targets change from the background. Please note the scale of these two are different and this explains most of the differences in colours.

Figure 16 shows the retrieved α using the two optimisations. Before interpreting these results, it is important to keep in mind that α is a scale invariant measurement and it will be retrieved regardless of the intensity of the eigenvalue. This makes α images look very noisy. In the following we propose a better way to visualise the α, however it is interesting to have a look at α everywhere for testing purposes.

Looking at DIFF and ParDIFF, the images show a reduction of surface scattering on the fields. As expected, the α retrieved by the ratio, does not seem to be interpretable as a scattering mechanism.

Figure 17 is build by using α as the colour and the eigenvector as the saturation. A pixel will keep the colour given by α, but appear very dark if the eigenvalue is low. Looking at the moved targets we can observe how the removed scattering mechanisms (smallest eigenvalue) resemble in colours the Pauli RBG of the first acquisition. The corner reflectors appear mostly blue, although the northernmost one has a strange behaviour when redeployed showing different colours. The CR in the forest is very hard to identify, due to the large clutter represented by forest changes. From the simulation we saw that changes of depolarised targets seem to be enlarged by DIFF.

Finally Figure 19 presents the same colour image where the RATIO optimisation is performed. Interestingly, the detection of point targets is improved using the ratio, but the colours do not suggest any possibility to perform classification.

Figure 18 presents the same colour image where the ParDIFF model is performed. Interestingly the RGB seems very similar to a mix of the Largest and Smallest RGBs of the DIFF. Simulations did show that both DIFF and ParDIFF are able to identify the changing scattering mechanisms. Finally Figure 18 also presents a change detection image obtained using the Wishart change detector tested in the simulation section. Obviously, the Wishart detector will not be able to classify the change (it is black&white). Additionally some of the targets, as the ones in the forest are almost invisible in the image, showing that in real data RATIO may have some advantage to the Wishart. This could be linked to the fact that simulations were assuming a Wishart distribution of the target and clutter, while real data may have a different distribution due to texture.

Finally we want to investigate further the similarity of DIFF and ParDIFF on real data. To do so, Figure 20 plots a 2D density distribution of DIFF against ParDIFF estimations. If the density is on the bisect then the two results are identical. The three plots represent the three elements of the Pauli decomposition that we used for the RGB. The estimations are in pretty good agreement, as indicated by the RMS and bias. Although ParDIFF tends to estimate more surface change than DIFF. From this comparison we can conclude that indeed the two algorithms are looking at very similar scattering mechanisms.

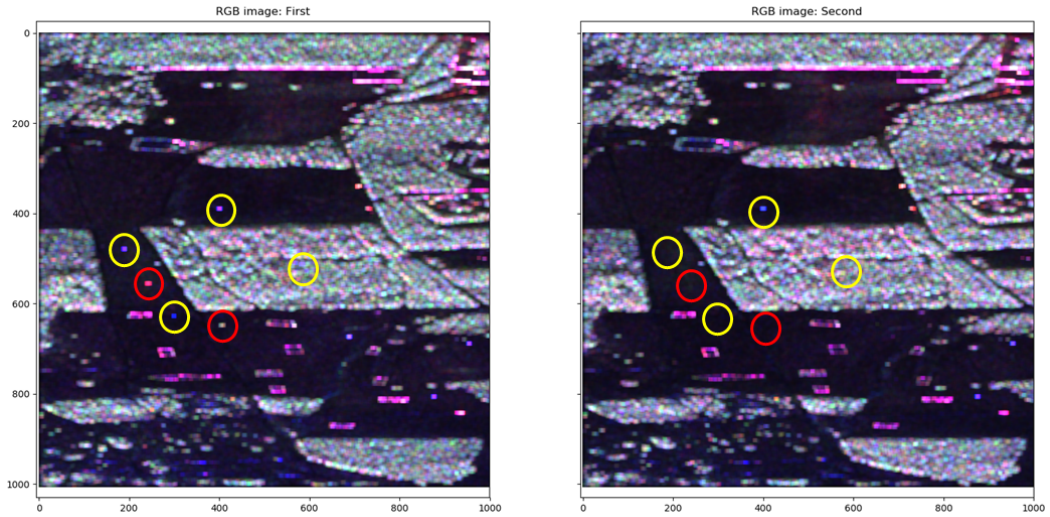


Fig. 14. RGB Pauli images of the SARTOM test site with 4 days of temporal baseline: (a) Master acquisition (b) Slave acquisition (4 days after). Circles indicate targets that were removed/moved. Yellow is for corner reflectors, while Red were jeeps (DLR E-SAR L-band SARTOM2006 Campaign).

C. Data Presentation: ALOS-1 (JAXA)

The second dataset was acquired by the PALSAR instrument on board of the JAXA ALOS satellite. The images analysed here were acquired in 2007 in the west of the UK over the Morecombe Bay.

The first image was acquired on the 1^{st} of April, while the second was on the 17^{th} of May 2006. The area presents a very dynamic portion of the coast (with a salt marsh in the red rectangle) and a caravan park included in the red circle.

The Pauli RGB of the two images are presented in Figure 21. It is possible to observe how the sea region is showing some change due to different wind conditions. Additionally a large area of agricultural fields near the coast shows an increased green colour (indicating an increase in volume scattering). The caravan park also presents several point target changes.

C.1 Change detection with ALOS data

The resolution is $30 \times 10m$ in slant range and azimuth. The coherency matrices were filtered with a 3x9 boxcar and multilooked to produce a pixel on the ground that is more squared. For the sake of brevity we only present the final RGB images.

In Figure 22 we present the α angle saturated by the eigenvalues for the DIFF optimisation. Interestingly, DIFF is able to classify the changes to scattering mechanisms, for instance identifying an increase in surface scattering over the sea. Also DIFF identifies an increase of volume scattering over the agricultural areas. It monitors an increase and reduction of dihedral scattering over the caravan park (due to large moving vehicles).

Figure 23 presents the same colour image where the RATIO is performed. These images cannot be used to classify the changes. Additionally, they seem more appropriate for point target detection.

Figure 24 presents the same colour image where the ParDIFF is performed. Again the image seems to be a mix of the two DIFF images. The Wishart change detector is also presented and clearly it can detect but not classify the change.

To test again the similarity of outputs between DIFF and ParDIFF Figure 25 plots a density map of estimations. Again the three plots are the three elements of the Pauli decomposition. The RMS and bias show minimal differences of more than an order of magnitude smaller than the actual values. From this comparison we can again conclude that the simulations where realistic and DIFF and ParDIFF produce similar results because they are observing the same underlying physical phenomena.

CONCLUSIONS

In this work, a validation and comparison of three optimisations for change detection with PolSAR was carried out. One of these optimisations was also first published in this article. The algorithms are based on diagonalisations of change matrices and therefore they can be referred as spectral decompositions of the change matrix. These have the aim to separate the different components of signal representing a polarimetric change. The proposed decomposition is aimed at optimising the DIFFERENCE (or DIFF) of covariance matrices $[C_a] = [C_2] - [C_1]$ while the second considers the RATIO of powers which lead to the change matrix $[C_m] = [C_1]^{-1}[C_2]$. The last was recently proposed [19] and aims at constraining the change matrix of DIFF to be positive semi-definite. This decomposition is here referred as ParDIFF.

The DIFF and ParDIFF decompositions can be used to provide physical interpretation to additive changes. This is when scattering mechanisms are added or subtracted from the scene. The constrain applied to ParDIFF assumes that only one partial target is either added or removed, while DIFF does not need this assumption. Finally, RATIO could provide physical interpretation only when a transformation modifies the first partial target into the second one (e.g. a change in orientation).

It is important to notice that these are signal decompositions focused on Changes and not Targets. They therefore have different properties from ordinary Target Decompositions [13], [11].

The three decompositions were validated using Monte Carlo simulations. It was evident that DIFF can retrieve the domi-

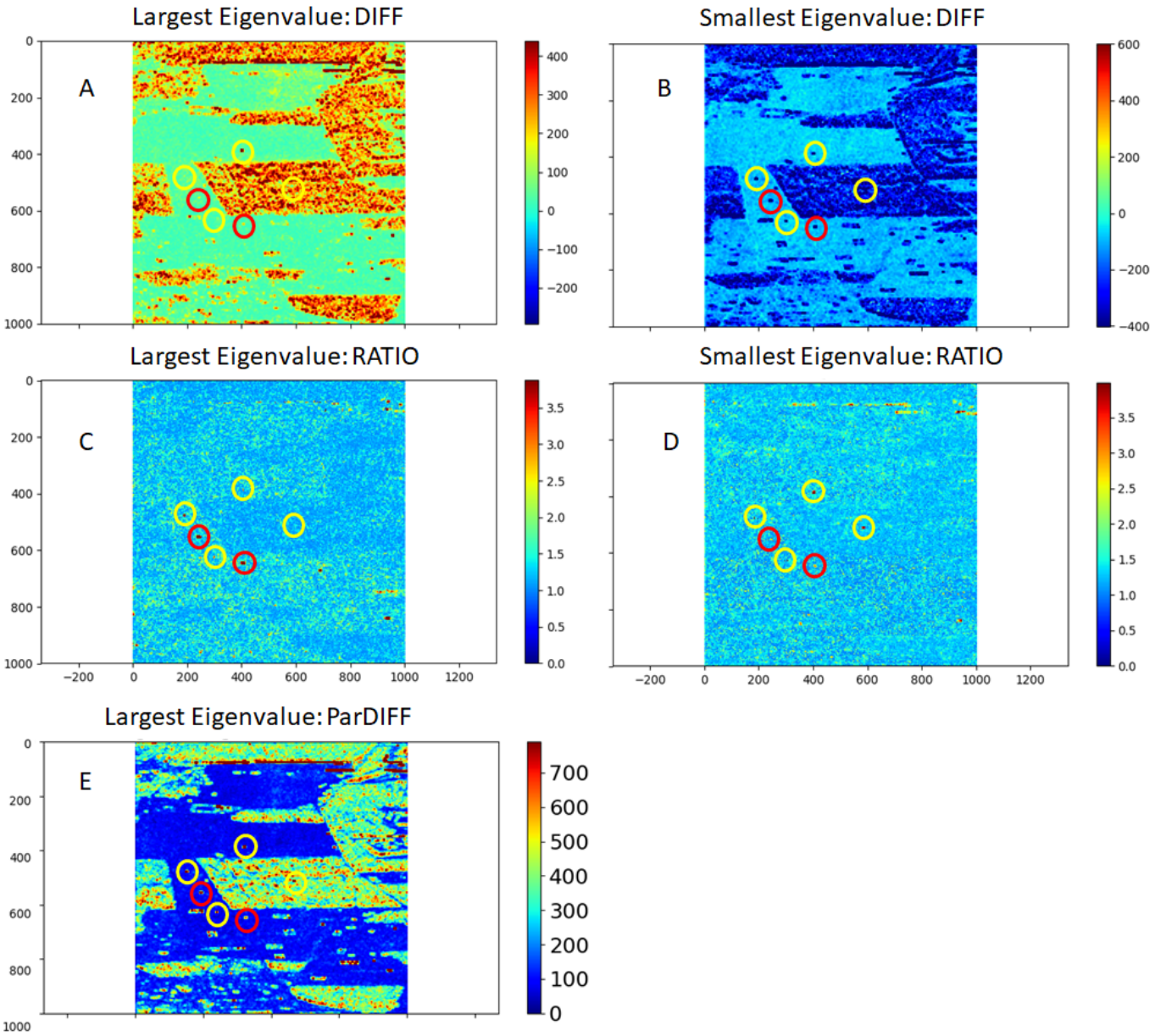


Fig. 15. Eigenvalues(SARTOM): (A) Largest Eigenvalue of DIFF (B) Smallest Eigenvalue of DIFF (C) Largest Eigenvalue of RATIO (D) Smallest Eigenvalue of RATIO (if smaller then 1, the inverse is shown), (E) Largest Eigenvalue of ParDIFF. Circles indicate targets that where removed/moved. Yellow is for corner reflectors, while Red were jeeps. Averaging: 9x9 boxcar.

nant scattering mechanisms that were added or removed, while RATIO was not providing this information. ParDIFF can also retrieve the two dominant SM changes, if applied twice to cover the two time directions ($1 \rightarrow 2$ and $2 \rightarrow 1$). On the other hand, RATIO and Conradsen detectors [18] provided better detection performance for polarised point targets.

Finally, the algorithms were tested on two different quad-polarimetric L-band datasets. The first was acquired during the E-SAR DLR SARTOM 2006 campaign and it is largely dedicated to point target detection. The second considers JAXA ALOS data acquired in 2006 over Morecombe Bay in the UK. This dataset was very beneficial to assess detection and classification of distributed targets. Interestingly, the eigenvectors results from DIFF and ParDIFF are very similar showing that both algorithms are able to identify the dominant changes in the scene.

- DIFF: When the additive hypothesis is true, DIFF can represent the dominant scattering mechanisms both added and removed. In case that there are multiple scattering mechanisms added or removed, the decomposition is focusing on the unique minimal change that can lead from the first to the second target. This is ensured by the property of orthogonality of the eigenvectors. Since it focuses on the minimal change, its information is complementary to applying a spectral decomposition of $[C_1]$ and $[C_2]$ separately and looking at differences in parameters.
- ParDIFF: When the additive hypothesis is true and only one partial target is either added or removed, ParDIFF is equivalent to a target decomposition because the change matrix will rep-

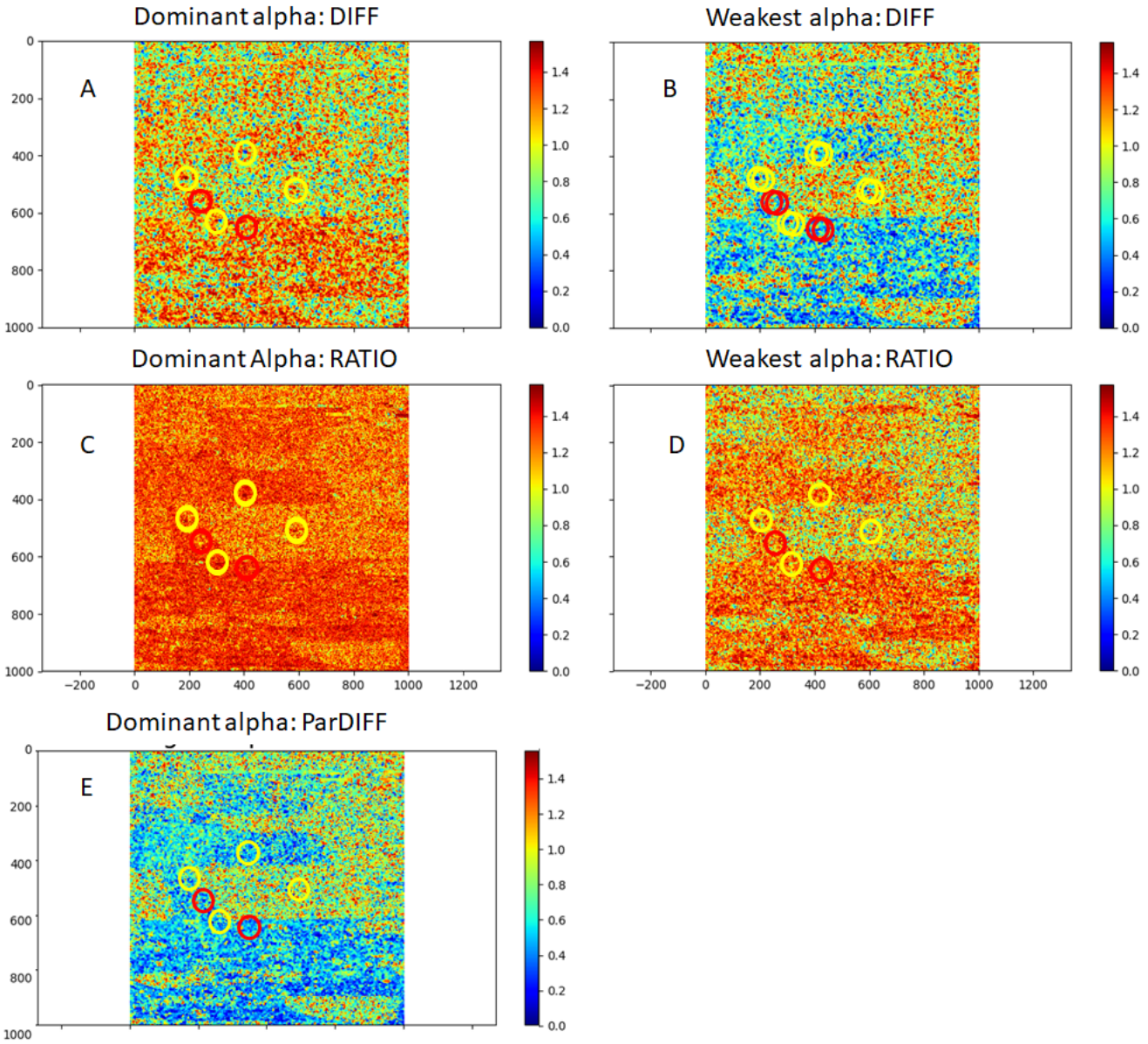


Fig. 16. α (SARTOM): (A) α for largest eigenvector of DIFF (B) α for lowers eigenvector of DIFF (C) α for largest eigenvector of RATIO (D) α for lowest eigenvector of RATIO (E) α for largest eigenvector of ParDIFF. Averaging: 9x9 boxcar.

resent a physically feasible partial target. However, if there are multiple partial targets added or removed, the algorithm will not be able to follow both changes and it needs to be applied twice in the two time directions. Interestingly, when the assumptions are not met, the algorithm is able to deal with the situation by converging toward the covariance matrices of destination (first or second depending on time direction). Although the result is equivalent to apply a diagonalisation of the original covariance matrix, ParDIFF is able to automatically recognise when this is needed. As a note of caution, applying the diagonalisation in both direction may result in one of the matrices not being positive semi-definite.

- RATIO provides eigenvectors that cannot be interpreted as scattering mechanisms, but directions in which we stretch the ellipsoid of the first partial target to make it into the second partial target. They are therefore very different from Target Decompositions and provide complementary information.

In terms of computational cost, DIFF presents the quickest algorithm with one diagonalisation of a symmetric matrix per pixel. RATIO performs one diagonalisation of a Normal, but asymmetric matrix and ParDIFF requires two diagonalisations if applied with the automatic selection of time direction (1 symmetric and 1 symmetric) or three diagonalisations if applied twice.

To summarise, the DIFF and ParDIFF decompositions are suited to classify the type of change (if the additive hypothesis is fulfilled), while RATIO is better in detecting polarised point targets.

As a future work, we want to explore more optimisations of different change matrices. We want also to investigate how to combine physical models to interpret the changes (e.g. for agri-

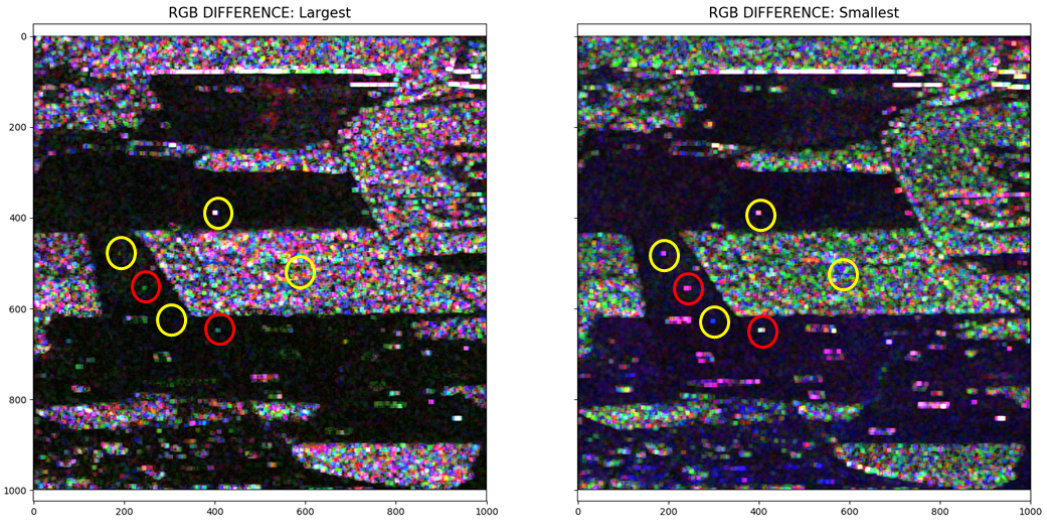


Fig. 17. DIFF optimisation(SARTOM): (Left) Largest eigenvalue; (Right) Lowers eigenvalue. Averaging: 9x9 boxcar.

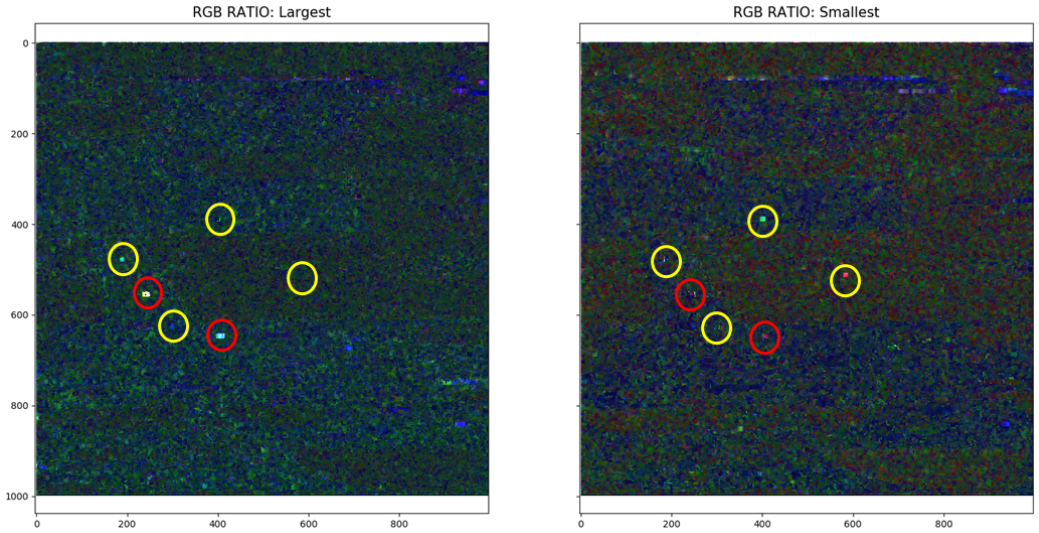


Fig. 18. RATIO optimisation (SARTOM): (Left) Largest eigenvalue; (Right) Lowers eigenvalue. Averaging: 9x9 boxcar.

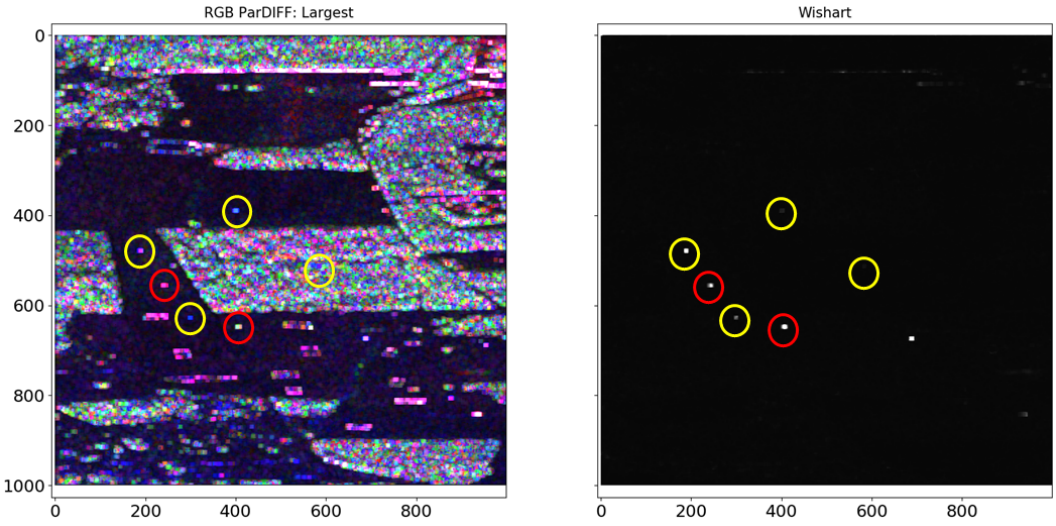


Fig. 19. Changed detection (SARTOM): (Left) ParDIFF optimisation largest eigenvalue; (Right) Wishart. Averaging: 9x9 boxcar.

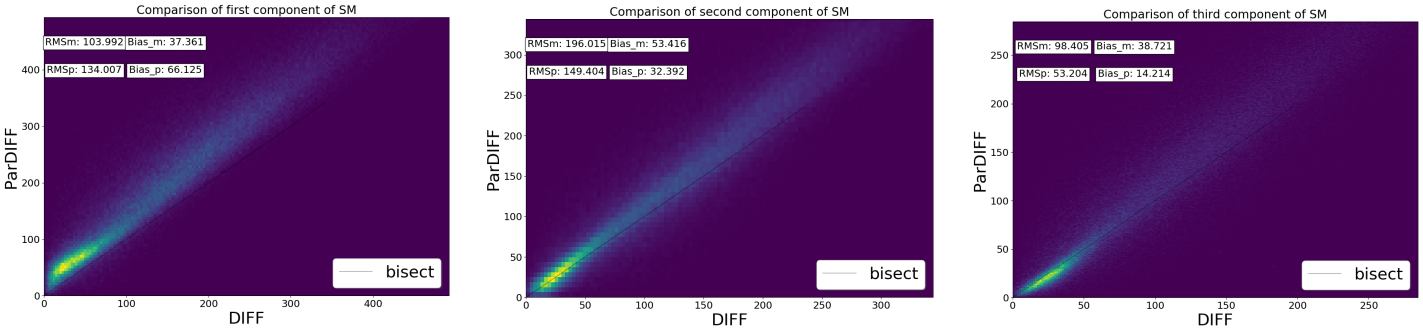


Fig. 20. Comparison(SARTOM): (Left) Largest eigenvalue; (Right) Lowers eigenvalue. Averaging: 9x9 boxcar.

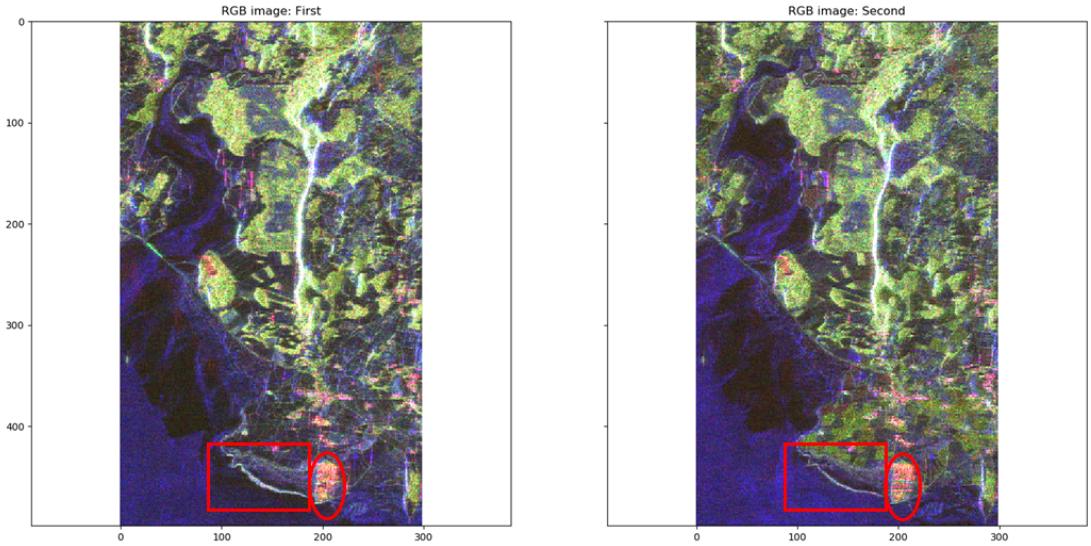


Fig. 21. RGB Pauli images of the JAXA test site with 1 and a half month of temporal baseline: (a) 1st of April 2007 (b) 17th of May 2007. (JAXA L-band).

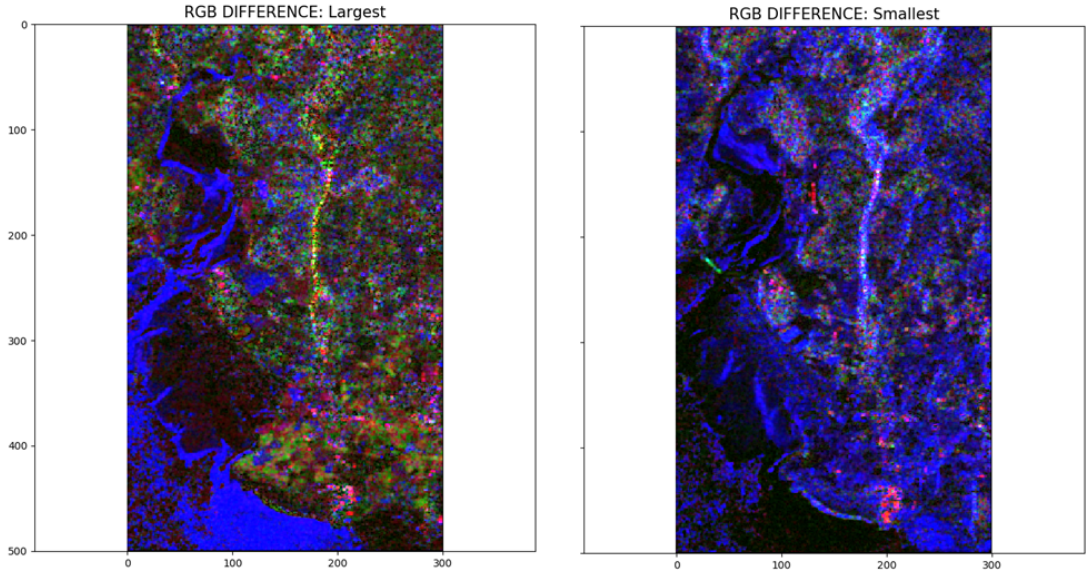


Fig. 22. DIFF optimisation (JAXA): (Left) Largest eigenvalue; (Right) Lowers eigenvalue. Averaging: 3x9 boxcar.

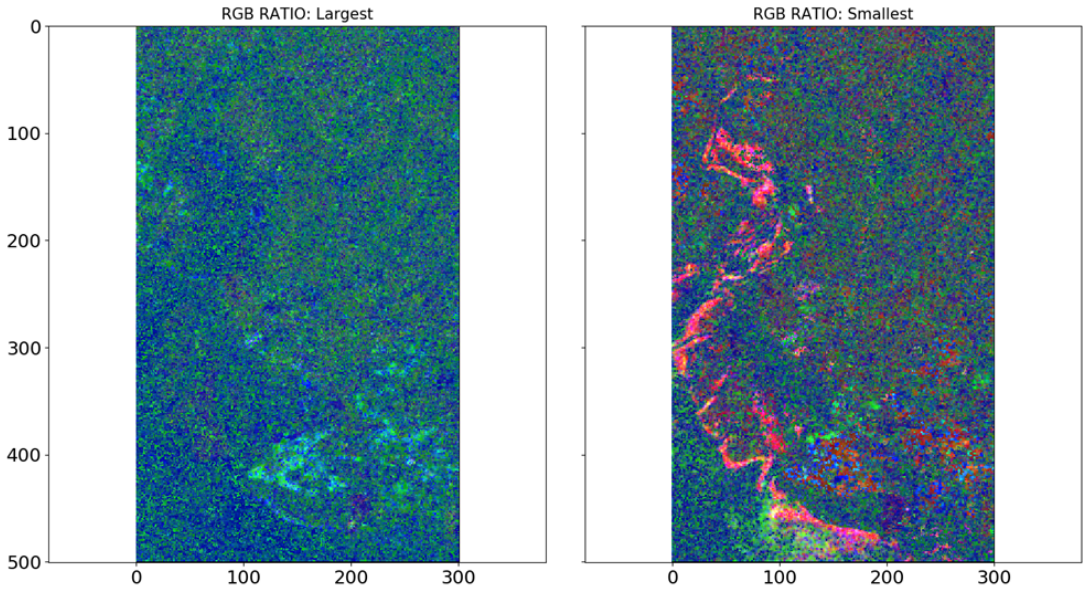


Fig. 23. Ratio optimisation (JAXA): (Left) Largest eigenvalue; (Right) Lowers eigenvalue. Averaging: 3x9 boxcar.

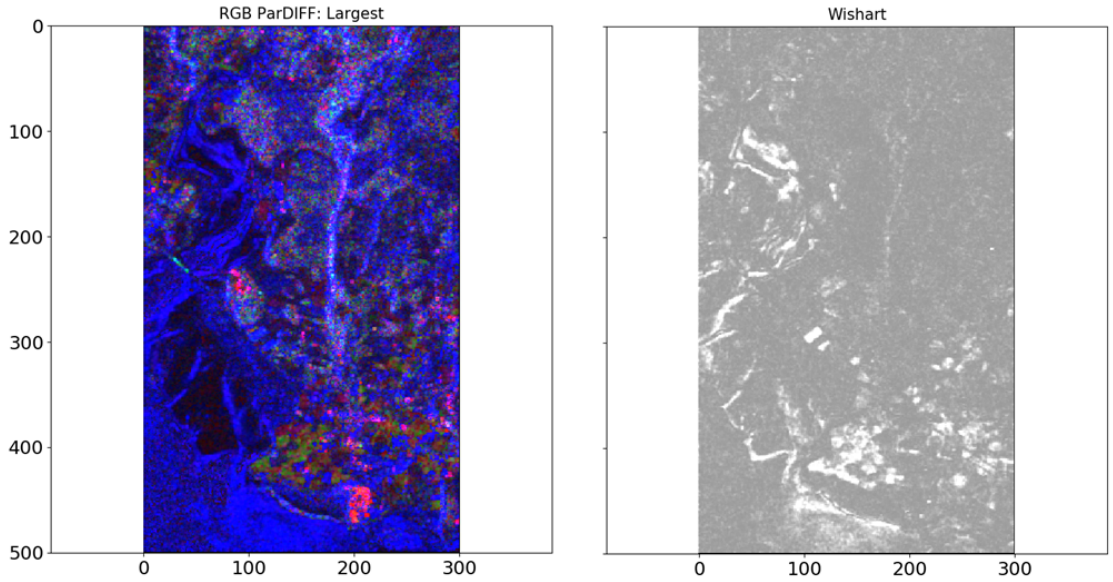


Fig. 24. ParDIFF optimisation (JAXA): (Left) Averaging: 3x9 boxcar.

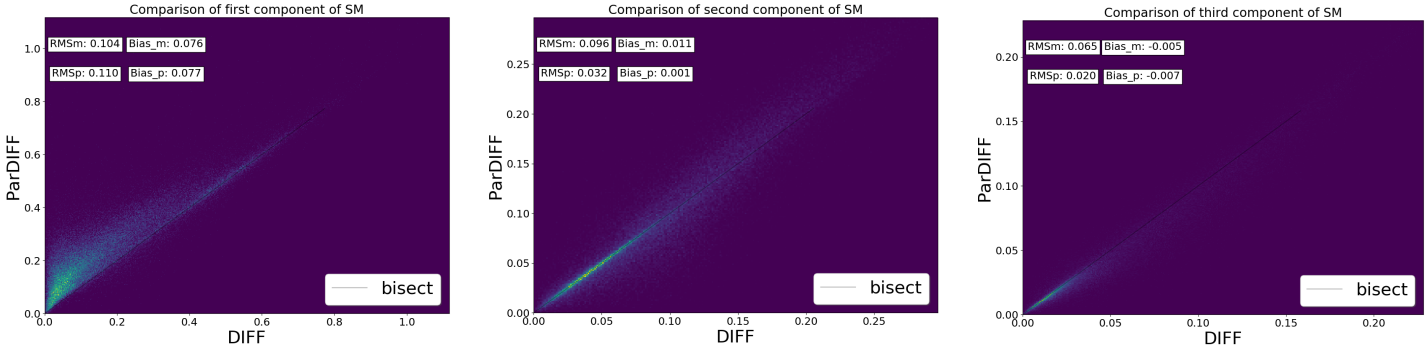


Fig. 25. Comparison (ALOS): (Left) Largest eigenvalue; (Right) Lowers eigenvalue. Averaging: 9x9 boxcar.

cultural purposes) and we want to extend these decompositions to time series of PolSAR data.

ACKNOWLEDGMENTS

The SARTOM2006 were acquired by the E-SAR airborne system of DLR. The authors would like to thank the DLR team that participated to the SARTOM campaign in 2006, in particular Ralf Horn and Rolf Scheiber.

The ALOS data were provided by JAXA under the project number 1151. ALOS Product-JAXA 2006, all rights reserved.

REFERENCES

[1] M. Preiss and N. J. S. Stacy, *Coherent Change Detection: Theoretical Description and Experimental Results*, Defence Science and Technology Organisation, 2006.

[2] E. J. M. Rignot and J. J. Van Zyl, "Change detection techniques for ERS-1 SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31(3), pp. 896–906, 1993.

[3] R. Dekker, "SAR change detection techniques and applications," *25th EARSeL Symposium Global Developments in Environmental Earth Observation from Space Porto, Portugal*, 2005.

[4] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Transactions on Signal Processing*, vol. 14 Issue:3, pp. 294 – 307, 2005.

[5] J. F. Mas, "Monitoring land-cover changes: A comparison of change detection techniques," *International Journal of Remote Sensing*, vol. 20(1), pp. 139–152, 1999.

[6] L. M. Novak, "Coherent change detection for multi-polarization SAR," *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, 2005.*, pp. 568 – 573, 2004.

[7] A. Marino, S. R. Cloude, and J. M. Lopez-Sanchez, "A new polarimetric change detector in radar imagery," *IEEE Transactions on Gescience and Remote Sensing*, vol. 51, no. 5, pp. 2986 – 3000, 2013.

[8] A. A. Nielsen, K. Conradsen, and H. Skriver, "Change detection in full and dual polarization, single- and multifrequency SAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 4041–4048, 2015.

[9] A. A. Nielsen, H. Skriver, and K. Conradsen, "The Loewner Order and direction of detected change in Sentinel-1 and Radarsat-2 data," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 242–246, 2020.

[10] A. A. Nielsen, "Fast matrix based computation of eigenvalues and the loewner order in polsar data," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, pp. 1727–1731, 2020.

[11] Y. Yamaguchi, T. Moriyama, M. Ishido, and H. Yamada, "Four-component scattering model for polarimetric sar image decomposition," *IEEE Trans. Geosci. Remote Sens*, vol. 43, pp. 1699–1706, 2005.

[12] A. Freeman and S. L. Durden, "A threecomponent scattering model for polarimetricsar data," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, pp. 963–973, 1998.

[13] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, pp. 498–518, 1996.

[14] S. R. Cloude, *Polarisation: Applications in Remote Sensing*, Oxford University Press, Oxford, UK, 2009.

[15] J. S. Lee and E. Pottier, *Polarimetric radar imaging: from basics to applications*, CRC Press, Taylor & Francis Group, 2009.

[16] A. Marino and A. Alonso-González, "Optimisations for different change models with polarimetric SAR," *EUSAR Conference 2018*.

[17] A.A. Swartz, H.A. Yueh, J.A. Kong, L.M. Novak, and R.T. Shin, "Optimal polarizations for archiving maximum contrast in radar images.," *Journal of Geophysical Research*, pp. 252 – 260, 1988.

[18] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver, "A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data," *IEEE Trans. on Geos. & Rem. Sen.*, vol. 41, 2003.

[19] S Cloude, "A physical approach to POLSAR time series change analysis," *IEEE Geoscience and Remote Sensing Letters*, 14 September 2020.

[20] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 68–78, 1997.

[21] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall, Upper Saddle River, US, 1993.

[22] A. Marino and I. Hajnsek, "A change detector based on an optimization with polarimetric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4781–4798, Aug. 2014.

[23] A. Alonso-Gonzalez, C. Lopez-Marinez, and P. Salembier, "PolSAR time series processing with binary partition trees," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3553–3567, 2014.

[24] A. Alonso-González, K. Papathanassiou, and I. Hajnsek, "Physical interpretation of the polarimetric SAR changes observed over agricultural time series.," *BioGeo SAR, Harwell, UK*, 2015.

[25] L.M. Novak, M.C. Burl, and Irving W.W., "Optimal polarimetric processing for enhanced target detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, pp. 234–244, 1993.

[26] R. O. Duda, P. E. Hart, and D. H. Stork, *Pattern Classification*, Wiley Interscience, 2000.

[27] G. Strang, *Linear Algebra and its Applications*, Thomson Learning, 1988.

[28] A. Marino, "Comparison of signal models for change detection with polarimetric sar," *ESA POLinSAR Workshop 2018*, 2018.

[29] A. Marino and A. Alonso-Gonzalez, "An optimization of the difference of covariance matrices for POLSAR change detection," *IGARSS symposium 2017*, 2017.

[30] C. Silva, A. Marino, JM Lopez-Sanchez, and I. Cameron, "Monitoring agricultural fields using an optimisation of the difference of covariance matrices for POLSAR," in *IEEE International Geoscience and Remote Sensing Symposium*, 2018.

[31] R. Horn, M. Nannini, and M. Keller, "SARTOM airborne campaign 2006: Data acquisition report," *DLR-HR-SARTOM-TR-001*, 2006.