IMAGE PROCESSING ON LINE

Published in Image Processing On Line on 2022–12–28.
Submitted on 2022–06–30, accepted on 2022–11–01.
ISSN 2105–1232 © 2022 IPOL & the authors CC–BY–NC–SA
This article is available online with supplementary materials,
software, datasets and online demo at
https://doi.org/10.5201/ipol.2022.416

Detection and Interpretation of Change in Registered Satellite Image Time Series

Tristan Dagobert[1], Rafael Grompone von Gioi[1],
Carlo de Franchis[1,2], Charles Hessel[1,2]

[1]Centre Borelli, ENS Paris-Saclay, France
{tristan.dagobert, grompone, carlo.de-franchis, charles.hessel}@ens-paris-saclay.fr
[2]Kayrros, France

Communicated by Luis Álvarez   *Demo edited by* Tristan Dagobert

Abstract

Time series of satellite images are now massively available thanks to the existence of several constellations of recurrent satellites. We propose a method for detecting and measuring the duration of changes on such series. This approach is intended to be generic and independent of the type of satellite used, whether band limited or multispectral. It is based on a global analysis of the sequence. The statistical detection method is applied to a residual sequence computed from backward and forward novelty filters applied to all images in the series. Significant changes are computed with a guarantee on their number of false alarms (NFA). To establish the efficiency of the method, we have created an open database of 28 sequences of 20 images acquired by the Sentinel-2 satellite, in different regions of the world. We obtain satisfactory results, which are consistent with the visual observations.

Source Code

The C source code implementing the algorithm described in the paper, an online demo, and some of the data used in the experiments are accessible at the associated web page[1].

Supplementary Material

A reference dataset, to be used for further comparisons, is provided with the article and available at the address `https://www.ipol.im/data/algo/416/DICRSITS.zip`.

Keywords: change detection; multi-temporal; time series; satellite; dataset

---

[1] https://doi.org/10.5201/ipol.2022.416

Tristan Dagobert, Rafael Grompone von Gioi,
Carlo de Franchis, Charles Hessel , *Detection and Interpretation of Change in Registered Satellite Image Time Series*, Image Processing On Line, 12 (2022), pp. 625–651. https://doi.org/10.5201/ipol.2022.416

1 Introduction

In an image series, the detection of land cover or urban changes is an essential preliminary step of any optical satellite image interpretation. Over the past decade, an increasing number of missions involve large satellite constellations which considerably reduce the revisit time. Furthermore, the revisit is repeatedly made from the same point of view, often with heliosynchronous orbits ensuring slowly varying illumination. This opens the way to reliable and frequent change detection. While bi-temporal comparisons only give a set of differences, a multi-temporal approach enables a more complete interpretation of time variations.

We propose an unsupervised approach which is an enhanced version of the algorithms described in [13] and [14]. It derives from the ideas of [6] where an estimator is computed for each time and pixel thanks to a backward-forward summation. This leads to search, over time and for each pixel, for significantly high values when several changes occur. The estimator is based on residuals computed from the novelty filter of [25] so that it is more robust against linear variations of illuminations. The change detector combines hue and illumination variations. We control the false detection rate by applying an *a contrario* statistical test [16]. In order to test the duration of the changes, we use a robust estimator based on the cross-correlation (see Figure 1). To evaluate our method, we created and made available an open dataset of multi-temporal registered Sentinel-2 time series. Our paper makes the following contributions:

- a general change detection method, applicable to any observation satellite with any number of spectral bands;
- a mathematically founded change estimator derived from the novelty filter;
- a database of 28 Sentinel-2 time series, representing different regions of the world observed during 4 years and in which significant changes occur.

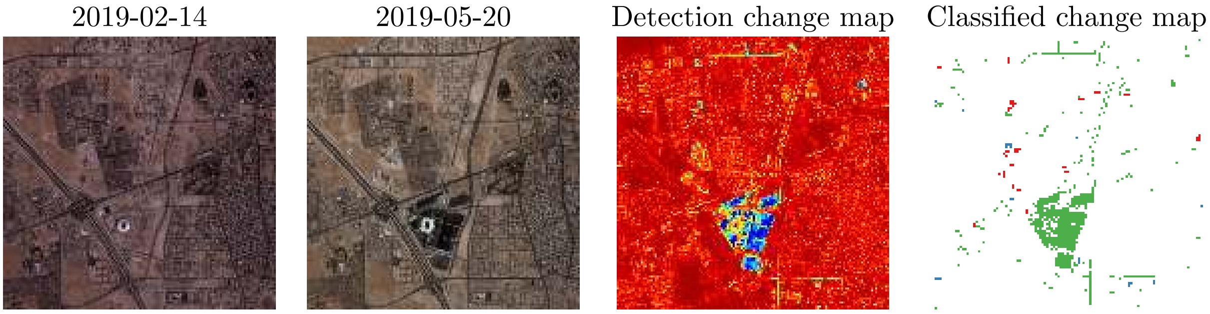

Figure 1: Example of a detection map (left) and a classified change map (right) obtained by our method over the region of Al Wakrah (World cup soccer). The color code (see Figure 5) characterizes the duration of the change.

2 Related Work

This paper stands at the intersection of the fields of change detection in time series and of remote sensing. Exhaustive surveys on this topic can be found in [1] and [40] for the change point detection and satellite change detection problems respectively. Some methods are called "multi-temporal" in remote sensing because they process time series, yet they only estimate the changes in pairs. We call such methods "pair based", while we shall call "multi-temporal", the algorithms which rely on more than two frames to detect the changes occurred in one of them. These methods can be classified according to four independent modalities (see Table 1) namely:

- Time exploitation: bi-temporal (BT) vs. multi-temporal (MT);
- Spectrum exploitation: band limited (BS) vs. multi-spectral (MS);
- Input image resolution: low (LR) vs. high (HR);
- Change localization: pixel-based (P), region-based (R), object-based (O).

Publication	Temporality	Spectrum	Resolution	Localization
[49], [48], [10]	BT	MS	HR	R
[46], [47], [31], [17]	BT	MS	LR	P
[9], [45], [41]	BT	MS	LR	R
[8], [7]	BT	MS	LR	O
[37], [28]	BT	BS	HR	P
[30]	BT	BS	HR	R
[23], [36]	BT	BS	HR	O
[38],[35]	MT	MS	LR	R
[2]	MT	PS	LR	P
[42], [43], [5], [20], [21], [50], [29], [4], [15]	MT	MS	LR	P

Table 1: Classification of image change detection algorithms according to four features.

Change detection on image pairs. The authors of [45] describe a method based on an iterative slow feature analysis (ISFA) followed by a Bayesian soft fusion for bi-temporal multispectral remote sensing data. In [46], Zanetti et al. consider the difference image and assume that its pixels are split into two classes w_n (no change) and w_c (change); the distribution associated to a pixel ρ is $p(\rho) = p(w_n)p(\rho|w_n) + p(w_c)p(\rho|w_c)$. Assuming that the residual for no-change follows $\mathcal{N}(0, \sigma_n^2)$ and for changes $\mathcal{N}(\mu_b, \sigma_n^2)$ (b is the band), then $p(\rho|w_n)$ follows a Rayleigh and $p(\rho|w_c)$ a Rice distribution respectively. The parameters of the distributions are estimated using an EM algorithm. Then a threshold is computed to decide between no-change and change. In [47], Zanetti and Bruzzone enhanced the method via a Bayesian theoretical framework to describe the statistical distribution of the difference image. In [9], Celik et al., dealing with the image of the differences, analyze a feature space built by PCA and then cluster it by the k-means method. Luo et al. [31] use three unsupervised change detection methods such as CVA, IRMAD and ISFA, then fuse their results by the Dempster-Schafer theory. In [30], Liu et al. compute global and local mapping functions between two images, estimated through SIFT keypoints extraction and matching. Based on these mappings, keypoint matchings are then used to detect and group regions of changes. Both steps are defined through an *a contrario* framework, which simplifies the parameter setting.

Zhan et al. [48] propose a deep siamese network to create feature vectors, then classify them according to Euclidean distance between the 1st and the 2nd image. Zhang et al. propose in [49] a solution accepting pairs of different resolutions and modalities (e.g. SAR and RGB) with at first an autoencoder for feature learning followed by a mapping function learning. Both are fully connected networks. In [17], Du et al. use two symmetric deep networks. Then a Slow Feature Analysis module is deployed to suppress the unchanged components and highlight the changed components of the transformed features. Finally, the change intensity is calculated with a χ-square distance and the changes are determined by threshold algorithms.

Chen et al. [10] present two deep siamese convolutional networks with a multi-scale approach. To alleviate the effect of unregistered objects in the pair, Ren et al. [37] propose to compare the two latent images produced by an optimized GAN model. The largest values of the absolute difference between these last ones compose the change map. Oriented to change object-based detection in HR images, Ji et al. [23] propose an architecture composed by a Mask R-CNN network for object-based instance segmentation and a multi-scale full convolutional network for pixel-based semantic segmentation while Peng et al. [36] present UNet++ an encoder-decoder, where change maps could be learned from scratch using available annotated datasets. Daudt et al. [8] show two fully connected deep siamese networks trained on the Onera satellite change detection dataset [7] composed of Sentinel-2 pairs.

Change detection on image time series. In the domain of the data transfer, Kohonen and Oja [25] proposed an orthogonalizing filter that extracts the updated parts of an input image, with respect to previously learnt patterns. They applied this method on a set of typographic letters composed by binary black and white patterns.

Verbesselt et al. [42, 43] define a season-trend model, called BFAST, with a linear trend and a harmonic season so that the time series data are modeled by a discrete sine truncated series. To measure the discrepancy between the model and the measurements for the monitored period, the BFAST resorts to moving sums. A break is declared if the moving sum statistics (MOSUM) of Chu et al. [12] exceeds some boundary that is asymptotically only crossed with a 5% probability. In [20], Hamunyela et al. adapt the BFAST method to take into account the spatial coherence.

With a similar approach, Zhu et al. [50] propose a regression model based on the first Fourier coefficients called the Continuous Change Detection (CCD) algorithm, which takes into account the seasonality and the trend of the surface reflectance over years. The model is fitted to a Landsat multi-spectral dataset with a RMSE loss. When the residual of new fitted observations exceeds three consecutive times the learned RMSE, the pixel is considered as a change. Several algorithms reuse this method. Liu et al. [29] use spatio-temporal rules and dense time series stacks of a Landsat dataset. First, a stable area mask based on image classification in the start and the end years is generated to remove pixels that are persistent or spatially irrelevant. Then the CCD method is employed to determine the change points when non-impervious cover converts to an impervious surface based on the property of temporal irreversibility.

In [2] Brooks et al. first apply a harmonic regression model to the data in the CCD fashion, then test the deviation of the fitted residuals aggregated by an exponentially weighted moving average. In order to robustify the CCD method, Bullock et al. [4] combine it with the change point CUSUM method of Brown et al. [3], then remove their false positives with the Chow test [11]. In [15] Deng et al. use the CCD as an encoder, extracting the time series model coefficients. These features then serve as inputs for a random forest classifier.

Using a combinatorial technique, Hermosilla et al. [21] obtain a global trend of the variations by defining first the segments between each temporal sample pixel, then by merging each pair of adjacent segments according to a bottom-up strategy. The changes are identified as the segment extremities. In [5] Cai et al. propose a non parametric method with a two time scale strategy to detect land cover changes. In the first step a pair of time series of neighboring years are compared thanks to the two-sample Kolmogorov-Smirnov (KS) test. In the second step they detect changes within years with a threshold on the difference images. In [35], Papadomanolaki et al. present a deep learning framework for urban change detection which combines a U-Net-like network for the feature representation and a recurrent network for the temporal modeling. While learning is made with several dates, testing is performed on Sentinel-2 pairs.

Change detection in multivariate time series. In [6], Carlstein proposes a method to estimate the change time θ, assumed unique, of a time series $(X_n)_{1\leq n\leq N}$. The only requirement is that the cumulative distribution functions F and G, representing respectively the data behaviour before and after the change, differ on a set of positive probability. Defining an estimator $D(t_n)$ based on the absolute differences $|F_{t_n}(X_n) - G_{t_n}(X_n)|$, where the empirical CDF F_{t_n} and G_{t_n} are built for all $1 \leq t_n \leq N$, the author shows that the t_n that maximizes $D(t_n)$ is the closest estimator of θ.

Nielsen et al. [34] introduce the multivariate alteration detection (MAD) transformation over multispectral (k bands) data, which is based on the established canonical correlations analysis. It also proposes using postprocessing of the change detected by the MAD variates using maximum autocorrelation factor (MAF) analysis. The goal is to concentrate information on change, by an adequate linear transformations of the image data difference made thanks to the MAD transformation. This ensures a spatial coherence thanks to the MAF transformation.

In [26] Lavielle and Teyssière consider the multiple change-point problem for a $m-$variate time series $\{\mathbf{Y} = Y_{1,t}, \ldots, Y_{m,t}\}$, including strongly dependent processes, with an unknown number K of change-points. They assume that the covariance structure of the series changes abruptly at some unknown common change-point times $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_K)$. Defining the contrast function to minimize as $U(\boldsymbol{\tau}) = J(\boldsymbol{\tau}, \mathbf{Y}) + \beta \mathrm{R}(\boldsymbol{\tau})$ with J the data-fit function and R the penalization function, they propose a method able to detect changes in multivariate i.i.d., weakly and strongly dependent series with adaptive approaches to estimate both β and K.

In [33] Matteson and James propose an a-posteriori (i.e. offline) method to estimate the K different change times $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_K)$ of a m-variate time series $\{\mathbf{X} = X_{1,t}, \ldots, X_{m,t}\}$. They do not make any assumptions regarding the nature of the change in distribution or any distribution assumptions beyond the existence of the α^{th} absolute moment, for some $\alpha \in]0, 2[$. Estimation is based on a hierarchical clustering where they first iteratively divide, then agglomerate intervals where one change-point was located. In [18], Eichinger and Kirch investigate statistical properties of change point estimators based on the MOSUM statistics [12]. They consider a multiple mean change model with possible time series errors and prove that the number and location of change points are estimated consistently.

Our approach. Among the methods considered above, a major drawback of methods based on pairwise comparison is that they generally use an arbitrarily chosen reference image. The relationships between all the images are therefore not taken into account, which constitutes a serious information loss. On the contrary, the advantage of the novelty filter of Kohonen et al. [25] is that it compares each single image to a set of reference images; we borrow this powerful idea for our elementary estimator. We also strengthen change detection, by taking up the idea of Carlstein [6] of temporally symmetrizing the change estimator.

The other methods dealing with image time series are essentially based on techniques for detecting change in univariate series such as the MOSUM [12] or CUSUM [3] statistics. Roughly speaking, this leads to define a null hypothesis testing $\mathcal{H}_0$ specific to each pixel x, where one assumes that a pixel observation Y_x follows a distribution f_x. This amounts *in fine* to assume that f_x has an analytic shape (generally Gaussian) and that the N times series is long enough to estimate its parameters μ_x, σ_x from the samples $(y_x)_{2 \leq n \leq N}$. The implicit assumption is that a pixel x undergoes few variations during the considered time period, so that the estimates μ_x, σ_x are reliable. This approach has a major limitation because:

- the sample length is in practice quite small to finely estimate the parameters. For example, in temperate or tropical environments image samples are too often perturbed by large cloud cover. In more favourable regions, one observes only between 10 and 30 cloud free samples over a year;
- one does not know *a priori* where the region changes are located;
- one ignores the frequency of changes given the position.

This sample scarcity leads to consider a distribution f applicable to all pixels and where its parameters, if any, can be computed from a large sample. This is what we will show in the next section by defining a $\mathcal{H}_0$ distribution based on the maximal acceptable number of changes per pixel over the time.

3 Algorithm Description

3.1 Overview

Our algorithm consists of two parts. The first one performs the actual detection from a time series. The detection is based on an estimator constructed from residuals calculated by the novelty filter. The second part classifies the changes according to their nature: unique or permanent.

The algorithm is summarized in Algorithm 1 and its parameters in Table 2. We consider an input multi-spectral time series $(\mathbf{u}_{c,n}^{\text{raw}})_{1\leq n\leq N}^{1\leq c\leq C}$ of $N \geq 3$ images composed by C channels and defined on the same discrete domain $\Omega = [W] \times [H]$ where W (resp. H) are the images width (resp. height). We suppose in addition that these images have been jointly subpixel registered; for this task we used the method of Hessel et al. [22]. Pixel positions are denoted by $x = (x_1, x_2)^\top$ for all $x_1 \in [H]$ and $x_2 \in [W]$. To each image $\mathbf{u}_{c,n}$ we associate its vector representation $\mathbf{v}_{c,n}$ such that $\mathbf{v}_{c,n}(x) = \mathbf{v}_{c,n}(x_1 W + x_2) = \mathbf{u}_{c,n}(x_1, x_2)$.

As the noise of satellite images naturally follows a Poisson law, we apply a pseudo gamma correction to make its standard deviation approximately constant. So we set

$$\mathbf{v}_{c,n}^{\gamma} = \sqrt{\mathbf{v}_{c,n}}. \tag{1}$$

Notice that in Equations (2), (3) and (4) which follow, the expressions are general enough to be applied indifferently to the series $(\mathbf{v}_{c,n})_{1\leq n\leq N}^{1\leq c\leq C}$ or $(\mathbf{v}_{c,n}^{\gamma})_{1\leq n\leq N}^{1\leq c\leq C}$.

3.2 Novelty Filtering

The change detection or change point detection tries to identify times when the probability distribution of a stochastic process or time series changes. Changes can be detected by linear novelty filters [32], first introduced by Kohonen and Oja [25]. We present a variant of the original formulation, described as follows:

Definition 1. *Let $\mathbf{v}_1, \ldots, \mathbf{v}_M \in \mathbb{R}^{|\Omega|}$ be M distinct Euclidean vectors spanning a positive convex cone $\mathcal{L}^c$, i.e. the sub-space such that for all α and β in $\mathbb{R}_+^*$, and for all $\mathbf{v}$ and $\mathbf{w}$ in $\mathcal{L}^c$, $\alpha\mathbf{v} + \beta\mathbf{w} \in \mathcal{L}^c$. Then any vector $\mathbf{v} \in \mathbb{R}^{|\Omega|}$ can be uniquely written as*

$$\mathbf{v} = \hat{\mathbf{v}} + \mathbf{r}, \tag{2}$$

where $\hat{\mathbf{v}} \in \mathcal{L}^c$ is the closest vector to $\mathbf{v}$ in $\mathcal{L}^c$ and $\mathbf{r}$ is the residual i.e. the novelty.

The computation of the novelty $\mathbf{r}$ corresponds to the non-negative least-square (NNLS) problem

$$\begin{aligned} \mathbf{r} &= \mathbf{v} - \mathbf{A}\mathbf{z}^*, \\ \text{s.t. } \mathbf{z}^* &= \arg\min_{\mathbf{z}} \|\mathbf{v} - \mathbf{A}\mathbf{z}\|_{\ell_2}, \mathbf{z} \geq \mathbf{0}. \end{aligned} \tag{3}$$

where the vector $\mathbf{A}\mathbf{z}^*$ represents $\hat{\mathbf{v}}$ and the matrix $\mathbf{A} = [\mathbf{v}_1 \ldots \mathbf{v}_M]$ is the concatenation of the column vectors $(\mathbf{v}_m)_{1\leq m\leq M}$. The resolution of (3) can be made by the Lawson and Hanson [27] numerical method. Definition 1 differs from the original version [25] where there was no constraint on $\mathbf{z}$. In particular, some of its components could be negative. The reason why we impose the non negativity is that, at least in visible bands, landscape images should only result in a superimposition of the past component images with positive contrasts.

3.3 Sliding Change Estimator

General formulation. We consider the multi-spectral series $(\mathbf{v}_{c,n})_{1\leq n\leq N}^{1\leq c\leq C}$ in its vector representation. We denote by $\mathbf{r}^b_{c,n,V}$ (resp. $\mathbf{r}^f_{c,n-1,V}$) the time backward (resp. forward) residual vector computed according to (3) where the vector $\mathbf{v}$ of this equation corresponds to $\mathbf{v}_{c,n}$ (resp. $\mathbf{v}_{c,n-1}$), the matrix $\mathbf{A}$ to the matrix $[\mathbf{v}_{c,n-V}\dots\mathbf{v}_{c,n-1}]$ (resp. $[\mathbf{v}_{c,n}\dots\mathbf{v}_{c,n+V-1}]$) and V to the number of basis vectors composing the matrix $\mathbf{A}$. This leads to define for all $n\geq 2$ the scalar estimator

$$r_{c,n-\frac{1}{2},V}(x) = \frac{1}{2}\left(|\mathbf{r}^b_{c,n,V}(x)| + |\mathbf{r}^f_{c,n-1,V}(x)|\right). \tag{4}$$

By convention the vectors $\mathbf{v}_{c,1}$ and $\mathbf{v}_{c,N}$ are replicated as many times as necessary when the windows depending on V exceed the boundaries of $[N]$. We use the half index $n-\frac{1}{2}$, first to emphasize that the change occurs between times $n-1$ and n, secondly to have a symmetric formula according to the backward and forward temporal directions (see Figure 2). Concerning the parameter V, the larger the temporal window V, the more it captures the variations before or after the change. In addition, it is easy to check that if a single change occurs at time n_0, then the residual $r_{c,n_0-\frac{1}{2},V}(x)$ is a global maximum of the time series.

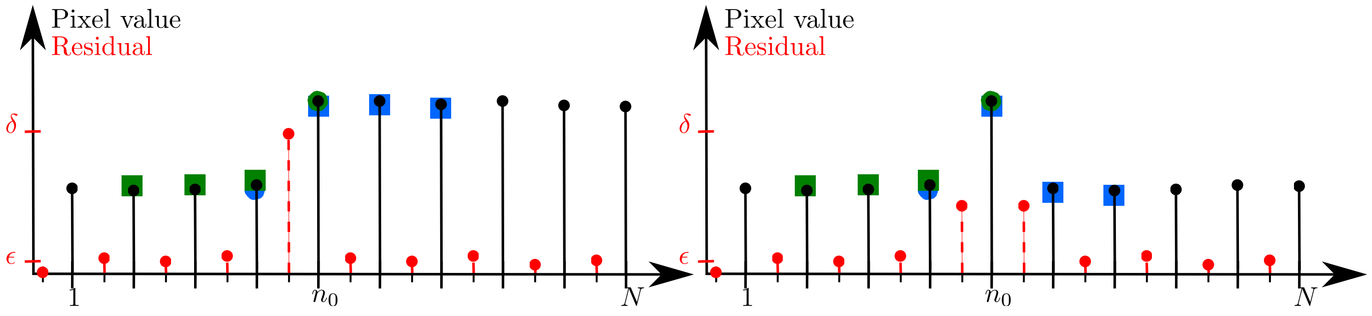

Figure 2: On the left, in black, temporal value profile of a pixel which belongs to a permanent surface change ω appearing from n_0; in red, its temporal residual value. On the right, in black, temporal value profile of the pixel which belongs to the same surface ω appearing only once at n_0. The squares (resp. the discs) indicate the pixels belonging to the basis (resp. the projected) vectors, for $V=3$. In green (resp. in blue), those implied in the left (resp. right) elementary residual of (4).

Experiments conducted by Dagobert et al. [13] show that two complementary estimators can be defined: a hue-change estimator which captures mainly the landscape variations and a contrast-based estimator which captures mainly the urban changes. Sticking to this idea, we propose here an enhancement of both estimators as well as an analytic interpretation.

Hue-change estimators. To justify our purpose we first clarify the meaning of the terms "hue", "chrominance" and "lightness". According to Fairchild [19], the hue of a color is "the degree to which a stimulus can be described as similar to or different from stimuli that are described as red, orange, yellow, green, blue, violet" [19, p. 85], the chrominance (or chroma) is the "colorfulness of an area judged as a proportion of the brightness of a similarly illuminated area that appears white or highly transmitting" [19, p. 87] and the lightness is "the brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting" [19, p. 86]. Consequently colors can keep the same hue while their lightness and chrominance may vary, which one can see on Figure 3 where the color set is represented geometrically by a bicone.

From these considerations and in order to catch better the hue variations, we first change of color space: we transform the RGB (or the RGBI) representation of the image $\mathbf{v}^\gamma_n$ into a luminance (i.e lightness) plus chrominance representation, which can be seen as a YUV transformation generalized to the infrared channel and such that

$$\begin{pmatrix} \mathbf{v}_n^{\text{lum}}(x) \\ \mathbf{v}_{R,n}^{\text{chr}}(x) \\ \mathbf{v}_{B,n}^{\text{chr}}(x) \\ \mathbf{v}_{I,n}^{\text{chr}}(x) \end{pmatrix} = \begin{pmatrix} 1/C & 1/C & 1/C & 1/C \\ 1-1/C & -1/C & -1/C & -1/C \\ -1/C & -1/C & 1-1/C & -1/C \\ -1/C & -1/C & -1/C & 1-1/C \end{pmatrix} \begin{pmatrix} \mathbf{v}_{R,n}^{\gamma}(x) \\ \mathbf{v}_{G,n}^{\gamma}(x) \\ \mathbf{v}_{B,n}^{\gamma}(x) \\ \mathbf{v}_{I,n}^{\gamma}(x) \end{pmatrix}. \tag{5}$$

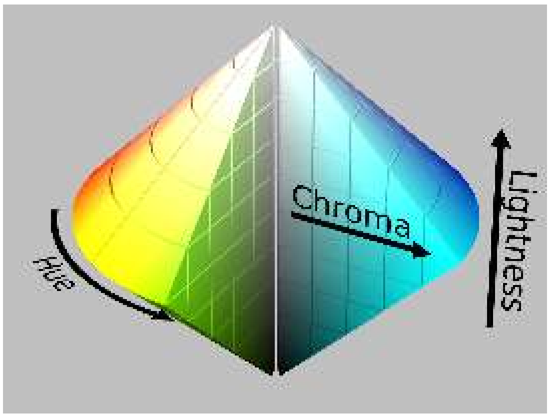

Figure 3: Representation of the colors according to their hue, chrominance and lightness properties such as described by Joblove and Greenberg [24, fig.5].

From this transformation, we define two related estimators : a luminance-based and a chrominance-based. Concerning the luminance, we compute the residuals related to $(\mathbf{v}_n^{\text{lum}})_{1\le n\le N}$ according to (3) then apply (4) which yields the series $(r_{n-1/2,V}^{\text{lum}})_{1\le n\le N}$.

For the chrominance, the computation is more subttle. We first form the vectors by concatenating the chrominance channels together so that

$$\mathbf{v}_n^{\text{chr}} = [\mathbf{v}_{R,n}^{\text{chr}\top}, \mathbf{v}_{B,n}^{\text{chr}\top}, \mathbf{v}_{I,n}^{\text{chr}\top}]^\top. \tag{6}$$

From these concatenated vectors, we then compute the residuals according to (3) then split them again into their channels and apply the relation (4) to obtain the residual series of each channel namely $(r_{R,n-1/2,V}^{\text{chr}})_{1\le n\le N}$, $(r_{B,n-1/2,V}^{\text{chr}})_{1\le n\le N}$ and $(r_{I,n-1/2,V}^{\text{chr}})_{1\le n\le N}$.

As all the chrominance channels of a basis vector $\mathbf{v}_k^{\text{chr}}$ are scaled by the same scalar z_k according to (3) (i.e. keep the same direction in the bicone), then low values of the residuals indicate only change of chrominance. On the contrary, if we observe significant values of these residuals, then a local change in the orientation of the chrominance occurred, which can be interpreted as an evolution of the hue. That's why we call them "hue-change" estimators.

We can interpret the novelty filter applied on the images $(\mathbf{v}_n^{\gamma})_{1\le n\le N}$ as an optimal adjustment, in the sense of the NNLS, of the luminance and the chrominance of the basis vectors according to those of $\mathbf{v}_n^{\gamma}$. It follows from this observation that the estimators $(r_{n-1/2,V}^{\text{lum}})_{1\le n\le N}$ are invariant to positive independent multiplicative factors of the luminance for each time, while the others are invariant to independent positive multiplicative factors of the chrominance.

To homogenize the notation in Algorithm 1, we set

$$r_{2,n-\frac{1}{2},V}^{\text{hue}} \equiv r_{R,n-\frac{1}{2},V}^{\text{chr}}, r_{3,n-\frac{1}{2},V}^{\text{hue}} \equiv r_{B,n-\frac{1}{2},V}^{\text{chr}}, r_{4,n-\frac{1}{2},V}^{\text{hue}} \equiv r_{I,n-\frac{1}{2},V}^{\text{chr}} \text{ and } r_{1,n-\frac{1}{2},V}^{\text{hue}} \equiv r_{n-\frac{1}{2},V}^{\text{lum}} \tag{7}$$

knowing that the last relation is an abuse of notation because luminance and hue are conceptually different (we set the same exponents because luminance and chrominance terms where computed at the same time in (5)).

Detecting local chrominance changes is interesting in the landscape context, but may be less effective in the urban context. Indeed, local variations of the greyish urban environment mainly impact their luminance.

Contrast-based estimator. Because the two previous estimators are not well designed to measure single channel variations between times, we introduce another estimator. We consider here the RGB (or RGBI) space color, and process independently the channels as we do not want to miss changes occurring in a single one. Starting from the gamma corrected series $(\mathbf{v}^{\gamma}_{c,n})^{1\leq c\leq C}_{1\leq n\leq N}$ defined in (1), we rename it as the series $(\mathbf{v}^{\text{con}}_{c,n})^{1\leq c\leq C}_{1\leq n\leq N}$ and consider the decomposition $\mathbf{v}^{\text{con}}_{c,n} = \mathbf{s}^{\text{con}}_{c,n} + \bar{\mathbf{v}}^{\text{con}}_{c,n}$ where $\bar{\mathbf{v}}^{\text{con}}_{c,n}$ is the spatial mean. We then define the time backward residual by

$$\mathbf{r}^{\text{con},b}_{c,n,V} = \left(\bar{\mathbf{v}}^{\text{con}}_{c,n} - \frac{1}{V}\sum_{k=n-V}^{n-1} \bar{\mathbf{v}}^{\text{con}}_{c,k}\right) + \left(\mathbf{s}^{\text{con}}_{c,n} - \mathbf{A}\mathbf{z}^{*}\right), \tag{8}$$

where $\mathbf{z}^{*} = \arg\min ||\mathbf{s}_{c,n} - \mathbf{A}\mathbf{z}||$ and $\mathbf{A} = [\mathbf{s}^{\text{con}}_{c,n-V}, \ldots, \mathbf{s}^{\text{con}}_{c,n-1}]^{\top}$. We express the time forward residual in the same way and from (4) we construct the contrast-based estimator series $(r^{\text{con}}_{c,n-\frac{1}{2},V})^{1\leq c\leq C}_{1\leq n\leq N}$. On the one hand, we can interpret the right term of the right member, the novelty filter applied on the images $\mathbf{s}_{c,n}$ as an optimal adjustment, in the sense of the NNLS, of the spatial contrast of the basis vectors according to that of $\mathbf{s}^{\text{con}}_{c,n}$. Indeed as the basis vector $\mathbf{s}^{\text{con}}_{c,k}$ is zero-centered and scaled by the scalar z_k, we update in some way the variance of the image $\mathbf{v}^{\text{con}}_{c,k}$. On the other hand, the left term of the right member quantifies a mean luminance variation between the tested vector and the averaged basis vectors.

Notice that in our implementation one can use both the hue-change and contrast-based estimators together or only one of the two. Indeed as they capture variations of quite different natures, their individual results also are different.

3.4 Multi-Scale and Shift Patch Strategy

When there is no change in the considered domain Ω, then one can roughly consider that the residuals are of the same order of magnitude as the ones caused by noise. By noise, we mean not only Poisson observation noise, but also the casual small luminance and hue variations at different dates. Let ϵ denote their order of magnitude. Inherently, the novelty filter incorrectly calculates the residuals when there is an area ω with significant change in the domain Ω. Indeed, when there is a perturbation in ω its residual has an order of magnitude of $\delta \gg \epsilon$ and in $\Omega \setminus \omega$ of order $\eta \gtrsim \epsilon$ as showed Figure 4-left. Moreover, if the perturbation is not negligible, then $\delta \geq \eta \gg \epsilon$ so that we cannot discriminate the changes anymore (see Figure 4-right).

In order to attenuate this behavior and increase the precision of the residuals of the unchanged domain $\Omega \setminus \omega$, we propose to compute several residuals over patches of different sizes and positions. The idea is that among the different patches, one contains the lowest possible change surface ω. Consequently for a given position x we take the minimal residual computed over the set of patches containing x. Formally, we denote by $\mathcal{T}_{q,d_1,d_2}$ the tiling of Ω as the set of the patches $\{\omega_{i+d_1,j+d_2}\}$ of size $2^q \times 2^q$

$$\begin{aligned}\mathcal{T}_{q,d_1,d_2} = \{\omega_{i+d_1,j+d_2}\} \text{ for all } i &\in \{0+d_1, \ldots, 2^q+d_1, \ldots, W-(W \mod 2^q)+d_1\},\\ j &\in \{0+d_2, \ldots, 2^q+d_1, \ldots, H-(H \mod 2^q)+d_2\},\end{aligned} \tag{9}$$

where d_1 and d_2 are two shifts and with the convention that Ω has a toric extension. Let $r^{q,d_1,d_2}_{c,n-\frac{1}{2},V}(x)$ be the residual computed at position x from the tiling $\mathcal{T}_{q,d_1,d_2}$ according to (4) and (8). We then assign the minimal value among the set of residuals obtained by the various scales and shifts to form the multi-scale and shift (MSS) residual

$$r^{\text{MSS}}_{c,n-\frac{1}{2},V}(x) = \min_{q,d_1,d_2} r^{q,d_1,d_2,V}_{c,n-\frac{1}{2}}(x). \tag{10}$$

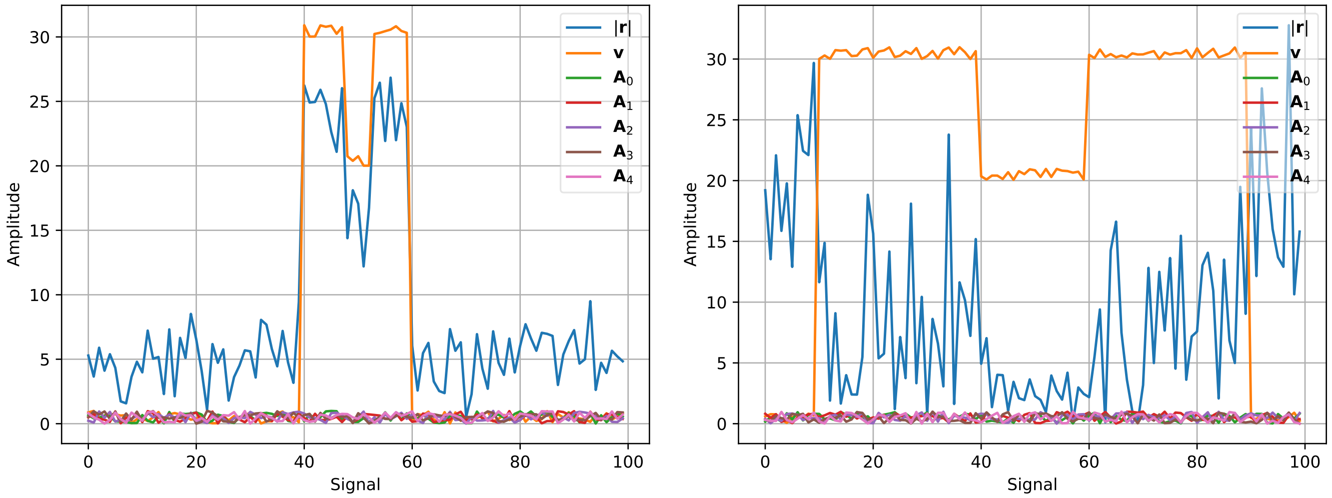

Figure 4: Effect of the global approach on the residual estimation. On the left, a profile where the vector $\mathbf{v}$ for which novelty is estimated is different on a short interval with width (20), relatively to the width (100) of the whole tile. The profile vectors $\mathbf{A}_i$ represent the ground at different times; they fluctuate so that their variations are of order ϵ around the average. The residual vector $|\mathbf{r}|$ located along the change part, i.e. in $[40; 60]$ has an amplitude of order δ while the ground, i.e. $[0; 40] \cap [60; 100]$, is of order $\eta < \delta$. On the right, when the change profile is large according to the width tile, the amplitude of the residual is no longer distinctive as $\eta \sim \delta$.

This leads to define the set of tilings $\mathfrak{T} = \{\mathcal{T}_{q,d_1,d_2}, \forall q_0 \leq q \leq q_{\max}, \forall d_1 \in D_1, \forall d_2 \in D_2\} \cup \mathcal{I}$ where $\mathcal{I}$ is the image tiling, q_0 is a parameter, $q_{\max}$ is such that $2^{q_{\max}} \leq \min(H, W)$, $D_1 = \{2^q i_1/S, \forall i_1 \in \{0, \dots, S-1\}\}$ and $D_2 = \{2^q i_2/S, \forall i_2 \in \{0, \dots, S-1\}\}$ with S the number of shifts. From the estimator formulation (10) we then define the hue-change and contrast-based versions which are respectively $r^{\text{MSS, hue}}_{i,n-\frac{1}{2},V}$ for $1 \leq i \leq C$ and $r^{\text{MSS, con}}_{c,n-\frac{1}{2},V}$ for $1 \leq c \leq C$.

3.5 Hypothesis Testing

The NFA Framework. For each pixel of Ω, and times $n = 2, \dots, N$ we consider an *a contrario* ground stochastic model

$$\mathcal{H}_0 : Y_{n,x} \sim f, \tag{11}$$

where $Y_{n,x}$ is the random variable associated to x at time n and f is the probability density function modeling the pixel behavior in the absence of change. In the case of a simple binary decision, hypothesis testing would lead to threshold the p-value defined by

$$\mathbb{P}(Y_{n,x} \geq y_n | \mathcal{H}_0) = 1 - F(y_n) \tag{12}$$

where y_n is the observed value of $Y_{n,x}$ and F the cumulative distribution function related to f. However, as this test is applied separately on each pixel, i.e. more than 100,000 times per image, chances that false rejections of $\mathcal{H}_0$ appear are huge. To overcome this situation by controlling the total number of false rejections several approaches exist (see [44, § 10.7]). Our method computes a *number of false alarms* (NFA)

$$\text{NFA}_{n-\frac{1}{2}}(x) = \#T \times \mathbb{P}(Y_{n,x} \geq y_n | \mathcal{H}_0) = \#T \times (1 - F(y_n)) \tag{13}$$

where $\#T$ is the number of tests; in the present case $\#T = \Omega$.

Law under the null hypothesis. In order to estimate in a robust way the changes, we consider independently the hue and contrast based estimators, then combine their observable residuals before applying the hypothesis testing (13). We therefore proceed as follows. First, for each channel c, we define the empirical probability density functions $\hat{f}_c$ related to the null hypothesis. Assuming that there is at least, for each x, a proportion Q of times where there is no change, then it is quite reasonable to assume that the Q smallest values of the residuals correspond to the stable states, while the great values to change states. Consequently, we consider that $\hat{f}_c$ can be represented by the set

$$\{Q((r_{c,n-\frac{1}{2},V}(x))_{c,2\leq n\leq N}), x \in \Omega\}. \tag{14}$$

Introducing the random variable $X_{c,n,x}$ associated with the observable value of the residual $r_{c,n}(x)$, we have, under the null hypothesis:

$$X_{c,n,x} \sim \hat{f}_c, \text{ for } c = 1, \dots, C. \tag{15}$$

As the shapes of the empirical cumulative distribution function $\hat{F}_c$ associated with the probability density functions $\hat{f}_c$ may vary according to c, we transform first the random variables into uniform variables by setting

$$Y_{c,n,x} = \hat{F}_c(X_{c,n,x}). \tag{16}$$

As a result the distributions of $Y_{c,n,x}$ are uniform and defined on $[0, 1]$.

Estimator combination. Among the different possible strategies to mix together the channels of the hue and contrast estimators one catches our attention. First, if one channel among the others presents a significant change, then this change should be highlighted whatever the residual values of the others. While the averaging of the channel residuals would not express a significant variation, the choice of the maximum would. Secondly, the hue and constrast based estimators play a complementary role for detecting at the same time the landscape and urban changes. Consequently, the choice of the maximum operator seems pertinent. These two reasons lead to define the formulation

$$Y_{n,x} = \max\{Y_{1,x}^{\text{hue}}, \dots, Y_{C,x}^{\text{hue}}, Y_{1,x}^{\text{con}}, \dots, Y_{C,x}^{\text{con}}\}. \tag{17}$$

Under the null hypothesis and according to (17), the random variable $Y_{n,x}$ follows the law whose associated cumulative distribution F is defined by

$$F(t) = t^{2C} \tag{18}$$

so that the final formulation of the NFA is

$$\text{NFA}_{n-\frac{1}{2}}(x) = |\Omega|(1 - Y_{n,x}^{2C}). \tag{19}$$

Once the NFA series is computed, the meaningful changes are detected thanks to a threshold ε which becomes our decision parameter; for commodity we use the log of the NFA. Finally the Boolean detection masks $(B_{n-\frac{1}{2}})_{2\leq n\leq N}$ are defined by

$$B_{n-\frac{1}{2}}(x) = \mathbf{1}\{\log \text{NFA}_{n-\frac{1}{2}}(x) \leq \log \varepsilon\}, \tag{20}$$

where $\mathbf{1}$ is the indicator function.

3.6 Morphological Post-Processing

A number of false or no detection areas may appear spatially isolated in the Boolean maps $(B_{n-\frac{1}{2}})_{2\leq n\leq N}$ defined by Equation (20). These are generally detection errors that can be easily eliminated by considering a surface threshold $\alpha \leq |\Omega|$ consistent with the image resolution and the minimal acceptable surface of change. Denoting $\mathcal{R}$ the set of all connected regions ω of image $B_{n-\frac{1}{2}}$, we fill the undesirable holes according to

$$\forall \omega \in \mathcal{R}, \forall \mathbf{x} \in \omega, \tilde{B}_{n-\frac{1}{2}}(\mathbf{x}) = \begin{cases} 1 - B_{n-\frac{1}{2}}(\mathbf{x}) \text{ if } |\omega| < \alpha, \\ B_{n-\frac{1}{2}}(\mathbf{x}) \text{ otherwise.} \end{cases} \tag{21}$$

This task is performed thanks to the Boolean image partitioning approach of Tarjan and Endre [39]. Numerically it is computed by the module `ccproc` of the `imscript` library[2].

V	The number of vectorized images of the time series.
q_0	The exponent of the minimal dyadic square patch of side 2^{q_0}.
S	The number of shifts in each direction.
Q	The quantile of sample used to form the empirical law.
ε	The NFA threshold.
α	The surface threshold of the change detection post-processing.

Table 2: Table summarizing the parameters of the proposed method.

3.7 Change Classification

The classification of changes consists in attributing a duration label to each of the related regions ω defined as a connected set of pixels x where $\tilde{B}_{n_0-\frac{1}{2}}(x) = 1$ at time n_0. To realize this classification we compare the values of the pixels of ω at time n_0 with their past and future values $(\mathbf{u}^{\text{raw}}_{c,n})^{1 \le c \le C}_{1 \le n \le N}$. We first introduce a metric based on the zero-normalized cross-correlation (ZNCC)

$$\kappa_\omega(m,n) = \frac{1}{C|\omega|} \sum_{c=1}^{C} \sum_{x \in \omega} \frac{(u_{c,m}(x) - \bar{u}_{c,m})(u_{c,n}(x) - \bar{u}_{c,n})}{\sigma_{u_{c,m}} \sigma_{u_{c,n}}}, \tag{22}$$

where $\bar{u}_{c,m}$, $\bar{u}_{c,n}$ are the average values computed on the domain ω. We then test the times n for $n_0 < n \le N$ while

$$\kappa_\omega(n_0, n) \ge \vartheta. \tag{23}$$

In this study, the threshold ϑ is set to $\frac{1}{2}$ which is a rough but intuitive value. We make no difference between the negative values of $\kappa_\omega(m,n)$ and the zero values because they carry the same idea, namely a great dissimilarity between the surfaces ω at times m and n. This method then gives us the duration of each stable texture. We show the associated color code in Figure 5.

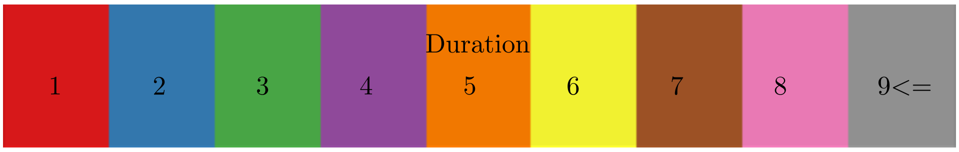

Figure 5: Color code of the estimated duration appearance of the surface ω.

---

[2] https://github.com/mnhrdt/imscript

Algorithm 1: Computation of the change detection NFA time series.

```
Input: $(\mathbf{v}_{c,n})_{1\leq n\leq N}^{1\leq c\leq C}$ the multi-spectral time series of $N \geq 3$ images, composed by $C$ channels and defined
       over the discrete domain $\Omega = [\![1, H]\!], \times [\![1, W]\!]$.
Input: $V$ the number of vector images composing the basis.
Input: $q_0$ The exponent of the minimal dyadic square patch of side $2^{q_0}$.
Input: $S$ The number of shifts in each direction.
Input: $Q$ The quantile of sample used to form the empirical law.
Input: $\log \varepsilon$ the log NFA threshold.
Input: $\alpha$ the post-processing connected component area threshold.
Output: $(\tilde{B}_{n-\frac{1}{2}})_{2\leq n\leq N}$ the change detection Boolean series.
// PREPROCESSING
for $c = 1, \dots, C$ do
    for $n = 1, \dots, N$ do
        Pseudo gamma correction
        $\mathbf{v}_{c,n}^{\gamma} \leftarrow \sqrt{\mathbf{v}_{c,n}}$
        Pseudo YUV transformation
        $\mathbf{v}_{c,n}^{\text{hue}} \leftarrow$ according to (5) $[\mathbf{v}_{c,n}^{\gamma})]$
        Renaming
        $\mathbf{v}_{c,n}^{\text{con}} \leftarrow \mathbf{v}_{c,n}^{\gamma}$
// COMPUTATION OF THE RESIDUAL SERIES
Computation of the tile set
$\mathfrak{T} \leftarrow$ according to (9)$[q_0, S]$
for $c = 1, \dots, C$ do
    for $n = 2, \dots, N$ do
        Computation of the estimators per tile
        for $\mathcal{T} \in \mathfrak{T}$ do
            for $x = 1, \dots, |\Omega|$ do
                $r_{c,n}^{\text{hue},q,d_1,d_2}(x) \leftarrow$ according to (4) $[\mathbf{v}_{c,n}^{\text{hue}}, V]$
                $r_{c,n}^{\text{con},q,d_1,d_2}(x) \leftarrow$ according to (8)$[\mathbf{v}_{c,n}^{\text{con}}, V]$
        Computation of the global estimators (10)
        $r_{c,n-\frac{1}{2},V}^{\text{MSS, hue}}(x) \leftarrow \min_{q,d_1,d_2} r_{c,n}^{\text{hue},q,d_1,d_2}(x)$
        $r_{c,n-\frac{1}{2},V}^{\text{MSS, con}}(x) \leftarrow \min_{q,d_1,d_2} r_{c,n}^{\text{con},q,d_1,d_2}(x)$
// HYPOTHESIS TESTING
for $c = 1, \dots, C$ do
    Computation of the empirical $H_0$ law
    $\hat{f}_c$ according to $Q$ and (14)
    Uniformization of the random variables
    for $n = 2, \dots, N$ do
        for $x = 1, \dots, |\Omega|$ do
            for type=hue, con do
                $y_{c,n}^{\text{type}} \leftarrow \hat{F}_c(r_{c,n-\frac{1}{2},V}^{\text{MSS, type}}(x))$ according to (16)
for $n = 2, \dots, N$ do
    for $x = 1, \dots, |\Omega|$ do
        Computation of the combined estimator
        $y_{n,x} \leftarrow \max\{y_{1,x}^{\text{hue}}, \dots, y_{C,x}^{\text{hue}}, y_{1,x}^{\text{con}}, \dots, y_{C,x}^{\text{con}}\}$
        Computation of the NFA map series
        $\text{NFA}_{n-\frac{1}{2}}(x) \leftarrow |\Omega| \times (1 - y_{n,x}^{2C})$
        Computation of the unfiltered Boolean map series
        $B_{n-\frac{1}{2}}(x) \leftarrow \mathbf{1}\{\log \text{NFA}_{n-\frac{1}{2}}(x) \leq \log \varepsilon\}$
        Computation of the filtered Boolean map series
        $\tilde{B}_{n-\frac{1}{2}}(x) \leftarrow$ according to (21)$[B_{n-\frac{1}{2}}, \alpha]$
return $\tilde{B}$
```

3.8 Time Complexity Analysis

To simplify our estimation, we restrict our time complexity analysis to the trivial case where there is no tiling split, i.e. $\mathcal{T} = \mathcal{I}$. The computation of the Boolean series defined by Equation (21) requires four steps. The time complexity of the first step, namely the luminance centering is in $O(C(N-1)|\Omega|\log|\Omega|)$. The second step is mainly based on the non negative linear square (NNLS) method by Lawson and Hanson [27, chap. 23]. It is difficult to give a precise time complexity of this part because theoretically, the NNLS requires a QR matrix decomposition in $O(|\Omega|^3)$ at each of its iterations, and would need L^2 iterations for an input matrix $\mathbf{A} \in \mathcal{M}_{|\Omega|,L}(\mathbb{R})$. However, the authors observed that the NNLS converges on average in $L/2$ iterations, and they carry out the QR decomposition only once, recomputing partially some terms over the iterations. As NNLS is computed according to V at each time n and channel c, we can roughly estimate the complexity of this step to be in $O(C(N-1)(|\Omega|^3 + V/2))$. The third step is the hypothesis testing whose complexity is globally $O(C(N-1)|\Omega|)$. The fourth step computes the NFA maps in $O((N-1)|\Omega|C\log C)$, their binarization in $O((N-1)|\Omega|)$ and the region removal takes constant time (see [39]), then in $O((N-1)|\Omega|)$. Finally, the overall Boolean masks creation has complexity in $O(CN(|\Omega|^3 + |\Omega|\log|\Omega| + |\Omega|\log C))$ and for $N \ll |\Omega|$ is dominated by $O(CN|\Omega|^3)$.

4 Experiments

4.1 Our Testing Dataset

We built a dataset of 28 series of 20 images each, of size 480 × 480, composed of the four bands red (B04), green (B03), blue (B02) and near-infrared (B08) from Sentinel-2 data. The series represent different locations over the world between 01/01/2016 and 01/01/2020. Some of the locations coincide with those of the dataset built by Daudt et al. [8]. Images were subpixel registered, and are cloudless and shadow free. In Appendix B we give an overview of the scenes (Figure 12) as well as the geographic coordinates of the image centers before the subpixel registration (Table 3).

Since the notion of change can have different semantic levels (pixel difference, texture or object based, etc.) depending on the objectives to be achieved, this database does not contain ground truth.

4.2 Qualitative Results

We compared our detection method with the MOSUM approach of Eichinger and Kirch [18]. This algorithm can be found as a `Gnu R` package[3]. As the implemented version does not take into account the image colors and processes each pixel independently as a univariate time series, we proceed as follows. We first split the image channels red, blue and green into three gray-level series, from which we compute three change map series. We finally merge these series with a median filter over the three channels. We used the method MOSUM as it was, applying, if possible, the parameter values indicated in the papers. In order to compare only the raw temporal change detectors, we did not apply our morphological post-processing. The results are shown in Figures 6 and 7. On Figure 6 we present the five first images of the Al Wakrah time series (300 × 300 cropped areas). At times 2, 3 and 4, our method detects the real important changes, in particular at the image center where several building were demolished and in the highway region. The MOSUM is totally blind at times 2, 3 and under-detects strongly at times 4 and 5. The same behaviour of both algorithms can be observed on Figure 7 with the Las Vegas time series. At times 2 and 4, our method detects the most relevant changes while MOSUM detects nothing. Moreover, at time 5, MOSUM detects doubtful changes (bottom left).

---

[3] `https://CRAN.R-project.org/package=mosum`

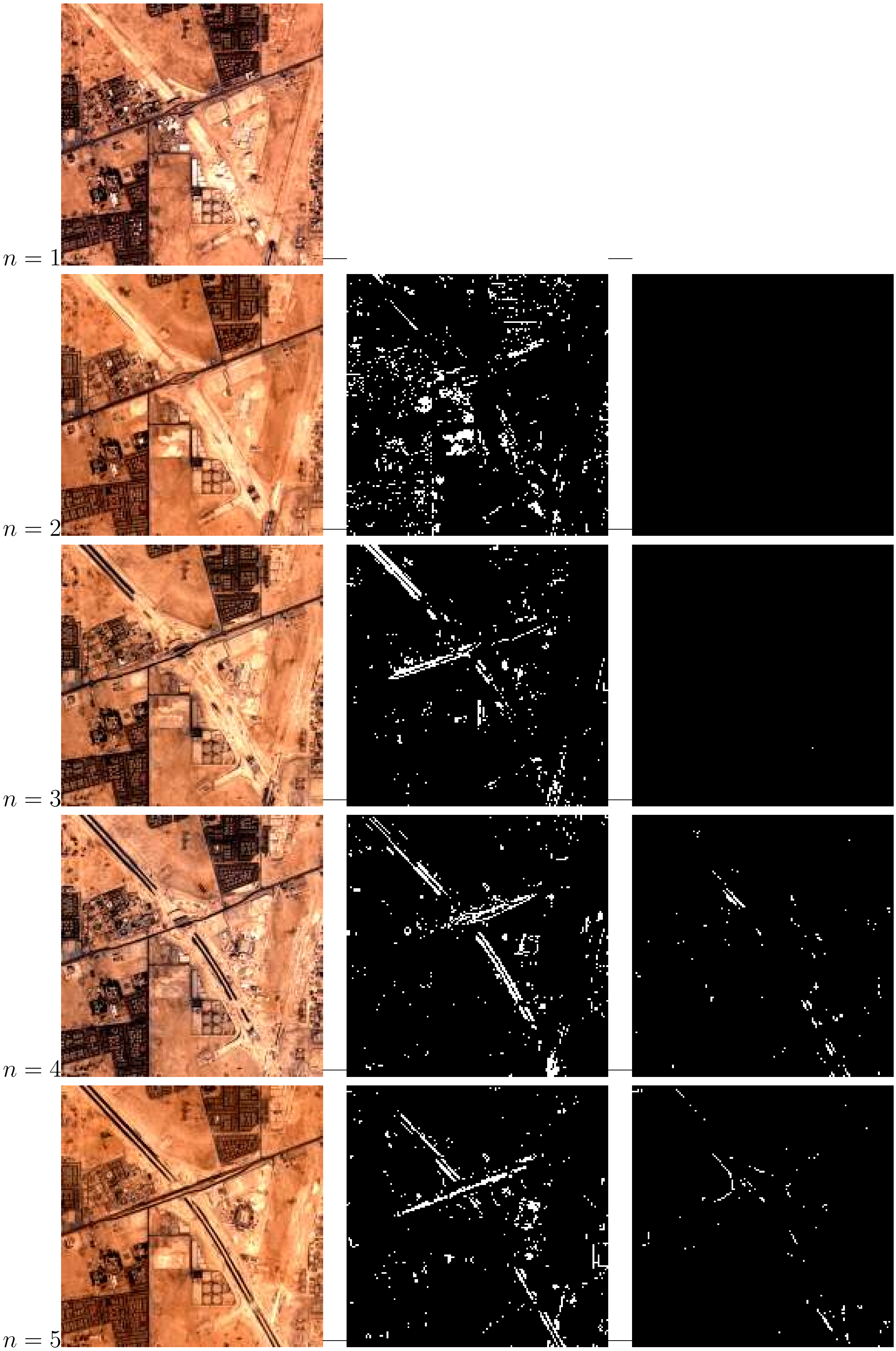

Figure 6: Comparison between our method (second column) and MOSUM (third column) on the five first images of the Al Wakrah time series (300×300 cropped areas). For our method, no morphological post-processing was applied, $V = 5$, $\mathcal{T} = \mathcal{I}, S = 2, Q_{50}$ and $\log \varepsilon = 1$. For the MOSUM $G_{\text{MOSUM}} = 0.2$ and $\alpha_{\text{MOSUM}} = 0.001$ (the empirical best value found).

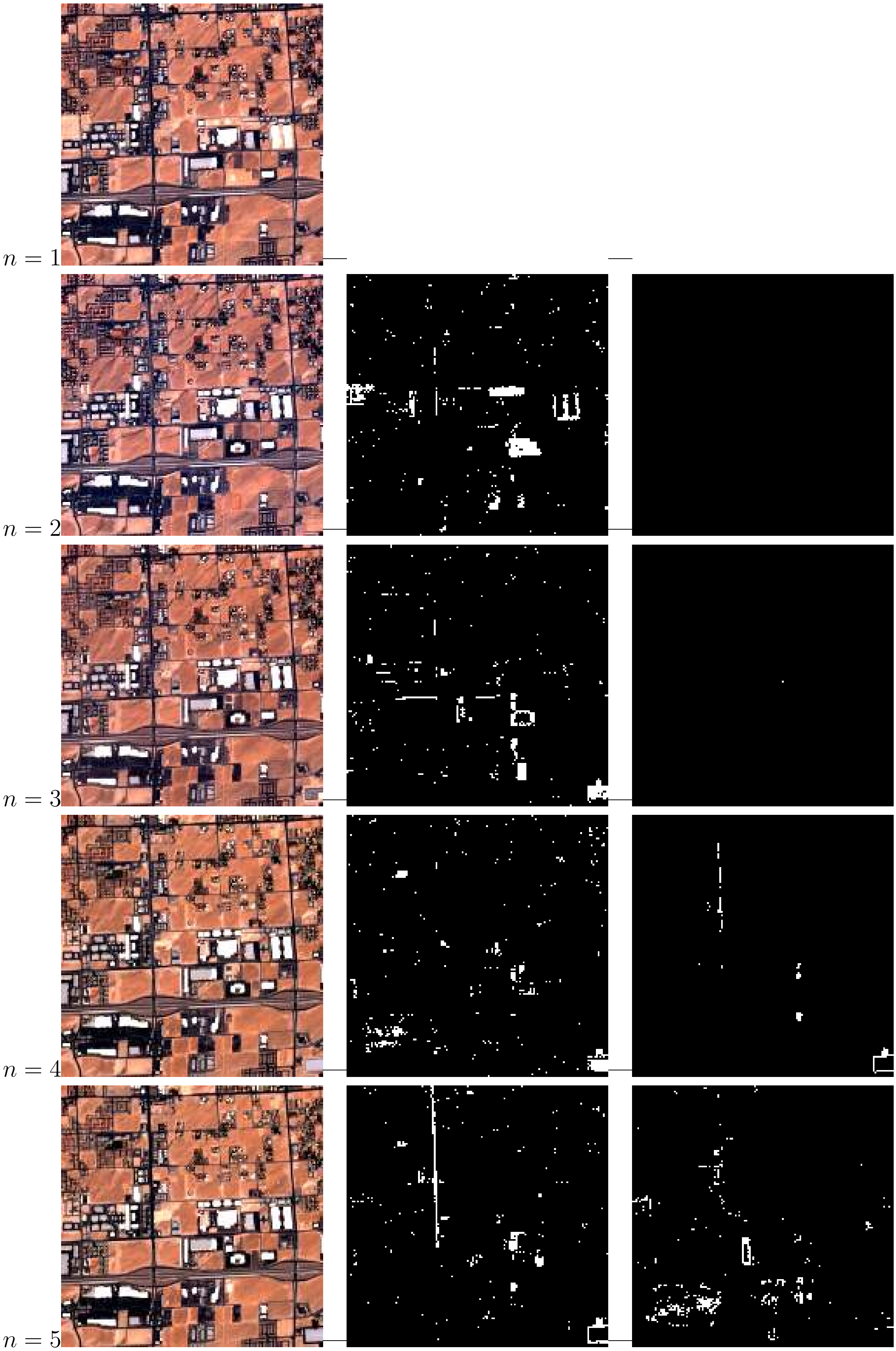

Figure 7: Comparison between our method (second column) and MOSUM (third column) on the five first images of the Las Vegas time series (300×300 cropped areas). For our method, no morphological post-processing was applied, $V = 5$, $\mathcal{T} = \mathcal{I}, S = 2, Q_{50}$ and $\log \varepsilon = 1$. For the MOSUM $G_{\text{MOSUM}} = 0.2$ and $\alpha_{\text{MOSUM}} = 0.001$ (the empirical best value found).

We also illustrate the efficiency of the multi-scale patch residual strategy discussed in Section 3.4 in Figure 8. It shows the Chaco time series at times 6 to 10 where we have to deal with extreme cases where changes are global and very frequent. When the changes are spatially global and also the tiling i.e. $\mathcal{T} = \mathcal{I}$ (middle column), some no change regions are considered as such: see the two central rectangles at times 8 and 9. When the domain is tiled with 64×64 patches, the resulting log NFA maps (right columns) estimate best the unchanged regions.

We present the influence of the number V of vectors composing the basis on Figure 11 on the 9^{th} frame of the Las Vegas time series. We choose this frame to avoid the influence of the frame boundary replication. We observe that globally the number of change surfaces decrease when the number of vectors increase. We can explain this by the fact that the more we introduce basis vectors, the more their linear combination fits with the tested frame. The residual values would decrease both for real changes and for noisy variations. Consequently, if increasing the number of basis vectors is a good thing to remove the noise changes, increasing it too much would attenuate the detection of the real changes which are weakly stable. Only the stable changes would stay observable.

Concerning the change duration estimation, we show examples on Figures 9 and 10. Figure 9 deals with the five first images of the Piraeus time series (300×300 cropped areas). The durations are globally well estimated; some duration errors can appear when two ships are merged into the same region. In absence of a reference image, the return to the initial state (i.e. the sea) is considered as a change itself, what we can observe with the grey region at time $n = 4$. Figure 10 shows the evolutions of the Las Vegas time series (300×300 cropped areas) at times 1 to 5. The results are coherent with the visual observations: most of the buildings are definitive so that their color are grey while transitory constructions are in red.

5 Conclusion

In this work we first proposed to detect changes in satellite time series thanks to a statistics defined by the novelty filter and applied on the full sequence. This statistics is composed by a backward and a forward novelty term which summarizes the temporal information of the whole time series. Significant changes are then detected by hypothesis testing. The study of the influence of the time series length and qualitative results lead to conclude that the approach is efficient for all scenes when handling at least ten images, provided the changes are spatially sparse in relation to the size of the tiles. We secondly proposed to classify the changes according to their duration, thanks to a metric based on the zero-normalized cross-correlation. The results are also coherent with the visual observations. The last innovation of this paper consists in the creation of an open dataset of time series fairly representative of the regions of the world, subpixel registered and not containing atmospheric alterations.

Future work will focus on the extension of the method to sequences composed of images acquired by satellites with different spectra and resolutions.

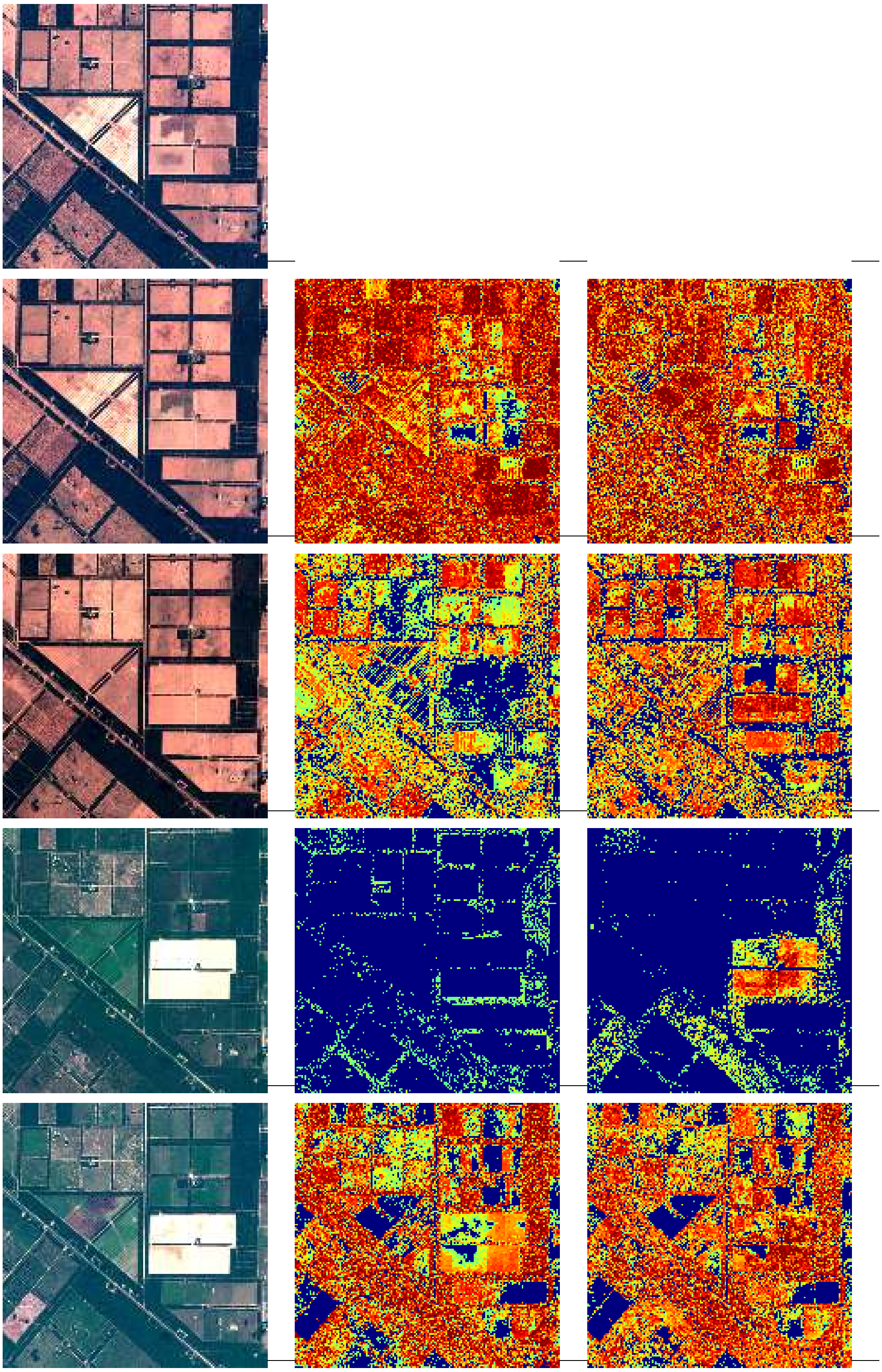

Figure 8: Example of the Chaco time series at times 6 to 10 where the log NFA maps represent the changes in blue tones. We used $V = 3$, $\mathcal{T} = \mathcal{I}, S = 2, Q_{25}$ (middle column) and $V = 3$, $\mathcal{T}_{64}, S = 2, Q_{25}$ (right column).

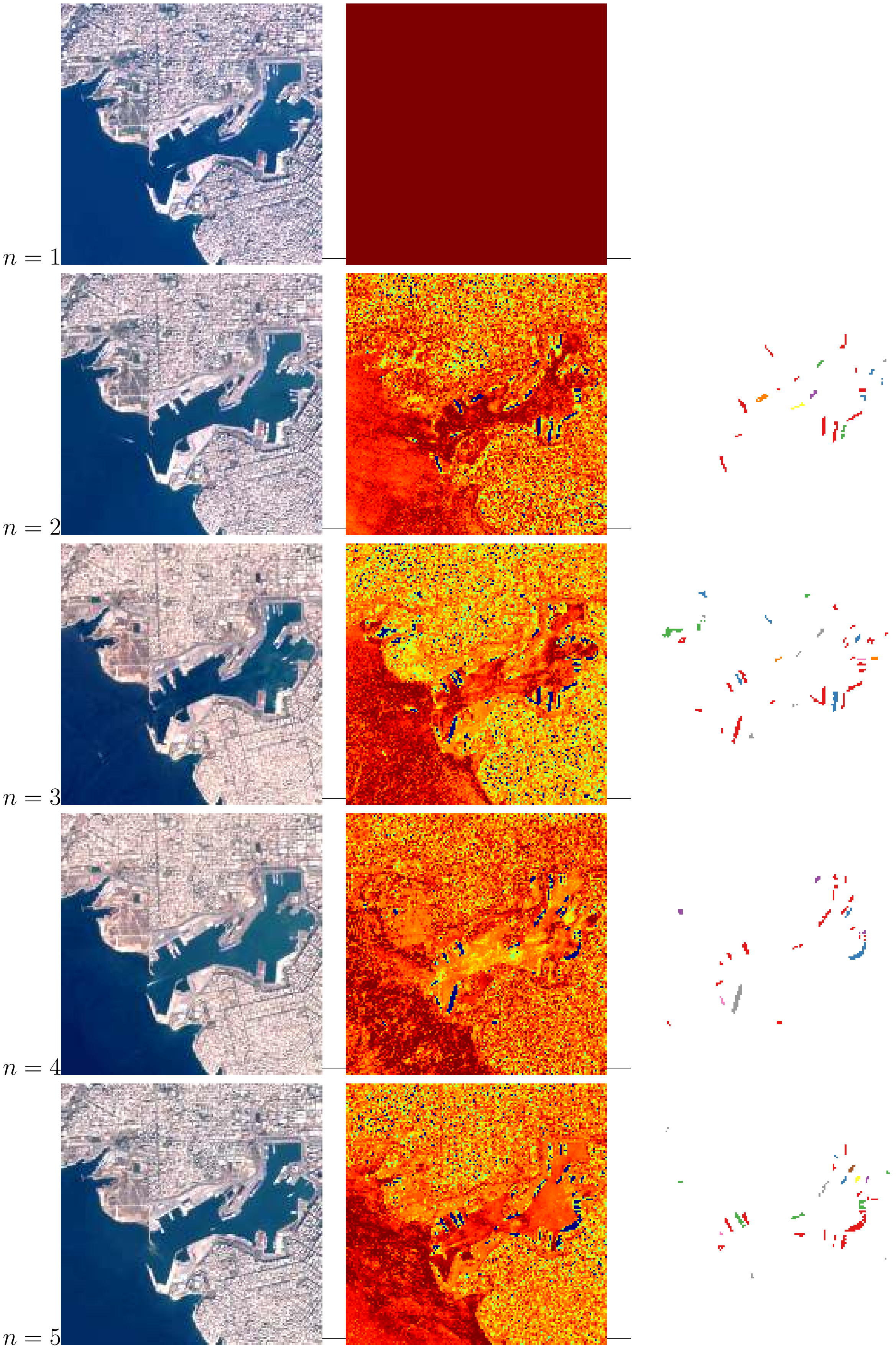

Figure 9: Log NFA maps (second column) and change maps (third column) obtained on the five first images of the Piraeus time series (300×300 cropped areas) according to the color code of the Figure 5 with the parameters $V = 5$, $\mathcal{T} = \mathcal{I}, S = 2, Q_{25}$ and $\log \varepsilon = 1$.

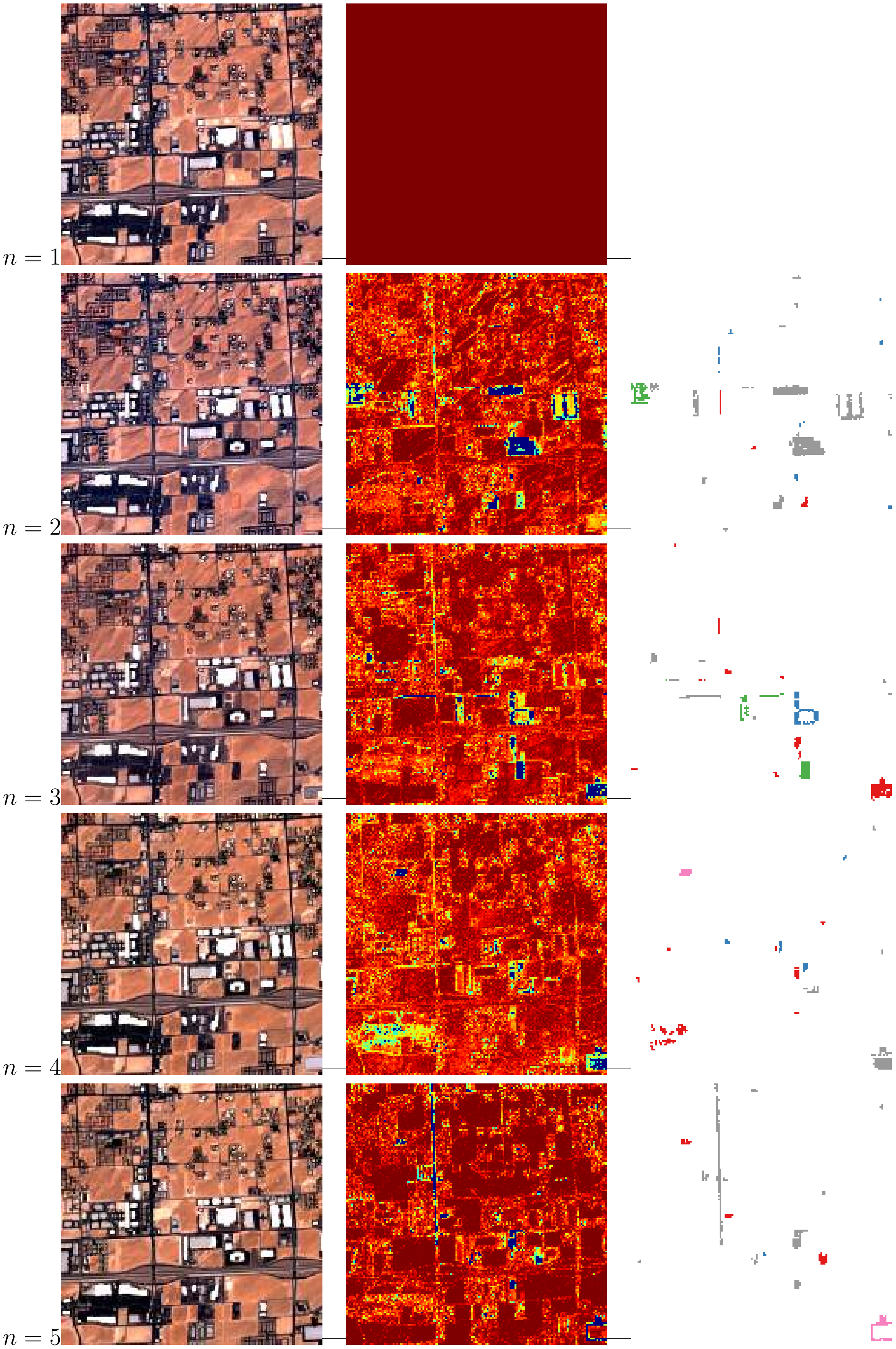

Figure 10: Log NFA maps (second column) and change maps (third column) obtained on the five first images of the Las Vegas time series (300×300 cropped areas) according to the color code of the Figure 5 with the parameters $V = 5$, $\mathcal{T} = \mathcal{I}, S = 2, Q_{25}$ and $\log \varepsilon = 1$.

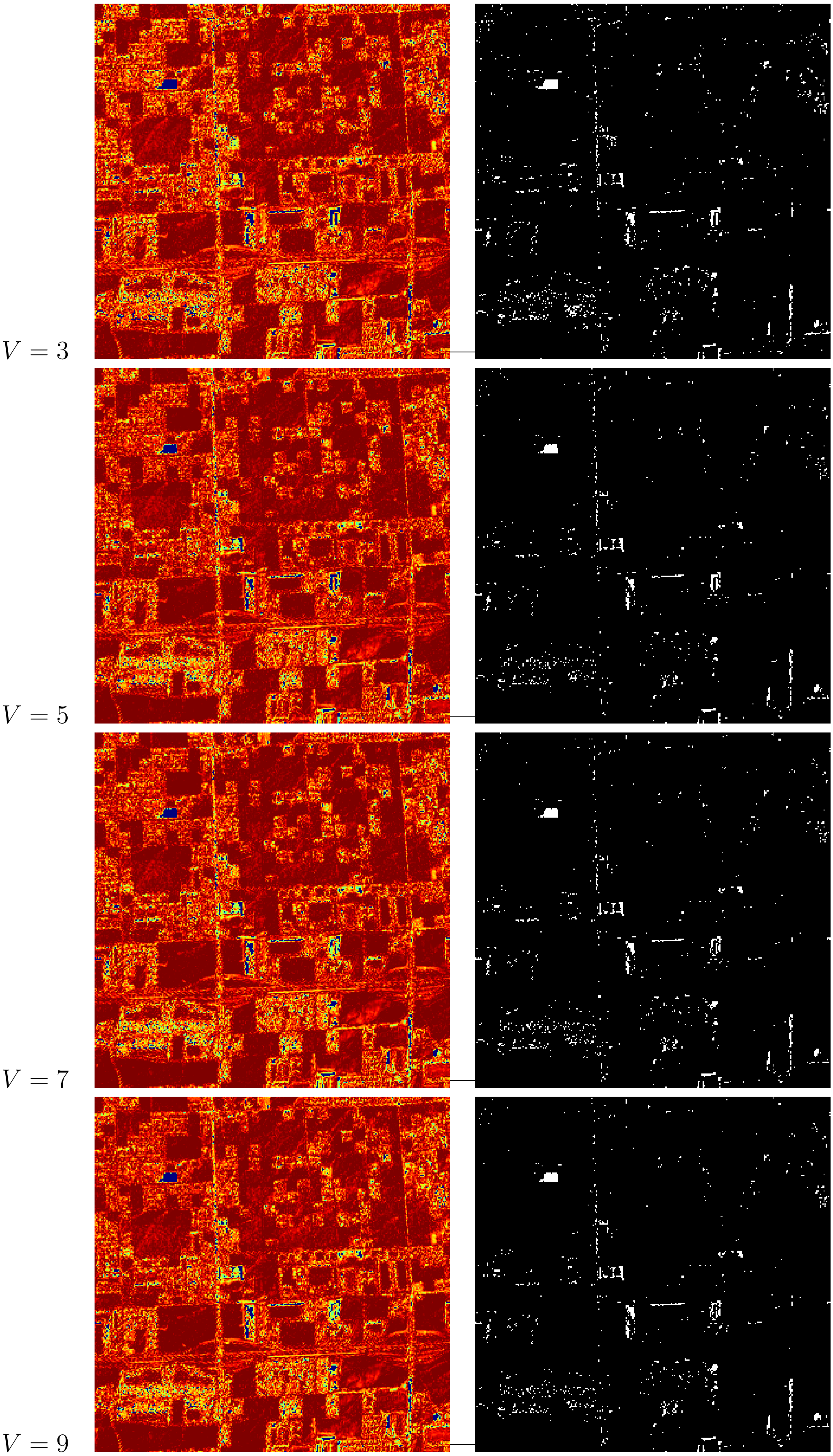

Figure 11: Log NFA maps (left column) and change maps (right column) obtained on the 9th frame of the Las Vegas time series (300×300 cropped areas) for different number of basis vectors V. To highlight the change detected surfaces, no morphological post-processing was applied, $\mathcal{T}_6$ (64×64 tiles), $S = 2, Q_{50}$ and $\log \varepsilon = 1$.

A Synthetic Dataset

To observe more easily the influence of the various parameters, in particular the hue-change and contrast estimators, we added in the IPOL demonstrator some toy examples made of synthetic examples. These images do not contain infrared bands and have values in $[0, 255]$: one should neither apply the gamma correction nor the normalization.

The sets named A to H are composed by images of size 512×512 containing 64 time varying patches. The set A is made of grayscale images while the set B contains luminance variations with the same hue. In both cases, the multi-scale strategy, with the 64 choice and no shift, should fail to detect if the hue-change estimator is selected solely and succeed if the contrast-based estimator is. The set C contains hue variations with the same luminance: the hue-change estimator should detect these variations better than that of contrast. The set D is intended to test the variation of patch size in the multi-scale strategy. The set E contains variable size changes with some shifts. The sets F, G and H have various change durations to check the classification.

B Real Dataset

We show on Figure 12 a frame representing each time series of our real RGBI (bands 2, 3, 4 and 8) Sentinel-2 dataset.

Figure 12: First images of the 28 time series composing the dataset.

Location	Longitude	Latitude	Location	Longitude	Latitude
Abudhabi	54°34′38.03″ E	24°19′44.22″N	Livermore	121°42′52.54″W	37°40′38.96″N
Al Wakrah	51°34′38.41″ E	25°10′04.64″N	Ljubljana	14°33′06.31″ E	46°03′43.72″N
Austin	97°52′40.40″W	30°13′41.70″N	Los Paseos	121°46′31.10″W	37°13′52.22″N
Beijing airport	116°35′51.38″ E	40°04′35.87″N	Montevideo	56°16′31.05″W	34°51′08.17″ S
Cachan	02°19′42.80″ E	48°47′49.94″N	Morandi	08°53′18.84″ E	44°25′33.42″N
Cape Town	18°26′36.86″ E	34°04′19.89″ S	Musashino	139°31′28.10″ E	35°39′55.76″N
Chaco	60°18′22.83″W	22°17′35.30″ S	Phoenix	112°12′28.67″W	33°33′30.32″N
Cupertino	122°00′48.71″W	37°20′57.84″N	Piraeus	23°37′54.44″ E	37°56′49.47″N
Cushing	102°21′20.42″W	32°03′05.63″N	Pohang	129°22′49.44″ E	35°57′52.92″N
Darwin	130°54′41.49″ E	12°25′32.41″ S	Rennes	01°42′13.84″W	48°05′20.57″N
Dublin	06°19′57.41″W	53°23′53.08″N	Saclay	02°10′28.22″ E	48°43′30.82″N
Karalee	152°48′23.35″ E	27°34′46.60″ S	Sanshui	112°52′21.49″ E	23°12′00.84″N
Koto	139°48′54.10″ E	35°39′03.05″N	Sant Jordi	02°09′02.70″ E	41°21′48.01″N
Las Vegas	115°13′37.27″W	36°04′49.41″N	Tuzla	36°34′16.70″ E	45°15′03.90″N

Table 3: Geographic coordinates of the 28 locations (image centers) constituting the dataset.

Image Credits

All images in this manuscript (in particular Figures 1, 6, 7, 8, 9, 10 and 12) were produced by the authors (license CC-BY-SA) except the scheme of Figure 3 (license CC-BY-SA 3.0).

References

[1] M. Basseville and I. Nikiforov, *Detection of abrupt changes. Theory and application*, vol. 104, 1993. ISBN 0-13-126780-9, `https://doi.org/10.1016/0967-0661(94)90196-1`.

[2] E.B. Brooks, R.H. Wynne, V.A. Thomas, C.E. Blinn, and J.W. Coulston, *On-the-Fly Massively Multitemporal Change Detection Using Statistical Quality Control Charts and Landsat Data*, IEEE Transactions on Geoscience and Remote Sensing, 52 (2014), pp. 3316–3332. `https://doi.org/10.1109/TGRS.2013.2272545`.

[3] R.L. Brown, J. Durbin, and J.M. Evans, *Techniques for testing the constancy of regression relationships over time*, Journal of the Royal Statistical Society. Series B (Methodological), 37 (1975), pp. 149–192. `http://www.jstor.org/stable/2984889`.

[4] E.L. Bullock, C.E. Woodcock, and C.E. Holden, *Improved change monitoring using an ensemble of time series algorithms*, Remote Sensing of Environment, 238 (2020), p. 111165. `https://doi.org/10.1016/j.rse.2019.04.018`.

[5] S. Cai and D. Liu, *Detecting change dates from dense satellite time series using a sub-annual change detection algorithm*, Remote Sensing, 7 (2015), pp. 8705–8727. `https://doi.org/10.3390/rs70708705`.

[6] E. Carlstein, *Nonparametric change-point estimation*, Annals of Statistics, 16 (1988), pp. 188–197. `https://doi.org/10.1214/aos/1176350699`.

[7] R. Caye Daudt, B. Le Saux, and A. Boulch, *Fully convolutional siamese networks for change detection*, in IEEE International Conference on Image Processing (ICIP), 2018, pp. 4063–4067. `https://doi.org/10.1109/ICIP.2018.8451652`.

[8] R. Caye Daudt, B. Le Saux, A. Boulch, and Y. Gousseau, *Urban change detection for multispectral earth observation using convolutional neural networks*, in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2018. `https://doi.org/10.1109/IGARSS.2018.8518015`.

[9] T. Celik, *Unsupervised change detection in satellite images using principal component analysis and k-means clustering*, IEEE Geoscience and Remote Sensing Letters, 6 (2009), pp. 772–776. `https://doi.org/10.1109/LGRS.2009.2025059`.

[10] H. Chen, C. Wu, B. Du, and L. Zhang, *Deep siamese multi-scale convolutional network for change detection in multi-temporal VHR images*, in International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), 2019, pp. 1–4. `https://doi.org/10.1109/Multi-Temp.2019.8866947`.

[11] G.C. Chow, *Tests of equality between sets of coefficients in two linear regressions*, Econometrica, 28 (1960), pp. 591–605. `http://www.jstor.org/stable/1910133`.

[12] C-S.J. Chu, K. Hornik, and C-M. Kuan, *MOSUM Tests for Parameter Constancy*, Biometrika, 82 (1995), pp. 603–617. `https://doi.org/10.2307/2337537`.

[13] T. Dagobert, C. Hessel, J. Anger, and T. Ehret, *Multiscale a contrario change detection on Planetscope timeseries*, in IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR), 2021.

[14] T. Dagobert, R. Grompone Von Gioi, C. Hessel, J.-M. Morel, and C. De Franchis, *Change analysis in registered satellite image time series*, in 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021, pp. 4360–4363.

[15] C. Deng and Z. Zhu, *Continuous subpixel monitoring of urban impervious surface using Landsat time series*, Remote Sensing of Environment, 238 (2020), p. 110929. `https://doi.org/10.1016/j.rse.2018.10.011`.

[16] A. Desolneux, L. Moisan, and J.-M. Morel, *A grouping principle and four applications*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25 (2003), pp. 508–513. `https://doi.org/10.1109/TPAMI.2003.1190576`.

[17] 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, 57 (2019), pp. 9976–9992. `https://doi.org/10.1109/TGRS.2019.2930682`.

[18] B. Eichinger and C. Kirch, *A MOSUM procedure for the estimation of multiple random change points*, Bernoulli, 24 (2018), pp. 526 – 564. `https://doi.org/10.3150/16-BEJ887`.

[19] M.D. Fairchild, *Color Appearance Models*, John Wiley and Sons Ltd, 2005. ISBN 9781119967033, `https://doi.org/10.1002/9781118653128`.

[20] E. Hamunyela, J. Verbesselt, and M. Herold, *Using spatial context to improve early detection of deforestation from Landsat time series*, Remote Sensing of Environment, 172 (2016), pp. 126 – 138. `https://doi.org/10.1016/j.rse.2015.11.006`.

[21] T. Hermosilla, M.A. Wulder, J.C. White, N.C. Coops, and G.W. Hobart, *An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites*, Remote Sensing of Environment, 158 (2015), pp. 220 – 234. `https://doi.org/10.1016/j.rse.2014.11.005`.

[22] C. Hessel, C. De Franchis, G. Facciolo, and J-M. Morel, *A global registration method for satellite image series*, in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2021, pp. 3121–3124. https://doi.org/10.1109/IGARSS47720.2021.9554786.

[23] S. Ji, Y. Shen, M. Lu, and Y. Zhang, *Building instance change detection from large-scale aerial images using convolutional neural networks and simulated samples*, Remote Sensing, 11 (2019). https://doi.org/10.3390/rs11111343.

[24] G.H. Joblove and D. Greenberg, *Color spaces for computer graphics*, ACM SIGGRAPH Computer Graphics, 12 (1978), p. 2025. https://doi.org/10.1145/965139.807362.

[25] T. Kohonen and E. Oja, *Fast adaptive formation of orthogonalizing filters and associative memory in recurrent networks of neuron-like elements*, Biological Cybernetics, 21 (1976), pp. 85–95. https://doi.org/10.1007/BF01259390.

[26] M. Lavielle and G. Teyssière, *Detection of multiple change-points in multivariate time series*, Lithuanian Mathematical Journal, (2006). https://doi.org/10.1007/s10986-006-0028-9.

[27] C. Lawson and R. Hanson, *Solving Least Squares Problems*, Society for Industrial and Applied Mathematics, 1995. ISBN 9780898713565, https://doi.org/10.1137/1.9781611971217.

[28] J.L. Lisani and J-M. Morel, *Detection of major changes in satellite images*, in IEEE International Conference on Image Processing (ICIP), 2003, pp. I–941. https://doi.org/10.1109/ICIP.2003.1247119.

[29] C. Liu, Q. Zhang, H. Luo, S. Qi, S. Tao, H. Xu, and Y. Yao, *An efficient approach to capture continuous impervious surface dynamics using spatial-temporal rules and dense Landsat time series stacks*, Remote Sensing of Environment, 229 (2019), pp. 114 – 132. https://doi.org/10.1016/j.rse.2019.04.025.

[30] G. Liu, Y. Gousseau, and F. Tupin, *A contrario comparison of local descriptors for change detection in very high spatial resolution satellite images of urban areas*, IEEE Transactions on Geoscience and Remote Sensing, 57 (2019), pp. 3904–3918. https://doi.org/10.1109/TGRS.2018.2888985.

[31] H. Luo, C. Liu, C. Wu, and X. Guo, *Urban Change Detection Based on Dempster–Shafer Theory for Multitemporal Very High-Resolution Imagery*, Remote Sensing, 10 (2018), p. 980. https://doi.org/10.3390/rs10070980.

[32] Stephen Marsland, *Novelty Detection in Learning Systems*, Neural Computing Surveys, 3 (2003).

[33] D.S. Matteson and N.A. James, *A nonparametric approach for multiple change point analysis of multivariate data*, Journal of the American Statistical Association, 109 (2014), pp. 334–345. https://doi.org/10.1080/01621459.2013.849605.

[34] A.A. Nielsen, K. Conradsen, and J.J. Simpson, *Multivariate Alteration Detection (MAD) and MAF Postprocessing in Multispectral, Bitemporal Image Data: New Approaches to Change Detection Studies*, Remote Sensing of Environment, 64 (1998), pp. 1 – 19. https://doi.org/10.1016/S0034-4257(97)00162-4.

[35] M. Papadomanolaki, S. Verma, M. Vakalopoulou, S. Gupta, and K. Karantzalos, *Detecting Urban Changes with Recurrent Neural Networks from Multitemporal Sentinel-2 Data*, in IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2019, pp. 214–217. https://doi.org/10.1109/IGARSS.2019.8900330.

[36] D. Peng, Y. Zhang, and H. Guan, *End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet++*, Remote Sensing, 11 (2019). https://doi.org/10.3390/rs11111382.

[37] C. Ren, X. Wang, J. Gao, and H. Chen, *Unsupervised change detection in satellite images with generative adversarial network*, 2020. arXiv preprint, https://doi.org/10.48550/arXiv.2009.03630.

[38] A. Schneider, *Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach*, Remote Sensing of Environment, 124 (2012), pp. 689 – 704. https://doi.org/10.1016/j.rse.2012.06.006.

[39] R.E. Tarjan, *Efficiency of a good but not linear set union algorithm*, Journal of the ACM, 22 (1975), pp. 215–225. https://doi.org/10.1145/321879.321884.

[40] A.P. Tewkesbury, A.J. Comber, N.J. Tate, A. Lamb, and P.F. Fisher, *A critical synthesis of remotely sensed optical image change detection techniques*, Remote Sensing of Environment, 160 (2015), pp. 1–14. https://doi.org/10.1016/j.rse.2015.01.006.

[41] Harikrishnan V. and A. Paulose, *A deep learning kernel k-means method for change detection in satellite image*, in International Journal of Science and Research (IJSR), vol. 3, 2014, pp. 1220 – 1226.

[42] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, *Detecting trend and seasonal changes in satellite image time series*, Remote Sensing of Environment, 114 (2010), pp. 106–115. https://doi.org/10.1016/j.rse.2009.08.014.

[43] J. Verbesselt, A. Zeileis, and M. Herold, *Near real-time disturbance detection using satellite image time series*, Remote Sensing of Environment, 123 (2012), pp. 98 – 108. https://doi.org/10.1016/j.rse.2012.02.022.

[44] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, Springer Publishing Company, Inc., 2010. ISBN 9780387217369.

[45] C. Wu, B. Du, X. Cui, and L. Zhang, *A post-classification change detection method based on iterative slow feature analysis and Bayesian soft fusion*, Remote Sensing of Environment, 199 (2017), pp. 241 – 255. https://doi.org/10.1016/j.rse.2017.07.009.

[46] M. Zanetti, F. Bovolo, and L. Bruzzone, *Rayleigh-Rice Mixture Parameter Estimation via EM Algorithm for Change Detection in Multispectral Images*, IEEE Transactions on Image Processing, 24 (2015), pp. 5004–5016. https://doi.org/10.1109/TIP.2015.2474710.

[47] M. Zanetti and L. Bruzzone, *A Theoretical Framework for Change Detection Based on a Compound Multiclass Statistical Model of the Difference Image*, IEEE Transactions on Geoscience and Remote Sensing, 56 (2018), pp. 1129–1143. https://doi.org/10.1109/TGRS.2017.2759663.

[48] Y. ZHAN, K. FU, M. YAN, X. SUN, H. WANG, AND X. QIU, *Change detection based on deep siamese convolutional network for optical aerial images*, IEEE Geoscience and Remote Sensing Letters, 14 (2017), pp. 1845–1849. https://doi.org/10.1109/LGRS.2017.2738149.

[49] P. ZHANG, M. GONG, L. SU, J. LIU, AND Z. LI, *Change detection based on deep feature representation and mapping transformation for multi-spatial-resolution remote sensing images*, ISPRS Journal of Photogrammetry and Remote Sensing, 116 (2016), pp. 24 – 41. https://doi.org/10.1016/j.isprsjprs.2016.02.013.

[50] Z. ZHU AND C.E. WOODCOCK, *Continuous change detection and classification of land cover using all available Landsat data*, Remote Sensing of Environment, 144 (2014), pp. 152 – 171. https://doi.org/10.1016/j.rse.2014.01.011.