

Wavelet Spatio-Temporal Change Detection on multi-temporal SAR images

Rodney V. Fonseca, Rogério G. Negri, Aluísio Pinheiro, and Abdourrahmane Atto, *Senior Member, IEEE*

Abstract—We introduce WECS (Wavelet Energies Correlation Screening), an unsupervised procedure to detect spatio-temporal change points on multi-temporal SAR images. The procedure is based on wavelet approximation for the multi-temporal images, wavelet energy apportionment, and ultra-high dimensional correlation screening for the wavelet coefficients. We show WECS performance on simulated multi-temporal image data. We also evaluate the proposed method on a time series of 84 satellite images in a forest region at the border of Brazil and the French Guiana. The proposed method displays good results in covering change regions, with the additional benefit of having simple and fast computation.

Index Terms—Change detection, Remote Sensing, multi-temporal images, simulated images, wavelets.

I. INTRODUCTION

CHANGE detection is an important task performed in remote sensing image that allows researchers and engineers to identify and evaluate modifications on land surfaces captured by multi-temporal satellite images. Analyzing problems such as deforestation [1], rapid urbanization [2] and glacier melting [3] are of great importance to study the dynamics of regions sensitive to climate changes and human activity. Furthermore, the increase on availability of satellite images in the past years raises the challenge of applying computationally cheap methods to images available at larger and larger from time to time. A review for change detection in multi-temporal remote sensing is given by [4].

Most methods used for change detection analysis can be classified either as supervised (training data is used to set up the method) or unsupervised (fully data-driven techniques). We shall focus in this work on unsupervised approaches, whose examples in the literature include the works of [?], [?], [5]–[8]. Many other proposals of methods vary in their motivations as well as in their applicability. Change detection in multi-temporal hyperspectral images is discussed in [9]–[11]. [12] pursue change detection techniques via non-local means and principal component analysis. Compressed projection and image fusion are employed by [13]. Deep learning by slow feature analysis for change detection is the subject of [14].

This work was supported by FAPESP grants 2016/24469-6 and 2018/04654-9 and CNPq grants 309230/2017-9 and 310991/2020-0.

R. Fonseca and A. Pinheiro are with the Department of Statistics, University of Campinas, Campinas 13083-859, Brazil (e-mails: ra192588@dac.unicamp.br; pinheiro@ime.unicamp.br).

A. Atto is with the LISTIC - Polytech Annecy-Chambéry, Université de Savoie, 74944 Annecy le Vieux Cedex, France (e-mail: Abdourrahmane.Atto@univ-savoie.fr).

R. G. Negri is with the Department of Environmental Engineering, São Paulo State University, São José dos Campos, Brazil (e-mail: rogerio.negri@unesp.br).

[15] proposes a change detection method driven by adaptive parameter estimation.

A special attention has been given for multi-temporal change detection using Synthetic Aperture Radar (SAR) images. Some of the references already mentioned are examples of change detection methods applied to SAR images [1]–[3], [7], [12], [13], [16]. Known to be not affected by weather, cloud and sunlight conditions, SAR images rise as an essential data source in change detection applications [16]. Conversely, due to its acquisition architecture, the speckle noise typically affects the images obtained by SAR sensors and demands additional pre-processing before its use.

Face with the complexity imposed by the high-dimensionality of multi-temporal datasets in addition to the presence of speckle noise, the change detection process using SAR images becomes a challenge task. In this context, the use of Wavelets rises as a convenient approach once such technique is robust when dealing with noise data and it allows an efficient computational treatment.

In more details, Wavelet methods present many advantages for a plethora of statistical applications [17] not only in remote sensing problems thanks to wavelet capabilities in capturing multi-scale/resolution information. Their computational efficiency and sparseness are specially relevant for large images and other high-dimensional data [18]. Analysis of SAR images have been investigated under different approaches using wavelet methods, such as [19], [20], [21], [22]. However, an application of screening ideas on wavelet coefficients is still a novel approach in change detection literature, and we show it has an interesting potential to provide good results even with simple algorithms.

In the light of presented discussions and motivations, using Wavelet and data screening concepts, we propose the Wavelet Energies Correlation Screening (WECS), a novel unsupervised multi-temporal change detection method for SAR images. The main idea behind WECS is built on ultra-high dimensional feature screening for the wavelet coefficients [23]. Such method is usually employed in high-dimensional regression models to reduce the problem's dimension by subsetting the available covariates in such a way that true covariates are among the chosen ones with high probability [24]. We show that by applying the feature screening idea on multi-temporal images, we obtain a fast and accurate method to cover change regions with good detection rates.

This paper is organized as follows: Section III introduces the problem and presents the proposed method. We show WECS performance on simulated multi-temporal image data in Section ?? . In Section ?? we apply the proposed method to a time

series of 85 satellite images in the border region of Brazil and the French Guiana, for images captured from November 08, 2015 to December 09 2017. Section V concludes the paper with a discussion.

II. THEORY BACKGROUND

Wavelet methods have been widely applied to analyze images in signal processing literature, specially for tasks such as signal denoising and compression [25]. The most common way of describing wavelet representations is as a multiresolution decomposition, where a signal is represented on approximation and detail coefficients, which provide coarse and finer details of the signal, respectively. In practice, the discrete wavelet transform of matrices (image) consists in applying low and high-pass convolution filters to its rows and columns [26]. In case of smoothing, applying such low-pass filter J times to rows and columns of a matrix \mathcal{I} , yields a smooth image \mathbf{X} . The number J is also called resolution level, a tuning parameter for wavelet smoothing. This wavelet smoothing is employed in the next section as a denoising step for image analysis.

Another idea that shall be employed in this work is the feature screening procedure, a method originally designed for ultra-high dimensional regression models [24]. In order to explain how it works, consider the usual linear regression framework where \mathbf{y} is a $n \times 1$ vector of observations from a response variable and $\{\mathbf{W}_1 \dots \mathbf{W}_p\}$ is a $n \times p$ matrix with explanatory variables on its columns, which are used in a linear model $\mathbf{y} = \sum_{i=1}^p \beta_i \mathbf{W}_i + \boldsymbol{\epsilon}$, where β_1, \dots, β_p are unknown parameters and $\boldsymbol{\epsilon}$ is a zero mean random noise. The problem setup is that p is much larger than n , what makes standard regression methods unfeasible, but only a handful of the available covariates are relevant for the model, i.e., have a nonzero corresponding parameter β_i . The feature screening idea consists in computing the sample correlation $\text{corr}(\mathbf{y}, \mathbf{W}_i)$ among response and explanatory variables, and then selecting those covariates whose correlation are among the highest values. Under suitable conditions, such method is known to select a set containing all true covariates with high probability. In the image change detection problem, our idea is that a similar approach could be used to detect change locations by taking an overall change measure as response variable and local (pixel) measures as potential covariates.

III. METHODOLOGY

Let $\mathcal{I}^{(1)}, \dots, \mathcal{I}^{(n)}$ be an image time series defined on a support $\mathcal{S} = \{1, \dots, u\} \times \{1, \dots, v\} \subset \mathbb{N}^2$, hence representing a region of interest over n distinct instants. These images may be relative to one SAR channel or a combination of channels; this will be specified when appropriate. Our goal is twofold: to find possible points in time where some relevant changes might have taken place at the region represented in $\mathcal{I}^{(m)}$, $m = 1, \dots, n$, and to find which regions are closely associated to the observed changes along time. We shall address these tasks by analyzing the bidimensional stationary discrete wavelet decomposition of $\mathcal{I}^{(m)}$. Stationary wavelets (also known as non-decimated or redundant wavelets) is a traditional de-noising

method that can be efficiently applied to two-dimensional signals such as images [19], [27], [28].

After application of this wavelet transform to $\mathcal{I}^{(m)}$ at some appropriate resolution level $J \geq 1$, one of its by-products is a matrix of so called approximation wavelet coefficients $\mathbf{X}^{(m)}$, a smoothed version of $\mathcal{I}^{(m)}$ defined on the same support \mathcal{S} . The higher $J \in \{1, \dots, \log_2(\min u, v)\}$ is, the smoother $\mathbf{X}^{(m)}$ gets.

Beyond many other aspects that can be involved in wavelet analysis of images, which may include different types of wavelet transform and basis as well the use of thresholding for detail coefficients, the current construction is focused on $\mathbf{X}^{(m)}$ to provide a simple wavelet smoothing. Nevertheless, extensions based on distinct aspects are straightforward.

We can then consider further apportioning the total \mathbb{L}_2 energy of the series $\{\mathbf{X}^{(m)}\}_{m=1, \dots, n}$ as:

$$\begin{aligned} \sum_{m=1}^n \|\mathbf{X}^{(m)}\|_F^2 &= n\|\bar{\mathcal{I}}\|_F^2 + 2n\langle \bar{\mathbf{X}} - \bar{\mathcal{I}}, \bar{\mathcal{I}} \rangle_F + \\ &\quad + \sum_{m=1}^n \|\mathbf{X}^{(m)} - \bar{\mathcal{I}}\|_F^2, \end{aligned} \quad (1)$$

where $\bar{\mathcal{I}} = n^{-1} \sum_{m=1}^n \mathcal{I}^{(m)}$ and $\bar{\mathbf{X}} = n^{-1} \sum_{m=1}^n \mathbf{X}^{(m)}$; $\|\cdot\|_F$ and $\langle \cdot, \cdot \rangle_F$ represents the Forbenius norm and inner product, respectively. Naturally, both $\bar{\mathcal{I}}$ and $\bar{\mathbf{X}}$ shares the support \mathcal{S} .

The most right-hand term in Equation (1) measures the deviations between the wavelet coefficients at distinct instants ($\mathbf{X}^{(m)}$; $m = 1, \dots, n$) regarding the *average image* (i.e., $\bar{\mathcal{I}}$). Such deviations may favor detecting relevant changes over the time. Since each element (i.e., a pixel/position $(k, l) \in \mathcal{S}$) of $\mathbf{X}^{(m)}$ also has a corresponding sequence of deviations in time, the local deviation to the overall measure may also allow detecting the spatial changes. Among several approaches able to detect change events, the Pearson correlation coefficient rises as an convenient alternative. Furthermore, such measure shares connections with the idea of feature screening employed in high-dimensional regression, as explained earlier.

Let $X_{kl}^{(m)}$ and \bar{I}_{kl} be the entry (k, l) of $\mathbf{X}^{(m)}$ and $\bar{\mathcal{I}}$, respectively. The matrix $\mathbf{D}^{(m)}$ embraces the squared differences between $\mathbf{X}^{(m)}$ and $\bar{\mathcal{I}}$, where $D_{kl}^{(m)} = (\mathbf{X}_{kl}^{(m)} - \bar{I}_{kl})^2$. Therefore, it is defined the temporal overall variation sequence $\{\mathbf{d}^{(m)}\}_{m=1, \dots, n}$, whose elements are given by:

$$\mathbf{d}(m) = \sum_{k=1}^u \sum_{l=1}^v D_{kl}^{(m)} \quad (2)$$

In this context, a instant m in $\{\mathbf{d}^{(m)}\}_{m=1, \dots, n}$ with high value stands for the image $\mathcal{I}^{(m)}$ where the most expressive changes take place.

In order to identify spatio-temporal changes, we employ the concept of ultra-high dimensional correlation screening [23] discussed in Section II. For each local squared deviation time series given by individual elements of $\mathbf{D}^{(m)}$ across $m = 1, \dots, n$, say $\mathbf{D}_{kl} = \{D_{kl}^{(1)}, \dots, D_{kl}^{(n)}\}$, consider the absolute

value of its Pearson correlation with the overall squared mean deviations, given by \mathbf{d} :

$$R_{kl} = |\text{corr}(\mathbf{D}_{kl}, \mathbf{d})|.$$

Consequently, rises the matrix \mathbf{R} from the elements R_{kl} assigned to each $(k, l) \in \mathcal{S}$.

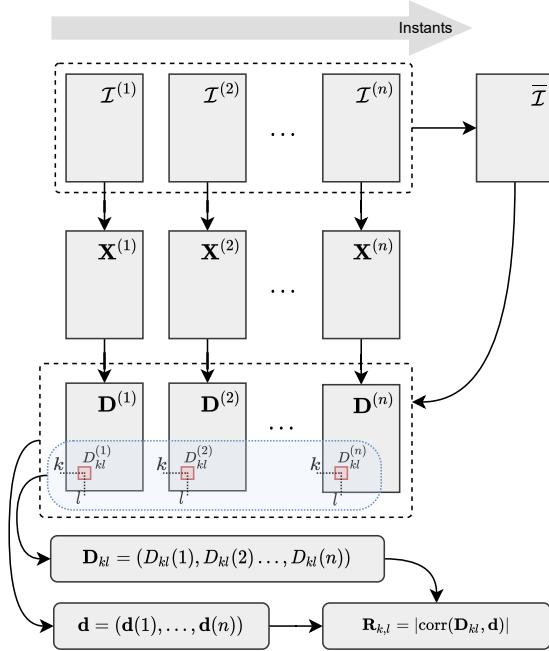


Fig. 1. Diagram of steps performed on WECS to compute an absolute correlation of a single point.

Let us define a mapping of *important* indices for changes over the image series with respect to $\bar{\mathcal{I}}$ as $\mathcal{M}^* \subseteq \mathcal{S}$ where *changes in $\{\mathcal{I}^{(m)}\}_{m=1,\dots,n}$ with respect to $\bar{\mathcal{I}}$ are affected by local changes in the images*.

An empirical mapping of change locations may be stated by:

$$\mathcal{M}_\tau = \{(k, l) \in \mathcal{S} : |R_{kl}| > \tau\}, \quad (3)$$

where $\tau \in \mathbb{R}_+^*$ is a convenient threshold value defined as function of n . Under some regularity conditions, the probability $P(\mathcal{M}_\tau \supseteq \mathcal{M}^*) \rightarrow 1$ as $n \rightarrow \infty$ [23]. In other words, the empirical set \mathcal{M}_τ has high probability of detecting the correct change locations in \mathcal{M}^* when the number of instants n is large enough.

IV. EXPERIMENTS

In this section we shall evaluate the performance of WECS on two types of data sets. Initially, we generate simulated data where the level of noise and true change regions are known. For these images, we analyze different aspects of concerning the application of WECS and compare its performance with a standard change detection method. In the following application, real satellite data is employed, where the performance of WECS is evaluated on predefined change regions.

A. Simulated data analysis

In this section we apply the change detection methods above on simulated data of multi-temporal images. The simulated multi-temporal images ($n = 80$) is obtained by repeating a sequence of four images with different types of changes plus a noise. Examples of the first four of these images are shown in Figure 2. These images are generated from the sum of two matrices: a signal matrix with ones on entries where ellipses occur and zero elsewhere, and a noise matrix with random variables following a standard Gaussian distribution. The first image, $\mathcal{I}^{(1)}$, presents three elongated ellipses. The second image $\mathcal{I}^{(2)}$ has shorter and larger ellipses added. Smaller ellipses are then added to form $\mathcal{I}^{(3)}$ and $\mathcal{I}^{(4)}$. All the changes made that occur among subsequent images can be seen in Figure 3(a), which displays a matrix of zeros and ones that correspond, respectively, to locations without and with changes along time. Applying WECS to these images, we obtain a matrix \mathbf{R} of correlations between deviations of each \mathcal{I} entry with the total squared mean deviation. An example of \mathbf{R} is presented in Figure 3(b). For some choice of threshold τ on absolute values of \mathbf{R} , we obtain a matrix of zeros and ones that can be compared with the total true changes displayed in Figure 3(a).

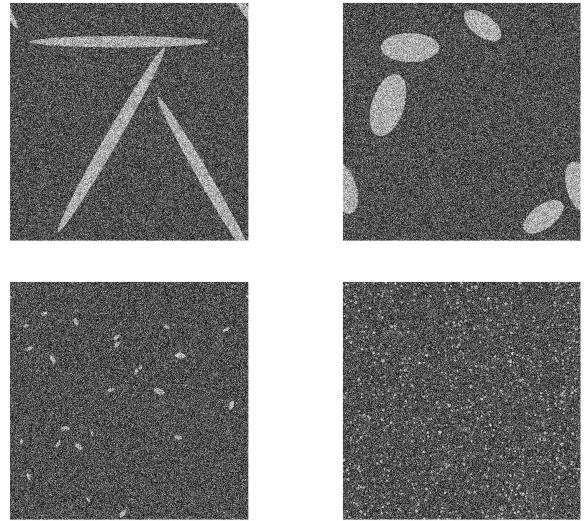


Fig. 2. Example of the first four simulated multi-temporal images. Features and changes come as white ellipses and dots.

Figure 3 illustrates a comparison of different approaches to detect accumulated changes. Panel (b) shows the result of WECS using Daubechies wavelet with two null moments (db2) and $J = 2$. Panel (c) presents the result of using aggregated absolute differences, a standard approach where the accumulation of changes are measured by a matrix $\mathbf{S} = \{\mathbf{S}_{k,l}\}$ with $S_{kl} = \sum_{m=2}^n |\mathcal{I}_{k,l}^{(m)} - \mathcal{I}_{k,l}^{(m-1)}|$. Finally, in Panel (d) we can see the result if $\mathbf{d}(m)$ is performed purely on the spatial domain, using $\mathcal{I}^{(m)}$ instead of $\mathbf{X}^{(m)}$ in the WECS formulation.

We compute ROC curves to compare the detection performance of different methods and to verify the influence of some tuning parameters of wavelet smoothing: the resolution level J and the choice of wavelet basis. Each detection method

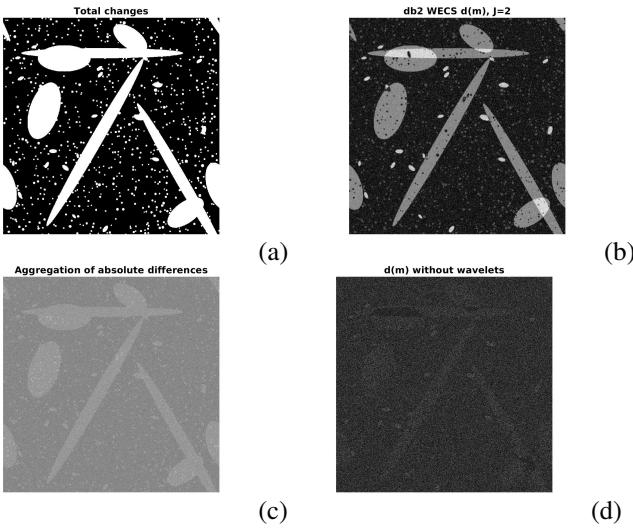


Fig. 3. Simulated images with changing ellipses. (a) Image composed by the total changes over time. (b) Proposed db2 WECS $\mathbf{d}(m)$ with $J = 2$; (c) Standard approach. (d) $\mathbf{d}(m)$ without wavelets.

generates a matrix of change detection measures (correlations in the case of WECS). The ROC curves present the performance of change detection by applying a threshold on these measures, in the following way:

- 1) Let R be the matrix of change measures. Compute the range $[r_{\min}, r_{\max}]$ of the values in R ;
- 2) Let $(r_{(1)}, \dots, r_{(100)})$ be equally space values between r_{\min} and r_{\max} ;
- 3) For each $k = 1, \dots, n$, check how many entries are there such that $R_{i,j} > r_{(k)}$ coincide with the element (i, j) where a change really occurs on the image of total changes. Dividing this number by the total number of changes gives the true positive rate.
- 4) For each $k = 1, \dots, n$, check how many entries are there such that $R_{i,j} > r_{(k)}$ do not coincide with the element (i, j) where a change really occurs. Dividing this number by the total number of entries where changes do not occur gives the false positive rate.
- 5) The ROC curve is the plot of true and false positive rates corresponding to each k .

Figure 4 presents the different ROC curves for change detection methods applied to the simulated data as follows. The effects of wavelet bases, level of decomposition, deep-learning feature extraction are shown on the ROC curves. We employ the following wavelet bases: Haar; Daubechies db2; Daubechies db4; Coiflets coif4; Symlets sym2 ; and Symlets sym4. Panel (c) presents the ROC curves for the proposed method under the aforementioned bases. On all instances $J = 2$ is employed. Comparing the ROC curves of all options, we can notice that Daubechies db2 and Symlets sym2 are the best choices. Panel (b) presents the ROC curves for five different levels of decomposition $J = \{1, 2, 3, 4, 5\}$ under the Daubechies db2 basis. Levels $J = 2, 3$ have a clear better performance, with a slight advantage to $J = 2$, since it is uses less levels on the decomposition. The overall performance of $J = 2$ warrants its use for the rest of the comparisons. Panel (d) shows

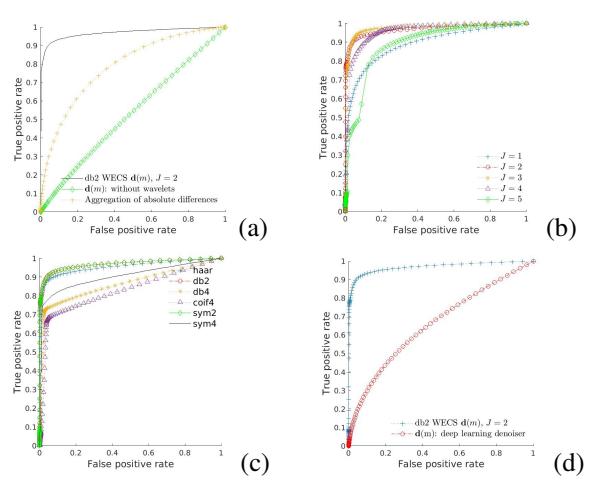


Fig. 4. ROC curves for detection of changing ellipses in simulated images and different methods. (a) The proposed methods in black (db2 WECS $\mathbf{d}(m)$) vs two non-wavelet methods: standard log-ratio aggregation (red stars); and $\mathbf{d}(m)$ (blue). (b) db2 $\mathbf{d}(m)$ with different levels. (c) $\mathbf{d}(m)$ with different wavelet bases and $J=2$; (d) the proposed db2 WECS $\mathbf{d}(m)$ (red circles) and $\mathbf{d}(m)$ with deep-learning feature extraction and without wavelets (yellow line).

how the proposed method performs in comparison to a deep learning feature extraction from a residual learning network [29]. WECS is applied with db2 wavelets and $J = 2$, whereas the other method applies WECS's idea, the only difference being that $\mathbf{X}^{(m)}$ is replaced by another smooth version of $\mathcal{I}^{(m)}$ that employs neural-networks trained to compete with wavelet smoothing. We can see that the ROC curves for images treated with a deep-learning method has a performance worse than that obtained with WECS. We finally have in Panel (a) the proposed WECS with db2 wavelets and $J = 2$ compared to two other non-wavelet methods. The first involves computing $\mathbf{d}(m)$ without wavelet smoothing, i.e., the squared deviations are computed using $\{\mathcal{I}^{(m)}\}$ instead of $\{\mathbf{X}^{(m)}\}$, and the classic method of analyzing aggregated absolute differences of $\{\mathcal{I}^{(m)}\}$. The ROC curves in Panel (a) show that the proposed WECS outperforms both the other two methods.

B. Actual remote sensing application

We employed the proposed change detection method on a series of 84 multi-date satellite images. The images were taken on a forest region at the border of Brazil and the French Guiana from 2015-12-26 to 2017-12-3. Each image has two channels (VV and VH) and 1538 by 1556 pixels. We perform a change detection wavelet analyses on the combined image by considering each observed entry as $\mathcal{I}_{k,l}^{(m)} = ((\text{VV}_{k,l}^{(m)})^2 + (\text{VH}_{k,l}^{(m)})^2)^{1/2}$, where VV and VH represent the matrices of observations from VV and VH channels, respectively.

A multi-resolution analysis based on a Symlet basis with filter of length 16 (symlet 8) is built. In order to have dimensions as power of 2, the matrices $\mathcal{I}_{k,l}$ are extended to a 2048×2048 matrix with $\mathcal{I}_{k,l}$ at the center and the remaining parts being completed with mirrored values at the borders. The wavelet transform at resolution level $J = 2$ is applied to these matrices and only the portion corresponding to the 1538×1556

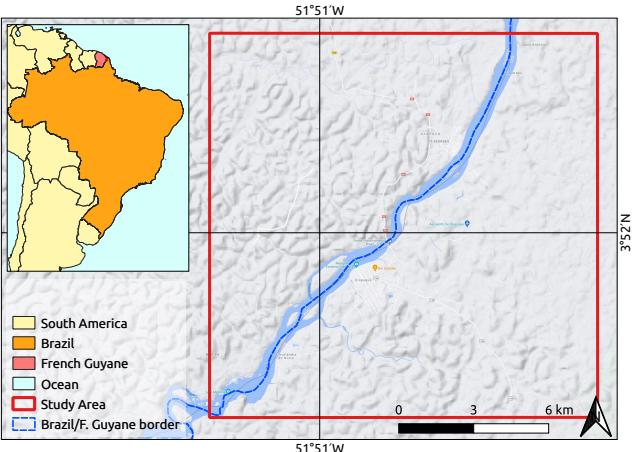


Fig. 5. Study area location. (preliminar)

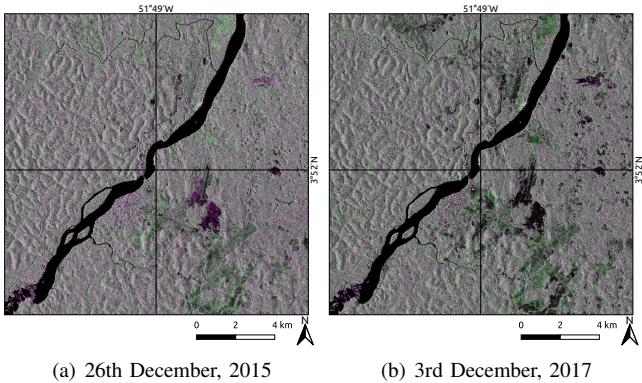
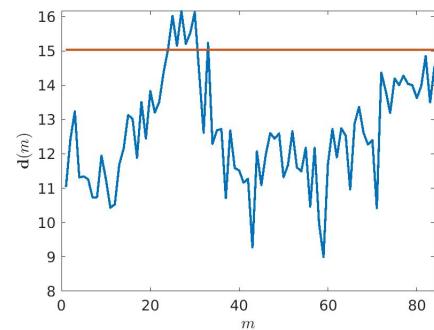
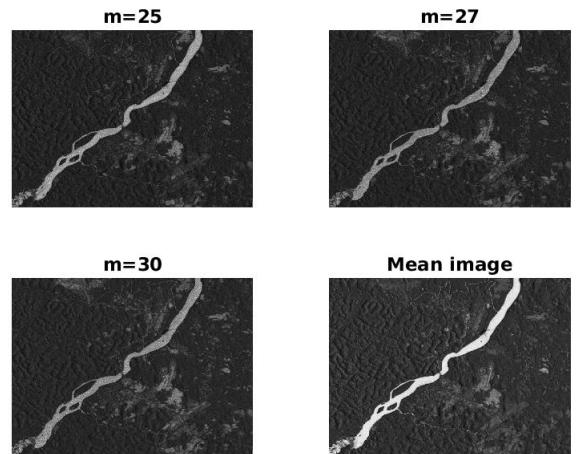


Fig. 6. Inserir caption...

is kept for further processing. Then, we are able to compute the squared mean differences vector \mathbf{d} and the matrix of absolute correlations \mathbf{R} .

The time change measures computed with \mathbf{d} are shown in Figure 7. The orange line in Figure 7a represents the median absolute deviation of \mathbf{d} , which allow us to notice times that differ expressively from the others. We can notice that times $m = 25, 27, 30$ are highlighted as having the most expressive changes. The images corresponding to these times can be seen in Figure 8.

The changes in space can be analyzed using the image obtained with \mathbf{R} , which is displayed in Figure 7c. Denoting the images' dimension as $N = 1538 \times 1556$, if we take the $[N / \log N]$ largest absolute correlations as those corresponding to possible change points, we obtain a matrix of zeros and

Fig. 7. Plot of $\mathbf{d}(m)$, $m = 1, \dots, 85$ with orange horizontal line indicating two times its median absolute deviation.Fig. 8. Images $\mathbf{X}^{(m)}$ from times m correspond to top-3 highest values of $\mathbf{d}(m)$ and the mean image $\bar{\mathbf{I}}$.

ones that is presented in Figure 7e. The white regions in Figure 7c (entries with value one) represent the change points, which seem to concentrate mainly on three regions: at the center, to the right of the river; at the top, on the left border of the river; and at the top left. Computing aggregated absolute differences to measure changes in space, we obtain Figure 7d. An image of detected changes corresponding to measures in Figure 7d can be obtained applying a thresholding method for grayscale images. Figure 7f shows the result of using Otsu's thresholding method [30] for aggregation of absolute differences.

The comparison of performance for the two change detection measures considered here can be checked in Figure 9, where a ROC curve is computed to check the correct detection of changes. WECS attains high true positive rates before the aggregation method, but both methods design competitive results. This performance differs from the ROC curve results in the previous section because the real images have a mean signal to noise ratio (SNR) of 3.9327, whereas the simulated images were generated with SNR of 0.2309. Hence, WECS is expected to have performance closer to the aggregation method in less noisy applications. The reference of correct change regions is shown in Figure 9a, which are determined using [XXXXX].

We can also compare the change detection methods using the F_1 -score accuracy measure. Denoting TP as the number of true positives (change pixels correctly detected), FP as the number of false positives (nonchange pixels flagged as change point) and FN as the total of false negatives (change points flagged as nonchange point), the F_1 -score is defined as

$$F_1 = 2 \frac{\text{Pr} \cdot \text{Re}}{\text{Pr} + \text{Re}},$$

where

$$\text{Pr} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{and} \quad \text{Re} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

Results of these three accuracy measures are presented in Table I. Results for aggregation of absolutes differences correspond to two types of thresholds: Otsu's method and Kittler-Illingworth [31]. We can observe that WECS presents the highest value of F_1 -score, which means it has a better performance than the competing method.

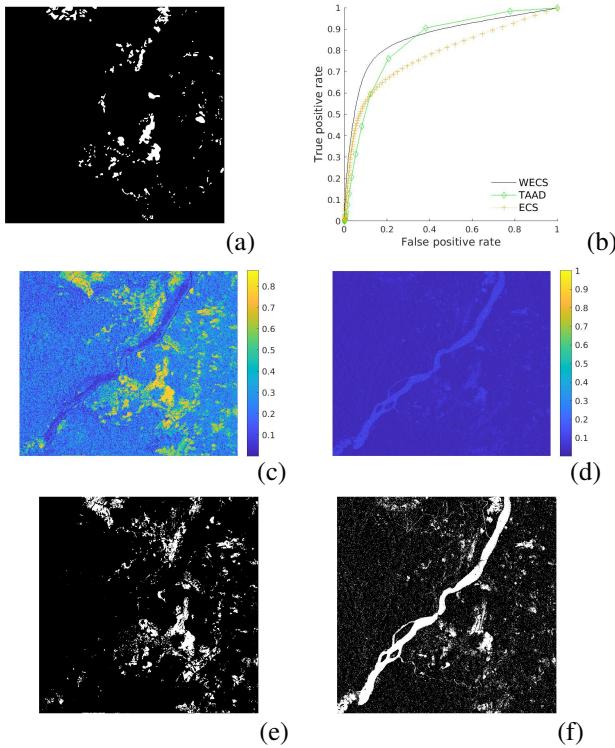


Fig. 9. Change points in space of the forest data. (a) Regions (in white) where changes should be detected. (b) ROC curve for detection of changing regions. (c) Matrix of absolute correlations obtained with WECS. (d) Change measures from aggregation of absolute differences. (e) Change regions detected by highlighting the $[N / \log N]$ largest correlations in \mathbf{R} (N denoting the image's dimension). (f) Change regions detected using aggregated absolute differences and Otsu's thresholding.

V. DISCUSSION

We present a novel way of detecting changes in multi-temporal satellite images, WECS. The procedure is based on wavelet energies from both the estimated individual coefficients as well as the whole mean image. It makes use of correlation screening for ultra-high dimensional data. Thereupon, WECS is expected to provide a sample of points in space in a way that

TABLE I
ACCURACY MEASURES F_1 -SCORE, PRECISION (Pr) AND RECALL (Re)
COMPUTED FROM DETECTION OF CHANGING REGIONS IN THE FOREST DATA
USING WECS (TAKING THE $[N / \log N]$ LARGEST CORRELATIONS AS
CORRESPONDING TO CHANGE) AND THE AGGREGATION OF ABSOLUTE
DIFFERENCES WITH OTSU AND KITTLER-ILLINGWORTH'S (KI)
THRESHOLDS.

	Aggregated - Otsu	Aggregated - KI	WECS
F_1 -score	0.2231	0.2163	0.3253
Pr	0.1369	0.1296	0.2390
Re	0.6022	0.6526	0.5094

such set, contains real change points with high probability. The proposed method's performance is shown using both simulated and real data. The proposed method is useful to detect spatio-temporal change points, which is illustrated on data analyses. The method is employed to analyze a time series of 84 images of a forest.

REFERENCES

- [1] T. L. Barreto, R. A. Rosa, C. Wimmer, J. B. Nogueira, J. Almeida, and F. A. M. Cappabianco, "Deforestation change detection using high-resolution multi-temporal X-band SAR images and supervised learning classification," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. Beijing, China: IEEE, 2016, pp. 5201–5204.
- [2] Y. Ban and O. A. Yousif, "Multitemporal spaceborne SAR data for urban change detection in China," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1087–1094, 2012.
- [3] C. Scher, N. C. Steiner, and K. C. McDonald, "Mapping seasonal glacier melt across the Hindu Kush Himalaya with time series synthetic aperture radar (SAR)," *The Cryosphere*, vol. 15, no. 9, pp. 4465–4482, 2021.
- [4] Y. Ban and O. Yousif, "Change detection techniques: A review," in *Multitemporal Remote Sensing*, Y. Ban, Ed. Cham: Springer, 2016, pp. 19–43.
- [5] L. Bruzzone and D. F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1171–1182, 2000.
- [6] T. Celik, "Change detection in satellite images using a genetic algorithm approach," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 2, pp. 386–390, 2010.
- [7] G. Quin, B. Pinel-Puyssegur, J.-M. Nicolas, and P. Loreaux, "MIMOSA: An automatic change detection method for SAR time series," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5349–5363, 2014.
- [8] S. Saha, F. Bovolo, and L. Bruzzone, "Change detection in image time-series using unsupervised LSTM," *IEEE Geoscience and Remote Sensing Letters*, 2020.
- [9] F. Bovolo and L. Bruzzone, "The time variable in data fusion: A change detection perspective," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 8–26, 2015.
- [10] S. Liu, D. Marinelli, L. Bruzzone, and F. Bovolo, "A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 2, pp. 140–158, 2019.
- [11] T. Matsunaga, A. Iwasaki, S. Tsuchida, K. Iwao, J. Tanii, O. Kashimura, R. Nakamura, H. Yamamoto, S. Kato, K. Obata, K. Moura, and T. Tachikawa, "Current status of hyperspectral imager suite (HISUI) onboard International Space Station (ISS)," in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. Fort Worth, USA: IEEE, 2017, pp. 443–446.
- [12] M. Jia and L. Wang, "Novel class-relativity non-local means with principal component analysis for multitemporal SAR image change detection," *International Journal of Remote Sensing*, vol. 39, no. 4, pp. 1068–1091, 2018.
- [13] B. Hou, Q. Wei, Y. Zheng, and S. Wang, "Unsupervised change detection in SAR image based on Gauss-log ratio image fusion and compressed projection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 8, pp. 3297–3317, 2014.

- [14] B. Du, L. Ru, C. Wu, and L. Zhang, "Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9976–9992, 2019.
- [15] Y. Chen, Z. Ming, and M. Menenti, "Change detection algorithm for multi-temporal remote sensing images based on adaptive parameter estimation," *IEEE Access*, vol. 8, pp. 106 083–106 096, 2020.
- [16] F. Bovolo and L. Bruzzone, "A detail-preserving scale-driven approach to change detection in multitemporal SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 12, pp. 2963–2972, 2005.
- [17] B. Vidakovic, *Statistical Modeling by Wavelets*. New York: John Wiley & Sons, 1999.
- [18] P. A. Morettin, A. Pinheiro, and B. Vidakovic, *Wavelets in Functional Data Analysis*. Cham: Springer, 2017.
- [19] A. M. Atto, E. Trouvé, Y. Berthoumieu, and G. Mercier, "Multidate divergence matrices for the analysis of SAR image time series," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 4, pp. 1922–1938, 2012.
- [20] N. Bouhlel, G. Ginolhac, E. Jolibois, and A. Atto, "Multivariate statistical modeling for multi-temporal SAR change detection using wavelet transforms," in *2015 8th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi-Temp)*. Annecy, France: IEEE, 2015, pp. 1–4.
- [21] T. Celik, "Multiscale change detection in multitemporal satellite images," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 820–824, 2009.
- [22] S. Cui and M. Datcu, "Statistical wavelet subband modeling for multi-temporal SAR change detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 4, pp. 1095–1109, 2012.
- [23] J. Fan, R. Li, C.-H. Zhang, and H. Zou, *Statistical Foundations of Data Science*. Boca Raton: CRC Press, 2020.
- [24] J. Fan and J. Lv, "Sure independence screening for ultrahigh dimensional feature space," *Journal of the Royal Statistical Society: Series B*, vol. 70, no. 5, pp. 849–911, 2008.
- [25] S. G. Mallat, *A Wavelet Tour of Signal Processing*. San Diego: Academic Press, 1998.
- [26] ———, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989.
- [27] R. R. Coifman and D. L. Donoho, "Translation-invariant de-noising," in *Wavelets and statistics*, A. Antoniadis and G. Oppenheim, Eds. New York: Springer, 1995, pp. 125–150.
- [28] A. M. Atto, E. Trouvé, J.-M. Nicolas, and T. T. Lê, "Wavelet operators and multiplicative observation models—application to SAR image time-series analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6606–6624, 2016.
- [29] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142–3155, 2017.
- [30] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [31] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, vol. 19, no. 1, pp. 41–47, 1986.