A full non-parametric approach for SAR Coherent Change Detection

Giovanni Costa[a], Andrea Virgilio Monti-Guarnieri[a], Alessio Rucci[b], and Marco Manzoni[a]
[a]Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy
[b]TRE ALTAMIRA s.r.l., Ripa di Porta Ticinese 79, 20143 Milan, Italy

Abstract

Synthetic Aperture Radar (SAR) is widely used in heterogeneous fields with aims strictly dependent on the objectives of the application. One of the most common is the exploitation of the Interferometric-SAR (InSAR) to measure millimeter movements on the Earth's surface, aiming to monitor failures or measure infrastructures' health state. In this context, developing algorithms to detect temporal and spatial changes in the radar targets becomes very important. This paper focuses on the temporal change detection framework, proposing a non-parametric Coherent Change Detection (CCD) algorithm called Permutational Change Detection (PCD). The PCD estimates the temporal Change Points (CPs) of a radar target recognizing *blocks structure* in the coherence matrix without making any assumptions. The performance analysis on simulated data is accomplished, considering a realistic scenario where the geometrical and temporal decorrelation are properly modeled. Finally, the PCD is compared with a parametric CCD algorithm based on the Generalized Likelihood Ratio Test (GLRT).

1 Introduction

Over the last two decades, Synthetic Aperture Radar (SAR) has been used successfully for the study of surface deformation phenomena and for monitoring various assets and infrastructures by applying interferometric techniques (InSAR) to long temporal series of radar images acquired over the same area of interest. Radar targets can be point-wise, generally corresponding to man-made objects called PS [1] widely available in urban areas but also distributed DS [2], present mainly in non-urban areas. The main limitations of the algorithms used today are due to their difficulty in dealing with changes in the scenario under analysis. Coherent points, whether PS or DS, can change in time, sometimes disappearing and new measurement points can become available. The method proposed moves in the CCD framework, aiming to detect severe temporal changes by looking if the coherence matrix of the radar target presents a block structure, without any assumptions on the number of changes as well as the extension of the blocks. The algorithm is detailed in Section 2 after a theoretical introduction to the coherence matrix and the permutation tests. Performance analysis on simulated data is carried out in Section 3, considering a scenario where the temporal and geometrical decorrelation are properly modeled. Finally, PCD is compared with the parametric CCD algorithm proposed in [3].

2 Permutational Change Detection

The *Permutational Change Detection* (PCD) is a non-parametric CCD algorithm that estimates changes in coherence of a radar target and block structure in the modulus of the coherence matrix $\hat{\Gamma}$.

It consists of two steps:

- *Change Point Detection* (CPD): detection of the *candidate change point* (C-CP);
- *Change Point Validation* (CPV): validation of the candidate;

and gives as outputs:

- *Change Detection Matrix* (CDM): a matrix of the same dimension of $\hat{\Gamma}$ showing the estimated model for the target under analysis;
- *Change Vectors* (CV): vectors of length equal to the dimension of the dataset, storing the temporal location of the changes;

To better understand the mathematical details of the PCD steps (Section 2.2 and Section 2.3), an overview of the most important theoretical concepts is carried out in 2.1.

2.1 Coherence Matrix and Permutation Tests

2.1.1 Coherence Matrix

Given a SAR data stack of NI images, for each pixel, $\hat{\Gamma}$ is the $[NI, NI]$ matrix containing the modulus of the interferometric coherence [4] $\hat{\gamma}_{ij}$ between the i_{th} and j_{th} image, estimated over a window of dimension L, number of looks. In formulas

$$\hat{\Gamma} = \frac{1}{L} \left| \sum_{l=1}^{L} \frac{y(l) y^H(l)}{\sqrt{|y(l)|^2 \, |y^H(l)|^2}} \right|, \qquad (1)$$

where $\mathbf{y} \in \mathbb{C}^{NI}$ is the vector collecting the temporal samples of the l-th pixel. By following [4] [5] [6], (1) can be

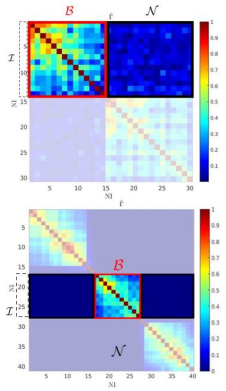

Figure 1 Example of blocks structures. The red and black squares indicate the block $\mathcal{B}$ and $\mathcal{N}$, respectively. The black dashed line defines the group of images $\mathcal{I}$.

approximated as

$$\hat{\Gamma} \simeq \gamma_0 \gamma_t \gamma_b \gamma_n, \tag{2}$$

where

- γ_0: maximum theoretical coherence;
- $\gamma_t = e^{-\frac{|B_{t_{ij}}|}{\tau}}$: temporal decorrelation with time constant τ [5];
- $\gamma_b = \left(1 - \frac{B_{n_{ij}}}{B_c}\right)$: geometrical decorrelation with B_c *critical baseline* [7];
- γ_n: other noise sources.

Starting from the theoretical analysis in [8] [9], it is possible to statistically model a pure noise realization. Because the interest of the PCD are small pieces of $\hat{\Gamma}$ that can be modeled as noise (Section 2.2 and Section 2.3), for simplicity it is assumed to be distributed like a Rayleigh [10] with scale parameter $\sqrt{\frac{1}{2L}}$.
The focus of the paper is the investigation of changes implying significant alteration of the electromagnetic properties, such that a target can evolve into new objects. Whenever a new radar object is born, $\hat{\Gamma}$ presents *blocks structure* (Figure 1), implying that the change is not reversible. Mathematically, a block $\mathcal{B}$ is a submatrix of $\hat{\Gamma}$ determined by a group of consecutive coherence estimates much greater than the ones outside the block. Being

$$\mathcal{D} = \{1, 2, ..., NI\},$$

the set of the observations' indexes, and

$$\mathcal{I} = \left\{ k \in \mathcal{D} \,|\, \frac{1}{L} \left| \sum_{l=1}^{L} y_k(l) {y_h}^H(l) \right| > \epsilon \, \forall h \in \mathcal{D} \text{ and } h \neq k \right\}$$

the groups of observations' indexes satisfying the condition above, then a block $\mathcal{B}$ is

$$\mathcal{B} = \{\hat{\gamma}_{ij} \in \hat{\Gamma} \,|\, i, j \in \mathcal{I}\}, \tag{3}$$

and the relative noise block

$$\mathcal{N} = \{\hat{\gamma}_{ij} \in \hat{\Gamma} \,|\, i \in \mathcal{I},\, j \in \mathcal{D} \setminus \mathcal{I}\} = \mathcal{B}^c. \tag{4}$$

Under these conditions, it is possible to write the $ij-th$ element of $\hat{\Gamma}$ as

$$\hat{\gamma}_{ij} = \frac{1}{L} \left| \sum_{l=1}^{L} \frac{y_i(l) y_j^H(l)}{\sqrt{|y_i(l)|^2 \left|y_j^H(l)\right|^2}} \right| \simeq \gamma_{\mathcal{B}_{ij}} \gamma_{t_{ij}} \gamma_{b_{ij}} \gamma_{n_{ij}}, \tag{5}$$

where $\gamma_{t_{ij}}, \gamma_{b_{ij}}$ and $\gamma_{n_{ij}}$ are respectively the values of the temporal and geometrical decorrelation and of the noise for the given interferometric pair. Instead, $\gamma_{\mathcal{B}_{ij}}$ is the $ij-th$ element of the block model of $\hat{\Gamma}$, which can be theoretically modelled as

$$\gamma_{\mathcal{B}_{ij}} = \begin{cases} 1 & \text{if } i, j \in \mathcal{I}, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

2.1.2 Permutation Tests

Fisher introduced the permutation test in 1936 [11]. The idea is very simple: given two statistical populations A and B with any cardinality, it is possible to perform a test of significance to verify a certain hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$, by properly designing a test statistic $T(X)$ and estimating the $p-value(\hat{p})$. The estimation is carried out repeating $\mathcal{N}_p$ times the same measurement $T(X)$ over new populations with the same cardinalities of the original ones, $\tilde{A}$ and $\tilde{B}$, randomly drawn from the pooled ensemble. Then $\hat{p}$ can be expressed as the number of times the measurements T^* over the permuted populations are greater than the value of $T(X)$ over the original population (T_0) [12]. In formulas

$$\hat{p} = \frac{\#(T^* > T_0)}{\mathcal{N}_p}. \tag{7}$$

The *caveat* on the accepted transformations of the data is that they must be likelihood-invariant; this means that if a paired dependence between the units of the populations exists, it must be preserved by the transformation so that the test is exact. Finally, the significance test demonstrates H_0 with a significance level α if $\hat{p} < \alpha$, where α can be fixed equal to a certain value or *adaptively* updated like it will be shown later on (Section 2.2). It is worth noticing that $\hat{p}$ asymptotically tends to the true one, meaning that the higher $\mathcal{N}_p$, the lower the *bias* affecting the estimation. Of course, not all the possible permutations are needed to reach a good estimation, but a number dependent on the type of application as well as the type of test statistic that represents a crucial aspect in designing this kind of operation: a general approach is to define a test statistic being *stochastically* greater under H_1 than under H_0 [12].

2.2 Change Point Detection

Before starting the detection on the i_{th} line of $\hat{\Gamma}$, a preliminary check is performed considering a noise threshold th_n. The threshold is obtained by modeling the probability distribution of N_r independent realizations of the maximum of the noise coherence matrix, as Weibull distributions from the *Generalized Extreme Value distribution* theory [13]. Taking $\gamma_n \in [0, 1]^W$, the vector containing W equispaced values between 0 and 1, it is possible to write

the probability density function of a single realization r as

$$\begin{cases} f_r(\gamma_n; L) = k\sqrt{L}\left(\sqrt{L}\gamma_n\right)^{k-1} e^{-\left(\sqrt{L}\gamma_n\right)^k}, \\ k = \left|2 - e^{-L+5}\right|, \end{cases}$$

for large L ($L > 5$). Considering now N_r independent realizations, the joint probability density function will be the product of the marginal probability density functions.
By setting a tolerance p_e, then the threshold th_n is the value satisfying

$$p_e = \int_0^{th_n} f(\gamma_n; L)\, d\gamma_n. \tag{8}$$

Defining $\hat{\gamma}_i \in [0,1]^{NI}$ the i_{th} line of $\hat{\Gamma}$, if $\max(\hat{\gamma}_i) > th_n$, it performs the following:

- *Screening and Candidates Selection*: calculation of a screening function s for the selection of the candidate change points (C-CPs). Supposing that $\hat{\gamma}_i$ is not affected by noise, an abrupt change can be directly identified by separating the vector progressively into two populations and by comparing the maxima of the two. In formulas, $s \in [0,1]^{(NI-i)-1}$ is defined as

$$s(j-i) = \max_{k=i+1,\dots,j} \hat{\gamma}_i(k) \; - \max_{h=j+1,\dots,NI} \hat{\gamma}_i(h)\,, \quad \forall j = i+1, i+2, ..., NI-1. \tag{9}$$

  In the ideal case, the CP is uniquely identified by the sample where s saturates. The only exploitation of s is not enough in a more realistic case when the noise strongly affects $\hat{\Gamma}$, but it is possible to recover some useful properties of the function in correspondence of the candidate changes; in fact, they can be recognized as the local maxima and flex points. Considering $\hat{\gamma}_i$ and defining $\mathcal{Z} \subseteq \mathcal{D}$ as $\mathcal{Z} = \{i+1, ..., NI-1\}$

$$\text{C-CPs} = \{z \in \mathcal{Z} \,\big|\, \Delta s(z) = 0\,,\, \Delta^2 s(z) = 0\}. \tag{10}$$

  If #C-CPs $\neq 0$, the CPD goes on; otherwise, it repeats the same operations on the next line;

- *Candidate Election*: election of the C-CP among the C-CPs, by unpaired permutation tests. The choice of using this kind of test instead of the classical statistical hypothesis tests is due to its non-parametric framework, its flexibility in designing an *ad hoc* test statistic in dependence on the application, and, in particular, to the evaluation of the $p - value$ which substitutes the classical integral calculus with a Monte Carlo simulation (7). Given the C-CPs at $\hat{\gamma}_i$, if a real abrupt change happens at index $z^* \in$ C-CPs then

$$\hat{\gamma}_{i+z^*}^{(nr)} = \hat{\gamma}^{(nr)} = \begin{bmatrix} \hat{\gamma}_{i+z^*}(i) \\ \hat{\gamma}_{i+z^*}(i+1) \\ \vdots \\ \hat{\gamma}_{i+z^*}(i+z^*-1) \end{bmatrix} \sim Rayleigh\left(\sqrt{\frac{1}{2L}}\right), \tag{11}$$

  meaning that the observed samples before z^* at the $(i+z^*)_{th}$ line, namely the *noise restricted*, is identically distributed like noise ($\mathcal{H}_0$). Under this reasoning, one good example of $T(X)$ is the two-sided *Kolmogorv-Smirnov statistic* (12)

$$T(X) = \sup_X \left|\mathcal{F}_{\hat{\gamma}^{(nr)}}(X) - \mathcal{F}_{th}(X)\right|. \tag{12}$$

  The criterion followed to set $\mathcal{N}_p$ is based on the sensitivity with which CPD must work. A wise choice is to privilege the cases when $\#\hat{\gamma}^{(nr)}$ is low, because they are much more difficult to detect and the exact $p - value$ makes the step much more robust. In the opposite cases, lower $\mathcal{N}_p$ can be accepted because of the effectiveness of the designed test statistic and because of the other PCD's steps. It can also be shown that the estimation of *p* by permutation tests in terms of MSE is much more affected by L than $\mathcal{N}_p$. By this reasoning, the decision to rule $\mathcal{N}_p$ as follows

$$\mathcal{N}_p = \begin{cases} 20 + \left\lceil \frac{NI}{2} \right\rceil & \text{if } \begin{pmatrix} 2\#\hat{\gamma}^{(nr)} \\ \#\hat{\gamma}^{(nr)} \end{pmatrix} > 10^3, \\ \begin{pmatrix} 2\#\hat{\gamma}^{(nr)} \\ \#\hat{\gamma}^{(nr)} \end{pmatrix} & \text{otherwise.} \end{cases}$$

  After the estimation of all the $\hat{p} - values$, are kept only the points such that

$$\text{C-CPs}_e = \left\{ z \in \text{C-CPs} \,\big|\, \hat{p}(z) < \alpha = (\#\text{C-CPs} + 1)^{-2} \right\}.$$

  The significance level α is adaptive because of the noise's impact on the calculation of the C-CPs (10): the lower L, the higher the impact of the noise, and the higher will be the number of elements in C-CPs so that, putting in α a dependence on the cardinality of C-CPs, will imply more stringent condition for the demonstration of $\mathcal{H}_0$. Finally, the C-CP is the point having the maximum $\hat{p}$ because it is supposed to be at the border of demonstrating $\mathcal{H}_1$;

- *Cross Validation*: cross validation of the C-CP and eventual re-election, by applying classical KS test moving along the diagonal. It has a double function because, in case of short-term changes, it will simply re-validate the CP; on the other, it will prevent the misclassification of CP by checking if it is due to decorrelation phenomena.

2.3 Change Point Validation

Given the i_{th} tested line and the related C-CP, the CPD isolates the relative $\mathcal{N}$ (4) and tests if it is distributed like noise. The strength of the validation step is due to the double information about the starting and ending point of the block. It performs the following:

- *Noise Block Validation*: after the extraction of $\mathcal{N}$, the validation is carried out through a two sided *Anderson-Darling test* where the probability distribution of $\mathcal{N}$ is tested against the theoretical one, to demonstrate $\mathcal{H}_0$. The choice of the test is due to its

statistical power [14]. Because of the robustness of CPD, it is possible to relax the condition on the significance level α, putting it equal to 0.05;

- *Cross Validation*: after the Noise Block Validation, the same CPD's cross-check is performed.

2.4 PCD Outputs

The PCD produces two different outputs:

- *Change Detection Matrix* (CDM): CDM is the $[NI, NI]$ matrix, showing the estimated block model for $\hat{\Gamma}$ (Figure 2). It is filled following these rules:
  - Every time the condition on th_n is not met as well as no C-CPs have been found during CPD, the correspondent line and column are equal to 0.5;
  - Every time a CP is found, the PCD fills the correspondent block $\mathcal{B}$ in the CDM with 2 or 1, respectively, if cross validation was or was not needed;
- *Change Vectors* (CV): the CV is simply a $[NI, 1]$ vector where the entries are one if a change happens at the correspondent image.

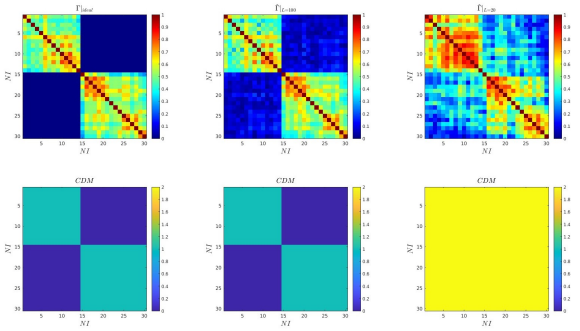

Figure 2 Example of CDM for the matrix Γ_{ideal} with its $\hat{\Gamma}_{L=100}$ and $\hat{\Gamma}_{L=20}$.

3 Performance Analysis

This section provides a detailed description of the performance of the PCD varying the number of looks L and NI, in an ideal setting where the extension of the blocks is $\left\lceil \frac{NI}{\#\mathcal{B}} \right\rceil$.

3.1 Metrics

The assessment is carried out by defining metrics that summarize true and false, positive and negative classes. In particular:

- *Accuracy* (ACC): It is the overall correctness

$$\text{ACC} = \frac{\text{TP+TN}}{\text{TN+FP+TP+FN}};$$

- *F1 score* (F1): It is a summary metric

$$\text{F1} = \frac{2 * \text{PRE} * \text{REC}}{\text{PRE+REC}},$$

where

  - *Precision* (PRE): It is the capability of predicting a specific category

$$\text{PRE} = \frac{\text{TP}}{\text{TP+FP}};$$

  - *Recall* (REC): It is the ability to detect a specific category

$$\text{REC} = \frac{\text{TP}}{\text{TP+FN}}.$$

3.2 Simulation settings and results

The performance has been measured out of 5000 Monte Carlo simulations assuming blocks of equal extension. The settings are summarized in Table 1. Figure 3 shows the results of the simulations as a function of L and NI, for the cases when $\#\mathcal{B}$ is 2, on the left, and 3, on the right. The performances are better increasing L as expected, passing from 0.57 to 0.74 in F1 in case of $\#\mathcal{B}$ equal two, and from 0.70 to 0.82 in the other, preserving overall a good accuracy. By comparing the two analyzed cases, it follows that the improvement of the performance is also correlated with the length of the dataset, meaning that the higher the number of blocks, the better the performance at higher NI. The deductions and comments drawn about the PCD are extremely useful and preparatory to understanding the comparison carried out with the GLRT-CCD algorithm presented in [3]. This approach is based on the Generalized

Table 1 Simulation Parameters in a realistic scenario.

t_r [days]	τ [days]	B_n [m]	B_c [m]	L	$\#\mathcal{B}$	NI
12	365	[-200, 200]	1300	5, ...,55	2,3	30, ...,60

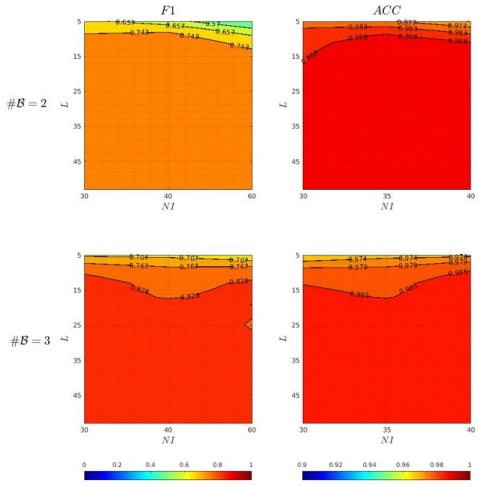

Figure 3 F1 and ACC for $\#\mathcal{B} = 2$ and 3. The performances improve by increasing L and NI.

Likelihood ratio Test (GLRT), which compares the posterior probabilities of the observations given the null hypothesis $\mathcal{H}_0$ if no changes occurred, with the alternative one, i.e., if a change occurred after NI_c acquisitions. $\mathcal{H}_1$. The detection of the changes is triggered if the LR function presents a single minimum or different minima below a threshold whose choice is fundamental. In order to make a comparison between the PCD and the GLRT-CCD, an adaptive evaluation of the threshold is implemented following the same reasoning behind 8. In fact, given the indexes of the minima of the LR function, it is possible to validate the detected change points by simply comparing the maximum value in the detected $\mathcal{N}$ with th_n. After the definition

Table 2 Simulation Parameters in an ideal scenario.

t_r [days]	τ [days]	B_n [m]	B_c [m]	L	$\#\mathcal{B}$	NI
12	$\frac{NI}{2}t_r$	[-1,1]	1300	20, ...,55	2,3	30, 60, 80

of the threshold, the comparison between the algorithms is performed by considering blocks of equal extension in two different settings:

- Ideal scenario: the impact of γ_b and γ_t is basically absent, simulating the cases where an optimum phase calibration has been performed. The simulation parameters are summarized in Table 2;
- Real scenario: the same simulation parameters summarized in Table 2, where $\hat{\Gamma}$ is affected by all the components in 2.

A first qualitative comparison between the two approaches is carried out through the analysis of the CDM, considering $NI = 60$, $\#\mathcal{B} = 3$ of equal extension and $\hat{\Gamma}$ estimated by means of $L = 55, 20$. As it is possible to notice in Figure 4, in the case of a realistic scenario, the impact of the estimation's noise of the coherence matrix much more affects the model estimated through GLRT-CCD. In fact, if in the case $L = 55$ the two CDM are the same, in the case $L = 20$, only the PCD estimate is preserved except for the extension of the first block. Instead, the GLRT-CCD estimate is only spotting the pieces of the blocks defined by coherence values $\hat{\gamma}_{ij}$ that are basically constant, in line with what is stated in [3]. In addition to that, the GLRT-

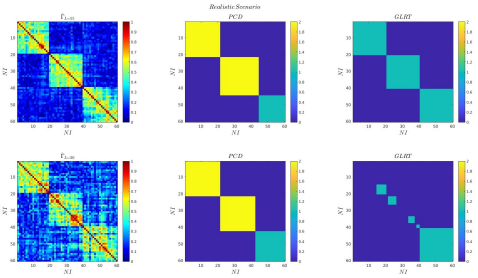

Figure 4 Example of CDM out of PCD and GLRT-CCD in a realistic scenario. PCD shows higher robustness in its estimate even lowering L.

CCD estimate is strongly biased by the value of th_n. In fact, by putting $th_n = \{0.3, 0.5, 0.7\}$ without proceeding through the proposed adaptive evaluation, the GLRT-CCD completely changes its estimate (Figure 5).
The GLRT-CCD starts detecting changes by increasing the value of th_n. Instead, the CDM produced by PCD remains unchanged except for the presence of the 0.5 crosses in the case $th_n = 0.7$ (Section 2.4). The reason behind this result

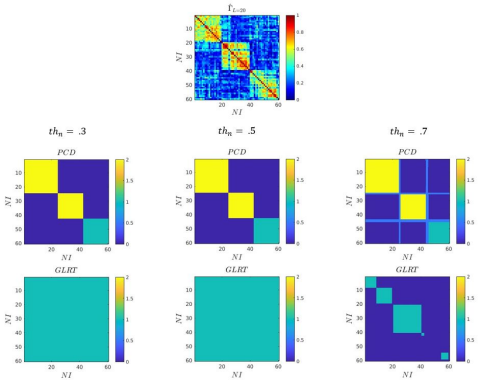

Figure 5 CDM out of PCD and GLRT-CCD, obtained by varying th_n. The GLRT-CCD estimate is strongly affected by the th_n. On the contrary, the non-parametric nature of PCD implies an unchanged estimated model.

is related to the different roles played by th_n: in the GLRT-CCD algorithm, the threshold is a fundamental parameter for the validation of the detected change points instead, the non-parametric nature of PCD, relegates th_n to being only a simple threshold useful for deciding whether or not to test a line of $\hat{\Gamma}$.
Considering instead the example shown in Figure 6, in the case of ideal scenario emerges that the GLRT-CCD is now more able to correctly detect the CPs regardless L, that on the contrary can generally affect more PCD as happens in the shown example. In the case of $L = 20$, the PCD merges the first and second blocks because the coherence values defined by the first 20 images and those at the end of the second block are not negligible. To give a quantitative

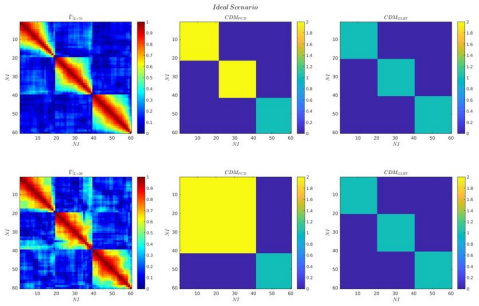

Figure 6 CDM out of PCD and GLRT-CCD in an ideal scenario. The GLRT-CCD is more able to correctly detect the CPs regardless L, that can generally affect more the PCD.

measure of the comparison, the F1 for both algorithms is measured through 5000 Monte Carlo simulations and com-

bined into a single one called F1 Ratio (13)

$$F1R = \frac{F1_{PCD}}{F1_{GLRT-CCD}}. \quad (13)$$

The results of the simulations are shown in Figure 7. The

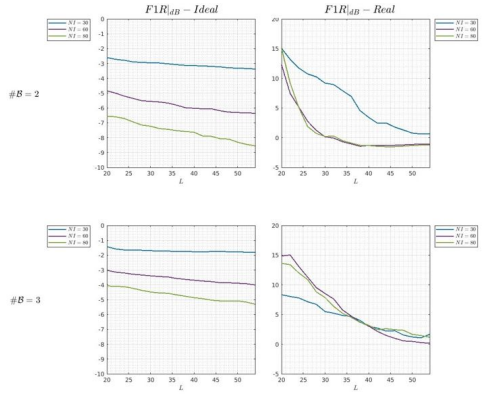

Figure 7 $F1R$ [dB] in the case of ideal and real scenario, respectively on the left and on the right, considering $\#\mathcal{B} = 2, 3$ of equal extension.

GLRT-CCD, as expected, performs better in the ideal scenario, with a gain increasing exponentially with L. Instead, in a more realistic scenario, PCD shows much more robustness, particularly in the case of lower L, where it reaches very high $F1R$ values. Moreover, the $F1R$ in the ideal case increases with $\#\mathcal{B}$ in both scenarios. Focusing the attention on the $\#\mathcal{B} = 2$ cases, $L = 30$ represents a turning point: in fact, the shape of the function diminishes the slope. This change translates in the realistic scenario for the cases $NI = 60, 80$, in a gain of 1dB of the GLRT-CCD. This is because of the effect of temporal decorrelation together with the block extension. Instead, by increasing the number of blocks, the change in the slope in the ideal scenario is much less pronounced, implying overall better performance of the PCD. In conclusion, increasing $\#\mathcal{B}$, the behavior of $F1R$ in the realistic settings seems to approach the one in the ideal in a dual sense: in fact, PCD exponentially gains in performance by decreasing L.

4 Conclusions

This paper presents a non-parametric algorithm for Coherent Change Detection, the Permutational Change Detection. The novelty introduced allows the users to apply PCD without having the need to make strong assumptions about the data and without the need to model parameters that can strongly affect the final results when the model's hypotheses are not perfectly met, as happened for the GLRT-CCD. The algorithm's strength is represented by its robustness in detecting changes in coherence in any scenario and the flexibility with which it can be adapted for any application by simply modifying the test statistic $T(X)$ to fit the users' needs. The main limitation can be represented by the computational time, dependent on NI and how dynamic a target is. It can be reduced by means of parallelization or by strongly reducing the bandwidth of considered images or implementing simplest $T(X)$.

5 Literature

[1] A. Ferretti, C. Prati, and F. Rocca, " Non-uniform motion monitoring using the permanent scatterers technique", FRINGE'99: Advancing ERS SAR Interferometry from Applications towards Operations, 2000.

[2] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci, "A new algorithm for processing interferometric data-stacks: SqueeSAR", IEEE Transactions on geoscience and remote sensing, vol. 49, no. 9, pp. 3460–3470, 2011.

[3] A. V. Monti-Guarnieri, M. A. Brovelli, M. Manzoni, M. M. d'Alessandro, M. E. Molinari, and D. Oxoli, "Coherent change detection for multipass SAR", IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 11, pp. 6811–6822, 2018.

[4] A. M. Guarnieri and S. Tebaldini, "On the exploitation of target statistics for SAR interferometry applications", IEEE Transactions on Geoscience and Remote Sensing, vol. 46, no. 11, pp. 3436–3443, 2008.

[5] A. M. Guarnieri and S. Tebaldini, "On the exploitation of target statistics for SAR interferometry applications", IEEE Transactions on Geoscience and Remote Sensing, vol. 46, no. 11, pp. 3436–3443, 2008.

[6] A. Ferretti, "Satellite InSAR Data: reservoir monitoring from space (EET 9)", Earthdoc, 2014.

[7] A. Ferretti, A. Monti-Guarnieri, C. Prati, F. Rocca, and D. Massonet, "InSAR principles-guidelines for SAR interferometry processing and interpretation", 2007, vol. 19.

[8] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry", Inverse problems, vol. 14, no. 4, p. R1, 1998.

[9] R. Touzi, A. Lopes, J. Bruniquel, and P. W. Vachon, "Coherence estimation for SAR imagery", IEEE Transactions on geoscience and remote sensing, vol. 37, no. 1, pp. 135–149, 1999.

[10] E. E. Kuruoglu and J. Zerubia, "Modeling SAR images with a generalization of the Rayleigh distribution", IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 527-533, 2004.

[11] R. A. Fisher, " "The coefficient of racial likeness" and the future of craniometry", The Journal of the Royal Anthropological Institute of Great Britain and Ireland, vol. 66, pp. 57–63, 1936.

[12] B. Brown, "Permutation tests for complex data: Theory, applications and software by F. Pesarin and L. Salmaso", 2012.

[13] S. Kotz and S. Nadarajah, "Extreme value distributions: theory and applications", World Scientific, 2000.

[14] N. M. Razali, Y. B. Wah et al., "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests", Journal of statistical modeling and analytics, vol. 2, no. 1, pp. 21–33, 2011.