

Cross-Sensors Comparison of Popular Vegetation Indexes From Landsat TM, ETM+, OLI, and Sentinel MSI for Time-Series Analysis Over Europe

Francesca Trevisiol¹, Pietro Mattivi, Emanuele Mandanici¹, and Gabriele Bitelli¹, *Member, IEEE*

Abstract—In the perspective of long and dense time-series analyses for environmental monitoring applications, this article discusses a cross-comparison analysis between the different instruments of Landsat and Sentinel missions [thematic mapper (TM), enhanced thematic mapper plus (ETM+), operational land imager (OLI), and multispectral instrument (MSI)]. The level-2 surface reflectance (SR) products were considered (in particular, the reprocessed Collection-2 for Landsat). The calibration coefficients for four of the most popular vegetation indexes [normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), soil adjusted vegetation index (SAVI), and normalized difference moisture index (NDMI)] were estimated, with the aim of harmonizing and minimizing radiometric differences for the combined use of these sensors. For this purpose, more than 20 000 pairs of images almost simultaneously acquired (± 1 -day tolerance window) were selected over a period of several years (depending on the lifespan overlap of every sensor pair). Vegetation indices (VIs) were computed for each image, and for each cross comparison, 100 random extractions of 300 000 sample pixels were performed all over the European continent. Linear transformation functions for each VI and between each sensor couple were computed by regression analyses, also assessing the repeatability of the estimation. Furthermore, the stability over time of the obtained coefficients was assessed when enough years of corresponding observations are available.

Index Terms—Cross-sensor transformation function, enhanced vegetation index (EVI) harmonization, Landsat, NDMI, normalized difference vegetation index (NDVI), Sentinel, soil adjusted vegetation index (SAVI).

I. INTRODUCTION

EARTH observation (EO) represents the most efficient approach to detect and monitor Earth surface dynamics and changes [1], [2], [3]. Nowadays, the available open-access remote sensing datasets represent a massive source of information that has not been completely exploited yet to its full potential and requires new methodologies, supported by technological advances [4]. Among these, long time-series analysis (TSA) is undoubtedly promising to better understand Earth's surface change processes, making possible to study

not only drastic changes, but also long-term or more subtle phenomena, crucial in climate change studies. In this context, multispectral missions, such as Landsat and Sentinel, that capture images of our planet since the early 1970s, represent a precious data source that deserves deeper examination.

The Landsat program, started officially in 1967 from a partnership between the National Aeronautics and Space Administration (NASA) and the U.S. Geological Survey (USGS), is the longest running and the first medium-resolution EO mission. Its programmatic development focuses on program continuity providing high-quality data without interruptions because of continuous turnover and updating of satellites [5], [6]. One major turning point for the program and for the EO industry was in 2008, when the Landsat archive policy turned into open access, delivering data with different processing and complexity levels [7]. The same programmatic path has been followed in the design and maintenance of the open-access Sentinel-2 mission, managed by the European Space Agency (ESA) in the framework of the European EO program “Copernicus.” Some improvements with respect to the Landsat program are in the spatial and temporal resolution [8].

In this context, the development of new technologies in the era of big EO data is crucial to enable the full exploitation of the data potentiality [9]. This demand has been successfully accomplished with the development of cloud-computing technologies, such as Google Earth engine (GEE), which was used in this study [10]. GEE is a platform for geospatial analysis consisting in a multipetabyte analysis-ready data (ARD) catalog colocated with a high-performance, intrinsically parallel computation service powered by Google [11].

Because of this great amount of free data accessible through cloud computing platforms, traditional remote sensing applications, such as change detection, have successfully moved from standard approaches involving image pairs to TSA of remotely sensed data [12], [13]. The pixel-based TSA for monitoring pixel trajectories over time represents, nowadays, a well-established methodological trend [14]. An overview regarding potentiality, methodologies, and challenges of remote sensing time series can be found in [15].

Moreover, this approach to the analysis of multispectral data is rapidly becoming a standard practice in vegetation-related studies, including monitoring of both forest and agricultural environments [16], [17]. The reviews by Bégué et al. [18] and Gómez et al. [19] cover a variety of examples where

Manuscript received 5 August 2023; revised 26 October 2023 and 28 November 2023; accepted 7 December 2023. Date of publication 14 December 2023; date of current version 29 December 2023. (Corresponding author: Francesca Trevisiol.)

The authors are with the Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, 40136 Bologna, Italy (e-mail: francesca.trevisiol2@unibo.it; pietro.mattivi2@unibo.it; emanuele.mandanici@unibo.it; gabriele.bitelli@unibo.it).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TGRS.2023.3343071>, provided by the authors.

Digital Object Identifier 10.1109/TGRS.2023.3343071

this methodology proved to be effective in identifying seasonal phenological variation, allowing, for example, a better classification of crop types. In this context, information on the biophysical characteristics of vegetation collected by multispectral cameras can be successfully synthesized because of vegetation indices (VIs), which are obtained through the combination of spectral bands [20], [21]. Four of the most widely used VIs are considered in this study: normalized difference vegetation index (NDVI), soil adjusted vegetation index (SAVI), enhanced vegetation index (EVI), and normalized difference moisture index (NDWI). In the last years, these VIs has been successfully implemented in a large variety of applications for their simplicity and effectiveness [22], [23], [24], [25], [26].

However, multispectral derived datasets have some known limits that need to be overcome to improve TSA applications. For example, a recent study conducted by USGS [5], investigating user needs for future Landsat missions, raised attention on desired cloud-free observation frequency: the survey showed that in the 71% of cases, subject matter experts believe that a weekly cloud-free observation frequency is a breakthrough requirement. Having cloud-free data information on a weekly basis would actually lead to a significant improvement in data effectiveness for their applications. One possible answer to this need is the harmonization of several sensors, which would allow to increase the number of acquisitions. Furthermore, since the main limitation of passive optical sensors is the dependence on weather conditions, harmonization would increase the probability of collecting cloud-free images. Looking at the present, as demonstrated by the analysis of Li and Chen [27], the harmonization of Sentinel-2 MSI and Landsat OLI/OLI+ (equipped on Landsat-8 and Landsat-9) would highly increase the revisit time, up to a 2.3-day global average, giving satellite remote sensing a new perspective for land surface monitoring.

In order to achieve these purposes, some authors proposed statistical calibration parameters to adjust spectral reflectance values across similar instruments and build synthetic multisatellite constellation. Mandanici and Bitelli [28] analyzed the correlation between the corresponding bands of Sentinel-2A and Landsat-8 on selected but limited sites, also evaluating the effects of spatial heterogeneity. Chastain et al. [29] proposed a cross-sensors comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere (TOA) spectral bands, providing the regression coefficients that allow to integrate MSI data with ETM+ and OLI over the conterminous United States (CONUS). Denaro and Lin [30] proposed the use of nonlinear models for the cross-sensor normalization of Landsat-7 and Landsat-8 imagery. More recently, Xie et al. [31] estimated cross-sensors linear transformations between Landsat-8 and Sentinel-2 based on only 76 image pairs spread all over the world. Cao et al. [32] proposed a similar analysis including Landsat-7 and deriving the transformations from a larger dataset but limited to the Chinese territory.

Other authors explicitly addressed the problem of vegetation index comparison. Li et al. [33] conducted a cross comparison of four VIs derived from six pairs of Landsat-7 and Landsat-8

images over Myanmar. Roy et al. [34] compared Landsat-7 ETM+ and Landsat-8 OLI and computed transformation coefficients for the integration of their spectral bands and the NDVI obtained from them. Chen et al. [35] proposed transformations to harmonize the NDVI computed from Landsat-4-5 multispectral scanner and thematic mapper (TM), elaborating on simulated data derived from Hyperion hyperspectral images. Mancino et al. [36] presented a specific case study in Italy comparing six VIs between Landsat-7 and Landsat-8 and finding statistically significant differences among four different land-cover classes.

In the last years, NASA and USGS developed the harmonized Landsat-8 and Sentinel-2 (HLS) surface reflectance (SR) dataset currently at version 2.0, which provides harmonized multispectral bands with global coverage [37]. However, at the time of the writing, there are no “ARD” for vegetation index TSA including the full Landsat constellation and Sentinel data. Such datasets would further expand and facilitate the use of EO by even technicians who are not experts in remote sensing.

Therefore, this study aims to propose a set of cross-sensor transformation coefficients, which are valid on a continental scale and easy to implement for end-users, in order to create harmonized vegetation index time series. Starting from the approach implemented by Chastain et al. [29] and Roy et al. [34], this work introduces four major novelties. First, from a methodological point of view, sampling strategy was improved to provide an estimation of the repeatability of the computed coefficients. Second, the analysis was performed using the recently released Landsat Collection-2 dataset, replacing the Collection-1 used in previous study. Moreover, the cross comparison was performed on vegetation indexes and not on the single bands. Third, the paired observations samples were gathered for the first time from the entire European continent. Finally, this study encompasses data from Landsat-5 (L5), Landsat-7 (L7), Landsat-8 (L8), and Sentinel-2 (S2) altogether, allowing the extraction of VIs time series starting from 1984.

In the following sections, we provide a summary of L5, L7, L8, and S2 characteristics and their products (Section II); a description of the methodology include preprocessing, data gathering, sampling, and statistical analysis (Section III), and finally, the results and the transformation coefficients are presented (Section IV).

II. MATERIALS

A. Landsat

Landsat program has been collecting Earth surface multispectral images since 1972, and it represents the longest living remote sensing mission in the world. It is a joint project between NASA, responsible for the satellite construction and launch, and USGS, which manages the archive and the distribution of data. The mission was designed to achieve efficient multispectral monitoring of land surface, with 30-m pixel average resolution and a revisit cycle of 16 days on every point on Earth. Every new Landsat satellite has been designed to create a pair constellation with the previous one still operating, resulting in an eight-day revisit coverage [6].

This study included datasets acquired by three different Landsat instruments: L5 TM, L7 enhanced TM Plus (ETM+), and L8 operational land imager (OLI).

After its launch in March 1984, L5 was operated by USGS until January 2013. For more than 29 years, it acquired over 2.5 million images of the Earth, largely exceeding its original three-year designed life. Its TM sensor captured the Earth surface spectral reflectance in six bands at 30-m spatial resolution and 120-m spatial resolution for the thermal band [38].

The L7 was launched in April 1999 and carried the ETM+ instrument. The ETM+ represents an improvement with respect to the TM sensors, with the addition of the panchromatic 15-m resolution band. From June 2003, when the scan line corrector (SLC) failed, L7 images were acquired and delivered with gaps, producing a loss of information up to the 22% [39]. The decommissioning of L7 began in mid-2021, leaving its orbit to the new Landsat-9.

L8 was launched in February 2013 equipped with the OLI and the thermal infrared sensor (TIRS). OLI measures the visible, NIR, and SWIR part of the electromagnetic spectrum, while TIRS operates in the thermal region. Following the spectral improvements achieved with ETM+ instrument, OLI was designed with a panchromatic 15-m band and eight 30-m spectral bands. In OLI, the new ultrablue band (Band 1) and the Band 9 (1.36–1.38 μm) are useful in coastal/aerosol studies and cirrus cloud detection, respectively. Table I gives a summary of L5, L7, and L8 spectral characteristics.

In order to improve Landsat data consistency and interoperability across sensors, the archive was reprocessed twice. After the first global Landsat-1 to Landsat-8 reprocessing into Collection-1 data [40], the second major reprocessing of the archive, performed in 2021, led to the release of Collection-2 data, which replaced Collection-1 starting from January 2022. Collection-2 major improvement regards geometry accuracy: for a better exploitation of archive interoperability, the L8 Ground Control Points were rebaselined to the ESA S2 Global Reference Image. In addition, the digital elevation model sources were updated, and the accessibility from commercial cloud-based environment was improved [41].

Since the release of the Collection-1 data in 2016, the Landsat products are structured in a three-level hierarchical quality inventory to grant consistency in Landsat data processing and traceability of data quality records [42]. The highest data quality products with high geolocation accuracy (root mean square error (RMSE) ≤ 12 m) are included in Tier-1.

In addition, three different product levels are delivered: the global Level-1 data [top-of-atmosphere (TOA)], Level-2 data (SR), and the U.S. ARD. Level-2 data products are obtained applying atmospheric correction to Level-1 products (standard Landsat products) with a Solar Zenith Angle lower than 76° [43]. In particular, the following algorithms are used: the land SR code (LaSRC) algorithm (version 1.5.0) [44] was applied to L8 OLI scenes, while L5 TM and L7 ETM+ SR products are generated using the Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm (version 3.4.0) [45].

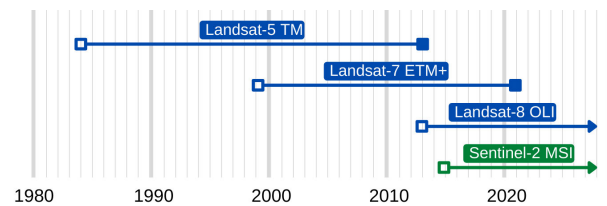


Fig. 1. Timeline of operational periods of Landsat and Sentinel satellites.

In this study, the SR Collection-2 Tier-1 datasets were used for all the Landsat instruments considered.

B. Sentinel-2

S2 mission by ESA is a twin polar-orbiting satellite phased at 180° to each other: Sentinel-2A and Sentinel-2B were launched in 2015 and 2017, respectively. Their orbit is angled of 98.62° and acquires images over land and coastal areas with a 290-km width swath, covering the Earth surface between the latitudes 56° South and 83° North. The mean local solar time at the descending node is 10:30 A.M. [46].

The two satellites are equipped with the multispectral instrument (MSI), a sensor that measures the Earth’s reflected radiance in 13 spectral bands, from VNIR to SWIR, providing imagery at different spatial resolution, ranging between 10 and 60 m, as summarized in Table I.

Users can freely download two different product levels. Level-1C products provide TOA reflectance data. The atmospheric correction of Level-1C scenes, by means of the Sen2Cor processor, produces the bottom-of-atmosphere (BOA) Level-2A products, i.e., the SR image [47], [48]. For the study, the SR Level-2A dataset was used.

Fig. 1 shows the operational timeline of the satellites considered in the study and gives a comparison of main mission characteristics.

C. Vegetation Indices

Four VIs were considered in this study: NDVI, EVI, SAVI, and NDMI. These VIs are obtained from the following bands: green, red, and red-edge bands (highly correlated with chlorophyll and other pigments contents); NIR band (sensitive to leaf structure), and the SWIR band (sensitive to water content) [20], [21]. NDVI is the most widespread index sensitive to chlorophyll computed from NIR and red bands [49], [50]. Being one of the most stable indexes, NDVI allows comparisons of seasonal and interannual changes in vegetation growth. Some improvements to NDVI were implemented to reduce the environmental effects to index variations. SAVI proposed by Huete [51] minimizes background soil brightness influences of NDVI. On the other hand, the EVI is used to reduce atmospheric effects that could lead to high biomass saturation [52]. Finally, NDMI consists in the normalized difference between NIR and SWIR, and it helps in vegetation water content assessment, useful, for example, when dealing with irrigations systems [53], [54]. In this study, these VIs were calculated according to (1)–(4), using the Landsat Collection-2 and Sentinel Level-2A SR datasets; the bands

TABLE I
SPECTRAL BANDS CHARACTERISTICS OF TM, ETM+, L8, AND MSI SENSORS. IN BOLD, THE BANDS CONSIDERED FOR THE VIS COMPUTATION

Landsat-5 TM			Landsat-7 ETM+			Landsat-8 OLI			Sentinel-2 MSI		
Bands	Wavelength (μm)	Res. (m)	Bands	Wavelength (μm)	Res. (m)	Bands	Wavelength (μm)	Res. (m)	Bands	Wavelength (μm)	Res. (m)
			B8 Pan	0.515 - 0.896	15	B8 Pan	0.503 - 0.676	15			
B1 Blue	0.45 - 0.52	30	B1 Blue	0.441 - 0.514	30	B1 Coastal/Aerosol	0.435 - 0.451	30	B1 Coastal aerosol	0.433 - 0.453	60
B2 Green	0.52 - 0.60	30	B2 Green	0.519 - 0.601	30	B2 Blue	0.452 - 0.512	30	B2 Blue	0.458 - 0.523	10
B3 Red	0.63 - 0.69	30	B3 Red	0.631 - 0.692	30	B3 Green	0.533 - 0.590	30	B3 Green	0.543 - 0.578	10
						B4 Red	0.636 - 0.673	30	B4 Red	0.650 - 0.680	10
									B5 Red Edge 1	0.698 - 0.713	20
									B6 Red Edge 2	0.733 - 0.748	20
									B7 Red Edge 3	0.773 - 0.793	20
									B8 NIR	0.785 - 0.900	10
B4 NIR	0.76 - 0.90	30	B4 NIR	0.772 - 0.898	30	B5 NIR	0.851 - 0.879	30	B8a NIRn	0.855 - 0.875	20
						B9 Cirrus	1.363 - 1.384	30	B9 Water Vapour	0.935 - 0.955	60
						B6 SWIR-1	1.566 - 1.651	30	B10 Cirrus	1.360 - 1.390	60
B7 SWIR-2	2.08 - 2.35	30	B7 SWIR-2	2.064 - 2.345	30	B7 SWIR-2	2.107 - 2.294	30	B11 SWIR	1.565 - 1.655	20
B6 TIR	10.4 - 12.5	120	B6 TIR	10.31 - 12.36	60	B10 TIR-1	10.60 - 11.19	100	B12 SWIR	2.100 - 2.280	20
						B11 TIR-2	11.50 - 12.51	100			

used are those highlighted in bold in Table I, which are the most similar across sensors. The coefficients in the formulas for EVI and SAVI are those suggested by USGS for the computation of on-demand vegetation indexes [44], [45], [55]

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \quad (1)$$

$$\text{EVI} = 2.5 \cdot \frac{\text{NIR} - \text{Red}}{\text{NIR} + 6 \cdot \text{Red} - 7.5 \cdot \text{Blue} + 1} \quad (2)$$

$$\text{SAVI} = 1.5 \cdot \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red} + 0.5} \quad (3)$$

$$\text{NDMI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}} \quad (4)$$

The coefficients in the formulas for EVI and SAVI are those suggested by USGS for the computation of on-demand vegetation indexes. Indeed, these were selected in this study to ensure, as much as possible, a generalized analysis able to accommodate most land-cover types. For example, in (3), the L soil correction factor parameter, which varies with vegetation amount between 0 and 1, was set to 0.5, an intermediate value that is documented to work well in most conditions [20], [23], [56].

III. METHODOLOGY

The vegetation indexes derived from the spectral bands of the TM, ETM+, OLI, and MSI sensors were compared. The sensors were compared in pairs, by randomly sampling the indexes values from overlapping images acquired with a maximum delay of one day, as detailed in the following sections. As a consequence, to ensure statistically robust samples of data, the comparison was possible just between those sensors, which were actively operating for a common period of time of at least two years (at the time of the writing). For this reason, the newer Landsat-9 sensor is not included in this study.

A. Study Area

The study area covers the entire European continent (Fig. 2), which was tiled in 100 subregions for computational reasons. This area comprehends a wide spectrum of land-cover types

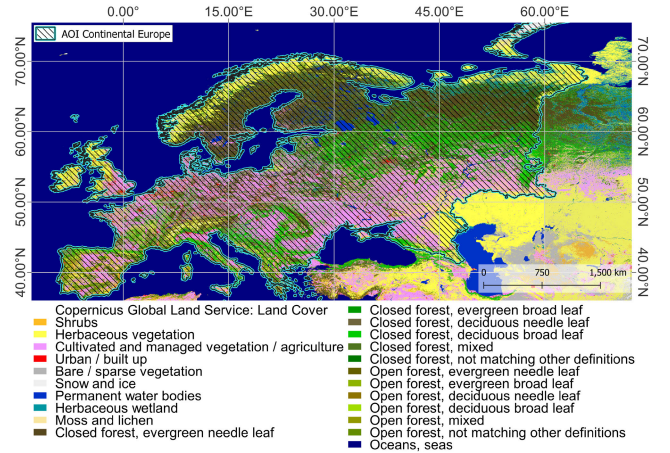


Fig. 2. AOI covered by the analysis including continental Europe. The base map shows the different land-cover classes as of 2018 [57].

and ecosystems, encompassing cultivated land (25%), natural vegetation (up to 65%), water bodies and wetland (6%), and urban area (2%) according to the Copernicus Global Land Cover 2018 [57]; thus, it provides an exhaustive and varied set of data.

B. Data Gathering

The data sampling was completely performed in GEE accessing the Landsat (Collection-2) and Sentinel-2 SR datasets available in its catalog.

First of all, the image collections were filtered based on time, space, and cloud cover. Specifically, at least two common years of acquisition were selected between the two missions; the search area was limited to Europe, and the maximum image cloud-cover percentage was set to 1% for Landsat data and to 0.1% for S2 data. This difference in percentages was derived empirically after several tests, and it is intended to compensate for the higher revisit time of S2, resulting in higher data availability and for the lower effectiveness of the cloud masking algorithm of S2 [58]. In order to obtain balanced datasets concerning images acquired in different seasons and, thus, in different vegetative states, the dataset was split into two different time spans: one from October to

March (autumn-wintertime) and one from April to September (spring-summertime). Because of cloud-free images shortage during wintertime and to balance the datasets between autumn-wintertime images and spring-summertime images, the search time span for the autumn-wintertime images was doubled. In addition, the search time span was carefully adjusted to ensure a similar population of valid pixels for each sensors couple, considering also the operational period of each sensor and some specific peculiarities, such as the SLC issues of L7. Table II provides specific information about the metadata filters applied to produce the dataset used for the analysis. In particular, time span, covering summer and winter period separately, and cloud-cover percentage are the filters applied to a single satellite collection. The total record (Tot records) is the number images, satisfying those filters for each satellite. The Joined collection is the number of pairs of images satisfying clouds and time span filters, which were acquired over the same area almost in the same date (± 1 day) by the two satellites considered in the cross-sensor analysis (the maps of the footprint intersections of these acquisitions are provided in the Supplementary Materials).

C. Pixel Masking

At this point, the selected images were pixelwise masked on the basis of the pixel quality assessment (QA) bitmask band. For Landsat products, it was generated from the C function of mask (CFMask) algorithm. The CFMask derives from the function of mask (FMask), which is able to label the scene pixels as cloud, cloud shadow, cirrus, snow/ice, or water, and provides a bit-mapped values output [59], [60]. This product was used to remove high- and medium-confidence clouds, dilated clouds, cloud shadows, snow/ice, and water pixels, in order to include in the analysis only clear land pixels. Only for L8 products, it was possible to mask also pixel marked as high confidence cirrus.

The same process was performed for the S2 images by means of the scene classification map (SCL) QA band, which, likewise the one for the Landsat products, labels the pixels on the basis of a classification process and, thus, allows the user to easily perform pixelwise masking. High- and medium-probability clouds, cloud shadows, cirrus, water, and snow/ice pixels were removed, accordingly with the masking process done for the Landsat products [61].

Furthermore, saturated and out-of-range pixels were masked using the radiometric saturation QA bands and valid value range. This means that all the saturated pixels and the pixels with a value of the vegetation index outside the range of interest, which is [0, 1] for NDVI, EVI and SAVI, and [-1, 1] for NDMI, were discarded.

D. Image Coupling, Coregistration, and Reprojection

The images, or portion of images, of two different sensors, filtered and masked as described above, which are spatially overlapping and acquired within 24 h, were paired, finely co-registered with each other, reprojected, to make sure the images shared the same coordinate reference system, and resampled at the coarsest resolution (30 m). This was done to

avoid differences in the VI values due to land-cover changes, bad spatial overlap, or differences in pixel size.

Despite the maximum time difference of 24 h between the two images of each pair, which should ensure no land-cover changes occurred, there are still some pixels showing huge reflectance differences. Therefore, the paired images were further masked following the methodology proposed by Roy et al. [34], which is based on the pixelwise difference in the blue band values. However, using here images already corrected at SR, pixels with a difference greater than the 50% of the average were discarded.

E. Sampling

For computational limit reasons, cross-sensors comparisons were not performed on the entire pixel population but on statistical samples randomly extracted from the population of valid pixels (after masking) belonging to the paired images. For each couple of sensors, samples were independently selected on a purely random basis. A statistical analysis was performed to assess the optimal sample size looking for a trade-off between computational complexity and statistical significance.

Different sample sizes (in the range between 1000 and 500 000 pixels) were tested by repeating the extraction 100 times and evaluating the variance of the cross-sensors parameters (described in Section III-F).

Fig. 3 shows the analysis performed on the L7 and L8 NDVI pair, here presented as an example. As it can be seen from these plots, the decrease of the variance with the increase of the sample size is asymptotic, and higher number of pixels would result only in an unfruitful increase of the computational burden. The optimal sample size was, therefore, set to 300 000 pixels, which corresponds—for the NDVI index—to a standard deviation of the linear regression coefficients lower than 0.0004 (i.e., intercept equal to 0.0002 and slope equal to 0.0003).

The case of the comparison between L5 and L7 is singular because of the exceptionally long period in which both the satellites were contemporary operational. This circumstance offers the opportunity to further analyze the linear relationships, in particular to investigate possible fluctuations of the estimated coefficients during time. For this purpose, further samples of the same size were extracted, one for each year in the period 1999–2011.

F. Cross-Sensor Analysis

For each couple of sensors, two ordinary least square (OLS) regressions and a reduced major axis (RMA) regression were computed. The OLS regression allows to find a transformation function from a sensor to the other: the slope and intercept parameters change depending on which variable (i.e., which sensor) is defined as dependent or independent; thus, the OLS regression was performed twice, inverting dependent and independent variable each time, in order to provide transformation functions from a sensor to the other and vice versa [34]. On the contrary, the RMA regression is performed only once, since the relationship between the interchanged variables can be obtained with a simple algebraic operation [62]. In any

TABLE II
SENTINEL AND LANDSAT DATA FILTERING. (a) L5–L7. (b) L8–S2. (c) L7–S2. (d) L7–L8

(a)						
	Start	Stop	Month span	Cloud cover	Records	Tot records
L7 summer	01/01/1999	01/01/2007	Apr - Sep	1%	7721	12148
L7 winter	01/01/1999	01/01/2012	Oct - Mar	1%	4427	
L5 summer	01/01/1999	01/01/2007	Apr - Sep	1%	4821	7344
L5 winter	01/01/1999	01/01/2012	Oct - Mar	1%	2523	
Joined collection						3812

(b)						
	Start	Stop	Month span	Cloud cover	Records	Tot records
S2 summer	01/01/2018	01/01/2020	Apr - Sep	0.10%	14824	25403
S2 winter	01/01/2018	01/01/2020	Oct - Mar	0.10%	10579	
L8 summer	01/01/2018	01/01/2020	Apr - Sep	1%	4141	7121
L8 winter	01/01/2016	01/01/2020	Oct - Mar	1%	2980	
Joined collection						4924

(c)						
	Start	Stop	Month span	Cloud cover	Records	Tot records
S2 summer	01/01/2018	01/01/2020	Apr - Sep	0.10%	14824	25403
S2 winter	01/01/2016	01/01/2020	Oct - Mar	0.10%	10579	
L7 summer	01/01/2018	01/01/2020	Apr - Sep	1%	3294	5275
L7 winter	01/01/2016	01/01/2020	Oct - Mar	1%	1981	
Joined collection						4520

(d)						
	Start	Stop	Month span	Cloud cover	Records	Tot records
L8 summer	01/01/2017	01/01/2021	Apr - Sep	1%	7994	13680
L8 winter	01/01/2013	01/01/2021	Oct - Mar	1%	5686	
L7 summer	01/01/2017	01/01/2021	Apr - Sep	1%	6366	10321
L7 winter	01/01/2013	01/01/2021	Oct - Mar	1%	3955	
Joined collection						9608

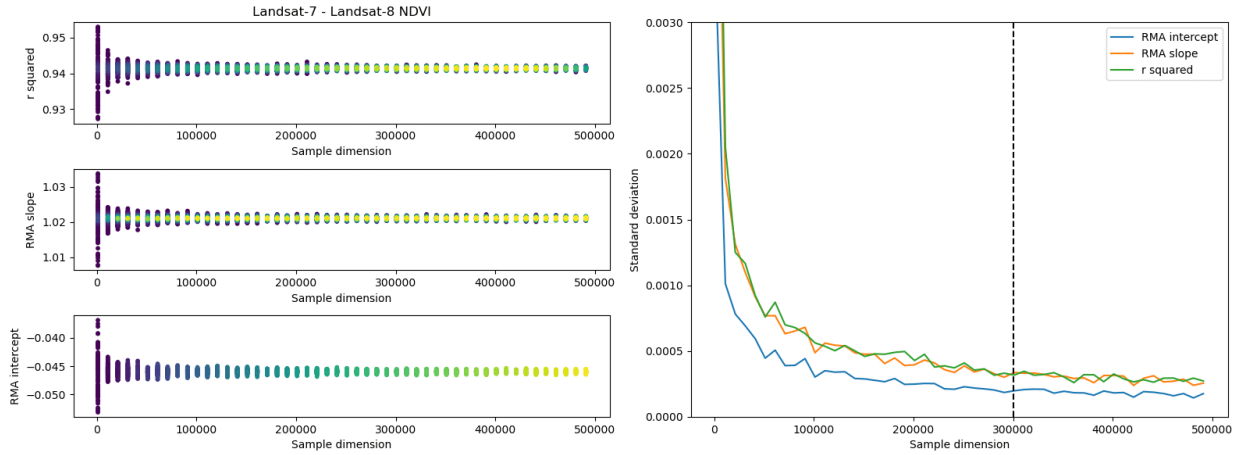


Fig. 3. Variation of RMA coefficients for L7 and L8 NDVI image pairs. (a) RMA intercept, slope, and R^2 coefficients values are plotted against the sample size; the color ramp is defined by values frequency (points represented in yellow are those with the most frequent values). (b) Standard deviation of RMA intercept, slope, and R^2 coefficients for the 100 random extractions for each sample size.

case, it is assumed that both the dependent and independent variables are subject to errors, which is appropriate because of the possible residual errors that the data may have, such as atmospheric correction and sensor calibration errors [29], [34].

The goodness of the fit of the regressions was evaluated with the coefficient of determination (R^2), while the significance of the regressions was defined by the overall F -statistic p value [34].

In order to provide an overall measure of similarity between the datasets, three different difference metrics were derived as follows:

$$MD = \sum_i^n \frac{v_i^A - v_i^B}{n} \quad (5)$$

$$RMSD = \sqrt{\frac{\sum_i^n (v_i^A - v_i^B)^2}{n}} \quad (6)$$

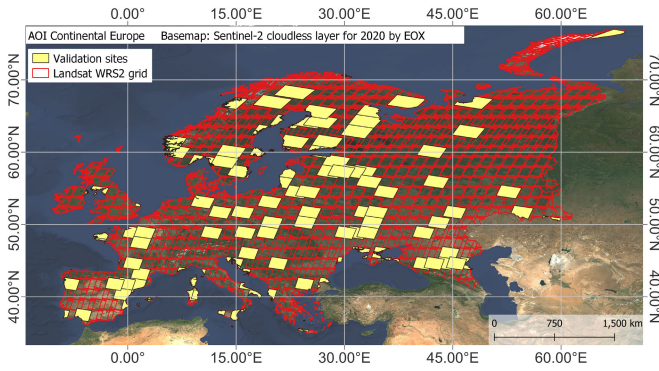


Fig. 4. Landsat standard full scene footprints. The ones randomly selected as validation areas are highlighted in yellow.

$$\text{MRD} = \frac{\sum_i^n \frac{v_i^A - v_i^B}{0.5(v_i^A + v_i^B)}}{n} 100 \quad (7)$$

where MD, RMSD, and MRD are the mean difference (MD), the root-mean-square deviation (RMSD), and the mean relative difference (MRD) between corresponding values of the generic sensor A and sensor B for n pixels, respectively [29], [34].

G. Validation

The effectiveness of the suggested linear transformations was checked by evaluating the differences in the VIs values between every couple of sensors before and after the harmonization. These differences are computed on fully independent samples, obtained as follows. Exploiting the Landsat Worldwide Reference System, the 30% of the nominal scene centers over Europe were randomly selected. The areas belonging to these Landsat standard full scenes (tiles defined by their paths and rows) were excluded from the extraction process of the samples used for the cross-sensor analyses described in Section III-F (Fig. 4). This 30% of the tiles was used for the extraction of the independent samples for validation. For each couple of sensors, one validation sample was created, consisting of at least two million random points. These are the points used to statistically analyze the differences in the VIs values. The balance between the validation and training tiles, in terms of land-cover classes, was verified (average difference around 1% and in any case lower than 6%).

IV. RESULTS

The results obtained from the comparison of the vegetation indexes derived from different sensors are reported in Table III and Figs. 5–11, and they are organized by couple of sensors: OLI and MSI, ETM+ and MSI, ETM+ and OLI, and TM and ETM+. As explained above, for every sensor couple, the results are obtained through 100 independent extractions of 300 000 paired observations. These analyses highlighted both differences and similarities between these products and allowed to derive transformation coefficients to be used for a harmonized integration of the different datasets.

A. OLI and MSI

In Table III(a) and Fig. 5, the results of the analysis involving the OLI and MSI sensors are presented. Consider-

ing the paired observations collected between 2016 and 2020, the lowest MD between corresponding indices was found in the NDVI, equal to -0.0004 , while the highest in NDMI, equal to 0.0248 . The RMSD values are quite similar for all the VIs, ranging from 0.0455 (SAVI) to 0.0586 (NDMI). All the regression models are highly significant, all showing R^2 values higher than 0.90 (p values < 0.0001). In this case, the highest deviation from the slope identity is given by the EVI, with an RMA slope value equal to 1.0835 . As expected, the computed coefficients of the regression are quite stable in the 100 independent sample extractions, as can be observed by the standard deviations reported in Table III(a). Indeed, the slope coefficients of EVI show the highest standard deviation for the regression, having OLI as the independent variable (0.0007 for both the OLS the RMA).

To evaluate the effectiveness of the transformations, the RMA coefficients were applied to the sampled OLI pixels belonging to the validation set, to compute the harmonized hOLI values that are expected to compare with MSI better. The hOLI values were obtained using OLI as independent variable and applying the coefficient in Table III(a). The differences between the hOLI VI values and the original values from the paired MSI are then computed (hDiff), and their histograms are shown in Fig. 6. The harmonization decreased the residuals for every index: NDMI has the highest improvement with a decrease of the MD of 0.0247 because of harmonization followed by the EVI with a decrease of 0.0109 . Overall, the means of the harmonized residuals are very low, between -0.0008 (NDMI) and 0.0002 (SAVI).

B. ETM+ and MSI

The comparison between ETM+ and MSI instruments acquisition is showed in Table III(b) and Fig. 7. The MD of the sampled values ranges from -0.0286 (EVI) to 0.0059 (NDMI). The RMSD ranges from 0.0470 (SAVI) to 0.0600 (NDVI). The MRD values are all quite low, ranging from 0.2335 (NDMI) to -7.6555 (EVI). The R^2 values and p values indicate a high significance of the regression models ($R^2 > 0.92$ and p values < 0.0001). The index with the highest deviation from slope identity is the EVI, with a slope value of 1.1083 .

Also, in this case, the RMA transformation coefficients were applied to the samples belonging to the validation set, using ETM+ observations as independent variable, to produce the harmonized hETM+ VIs. The hETM+ and corresponding MSI residuals (hDiff) distribution and mean values are presented in Fig. 8. Also here, the harmonization decreased the residuals for every index: EVI has the highest improvement with a decrease of the MD of 0.0331 . Overall, the means of the harmonized residuals are low, between 0.0033 (NDMI) and 0.0069 (NDVI).

C. ETM+ and OLI

The MDs of the sampled ETM+ and OLI paired observations range from a minimum of 0.0147 , for the EVI, to a maximum of 0.0347 , for the NDVI. The RMSD values go from 0.0416 (SAVI) and 0.0657 (NDVI), while the MRD

TABLE III

VEGETATION INDEXES SENSOR TRANSFORMATION FUNCTIONS: SLOPE AND INTERCEPT WITH THEIR STANDARD DEVIATIONS IN BRACKETS, R^2 COEFFICIENT, MD, RMSD, AND MRD. (a) OLI TO MSI. (b) ETM+ TO MSI. (c) ETM+ TO OLI. (d) TM TO ETM+

(a)						
VI	Regression type	Linear transformation functions	R^2	MD	RMSD	MRD
NDVI	RMA	MSI = 1.0715(± 0.0003) OLI -0.0407(± 0.0002)	0.9417	-0.0004	0.0573	1.9412
	OLS	OLI = 0.9056(± 0.0003) MSI +0.0538(± 0.0002)				
	OLS	MSI = 1.0398(± 0.0004) OLI -0.0225(± 0.0003)				
EVI	RMA	MSI = 1.0835(± 0.0007) OLI -0.0176(± 0.0002)	0.9045	-0.0102	0.0552	-0.8342
	OLS	OLI = 0.8778(± 0.0006) MSI +0.0317(± 0.0002)				
	OLS	MSI = 1.0305(± 0.0007) OLI +0.0001(± 0.0002)				
SAVI	RMA	MSI = 1.0624(± 0.0005) OLI -0.0183(± 0.0001)	0.9108	-0.0021	0.0455	1.2376
	OLS	OLI = 0.8983(± 0.0005) MSI +0.0314(± 0.0002)				
	OLS	MSI = 1.0139(± 0.0005) OLI -0.0025(± 0.0002)				
NDMI	RMA	MSI = 1.0053(± 0.0003) OLI -0.0254(± 0.0001)	0.9426	0.0248	0.0586	1.3434
	OLS	OLI = 0.9658(± 0.0004) MSI +0.0279(± 0.0001)				
	OLS	MSI = 0.9761(± 0.0004) OLI -0.0221(± 0.0001)				

(b)						
VI	Regression type	Linear transformation functions	R^2	MD	RMSD	MRD
NDVI	RMA	MSI = 1.0454(± 0.0004) ETM+ -0.0016(± 0.0002)	0.9442	-0.0231	0.0600	-3.7007
	OLS	ETM+ = 0.9295(± 0.0004) MSI +0.0168(± 0.0002)				
	OLS	MSI = 1.0158(± 0.0004) ETM+ +0.0145(± 0.0002)				
EVI	RMA	MSI = 1.1083(± 0.0006) ETM+ -0.0059(± 0.0002)	0.9202	-0.0286	0.0586	-7.6555
	OLS	ETM+ = 0.8656(± 0.0005) MSI +0.0181(± 0.0002)				
	OLS	MSI = 1.0632(± 0.0006) ETM+ +0.0085(± 0.0002)				
SAVI	RMA	MSI = 1.0707(± 0.0005) ETM+ -0.0017(± 0.0001)	0.9235	-0.0203	0.0470	-5.7299
	OLS	ETM+ = 0.8975(± 0.0005) MSI +0.0137(± 0.0001)				
	OLS	MSI = 1.0289(± 0.0005) ETM+ +0.0113(± 0.0002)				
NDMI	RMA	MSI = 1.0044(± 0.0004) ETM+ -0.0063(± 0.0001)	0.9425	0.0059	0.0531	-0.2335
	OLS	ETM+ = 0.9666(± 0.0004) MSI +0.0087(± 0.0001)				
	OLS	MSI = 0.9751(± 0.0005) ETM+ -0.0037(± 0.0001)				

(c)						
VI	Regression type	Linear transformation functions	R^2	MD	RMSD	MRD
NDVI	RMA	ETM+ = 1.0218(± 0.0004) OLI -0.0465(± 0.0002)	0.9419	0.0347	0.0657	8.7660
	OLS	OLI = 0.9498(± 0.0003) ETM+ +0.0602(± 0.0002)				
	OLS	ETM+ = 0.9917(± 0.0004) OLI -0.0302(± 0.0002)				
EVI	RMA	ETM+ = 0.9985(± 0.0004) OLI -0.0143(± 0.0001)	0.9333	0.0147	0.0453	6.7041
	OLS	OLI = 0.9675(± 0.0005) ETM+ +0.0243(± 0.0001)				
	OLS	ETM+ = 0.9646(± 0.0004) OLI -0.0038(± 0.0001)				
SAVI	RMA	ETM+ = 1.0035(± 0.0004) OLI -0.0202(± 0.0001)	0.9383	0.0192	0.0416	8.3763
	OLS	OLI = 0.9653(± 0.0005) ETM+ +0.0292(± 0.0001)				
	OLS	ETM+ = 0.9721(± 0.0004) OLI -0.0106(± 0.0001)				
NDMI	RMA	ETM+ = 0.9966(± 0.0003) OLI -0.0249(± 0.0001)	0.9502	0.0252	0.0582	0.0162
	OLS	OLI = 0.9781(± 0.0004) ETM+ +0.0266(± 0.0001)				
	OLS	ETM+ = 0.9715(± 0.0004) OLI -0.0226(± 0.0001)				

(d)						
VI	Regression type	Linear transformation functions	R^2	MD	RMSD	MRD
NDVI	RMA	ETM+ = 1.0377(± 0.0003) TM +0.0012(± 0.0002)	0.9374	-0.0189	0.0604	-3.8367
	OLS	TM = 0.9330(± 0.0003) ETM+ +0.0138(± 0.0002)				
	OLS	ETM+ = 1.0047(± 0.0004) TM +0.0167(± 0.0002)				
EVI	RMA	ETM+ = 0.9929(± 0.0005) TM +0.0017(± 0.0001)	0.9189	0.0003	0.0468	0.1268
	OLS	TM = 0.9654(± 0.0006) ETM+ +0.0102(± 0.0001)				
	OLS	ETM+ = 0.9518(± 0.0005) TM +0.0135(± 0.0001)				
SAVI	RMA	ETM+ = 1.0052(± 0.0004) TM +0.0020(± 0.0001)	0.9291	-0.0034	0.0388	-1.3164
	OLS	TM = 0.9589(± 0.0005) ETM+ +0.0081(± 0.0001)				
	OLS	ETM+ = 0.9689(± 0.0005) TM +0.0119(± 0.0001)				
NDMI	RMA	ETM+ = 1.0137(± 0.0004) TM +0.0058(± 0.0001)	0.9301	-0.0066	0.0576	-1.5395
	OLS	TM = 0.9514(± 0.0005) ETM+ -0.0037(± 0.0001)				
	OLS	ETM+ = 0.9776(± 0.0006) TM +0.0077(± 0.0001)				

is lower than 9 for all the vegetation indexes [Table III(c)]. The R^2 values are very high for all the regressions, with values higher than 0.93. The significance of the regression is confirmed also by the very low values of the F -statistic p value (p value < 0.0001). The highest deviation from the slope identity is given by the NDVI, with a slope value equal to 1.0218. In this case, the standard deviations on the coefficients are small (for the slope lower than 0.0005 and for the intercept lower than 0.0002).

Over all the validation samples, the harmonized hOLI was computed by means of the RMA using OLI observations as independent variable. The residuals distributions and mean values of the paired hOLI and ETM+ difference (hDiff) are presented in Fig. 10. Again, the harmonization decreased the residuals for every index: in this case, NDVI has the highest improvement with a decrease of the MD of 0.0342. Overall, the means of the harmonized residuals are very low, between -0.0007 (NDMI) and -0.0002 (EVI).

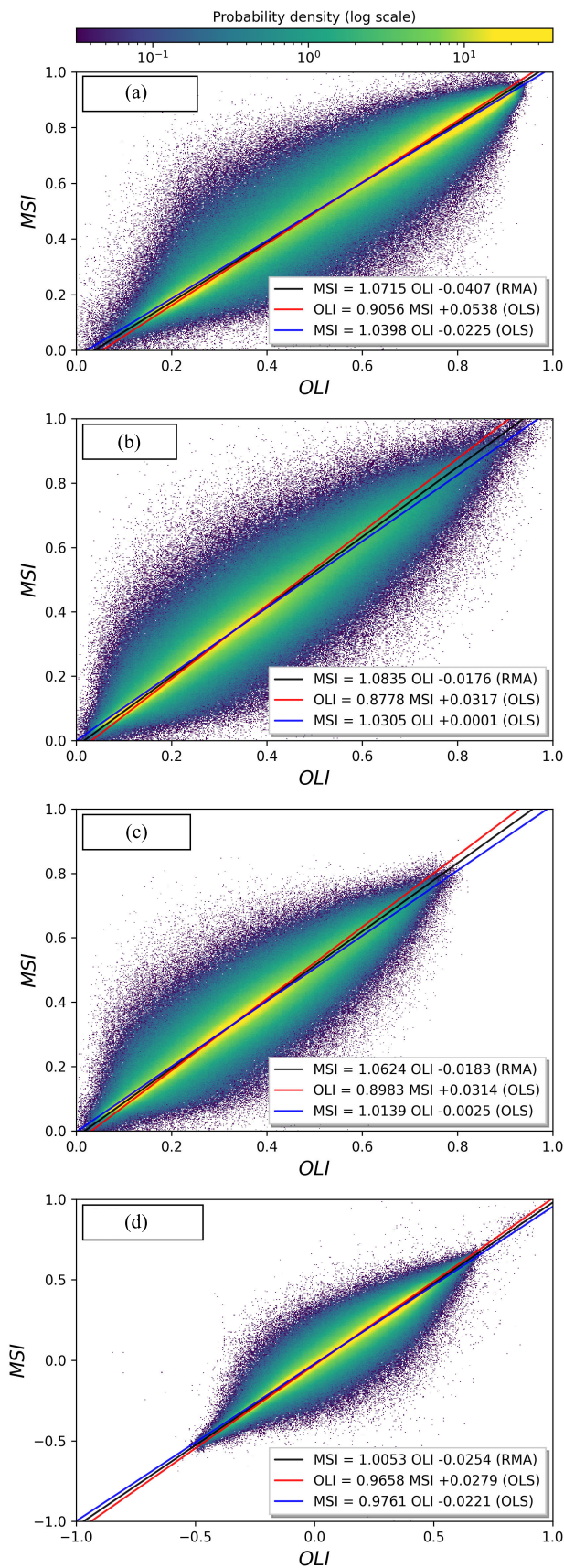


Fig. 5. Scatterplots of the VIs for S2 MSI (vertical axis) against L8 OLI (horizontal axis). (a) NDVI. (b) EVI. (c) SAVI. (d) NDMI. The plot colors illustrate the probability density of VIs values with logarithmic scale. The solid lines show the three regression fits.

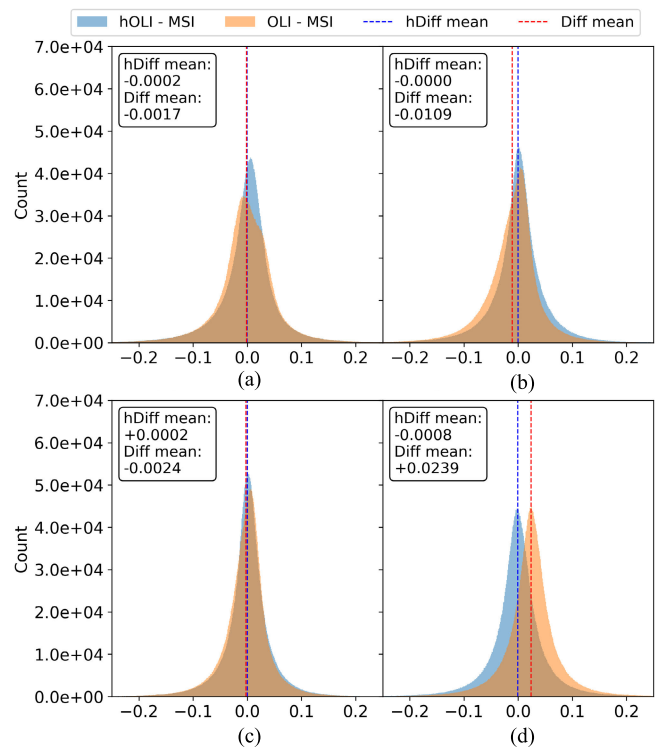


Fig. 6. Residuals distribution of MSI and OLI applying the RMA coefficients (OLI independent variable) for (a) NDVI, (b) EVI, (c) SAVI, and (d) NDMI. Red dashed lines represent the mean values.

D. TM and ETM+

The results of the comparison between TM and ETM+ are summarized in Table III(d) and graphically displayed in Fig. 11. The MD ranges from 0.0003 (EVI) to -0.0189 (NDVI), and the RMSD from 0.0604 (NDVI) to 0.0388 (SAVI). The MRD is lower than 4, in absolute value, for all the indexes. All the regression models show a high significance (R^2 values > 0.918 and p value < 0.0001), and the model parameters are very small in magnitude. The NDVI is the one showing the highest deviation from the slope identity, with a slope value equal to 1.0377.

Again, for all the validation samples, the harmonized hTM VIs were computed by the RMA using TM derived indexes as independent variables. The residuals distribution and mean values of the difference between hTM and ETM+ are presented in Fig. 12. Similar to the previous cases, the harmonization decreased the residuals for every index: NDVI has the highest improvement with a decrease of the MD of 0.0195. Overall, the means of the harmonized residuals are very low, between 0.0001 (NDMI) and 0.0013 (NDVI).

E. Time

Because of the longer period of contemporary acquisitions of these two sensors, an additional analysis was performed here to investigate the stability over time of the computed parameters for the transformations. The intercepts and slopes of the RMA transformations recomputed for all the VIs but using 12 samples, each one composed of 300 000 points and

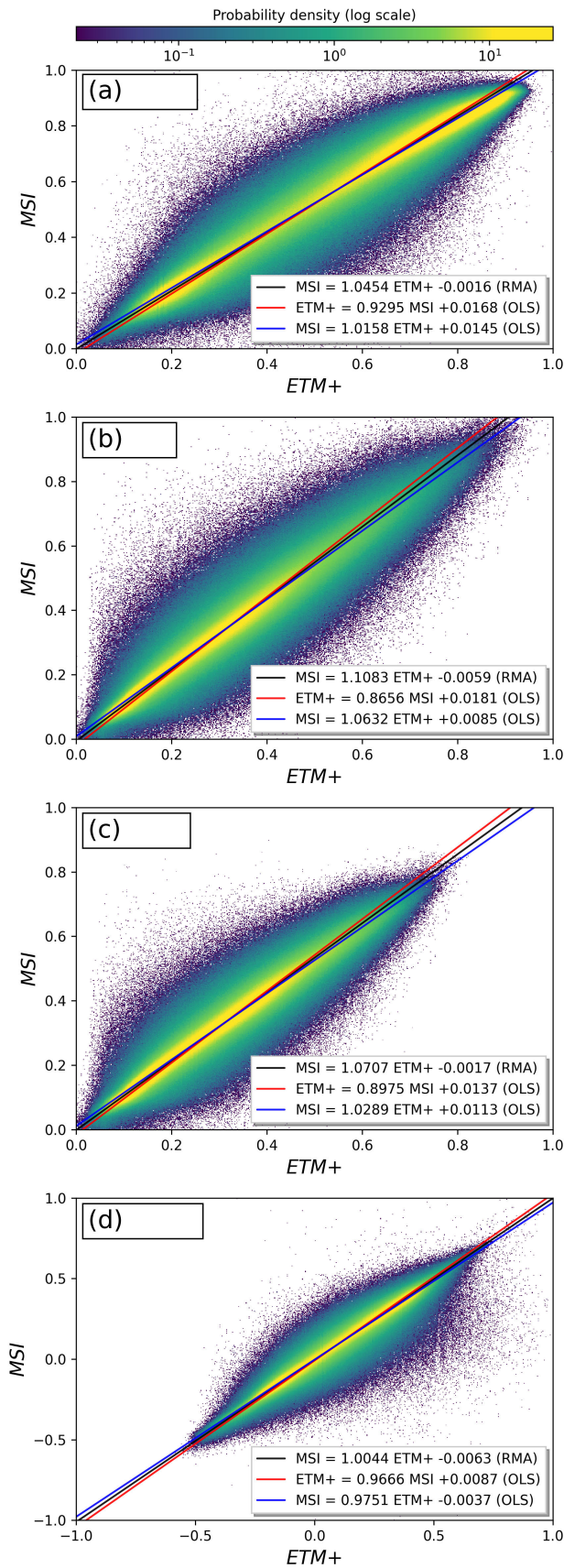


Fig. 7. Scatterplots for all of the vegetation indexes for S2 MSI (vertical axis) against L7 ETM+ (horizontal axis). (a) NDVI. (b) EVI. (c) SAVI. (d) NDMI. The plot colors illustrate the probability density of VIs values with logarithmic scale. The solid lines show the three regression fits.

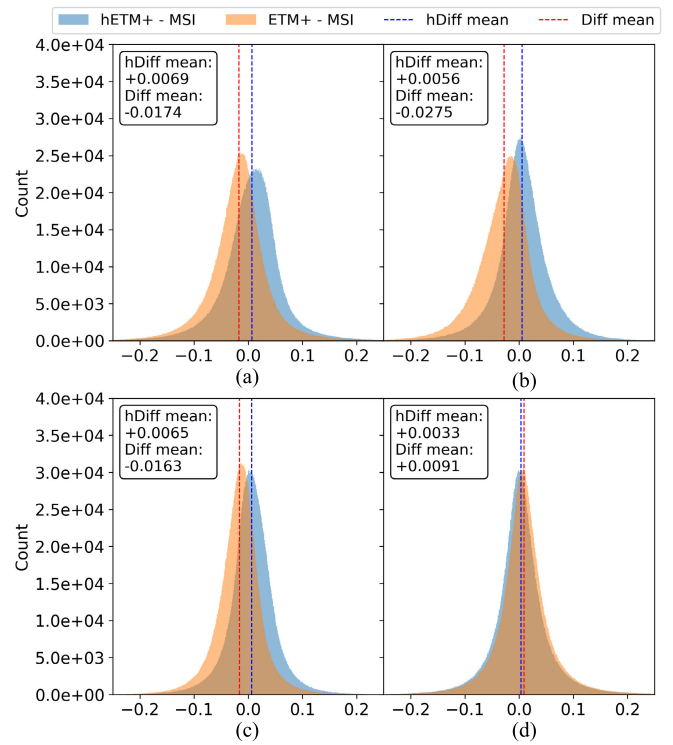


Fig. 8. Residuals distribution of MSI and ETM+ applying the RMA coefficients (ETM+ independent variable) for (a) NDVI, (b) EVI, (c) SAVI, and (d) NDMI. Red dashed lines represent the mean value.

extracted in a different year between 1999 and 2011 (as stated in Section III-E). Fig. 13 shows, for every examined VI, the resulting 12 sets of intercept and slope parameters and how their values change over years, compared with the suggested ones presented in Table III(d). A sensible fluctuation can be noted for both intercepts and slopes from year to year, even though no apparent trends are detectable.

V. DISCUSSION

In general, the presented results confirmed the data continuity expected from Landsat Collection-2. Furthermore, the calculated statistics validated its interoperability with the S2 mission. Indeed, the linear regression functions computed for each sensor pair showed a good agreement between the values of the indices, with R^2 always greater than 0.9 and the great majority of pairs falling close to the bisector of the scatter plots. Nevertheless, the application of a transformation seems advisable, when trying assessing changes and detecting anomalies in multisensor time-series analyses. Considering, for example, the NDVI computed on ETM+ and OLI (Fig. 10), a sensible improvement is observed in the MD (from -0.0337 to 0.0005 with the transformation). Despite the very little variation in the numbers, this improvement becomes sensible when dealing with applications that involve the assessment of subtle changes over time. Such applications, which are the target of this study, greatly benefit from a reduction in the noise in the time series, because it allows a clearer detection of the actual change [63]. This is the case

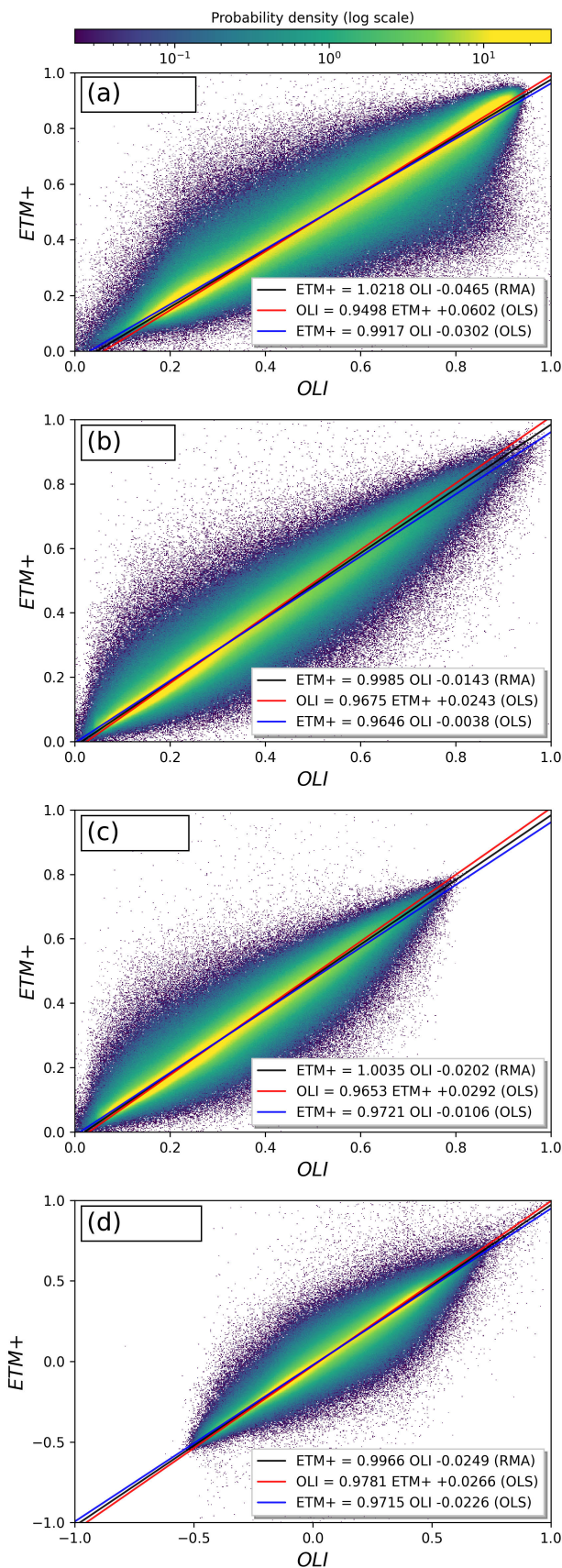


Fig. 9. Scatterplots for all of the vegetation indexes for L7 ETM+ (vertical axis) against L8 OLI (horizontal axis). (a) NDVI. (b) EVI. (c) SAVI. (d) NDMI. The plot colors illustrate the probability density of VIs values with logarithmic scale. The solid lines show the three regression fits.

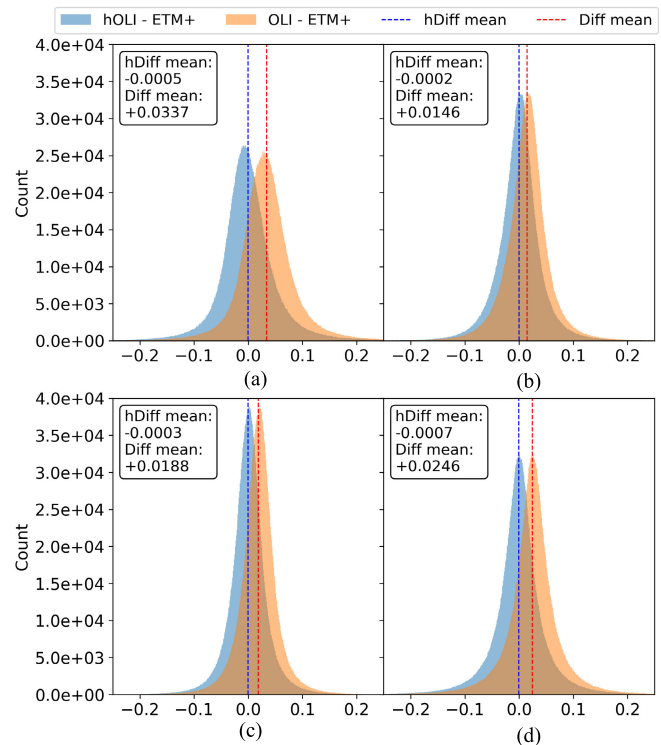


Fig. 10. Residuals distribution of ETM+ and OLI applying the RMA coefficients (OLI independent variable) for (a) NDVI, (b) EVI, (c) SAVI, and (d) NDMI. Red dashed lines represent the mean value.

of the study held by Gouveia et al. [64] where a change in the NDVI of 0.025 was found out to be correlated with vegetation stress due to drought.

Observing the graphs in Figs. 5, 7, 9, and 11, two considerations can be made. First, NDVI observations are the most disperse with respect to the linear function fits. Second, EVI and SAVI sample densities are higher for lower values of the indexes, while NDVI is more homogeneously distributed along the entire range, showing a great concentration also for very high values of the index. In addition, the NDVI scatter is the most dispersed with respect to its linear regression lines: this is especially true when L7 ETM+ is considered. Specifically, in the comparison between ETM+ and OLI, some points present sensibly higher values of NDVI from one sensor compared with the other.

Likely, the explanation of the observed behaviors can be found in a well-known limitation of the NDVI, i.e., the saturation, which occurs especially in the areas of high biomass [52], [65]. This would explain the higher point density for high values of NDVI, if compared with those of the other indexes [for example, compare Fig. 11(a) with 11(b) and (c)]. SAVI and EVI are instead able to reduce noise and saturation; this is confirmed by the values reported in Table III in which the NDVI difference metrics are higher, compared with the other indices, almost for every sensors pair. The signal-to-noise ratio of SAVI, for example, is reported to be up to five times higher than NDVI depending on the green cover percentage [66]. The peculiar dispersion observed in the NDVI scatter plot

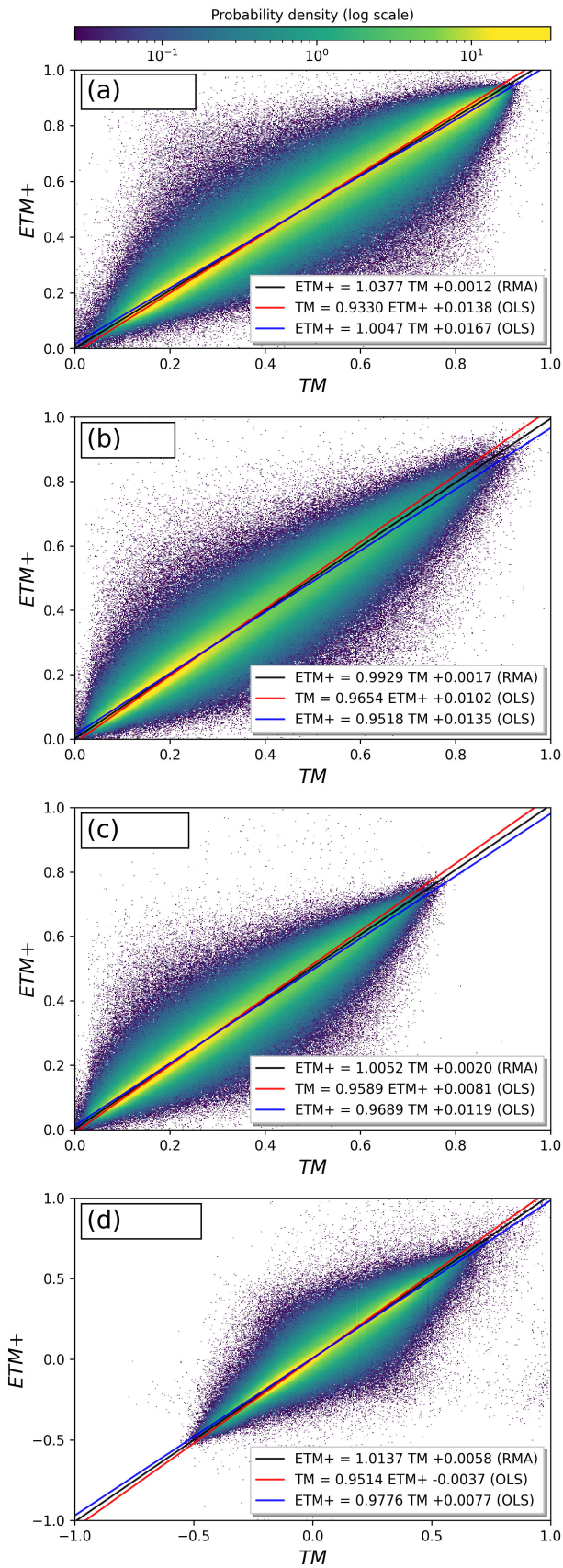


Fig. 11. Scatterplots for all of the vegetation indexes for L7 ETM+ (vertical axis) against L5 TM (horizontal axis). (a) NDVI. (b) EVI. (c) SAVI. (d) NDMI. The plot colors illustrate the probability density of VIs values with logarithmic scale. The solid lines show the three regression fits.

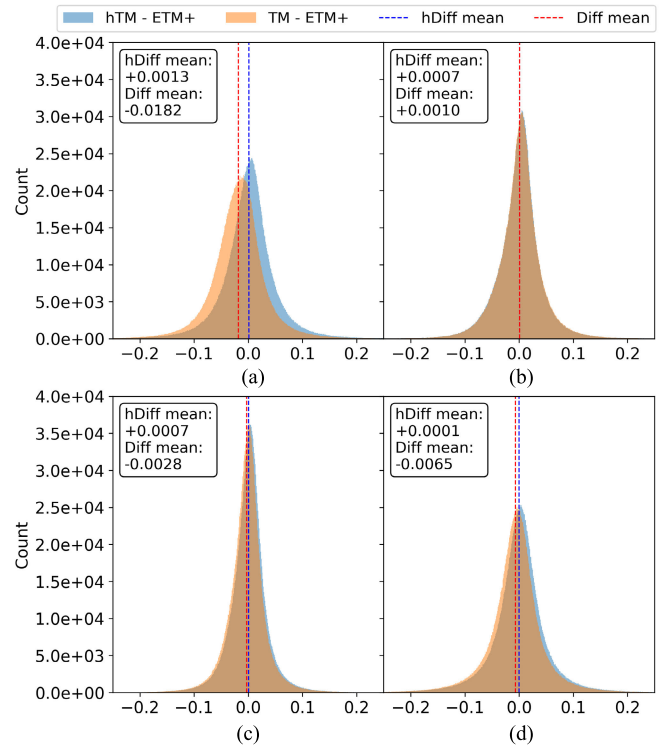


Fig. 12. Residuals distribution of ETM+ and TM applying the RMA coefficients (OLI independent variable) for (a) NDVI, (b) EVI, (c) SAVI, and (d) NDMI. Red dashed lines represent the mean value.

between ETM+ and OLI suggests an influence of the spectral response function on the saturation phenomenon. Indeed, some outliers seem due to the saturation of the OLI (NDVI greater than 0.75 for OLI and NDVI between 0.3 and 0.6 for ETM+), while others show the opposite pattern even though with lower saturation levels [Fig. 9(a)]. More details about spectral response function comparison between L7 and L8 are provided by Roy et al. [34]. In addition, as previously highlighted by Irons et al. [67], the NIR band of OLI was designed narrower to exclude the ETM+ water vapor absorption feature at $0.825 \mu\text{m}$ in Band 4. This slight difference between OLI and ETM+ NIR bands has been already highlighted by Chastain et al. [29], where the NIR bands exhibited the most variation in the L8–L7 cross-sensor comparison on Level-1 products, reaching the highest identity line deviation (slope of 1.1406).

Another aspect that can contribute to the different distribution of Vis values is calibration error propagation. According to the study by Miura et al. [68], assessing the impact of reflectance calibration uncertainties on VIs derived by MODIS Terra mission, such uncertainties propagation is index-dependent, and different patterns and magnitudes were found for NDVI, EVI, and SAVI.

However, it is important to remark that the majority of the pairwise values are concentrated around the regression lines, and the model well fit the data as demonstrated by the high R^2 values, indeed those which appears disperse have a very low

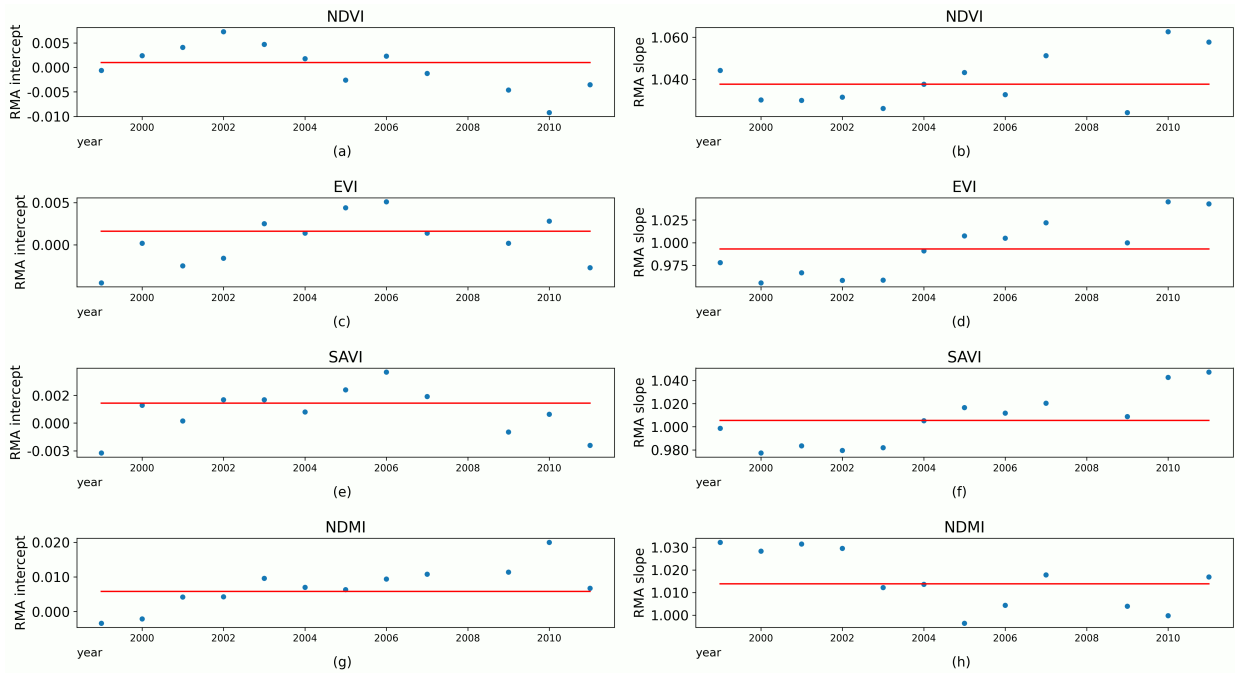


Fig. 13. RMA (a), (c), (e), (g) intercept and (b), (d), (f), (h) slope coefficients computed for each VI on yearly samples of L5 TM and L7 TM+ paired observations. The red lines represent the coefficients resulting from the RMA computed within the entire period [from Table III(d)].

probability density (lower than 0.1 on more than 10 million points).

In general, the NDMI performs similar to the other indices, despite the fact that the SWIR bands—which are the ones with the lowest overlap in the spectral response function [29]—are involved.

A direct comparison of the coefficients of the linear transformations proposed here with analogous values published in previous works is problematic, because in many cases, the estimation is based on limited samples [28], [33], or using different regression algorithms [32], or on a limited timespan [36]. On the one hand, the numerical values of the coefficients may vary depending on the choice of the area of interest (AOI), its land cover, and the considered time span; on the other hand, the observed RMSEs are comparable. The coefficient proposed here is averaged among 100 random extractions of a very large number of pixels, and the obtained standard deviations are relatively small, proving the robustness of the solution. For example, a standard deviation by 0.0003 for the slope coefficient of the NDVI in Table III(d) means a variation of 0.00015 on the corrected index, when the original value is 0.5. As a further test, it was verified that, when repeating the whole process on the area of intersection among all the four cross-sensor analyses (see the Supplementary Materials), the difference in the results is negligible. Thus, the proposed coefficients can be used and considered valid all over Europe, at regional as well as continental scale.

More sensible fluctuations emerge instead from the analysis performed year by year when comparing ETM+ with TM. Indeed, their entity is one order of magnitude higher than the standard deviations observed in the 100 extractions over the full period considered as a unique population. Observing the charts of Fig. 13, it seems not possible to ascertain a

common trend, nor a clear periodical oscillation. One possible justification for these fluctuations may be sought in the orbit drifting of L5. As pointed out by Roy et al. [24], L5 orbit was not maintained consistently over years, and a temporal pattern of increasing and then decreasing overpass times was observed, as the orbit was adjusted by periodic station keeping maneuvers. These resulted in changes in the illumination geometry at the moment of acquisition in a place. Anyway, further investigations are necessary to clarify this specific issue. All things considered, from the point of view of an end-user, it is recommended to use the average harmonization coefficients given in Table III.

As a final remark, using all the considered sensors together appears beneficial to two main scenarios: first, the creation of a very long TS from 1984 to present (for example, for climate change related studies) and, secondly, the generation of a dramatically denser TS henceforth (for near real-time monitoring applications). In both the cases, when assembling the TS including all these sensors, the estimated linear transformations can be used to produce a harmonized dataset. Due to its operational timespan, the authors recommend using L7 as the common reference and harmonizing all the other sensors with it.

VI. CONCLUSION

The goal of the study was to extract cross-sensors transformation coefficients for popular vegetation indexes computed from Landsat-5 TM, Landsat-7 ETM+, Landsat-8 OLI, and Sentinel-2 MSI. The aim was to enable long time series to increase the frequency of data because of different sensors harmonization. For each sensor pair, RMA and OLS linear regressions were computed on 300 000 pixels, randomly sampled from couples of almost simultaneous acquisitions

by different sensors, and the computations were repeated 100 times to check the repeatability. For the first time, a cross-comparison analysis on vegetation indexes (NDVI, SAVI, EVI, and NDMI) derived from Landsat Collection-2 and Sentinel SR (L2A) products acquired all over the European continent was performed. Furthermore, the study included data from L5, L7, L8, and S2 altogether, allowing the extraction of time series starting from 1984. This approach highly increases the acquisition frequency, combining the 16-day L8 repeat cycle with the five days of S2 and, thus, raising the chance to collect cloud-free images that enables effective vegetation monitoring. This study was able to compute coefficients that allow to create a consistent time series starting from 1984 to present, combining in a temporal sequence L5, L7, L8, and S2. The tests on validation datasets proved that the application of these coefficients reduces the average difference in VIs values between sensors by at least an order of magnitude.

Future studies will consider the integration of the recently launched Landsat-9 that can further increase data frequency. In addition, despite the intense effort by calibration and validation teams of both Landsat and Sentinel programs, only few studies, including in situ measurement for data QA, can be found in the literature about sensors harmonization. Collecting spectral signature measurements on the ground would help assessing data quality and comparing different products levels offered by these missions.

ACKNOWLEDGMENT

This research did not receive any specific grant from funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

- [1] J. Dash and B. O. Ogutu, "Recent advances in space-borne optical remote sensing systems for monitoring global terrestrial ecosystems," *Prog. Phys. Geography, Earth Environ.*, vol. 40, no. 2, pp. 322–351, Apr. 2016.
- [2] D. Phiri, M. Simwanda, S. Salekin, V. Nyirenda, Y. Murayama, and M. Ranagalage, "Sentinel-2 data for land cover/use mapping: A review," *Remote Sens.*, vol. 12, no. 14, p. 2291, Jul. 2020. [Online]. Available: <https://www.mdpi.com/2072-4292/12/14/2291>
- [3] A. Tsatsaris et al., "Geoinformation technologies in support of environmental hazards monitoring under climate change: An extensive review," *ISPRS Int. J. Geo-Inf.*, vol. 10, no. 2, p. 94, Feb. 2021.
- [4] H. West, N. Quinn, and M. Horswell, "Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities," *Remote Sens. Environ.*, vol. 232, Oct. 2019, Art. no. 111291, doi: [10.1016/j.rse.2019.111291](https://doi.org/10.1016/j.rse.2019.111291).
- [5] Z. Wu et al., "User needs for future Landsat missions," *Remote Sens. Environ.*, vol. 231, Sep. 2019, Art. no. 111214, doi: [10.1016/j.rse.2019.111214](https://doi.org/10.1016/j.rse.2019.111214).
- [6] M. A. Wulder et al., "Current status of Landsat program, science, and applications," *Remote Sens. Environ.*, vol. 225, pp. 127–147, May 2019.
- [7] C. E. Woodcock et al., "Free access to Landsat imagery," *Science*, vol. 320, no. 5879, p. 1011, May 2008, doi: [10.1126/science.320.5879.1011a](https://doi.org/10.1126/science.320.5879.1011a).
- [8] Z. Malenovský et al., "Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land," *Remote Sens. Environ.*, vol. 120, pp. 91–101, May 2012.
- [9] F. Casu, M. Manunta, P. S. Agram, and R. E. Crippen, "Big remotely sensed data: Tools, applications and experiences," *Remote Sens. Environ.*, vol. 202, pp. 1–2, Dec. 2017.
- [10] X. Yao et al., "Enabling the big earth observation data via cloud computing and DGGs: Opportunities and challenges," *Remote Sens.*, vol. 12, no. 1, p. 62, Dec. 2019. [Online]. Available: <https://www.mdpi.com/2072-4292/12/1/62>
- [11] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, "Google Earth engine: Planetary-scale geospatial analysis for everyone," *Remote Sens. Environ.*, vol. 202, pp. 18–27, Dec. 2017.
- [12] J.-F. Mas, "Monitoring land-cover changes: A comparison of change detection techniques," *Int. J. Remote Sens.*, vol. 20, no. 1, pp. 139–152, Jan. 1999. [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/014311699213659>
- [13] Z. Zhu, "Change detection using Landsat time series: A review of frequencies, preprocessing, algorithms, and applications," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 370–384, Aug. 2017, doi: [10.1016/j.isprsjprs.2017.06.013](https://doi.org/10.1016/j.isprsjprs.2017.06.013).
- [14] M. E. D. Chaves, M. C. A. Picoli, and I. D. Sanches, "Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review," *Remote Sens.*, vol. 12, no. 18, p. 3062, Sep. 2020.
- [15] C. Kuenzer, S. Dech, and W. Wagner, *Remote Sensing Time Series*, vol. 22. Berlin, Germany: Springer, 2015.
- [16] A. Banskota, N. Kayastha, M. J. Falkowski, M. A. Wulder, R. E. Froese, and J. C. White, "Forest monitoring using Landsat time series data: A review," *Can. J. Remote Sens.*, vol. 40, no. 5, pp. 362–384, Sep. 2014, doi: [10.1080/07038992.2014.987376](https://doi.org/10.1080/07038992.2014.987376).
- [17] Z. Xue, P. Du, and L. Feng, "Phenology-driven land cover classification and trend analysis based on long-term remote sensing image series," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1142–1156, Apr. 2014.
- [18] A. Bégué et al., "Remote sensing and cropping practices: A review," *Remote Sens.*, vol. 10, no. 2, p. 99, Jan. 2018.
- [19] C. Gómez, J. C. White, and M. A. Wulder, "Optical remotely sensed time series data for land cover classification: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 116, pp. 55–72, Jun. 2016.
- [20] J. Xue and B. Su, "Significant remote sensing vegetation indices: A review of developments and applications," *J. Sensors*, vol. 2017, pp. 1–17, Jan. 2017.
- [21] L. Zeng, B. D. Wardlow, D. Xiang, S. Hu, and D. Li, "A review of vegetation phenological metrics extraction using time-series, multispectral satellite data," *Remote Sens. Environ.*, vol. 237, Feb. 2020, Art. no. 111511, doi: [10.1016/j.rse.2019.111511](https://doi.org/10.1016/j.rse.2019.111511).
- [22] L. Liu et al., "Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth engine," *Remote Sens. Environ.*, vol. 239, Mar. 2020, Art. no. 111624.
- [23] A. Nagy, A. Szabó, O. D. Adeniyi, and J. Tamás, "Wheat yield forecasting for the Tisza river catchment using Landsat 8 NDVI and SAVI time series and reported crop statistics," *Agronomy*, vol. 11, no. 4, p. 652, Mar. 2021. [Online]. Available: <https://www.mdpi.com/2073-4395/11/4/652>
- [24] D. P. Roy, Z. Li, H. K. Zhang, and H. Huang, "A conterminous United States analysis of the impact of Landsat 5 orbit drift on the temporal consistency of Landsat 5 thematic mapper data," *Remote Sens. Environ.*, vol. 240, Apr. 2020, Art. no. 111701. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425720300705>
- [25] C. Wang, J. Li, Q. Liu, B. Zhong, S. Wu, and C. Xia, "Analysis of differences in phenology extracted from the enhanced vegetation index and the leaf area index," *Sensors*, vol. 17, no. 9, p. 1982, Aug. 2017. [Online]. Available: <http://www.mdpi.com/1424-8220/17/9/1982>
- [26] L. Zhong, L. Hu, and H. Zhou, "Deep learning based multi-temporal crop classification," *Remote Sens. Environ.*, vol. 221, pp. 430–443, Feb. 2019. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425718305418>
- [27] J. Li and B. Chen, "Global revisit interval analysis of Landsat-8 -9 and Sentinel-2A -2B data for terrestrial monitoring," *Sensors*, vol. 20, no. 22, p. 6631, Nov. 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/22/6631>
- [28] E. Mandanici and G. Bitelli, "Preliminary comparison of Sentinel-2 and Landsat 8 imagery for a combined use," *Remote Sens.*, vol. 8, no. 12, p. 1014, Dec. 2016. [Online]. Available: <http://www.mdpi.com/2072-4292/8/12/1014>
- [29] R. Chastain, I. Housman, J. Goldstein, M. Finco, and K. Tenneson, "Empirical cross sensor comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere spectral characteristics over the conterminous United States," *Remote Sens. Environ.*, vol. 221, pp. 274–285, Feb. 2019, doi: [10.1016/j.rse.2018.11.012](https://doi.org/10.1016/j.rse.2018.11.012). [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425718305212>
- [30] L. G. Denaro and C.-H. Lin, "Hybrid canonical correlation analysis and regression for radiometric normalization of cross-sensor satellite imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 976–986, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9016141/>

- [31] S. Xie, L. Sun, L. Liu, and X. Liu, "Global cross-sensor transformation functions for Landsat-8 and Sentinel-2 top of atmosphere and surface reflectance products within Google Earth engine," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–9, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9762921/>
- [32] H. Cao, L. Han, and L. Li, "Harmonizing surface reflectance between Landsat-7 ETM+, Landsat-8 OLI, and Sentinel-2 MSI over China," *Environ. Sci. Pollut. Res.*, vol. 29, no. 47, pp. 70882–70898, Oct. 2022. [Online]. Available: <https://link.springer.com/10.1007/s11356-022-20771-4>
- [33] P. Li, L. Jiang, and Z. Feng, "Cross-comparison of vegetation indices derived from Landsat-7 Enhanced Thematic Mapper Plus (ETM+) and Landsat-8 operational land imager (OLI) sensors," *Remote Sens.*, vol. 6, no. 1, pp. 310–329, Dec. 2013. [Online]. Available: <http://www.mdpi.com/2072-4292/6/1/310>
- [34] D. P. Roy et al., "Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity," *Remote Sens. Environ.*, vol. 185, pp. 57–70, Nov. 2016, doi: [10.1016/j.rse.2015.12.024](https://doi.org/10.1016/j.rse.2015.12.024).
- [35] F. Chen et al., "Normalized difference vegetation index continuity of the Landsat 4-5 MSS and TM: Investigations based on simulation," *Remote Sens.*, vol. 11, no. 14, p. 1681, Jul. 2019. [Online]. Available: <https://www.mdpi.com/2072-4292/11/14/1681>
- [36] G. Mancino, A. Ferrara, A. Padula, and A. Nolè, "Cross-comparison between Landsat 8 (OLI) and Landsat 7 (ETM+) derived vegetation indices in a Mediterranean environment," *Remote Sens.*, vol. 12, no. 2, p. 291, Jan. 2020. [Online]. Available: <https://www.mdpi.com/2072-4292/12/2/291>
- [37] M. Claverie et al., "The harmonized Landsat and Sentinel-2 surface reflectance data set," *Remote Sens. Environ.*, vol. 219, pp. 145–161, Dec. 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425718304139>
- [38] G. Chander, B. L. Markham, and J. A. Barsi, "Revised Landsat-5 thematic mapper radiometric calibration," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 490–494, Jul. 2007. [Online]. Available: <http://ieeexplore.ieee.org/document/4271443/>
- [39] S. Andrefouet et al., "Preliminary assessment of the value of Landsat 7 ETM+ data following scan line corrector malfunction," U.S. Geological Survey, EROS Data Center, Sioux Falls, SD, USA, Tech. Rep., Jul. 2003.
- [40] J. L. Dwyer, D. P. Roy, B. Sauer, C. B. Jenkinson, H. K. Zhang, and L. Lyburner, "Analysis ready data: Enabling analysis of the Landsat archive," *Remote Sens.*, vol. 10, no. 9, p. 1363, Aug. 2018.
- [41] US Geological Survey. (2021). *Landsat Collection 2*. Reston, VA, USA. [Online]. Available: <http://pubs.er.usgs.gov/publication/fs20213002>
- [42] US Geological Survey, *Landsat 4-7 Level 2 Science Product (L2SP) (Guide v. 4.0)*, US Geological Survey, Sioux Falls, South Dakota, Sep. 2021.
- [43] C. Teixeira Pinto, X. Jing, and L. Leigh, "Evaluation analysis of Landsat Level-1 and Level-2 data products using in situ measurements," *Remote Sens.*, vol. 12, no. 16, p. 2597, Aug. 2020. [Online]. Available: <https://www.mdpi.com/2072-4292/12/16/2597>
- [44] E. Vermote, C. Justice, M. Claverie, and B. Franch, "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product," *Remote Sens. Environ.*, vol. 185, pp. 46–56, Nov. 2016. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425716301572>
- [45] J. G. Masek et al., "A Landsat surface reflectance dataset for North America, 1990–2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 68–72, Jan. 2006.
- [46] ESA. (2015). *Sentinel-2 Processing Baseline*. [Online]. Available: <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/processing-baseline>
- [47] J. Louis et al., "Sentinel-2 global surface reflectance level-2A product generated with Sen2Cor," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2019, pp. 8522–8525. [Online]. Available: <https://ieeexplore.ieee.org/document/8898540/>
- [48] M. Main-Knorn, B. Pflug, J. Louis, V. Debaecker, U. Muller-Wilm, and F. Gascon, "Sen2Cor for Sentinel-2," *Proc. SPIE*, vol. 10427, pp. 37–48, Oct. 2017. [Online]. Available: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10427/2278218/Sen2Cor-for-Sentinel-2/10.1117/12.2278218.full>
- [49] S. Huang, L. Tang, J. P. Hupy, Y. Wang, and G. Shao, "A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing," *J. Forestry Res.*, vol. 32, no. 1, pp. 1–6, Feb. 2021, doi: [10.1007/s11676-020-01155-1](https://doi.org/10.1007/s11676-020-01155-1).
- [50] F. J. Krieglger, W. A. Malila, R. F. Nalepka, and W. Richardson, "Preprocessing transformations and their effects on multispectral recognition," in *Proc. 6th Int. Symp. Remote Sens. Environ.*, vol. 2. Ann Arbor, MI, USA: Infrared and Optics Laboratory, Willow Run Laboratories, Environmental Research Institute of Michigan, Oct. 1969, p. 97.
- [51] A. R. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, Aug. 1988. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/003442578890106X>
- [52] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, nos. 1–2, pp. 195–213, Nov. 2002. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425702000962>
- [53] S. Ferrant et al., "Detection of irrigated crops from Sentinel-1 and Sentinel-2 data to estimate seasonal groundwater use in south India," *Remote Sens.*, vol. 9, no. 11, p. 1119, Nov. 2017. [Online]. Available: <https://www.mdpi.com/2072-4292/9/11/1119>
- [54] B. C. Gao, "NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, Dec. 1996.
- [55] US Geological Survey. (2021). *Landsat Surface Reflectance-derived Spectral Indices*. [Online]. Available: <https://www.usgs.gov/Landsat-missions/Landsat-surface-reflectance-derived-spectral-indices>
- [56] M. A. Gilabert, J. González-Piqueras, F. J. Garcia-Haro, and J. Meliá, "A generalized soil-adjusted vegetation index," *Remote Sens. Environ.*, vol. 82, nos. 2–3, pp. 303–310, Oct. 2002. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0034425702000482>
- [57] M. Buchhorn, "Copernicus global land service: Land cover 100 m: Collection 3: Epoch 2018: Globe," Zenodo, Aug. 2020, doi: [10.5281/zenodo.3518038](https://doi.org/10.5281/zenodo.3518038).
- [58] L. Baetens, C. Desjardins, and O. Hagolle, "Validation of Copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and FMask processors using reference cloud masks generated with a supervised active learning procedure," *Remote Sens.*, vol. 11, no. 4, p. 433, Feb. 2019. [Online]. Available: <http://www.mdpi.com/2072-4292/11/4/433>
- [59] S. Foga et al., "Cloud detection algorithm comparison and validation for operational Landsat data products," *Remote Sens. Environ.*, vol. 194, pp. 379–390, Jun. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0034425717301293>
- [60] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0034425711003853>
- [61] J. Louis, A. Charantonis, B. Berthelot, and H. Lacoste-Francis, "Cloud detection for Sentinel-2," in *Proc. Living Planet Symp. Noordwijk*, The Netherlands: ESA Communications, 2010, p. 499. [Online]. Available: <https://www.tib.eu/de/suchen/id/BLCP%3ACN078005293>
- [62] W. V. Harper, "Reduced major axis regression," in *Wiley StatsRef: Statistics Reference Online*. Hoboken, NJ, USA: Wiley, Nov. 2016, pp. 1–6. [Online]. Available: <https://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat07912>
- [63] D. Moravec, J. Komárek, S. López-Cuerdo Medina, and I. Molina, "Effect of atmospheric corrections on NDVI: Intercomparability of Landsat 8, Sentinel-2, and UAV sensors," *Remote Sens.*, vol. 13, no. 18, p. 3550, Sep. 2021.
- [64] C. Gouveia, R. M. Trigo, and C. C. DaCamara, "Drought and vegetation stress monitoring in Portugal using satellite data," *Natural Hazards Earth Syst. Sci.*, vol. 9, no. 1, pp. 185–195, Feb. 2009.
- [65] S. Madonsela, M. A. Cho, A. Ramoelo, and O. Mutanga, "Remote sensing of species diversity using Landsat 8 spectral variables," *ISPRS J. Photogramm. Remote Sens.*, vol. 133, pp. 116–127, Nov. 2017, doi: [10.1016/j.isprsjprs.2017.10.008](https://doi.org/10.1016/j.isprsjprs.2017.10.008).
- [66] J. Qi, A. Chehbouni, A. R. Huete, Y. H. Kerr, and S. Sorooshian, "A modified soil adjusted vegetation index," *Remote Sens. Environ.*, vol. 48, no. 2, pp. 119–126, May 1994. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0034425794901341>
- [67] J. R. Irons, J. L. Dwyer, and J. A. Barsi, "The next Landsat satellite: The Landsat data continuity mission," *Remote Sens. Environ.*, vol. 122, pp. 11–21, Jul. 2012, doi: [10.1016/j.rse.2011.08.026](https://doi.org/10.1016/j.rse.2011.08.026).
- [68] T. Miura, A. R. Huete, and H. Yoshioka, "Evaluation of sensor calibration uncertainties on vegetation indices for MODIS," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1399–1409, May 2000.



Francesca Trevisiol received the master's degree in environmental engineering from the University of Bologna, Bologna, Italy, in 2018, and the Ph.D. degree from the Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, in 2023. Her Ph.D. thesis was on Earth observation.

She is currently a Post-Doctoral Researcher of geomatics with the University of Bologna.



Pietro Mattivi received the master's degree in Earth resources engineering from the University of Bologna, Bologna, Italy, in 2018.

He is a Research Fellow with the Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna. His research interests include remote sensing, DEM production and analysis, and UAV surveying for precision agriculture.

Emanuele Mandanici is an Associate Professor of geomatics with the Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Bologna, Italy. His research activity concerns the development of remote sensing methods for the acquisition and processing of geospatial information, to address problems of environmental monitoring and sustainability, energy efficiency in urban environments, study, and preservation of cultural heritage sites.



Gabriele Bitelli (Member, IEEE) is a Full Professor of geomatics with the Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Bologna, Italy. The prevailing research interests are the integration of different systems and sensors, surveying and representation for cultural heritage, Earth observation and satellite remote sensing for environmental monitoring (e.g., subsidence, risk management, and urban studies), 3-D technologies at different scales in multidisciplinary contexts, and GIS. He is engaged

in several national and international research projects; he collaborates with various public and private institutions.