Robust fusion algorithms for unsupervised change detection between multi-band optical images - A comprehensive case study

Vinicius Ferraris, Nicolas Dobigeon, Marie Chabert

To cite this version:

Vinicius Ferraris, Nicolas Dobigeon, Marie Chabert. Robust fusion algorithms for unsupervised change detection between multi-band optical images - A comprehensive case study. Information Fusion, 2020, 64, pp.293-317. 10.1016/j.inffus.2020.08.008 . hal-02949168

HAL Id: hal-02949168
https://hal.science/hal-02949168v1

Submitted on 28 Sep 2020

Robust fusion algorithms for unsupervised change detection between multi-band optical images – A comprehensive case study

Vinicius Ferraris[a,], Nicolas Dobigeon[a,b], Marie Chabert[a]

[a] *University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse, France*
[b] *Institut Universitaire de France, France*

Abstract

Unsupervised change detection techniques are generally constrained to two multi-band optical images acquired at different times through sensors sharing the same spatial and spectral resolution. In the case of the optical modality, largely studied in the remote sensing community, a straight comparison of homologous pixels such as pixel-wise differencing is suitable. However, in some specific cases such as emergency situations, punctual missions, defense and security, the only available images may be those acquired through different kinds of sensors with different resolutions. Recently some change detection techniques, dealing with images with different spatial and spectral resolutions, have been proposed. Nevertheless, they are focused on a specific scenario where one image has a high spatial and low spectral resolution while the other has a low spatial and high spectral resolution. This paper addresses the problem of detecting changes between any two multi-band optical images disregarding their spatial and spectral resolution disparities. To overcome resolution disparity, state-of-the art methods apply conventional change detection methods after preprocessing steps applied independently on the two images, e.g. resampling operations intended to reach the same spatial and spectral resolutions. Nevertheless, these preprocessing steps may waste relevant information since they do not take into account the strong interplay existing between the two images. Conversely, in this paper, we propose a method that more effectively uses the available information by modeling the two observed images as spatially and spectrally degraded versions of two (unobserved) latent images characterized by the same high spatial and high spectral resolutions. Covering the same scene, the latent images are expected to be globally similar except for possible changes in spatially sparse locations. Thus, the change detection task is envisioned through a robust fusion task which enforces the differences between the estimated latent images to be spatially sparse. We show that this robust fusion can be formulated as an inverse problem which is iteratively solved using an alternating minimization strategy. The proposed framework is implemented for an exhaustive list of applicative scenarios and applied to real multi-band optical images. A comparison with state-of-the-art change detection methods evidences the accuracy and the versatility of the proposed robust fusion-based strategy.

Keywords:
Image fusion, change detection, different resolutions, hyperspectral imagery, multispectral imagery.

---

1. Introduction

Remote sensing consists in collecting measurements, without any physical contact, about an object or phenomenon. This paper focuses on applications to Earth observation and surface monitoring [1, 2, 3, 4]. The type of acquired measurements, also referred to as modality, is intimately related to the sensor. Each modality provides a predefined amount and type of information about the scene. The technological growth and new data policies increase the availability of multi-temporal data (i.e., acquired at different time instants) [5], while simultaneously introducing new challenges. Notably, multi-temporal data acquired over the same geographical location can be used to detect changes or variations. Thus, analyzing multi-temporal data has culminated in the development of an extremely important area for the remote sensing community, namely, change detection (CD).

CD refers to the techniques used to detect areas where potential changes have occurred between multiple multi-temporal and possibly multi-source (i.e., from different sensors) images acquired over the same scene (geographical location) [6, 7]. CD is generally conducted within a supervised or unsupervised context [5]. The former requires prior ground-truth knowledge in order to train algorithms maximizing the detection rate while minimizing the false alarm rate [8, 9, 10]. Conversely the latter tries to infer changes after carefully designing a blind model-based distance operator. As ground-truth information is hardly available, significant efforts have been made so that unsupervised CD techniques reach the supervised CD performance. For instance, by combining the results of traditional CD techniques in order to obtain a more reliable change map [11]. Nevertheless, almost all unsupervised CD methods only focus on a particular scenario, actually the most favorable one, which considers two multi-band optical images with same spatial and spec-

---

Email addresses: `vinicius.ferraris@enseeiht.fr` (Vinicius Ferraris), `nicolas.dobigeon@enseeiht.fr` (Nicolas Dobigeon), `marie.chabert@enseeiht.fr` (Marie Chabert)

Preprint submitted to Information Fusion *September 28, 2020*

tral resolutions [8, 5, 11]. There are two main reasons for considering such a scenario: *i*) multi-band optical images represent the most commonly used remote sensing modality [12], and *ii*) images with same spatial and spectral resolutions are pixel-wisely comparable, which eases the use of distance operators.

Multi-band optical sensors provide a particular representation of the observed scene according to some of its intrinsic characteristics, particularly, its ability of reflecting the incoming light. Well suited to map horizontal structures like land-cover type at large scales [13], easy to interpret, optical imagery is widespread. Another important aspect of optical imagery is the commonly admitted Gaussian modeling of the sensor noise, which has lead to a massive development of least-squares like methods, specially for CD. Indeed, the properties of the noise model, for instance the symmetry of the Gaussian probability distribution function, justify the implementation of CD techniques through image differencing, as noticed in [6] and [5]. Although differencing methods have been adapted to handle multi-band images by considering spectral change vectors [14, 15] and transform analysis [16, 17], they generally rely on the crucial premise that the observed images share the same spatial and/or spectral resolutions.

However, the need for flexible and reliable CD techniques that are able to handle more scenarios is real. In some situations, for instance consecutive to natural disasters or within punctual imagery missions, the limited availability of the sensors and the time constraints may preclude the use of the same sensor at two distinct time instants. Thus, in these cases, observed images are possibly from different modalities and may not share the same spatial and/or spectral resolutions. To make existing conventional CD methods usable in these cases, one strategy, hereafter referred to as the worst-case (WC) method, consists in individually and independently, spatially and/or spectrally, resampling the images to reach the same spatial and spectral resolutions. Although this WC technique allows off-the-shelf CD techniques to be used directly, it may remain sub-optimal since *i*) resampling operations independently applied to each image do not take into account their joint characteristics and thus crucial information may be missed and *ii*) these spatial and spectral operations are generally from a higher to a lower resolution, which results in a significant loss of information. To overcome these limitations, the authors in [18] and [19] recently proposed two CD approaches specifically designed to deal with multi-band images with different spatial and spectral resolutions. Both approaches rely on the inference of a latent (i.e., unobserved) image which results from the fusion of the two observed images. Fusing information contained in remote sensing images has motivated a lot of research works in the literature [20, 21, 22, 23, 24, 11, 10]. Within a CD context, the underlying assumption is that most of pixels of the fused image, who remain unchanged during the time interval, produce consistent information while the few others, locating in the change regions, produce aberrant information. More precisely, the method proposed in [19] is based on a 3-step procedure (namely fusion, prediction and detection) which, instead of independently preprocess each observed image, recovers a latent high spatial and spectral resolution image containing changed and unchanged regions by fusing observed images. Then, it predicts pseudo-observed images by artificially degrading this estimated latent image using forward models underlying the actually observed images. The result is two pairs, each composed of a predicted image and an observed image with the same spatial and spectral resolutions. Finally, any classical multi-band CD method can be applied to estimate two change images, that can be thresholded to build the change maps. Conversely, the robust fusion-based CD technique proposed in [18] aims at recovering two high spatial and spectral resolution latent images related to the observed images via a double physically-inspired forward model. Even if they have shown significant improvements in the detection rate when compared to WC method, they are also still limited to a single scenario: one high spatial low spectral resolution image and one low spatial high spectral resolution image.

In this paper, capitalizing on the appealing results reported in [18], we show that the unsupervised CD problem can be formulated as a generic robust-fusion task for an exhaustive set of experimental scenarios, extending significantly the work in [18]. The two observed images are modeled as spatially and spectrally degraded versions of two (unobserved) latent images characterized by the same high spatial and high spectral resolutions. Covering the same scene, the latent images are expected to be globally similar except for possible changes in spatially sparse locations. The resulting objective function is solved through the use of an alternating minimization (AM) algorithm, which iteratively optimizes with respect to (w.r.t.) one latent image and the change image. This unifying framework has the major asset of encompassing all possible situations that can be encountered when detecting changes between two multi-band optical images. More precisely, we exhaustively identify ten different acquisition scenarios that differ by the respective spatial and spectral resolutions of the two images to be analyzed. In particular, one scenario will correspond to detecting changes between images of same spatial and spectral resolutions, as conveniently considered in the literature [8, 5, 11]. Another particular scenario will consist in conducting CD between two images with the same spatial resolution but different spectral resolution, considered in [16] and [17]. Among the ten scenarios, a third specific one will correspond to detecting changes between images of complementary spatial and spectral resolutions, as considered in our previous contributions [25, 18, 19, 26]. The proposed generic robust-fusion formulation is subsequently instantiated for each of these identified applicative scenarios[1]. Remarkably, regardless the considered scenario, each of the steps embedded in the resulting AM algorithms can be interpreted as well-documented image restoration procedures, namely multiband image fusion, spatial deblurring, spectral deblurring and denoising. Moreover, these steps can be efficiently conducted using state-of-the-art strategies from the literature. To summarize, one of the major contributions reported in this manuscript consists in demonstrating that a uni-

---

[1] Note that the solution introduced by [18] is a specific instance of the general framework developed in this paper dedicated to the sole third scenario highlighted in this introduction.

fying framework, that is robust fusion, offers the possibility of addressing the problem of CD, whatever the spatial and spectral resolutions of the observed images. To the best of our knowledge, no work from the literature has conducted such an exhaustive case study and has proposed such a versatile and efficient CD framework able to indistinctly handle all possible acquisition scenarios.

The paper is organized as follows. Section 2 formulates the problem of CD between multi-band optical images as a robust fusion task. It also draws an exhaustive list of ten possible scenarios that could be encountered in real-world applications. These scenarios mainly differ by the dissimilarities of the two images in terms of spatial and spectral resolutions. Section 3 presents the generic algorithm proposed to conduct this robust fusion task. It also discusses its particular instances with respect to the scenarios previously identified. In Section 4, the performance of the proposed solution is quantitatively assessed on a set of simulated datasets for the five most representative applicative scenarios. Section 5 allows the versatility and accuracy of the solution to be visually assessed based on experiments conducted on real images for all possible scenarios described in Section 2.3. Section 6 concludes the paper.

2. Problem formulation

2.1. *Generic forward model for multi-band optical images*

The image formation process, inherent to multi-band optical sensors, can be generally modeled as a sequence of successive transformations and degradations. These transformations are applied over the original scene and result in an output digital image, commonly referred to as the observed image and denoted $\mathbf{Y} \in \mathbb{R}^{m_\lambda \times m}$ where m_λ and m are the numbers spectral bands and of pixels in the observed image. $\mathbf{Y}$ corresponds to the particular limited representation of the original scene the sensor is able to acquire. The original scene cannot be exactly represented because of its continuous nature, but it can be conveniently approximated by an (unknown) latent digital image of higher spatial and spectral resolutions, $\mathbf{X} \in \mathbb{R}^{n_\lambda \times n}$, where $n_\lambda \geq m_\lambda$ and $n \geq m$ are the numbers of spectral bands and of pixels, respectively. In what follows, as a well-admitted approximation, the observed and latent images are assumed to be related according to the generic forward model [27, 28, 29]

$$\mathbf{Y} = \mathbf{L}\mathbf{X}\mathbf{R} + \mathbf{N} \qquad (1)$$

where

- $\mathbf{L} \in \mathbb{R}^{m_\lambda \times n_\lambda}$ is a spectral degradation matrix,
- $\mathbf{R} \in \mathbb{R}^{n \times m}$ is a spatial degradation matrix,
- $\mathbf{N}$ is an additive term comprising sensor noise and modeling errors.

In (1), the left-multiplying matrix $\mathbf{L}$ and right-multiplying matrix $\mathbf{R}$ spectrally and spatially degrade the latent image, respectively, by combining some spectral bands of each pixel or by combining neighboring pixel values in each spectral band. More precisely, the spectral degradation $\mathbf{L}$ represents a spectral resolution reduction with respect to the latent image $\mathbf{X}$, as already considered in [28], [29] and [30]. In practice, this matrix is fully defined by spectral filters characterizing the optical sensors. When the specifications of the sensor are available, these filters are known. Otherwise, they can be learned by cross-calibration, e.g., following the strategies proposed in [29] or [31]. On the other hand, the spatial degradation matrix $\mathbf{R}$ models the combination of different spatial transformations applied to the pixel values within each spectral band. These transformations are specific of the sensor architecture and include warp, blur, translation and decimation [31, 30]. In this work, geometrical transformations such as warp and translation are assumed to have been previously corrected, e.g., using image spatial alignment techniques. Thus, similarly to the model considered in [30], the spatial degradation matrix $\mathbf{R}$ only stands for a spatially invariant blurring, followed by a decimation (i.e., downsampling) operation. Thus, in what follows, the spatial degradation matrix $\mathbf{R}$ will be assumed of the form

$$\mathbf{R} = \mathbf{B}\mathbf{S}. \qquad (2)$$

The sparse symmetric Toeplitz matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ in (2) operates a cyclic convolution on each individual band to model a space-invariant blur associated with a symmetric convolution kernel. The decimation operation, denoted by the $n \times m$ matrix $\mathbf{S}$ in (2), corresponds to a uniform downsampling operator[2] of factor $d = d_r \times d_c$ with $m = n/d$ ones on the block diagonal and zeros elsewhere, such that $\mathbf{S}^T\mathbf{S} = \mathbf{I}_m$ [30].

The noise corrupting multi-band optical images is generally modeled as additive and Gaussian [5, 2, 32, 30]. Thus the noise matrix $\mathbf{N}$ in (1) is assumed to be distributed according to the following matrix normal distribution (see Appendix A)

$$\mathbf{N} \sim \mathcal{MN}_{m_\lambda, m}(\mathbf{0}_{m_\lambda \times m}, \mathbf{\Lambda}, \mathbf{\Pi}).$$

The row covariance matrix $\mathbf{\Lambda}$ carries information regarding the between-band spectral correlation. In what follows, similarly to the approach by [30], this covariance matrix $\mathbf{\Lambda}$ will be assumed to be diagonal, which implies that the noise is spectrally independent and characterized by a specific variance in each band. Conversely, the column covariance matrix $\mathbf{\Pi}$ models the noise correlation w.r.t. to the pixel locations. Following a hypothesis widely admitted in the literature, this matrix is assumed to be identity, $\mathbf{\Pi} = \mathbf{I}_m$, to reflect the fact that the noise is spatially independent. In real applications, both matrices $\mathbf{\Lambda}$ and $\mathbf{\Pi}$ can be estimated by calibration [31].

2.2. *Problem statement*

Let us consider two co-registered multi-band optical images $\mathbf{Y}_1 \in \mathbb{R}^{m_{\lambda_1} \times m_1}$ and $\mathbf{Y}_2 \in \mathbb{R}^{m_{\lambda_2} \times m_2}$ acquired by two sensors S_1 and S_2 at times t_1 and t_2, respectively. The chronological order of acquisitions is indifferent: either $t_2 < t_1$ or $t_2 > t_1$ are possible cases. The problem addressed in this paper consists in detecting significant changes between these two multi-band optical

---

[2] The operator $\mathbf{S}^T$ represents an upsampling transformation by zero-interpolation from m to n.

images. This is a challenging task mainly due to the possible spatial and/or spectral resolution dissimilarity (i.e., $m_{\lambda_1} \neq m_{\lambda_2}$ and/or $m_1 \neq m_2$), which prevents the use of any pixel-wise differencing operations [6, 5]. To alleviate this issue, this work proposes to generalize the CD framework introduced in [18] to handle all possible combinations (scenarios) of the multi-band optical image resolutions. More precisely, following the widely admitted forward model (1) described in Section 2.1 and adopting consistent notations, the observed images $\mathbf{Y}_1$ and $\mathbf{Y}_2$ can be related to two latent images $\mathbf{X}_1 \in \mathbb{R}^{n_\lambda \times n}$ and $\mathbf{X}_2 \in \mathbb{R}^{n_\lambda \times n}$ with the same spatial and spectral resolutions

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1 \tag{3a}$$
$$\mathbf{Y}_2 = \mathbf{L}_2\mathbf{X}_2\mathbf{R}_2 + \mathbf{N}_2. \tag{3b}$$

where $\mathbf{L}_1$ and $\mathbf{L}_2$ denote two spectral degradation operators and $\mathbf{R}_1$ and $\mathbf{R}_2$ denote two spatial degradation operators that can be decomposed according to (2). Note that (3a) and (3b) are a specific double instance of the model (1). In particular, the two multi-band latent images $\mathbf{X}_1$ and $\mathbf{X}_2$ share the same spectral and spatial resolutions, generally higher than those of the observed images:

$$n_\lambda \geq \max\{m_{\lambda_1}, m_{\lambda_2}\} \text{ and/or } n \geq \max\{m_1, m_2\}. \tag{4}$$

Thereby, after inferring the latent images, any classical differencing technique can be subsequently implemented to compute a change image $\Delta\mathbf{X} = [\Delta\mathbf{x}_1, \ldots, \Delta\mathbf{x}_n] \in \mathbb{R}^{n_\lambda, n}$ defined by

$$\Delta\mathbf{X} = \mathbf{X}_2 - \mathbf{X}_1 \tag{5}$$

where $\Delta\mathbf{x}_p \in \mathbb{R}^{n_\lambda}$ denotes the spectral change vector in the pth pixel ($p = 1, \ldots, n$). It is worth noting that, under the assumptions (4), these changes can be identified at a high spatial and spectral resolutions. Finally this change image can be further exploited by conducting a pixel-wise change vector analysis (CVA) which exhibits the polar coordinates (i.e., magnitude and direction) of the spectral change vectors [33]. Then, to spatially locate the changes, a natural approach consists in monitoring the information contained in the magnitude part of this representation, summarized by the change energy image [6, 14, 15]

$$\mathbf{e} = [e_1, \ldots, e_n] \in \mathbb{R}^n$$

with

$$e_p = \left\|\Delta\mathbf{x}_p\right\|_2, \quad p = 1, \ldots, n.$$

When the CD problem in the pth pixel is formulated as the binary hypothesis testing

$$\begin{cases} \mathcal{H}_{0,p} & : \text{ no change occurs in the } p\text{th pixel} \\ \mathcal{H}_{1,p} & : \text{ a change occurs in the } p\text{th pixel} \end{cases} \tag{6}$$

a pixel-wise statistical test can be written by thresholding the change energy image pixels

$$e_p \underset{\mathcal{H}_{0,p}}{\overset{\mathcal{H}_{1,p}}{\gtrless}} \tau. \tag{7}$$

The final binary CD map denoted $\mathbf{d} = [d_1, \ldots, d_n] \in \{0, 1\}^n$ can be derived as

$$d_p = \begin{cases} 1 & \text{if } e_p \geq \tau \quad (\mathcal{H}_{1,p}), \\ 0 & \text{otherwise} \quad (\mathcal{H}_{0,p}). \end{cases}$$

As a consequence, to solve the multi-band image CD problem, the key issue lies in the joint estimation of the pair of latent images $\{\mathbf{X}_1, \mathbf{X}_2\}$ from the joint forward model (3) or, equivalently, the joint estimation of one latent image and the difference image, i.e., $\{\mathbf{X}_1, \Delta\mathbf{X}\}$. Finally the next paragraph draws an exhaustive list of possible applicative scenarios that can be encountered in real-world problems. These scenarios differ by the corresponding spatial and spectral degradations relating the pair of observed images $\{\mathbf{Y}_1, \mathbf{Y}_2\}$ and the pair of latent images $\{\mathbf{X}_1, \mathbf{X}_2\}$ (or, equivalently, $\{\mathbf{X}_1, \Delta\mathbf{X}\}$).

2.3. Comprehensive taxonomy of applicative scenarios

The possible applicative scenarios covered by the generic formulation (3) differ by the combination of spectral degradations $\mathbf{L}_i$ ($i \in \{1, 2\}$) and spatial degradations $\mathbf{R}_j$ ($j \in \{1, 2\}$) actually involved in the joint forward model (3), i.e., when these operators are different from the identity matrix $\mathbf{I}$. Table 1 makes an exhaustive list of the 10 distinct scenarios according to the degradations operated on the two latent images $\mathbf{X}_1$ and $\mathbf{X}_2$. To cover all cases, these scenarios are denoted $\mathcal{S}_{\alpha\beta}$ where $\alpha \in \{\mathrm{s}, \mathrm{d}, \delta\}$ and $\beta \in \{\mathrm{s}, \mathrm{d}, \delta\}$ refer to the possible differences in terms of spectral and spatial resolutions, respectively, between the two observed images $\mathbf{Y}_1$ and $\mathbf{Y}_2$. The index s, related to either spectral or spatial resolution, indicates that the observed images $\mathbf{Y}_1$ and $\mathbf{Y}_2$ share the same corresponding resolution, i.e., there is no degradation or, equivalently, the corresponding degradation operators reduce to the identity matrix $\mathbf{I}$. When these two images are of different resolutions, two cases may occur. Either the underlying downsamplings are conducted on superimposable grids (scenario indexed by d), for which a unique actual degradation is involved while the other reduces to the identity operator. Conversely, the sampling can be conducted on non-superimposable sampling grids, which requires the simultaneous use of two degradation operators (scenario indexed by δ). Moreover, two different instances of the scenario $\mathcal{S}_{\mathrm{dd}}$ can be encountered. Recall that it corresponds to the acquisitions of two images of different spatial and spectral resolutions, yet acquired on the same spatial and spatial sampling grids, thus relying on one spatial and one spectral degradation operators. In this case, the so-called balanced scenario, denoted $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$, is characterized by one image of high spatial and low spectral resolutions while the other is of low spatial and high spectral resolutions. The dual unbalanced scenario, denoted $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$, occurs when one image is of high spatial and spectral resolutions while the other is of low spatial and spectral resolution. The specificities of these scenarios are also discussed in what follows.

$\mathcal{S}_{\mathrm{ss}}$ is devoted to a pair of observed images sharing the same spatial and spectral resolutions. In this case, CD can be conducted by pixel-wise comparisons, as classically addressed in the literature [6, 5].

	Forward model ♯1		**Forward model ♯2**		**Comments**
	Spectral degradation	**Spatial degradation**	**Spectral degradation**	**Spatial degradation**	
$\mathcal{S}_{\mathrm{ss}}$	–	–	–	–	Conventional CD framework – $\mathbf{Y}_1$ and $\mathbf{Y}_2$ of same spatial and spectral resolutions
$\mathcal{S}_{\mathrm{ds}}$	$\mathbf{L}_1$	–	–	–	$\mathbf{Y}_1$ of lower spectral resolution $\mathbf{Y}_1$ and $\mathbf{Y}_2$ of same spatial resolutions
$\mathcal{S}_{\delta\mathrm{s}}$	$\mathbf{L}_1$	–	$\mathbf{L}_2$	–	Generalization of $\mathcal{S}_{\mathrm{ds}}$ with non-superimposable spectral sampling grids
$\mathcal{S}_{\mathrm{sd}}$	–	$\mathbf{R}_1$	–	–	$\mathbf{Y}_1$ of lower spatial resolution $\mathbf{Y}_1$ and $\mathbf{Y}_2$ of same spectral resolutions
$\mathcal{S}_{\mathrm{s}\delta}$	–	$\mathbf{R}_1$	–	$\mathbf{R}_2$	Generalization of $\mathcal{S}_{\mathrm{sd}}$ with non-superimposable spatial sampling grids
$\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$	–	$\mathbf{R}_1$	$\mathbf{L}_2$	–	$\mathbf{Y}_1$ and $\mathbf{Y}_2$ of complementary resolutions
$\mathcal{S}_{\delta\mathrm{d}}$	$\mathbf{L}_1$	$\mathbf{R}_1$	$\mathbf{L}_2$	–	Generalization of $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ with non-superimposable spectral sampling grids
$\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$	$\mathbf{L}_1$	$\mathbf{R}_1$	–	–	$\mathbf{Y}_1$ of low spatial and spectral resolutions
$\mathcal{S}_{\mathrm{d}\delta}$	$\mathbf{L}_1$	$\mathbf{R}_1$	–	$\mathbf{R}_2$	Generalization of $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$ with non-superimposable spatial sampling grids
$\mathcal{S}_{\delta\delta}$	$\mathbf{L}_1$	$\mathbf{R}_1$	$\mathbf{L}_2$	$\mathbf{R}_2$	Generalization of $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ and $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$ with non-superimposable spatial and spectral sampling grids

Table 1: Overview of the spectral and spatial degradations w.r.t. to experimental scenarios. The symbol "–" stands for "no degradation" (i.e., corresponding operator replaced by the identity matrix **I**).

- $\mathcal{S}_{\mathrm{ds}}$ consists in conducting CD between two images with the same spatial resolution but different spectral resolutions, under the assumption that the underlying spectral sampling grids of the two observed images can be superimposed, as considered in [16] and [17].

- $\mathcal{S}_{\delta\mathrm{s}}$ generalizes the previous scenario $\mathcal{S}_{\mathrm{ds}}$ where the difference in spectral resolutions cannot be expressed using a unique spectral degradation matrix due to non-superimposable spectral sampling grids.

- $\mathcal{S}_{\mathrm{sd}}$ consists in conducting CD between two images with the same spectral resolution but different spatial resolutions. In this case, the underlying spatial sampling grids of the two observed images can be superimposed.

- $\mathcal{S}_{\mathrm{s}\delta}$ generalizes the scenario $\mathcal{S}_{\mathrm{sd}}$ where the difference in spatial resolutions cannot be expressed using a unique spatial degradation matrix due to non-superimposable spatial sampling grids. As a consequence, the latent images $\mathbf{X}_1$ and $\mathbf{X}_2$ are characterized by a common spatial resolution which is higher than those of both observed images. The choice of the virtual upsampling factors (from $\mathbf{Y}_1$ to $\mathbf{X}_1$ and from $\mathbf{Y}_2$ to $\mathbf{X}_2$) is based on the greatest common divisor between spatial resolutions.

- $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ relies on two complementary images of different spatial and spectral resolutions: the first image with high spectral and low spatial resolutions, the second image with low spectral and high spatial resolutions. This is the balanced CD scenario considered in [25, 19, 18]. When the two observed images have been acquired at the same time instants ($t_i = t_j$), this scenario corresponds to the multiband image fusion task considered in numerous works [27, 28, 29].

- $\mathcal{S}_{\delta\mathrm{d}}$ generalizes the scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$, but the difference in spectral resolutions cannot be expressed using a single degradation matrix due to different spectral sampling grid.

- $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$ represents the unbalanced counterpart of $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ where one image is of high spatial and spectral resolutions while the other is of low spatial and spectral resolutions.

- $\mathcal{S}_{\mathrm{d}\delta}$ generalizes the scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$, but the difference in spatial resolutions cannot be expressed using a single degradation matrix due to different spatial sampling grid.

- $\mathcal{S}_{\delta\delta}$ generalizes the scenarios $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ and $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$, but the difference in spatial and spectral resolutions cannot be expressed using unique spatial and spectral degradation matrices, due to non-aligned sampling grids.

The next section formulates the joint estimation of the latent image $\mathbf{X}_1$ and the difference image $\Delta\mathbf{X}$ as an optimization problem. It describes a generic alternating minimization scheme which solves this problem and discusses its particular instances with respect to these applicative scenarios.

3. Robust multi-band image fusion algorithm

3.1. Optimization problem

Following a Bayesian approach, the joint maximum a posteriori (MAP) estimator $\left\{\hat{\mathbf{X}}_{1,\mathrm{MAP}}, \Delta\hat{\mathbf{X}}_{\mathrm{MAP}}\right\}$ of the latent and change images can be derived by maximizing the posterior distribution

$$p(\mathbf{X}_1, \Delta\mathbf{X}|\mathbf{Y}_2, \mathbf{Y}_1) \quad \propto \quad p(\mathbf{Y}_2, \mathbf{Y}_1|\mathbf{X}_1, \Delta\mathbf{X})p(\mathbf{X}_1)p(\Delta\mathbf{X})$$

where $p(\mathbf{Y}_2, \mathbf{Y}_1|\mathbf{X}_1, \Delta\mathbf{X})$ is the joint likelihood function and $p(\mathbf{X}_1)$ and $p(\Delta\mathbf{X})$ correspond to the prior distributions associated with the latent and change images, respectively, assumed to be a priori independent. Because of the additive nature and statistical properties of the noise discussed in Section 2.1, this boils down to solve the following minimization problem

$$\left\{\hat{\mathbf{X}}_{1,\mathrm{MAP}}, \Delta\hat{\mathbf{X}}_{\mathrm{MAP}}\right\} \in \operatorname*{argmin}_{\mathbf{X}_1,\Delta\mathbf{X}} \mathcal{J}(\mathbf{X}_1, \Delta\mathbf{X}) \tag{8}$$

with

$$\begin{aligned}\mathcal{J}(\mathbf{X}_1, \Delta\mathbf{X}) = & \frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\mathbf{Y}_2 - \mathbf{L}_2\left(\mathbf{X}_1 + \Delta\mathbf{X}\right)\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 \\ & + \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ & + \lambda\phi_1\left(\mathbf{X}_1\right) + \gamma\phi_2\left(\Delta\mathbf{X}\right)\end{aligned} \tag{9}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. The regularizing functions $\phi_1(\cdot)$ and $\phi_2(\cdot)$ can be related to the negative log-prior distributions of the latent and change images, respectively, and the parameters λ and γ tune the amount of corresponding penalizations in the overall objective function $\mathcal{J}(\mathbf{X}_1, \Delta\mathbf{X})$. These functions should be carefully designed to exploit any prior knowledge regarding the parameters of interest. As discussed in Section 3.2.1, numerous regularizations can be advocated for the latent image $\mathbf{X}_1$. In this work, to maintain computational efficiency while providing accurate results [32], a Tikhonov regularization proposed in [27] has been adopted

$$\phi_1\left(\mathbf{X}_1\right) = \left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2$$

where $\bar{\mathbf{X}}_1$ refers to a crude estimate of $\mathbf{X}_1$.

Regarding the regularizing function $\phi_2(\cdot)$, as already mentioned in the previous section, it should reflect the fact that most of the pixels are expected to remain unchanged i.e., most of the columns of the change image $\Delta\mathbf{X}$ are expected to be null vectors. Thus, the regularizing function $\phi_2(\cdot)$ is chosen as in [34] as the $\ell_{2,1}$-norm of the change image

$$\phi_2\left(\Delta\mathbf{X}\right) = \|\Delta\mathbf{X}\|_{2,1} = \sum_{p=1}^{n}\left\|\Delta\mathbf{x}_p\right\|_2. \tag{10}$$

As mentioned above, the parameter γ adjusts the weight of the regularization (10) in the overall objective function (9), i.e., the expected level of spatial sparsity of the recovered change image $\Delta\mathbf{X}$. Interestingly, this parameter implicitly determines the operating point of the corresponding detector, in place of the conventional decision threshold τ defined in the hypothesis testing (6). Indeed, for $1/\gamma = 0$, the change image $\Delta\mathbf{X}$ will be only composed of 0's. All pixels are thus identified as unchanged. In this limit case (where γ goes to infinity), the operating point of the underlying detector is characterized by a probability of false alarm (or type I error) and a probability of detection (or sensitivity) of $\mathrm{PFA} = 0$ and $\mathrm{PD} = 0$, respectively. Conversely, for $1/\gamma$ going to infinity, no spatial sparsity will be promoted and the change image $\Delta\mathbf{X}$ will be only composed of 1's. All pixels are thus identified as changed with $\mathrm{PFA} = \mathrm{PD} = 1$. Intermediate values of $1/\gamma$ allows the full range of the detector operating points to be explored, reflecting the intrinsic trade-off between the PFA and the PD.

3.2. Generic resolution

Computing the joint MAP estimator of the latent image $\mathbf{X}_1$ at time t_1 and of the change image $\Delta\mathbf{X}$ can be achieved by solving the minimization problem in (8). However, no closed-form solution can be derived for this problem for all the scenarios of interest. Thus this section introduces a minimization algorithm which iteratively converges to this solution. This alternating minimization (AM) algorithm, summarized in Algo. 1, consists in iteratively minimizing the objective function (9) w.r.t. $\mathbf{X}_1$ and $\Delta\mathbf{X}$, within so-called *fusion* and *correction* step detailed below.

Algorithm 1 Algorithm for robust multi-band image fusion

```
Input: $\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{L}_1, \mathbf{L}_2, \mathbf{R}_1, \mathbf{R}_2, \mathbf{\Lambda}_1, \mathbf{\Lambda}_2$.
Set $\Delta\mathbf{X}_1$.
for $k = 1, \ldots, K$ do
    % Fusion step
    $\mathbf{X}_1^{(k+1)} = \arg\min_{\mathbf{X}_1} \mathcal{J}(\mathbf{X}_1, \Delta\mathbf{X}^{(k)})$
    % Correction step
    $\Delta\mathbf{X}^{(k+1)} = \arg\min_{\Delta\mathbf{X}} \mathcal{J}(\mathbf{X}_1^{(k+1)}, \Delta\mathbf{X})$
end for
Output: $\hat{\mathbf{X}}_{1,\mathrm{MAP}} \triangleq \mathbf{X}_1^{(K+1)}$ and $\Delta\hat{\mathbf{X}}_{\mathrm{MAP}} \triangleq \Delta\hat{\mathbf{X}}^{(K+1)}$
```

3.2.1. Fusion step

As mentioned above, the forward model (3) relying on the pair $\{\mathbf{X}_1, \mathbf{X}_2\}$ of latent images can be rewritten as a function of $\{\mathbf{X}_1, \Delta\mathbf{X}\}$, i.e.,

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1 \tag{11a}$$

$$\mathbf{Y}_2 = \mathbf{L}_2\left(\mathbf{X}_1 + \Delta\mathbf{X}\right)\mathbf{R}_2 + \mathbf{N}_2. \tag{11b}$$

Generalizing the strategy proposed in [18], given the change image $\Delta\mathbf{X}$ and the image $\mathbf{Y}_1$ observed at time t_1, a *corrected* image denoted $\tilde{\mathbf{Y}}_2$ that would have been acquired by the sensor S_2 at time t_1 can be defined as

$$\tilde{\mathbf{Y}}_2 = \mathbf{Y}_2 - \mathbf{L}_2\Delta\mathbf{X}\mathbf{R}_2. \tag{12}$$

With this notation, the forward model (11) can be easily rewritten, leading to

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1 \tag{13a}$$

$$\tilde{\mathbf{Y}}_2 = \mathbf{L}_2\mathbf{X}_1\mathbf{R}_2 + \mathbf{N}_2. \tag{13b}$$

Thus, the fusion step, at iteration k, consists in minimizing (9) w.r.t. $\mathbf{X}_1$, i.e.,

$$\hat{\mathbf{X}}_1^{(k+1)} = \operatorname*{argmin}_{\mathbf{X}_1} \mathcal{J}^{(\mathrm{f})}(\mathbf{X}_1)$$

with

$$\begin{aligned}\mathcal{J}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right)\\ &= \frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\mathbf{X}_1\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2\\ &+ \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2\\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{14}$$

The double forward model (13), as well as the optimization problem (14), underlie the estimation of an image $\mathbf{X}_1$ from an observed image $\mathbf{Y}_1$ and a pseudo-observed image $\tilde{\mathbf{Y}}_2$. Various instances of this pixel-level fusion problem have been widely considered in the literature [20, 22, 23, 24]. For instance, the authors in [35] and [36] have addressed the problem of single mono-band image superresolution from a single observed image $\mathbf{Y}_1$, i.e., with $\mathbf{L}_1 = \mathbf{I}_{m_{\lambda_1}}$ and $m_{\lambda_1} = n_\lambda = 1$. The problem of fusing several degraded mono-band images to recover a common high resolution latent image has been considered in [37]. Similarly, the model (13) generalizes the conventional observational model widely adopted by the remote sensing community to conduct multi-band image fusion [38, 39, 40, 28, 21, 29, 41, 27, 30]. Within this specific scenario corresponding to $\mathcal{S}_{\mathrm{dd}}^{(b)}$, a high spatial and high spectral resolution latent image $\mathbf{X}_1$ is estimated from two observed images, one of low spatial and high spectral resolutions (i.e., $\mathbf{L}_1 = \mathbf{I}_{m_{\lambda_1}}$) and the other of high spatial and low spectral resolutions (i.e., $\mathbf{R}_2 = \mathbf{I}_{n_2}$).

In this context, the CD task considered in this paper can be cast as a so-called *robust* fusion problem since the multi-band image fusion model (13) implicitly depends on the (unknown) change image $\Delta\mathbf{X}$. More precisely, since the two latent images $\mathbf{X}_1$ and $\mathbf{X}_2$ are related to the same scene observed at two time instants, they are expected to share a high level of similarity, i.e., the change image $\Delta\mathbf{X}$ is expected to be spatially sparse. Thus, this additional unknown change image $\Delta\mathbf{X}$ to be inferred can be considered as an outlier term, akin to those encountered in several robust factorizing models such as robust principal component analysis (RPCA) [42] and robust nonnegative factorization [34]. A particular instance of this strategy has been successfully adopted in [18] to detect changes between two complementary multi-band images, i.e., in the particular scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ where $\mathbf{L}_1 = \mathbf{I}_{m_{\lambda_1}}$ and $\mathbf{R}_2 = \mathbf{I}_{n_2}$. In this work, we propose to follow a similar route while significantly generalizing the approach to the much more generic model (3) to handle all the practical scenarios of CD identified in Section 2.3.

3.2.2. Correction step

Given the current state $\mathbf{X}_1$ of the latent image, the predicted image that would be observed by the sensor S_2 at time t_1 can be defined as

$$\check{\mathbf{Y}}_2^{(k)} = \mathbf{L}_2\mathbf{X}_1^{(k)}\mathbf{R}_2 \tag{15}$$

leading to the predicted change image

$$\Delta\check{\mathbf{Y}}_2 = \mathbf{Y}_2 - \check{\mathbf{Y}}_2. \tag{16}$$

Then, the correction step in Algo. 1 consists in solving

$$\Delta\hat{\mathbf{X}}^{(k+1)} = \operatorname*{argmin}_{\Delta\mathbf{X}} \mathcal{J}^{(\mathrm{c})}(\Delta\mathbf{X}) \tag{17}$$

with

$$\begin{aligned}\mathcal{J}^{(\mathrm{c})}(\Delta\mathbf{X}) &\triangleq \mathcal{J}\left(\mathbf{X}_1^{(k)}, \Delta\mathbf{X}\right)\\ &= \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\Delta\check{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\Delta\mathbf{X}\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned} \tag{18}$$

This correction can be interpreted as a joint spatial and spectral deblurring of the predicted change image $\Delta\check{\mathbf{Y}}_2^{(k)}$. Note that this ill-posed inverse problem is regularized through an $\ell_{2,1}$-norm penalization, which promotes the spatial sparsity of the change image $\Delta\mathbf{X}$.

It is worth noting that the difficulty of conducting the two steps of the AM algorithm detailed above is highly related to the presence or absence of spatial and/or spectral degradations operated on the two latent images, according to the applicative scenarios detailed in section 2.3. Interestingly, the following paragraph shows that, for each of these ten scenarios, these two steps generally reduce to ubiquitous (multi-band) image processing tasks, namely denoising, spectral deblurring or spatial super-resolution from a single or multiple images, for which efficient and reliable strategies have been already proposed in the literature.

3.3. Specific instances w.r.t. applicative scenarios

Specific instances of the generic AM algorithm proposed in Section 3.2 can be derived for each of the ten scenarios discussed in Section 2.3. Interestingly, these particular instances relate the embedded steps, namely fusion and correction, with ubiquitous image processing tasks that can be performed efficiently thanks to recent contributions proposed in the image processing literature. Table 2 provides a natural interpretation of each step encountered in each scenario as well as corresponding off-the-shelve algorithms to perform this task. More technical details on the implementation for all scenarios are reported in Appendix B.

Remark: Scenario $\mathcal{S}_{\mathrm{ss}}$ is the only one that considers two images of same spatial and spectral resolutions. Even if the proposed unifying framework is able to handle this simplest scenario, its corresponding instance does not pretend to achieve state-of-the-art detection performance. Since this situation frequently assumed in the literature is not of primary importance in this study, interested readers are invited to consult recent overviews dedicated to particular sensors, techniques or applications underlying images of same spatial and spectral resolutions [43, 5, 44, 45].

4. Experiments on simulated images

4.1. Simulation framework

Conducting a quantitative analysis of detection performance on real datasets requires these datasets to be accompanied by

	Fusion Step		Correction Step	
	Algorithm	Operation	Algorithm	Operation
$\mathcal{S}_{\mathrm{ss}}$	Least squares	Denoising	$\ell_{2,1}$-prox. mapping	Denoising
$\mathcal{S}_{\mathrm{ds}}$	Least squares	Spectral deblurring	$\ell_{2,1}$-prox. mapping	Denoising
$\mathcal{S}_{\delta\mathrm{s}}$	Least squares	Spectral deblurring	Forward-backward	Spectral deblurring
$\mathcal{S}_{\mathrm{sd}}$	[36]	Spatial super-resolution	$\ell_{2,1}$-prox. mapping	Denoising
$\mathcal{S}_{\mathrm{s}\delta}$	ADMM [36] [36]	Spatial super-resolution Spatial super-resolution	ADMM [36] $\ell_{2,1}$-prox. mapping	Spatial super-resolution Denoising
$\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$	[30]	Multi-band image fusion	Forward-backward	Spectral deblurring
$\mathcal{S}_{\delta\mathrm{d}}$	ADMM Least squares [36]	Spectral deblurring Multi-band image fusion	Forward-backward	Spectral deblurring
$\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$	ADMM Least squares [36]	Spectral deblurring Spatial super-resolution	$\ell_{2,1}$-prox. mapping	Denoising
$\mathcal{S}_{\mathrm{d}\delta}$	ADMM [30] [36]	Multi-band image fusion Spatial super-resolution	ADMM [36] $\ell_{2,1}$-prox. mapping	Spatial super-resolution Denoising
$\mathcal{S}_{\delta\delta}$	ADMM [30] [36] Least squares	Multi-band image fusion Spatial super-resolution Spectral deblurring	ADMM $\ell_{2,1}$-prox. mapping [36] Least squares	Denoising Spatial super-resolution Spectral deblurring

Table 2: Overview of the steps of the AM algorithm w.r.t. applicative scenarios.

reliable ground truth information. Unfortunately, as previously pointed out in [15] and [19], such information is generally hard and expensive to acquire, in particular for all scenarios listed in Section 2.3. To alleviate this issue, this section proposes to evaluate the performance of the proposed change detection techniques on real datasets affected by simulated yet realistic changes. More precisely, the evaluation framework proposed in [19] is adopted to generate pairs of before- and after-change multi-band optical images for which the ground truth (i.e., actual change maps) is available. Inspired by the well-known Wald's evaluation protocol dedicated to pansharpening algorithms [46], this framework only requires a single high-resolution hyperspectral (HR-HS) reference image $\mathbf{X}_{\mathrm{ref}}$ and generates a pair of latent HR-HS images $\mathbf{X}_1$ and $\mathbf{X}_2$ resulting from an unmixing-mixing procedure. In this work, the HR-HS reference image $\mathbf{X}_{\mathrm{ref}}$ is chosen as the HS image of Pavia University, Italy, acquired by the reflective optics system imaging spectrometer (ROSIS) sensor. This image is of size 610 × 340 with a spatial resolution of 1.3m and, after removing those affected by vapor water absorption, the number of spectral bands is 103 with a spectral coverage ranging from 0.43 to 0.86μm. This HR-HS reference image is submitted to the unmixing-mixing procedure proposed in [19], yielding two HR-HS latent images $\mathbf{X}_1$ and $\mathbf{X}_2$. These latent images only differ by some simulated changes generated according to predefined change rules with respect to a given HR binary change masks $\mathbf{D}_{\mathrm{HR}}$. Then, pairs of observed images $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are generated by spatially and/or spectrally degrading each HR-HS latent images following (3). Given the large number of scenarios handled by the proposed CD framework, only the most representative scenarios (e.g., $\mathcal{S}_{\mathrm{ss}}$, $\mathcal{S}_{\mathrm{ds}}$, $\mathcal{S}_{\mathrm{sd}}$, $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ and $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$) are considered in this section. Indeed, as emphasized in Table 1, the remaining scenarios ($\mathcal{S}_{\delta\mathrm{s}}$, $\mathcal{S}_{\mathrm{s}\delta}$, $\mathcal{S}_{\delta\mathrm{d}}$, $\mathcal{S}_{\mathrm{d}\delta}$ and $\mathcal{S}_{\delta\delta}$) are generalized counterparts of the previous ones when considering non-superimposable spatial and/or spectral grids. Note however that the performance of the proposed framework will be qualitatively assessed for all scenarios on real datasets in Section 5. According to the five considered scenarios, the real images result from *i*) a spatial degradation characterized by a 5 × 5 Gaussian kernel filtering followed by down-sampling with a decimation factor $d = 5$ and *ii*) a spectral degradation corresponding to a 4-band LANDSAT-like spectral response such that

- $\mathcal{S}_{\mathrm{ss}}$ considers a pair of HR-HS images,
- $\mathcal{S}_{\mathrm{ds}}$ considers a pair of HR-MS and HR-HS images,
- $\mathcal{S}_{\mathrm{sd}}$ considers a pair of LR-HS and HR-HS images,
- $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ considers a pair of LR-HS and HR-MS images
- $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$ considers a pair of LR-MS and HR-HS images

where LR means low resolution and MS denotes multispectral (i.e. lower spectral resolution than HS). To evaluate the robustness of the proposed method against noise, observed images for each simulated pair have been corrupted by a zero mean Gaussian noise whose variance has been adjusted to obtain a signal-to-noise ratio SNR= 30dB.

4.2. Compared methods

As previously exposed, the proposed robust fusion-based CD framework (referred to as RF) is able to deal with all combinations of mono- and multi-band optical images of different spatial and spectral resolutions. It is compared with the fusion-based CD framework (referred to as F) proposed in [19]. This state-of-the-art technique has been originally designed to handle scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ but its *fusion* and *prediction* steps can be easily adapted to fit the requirements of other scenarios. More precisely, the *fusion* step of the F method can be conducted by reusing the *fusion* step of the proposed RF framework (see section 3.2.1) after setting $\Delta\mathbf{X} = 0$. Then the *prediction* step of the F method can be achieved following the generic forward model (3) according to the specificities of each scenario. Finally, the *detection* step of the F method provides a HR estimated change map by using a conventional CD technique comparing observed and predicted HR images of same spatial and spectral resolutions.

Up to the authors' knowledge, there is no other CD technique in the literature with such a versatility, i.e., able to directly address all these scenarios. Thus, the F- and proposed RF-based CD techniques are also compared to a family of empirical CD methods whose rationale boils down to producing pairs of images of the same spatial and spectral resolutions, which can be subsequently compared by conventional pixel-wise CD techniques. More precisely, we consider the so-called worst-case (WC) method which consists in spatially and/or spectrally degrading the observed images to reach a pair of images of low spatial and low spectral resolutions. The second method, denoted SD, applies a spatial super-resolution followed by a spectral degradation of the image $\mathbf{Y}_i$ ($i = 1$ or $i = 2$) with the lowest spatial resolution, to finally produce a pair of images of high spatial and low spectral resolutions. The third method, denoted DS, follows the same process of the SD method but in a reversed order, by first applying a spectral degradation and then a spatial super-resolution.

For the F, WC, SD and DS methods, the comparisons between the images of same spatial and spectral resolutions are achieved thanks to two pixel-wise CD techniques, namely change vector analysis (CVA) [33] and the iteratively reweighted multivariate alteration detection (IRMAD) method [17]. Note that the spatial resolution of the resulting CD binary maps produced by the WC method is driven by the image $\mathbf{Y}_i$ with the lowest spatial resolution. Conversely, the RF, F, SD and DS methods lead to CD maps whose spatial resolution is defined by the image with the highest spatial resolution.

4.3. Figures-of-merit

The CD performance of the three methods has been assessed from empirical receiver operating characteristics (ROC) curves, representing the estimated pixel-wise probability of detection (PD) as a function of the probability of false alarm (PFA) [47, 48, 49]. ROC curves have been a privileged figure-of-merit for CD between remote sensing images [50, 9, 51]. By summarizing the detector sensitivity in a single graph, they provide a fair and comprehensive description of the detection performance. Besides, deriving these ROC curves avoids to determine a decision threshold beforehand, which is application-dependent, since this threshold directly sets the trade-off between good detections and false alarms. In practice, to evaluate the performance of F, WC, SD and DS methods, CVA and IR-MAD are conducted for a wide range of threshold values, e.g., from $\tau \in (0, +\infty)$. For each value τ of this threshold, the empirical probabilities of false alarm and detection are computed by comparing the ground truth and the estimated CD maps. In the limit cases, for $\tau \to +\infty$, all pixels are identified as unchanged (i.e., PFA $=$ PD $= 0$), whereas, for $\tau = 0$, all pixels are identified as changed (i.e., PFA $=$ PD $= 1$). Intermediate values of $\tau \in (0, +\infty)$ are associated with detector operating points characterized by PFA values ranging from 0 to 1. Conversely, for the proposed RF method, as explained in Section 3.1, a similar strategy is conducted using the regularizing parameter γ.

Moreover, two quantitative criteria derived from these ROC curves have been computed, namely, *i*) the area under the curve (AUC), corresponding to the integral of the ROC curve and *ii*) the distance between the no detection point ($\mathrm{PFA} = 1, \mathrm{PD} = 0$) and the equal error rate point (i.e., the intercept of the ROC curve with the diagonal line defined by $\mathrm{PD} = 1 - \mathrm{PFA}$). For both metrics, closer to 1 the criterion, better the detection. Note that the AUC, sometimes referred to as the c-statistic [52, Chap. 9], is a well-documented and statistically sound measure of performance for detection problems [53, 54, 55, 56].

4.4. Results

This paragraph provides the results associated with the five considered scenarios. For each scenario, according to the protocol described in Section 4.1, the performance measures have been averaged other 450 simulated pairs of observed images. The quantitative measures are reported in Table 3 for all methods, as well as the computational times. Note that the results of the SD and/or DS methods are not reported for scenarios $\mathcal{S}_{\mathrm{ss}}$, $\mathcal{S}_{\mathrm{ds}}$, $\mathcal{S}_{\mathrm{ss}}$ and $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$ since, in these cases, they are equivalent to the WC method. Complementary results, including alternative visualizations, are reported in the companion document [57].

4.4.1. Scenario $\mathcal{S}_{\mathrm{ss}}$ (HR-HS and HR-HS)

This scenario considers two HR-HS observed images of same spatial and spectral resolutions. As stated before, even if the proposed framework is able to handle this simplest case, it is of limited interest since the particular instance of the main steps of the AM algorithm boil down to naive pixel-wise denoising (see Table 2). Yet, the results reported here demonstrate that the proposed RF method provides good results, as illustrated by the ROC curves depicted in Fig. 1 and the quantitative metrics reported in Table 3 (first two rows). Indeed the proposed RF achieves results comparable to the F method. In this simple case, the WC method does not apply any spatial nor

		RF	F_{CVA}	F_{IRMAD}	WC_{CVA}	WC_{IRMAD}	SD_{CVA}	SD_{IRMAD}	DS_{CVA}	DS_{IRMAD}
$\mathcal{S}_{\text{ss}}$	AUC	**0.997685**	0.98429	0.988043	0.960281	0.960935	n/a	n/a	n/a	n/a
	Dist.	**0.988799**	0.985099	0.968997	0.913291	0.913191	n/a	n/a	n/a	n/a
	Time	1.41	0.45		**0.12**		n/a		n/a	
$\mathcal{S}_{\text{ds}}$	AUC	**0.975428**	0.966545	0.952212	0.969078	0.970259	n/a	n/a	n/a	n/a
	Dist.	**0.947595**	0.925193	0.89649	0.929393	0.929893	n/a	n/a	n/a	n/a
	Time	24.2	4.08		**0.22**		n/a		n/a	
$\mathcal{S}_{\text{sd}}$	AUC	**0.998422**	0.990585	0.989943	0.983735	0.954326	0.839961	0.852711	n/a	n/a
	Dist.	**0.991799**	0.961396	0.960496	0.941694	0.910691	0.772377	0.780978	n/a	n/a
	Time	63.5	13.5		**1.72**		7.6		n/a	
$\mathcal{S}^{\text{b}}_{\text{dd}}$	AUC	**0.993315**	0.985439	0.983908	0.943492	0.943578	0.859931	0.847857	0.839166	0.847457
	Dist.	**0.984298**	0.957096	0.954595	0.891189	0.892989	0.792179	0.776078	0.766377	0.775678
	Time	31.0	3.8		**0.04**		19.1		0.25	
$\mathcal{S}^{\text{u}}_{\text{dd}}$	AUC	**0.995969**	0.77304	0.821478	0.950671	0.950053	n/a	n/a	n/a	n/a
	Dist.	**0.987699**	0.69777	0.756476	0.89949	0.89719	n/a	n/a	n/a	n/a
	Time	442.9	337.8		**6.2**		n/a		n/a	

Table 3: Scenarios $\mathcal{S}_{\text{ss}}$, $\mathcal{S}_{\text{ds}}$, $\mathcal{S}_{\text{sd}}$, $\mathcal{S}^{\text{b}}_{\text{dd}}$ and $\mathcal{S}^{\text{u}}_{\text{dd}}$: quantitative detection performance (AUC and distance) and computational times (s).

spectral degradation, hence CVA and IRMAD are directly applied on the observed images. The better results obtained by the F and RF methods with respect to CVA and IRMAD can be explained by the intrinsic denoising operated by the latent image estimation.

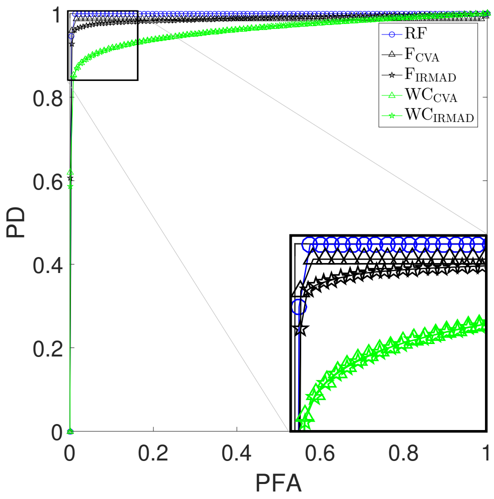


Figure 1: Scenario $\mathcal{S}_{\text{ss}}$ (HR-HS and HR-HS): ROC curves.

4.4.2. *Scenario $\mathcal{S}_{\text{ds}}$ (HR-MS and HR-HS)*

The ROC curves displayed in Fig. 2 with corresponding metrics in Table 3 (3rd and 4th rows) correspond to the CD results obtained from a pair of HR-MS and HR-HS observed images. The proposed robust fusion-based CD technique outperforms the other CD techniques. It is worth noting that the proposed RF approach and the F approach compute change maps from a pair of HR-HS images, while the WC approach, due to the degradations, only maps the changes from a pair of HR-MS images. The lower performance of the F approach can be explained by the difference of noise levels between the predicted and observed image in the *decision* step. This problem is overcome by the proposed RF approach thanks to the intrinsic denoising operation performed during the latent image estimation.

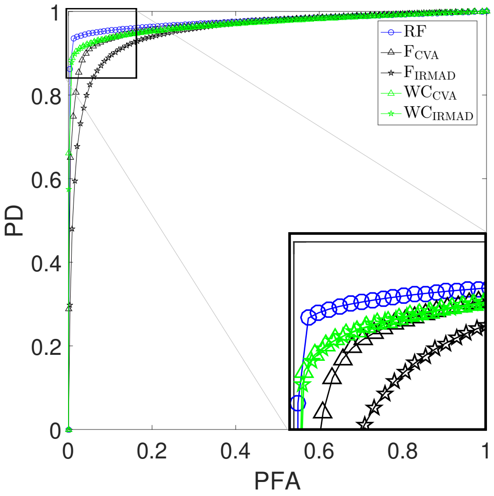


Figure 2: Scenario $\mathcal{S}_{\text{ds}}$ (HR-MS and HR-HS): ROC curves.

4.4.3. *Scenario $\mathcal{S}_{\text{sd}}$ (LR-HS and HR-HS)*

In Scenario $\mathcal{S}_{\text{sd}}$, the HR-MS observed image of Scenario $\mathcal{S}_{\text{ds}}$ is replaced by an LR-HS observed image. The ROC curves are displayed in Fig. 3 with corresponding metrics in Table 3 (5th and 6th rows). As in Scenario $\mathcal{S}_{\text{ds}}$, comparing curves in Fig. 3 shows that the proposed method offers a higher precision than the compared methods even when analyzing a LR observed image. In this scenario, the CD masks $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC method combined with CVA or IRMAD are defined at a LR which contributes to decrease the performance of this method. The CVA- and IRMAD-based SD methods fail to provide competitive results, despite the superresolution of the LR-HS image..

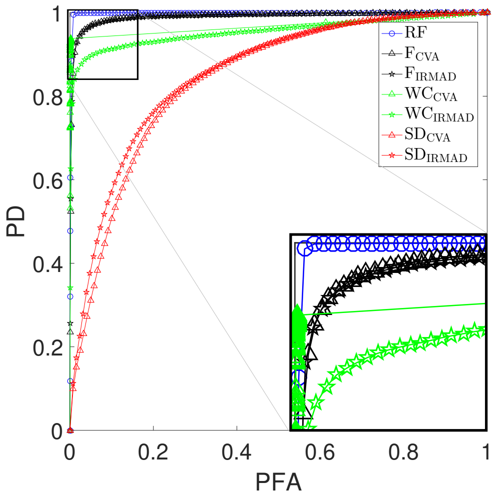


Figure 3: Scenario $\mathcal{S}_{\mathrm{sd}}$ (LR-HS and HR-HS): ROC curves.

4.4.4. Scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$ (LR-HS and HR-MS)

Following the same strategy, as for the previous scenarios, but now considering a LR-HS and a HR-MS observed images, the results lead to very similar overall performance. This scenario represents a more difficult situation for CD than the two previous ones due to the differences in both resolutions between observed images. The ROC curves are displayed in Fig. 4 with corresponding metrics in Table 3 (7th and 8th rows). Comparing curves in Fig. 4 shows that the proposed RF method offers reliable detection accuracy. The performance of the proposed RF method can be explained by the estimated HR-HS change maps, while the F approach estimates changes from HR-MS images and the WC approach estimates changes from LR-MS images. Again, in this scenario, the SD and DS methods, combined with CVA and IRMAD, fail to provide competitive results.

4.4.5. Scenario $\mathcal{S}^{\mathrm{u}}_{\mathrm{dd}}$ (LR-MS and HR-HS)

This scenario represents an unbalanced counterpart of Scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$ since one of the image is of lower spatial and spectral resolutions than the other. The corresponding ROC curves are depicted in Fig. 5 with quantitative metrics reported in Table 3 (last two rows). As for the other scenarios, the proposed RF method demonstrates its superiority with respect to the CVA- and IRMAD-based WC methods which operate on a pair of LR-MS images after degrading the HR-HS image, yielding to a significant loss of information. More interestingly, this scenario exhibits the low performance of the F method. Indeed, in this case, the fusion step fails to provide a relevant latent image since most of the information is brought by the HR-HS image.

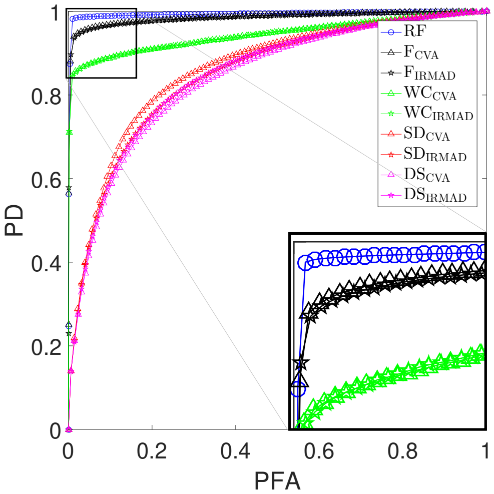


Figure 4: Scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$ (HR-MS and LR-HS): ROC curves.

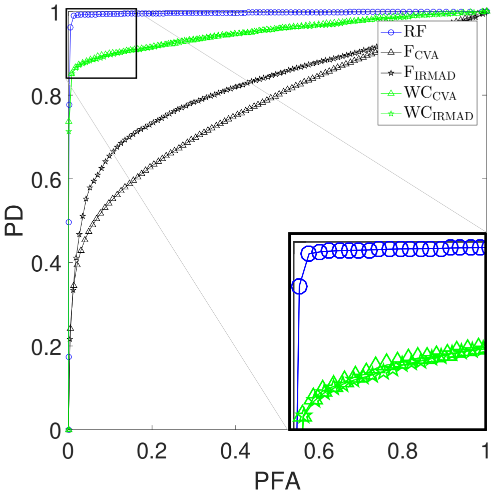


Figure 5: Scenario $\mathcal{S}^{\mathrm{u}}_{\mathrm{dd}}$ (LR-MS and HR-HS): ROC curves.

5. Experiments on real images

5.1. Reference images

To illustrate the performance of the proposed algorithmic framework using real multi-band optical images on each specific scenario discussed in Section 2.3, observed images from 4 largely studied open access multi-band sensors have been chosen, namely Landsat-8 from [58], Sentinel-2 from [59], Earth observing-1 Advanced Land Imager (EO-1 ALI) [60] and Airborne Visible Infrared Imaging Spectrometer (AVIRIS) from [61]. These images have been acquired over the same geographical location, i.e., the Mud Lake region in Lake Tahoe, CA, USA between June 8th, 2011 and October 29th, 2016. Unfortunately, no ground truth information is available for the chosen image pairs, as experienced in numerous experimental situations [15]. However, this region is characterized by interesting natural meteorological changes, e.g., drought of the Mud Lake, snow falls and vegetation growth, occurring along the seasons which help to visually infer the major changes between two dates and to assess the relevance of the detected changes. All considered images have been manually geographically and geometrically aligned to fulfill the requirements imposed by the considered CD setup.

In addition to the data provided by these sensors, complementary images have been synthetically generated by considering so-called *virtual* sensors derived from the real ones. The specifications of these virtual sensors, summarized in Fig. 6, are chosen such that all applicative scenarios previously discussed can be diversely represented. They are met by selecting a subset of the initial spectral bands or by artificially degrading the spatial resolution of the real sensors.

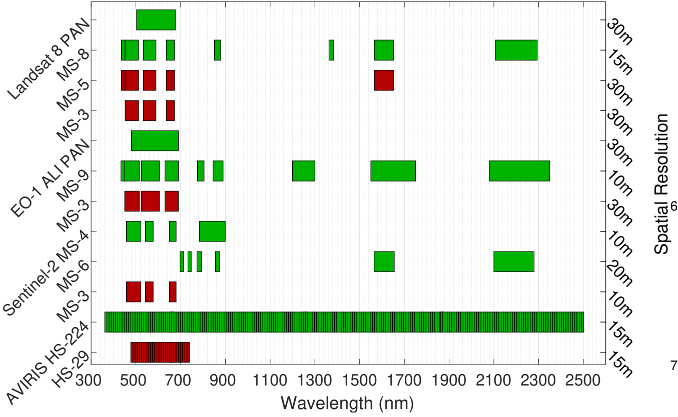


Figure 6: Spectral and spatial characteristics of real (green) and virtual (red) sensors.

5.1.1. *Landsat-8 images*

Landsat-8 is the eighth Earth observation satellite series of the US LANDSAT Program [58], launched on February 11th, 2013 with 16-days revisiting period. It is equipped with the Operational Land Imager (OLI) and the Thermal InfraRed Sensor (TIRS). In the conducted experiments, 3 sets of real images acquired at the dates 10/18/2013, 04/15/2015 and 09/22/2015 have been considered. For each acquisition, Landsat-8 provides

- one panchromatic image over the spectral range 0.503–0.676μm (band ♯8) at a 15m spatial resolution (denoted PAN),
- one multispectral image of 8 spectral bands (bands ♯1–♯7 and ♯9) at a 30m resolution (denoted MS-8).

To enrich this experimental study, as explained above, these real images are complemented with the following virtually acquired images

- one multispectral image of 5 spectral bands (bands ♯1–♯4 and ♯7) at a 30m spatial resolution (denoted MS-5),
- one red-green-blue (RGB) multispectral image of 3 spectral bands (bands ♯2–♯4) at a 30m spatial resolution (denoted MS-3).

5.1.2. *Sentinel-2 images*

Sentinel-2 is a series of two identical satellites for Earth observation missions developed by ESA [59] as part of the Copernicus Program launched in 2015 and 2017 with 5-days revisiting period. The multi-spectral instrument embedded on each platform is composed of two different sensors for acquisition in the visible and infrared spectral domains, respectively. The actual dataset used in the experiments is composed of two images acquired on 04/12/2016 and 10/29/2016 and, for each real scene, among all available spectral bands, one considers

- one multispectral image of 4 visible/near infrared (VNIR) spectral bands (bands ♯2–♯4 and ♯8) at a 10m spatial resolution (denoted MS-4)
- one multispectral image of 6 short wave infrared spectral range (SWIR) spectral bands (bands ♯5–♯8a and ♯11–♯12) at a 20m spatial resolution (denoted by MS-6)

and one additional virtually image, namely,

- one RGB multispectral image of 3 spectral bands (bands ♯2–♯4) at a 10m spatial resolution (denoted by MS-3).

5.1.3. *EO-1 ALI images*

Operated by NASA, EO-1 ALI is a Earth observation satellite part of the New Millennium Program launched in 2000 with 16-days repeat cycle and decommissioned in 2017 [60]. The main embedded sensor Advanced Land Imager (ALI) is complemented with the Hyperion spectrometer and the Linear Etalon Imaging Spectrometer Array (LEISA) for atmospheric correction. The considered dataset corresponds to 2 acquisition dates, 06/08/2011 and 08/04/2011, for

- one panchromatic image over the spectral range 0.48–0.69μm (band ♯1) at a 10m spatial resolution (denoted PAN),
- one multispectral image of 9 spectral bands (bands ♯2–♯10) at a 30m resolution (denoted MS-9),

in addition to the virtual acquisition of

- one RGB multispectral image of 3 spectral bands (bands ♯3–♯5) at a 30m spatial resolution (denoted MS-3).

5.1.4. *AVIRIS images*

AVIRIS is the second aircraft embedding an image spectrometer developed by Jet Propulsion Laboratory (JPL) for Earth remote sensing [61]. It delivers calibrated images in 224 contiguous 10nm-width spectral channels ranging from 0.4μm to 2.5μm. Since it is an airborne-dependent system, the spatial resolution is not a priori fixed and is designed for each individual acquisition. The dataset considered in the conducted experiments is composed by two real images acquired on 04/10/2014 and 09/19/2014. For each scene, one considers

- the original hyperspectral image of 224 spectral bands at a 15m spatial resolution (denoted HS-224)
- one virtual hyperspectral image of 29 spectral bands (corresponding to the RGB domain) at a 15m spatial resolution (denoted HS-29)

5.2. *Design of the spatial and spectral degradations*

The proposed model requires the prior knowledge of spectral and spatial degradation matrices $\mathbf{L}$ and $\mathbf{R} = \mathbf{BS}$, respectively. Regarding the spectral degradation matrices required in each simulation scenario, they can be easily derived from the intrinsic sensor characteristics freely available by averaging the spectral bands corresponding to the prescribed response. Conversely, the spatial degradation is not a sensor specification. It depends not only on the considered systems as well as external factors but also on the targeted resolution of the fused image. This work relies on commonly adopted assumptions by considering $\mathbf{R}$ as a Gaussian blur and by adjusting the down-sampling factor in $\mathbf{S}$ as an integer value corresponding to the relative ratio between spatial resolution of both observed images.

	Image ♯1			**Image ♯2**		
	Sensor	**Spatial resol.**	**Spectral resol.**	**Sensor**	**Spatial resol.**	**Spectral resol.**
$\mathcal{S}_{\mathrm{ss}}$	Landsat-8	15	PAN	Landsat-8	15	PAN
	Landsat-8	30	MS-3	Landsat-8	30	MS-3
	AVIRIS	15	HS-224	AVIRIS	15	HS-224
$\mathcal{S}_{\mathrm{sd}}$	Sentinel-2	10	MS-3	EO-1 ALI	30	MS-3
	Sentinel-2	10	MS-3	Landsat-8	30	MS-3
$\mathcal{S}_{\mathrm{s}\delta}$	EO-1 ALI	10	PAN	Landsat-8	15	PAN
$\mathcal{S}_{\mathrm{ds}}$	EO-1 ALI	10	PAN	Sentinel-2	10	MS-3
	Landsat-8	15	PAN	AVIRIS	15	HS-29
$\mathcal{S}_{\delta\mathrm{s}}$	Landsat-8	30	MS-8	EO-1 ALI	30	MS-9
$\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$	Landsat-8	15	PAN	Landsat-8	30	MS-3
	EO-1 ALI	10	PAN	Landsat-8	30	MS-3
	Landsat-8	15	PAN	EO-1 ALI	30	MS-3
$\mathcal{S}_{\delta\mathrm{d}}$	Landsat-8	30	MS-5	Sentinel-2	10	MS-4
$\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$	EO-1 ALI	30	MS-3	AVIRIS	15	HS-29
	Landsat-8	30	MS-3	AVIRIS	15	HS-29
$\mathcal{S}_{\mathrm{d}\delta}$	Sentinel-2	10	MS-3	Landsat-8	15	PAN
$\mathcal{S}_{\delta\delta}$	Sentinel-2	20	MS-6	EO-1 ALI	30	MS-9

Table 4: Pairs of real and/or virtual images, and their spatial and spectral characteristics, used for each applicative scenario.

5.3. *Results*

The following paragraphs discuss the CD performance of the WC, F and proposed RF methods for each applicative scenario detailed in paragraph 2.3 (see also Table 1). The WC and F methods are coupled with CVA as a detector since, from the results reported in Section 4, there is no a clear advantage of using IRMAD. The SD and DS methods are not considered for comparison since these results also showed that they were not able to provide interesting results w.r.t. the WC method. Depending on the considered scenario, pairs of real and/or virtual images described in paragraph 5.1, associated to different acquisition times but common acquisition location, are selected. Table 4 summarizes the pair of observed images provided by the real and/or virtual sensors used in each scenario. Note that several combinations of images can be made for some scenarios.

5.3.1. *Scenario* $\mathcal{S}_{\mathrm{ss}}$

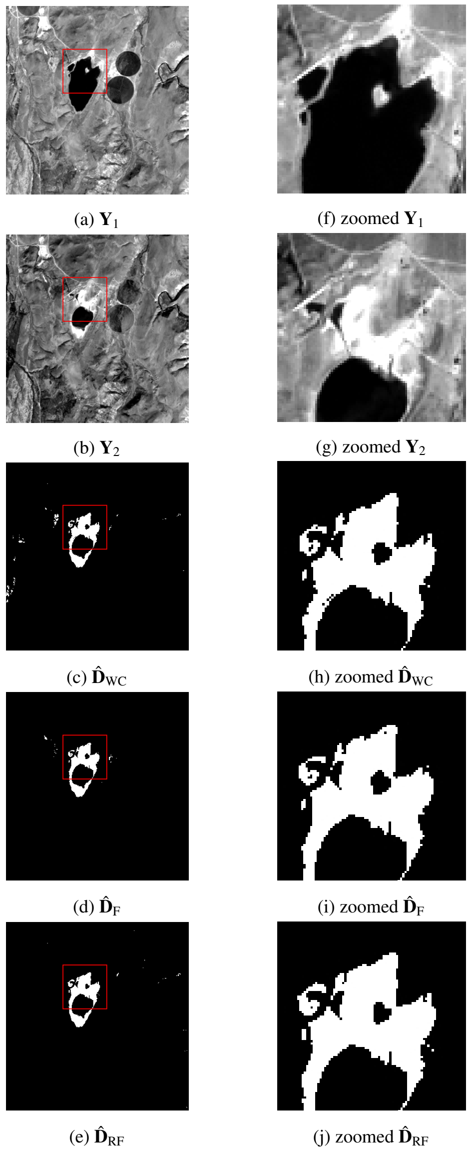
In the first scenario, CD is conducted on a pair of images of same spatial and spectral resolutions, which corresponds to the most favorable and commonly considered CD framework. Figures 7 to 9 present the CD binary masks recovered by the proposed RF-based CD method as well as by the F and WC methods for three pairs of panchromatic, multispectral and hyperspectral images, respectively. Note that, in this scenario, the WC boils down to conduct CVA directly on the observed images since they already share the same spatial and spectral resolutions and, thus, do not require to be degraded before pixel-wise comparison. Conversely, the *fusion* step of the F-based CD method only consists in a denoising. These change maps show that all CD methods detect the most significant changes, in particular the draught of the lake. However, for all configurations, the proposed RF method visually presents CD maps with better detection/false alarm rates, followed by the F and WC methods. This can be explained by the fact that both RF and F methods denoise the observed image while, in addition, the RF method jointly estimates the change image $\Delta\mathbf{X}$. Conversely, the WC method directly uses the observed images to derive the change image: the noise may introduce false alarms and misdetections. This is particularly visible in Fig. 9 depicting the results obtained from an hyperspectral image, known to be of lower signal-to-noise ratio.

5.3.2. *Scenario* $\mathcal{S}_{\mathrm{ds}}$

This CD scenario deals with observed images of same spatial resolution but different spectral resolutions. Figures 10 and 11 illustrate two possible situations and show the CD results of the proposed RF-based CD compared with the F and WC methods. In this scenario, similarly to scenario $\mathcal{S}_{\mathrm{ss}}$, both RF and F estimated change maps have the same spatial resolution as the observed image pair, which means that there is no loss of spatial resolution. Moreover, both methods deliver change maps estimated from $\Delta\mathbf{X}$ and the predicted pseudo-observed image, respectively, with the highest spectral resolution of the two observed images. Conversely, the WC method conducts CVA on a pair of images after spectral degradation to reach the lowest spectral resolution, which possibly results in loss of significant information. The resulting impact on the change/no-change decision is the visual reduction of false alarm rate for the RF and F methods, even if all change maps have the same spatial resolution.

5.3.3. *Scenario* $\mathcal{S}_{\delta\mathrm{s}}$

Scenario $\mathcal{S}_{\delta\mathrm{s}}$ is more challenging than the previous scenario $\mathcal{S}_{\mathrm{ds}}$, since it handles a pair of images with non-superimposable spectral sampling grids. This configuration requires the simultaneous use of two spectral degradation matrices in the proposed RF method. Fig. 12 provides one instance of this scenario. Due to the presence of non-overlapping bands, before conducting CVA, the WC requires to ignore the spectral bands which are not common to the two observed images. Conversely, by fully exploiting the whole available spectral information, the proposed method combines the overlapped bands and the non-overlapping bands to estimate a change image $\Delta\hat{\mathbf{X}}$ of higher



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

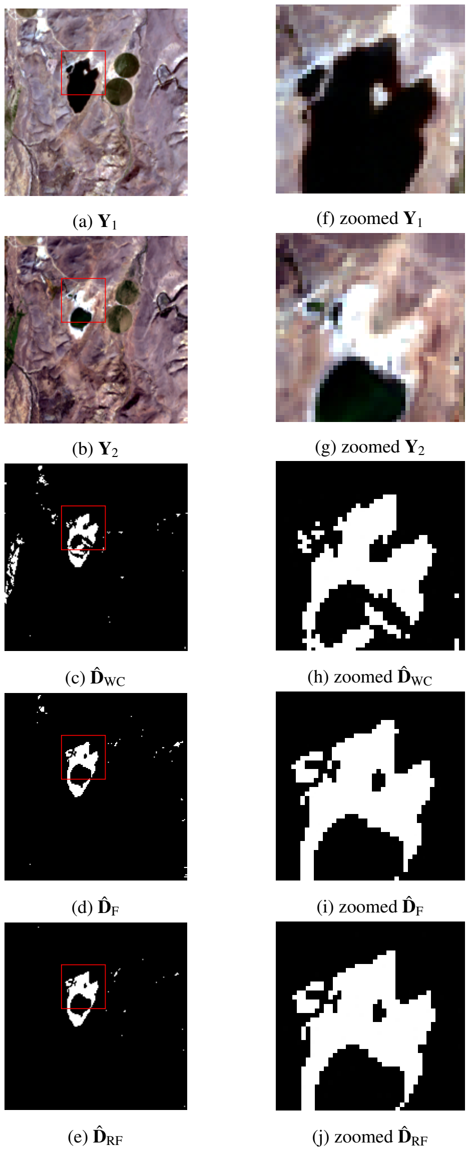
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 7: Scenario $\mathcal{S}_{\mathrm{ss}}$: (a) Landsat-8 15m PAN observed image $\mathbf{Y}_1$ acquired on 04/15/2015, (b) Landsat-8 15m PAN observed image $\mathbf{Y}_2$ acquired on 09/22/2015, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 15m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 15m PAN observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 15m PAN change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

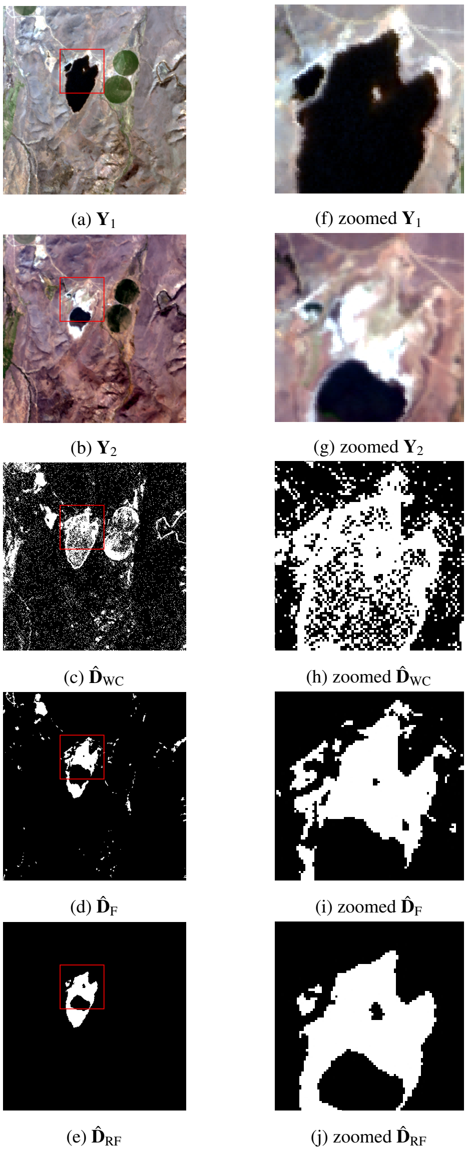
(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 8: Scenario $\mathcal{S}_{\mathrm{ss}}$: (a) Landsat-8 30m MS-3 observed image $\mathbf{Y}_1$ acquired on 04/15/2015, (b) Landsat-8 30m MS-3 observed image $\mathbf{Y}_2$ acquired on 09/22/2015, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m MS-3 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 30m MS-3 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 30m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

spectral resolution than the two observed images. This higher amount of information leads to visually more consistent results in Fig. 12(e). The F method, on one hand, outperforms the WC method by estimating the change map with the highest spectral resolution of the observed image pair, in this case 30m MS-9. On the other hand, the F-based CD method is not able to exploit the high resolution spectral content as the proposed RF method.

5.3.4. Scenario $\mathcal{S}_{\mathrm{sd}}$

In scenario $\mathcal{S}_{\mathrm{sd}}$, corresponding to the reverse situation encountered in scenario $\mathcal{S}_{\mathrm{ds}}$, observed images share the same spectral resolution but with different spatial resolution. Figures 13 and 14 present the results obtained for two possible real situations. Note that change maps obtained by the F and RF methods are of higher spatial resolutions than the ones estimated by the WC approach. Thus, this scenario is the first to illustrate a very important advantage of these two approaches, i.e., the higher spatial resolutions of the change maps. In scenario $\mathcal{S}_{\mathrm{ds}}$, the results have shown that the loss of spectral information inherent to the WC approach leads to an increase of false alarms and misdetections. Here, the loss of spatial information when applying the WC approach results in an inaccurate localization of the changes.


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

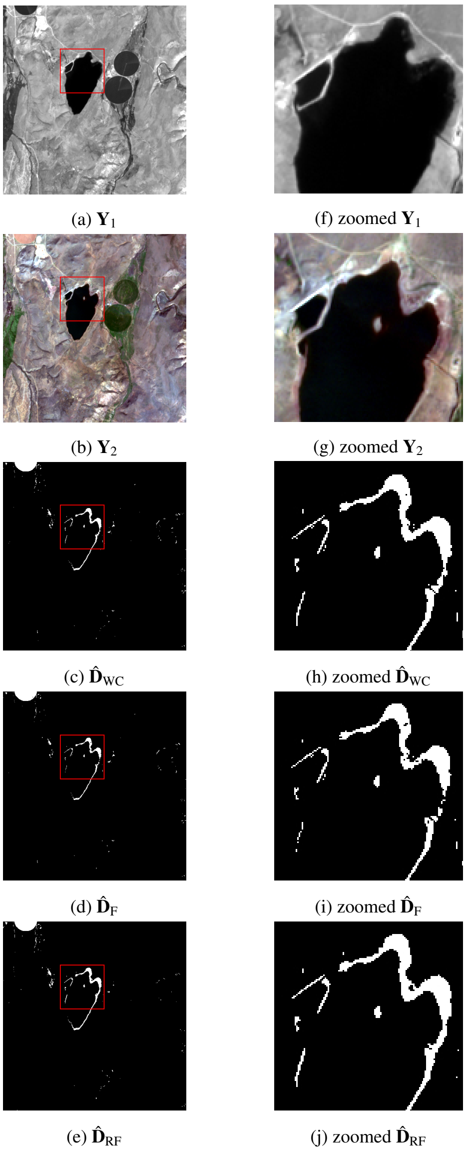
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 9: Scenario $\mathcal{S}_{\mathrm{ss}}$: (a) AVIRIS 15m HS-224 observed image $\mathbf{Y}_1$ acquired on 04/10/2014, (b) AVIRIS 15m HS-224 observed image $\mathbf{Y}_2$ acquired on 09/19/2014, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 15m HS-29 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 15m HS-29 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 30m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

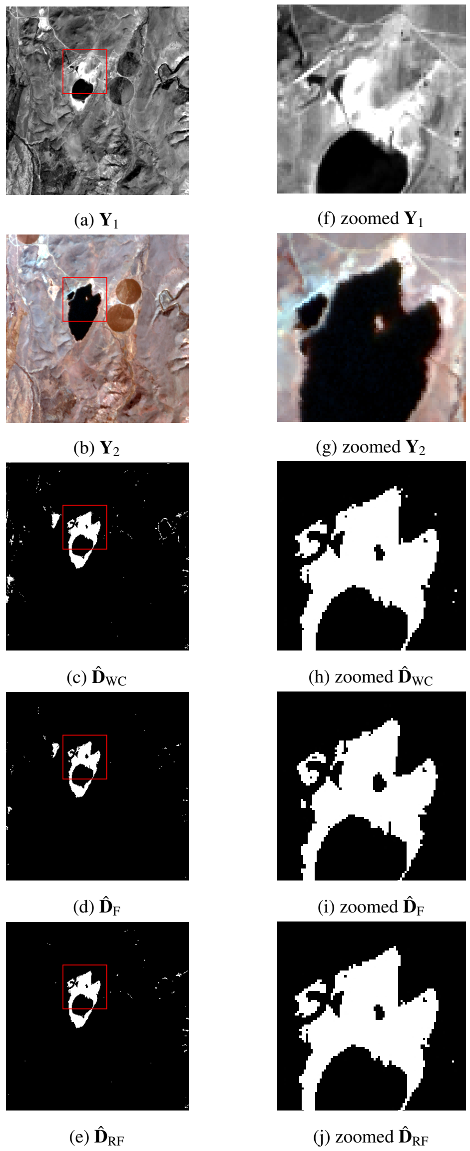
(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 10: Scenario $\mathcal{S}_{\mathrm{ds}}$: (a) EO-1 ALI 10m PAN observed image $\mathbf{Y}_1$ acquired on 06/08/2011, (b) Sentinel-2 10m MS-3 observed image $\mathbf{Y}_2$ acquired on 04/12/2016, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 10m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m MS-3 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 10m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

5.3.5. *Scenario $\mathcal{S}_{\mathrm{s}\delta}$*

This scenario is more challenging than the previously discussed scenario $\mathcal{S}_{\mathrm{sd}}$. Indeed, it handles two observed images of spatial resolutions related by a non-integer downsampling ratio (i.e., on non-superimposable spatial sampling grids), which precludes the use of a unique spatial degradation matrix in the RF-based CD method. As detailed in Appendix B.5, super-resolutions are conducted during the fusion and correction steps of the AM algorithm, which leads to a change image $\Delta\hat{\mathbf{X}}$ with a spatial resolution higher than the ones of the two observed images (defined as the greatest common divisor of the resolutions). For instance, Fig. 15 illustrates one possible configuration for which the observed images $\mathbf{Y}_1$ and $\mathbf{Y}_2$, depicted in Figures 15(a) and 15(b), are of 15m and 10m spatial resolutions, respectively. Thus the change image $\Delta\hat{\mathbf{X}}$ and change mask $\hat{\mathbf{d}}_{\mathrm{RF}}$ estimated by the RF method are at a 5m resolution. The F method produces a change map with the highest spatial resolution of the two observed images, in this case, 10m. Conversely, the WC method provides a change map at a spatial resolution equal to the least common multiple, which is, in this case, 30m. The significantly higher spatial resolution of the change map is clear in Fig. 15(e).


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

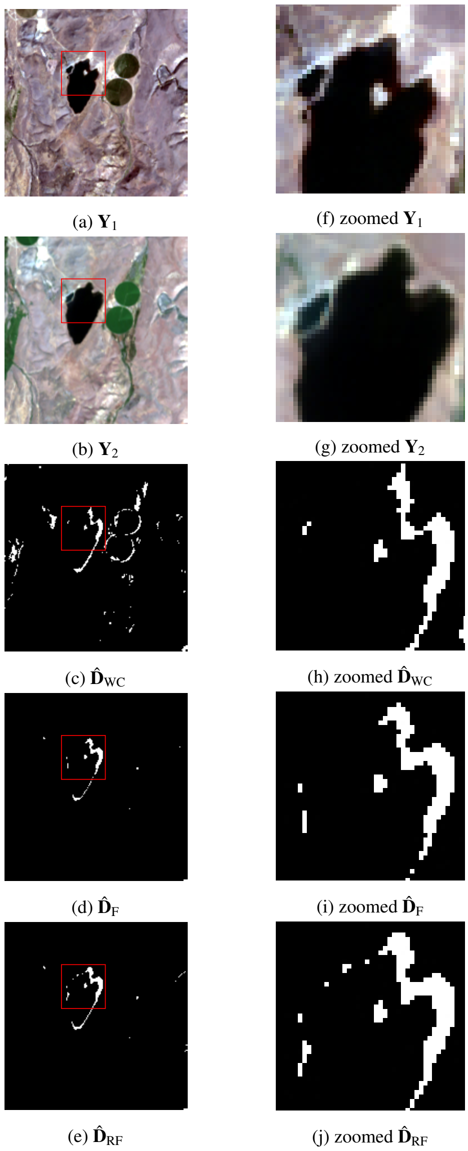
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\text{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\text{WC}}$

(d) $\hat{\mathbf{D}}_{\text{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\text{F}}$

(e) $\hat{\mathbf{D}}_{\text{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\text{RF}}$

Figure 11: Scenario $\mathcal{S}_{\text{ds}}$: (a) Landsat-8 15m PAN observed image $\mathbf{Y}_1$ acquired on 09/22/2015, (b) AVIRIS 15m HS-29 observed image $\mathbf{Y}_2$ acquired on 04/10/2014, (c) change mask $\hat{\mathbf{D}}_{\text{WC}}$ estimated by the WC approach from a pair of 15m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\text{F}}$ estimated by the F approach from a pair of 15m HS-29 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\text{RF}}$ estimated by the proposed RF approach from a 15m HS-29 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\text{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\text{WC}}$

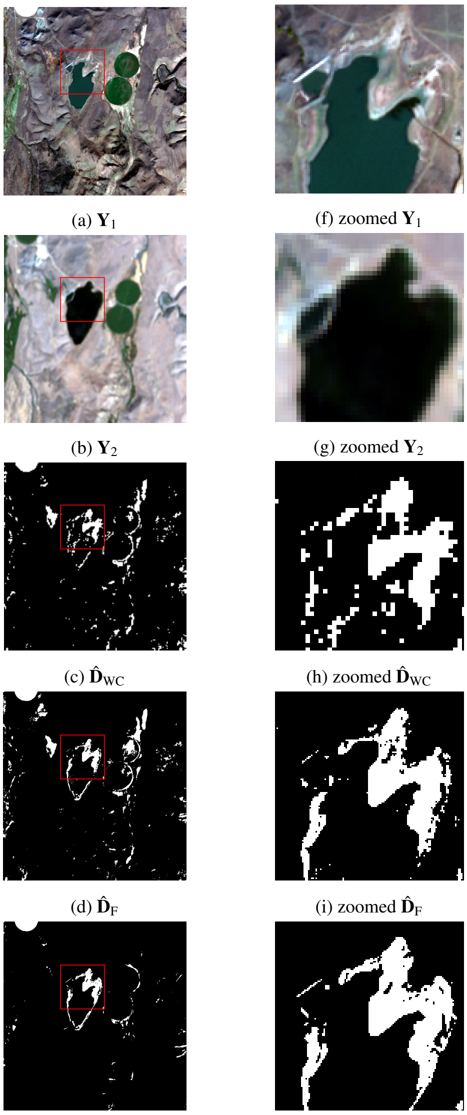
(d) $\hat{\mathbf{D}}_{\text{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\text{F}}$

(e) $\hat{\mathbf{D}}_{\text{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\text{RF}}$

Figure 12: Scenario $\mathcal{S}_{\delta\text{s}}$: (a) Landsat-8 30m MS-8 observed image $\mathbf{Y}_1$ acquired on 04/15/2015, (b) EO-1 ALI 30m MS-9 observed image $\mathbf{Y}_2$ acquired on 06/08/2011, (c) change mask $\hat{\mathbf{D}}_{\text{WC}}$ estimated by the WC approach from a pair of 30m MS-7 degraded images, (d) change mask $\hat{\mathbf{D}}_{\text{F}}$ estimated by the F approach from a pair of 30m MS-9 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\text{RF}}$ estimated by the proposed RF approach from 30m MS-10 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

5.3.6. Scenario $\mathcal{S}^{\text{b}}_{\text{dd}}$

This scenario has been deeply investigated in [18] who conducted a comprehensive analysis of the performance of the RF-based method. This scenario corresponds to a more difficult CD investigation than all previous ones since the pair of observed images have not the same spatial neither spectral resolutions. As a consequence, the conventional WC approach is constrained to compare a spatially degraded version of one observed image with a spectrally degraded version of the other observed image. Irredeemably, these degradations result in a loss of spectral information, essential to assess the presence of change, and a loss of spatial information, required to accurately localize the possible changes. On the contrary, the proposed RF method is able to derive the change mask from a change image characterized by the best of the spectral and spatial resolution of the observed images. The F method, while it estimates the change map with the same spatial resolution as the RF method, proceeds on a lower spectral resolution image. This results in a higher false-alarm/detection rate than the one obtained with the RF method, yet lower than the one obtained with the WC method. Figures 16 to 18 depict the CD results obtained for three common configurations and illustrate the superiority of the proposed RF-based CD method.



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

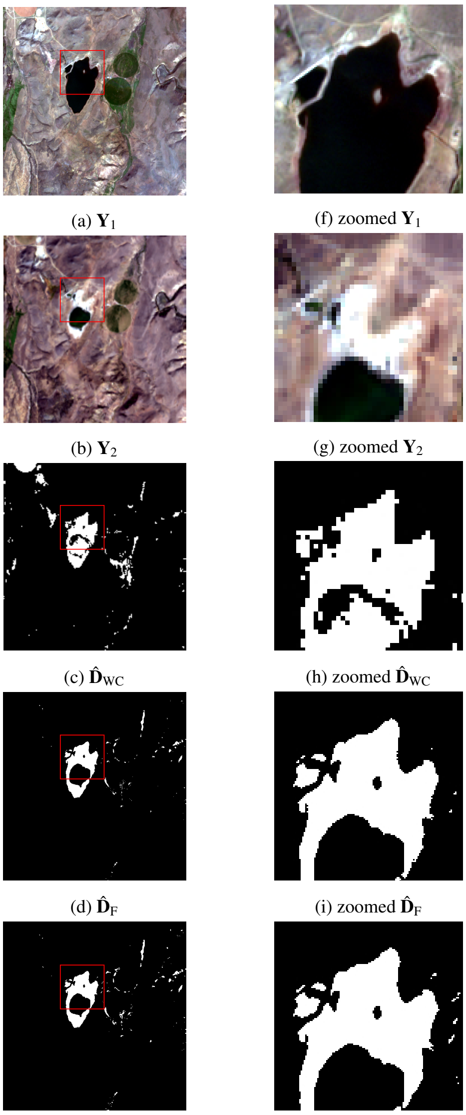
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 13: Scenario $\mathcal{S}_{\mathrm{sd}}$: (a) Sentinel-2 10m MS-3 observed image $\mathbf{Y}_1$ acquired on 10/29/2016, (b) EO-1 ALI 30m MS-3 observed image $\mathbf{Y}_2$ acquired on 08/04/2011, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m MS-3 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m MS-3 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 10m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

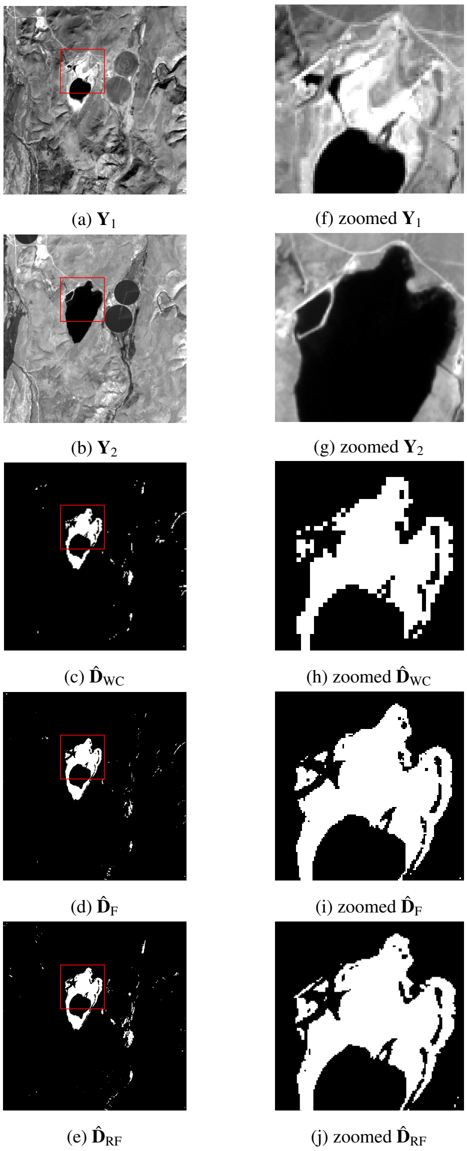
Figure 14: Scenario $\mathcal{S}_{\mathrm{sd}}$: (a) Sentinel-2 10m MS-3 observed image $\mathbf{Y}_1$ acquired on 04/12/2016, (b) Landsat-8 30m MS-3 observed image $\mathbf{Y}_2$ acquired on 09/22/2015, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m MS-3 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m MS-3 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 10m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

5.3.7. *Scenario $\mathcal{S}_{\delta\mathrm{d}}$*

This scenario corresponds to a modified instance of the previous scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$ (images of different spatial and spectral resolutions) where the two spectral sampling grids cannot be superimposed. The results obtained for one configuration are depicted in Fig. 19. In this case, the change image $\Delta\hat{\mathbf{X}}$ is characterized by a spatial resolution equal to the highest spatial resolution of the observed images and with a spectral resolution higher than the spectral resolution of both observed images. The F-based CD method produces a change map with the highest spatial resolution but with lower spectral resolution than the RF method, in this example 10m MS-4. Once again, the results show the accuracy of the proposed RF method in terms of detection and spatial resolution of the estimated change map.

5.3.8. *Scenario $\mathcal{S}^{\mathrm{u}}_{\mathrm{dd}}$*

This scenario also handles images which do not share the same spatial neither spectral resolutions. However, contrary to scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$, this scenario considers one of the two images of higher spatial and spectral resolutions. Again, the WC is expected to be less reliable (in terms of decision and localization) due to the loss of spectral and spatial information consecutive to the degradations before conducting CVA. Figures 20 and 21 present the results obtained from two possible real configura-


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

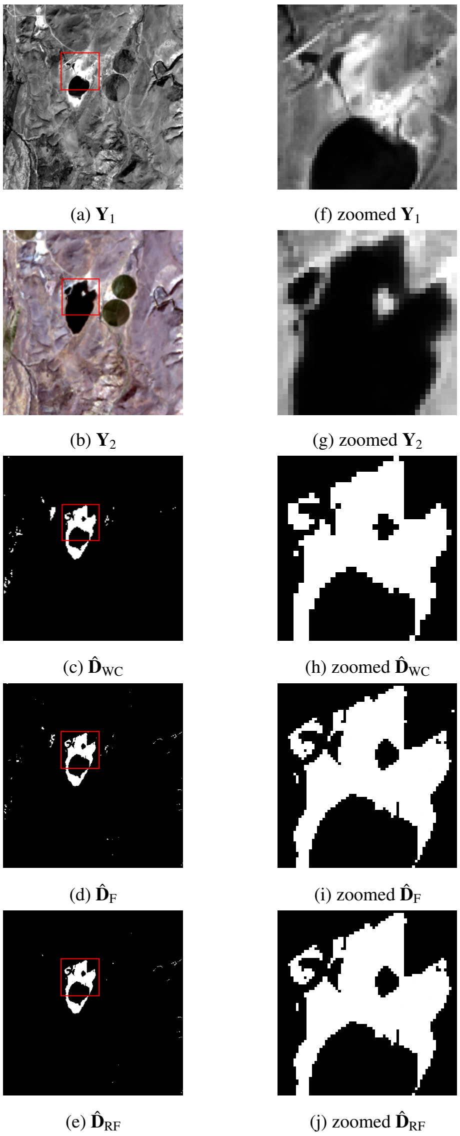
(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 15: Scenario $\mathcal{S}_{\mathrm{s}\delta}$: (a) Landsat-8 15m PAN observed image $\mathbf{Y}_1$ acquired on 10/18/2013, (b) EO-1 ALI 10m PAN observed image $\mathbf{Y}_2$ acquired on 08/04/2011, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m PAN observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from 5m PAN change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

tions. As expected the RF-based CD method provides visually more satisfactory results. The F-based method provides change maps of same spatial and spectral resolutions as the RF-based method. Nevertheless, as it compares an estimated image with the raw observed image, the SNR difference between them may increase the false alarm/detection rate compared to the RF-based method. Besides, as shown in Fig. 21, the WC method is unable to accurately localize the change due to the lake drought from the pair of multispectral and hyperspectral images.

5.3.9. Scenario $\mathcal{S}_{\mathrm{d}\delta}$

This scenario corresponds to a more challenging context than scenario $\mathcal{S}^{\mathrm{u}}_{\mathrm{dd}}$ since it handles images with non-superimposable spatial sampling grids. As before, the change image $\Delta\hat{\mathbf{X}}$ and the binary change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the RF-based method are defined at a higher spatial resolution than the observed images. Fig. 22 presents one example of this scenario. The F-based CD method outperforms the WC method because it estimates a change map with the highest spatial resolution of the pair of observed images. Nevertheless, this change map is of lower spectral resolution than the one estimated by the RF-based method. This explains the observed differences on the false alarm/detection rates for the three methods.


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

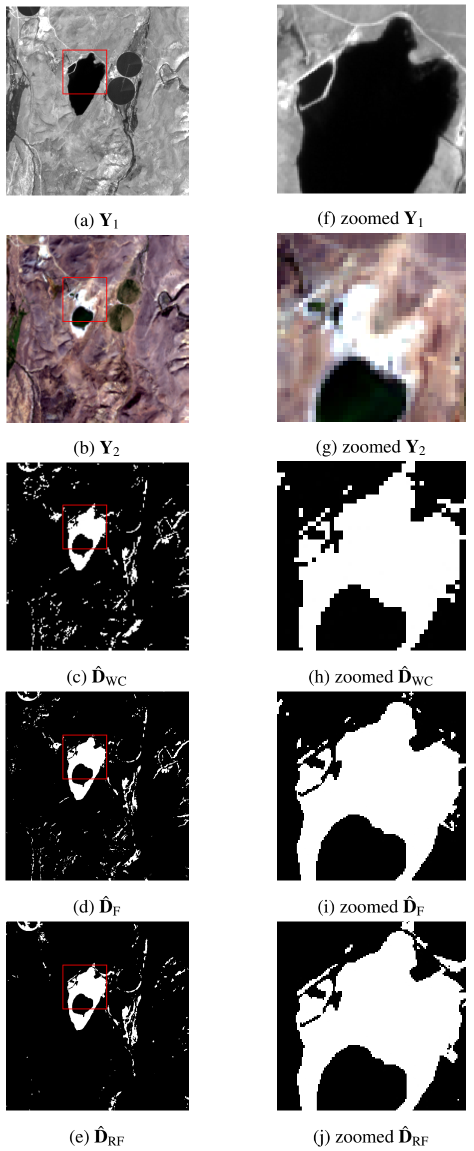
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 16: Scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$: (a) Landsat-8 15m PAN observed image $\mathbf{Y}_1$ acquired on 09/22/2015, (b) Landsat-8 30m MS-3 observed image $\mathbf{Y}_2$ acquired on 04/15/2015, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 15m PAN observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 15m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

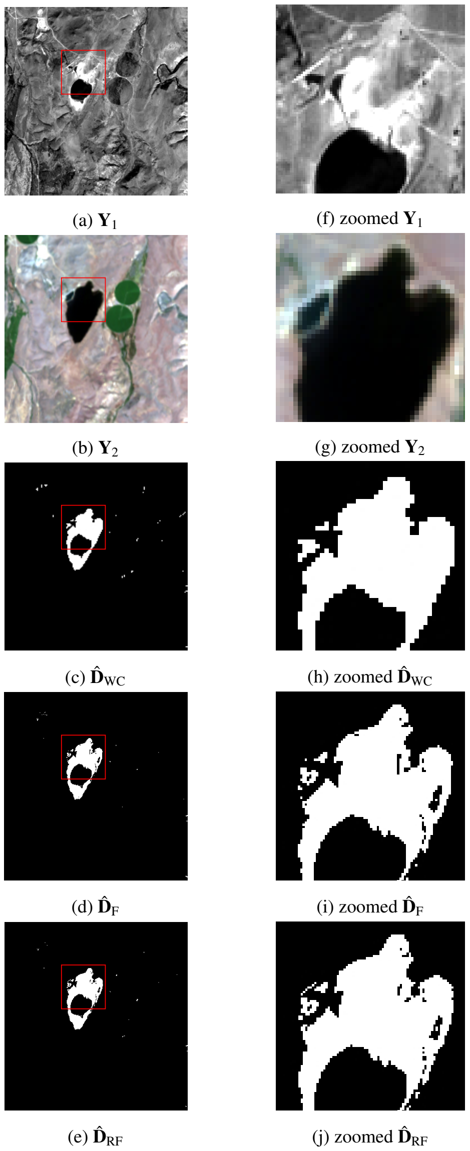
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 17: Scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$: (a) EO-1 ALI 10m PAN observed image $\mathbf{Y}_1$ acquired on 06/08/2011, (b) Landsat-8 30m MS-3 observed image $\mathbf{Y}_2$ acquired on 09/22/2015, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m PAN observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 10m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

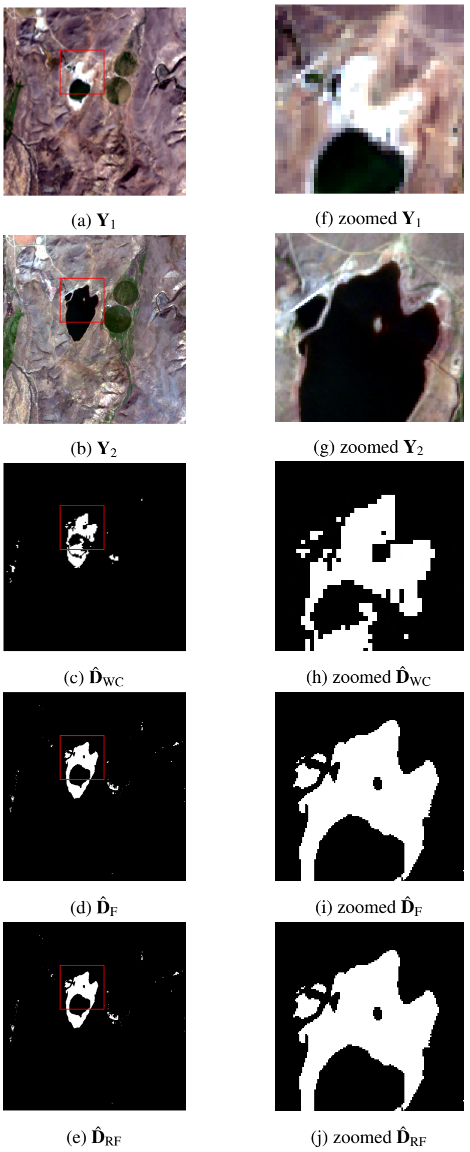
Figure 18: Scenario $\mathcal{S}^{\mathrm{b}}_{\mathrm{dd}}$: (a) Landsat-8 15m PAN observed image $\mathbf{Y}_1$ acquired on 09/22/2015, (b) EO-1 ALI 30m MS-3 observed image $\mathbf{Y}_2$ acquired on 06/08/2011, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 15m PAN observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 15m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

5.3.10. Scenario $\mathcal{S}_{\delta\delta}$

The last scenario combines the difficulties previously encountered: images of different spatial and spectral resolution, characterized by a non-integer relative downsampling factor and non-superimposable spectral sampling grids. As for scenarios $\mathcal{S}_{\mathrm{s}\delta}$ and $\mathcal{S}_{\mathrm{d}\delta}$, the change image $\Delta\hat{\mathbf{X}}$ and change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ recovered by the RF-based method is of higher spatial resolution than the two observed images. In addition, as for scenarios $\mathcal{S}_{\delta\mathrm{s}}$ and $\mathcal{S}_{\delta\mathrm{d}}$, the change image is also defined at a higher spectral resolution. The F-based CD method produces a change map $\hat{\mathbf{D}}_{\mathrm{F}}$ with the highest spatial resolution of the pair of observed images but with lower spectral resolution than the RF-based CD method. Conversely, the WC approach derives a change image of lower spatial and spectral resolutions before conducting CVA. Fig. 23 depicts the results obtained by all methods. In this particularly challenging scenario, the proposed approach demonstrates its superiority in recovering relevant changes and in localizing them accurately.

6. Conclusion

This paper derived a robust fusion framework to perform change detection between optical images of different spatial



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

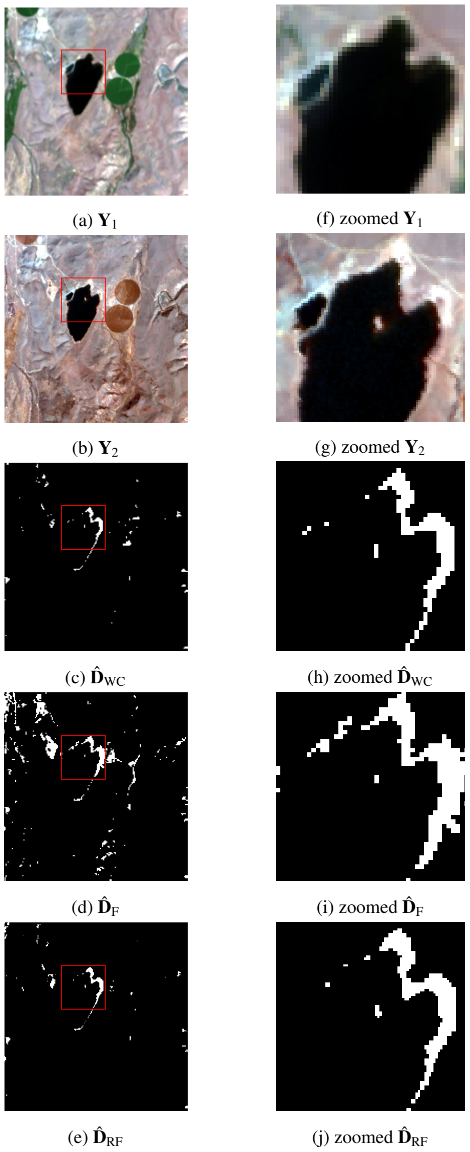
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 19: Scenario $\mathcal{S}_{\delta d}$: (a) Landsat-8 30m MS-5 observed image $\mathbf{Y}_1$ acquired on 09/22/2015, (b) Sentinel-2 10m MS-4 observed image $\mathbf{Y}_2$ acquired on 04/12/2016, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m MS-3 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m MS-4 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 10m MS-6 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).



(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

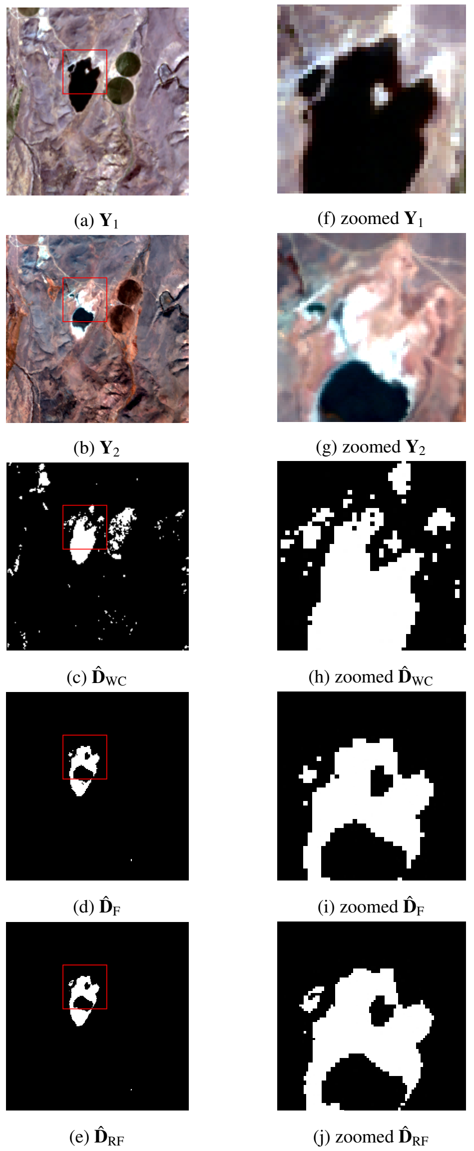
(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 20: Scenario $\mathcal{S}^{\mathrm{u}}_{\mathrm{dd}}$: (a) EO-1 ALI 30m MS-3 observed image $\mathbf{Y}_1$ acquired on 08/04/2011, (b) AVIRIS 15m HS-29 observed image $\mathbf{Y}_2$ acquired on 04/10/2014, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m MS-3 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 15m HS-29 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 15m HS-29 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

and spectral resolutions. The versatility of the proposed approach allowed all possible real scenarios to be handled efficiently. The technique was based on the definition of two high spatial and spectral resolution latent images related to the observed images via a double physically-inspired forward model. The difference image between these two latent images was assumed to be spatially sparse, implicitly locating the changes at a high resolution scale. Inferring these two latent images was formulated as an inverse problem which was solved within a 2-step iterative scheme. Depending on the considered scenario, these 2 steps can be interpreted as ubiquitous image processing problems (namely spatial super-resolution, spectral deblurring, denoising or multi-band image fusion) for which closed-form solutions or efficient algorithms had been recently proposed in the literature. A simulation protocol allowed the performance of the proposed technique in terms of detection and precision to be assessed and compared with the performance of state-of-the-art algorithms. Real images acquired by four different sensors were used to illustrate the accuracy and the flexibility of the proposed method, as well as its superiority with respect to the state-of-the-art change detection methods. Future works will assess the robustness of the proposed technique w.r.t. nonlinear effects (e.g., due to atmospheric effects, geometric and radiometric distortions). Detecting changes between optical and non-optical data is also under investigation and preliminary re-


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 21: Scenario $\mathcal{S}^{\mathrm{u}}_{\mathrm{dd}}$: (a) Landsat-8 30m MS-3 observed image $\mathbf{Y}_1$ acquired on 04/15/2015, (b) AVIRIS 15m HS-29 observed image $\mathbf{Y}_2$ acquired on 09/19/2014, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m MS-3 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 15m HS-29 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 15m HS-29 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

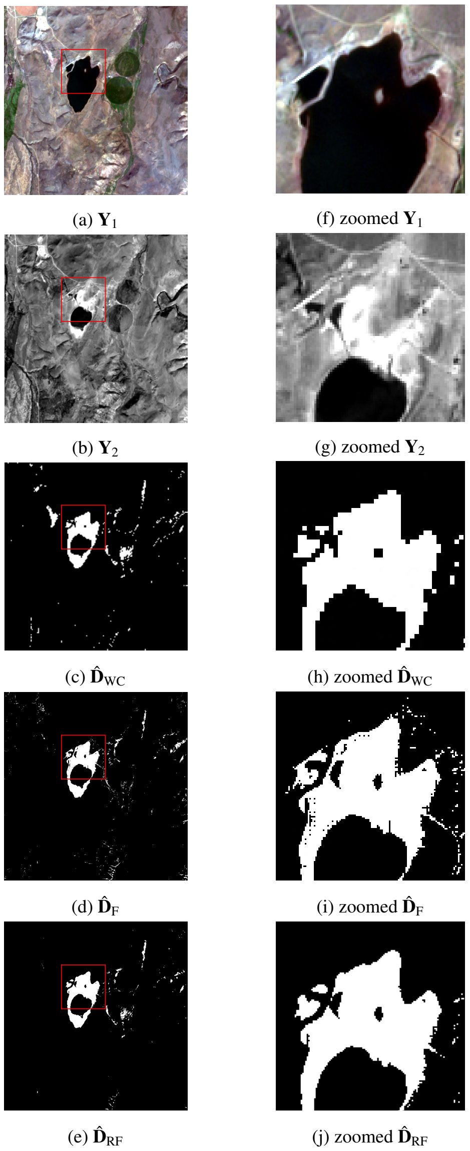
sults are discussed in [62].

Appendix A. Matrix normal distribution

The probability density function $p(\mathbf{X}|\mathbf{M}, \boldsymbol{\Sigma}_r, \boldsymbol{\Sigma}_r)$ of a matrix normal distribution $\mathcal{MN}_{r,c}(\mathbf{M}, \boldsymbol{\Sigma}_r, \boldsymbol{\Sigma}_c)$ is given by [63]

$$p\left(\mathbf{X}|\mathbf{M}, \boldsymbol{\Sigma}_r, \boldsymbol{\Sigma}_r\right) = \frac{\exp\left(-\frac{1}{2} tr\left[\boldsymbol{\Sigma}_c^{-1}(\mathbf{X}-\mathbf{M})^T \boldsymbol{\Sigma}_r^{-1}(\mathbf{X}-\mathbf{M})\right]\right)}{(2\pi)^{rc/2}|\boldsymbol{\Sigma}_c|^{r/2}|\boldsymbol{\Sigma}_r|^{c/2}}$$

where $\mathbf{M} \in \mathbb{R}^{r\times c}$ is the mean matrix, $\boldsymbol{\Sigma}_r \in \mathbb{R}^{r\times r}$ is the row covariance matrix and $\boldsymbol{\Sigma}_c \in \mathbb{R}^{c\times c}$ is the column covariance matrix.


(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$

(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$

(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$

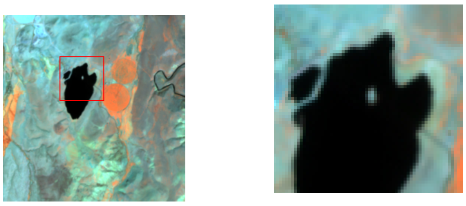
(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$

(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

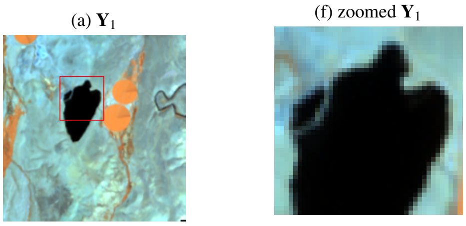
Figure 22: Scenario $\mathcal{S}_{\mathrm{d}\delta}$: (a) Sentinel-2 10m MS-3 observed image $\mathbf{Y}_1$ acquired on 04/12/2016, (b) Landsat-8 15m PAN observed image $\mathbf{Y}_2$ acquired on 09/22/2015, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 30m PAN degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 10m MS-3 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from 5m MS-3 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

Appendix B. Detailed implementations of the AM algorithm

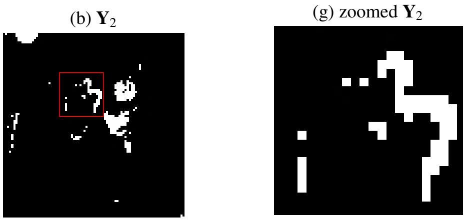
This appendix details the specific implementations of the generic AM algorithm introduced in paragraph 3.2 when considering each of the 10 scenarios identified in paragraph 2.3 and summarized in Table 1. For each scenario, we show that the fusion and correction steps can be cast as standard image processing tasks (see also Table 2).



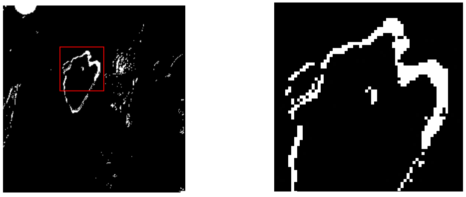
(a) $\mathbf{Y}_1$ (f) zoomed $\mathbf{Y}_1$



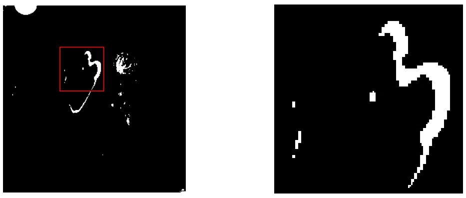
(b) $\mathbf{Y}_2$ (g) zoomed $\mathbf{Y}_2$



(c) $\hat{\mathbf{D}}_{\mathrm{WC}}$ (h) zoomed $\hat{\mathbf{D}}_{\mathrm{WC}}$



(d) $\hat{\mathbf{D}}_{\mathrm{F}}$ (i) zoomed $\hat{\mathbf{D}}_{\mathrm{F}}$



(e) $\hat{\mathbf{D}}_{\mathrm{RF}}$ (j) zoomed $\hat{\mathbf{D}}_{\mathrm{RF}}$

Figure 23: Scenario $\mathcal{S}_{\delta\delta}$: (a) Sentinel-2 20m MS-6 observed image $\mathbf{Y}_1$ acquired on 04/12/2016, (b) EO-1 ALI 30m MS-9 observed image $\mathbf{Y}_2$ acquired on 06/08/2011, (c) change mask $\hat{\mathbf{D}}_{\mathrm{WC}}$ estimated by the WC approach from a pair of 60m MS-4 degraded images, (d) change mask $\hat{\mathbf{D}}_{\mathrm{F}}$ estimated by the F approach from a pair of 20m MS-6 observed and predicted images and (e) change mask $\hat{\mathbf{D}}_{\mathrm{RF}}$ estimated by the proposed RF approach from a 10m MS-11 change image $\Delta\hat{\mathbf{X}}$. From (f) to (j): zoomed versions of the regions delineated in red in (a)–(e).

Appendix B.1. Scenario $\mathcal{S}_{\mathrm{ss}}$

Considering the degradation matrices specified in Table 1 for this scenario, the forward model (11) can be rewritten as

$$\mathbf{Y}_1 = \mathbf{X}_1 + \mathbf{N}_1 \tag{B.1a}$$

$$\mathbf{Y}_2 = (\mathbf{X}_1 + \Delta\mathbf{X}) + \mathbf{N}_2 \tag{B.1b}$$

As expected, for this scenario, the observed, latent and change images share the same spatial and spectral resolutions. The resulting objective function, initially in (9), is simplified as

$$\begin{aligned}\mathcal{J}_{\mathrm{ss}}(\mathbf{X}_1, \Delta\mathbf{X}) = &\frac{1}{2}\left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}(\mathbf{Y}_2 - (\mathbf{X}_1 + \Delta\mathbf{X}))\right\|_{\mathrm{F}}^2 \\ &+ \frac{1}{2}\left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{X}_1)\right\|_{\mathrm{F}}^2 \\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\|\Delta\mathbf{X}\|_{2,1}.\end{aligned} \tag{B.2}$$

The two steps of the AM algorithm are detailed below.

Appendix B.1.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

At the kth iteration of the AM algorithm, let assume that the current value of the change image is denoted by $\Delta\mathbf{X}^{(k)}$. As suggested in Section 3.2.1, a corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ that would be observed at time t_1 by the sensor S_2 given the image $\mathbf{Y}_2$ observed at time t_2 and the change image $\Delta\mathbf{X}^{(k)}$ can be introduced as

$$\tilde{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \Delta\mathbf{X}^{(k)}. \tag{B.3}$$

Updating the latent image $\mathbf{X}_1$ consists in minimizing w.r.t. $\mathbf{X}_1$ the partial function

$$\begin{aligned}\mathcal{J}_{\mathrm{ss}}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\mathrm{ss}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{X}_1)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned}$$

This formulation shows that recovering $\mathbf{X}_1$ in Scenario $\mathcal{S}_{\mathrm{ss}}$ reduces to a denoising problem from an observed image $\mathbf{Y}_1$ and a pseudo-observed image $\tilde{\mathbf{Y}}_2^{(k)}$. A closed-form solution of this ℓ_2-penalized least-square problem can be easily and efficiently computed.

Appendix B.1.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

Following the strategy proposed by [18], let $\check{\mathbf{Y}}_2^{(k)}$ denote the *predicted* image that would be observed by the sensor S_2 at time t_1 given the current state of the latent image $\mathbf{X}_1^{(k)}$. Since the two sensors share the same spatial and spectral characteristics, one has

$$\check{\mathbf{Y}}_2^{(k)} = \mathbf{X}_1^{(k)}. \tag{B.4}$$

Similarly to (5), the *predicted* change image can thus be defined as

$$\Delta\check{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \check{\mathbf{Y}}_2^{(k)}. \tag{B.5}$$

The objective function (B.2) w.r.t. $\Delta\mathbf{X}$ is then rewritten by combining (B.4) and (B.5) with (B.2), leading to

$$\begin{aligned}\mathcal{J}_{\mathrm{ss}}^{(\mathrm{c})}(\Delta\mathbf{X}) &\triangleq \mathcal{J}_{\mathrm{ss}}(\mathbf{X}_1^{(k)}, \Delta\mathbf{X}) \\ &= \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}\left(\Delta\check{\mathbf{Y}}_2^{(k)} - \Delta\mathbf{X}\right)\right\|_{\mathrm{F}}^2 + \gamma\|\Delta\mathbf{X}\|_{2,1}.\end{aligned} \tag{B.6}$$

Again, since the observed, latent and change images share the same spatial and spectral resolutions, this correction step reduces to a denoising task of the predicted change image $\Delta\check{\mathbf{Y}}_2^{(k)}$. With the particular CD-driven choice of $\phi_2(\cdot)$ in (10), minimizing $\mathcal{J}_{\mathrm{ss}}^{(\mathrm{c})}(\Delta\mathbf{X})$ is an $\ell_{2,1}$-penalized least square problem. Minimizing (B.6) also defines the proximal operator associated with the $\ell_{2,1}$-norm and can be directly achieved by applying a group-soft thresholding on the predicted change image $\Delta\check{\mathbf{Y}}_2^{(k)}$.

Appendix B.2. Scenario $\mathcal{S}_{\mathrm{ds}}$

In this scenario, the two observed images are of same spatial resolution (as for scenario $\mathcal{S}_{\mathrm{ss}}$) but with different spectral resolution, which precludes a simple comparison between pixels. For this scenario, the joint forward observation model derived from (11) can be written as

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1 + \mathbf{N}_1, \tag{B.7a}$$

$$\mathbf{Y}_2 = (\mathbf{X}_1 + \Delta\mathbf{X}) + \mathbf{N}_2, \tag{B.7b}$$

which results in the objective function

$$\begin{aligned}\mathcal{J}_{\mathrm{ds}}(\mathbf{X}_1, \Delta\mathbf{X}) = &\ \frac{1}{2}\left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}(\mathbf{Y}_2 - (\mathbf{X}_1 + \Delta\mathbf{X}))\right\|_{\mathrm{F}}^2 \\ &+ \frac{1}{2}\left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1)\right\|_{\mathrm{F}}^2 \\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned}$$

Within an AM algorithmic scheme, the two sub-problems of interest are detailed below.

Appendix B.2.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

The same strategy as for scenario $\mathcal{S}_{\mathrm{ss}}$ in paragraph Appendix B.1.1 is adopted. Since model (B.7b) is the the same as model (B.1b) associated with scenario $\mathcal{S}_{\mathrm{ss}}$, the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ is defined following (B.3). Then, updating the latent image $\mathbf{X}_1$ consists in minimizing the partial objective function

$$\begin{aligned}\mathcal{J}_{\mathrm{ds}}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\mathrm{ds}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.8}$$

This problem can be interpreted as a spectral deblurring of the observed image $\mathbf{Y}_1$ where the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ plays the role of prior information. Minimizing (B.8) can be easily conducted by computing the standard least square solution.

Appendix B.2.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

As both models (B.7b) and (B.1b) are the same, optimizing w.r.t $\Delta\mathbf{X}$ can be conducted following the procedure detailed in paragraph Appendix B.1.2 (i.e., denoising of the predicted change image).

Appendix B.3. Scenario $\mathcal{S}_{\delta\mathrm{s}}$

This scenario is similar to the Scenario $\mathcal{S}_{\mathrm{ds}}$ described in paragraph Appendix B.2. It relies on two images of same spatial resolution but of distinct spectral resolution. However, contrary to Scenario $\mathcal{S}_{\mathrm{ds}}$, this difference in spectral resolutions cannot be expressed with a unique spectral degradation matrix, e.g., due to non-superimposable spectral sampling grids. In this case the joint forward observation model is

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1 + \mathbf{N}_1. \tag{B.9a}$$

$$\mathbf{Y}_2 = \mathbf{L}_2(\mathbf{X}_1 + \Delta\mathbf{X}) + \mathbf{N}_2. \tag{B.9b}$$

with the resulting objective function

$$\begin{aligned}\mathcal{J}_{\delta\mathrm{s}}(\mathbf{X}_1, \Delta\mathbf{X}) = &\ \frac{1}{2}\left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}(\mathbf{Y}_2 - \mathbf{L}_2(\mathbf{X}_1 + \Delta\mathbf{X}))\right\|_{\mathrm{F}}^2 \\ &+ \frac{1}{2}\left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1)\right\|_{\mathrm{F}}^2 \\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned}$$

The choice of which degradation matrices applies to the change image $\Delta\mathbf{X}$ is driven by considering the matrix with larger number of bands, which results in a change image of higher spectral resolution. The associated sub-problems are described in what follows.

Appendix B.3.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

By defining the corrected image as $\tilde{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \mathbf{L}_2\Delta\mathbf{X}^{(t)}$, updating the latent image $\mathbf{X}_1$ consists in minimizing

$$\begin{aligned}\mathcal{J}_{\delta\mathrm{s}}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\delta\mathrm{s}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.10}$$

Minimizing (B.10) formulates a joint spectral deblurring problem from an observed image $\mathbf{Y}_1$ and a pseudo-observed image $\tilde{\mathbf{Y}}_2^{(k)}$. Thanks to its quadratic form, this least-square problem can be easily solved.

Appendix B.3.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

The predicted image that would be observed by sensor S_2 at time t_1 can be defined as

$$\check{\mathbf{Y}}_2^{(k)} = \mathbf{L}_2\mathbf{X}_1^{(k)} \tag{B.11}$$

with the resulting predicted change image

$$\Delta\check{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \check{\mathbf{Y}}_2^{(k)}. \tag{B.12}$$

The objective function (9) w.r.t. $\Delta\mathbf{X}$ is then rewritten by combining (B.11) and (B.12) with (Appendix B.3), leading to

$$\begin{aligned}\mathcal{J}_{\delta\mathrm{s}}^{(\mathrm{c})}(\Delta\mathbf{X}) &\triangleq \mathcal{J}_{\mathcal{S}_{\delta\mathrm{s}}}(\mathbf{X}_1^{(k)}, \Delta\mathbf{X}) \\ &= \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}}\left(\Delta\check{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\Delta\mathbf{X}\right)\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned} \tag{B.13}$$

Minimizing (B.13) is a spectral deblurring of the predicted change image $\Delta\check{\mathbf{Y}}_2^{(k)}$, which can be achieved using a forward-backward algorithm as proposed by [18].

Appendix B.4. Scenario $\mathcal{S}_{\mathrm{sd}}$

In this scenario, the two observed images share the same spectral resolution but differ by their spatial resolutions. These spatial resolutions are related by an integer relative downsampling factor, i.e., based on superimposable sampling grids, which allows a unique spatial degradation matrix $\mathbf{R}_1$ to be used[3]. The joint forward observation model derived from (11) using the specific degradation matrices presented in Table 1 can be written as

$$\mathbf{Y}_1 = \mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1 \tag{B.14a}$$
$$\mathbf{Y}_2 = (\mathbf{X}_1 + \Delta\mathbf{X}) + \mathbf{N}_2 \tag{B.14b}$$

with the objective function

$$\begin{aligned}\mathcal{J}_{\mathrm{sd}}(\mathbf{X}_1, \Delta\mathbf{X}) = &\frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}(\mathbf{Y}_2 - (\mathbf{X}_1 + \Delta\mathbf{X}))\right\|_{\mathrm{F}}^2 \\ &+ \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{R}_1)\right\|_{\mathrm{F}}^2 \\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\|\Delta\mathbf{X}\|_{2,1}.\end{aligned}$$

Appendix B.4.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

The same strategy as for previous scenarios is adopted here. As the model (B.14b) is the same as model (B.1b), the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ is defined following (B.3). Then, updating the latent image consists in minimizing w.r.t. $\mathbf{X}_1$ the partial function

$$\begin{aligned}\mathcal{J}_{\mathrm{sd}}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\mathrm{sd}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{R}_1)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned}$$

This fusion task can be interpreted as a set of n_λ super-resolution problems associated with each band of the observed image $\mathbf{Y}_1$, where the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ acts here as a prior information. Closed-form expressions of these n_λ solutions are given by [36].

Appendix B.4.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

As the model (B.14b) is the same as the model (B.1b) of scenarios $\mathcal{S}_{\mathrm{ss}}$, optimizing w.r.t. $\Delta\mathbf{X}$ can be conducted following the procedure detailed in paragraph Appendix B.1.2 (i.e., denoising of the predicted change image).

Appendix B.5. Scenario $\mathcal{S}_{\mathrm{s}\delta}$

As for scenario $\mathcal{S}_{\mathrm{sd}}$, scenario $\mathcal{S}_{\mathrm{s}\delta}$ considers two observed images of same spectral resolutions but with distinct spatial resolutions. However, contrary to scenario $\mathcal{S}_{\mathrm{sd}}$, this difference in spatial resolutions cannot be expressed thanks to a unique spatial degradation matrix $\mathbf{R}_1$ due to non-superimposable spatial sampling grids (i.e., corresponding to a non-integer relative downsampling factor). Thus the forward model is written

$$\mathbf{Y}_1 = \mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1 \tag{B.15a}$$
$$\mathbf{Y}_2 = (\mathbf{X}_1 + \Delta\mathbf{X})\,\mathbf{R}_2 + \mathbf{N}_2 \tag{B.15b}$$

with the following objective function

$$\begin{aligned}\mathcal{J}_{\mathrm{s}\delta}(\mathbf{X}_1, \Delta\mathbf{X}) = &\frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}(\mathbf{Y}_2 - (\mathbf{X}_1 + \Delta\mathbf{X})\,\mathbf{R}_2)\right\|_{\mathrm{F}}^2 \\ &+ \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{R}_1)\right\|_{\mathrm{F}}^2 \\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\|\Delta\mathbf{X}\|_{2,1}.\end{aligned} \tag{B.16}$$

In (B.15), both latent images are supposed to suffer from spatial degradations. Thus, choosing which spatial degradation affects the change image $\Delta\mathbf{X}$ results in a particular spatial resolution for this change map. To derive a CD map at a high spatial resolution, the spatial degradation applied to $\Delta\mathbf{X}$ should be chosen as the one with the lowest virtual downsampling factor. The minimization of (B.16) according to the AM strategy is addressed in the following paragraphs.

Appendix B.5.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

For this scenario, the corrected image in (12) is defined as

$$\tilde{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \Delta\mathbf{X}^{(k)}\mathbf{R}_2.$$

Then, updating the latent image $\mathbf{X}_1$ consists in minimizing w.r.t. $\mathbf{X}_1$ the partial function

$$\begin{aligned}\mathcal{J}_{\mathrm{s}\delta}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\mathrm{s}\delta}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{R}_1)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.17}$$

As for scenario $\mathcal{S}_{\mathrm{sd}}$, minimizing (B.17) can be interpreted as recovering a spatially super-resolved image $\mathbf{X}_1$ from the observed image $\mathbf{Y}_1$ and the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$. However, contrary to scenario $\mathcal{S}_{\mathrm{sd}}$, here, $\tilde{\mathbf{Y}}_2^{(k)}$ rather defines an additional data-fitting term instead of a prior information [37]. Moreover, this subproblem cannot be solved directly since no closed-form solution can be efficiently derived, mainly due to the simultaneous presence of the two spatial degradation operators. Thus, one proposes to resort to an iterative algorithm, namely the alternating direction method of multipliers (ADMM). It consists in introducing the splitting variable $\mathbf{U} \in \mathbb{R}^{n_\lambda \times n} = \mathbf{X}_1$. The resulting scaled augmented Lagrangian can be written as

$$\begin{aligned}\mathcal{L}_\mu(\mathbf{X}_1, \mathbf{U}, \mathbf{V}) = &\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}(\mathbf{Y}_1 - \mathbf{U}\mathbf{R}_1)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 \\ &+ \frac{\mu}{2}\|\mathbf{X}_1 - \mathbf{U} + \mathbf{V}\|_{\mathrm{F}}^2.\end{aligned} \tag{B.18}$$

The ADMM iteratively minimizes $\mathcal{L}_\mu$ w.r.t. $\mathbf{U}$ and $\mathbf{X}_1$ and updates the dual variable $\mathbf{V}$. The two minimizations of (B.18) w.r.t. $\mathbf{U}$ and $\mathbf{X}_1$ can be conducted band-by-band following the strategy proposed by [36], which provides closed-form solutions of the underlying single-image super-resolution problems and also ensures the convergence of the AM algorithm.

---

[3] The case of observed images with non-integer relative spatial downsampling factor is discussed in the next presented scenario.

Appendix B.5.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

For this scenario, a predicted image that would be observed by the sensor S_2 at time t_1 can be defined as

$$\check{\mathbf{Y}}_2^{(k)} = \mathbf{X}_1^{(k)}\mathbf{R}_2 \tag{B.19}$$

with the resulting predicted change image

$$\Delta\check{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \check{\mathbf{Y}}_2^{(k)}. \tag{B.20}$$

The objective function w.r.t. $\Delta\mathbf{X}$ is then rewritten by combining (B.19) and (B.20) with (B.16), leading to

$$\begin{aligned}\mathcal{J}_{s\delta}^{(c)}(\Delta\mathbf{X}) &\triangleq \mathcal{J}_{s\delta}(\mathbf{X}_1^{(k)}, \Delta\mathbf{X})\\ &= \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\Delta\check{\mathbf{Y}}_2^{(k)} - \Delta\mathbf{X}\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned} \tag{B.21}$$

The minimization of (B.21) can be interpreted as a super-resolution problem. Even if a forward-backward algorithm could be used to iteratively minimize this objective function, the size of the spatial degradation matrix $\mathbf{R}_2$ suggests to resort to an ADMM. By introducing the splitting variable $\mathbf{W} \in \mathbb{R}^{n_\lambda \times m_2} = \Delta\mathbf{X}\mathbf{R}_2$, the resulting scaled augmented Lagrangian for the problem is expressed as

$$\begin{aligned}\mathcal{L}_\mu(\Delta\mathbf{X}, \mathbf{W}, \mathbf{V}) = &\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\Delta\check{\mathbf{Y}}_2^{(k)} - \mathbf{W}\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\Delta\mathbf{X}\right\|_{2,1}\\ &+ \frac{\mu}{2}\left\|\Delta\mathbf{X}\mathbf{R}_1 - \mathbf{W} + \mathbf{V}\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.22}$$

Closed-form expressions of the minimizers of (B.22) w.r.t. $\Delta\mathbf{X}$ and $\mathbf{W}$ can be derived, following a group soft-thresholding operation and the technique proposed by [36], respectively.

Appendix B.6. Scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$

Scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ is specifically addressed by [18] with the joint forward model

$$\begin{aligned}\mathbf{Y}_1 &= \mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1,\\ \mathbf{Y}_2 &= \mathbf{L}_2\left(\mathbf{X}_1 + \Delta\mathbf{X}\right) + \mathbf{N}_2.\end{aligned}$$

The two observed images have complementary information since $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are of high spectral and spatial resolutions, respectively. The resulting objective function writes

$$\begin{aligned}\mathcal{J}_{\substack{\mathrm{b}\\ \mathrm{dd}}}\left(\mathbf{X}_1, \Delta\mathbf{X}\right) = &\frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\mathbf{Y}_2 - \mathbf{L}_2\left(\mathbf{X}_1 + \Delta\mathbf{X}\right)\right)\right\|_{\mathrm{F}}^2\\ &+ \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2\\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned} \tag{B.23}$$

When these images have been acquired at the same time instant, the change image is $\Delta\mathbf{X} = \mathbf{0}$ and this configuration boils down to a multiband image fusion problem addressed by [30]. Thus, minimizing (B.23) can be conducted following the AM strategy by combining a multiband image fusion step [30] and a spectral deblurring step of the predicted change image. The interested reader is invited to consult the work by [18] for a comprehensive description of the resolution.

Appendix B.7. Scenario $\mathcal{S}_{\delta\mathrm{d}}$

This scenario generalizes the previous scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ for situations where the spectral sampling grids associated with the two observed image cannot be superimposed, which requires the simultaneous use of two spectral degradation operators. The joint forward observation model is then

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1. \tag{B.24a}$$

$$\mathbf{Y}_2 = \mathbf{L}_2\left(\mathbf{X}_1 + \Delta\mathbf{X}\right) + \mathbf{N}_2. \tag{B.24b}$$

which yields the objective function

$$\begin{aligned}\mathcal{J}_{\delta\mathrm{d}}\left(\mathbf{X}_1, \Delta\mathbf{X}\right) = &\frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\mathbf{Y}_2 - \mathbf{L}_2\left(\mathbf{X}_1 + \Delta\mathbf{X}\right)\right)\right\|_{\mathrm{F}}^2\\ &+ \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2\\ &+ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned}$$

Note that the estimated latent and change images are defined at the highest spatial resolution while benefiting from the spectral resolutions of both observed images. The choice of assuming that the image acquired by sensor S_2 does not suffer from spatial degradation has been motivated by an easier and accurate estimation of the change image $\Delta\mathbf{X}$ by avoiding additional spatial super-resolution steps. The resulting sub-problems involved in the AM algorithm are detailed below.

Appendix B.7.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

As for scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$, the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ can be defined as $\tilde{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \mathbf{L}_2\Delta\mathbf{X}^{(k)}$. Thus, updating the latent image $\mathbf{X}_1$ consists in minimizing

$$\begin{aligned}\mathcal{J}_{\delta\mathrm{d}}^{(\mathrm{f})}\left(\mathbf{X}_1\right) &\triangleq \mathcal{J}_{\delta\mathrm{d}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right)\\ &= \left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2\\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.25}$$

Minimizing (B.25) is challenging mainly due to the simultaneous presence of spatial and spectral degradation matrices $\mathbf{R}_1$ and $\mathbf{L}_2$ with an additional spatial degradation $\mathbf{L}_1$. Therefore, there is no closed-form solution for this problem, which can be eventually solved thanks to ADMM, by introducing the splitting variable $\mathbf{U} \in \mathbb{R}^{m_\lambda \times m_1} = \mathbf{X}_1\mathbf{R}_1$. The resulting scaled augmented Lagrangian is

$$\begin{aligned}\mathcal{L}_\mu(\mathbf{X}_1, \mathbf{U}, \mathbf{V}) = &\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{U}\right)\right\|_{\mathrm{F}}^2\\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2\\ &+ \frac{\mu}{2}\left\|\mathbf{X}_1\mathbf{R}_1 - \mathbf{U} + \mathbf{V}\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.26}$$

Closed-form expressions of the minimizers of (B.26) w.r.t. $\mathbf{X}_1$ and $\mathbf{U}$ can be derived, following the approach proposed by [30] and a least-square formulation, respectively.

Appendix B.7.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

As both models (B.24b) and (B.9b) are the same, optimizing w.r.t. $\Delta\mathbf{X}$ can be achieved following the strategy detailed in paragraph Appendix B.3.2, i.e., by spectrally deblurring a predicted change image $\Delta\check{\mathbf{Y}}_2^{(k)}$ thanks to the forward-backward algorithm proposed by [18].

Appendix B.8. Scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$

Under this scenario, the observed image $\mathbf{Y}_2$ is of higher spatial and spectral resolutions than the observed image $\mathbf{Y}_1$. Within a conventional fusion context, one would probably discard $\mathbf{Y}_1$ since it would not bring additional information to the one provided by $\mathbf{Y}_2$. Conversely, within a CD context, both observed images are of interest and should be exploited. More precisely, here, the joint forward observation model derived from (11) is specifically written

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1, \tag{B.27a}$$
$$\mathbf{Y}_2 = (\mathbf{X}_1 + \Delta\mathbf{X}) + \mathbf{N}_2, \tag{B.27b}$$

with the resulting objective function

$$\begin{aligned}\mathcal{J}_{\mathrm{u}_{\mathrm{dd}}}(\mathbf{X}_1, \Delta\mathbf{X}) = \frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\mathbf{Y}_2 - (\mathbf{X}_1 + \Delta\mathbf{X})\right)\right\|_{\mathrm{F}}^2 \\ + \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned}$$

Its minimization relies on the two steps detailed below.

Appendix B.8.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

After defining the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ by (B.3), updating the the latent image $\mathbf{X}_1$ consists in minimizing

$$\begin{aligned}\mathcal{J}_{\mathrm{u}_{\mathrm{dd}}}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\mathrm{u}_{\mathrm{dd}}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.28}$$

Minimizing (B.28) can be interpreted as a simultaneous spatial super-resolution and spectral deblurring of the multiband image $\mathbf{Y}_1$, with prior information brought by $\tilde{\mathbf{Y}}_2^{(k)}$. This minimization is a much more challenging task than the fusion steps previously encountered for scenarios $\mathcal{S}_{\mathrm{ss}}$ or $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$. Indeed, the simultaneous spatial and spectral degradations applied to $\mathbf{X}_1$ prevents a closed-form solution to be efficiently computed. Thus, as for scenario $\mathcal{S}_{\mathrm{s}\delta}$, one resorts to the ADMM scheme by introducing the splitting variable $\mathbf{U} \in \mathbb{R}^{m_{\lambda_1} \times n} = \mathbf{L}_1\mathbf{X}_1$. The resulting scaled augmented Lagrangian for the problem is expressed as

$$\begin{aligned}\mathcal{L}_{\mu}(\mathbf{X}_1, \mathbf{U}, \mathbf{V}) = \left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{U}\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 + \\ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\right)\right\|_{\mathrm{F}}^2 + \\ \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \frac{\mu}{2}\left\|\mathbf{L}_1\mathbf{X}_1 - \mathbf{U} + \mathbf{V}\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.29}$$

By comparing the partial objective function (B.28) and its augmented counterpart (B.29), it clearly appears that the splitting strategy allows the spectral and spatial degradations to be decoupled. Thus, within the ADMM framework, minimizing $\mathcal{L}_{\mu}(\mathbf{X}_1, \mathbf{U}, \mathbf{V})$ w.r.t. $\mathbf{U}$ and $\mathbf{X}_1$ can be easily conducted. More precisely, optimizing w.r.t. $\mathbf{U}$ consists in conducting a super-resolution step achieved as for scenario $\mathcal{S}_{\mathrm{sd}}$ by resorting to the algorithm proposed by [36]. Conversely, optimizing w.r.t. $\mathbf{X}_1$ consists in solving a least-square problem whose closed-form solution can be computed (akin to scenario $\mathcal{S}_{\mathrm{ds}}$).

Appendix B.8.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

Again, as the forward model (B.27b) is the same as (B.1b) of scenario $\mathcal{S}_{\mathrm{ss}}$, optimizing w.r.t. $\Delta\mathbf{X}$ can be conducted following the procedure detailed in paragraph Appendix B.1.2 (i.e., denoising of the predicted change image).

Appendix B.9. Scenario $\mathcal{S}_{\mathrm{d}\delta}$

Scenario $\mathcal{S}_{\mathrm{d}\delta}$ generalizes scenario $\mathcal{S}_{\mathrm{dd}}^{\mathrm{b}}$ with the specific case of two observed images whose spatial sampling grids cannot be superimposed, which precludes the use of a unique spatial degradation matrix. The resulting joint observation model is

$$\mathbf{Y}_1 = \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1 + \mathbf{N}_1. \tag{B.30a}$$
$$\mathbf{Y}_2 = (\mathbf{X}_1 + \Delta\mathbf{X})\,\mathbf{R}_2 + \mathbf{N}_2 \tag{B.30b}$$

which leads to the following objective function

$$\begin{aligned}\mathcal{J}_{\mathrm{d}\delta}(\mathbf{X}_1, \Delta\mathbf{X}) = \frac{1}{2}\left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\mathbf{Y}_2 - (\mathbf{X}_1 + \Delta\mathbf{X})\,\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 \\ + \frac{1}{2}\left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 + \gamma\left\|\Delta\mathbf{X}\right\|_{2,1}.\end{aligned}$$

The choice of assuming that the image acquired by the sensor S_2 does not suffers from spectral degradation is motivated by an easier and more accurate estimation of the change image $\Delta\mathbf{X}$ by avoiding additional spectral deblurring steps. The two subproblems underlying the AM algorithm are detailed below.

Appendix B.9.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

By defining the corrected image as for Scenario $\mathcal{S}_{\mathrm{s}\partial}$, i.e.,

$$\tilde{\mathbf{Y}}_2^{(k)} = \mathbf{Y}_2 - \Delta\mathbf{X}^{(k)}\mathbf{R}_2,$$

updating the latent image $\mathbf{X}_1$ consists in minimizing the partial function

$$\begin{aligned}\mathcal{J}_{\mathrm{d}\delta}^{(\mathrm{f})}(\mathbf{X}_1) &\triangleq \mathcal{J}_{\mathcal{S}_{\mathrm{d}\delta}}\left(\mathbf{X}_1, \Delta\mathbf{X}^{(k)}\right) \\ &= \left\|\mathbf{\Lambda}_1^{-\frac{1}{2}}\left(\mathbf{Y}_1 - \mathbf{L}_1\mathbf{X}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\mathbf{\Lambda}_2^{-\frac{1}{2}}\left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \lambda\left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2.\end{aligned} \tag{B.31}$$

Unfortunately, it is not possible to derive a closed-form solution of the minimizer (B.31). As for Scenarios $\mathcal{S}_{\mathrm{dd}}^{\mathrm{u}}$ and $\mathcal{S}_{\mathrm{s}\delta}$, capitalizing on the convexity of the objective function, an ADMM strategy is followed. By defining the splitting variable $\mathbf{U} \in \mathbb{R}^{m_{\lambda_1} \times n} = \mathbf{L}_1\mathbf{X}_1$, the scaled augmented Lagrangian can be written

$$\begin{aligned}\mathcal{L}_\mu(\mathbf{X}_1, \mathbf{U}, \mathbf{V}) = &\left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}} \left(\mathbf{Y}_1 - \mathbf{U}\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}} \left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{X}_1\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \lambda \left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 \\ &+ \frac{\mu}{2} \left\|\mathbf{L}_1\mathbf{X}_1 - \mathbf{U} + \mathbf{V}\right\|_{\mathrm{F}}^2 .\end{aligned} \tag{B.32}$$

Iterative minimizations of (B.32) w.r.t. $\mathbf{U}$ and $\mathbf{X}_1$ can be conducted efficiently. More precisely, optimizing w.r.t. $\mathbf{U}$ consists in solving a set of super-resolution problems whose closed-form solutions are given band-by-band by [36]. Regarding the minimization w.r.t. $\mathbf{X}_1$, it consists in solving a ℓ_2-penalized super-resolution problem, whose closed-form solution is given by [30].

Appendix B.9.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

Since the observation model (B.30b) related to $\Delta\mathbf{X}$ is the same as the one of Scenario $\mathcal{S}_{\mathrm{s}\delta}$ (see (B.15b)), optimizing w.r.t. $\Delta\mathbf{X}$ can be achieved thanks to ADMM, as described in paragraph Appendix B.5.2 (spatial super-resolution of the predicted change image).

Appendix B.10. Scenario $\mathcal{S}_{\delta\delta}$

This scenario generalizes all the previous scenarios with the particular difficulties of non-superimposable spatial and spectral sampling grids associated with the two observed images. The joint forward observation model is given by (11), which results in the objective function $\mathcal{J}_{\delta\delta}$ in (9). Again, as for scenarios $\mathcal{S}_{\mathrm{d}\delta}$ and $\mathcal{S}_{\delta\mathrm{d}}$, the choice of the spatial and spectral degradations applied to the change image $\Delta\mathbf{X}$ should be motivated by reaching the highest spatial and spectral resolutions of this change image. The optimization sub-problems are finally discussed below.

Appendix B.10.1. Fusion: optimization w.r.t. $\mathbf{X}_1$

For this scenario, the corrected image $\tilde{\mathbf{Y}}_2^{(k)}$ is given by (12), leading to an updating rule w.r.t. $\mathbf{X}_1$ which consists in minimizing (14). This minimization cannot be conducted in a straightforward manner, since it requires to conduct a spectral deblurring and a spatial super-resolution simultaneously. However, the optimal solution can be reached by resorting to a ADMM with two splitting variables $\mathbf{U}_1 = \mathbf{L}_1\mathbf{X}_1 \in \mathbb{R}^{m_{\lambda_1} \times n}$ and $\mathbf{U}_2 = \mathbf{X}_1\mathbf{R}_2 \in \mathbb{R}^{n_\lambda \times m_2}$. The resulting scaled augmented Lagrangian for the problem is expressed as

$$\begin{aligned}\mathcal{L}_\mu(\mathbf{X}_1, \mathbf{U}_1, \mathbf{U}_2, \mathbf{V}_1, \mathbf{V}_2) = &\left\|\boldsymbol{\Lambda}_1^{-\frac{1}{2}} \left(\mathbf{Y}_1 - \mathbf{U}_1\mathbf{R}_1\right)\right\|_{\mathrm{F}}^2 \\ &+ \left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}} \left(\tilde{\mathbf{Y}}_2^{(k)} - \mathbf{L}_2\mathbf{U}_2\right)\right\|_{\mathrm{F}}^2 + \lambda \left\|\mathbf{X}_1 - \bar{\mathbf{X}}_1\right\|_{\mathrm{F}}^2 \\ &+ \frac{\mu}{2} \left\|\mathbf{L}_1\mathbf{X}_1 - \mathbf{U}_1 + \mathbf{V}_1\right\|_{\mathrm{F}}^2 + \frac{\mu}{2} \left\|\mathbf{X}_1\mathbf{R}_2 - \mathbf{U}_2 + \mathbf{V}_2\right\|_{\mathrm{F}}^2 .\end{aligned} \tag{B.33}$$

Closed-form expressions of the minimizers of (B.33) w.r.t. $\mathbf{X}_1$, $\mathbf{U}_1$ and $\mathbf{U}_2$ can be derived as proposed by [30], [36] and following a least-square formulation, respectively.

Appendix B.10.2. Correction: optimization w.r.t. $\Delta\mathbf{X}$

For this scenario, given the current state $\mathbf{X}_1^{(k)}$ of the latent image, the predicted image that would be observed by the sensor S_2 at time t_1 can be defined as in (15) leading to the predicted change image (16). Then, the correction step consists in minimizing the objective function $\mathcal{J}_{\delta\delta}^{(\mathrm{c})}(\Delta\mathbf{X})$ in (18). It consists in conducting a spectral deblurring and spatial super-resolution jointly. This problem has no closed-form solution. Therefore, the objective function is iteratively minimized using an ADMM with two splitting variables $\mathbf{W}_1 = \mathbf{L}_1\Delta\mathbf{X} \in \mathbb{R}^{m_{\lambda_1} \times n}$ and $\mathbf{W}_2 = \Delta\mathbf{X} \in \mathbb{R}^{n_\lambda \times n}$. The resulting scaled augmented Lagrangian for the problem is expressed as

$$\begin{aligned}&\mathcal{L}_\mu(\Delta\mathbf{X}, \mathbf{W}_1, \mathbf{W}_2, \mathbf{V}_1, \mathbf{V}_2) = \\ &\left\|\boldsymbol{\Lambda}_2^{-\frac{1}{2}} \left(\Delta\check{\mathbf{Y}}_2^{(k)} - \mathbf{W}_1\mathbf{R}_2\right)\right\|_{\mathrm{F}}^2 + \gamma \left\|\mathbf{W}_2\right\|_{2,1} \\ &+ \frac{\mu}{2} \left\|\mathbf{L}_1\Delta\mathbf{X} - \mathbf{W}_1 + \mathbf{V}_1\right\|_{\mathrm{F}}^2 + \frac{\mu}{2} \left\|\Delta\mathbf{X} - \mathbf{W}_2 + \mathbf{V}_2\right\|_{\mathrm{F}}^2 .\end{aligned} \tag{B.34}$$

Closed-form expression of the minimizers of (B.34) w.r.t. $\Delta\mathbf{X}$, $\mathbf{W}_1$ and $\mathbf{W}_2$ can be derived, following a least-square formulation, the computation proposed by [36] and a group soft-thresholding, respectively.

Appendix C. Acknowledgments

Part of this work has been supported by Coordenação de Aperfeiçoamento de Ensino Superior (CAPES), Brazil, EU FP7 through the ERANETMED JC-WATER program [MapInvPlnt Project ANR-15-NMED- 0002-02] and the ANR-3IA Artificial and Natural Intelligence Toulouse Institute (ANITI).

References

[1] T. E. Bell, Remote sensing, IEEE Spectrum 32 (3) (1995) 24–31.

[2] C. Elachi, J. Van Zyl, Introduction to the physics and techniques of remote sensing, 2nd Edition, Wiley series in remote sensing, Wiley-Interscience, Hoboken, N.J., 2006.

[3] J. A. Richards, X. Jia, Remote sensing digital image analysis: an introduction, 4th Edition, Springer, Berlin, 2006.

[4] J. B. Campbell, R. H. Wynne, Introduction to remote sensing, 5th Edition, Guilford Press, New York, 2011.

[5] F. Bovolo, L. Bruzzone, The time variable in data fusion: A change detection perspective, IEEE Geosci. Remote Sens. Mag. 3 (3) (2015) 8–26.

[6] A. Singh, Review Article Digital change detection techniques using remotely-sensed data, Int. J. Remote Sens. 10 (6) (1989) 989–1003.

[7] P. Du, S. Liu, J. Xia, Y. Zhao, Information fusion techniques for change detection from multi-temporal remote sensing images, Information Fusion 14 (1) (2013) 19–27.

[8] G. Xian, C. Homer, J. Fry, Updating the 2001 national land cover database impervious surface products to 2006 using Landsat imagery change detection methods, Remote Sensing of Environment 113 (6) (2009) 1133–1147.

[9] J. Prendes, M. Chabert, F. Pascal, A. Giros, J.-Y. Tourneret, A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors, IEEE Trans. Image Process. 24 (3) (2015) 799–812.

[10] C. Wu, B. Du, X. Cui, L. Zhang, A post-classification change detection method based on iterative slow feature analysis and Bayesian soft fusion, Remote Sensing of Environment 199 (2017) 241–255.

[11] H. Luo, C. Liu, C. Wu, X. Guo, Urban change detection based on Dempster–Shafer theory for multitemporal very high-resolution imagery, Remote Sensing 10 (7) (2018) 980.

[12] Union of Concerned Scientists, UCS Satellite Database (2017).
URL https://www.ucsusa.org/nuclear-weapons/space-weapons/satellite-database#.WigLykqnFPY

[13] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, J. A. Benediktsson, Challenges and opportunities of multimodality and data fusion in remote sensing, Proc. IEEE 103 (9) (2015) 1585–1601.

[14] F. Bovolo, L. Bruzzone, A theoretical framework for unsupervised change detection based on change vector analysis in the polar domain, IEEE Trans. Geosci. Remote Sens. 45 (1) (2007) 218–236.

[15] F. Bovolo, S. Marchesi, L. Bruzzone, A framework for automatic and unsupervised detection of multiple changes in multitemporal images, IEEE Trans. Geosci. Remote Sens. 50 (6) (2012) 2196–2212.

[16] A. A. Nielsen, K. Conradsen, J. J. Simpson, Multivariate alteration detection (MAD) and MAF postprocessing in multispectral, bitemporal image data: New approaches to change detection studies, Remote Sens. Environment 64 (1) (1998) 1–19.

[17] A. A. Nielsen, The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data, IEEE Trans. Image Process. 16 (2) (2007) 463–478.

[18] V. Ferraris, N. Dobigeon, Q. Wei, M. Chabert, Robust fusion of multiband images with different spatial and spectral resolutions for change detection, IEEE Trans. Computational Imaging 3 (2) (2017) 175–186.

[19] V. Ferraris, N. Dobigeon, Q. Wei, M. Chabert, Detecting changes between optical images of different spatial and spectral resolutions: A fusion-based approach, IEEE Trans. Geosci. Remote Sens. 56 (3) (2018) 1566–1578.

[20] K. Kotwal, S. Chaudhuri, A novel approach to quantitative evaluation of hyperspectral image fusion techniques, Information Fusion 14 (1) (2013) 5–18.

[21] K. Kotwal, S. Chaudhuri, A bayesian approach to visualization-oriented hyperspectral image fusion, Information Fusion 14 (4) (2013) 349–360.

[22] H. Song, B. Huang, K. Zhang, H. Zhang, Spatio-spectral fusion of satellite images based on dictionary-pair learning, Information Fusion 18 (Supplement C) (2014) 148–160.

[23] H. Ghassemian, A review of remote sensing image fusion methods, Information Fusion 32 (Part A) (2016) 75–89.

[24] S. Li, X. Kang, L. Fang, J. Hu, H. Yin, Pixel-level image fusion: A survey of the state of the art, Information Fusion 33 (Supplement C) (2017) 100–112.

[25] V. Ferraris, N. Dobigeon, Q. Wei, M. Chabert, Change detection between multi-band images using a robust fusion-based approach, in: Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP), 2017, pp. 3346–3350.

[26] V. Ferraris, N. Yokoya, N. Dobigeon, M. Chabert, A comparative study of fusion-based change detection methods for multi-band images with different spectral and spatial resolutions, in: Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS), Valencia, Spain, 2018, pp. 5021–5024.

[27] Q. Wei, N. Dobigeon, J.-Y. Tourneret, Bayesian Fusion of Multi-Band Images, IEEE J. Sel. Topics Signal Process. 9 (6) (2015) 1117–1127.

[28] N. Yokoya, T. Yairi, A. Iwasaki, Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion, IEEE Trans. Geosci. Remote Sens. 50 (2) (2012) 528–537.

[29] M. Simões, J. Bioucas Dias, L. Almeida, J. Chanussot, A convex formulation for hyperspectral image superresolution via subspace-based regularization, IEEE Trans. Geosci. Remote Sens. 6 (53) (2015) 3373–3388.

[30] Q. Wei, N. Dobigeon, J.-Y. Tourneret, Fast Fusion of Multi-Band Images Based on Solving a Sylvester Equation, IEEE Trans. Image Process. 24 (11) (2015) 4109–4121.

[31] N. Yokoya, N. Mayumi, A. Iwasaki, Cross-Calibration for Data Fusion of EO-1/Hyperion and Terra/ASTER, IEEE J. Sel. Topics Appl. Earth Observations Remote Sens. 6 (2) (2013) 419–426.

[32] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J.-Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, N. Yokoya, Hyperspectral pansharpening: A review, IEEE Geosci. Remote Sens. Mag. 3 (3) (2015) 27–46.

[33] R. D. Johnson, E. S. Kasischke, Change vector analysis: A technique for the multispectral monitoring of land cover and condition, Int. J. Remote Sens. 19 (3) (1998) 411–426.

[34] C. Févotte, N. Dobigeon, Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization, IEEE Trans. Image Process. 24 (12) (2015) 4810–4819.

[35] Jianchao Yang, J. Wright, T. S. Huang, Yi Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2861–2873.

[36] N. Zhao, Q. Wei, A. Basarab, N. Dobigeon, D. Kouame, J.-Y. Tourneret, Fast Single Image Super-Resolution Using a New Analytical Solution for $\ell_2 - \ell_2$ Problems, IEEE Trans. Image Process. 25 (8) (2016) 3683–3697.

[37] M. Elad, A. Feuer, Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images, IEEE Trans. Image Process. 6 (12) (1997) 1646–1658.

[38] R. C. Hardie, M. T. Eismann, G. L. Wilson, MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor, IEEE Trans. Image Process. 13 (9) (2004) 1174–1184.

[39] M. T. Eismann, R. C. Hardie, Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions, IEEE Trans. Image Process. 43 (3) (2005) 455–465.

[40] Y. Zhang, S. De Backer, P. Scheunders, Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images, IEEE Trans. Geosci. Remote Sens. 47 (11) (2009) 3834–3843.

[41] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, Hyperspectral and multispectral image fusion based on a sparse representation, IEEE Trans. Geosci. Remote Sens. 53 (7) (2015) 3658–3668.

[42] E. J. Candés, X. Li, Y. Ma, J. Wright, Robust principal component analysis?, Journal of the ACM (JACM) 58 (3) (2011) 11.

[43] M. J. Canty, Image analysis, classification and change detection in remote sensing: with algorithms for ENVI/IDL and Python, CRC Press, 2014.

[44] P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, A. Plaza, Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art, IEEE Geosci. Remote Sensing Mag 5 (4) (2017) 37–78.

[45] S. Liu, D. Marinelli, L. Bruzzone, F. Bovolo, A review of change detection in multitemporal hyperspectral images: Current techniques, applications, and challenges, IEEE Geosci. Remote Sensing Mag. 7 (2) (2019) 140–158.

[46] L. Wald, T. Ranchin, M. Mangolini, Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images, Photogrammetric engineering and remote sensing 63 (6) (1997) 691–699.

[47] W. W. Peterson, T. G. Birdsall, W. Fox, The theory of signal detectability, IRE Trans. Inf. Theory 4 (4) (1954) 171–212.

[48] M. S. Pepe, Receiver operating characteristic methodology, J. Am. Stat. Ass. 95 (449) (2000) 308–311.

[49] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Lett. 27 (2006) 861–874.

[50] J. Inglada, G. Mercier, A new statistical similarity measure for change detection in multitemporal SAR images and its extension to multiscale change analysis, IEEE Trans. Geosci. Remote Sens. 45 (5) (2007) 1432–1445.

[51] M.-T. Pham, G. Mercier, J. Michel, Change detection between SAR images using a pointwise approach and graph theory, IEEE Trans. Geosci. Remote Sens. 54 (4) (2016) 2020–2032.

[52] T. Hastie, R. Tibshirani, J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction, 2nd Edition, Springer, New York, 2009.

[53] J. A. Hanley, B. J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, Radiology 143 (1) (1982) 29–36.

[54] A. P.Bradley, The use of the area under the roc curve in the evaluation of machine learning algorithms, Pattern Recognition 30 (7) (1997) 1145–1159.

[55] S. J. Mason, N. E. Graham, Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation, Quarterly J. the Royal Meteorological Society 128 (2002) 2145–2166.

[56] P. A. Flach, J. Hernandez-Orallo, C. Ferri, A coherent interpretation of AUC as a measure of aggregated classification performance, in: Proc. Int. Conf. Machine Learning (ICML), 2011, pp. 657–664.

[57] V. Ferraris, N. Dobigeon, M. Chabert, Robust fusion algorithms for unsupervised change detection between multi-band optical images – A comprehensive case study – Complementary results (June 2020).
URL http://dobigeon.perso.enseeiht.fr/papers/Ferraris_TechReport_2020.pdf

[58] United States Geological Survey, Landsat-8 (2017).
URL `https://landsat.usgs.gov/landsat-8`

[59] European Space Agency, Sentinel-2 (2017).
URL `https://sentinel.esa.int/web/sentinel/missions/sentinel-2`

[60] United States Geological Survey, EO-1 Advanced Land Imager (ALI) (2017).
URL `https://eo1.usgs.gov/sensors/ali`

[61] Jet Propulsion Laboratory, Airborne visible / infrared imaging spectrometer (AVIRIS) (2017).
URL `https://aviris.jpl.nasa.gov`

[62] V. Ferraris, N. Dobigeon, Y. C. Cavalcanti, T. Oberlin, M. Chabert, Coupled dictionary learning for unsupervised change detection between multi-sensor remote sensing images, Computer Vision and Image Understanding 189 (102817) (Dec. 2019).

[63] A. K. Gupta, D. K. Nagar, Matrix Variate Distribution, no. 104 in Monographs and Surveys in Pure and Applied Mathematics, Chapman and Hall, 1999.