

© 2018. This manuscript version is made available under the CC
BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Article

Assessment of atmospheric correction methods for Sentinel-2 images in Mediterranean landscapes

Ion Sola ^{1*}, Alberto García-Martín^{2,3}, Leire Sandonis-Pozo ⁴, Jesús Álvarez-Mozos ¹, Fernando Pérez-Cabello ^{3,4} María González-Audicana ¹ and Raquel Montorio Llovería ^{3,4}

¹ Universidad Pública de Navarra. Departamento de Proyectos e Ingeniería Rural, c/Campus Arrosadía s/n 31006 Pamplona.

² Centro Universitario de la Defensa de Zaragoza, Academia General Militar. Ctra. Huesca s/n, 50090 Zaragoza, Spain; algarcia@unizar.es

³ GEOFOREST research group, Environmental Sciences Institute (IUCA), University of Zaragoza, C/ Pedro Cerbuna 12, 50009 Zaragoza, Spain.

⁴ Department of Geography and Spatial Management, University of Zaragoza, C/ Pedro Cerbuna 12, 50009 Zaragoza, Spain; l.sandonisp@gmail.com; fcabello@unizar.es; montorio@unizar.es

* Correspondence: ion.sola@unavarra.es; Tel.: 948168401

Academic Editor: name

Received: date; Accepted: date; Published: date

Abstract: Atmospheric correction of optical satellite imagery is an essential pre-processing for modelling biophysical variables, multi-temporal analysis, and digital classification processes. Sentinel-2 products available for users are distributed by the European Space Agency (ESA) as Top Of Atmosphere reflectance values in cartographic geometry (Level-1C product). In order to obtain Bottom Of Atmosphere reflectance images (Level-2A product) derived from this Level-1C products, ESA provides the SEN2COR module, which is implemented in the Sentinel Application Platform. Alternatively, ESA recently distributes Level-2A products processed by SEN2COR with a default configuration. On the other hand, the conversion from Level-1C to Level-2A product can be generated using alternative atmospheric correction methods, such as MAJA, 6S, or iCOR. In this context, this paper aims to evaluate the quality of Level-2A products obtained through different methods in Mediterranean shrub and grasslands by comparing data obtained from Sentinel-2 imagery with field spectrometry data. For that purpose, six plots with different land covers (asphalt, grass, shrub, pasture, and bare soil) were analyzed, by using

synchronous imagery to fieldwork (from July to September 2016). The results suggest the suitability of the applied atmospheric corrections, with coefficients of determination higher than 0.90 and root mean square error lower than 0.04 achieving a relative error in bottom of atmosphere reflectance of only 2-3%. Nevertheless, minor differences were observed between the four tested methods, with slightly varying results depending on the spectral band and land cover.

Keywords: Sentinel 2A-MSI; Atmospheric correction; SEN2COR; MAJA; 6S; iCOR; Field spectrometry; Mediterranean shrub and grasslands

1. Introduction

The pair of Sentinel-2 (S2) satellites provides improved continuity for Spot- and Landsat-type observations due to their complete interoperability with these programs (Drusch et al., 2012), increasing the quantity of available images to use in several applications where medium-high resolution optical images have been shown useful. In addition, S2 presents some advantages over Landsat and SPOT space programs (Astrium, 2013; Gascon et al., 2017; Zanter, 2015): (i) a higher spectral resolution in the optical spectrum than Landsat 8 (L8) and SPOT 5; (ii) a greater spatial resolution along the entire optical spectrum, compared to L8; (iii) a more extensive swath (around 280 km² compared to the 185 km² of Landsat and the 60 km² of SPOT 5) and; (iv) a higher temporal resolution than L8 (Sentinel-2A and Sentinel-2B together will provide a global median average revisit interval of 5 days overtaking significantly the 16 days of Landsat).

In the short period of time since Sentinel-2A, and more recently Sentinel-2B, became fully operative, many studies have been performed testing the application of these observations to land cover classification (Borràs et al., 2017; Shoko and Mutanga, 2017), crop mapping (Immitzer et al., 2016; Lebourgeois et al., 2017), biophysical parameter estimation (Clevers et al., 2017; Chrysafis et al., 2017; Sibanda et al., 2016; Vuolo et al., 2016), forest fires (Navarro et al., 2017; Quintano et al., 2016; Verhegghen et al., 2016), water resources (Dörnhöfer et al., 2016; Toming et al., 2016;

Traganos and Reinartz, 2017) or urban areas characterization (Lefebvre et al., 2016; Pesaresi et al., 2016; Yang and Chen, 2017). In addition, there is also an increasing amount of scientific literature combining or comparing S2 with L8 images (Colkesen and Kavzoglu, 2017; Korhonen et al., 2017; Mallinis et al., 2017; Munyati, 2017; Novelli et al., 2016; Paul et al., 2016; van der Meer et al., 2014).

Since 2015, the European Space Agency (ESA) provides Level-1C products derived from Sentinel-2A and from 29 June 2017 also for Sentinel-2B. Level-1C data corresponds to Top-Of-Atmosphere (TOA) reflectance values after the application of radiometric and geometric corrections (including orthorectification and spatial registration). Besides, the Copernicus Open Access Hub provides Level-2A products of S2 imagery data over Europe (EEA-39) from 28 March 2017. Level-2A corresponds to Bottom-Of-Atmosphere reflectances (BOA). The production of S2 Level-2A products is done using SEN2COR processor (version 2.3.1) (ESA, 2015; Gascon et al., 2017; Louis et al., 2016). This processor developed by ESA performs the atmospheric correction (AC), terrain and cirrus correction and scene classification applied to the mentioned TOA Level-1C input data. For this production the PlanetDEM Digital Elevation Model (DEM) is used, while cirrus and Bidirectional Reflectance Distribution Function (BRDF) corrections are deactivated.

ESA also distributes SEN2COR processor freely, as a plugin integrated in the Sentinel Application Platform software (SNAP). Besides of the BOA reflectance product, additional outputs such as Aerosol Optical Thickness (AOT) map, Water Vapor (WV) map, Scene Classification map and Quality Indicators data are generated. The atmospheric model of SEN2COR is based on the ATCOR model, which inverts the parameters previously mentioned through the use of transformation tables (LUT's, *Look Up Table*), obtained from Radiative Transfer Models (RTM) calculated over the type of sensor, solar geometry, topography of the terrain and atmospheric parameters. Further details on SEN2COR processor and its calibration can be obtained in (Müller-Wilm, 2016; Richter et al., 2017).

The accuracy of the AC applied to an optical image is a very relevant factor because it essentially affects the accuracy of: (i) multi-temporal studies; (ii) the inter-comparison of

measurements from different optical sensors; (iii) the improvement of digital classification processes; and (iv) the development of physical and empirical models to estimate biophysical variables (e.g.: Leaf Area Index, LAI; Fraction of Absorbed Photosynthetically Active Radiation by the green elements of the canopy, FAPAR; Canopy water content, CWC; Above Ground Biomass, AGB) (Chuvienco, 2010; Gascon et al., 2017; Lillesand et al., 2014; Marcello et al., 2016; Pons et al., 2014; Vuolo et al., 2016).

In particular, the performance of AC relies on an accurate estimation of the concentration and type of atmospheric particles at the acquisition time (Vermote and Kotchenova, 2008). The aerosol contribution still generates high uncertainties in the correction of atmospheric effects though (Vermote et al., 1997). In the last years, research institutes, aerospace agencies and private companies have developed AC methods for high spatial resolution (HSR) multispectral images (Hagolle et al., 2015a; Lantzanakis et al., 2017; Martins et al., 2017; Nazeer et al., 2014). Some of these methods have been successfully applied, and at present the selection of the best AC method to be applied to S2 Level-1C products is a hot topic under discussion, as demonstrated by the Atmospheric Correction Inter-comparison Exercise (ACIX), an international collaborative initiative to inter-compare a set of AC methods for HSR optical sensors, specially focused on L8 and S2 imagery (Doxani et al., 2018).

The accuracy of the S2 Level-2A products obtained through different AC processors must be validated. In the case of SEN2COR, the Level-2A product was calibrated and validated by (Gascon et al., 2017). In this work, the validation of the BOA reflectance product resulting from SEN2COR was based on 24 Aerosol Robotic Network (AERONET) test sites distributed worldwide covering different land cover types and atmospheric conditions. In this case, the validation focused on an area of 9 km² around each AERONET sun photometer location. The results obtained for different land uses (bare soils, forests and other vegetation covers) were considered consistent and of high-quality, with a divergence on BOA reflectance values of only a 5% (Gascon et al., 2017). It is accepted that field spectroscopy data is the ideal method to ensure

an accurate validation of AC methods (McCoy, 2005; Milton et al., 2009), and several studies relied on such measurements (Dörnhöfer et al., 2016; Vuolo et al., 2016). In (Vuolo et al., 2016), two field campaigns with a 350–2500nm range spectroradiometer were performed over an agricultural area in Austria with the objective of validating SEN2COR processor. Although the results obtained were deemed satisfactory, a quantitative measure of the agreement between the two reflectance data sets was not shown in the paper. However, recent studies also reported issues on the performance of these products. In Kukawska et al. (2017) a problem was reported related to the incorrect mask of clouds over artificial structures resulting from SEN2COR processor. According to the authors these errors had a direct influence on the overall accuracy of the classifications performed.

Recently, new inter-comparison studies of different AC methods applied to S2 imagery have been published. In Dörnhöfer et al. (2016) the results of three different AC processors were compared with in situ reflectance measurements in an Oligotrophic Lake. The comparison between the in situ reflectance and the reflectance obtained from SEN2COR yielded a root mean square error (RMSE) slightly worse than the other two AC methods. In Lantzanakis et al. (2017) the effectiveness of four AC methods on S2 imagery was compared, three of them physically-based (Second Simulation of a Satellite Signal in the Solar Spectrum (6S), FLAASH and SEN2COR) and one image-based (DOS). The authors concluded that the physically based methods performed better than the image-based one for S2 imagery, but the evaluation was only based on the comparison of BOA reflectances with spectral libraries (Baldrige et al., 2009). Generally, an image-based AC method does not require external data, although the use of the ancillary data required is not currently a drawback due to its increasing availability, easily accessible in an automatic way. In Martins et al. (2017) the assessment of three AC methods, 6SV (6S vector version), ACOLITE and SEN2COR, applied to a S2 image of Amazon Floodplain Lakes, with varying results depending on lake characteristics.

At present, one of the most popular algorithms to perform AC on S2 imagery is MAJA (Lonjou et al., 2016), the result of the joint effort of the French Centre National d'Études Spatiales (CNES) and the German Aerospace Center (DLR) to merge their algorithms Multi-Mission Atmospheric Correction and Cloud Screening (MACCS) and ATCOR (Richter and Schläpfer, 2015), respectively. MACCS is based on a multi-temporal algorithm that makes an optimized use of image time series to characterize the atmosphere and detect clouds. In (Hagolle et al., 2015b) the AOT retrieval, a key factor in AC, was validated by comparing its value at $0.55\ \mu\text{m}$ to AERONET in situ data. Moreover, BOA reflectances were also validated against in situ measurements (Rouquié et al., 2017) and MODIS BOA reflectance products. Finally, new alternatives to SEN2COR appeared in the last months, such as iCOR (VITO, 2017), implemented as a SNAP plugin to correct S2 and L8 satellite imagery. This processor was previously known as OPERA, and has been tested and validated extensively for MERIS, L8, S2 and PROBA-V in different research projects (ACIX, HIGHROC, INFORM and SPONGE) (Sterckx et al., 2015a). The performance of most of these mentioned processors (MAJA, iCOR, SEN2COR, etc.) have been recently compared in ACIX (Doxani et al., 2018). The comparison was based on BOA reflectances and AOT and WV values and the results varied depending on the sensors, products, and sites.

In this context, the aim of this study is to assess the quality of the Level-2A products obtained after the application of four AC processors (SEN2COR, MAJA, 6S and iCOR) using field spectrometry, concretely an Analytical Spectral Device (ASD) FieldSpec® 4. In particular, this work focuses on the accuracy assessment of these products in Mediterranean shrub and grasslands. For that purpose, six permanent plots representative of Mediterranean shrub and grasslands environments and anthropic covers were placed and ground reflectance measurements were acquired simultaneously to four Sentinel-2A overpasses taking into account the requirements to obtain high-quality data. Although, it will be interesting to have in the study area surfaces with reflectances bellow 0.1 or very bright objects in order to assessing better the performance of different atmospheric correction methods, they does not exist in the study area.

Additionally to the S2 spectral bands in the optical spectrum, the accuracy of the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Infrared Index (NDII) and the Normalized Burn Ratio (NBR) derived from them was also evaluated. The performance of cloud detection was not considered in this study due to the lack of ground truth data to compare with.

2. Materials and Methods

The methodology was divided into two phases (Fig. 1): (i) The retrieval of input data, consisting of the acquisition of field spectra data and the obtainment of Sentinel-2A data with four different AC and the subsequent calculation of selected spectral indices; and (ii) The validation consisted of a comparative statistical analysis of data obtained from both sources.

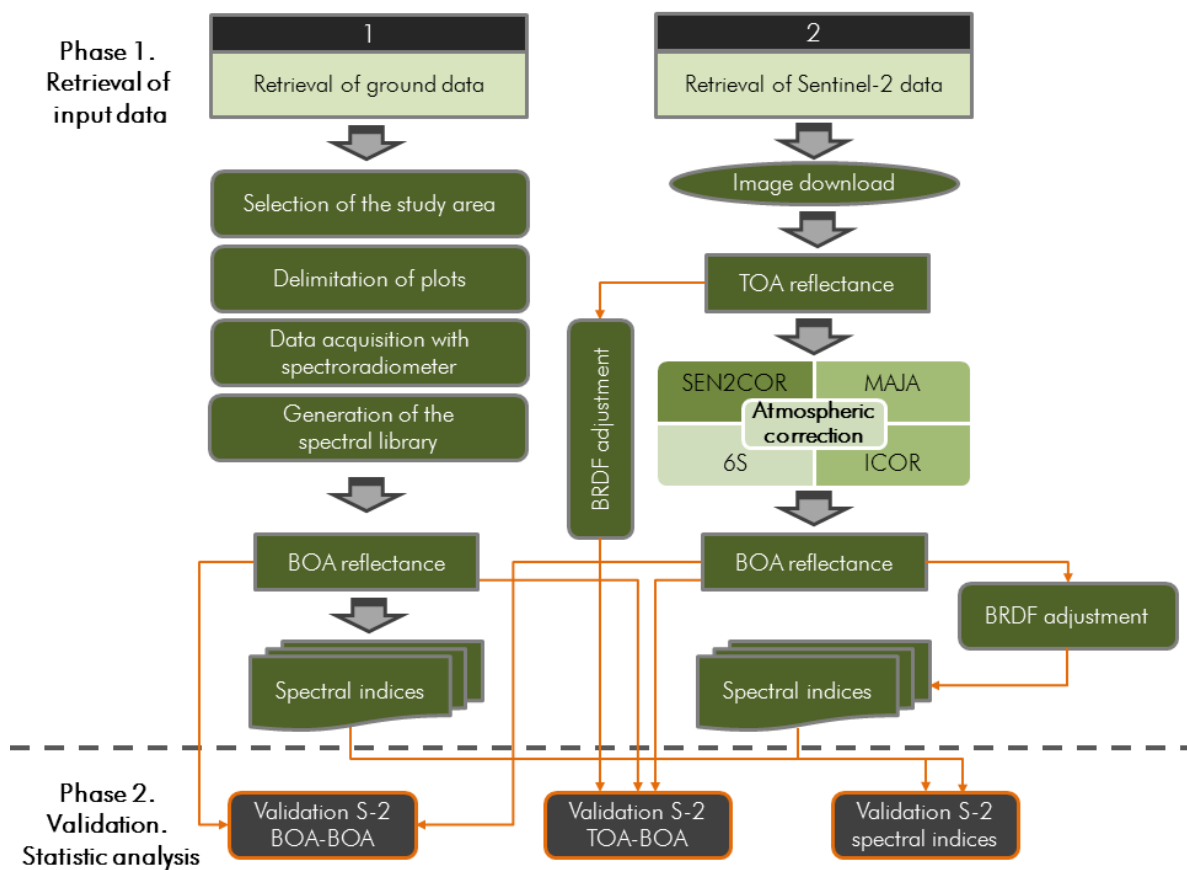


Figure 1. Methodological workflow of the study.

2.1. Field spectra data

The selected area to carry out the present study is located in the central sector of the Ebro Basin (Zaragoza, Spain). This area was considered optimal for the purpose of the study due to the presence of stable, homogeneous and extensive typical Mediterranean vegetation, as well as anthropic covers. The plots delimited meet the criteria established by the Committee on Earth Observation Satellites (CEOS) Working group on Calibration and Validation of being placed in sites with low probability of atmospheric variability and having high spatial homogeneity and weak directional effects (Patel et al., 2016), Table 1 contains the description of the six plots delimited, whereas Figure 2 shows their location.

Table 1. Description of the spectral sampling plots.

Plot name	Description
Asphalt	Asphalt area
Grass	Football field
Shrub 1	Main species: <i>Rosmarinus officinalis</i> L.
Shrub 2	Main species: <i>Rosmarinus officinalis</i> L.
Pasture	Main species: <i>Brachypodium retusum</i> Pers.
Bare soil	Area without vegetation

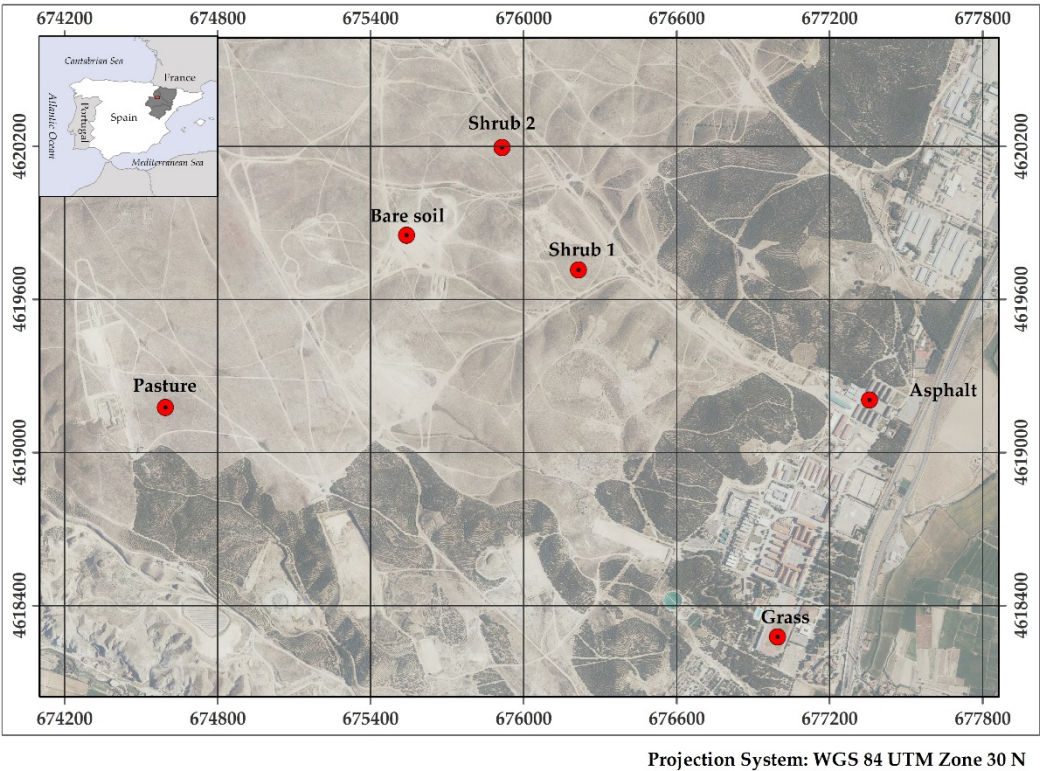


Figure 2. Location of the spectra sampling plots.

In order to ensure the agreement between the reflectance data obtained by true fieldwork and the S2-derived BOA reflectance values, plots were located according to the following criteria: (i) they were placed on a field path that would allow field spectrometry data acquisition within 2 hours of local solar noon (Milton et al., 2009); (ii) data collection was almost simultaneous to satellite overpasses over the study area (R051 and R094 orbit, between 10:50-11:10 solar time) and in no case exceeded a difference of 2 hours, and (iii) each plot was set at the centroid of the central pixel of a 3 x 3 pixel kernel corresponding to the 20 m grid defined in Sentinel-2A scenes. The 20 m grid of Sentinel-2A scenes was selected because when the field work campaigns were designed, the output of the plugin of Sen2Cor available to SNAP was an image of 10 m or 20 m or 60 m spatial resolution. The image of 20 m was chosen, as the majority of the bands MSI have this spatial resolution.

The plots were geo-positioned in field using a Leica VIVA GS15 CS10 real-time kinematic Global Navigation Satellite System, achieving an average planimetric accuracy of 0.50 m. Figure 3 shows the plots staked out in the field. Once the plots were geolocated, the next step was to define two opposing diagonals (NW-SE and NE-SO) with 11 meters length with the aid of a compass and a Vertex (Haglöf Sweden®). These transects marked the itinerary to perform field spectrometry.

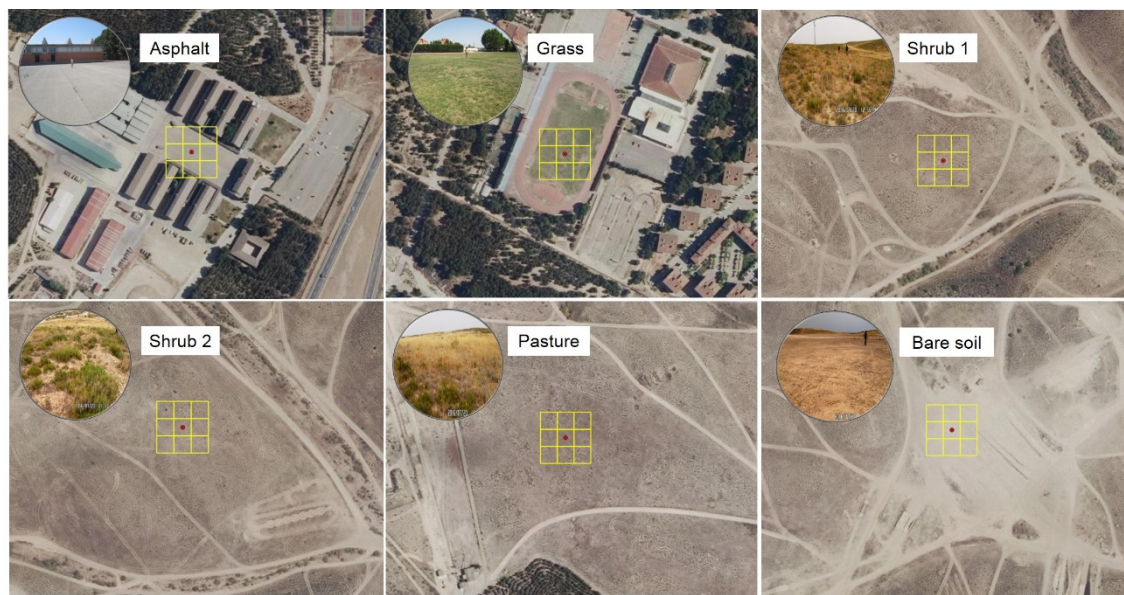


Figure 3. Plots staked out in field.

Measurements of the hemispherical conical reflectance factor (HCRF) (Schaeppman-Strub et al., 2006) were captured on clear-sky days using directly an Analytical Spectral Device (ASD) FieldSpec® 4. This instrument has been previously used in vicarious calibration, such as Abdou et al. (2001), Patel et al., (2016); calibration and validation of hyperspectral indices (Le Maire et al., 2008); and for understanding the optical responses of vegetation in field (Pacheco-Labrador et al., 2016). This device is being used in the vicarious calibration campaigns of the hyperspectral sensor EnMAP (Danner, et al., 2015). ASD designed to record the signal throughout the region between the visible (VIS) and near infrared (NIR) (350-1000 nm) and two regions corresponding to the short wave infrared (SWIR) (1001-1800 and 1801-2500 nm) with a sampling interval of 1.4 nm – 2 nm, and a spectral resolution Full-Width-Half-Maximum (FWHM) of 3 nm and 10 nm, respectively (ASD, 2015). Following the procedures established in the Spanish National research projects Fluxpec (http://www.lineas.cchs.csic.es/fluxpec/project_overview) and SynerTGE (http://www.lineas.cchs.csic.es/synertge/project_overview) described in Mendiguren et al., 2015; Pacheco-Labrador et al., 2016; Melendo-Vega et al., 2017, before measuring along each transect,

dark current was recorded, instrument settings were optimized and reflectance was calibrated using a using a Spectralon® 99% reflective reference panel (Labsphere Inc., North Sutton, NH, USA) 1 to ensure the perfect correspondence of illumination conditions. With a 1.5 m height and a 25° FOV (Field Of View) (rendering a sensor footprint diameter of about 66 cm), on-the-go measurements of each transect were obtained by averaging 25 individual reflectance spectra for each spectral measurement. Approximately 25 spectral measurements was registered for each transect, resulting 50 per plot. These measurements were used to calculate de average of of each plot every field work date. Table 2 shows the dates of the field work, the number of spectra collected each day and the corresponding Sentinel-2A orbit overpasses. The granules corresponding to the four dates were cloud-free for the study area.

Table 2. Field work dates,-number of spectra collected and Sentinel-2A orbit overpass.

Date	No. of spectra	Orbit
2016-07-23	300	R094
2016-08-22	315	R094
2016-09-01	302	R094
2016-09-28	304	R051

2.2. Calculation of Sentinel-2A bands and spectral indices

A spectral library was generated in ENVI using the field spectra obtained in each sampling day and plot. Every signature was resampled to the spectral response function (SRF) of Sentinel-2A. In this process, signatures following the approach presented in Veraverbeke et al. (2014), the MSI-Sentinel-2A specific noise was added to the simulated MSI using the MSI signal-to-noise ratios (SNR) in 20 m spatial resolution bands (<https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric>) as input to Eq. (1). The objective of this procedure was to ensure that the field spectra were more faithful to the spectral information registered by

MSI sensor. Final signatures were expressed in reflectance given as a fraction of unity. Recently, a new version of the Sentinel-2A SRF has been released by ESA, mainly changing the responses for bands B01 and B02.

$$r_{n,i} = r_i \times \left(1 + \frac{N(0,1)}{SNR_i}\right) \quad (1)$$

where r_i is the reflectance for band i and $N(0,1)$ is a random number generated from a normal distribution with a mean of zero and standard deviation of one.

Finally, the indices NDVI, NDII and NBR were calculated using the Sentinel-2A simulated signatures. These indices were selected due to their popularity in a plethora of applications (i.e., monitoring of vegetation dynamics (Anyamba and Tucker, 2005), estimation of biophysical parameters such as LAI or FCOVER (Carlson and Ripley, 1997) or land degradation assessment (Yengoh et al., 2015) in the case of NDVI, vegetation water content estimation with NDII (Davidson et al., 2006; Yilmaz et al., 2008) or burn severity estimation with NBR (Fraser et al., 2017), their ease of implementation and the fact that they consider spectral bands located on VIS, NIR and SWIR regions.

2.3. Obtaining S2 Level-2A product and calculation of spectral indices

The Level-1C products were acquired simultaneously to field work days (Table 2). Only the T30TXM granule corresponding to the study area was used (100 km²). Subsequently, Level-2A products were generated through four different AC methods at the native resolution of each spectral band. As a result, nine BOA reflectances were obtained for each date and method: three in the visible spectrum (B2, B3 and B4), four in the red edge and NIR (B5, B6, B7 and B8A) and two in the SWIR (B11 and B12).

2.3.1. SEN2COR

SEN2COR (2.4.0 version used, released on August 31 of 2017) is a processor for generating S2 Level-2A products from TOA reflectance Level-1C input data. In short, the following steps are

run through: i) Cloud Detection and Scene Classification, ii) Retrieval of AOT and WV content from the L1C image, and iii) TOA to BOA reflectance conversion. SEN2COR also includes optional cirrus correction, terrain correction and empirical BRDF-corrections. This processor requires a DEM and the LUTs as ancillary data. These LUTs are compiled using an atmospheric RTM based on libRadtran. The LUTs are generated for a wide variety of atmospheric conditions, solar geometries, and ground elevations and are calculated with a spectral resolution of 0.6 nm. This database is subsequently resampled with the S2 spectral responses, in order to obtain the sensor-specific functions needed for the AC. The aerosol optical thickness retrieval is based on the dense dark vegetation (DDV) algorithm. The water vapour retrieval over land is performed with the atmospheric pre-corrected differential absorption (APDA) algorithm. Besides of BOA reflectances, some additional output products are generated, in particular AOT, WV, Scene Classification Maps and Quality Indicators for cloud and snow probabilities. Its output product format is equivalent to the native resolutions of Level-1C imagery, that is, 60, 20 and 10 m. In this work the default parameters were used, neglecting topographic correction, cirrus correction and BRDF correction in order to provide a coherent comparison of AC methods, as the other methods do not include these corrections.

2.3.2. MAJA

The MAJA processor (MACCS ATCOR Joint Algorithm) (Lonjou et al., 2016) is a processor for cloud detection and AC, specifically designed to process time series of optical images at high resolution, acquired under quasi constant viewing angles, such as L8 or S2 images. Since 2016, it is progressively including methods taken from DLR's ATCOR processor. It is now the object of collaboration between CNES, DLR and Centre d'Etudes Spatiales de la BIOsphère (CESBIO). In the first step, MAJA corrects for absorption by atmospheric gas molecules, using the absorption part of the Simplified Model for Atmospheric Correction (SMAC) method and considering ozone, oxygen and water vapor concentrations obtained from either satellite data (ozone, water vapor)

or meteorological data (pressure). A second processing step detects the clouds and their shadows, using the multi-temporal cloud detection (MTCD) method. The third step is the AOT estimate, derived from the image, using a combination of three methods: the black pixel method, the dark dense vegetation method (DDV) and the multi-temporal (MT) method. The successive orders of scattering code (SOS) provides LUTs to convert the TOA reflectances already corrected for gas absorption into BOA reflectances. These BOA reflectances are provided with and without topographic correction, at the native spatial resolution of Level-1C products. In this work the latter was selected, with the objective of avoiding confronting factors to the AC. Additionally, MAJA generates a collection of masks including a cloud mask and a geophysical mask.

2.3.3. iCOR

Implemented as a plugin in SNAP, iCOR (previously known as OPERA) is an AC of S2 and L8 data based on the identification of water and land pixels (VITO, 2017). All input data required by the processor are derived from the image itself or delivered through pre-calculated LUTs. The AC comprises the following steps: i) land and water pixels are identified; ii) AOT is derived from land pixels based on an adapted version of the method developed by Guanter et al. (2007); iii) an adjacency correction is performed using SIMEC (Sterckx et al., 2015b) over water and fixed background ranges over land targets, and iv) the radiative transfer equation is solved. iCOR uses MODTRAN 5 (Berk et al., 2006) LUTs to perform the AC and needs information about the solar and viewing angles and a DEM. Besides of BOA reflectances no additional product is generated. The correction with default parameters has been tested in this work.

2.3.4. 6S

6S (Second Simulation of the Satellite Signal in the Solar Spectrum) (Vermote et al., 1997) is an open-source AC tool, established itself as one of the standard RTMs used for both remote sensing research and the creation of operational products. In this work AC of S2 imagery was performed with Google Earth Engine (GEE) using Py6S (Murphy, 2017; Wilson, 2013), an

interface to the 6S through the Python programming language. The 6S generates interpolated LUTs under different atmospheric conditions, considering altitude of the sensor and target, wavelength and ground reflectance (with the ability to use a number of built-in BRDF models). Essentially, these LUTs are used to calculate AC coefficients which convert TOA radiance to BOA reflectance. In contrast to the previous AC methods, the atmospheric parameters (WV, ozone and AOT) were not obtained from the image itself, but from GEE collections for the place, date and time of acquisition (see Table 3):

Table 3. Sources of atmospheric parameters obtained in 6S algorithm with GEE.

Parameter	GEE Collection	Description
Water vapor	NCEP_RE/surface_wv	Total column water vapor, kg/m ² at 6-hour temporal resolution (00:00, 06:00, 12:00, and 18:00 UTC) and 2.5 degree spatial resolution.
Ozone	TOMS/MERGED	Total column ozone (Dobson units). The Total Ozone Mapping Spectrometer (TOMS) data represent a merged ozone product from TOMS/EarthProbe, TOMS/Nimbus-7, TOMS/Meteor-3, OMI/Aura and USGS-interpolated data for dates with no data. The resolution ranges from 1.0 × 1.00 deg to 1.0 × 1.25 deg
AOT	MODIS/MOD08_M3_051	MOD08_M3 is a level-3 MODIS gridded atmosphere monthly global product. It contains monthly 1 × 1 degree grid average values of atmospheric parameters.

Level-1C products were downloaded from ESA Open Hub. SEN2COR and ICOR processors used this products as source to generate Level-2A products. Similarly, 6S method was applied to the S2 Level-1C image collection in Google Earth Engine (ID: COPERNICUS/S2). Finally, MAJA level-2A images were downloaded from Theia Land Data Centre.

2.3.5. BRDF adjustment

The four selected Sentinel 2 images correspond to adjoining paths (orbits R094 and R051). These observations were acquired with directional effects due to surface reflectance anisotropy and changes in the solar and viewing geometry. Therefore, it is required to adjust all images to a uniform nadir view. For that purpose a bidirectional reflectance distribution function (BRDF) correction was applied to TOA and BOA reflectances in order to provide a fair comparison with

nadiral ground measurements. The correction was applied following (Roy et al., 2017a; Roy et al., 2017b; Roy et al., 2016) using Google Earth Engine.

2.3.6. Calculation of spectral indices.

The BOA reflectances of Level-2A products were used as source to calculate three spectral indices: NDVI, one of the most popular index to identify vegetated areas and their condition; NDII, frequently used to estimate soil moisture storage; and NBR, originally designed to identify burned areas and to assess the severity of a burn. In order to calculate these indices from S2 data the following formulation was used:

$$NDVI = \frac{(\rho_{B8A} - \rho_{B4})}{(\rho_{B8A} + \rho_{B4})} \quad (2)$$

$$NDII = \frac{(\rho_{B8A} - \rho_{B11})}{(\rho_{B8A} + \rho_{B11})} \quad (3)$$

$$NBR = \frac{(\rho_{B8A} - \rho_{B12})}{(\rho_{B8A} + \rho_{B12})} \quad (4)$$

where, ρ_{B4} , ρ_{B8A} , ρ_{B11} and ρ_{B12} are BOA reflectances of band B4 (red), band B8A (NIR) and bands B11 and B12 (SWIR) respectively.

In the case of NDII and NBR the spatial resolution of both source bands was 20 meters, thus the indices were calculated at this resolution. In contrast, NDVI was calculated at a spatial resolution of 10 meters. Consequently a Nearest Neighbor resampling was applied to downscale band B8A from 20 to 10 meters.

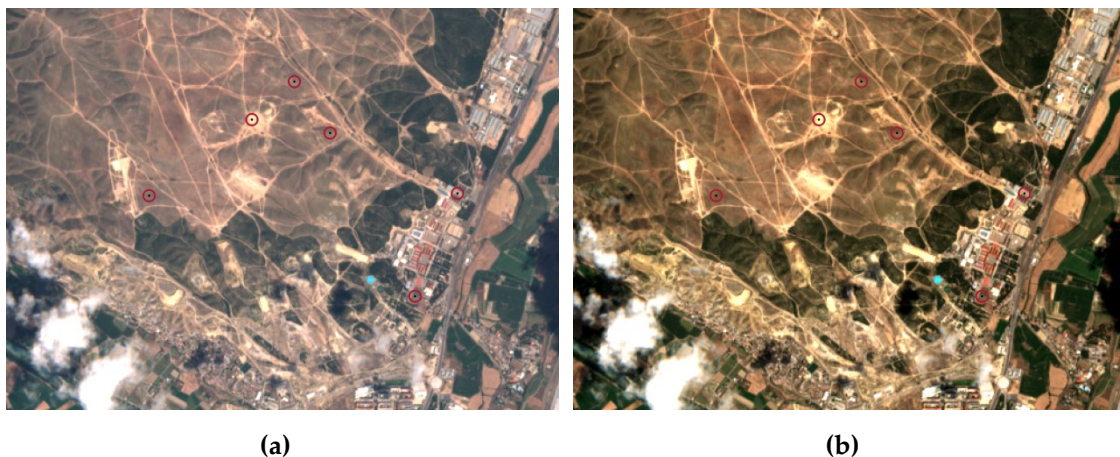
2.4. Statistical analysis for the validation of Level-2A products

The performance of AC methods producing BOA reflectances was evaluated by comparing these reflectances with the ground measurements obtained by the spectroradiometer. This analysis was performed per land cover and band wise in order to obtain more details on the performance of each method. The band 8 is not present in this analysis, as when the first version of the research was performed, this band was not an output at the 20 m spatial resolution obtained

from the plugin of Sen2Cor available to SNAP. The evaluation was based on statistical analyses of both datasets, i.e., ground measurement and Sentinel2-derived reflectances. The results of uncorrected Level-1C TOA reflectances were also provided as reference. In particular, the coefficient of determination (R^2), and the RMSE were analyzed, besides of spectral signatures of each land cover, and scatterplots of each spectral band. Finally, the three spectral indices mentioned in section 2.3.6. were compared to the results obtained from ground measurements.

3. Results and discussion

The visual assessment of AC shows a clear improvement in visible bands due to the subtraction of atmospheric scattering from the spectral bands with lower wavelength, especially blue band (B2). The RGB composition in true color (blue: B2, green: B3, red: B4) of TOA and BOA reflectances (Level-2A for MAJA algorithm) for each date is shown in Figure 4, with the six plots displayed as red crosses within the study area. The three other methods tested also provided similar visual enhancements.



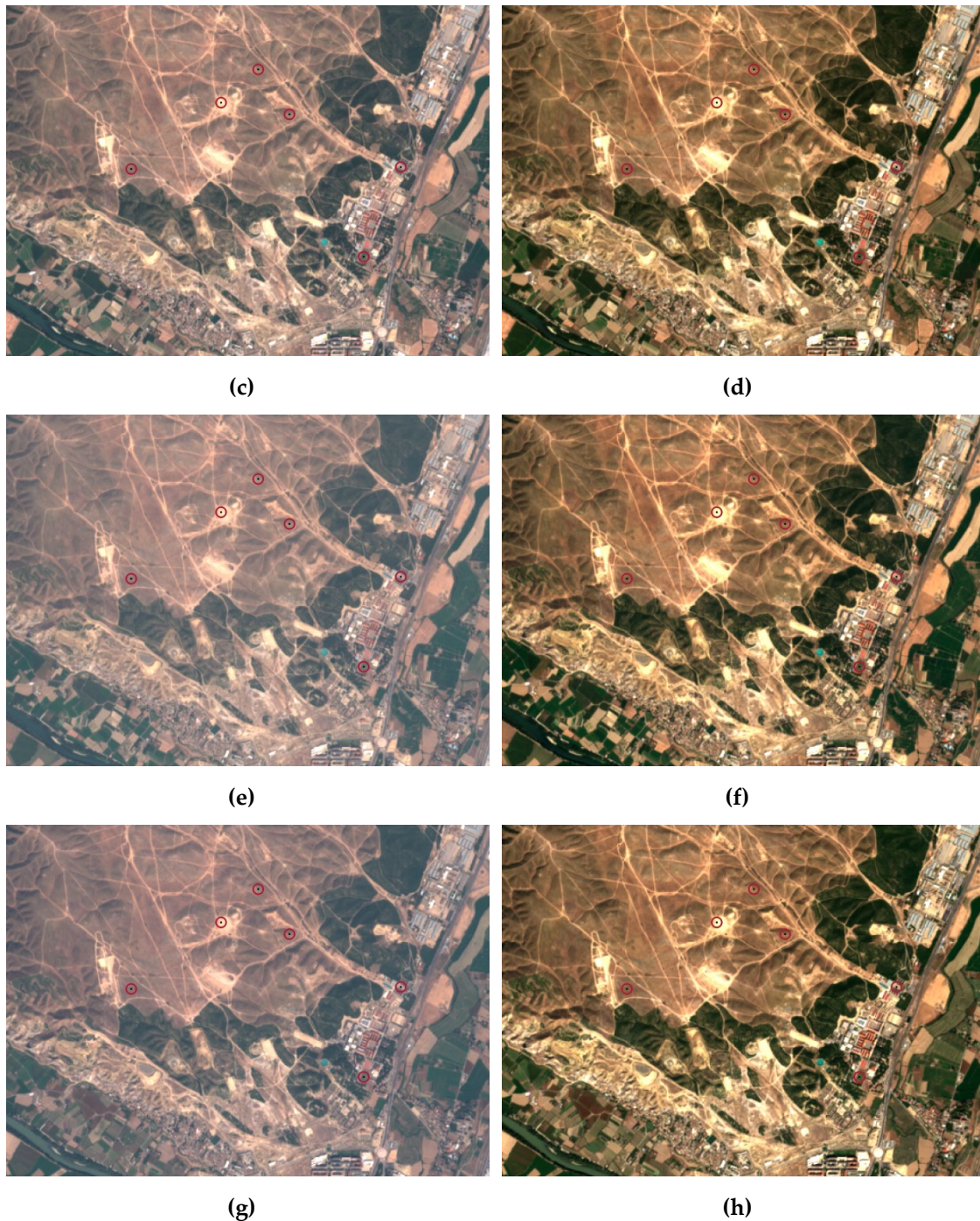


Figure 4. RGB composition of (a) TOA reflectance of 23/07 image, (b) BOA reflectance (MAJA) of 23/07 image, (c) TOA reflectance of 22/08 image, (d) BOA reflectance (MAJA) of 22/08 image, (e) TOA reflectance of 01/09 image, (f) BOA reflectance (MAJA) of 01/09 image, (g) TOA reflectance of 28/09 image, (h) BOA reflectance (MAJA) of 28/09 image.

3.1. Heterogeneity of land covers and radiometric uniformity of the 3 x 3 pixel kernel

As mentioned above, the results of the BRDF-adjusted BOA reflectances generated with each AC method were evaluated both per land cover and per spectral band. Prior to the comparison between ground measured and S2-derived BOA reflectances, a study on the heterogeneity of the field spectra on the different dates was performed, so as to assess the variability per land cover. To this, we used the approximately 200 spectra collected by plot. As illustrated in Table 4, all the plots showed a high homogeneity (i.e., low standard deviation) along the study period, being asphalt the most homogenous cover. This was expected because asphalt is considered to be a pseudo-invariant land cover, composed by a single artificial material and therefore no changes were observed between the four dates.

Table 4. Standard deviation of the spectra collected in the field along the study season (the spectra collected in the four field dates) per plot and Sentinel 2A band.

Plot name	B2	B3	B4	B5	B6	B7	B8A	B11	B12
Asphalt	0.003	0.003	0.003	0.004	0.008	0.007	0.007	0.004	0.005
Pasture	0.009	0.012	0.020	0.021	0.013	0.012	0.019	0.028	0.020
Shrub 1	0.013	0.013	0.012	0.012	0.015	0.014	0.014	0.009	0.009
Shrub 2	0.002	0.004	0.005	0.006	0.008	0.008	0.012	0.012	0.011
Grass	0.009	0.012	0.018	0.019	0.020	0.021	0.023	0.020	0.013
Bare soil	0.019	0.022	0.028	0.030	0.039	0.040	0.040	0.044	0.046

The category with higher heterogeneity was bare soil. This can be related to the fact that during the study period some off-road driving lessons and exercises were held in the plot, changing the conditions of stoniness and roughness. The rest of the categories (pasture, shrub and grass) showed intermediate values of homogeneity in all the bands. This fact may be due to their rich diversity of species, heterogeneous spatial pattern (presence of different shrubs and grass species and bare soil patches with different degrees of stoniness) and to the drying process of these plant formations throughout the summer. In most cases heterogeneity was higher in NIR and SWIR bands than in visible bands.

In order to ensure the radiometric uniformity of the 3×3 pixel kernel corresponding to the 20 m grid defined in Sentinel-2A scenes for each plot, the coefficient of variation (CV) was

calculated using the four L1C scenes together per plot and sensor band. As can be seen in Table 5, the CV values obtained are very low, being always lower than 0.05 in all bands, with the only exception of asphalt. As a result, the premise about radiometric uniformity of the 3 × 3 pixel kernel along all the study period is validated, and possible geometric correction errors in Sentinel-2A are avoided. The reason of higher CV values of the asphalt 3x3 kernel is that it includes two pixels that are contaminated by parking roofs spectral response (see Figure 4). This must be in mind in order to read the results obtained to this land cover in the present study.

Table 4. Coefficient of variation (CV) of the 3 × 3 pixel kernel corresponding to the 20 m grid defined in Sentinel-2A scenes for each plot using the four L1C scenes together per plot and sensor band.

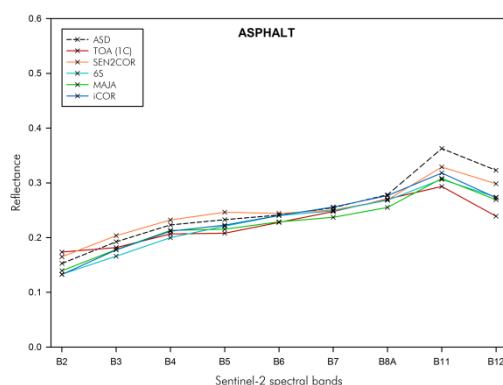
Plot name	B2	B3	B4	B5	B6	B7	B8A	B11	B12
Asphalt	0.05	0.07	0.08	0.08	0.07	0.07	0.07	0.09	0.10
Pasture	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.01
Shrub 1	0.01	0.01	0.01	0.03	0.03	0.03	0.03	0.02	0.02
Shrub 2	0.02	0.03	0.03	0.01	0.01	0.00	0.01	0.01	0.01
Grass	0.01	0.02	0.03	0.03	0.04	0.02	0.01	0.02	0.02
Bare soil	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.01	0.02

3.2. Spectral signatures

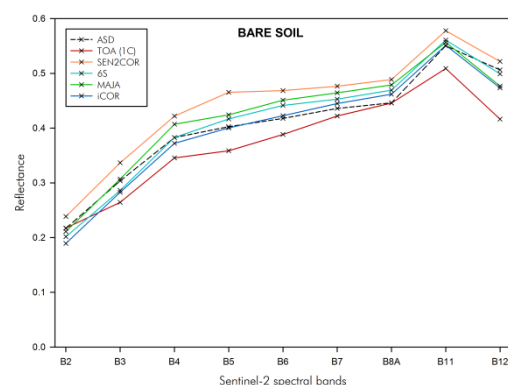
The comparison between the field BOA reflectances and S2 images per spectral band and land cover is evaluated in Figure 5, showing the average spectral signature of the four observation dates. In general terms, the highest difference between TOA and BOA reflectances was observed in visible bands, due to the exponential decay of atmospheric scattering effects at longer wavelengths (Martins et al., 2017). The asphalt plot showed a clear underestimation of TOA reflectance in SWIR bands, which was corrected to a certain degree with AC methods, especially with SEN2COR. On the contrary, minor differences were observed between methods in visible and red edge bands. SEN2COR was the method performing better in the visible bands. In the case of bare soil, higher reflectances were observed in all the spectral bands, as it was expected

from this land cover. In the visible bands 6S and MAJA obtained BOA reflectances closer to the ground measurements, whereas iCOR performed better in the red edge and NIR regions. In contrast, 6S showed the best performance in B12 (SWIR). In most of the bands the underestimation of Level-1C TOA reflectance was corrected for this land cover, yet many of the methods tended to over-correct atmospheric effects yielding BOA reflectance values higher than observed field spectral, especially in the NIR.

The grass plot showed a typical spectral signature of a vegetated cover, with a clear increase of reflectance from red (B4) to NIR (B8A). Besides, the reduction of reflectance in visible bands after AC was especially clear in this cover with all AC methods that only showed minor differences. On the opposite, red edge and NIR band reflectances were overestimated with the four methods, while SWIR showed better results in all the cases. Finally, in the case of pastures and shrubs (Figures 5d and 5e) the TOA reflectance was clearly overestimated in all the spectral bands and AC methods were not completely successful in correcting these differences. The shrub plot showed better results in the visible bands, where agreement with field spectra improved after AC, especially with MAJA, although in NIR bands the results varied depending on the method. On the contrary, SWIR bands showed a poor performance of Level-2A products, with overestimated reflectances in B11 and B12. In fact, the agreement with the field spectra was worse for the obtained Level-2A products than for the original Level-1C product. Notwithstanding that, the performance of each method could vary if different plots were selected in the comparison.



(a)



(b)

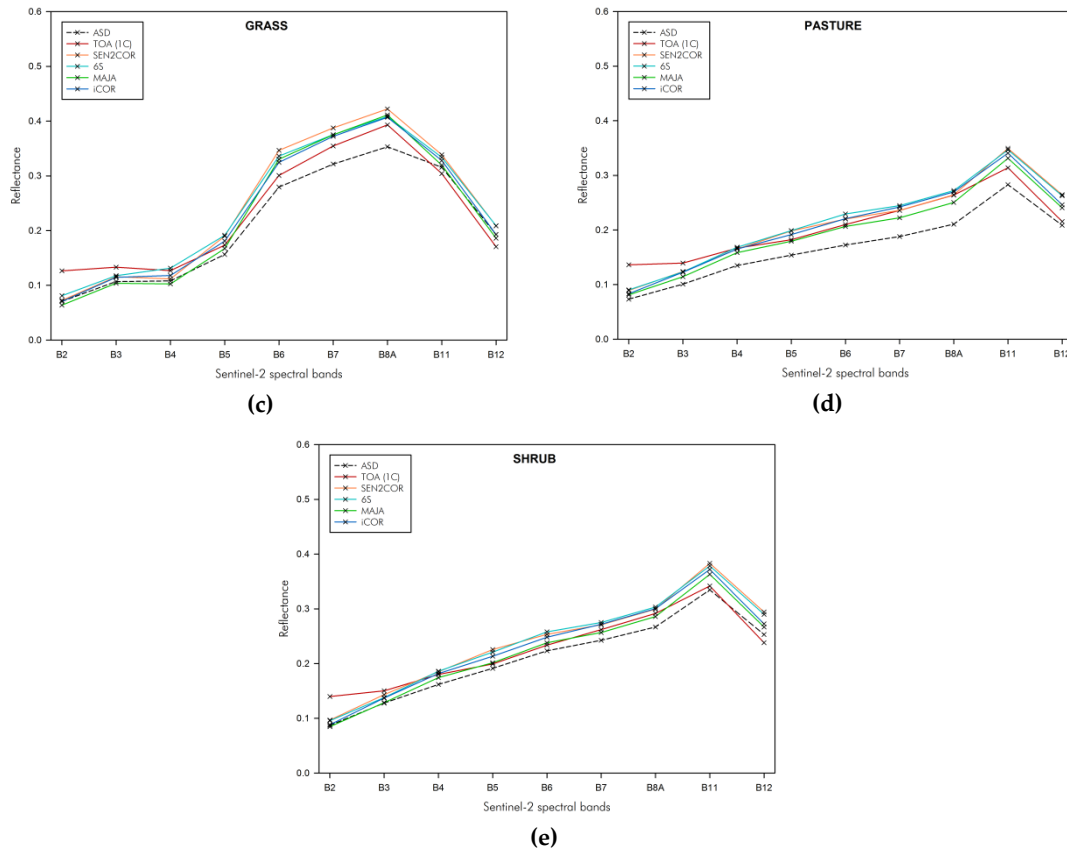


Figure 5. Spectral signatures of five land cover plots for each Level-2A product besides Level-1C uncorrected product. Average of the four dates considered. (a) Asphalt plot, (b) Bare soil plot, (c) Grass plot, (d) Pasture and (e) Shrub, average of the two shrub plots.

3.3. Statistical agreement of the different methods

The quantitative agreement between field spectra and BOA reflectance values obtained with each AC method was evaluated using R^2 and RMSE as figures of merit (Table 5):

Table 5. R^2 and RMSE obtained per AC method for all the spectral bands and dates. Best method is highlighted in bold.

	1C	SEN2COR	6S	MAJA	iCOR
R^2	0.919	0.954	0.920	0.947	0.939
RMSE	0.036	0.038	0.037	0.028	0.031

The results of Table 5 depict a clear increase in correlation between ground measurements and S2-derived reflectances after AC, with minor differences between methods. Similarly, iCOR

and MAJA reduced the RMSE achieved by Level-1C product. However, SEN2COR and 6S only showed slight or no improvement at all.

3.3.1. Analysis per land cover

Table 6 shows the R^2 obtained for each land cover, taking into account the four dates and the nine spectral bands. As it is observed, there is a clear correlation between field spectra and BOA reflectance in all methods, with R^2 values equal or higher than 0.90 in all cases, values comparable to those found in another recent study (Lange et al., 2017). Correlations improved clearly after AC correction with higher R^2 values achieved by Level-2A products compared to the Level-1C product. Among the land covers, the highest correlation was obtained for grass and shrub, two of the most heterogeneous covers, with the latter showing a great performance with the four AC methods, with minor differences between them. Yet, the degree of improvement after AC (i.e., comparison between R^2 values obtained after and before AC correction), was highest for asphalt. In his case, although differences between methods were stronger (ranging from 0.89 with 6S to 0.96 with MAJA), all the methods improved the correlation with the original Level-1C TOA reflectance. Altogether, the comparison between methods does not enable a clear identification of the best method for all cases, although in general ICOR, MAJA and SEN2COR achieved each the highest R^2 in two out of five covers.

Table 6. R^2 obtained per land cover and AC method. Best method for each cover highlighted in bold

Land cover	L1C	SEN2COR	6S	MAJA	iCOR
ASPHALT	0.810	0.948	0.896	0.961	0.913
BARE SOIL	0.923	0.932	0.909	0.911	0.950
GRASS	0.936	0.974	0.962	0.970	0.974
PASTURE	0.897	0.962	0.953	0.944	0.949
SHRUB	0.933	0.976	0.969	0.987	0.981

Similarly, Table 7 shows the RMSE obtained per land cover using the four AC methods evaluated. In all the cases the RMSE values were lower than 0.05, with values below 0.02 in the best cases.

Generally, the error achieved by Level-1C products decreased after AC, with reductions close to 50% of the original error in the best case. The comparison between methods again did not show a clear superior performance of a single method for all the land covers. Asphalt plot shows the lowest RMSE with SEN2COR, while iCOR performed the best with bare soil and grass plots. Finally, MAJA achieved a higher performance with pasture and shrubs. In conclusion, different AC methods showed different results for specific land cover types, in line with Nazeer et al. (2014). Notwithstanding that, further research with a higher number of samples is required to obtain a more complete understanding of the performance of each processor.

Table 7. RMSE obtained by land cover use. Best method for each cover highlighted in bold

Land cover	L1C	SEN2COR	6S	MAJA	iCOR
ASPHALT	0.039	0.019	0.030	0.030	0.025
BARE SOIL	0.043	0.045	0.032	0.032	0.024
GRASS	0.033	0.043	0.040	0.034	0.033
PASTURE	0.043	0.047	0.050	0.034	0.044
SHRUB	0.026	0.033	0.034	0.017	0.026

3.3.2. Band wise analysis

Table 8 and 9 show the band wise results obtained with the different AC methods. These results were obtained for the four studied dates and for the five land covers. First, R^2 values are analyzed (Table 8), suggesting a slight improvement in correlation for all methods if compared to the Level-1C product, with exceptions. Highest correlations were obtained for visible bands (B2, B3 and B4), with R^2 values above 0.95 in most cases. Then, R^2 values showed a decreasing trend as wavelengths increased, with lower values achieved for Red Edge (B5, B6 and B7), NIR (B8A) and SWIR bands (B11 and B12). The comparison between methods showed minor differences except in the case of the blue band (B2), where R^2 ranged from 0.90 in the worst scenario (6S) to 0.97 in the best one (MAJA and iCOR). In contrast, other spectral bands showed very slight differences between methods, and in some bands the highest correlating option was

the uncorrected Level-1C product (B6, B7, B8A and B11). The highest correlating method in three out of nine bands was SEN2COR, but differences were minor.

Table 8. R² obtained per spectral band. Best method for each band highlighted in bold.

Spectral band	1C	SEN2COR	6S	MAJA	iCOR
B2	0.946	0.949	0.900	0.971	0.975
B3	0.973	0.972	0.928	0.973	0.977
B4	0.974	0.977	0.937	0.969	0.972
B5	0.967	0.967	0.926	0.959	0.964
B6	0.939	0.926	0.873	0.919	0.927
B7	0.928	0.916	0.874	0.913	0.918
B8A	0.905	0.895	0.843	0.886	0.893
B11	0.840	0.836	0.772	0.835	0.833
B12	0.920	0.926	0.881	0.924	0.921

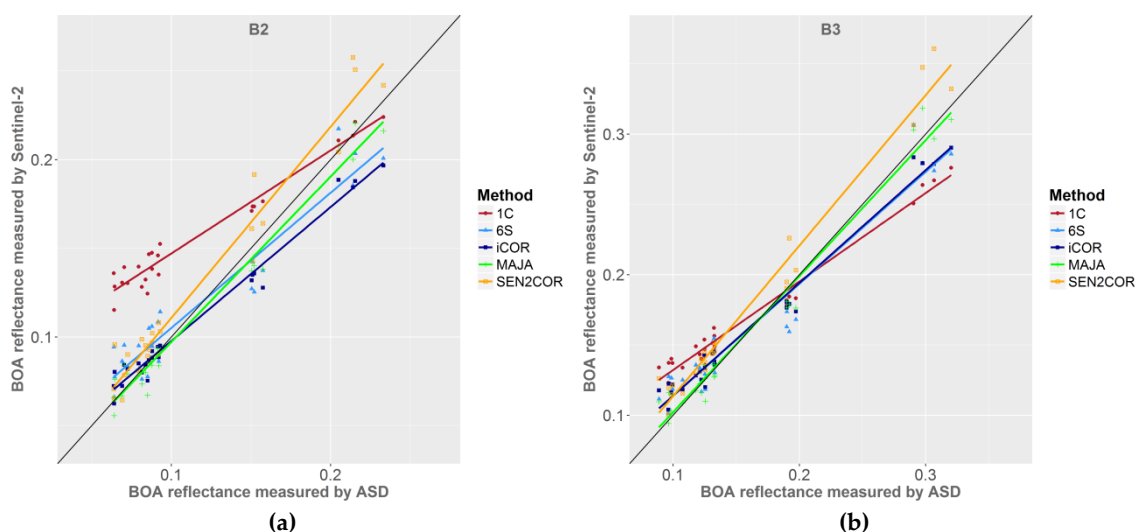
The comparison of the band wise performance in terms of RMSE showed better results for visible bands, maybe due to the higher spatial resolution, intermediate results for red edge and NIR bands, while SWIR bands showed a poorer performance, with RMSE as high as 0.10 in B12 (see Table 9). The RMSE values achieved by the different AC methods showed also minor differences between methods. In some bands, RMSE values after AC were even larger than before AC (bands B6, B7 and B8A). Although differences were not marked, MAJA achieved the lowest overall RMSE with minimum values in bands B2-B5 and also in B11-B12.

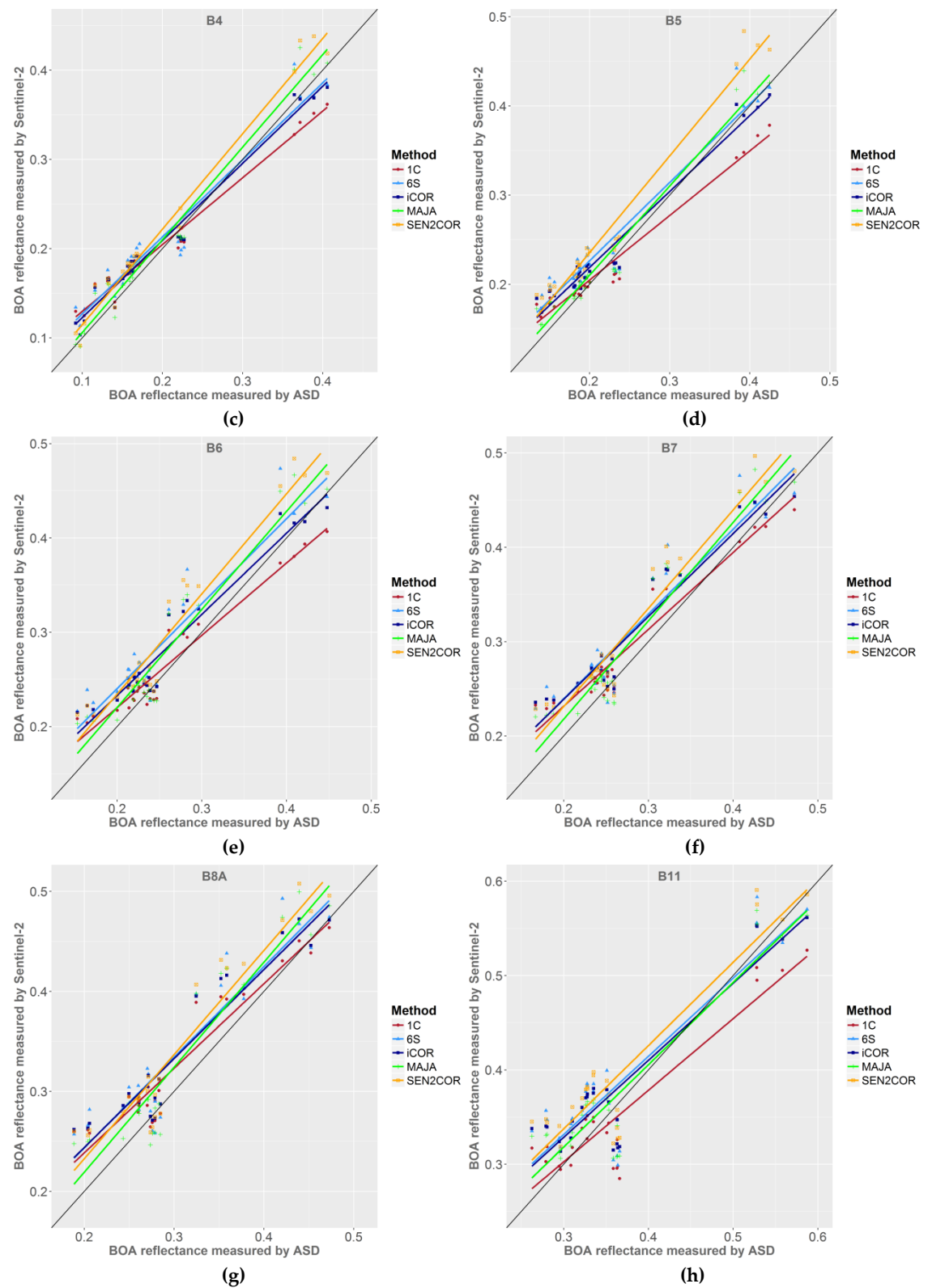
Table 9. RMSE obtained per spectral band. Best method for each band highlighted in bold.

	1C	SEN2COR	6S	MAJA	iCOR
B2	0.047	0.018	0.019	0.010	0.016
B3	0.029	0.023	0.022	0.012	0.017
B4	0.026	0.027	0.028	0.020	0.021
B5	0.026	0.042	0.034	0.020	0.025
B6	0.025	0.045	0.044	0.033	0.033
B7	0.029	0.043	0.042	0.034	0.037
B8A	0.033	0.046	0.046	0.038	0.041
B11	0.039	0.047	0.047	0.038	0.040
B12	0.099	0.104	0.101	0.076	0.083

3.4. Scatterplots

Scatterplots of ground reflectance measurements versus BOA reflectances obtained from S2 Level-2A products (besides of TOA reflectances from Level-1C products as reference) provide help interpreting the values given in tables 5-9. Figure 6 shows the scatterplots of the nine spectral bands considered. In general, Level-1C products overestimated low reflectance values in all spectral bands except B11 and B12. This behavior was especially noticeable in the case of visible bands (B2, B3 and B4), and the performance clearly improved after AC (see Figures 6a, 6b and 6c). On the contrary, a strong underestimation of higher BOA reflectances was observed in all the spectral bands except B8A for the Level-1C product. This effect was higher in SWIR bands (see Figures 6h and 6i), although Level-2A products attenuated this underestimation, especially with SEN2COR and 6S. Again, the comparison between methods did not show a clear superior method for all the spectral bands and all of them balanced this difference between high and low reflectances. Band B2 (blue) showed a better performance with MAJA with a regression line close to the 1:1 line, while ICOR showed a poor performance due to a clear underestimation of high reflectance values. Bands B3 and B4 followed a similar trend of overestimation of low values and underestimation of high values with Level-1C, which improved after AC, with a superior performance of MAJA. The comparison of methods for red edge spectral bands (see Figures 6d, 6e and 6f) showed a slightly better performance of MAJA, 6S and iCOR methods, in line with the results of Table 9.





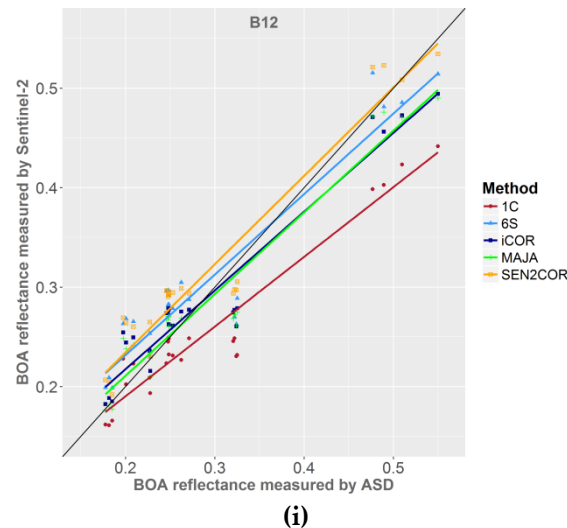


Figure 6. Scatterplot of each spectral band. BOA reflectance measured by ASD in X axis versus BOA reflectance measured by S2 in Y axis.

Finally, the poor performance of Level-1C product at SWIR bands improved clearly after AC. At these bands MAJA showed the best performance at low reflectance values, whereas SEN2COR and 6S achieved the best results at high reflectance values. The underestimation of SWIR BOA reflectances could be due to the results obtained for the asphalt and bare soil plots (see Figure 5a and 5b), where the strong underestimation obtained by Level-1C product, overall in the asphalt plot, was only partially solved by Level-2A products. Possible geometric correction errors in Sentinel 2A scenes could explain this fact, as the spectral response can be partially or totally identified with the roofs includes in two of the 3x3 kernel of this cover.

3.5. Assessment of S2-derived spectral indices

3.5.1. Normalized Difference Vegetation Index (NDVI)

The BOA reflectances of bands B4 and B8A were used as source to calculate NDVI scenes, and subsequently the statistics of the six plots were obtained for the different dates, so as to compare them with the ground measurements. The results of Figure 7a show typical low values of NDVI (lower than 0.3) in pasture, shrub, and specially asphalt and bare soil, consequence of the dry conditions of these dates in the region in the case of vegetated covers, and due to the

typical spectral behavior of man-made surfaces in the case of asphalt. In contrast, grass cover showed a much higher NDVI value, ~0.55 depending the AC method.

Moreover, figure 7b shows the difference in NDVI between the different AC methods and ground measurements-derived NDVI. These results show a slight overestimation of NDVI in most land covers, due to the overestimation of NIR band observed in Figure 5, with the exception of the asphalt cover. These results are in agreement with previous findings, such as Bru et al. (2017), where atmospherically corrected products provided by MACCS (the precursor of MAJA) and 6SV (the vector version of 6S) simulations showed a large overestimation of BOA reflectances, although that work was focused on costal land covers. The comparison of AC methods in Figure 7 shows a better performance of MAJA in every cover except for grass. These results are consistent with Figures 6c and 6g, where this method showed that for high reflectance values band B8A was overestimated. Nevertheless, these higher NDVI values for grass had a direct impact on the quantitative evaluation of this land cover.

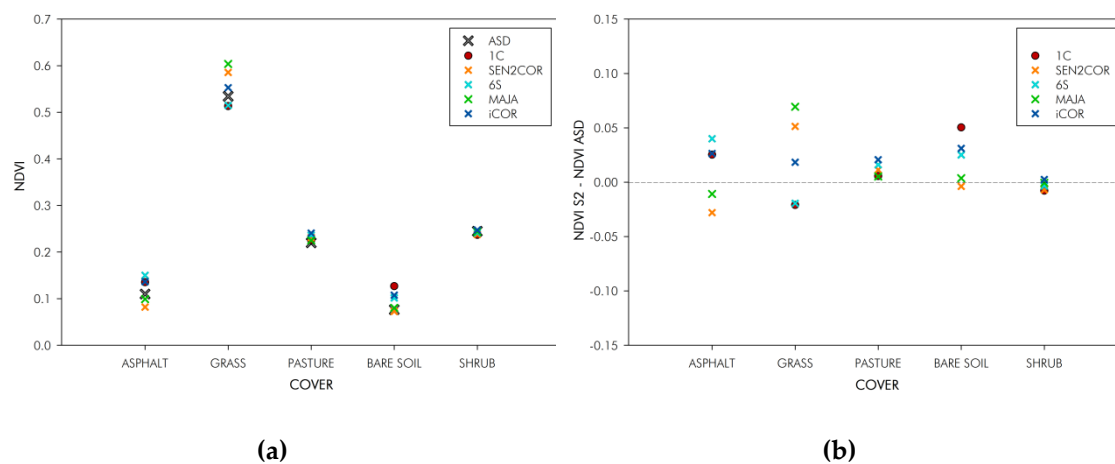


Figure 7. **a)** NDVI measured from BOA reflectance of ASD and Sentinel-2 products, **b)** Difference of NDVI measured from ASD and from Sentinel-2 products ($NDVI_{S2} - NDVI_{ASD}$).

3.5.2. Normalized Difference Infrared Index (NDII)

As mentioned above, NDII index is frequently used for estimating vegetation water content. Most vegetated covers in the study period were senescent, so, the negative values achieved by

most covers were expected (Fig. 8a) (Yilmaz et al., 2008). On the contrary, values of 0.05 to 0.15 were obtained for grass, since it was wetter. Figure 8b shows an overestimation of this index in all covers except for shrubs. The use of AC methods clearly improved the results obtained by Level-1C products, and no clear superior performance of a single method for all the land covers was observed. On the contrary MAJA did not succeed in reducing the overestimation of NDII index of Level-1C product on the grass plot.

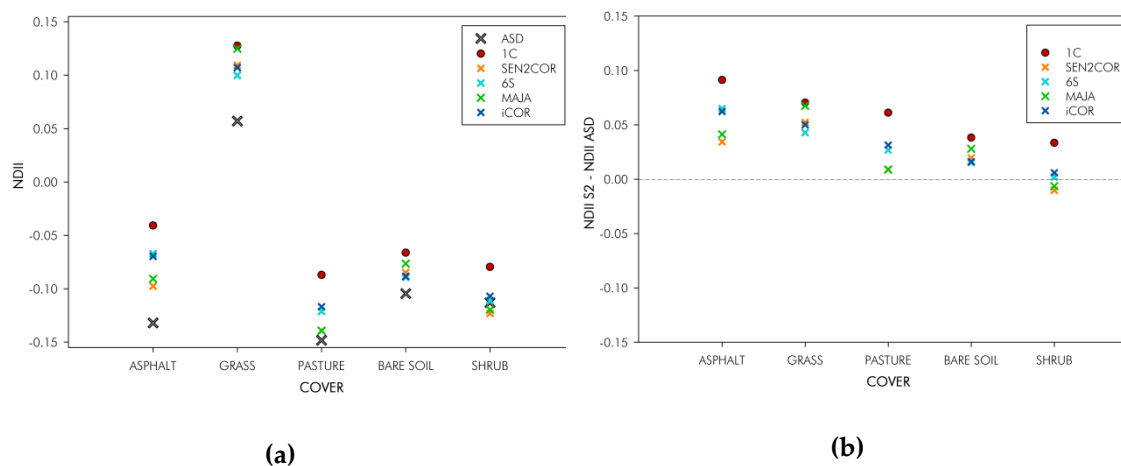


Figure 8. **a)** NDII measured from BOA reflectance of ASD and Sentinel-2 products, **b)** Difference of NDII measured from ASD and from Sentinel-2 products ($NDII_{S2} - NDII_{ASD}$).

3.5.3. Normalized Burn Ratio (NBR)

In line with the other two spectral indices, NBR (Figure 9a) showed higher values for the grass plot. Additionally, most AC methods showed a clear overestimation on this index too, yet AC improved the fit with ground data in most cases. SEN2COR performed slightly better, with a difference of NBR lower than 0.04 in all land covers. On the contrary, iCOR only partially reduced the overestimation of NBR.

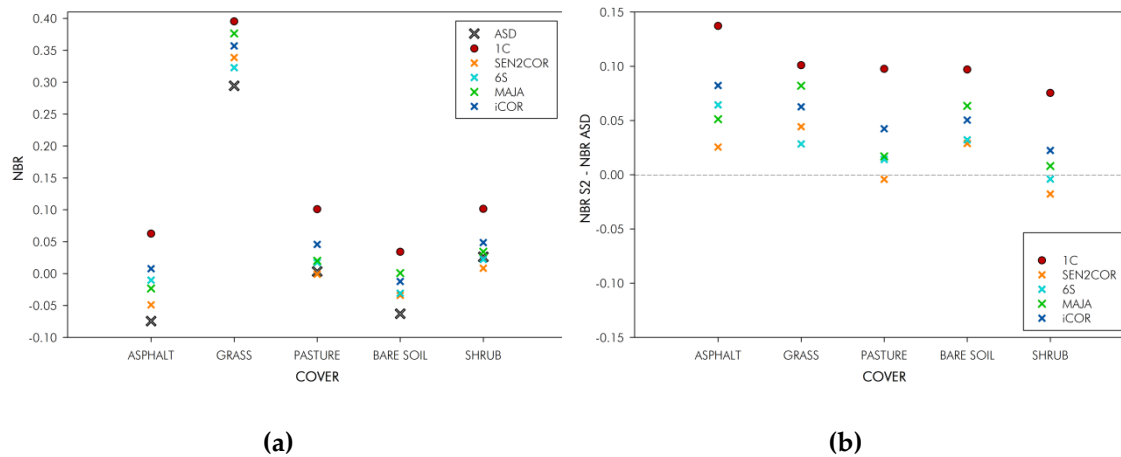


Figure 9. **a)** NBR measured from BOA reflectance of ASD and Sentinel-2 products, **b)** Difference of NBR measured from ASD and from Sentinel-2 products ($NBR_{S2} - NBR_{ASD}$)

3.5.4. RMSE of S2- derived spectral indices

Table 10 shows the RMSE between S2-derived spectral indices compared to the ones obtained from ground measurements. Depicted values from Table 10 were consistent with Figures 7, 8 and 9 and agreed with the higher performance of Level-2A-derived spectral indices, compared to the Level-1C product. This improvement was greater in the case of NBR index, probably due to the better performance of BOA reflectances on B12, as seen in Figure 6i. In contrast with the results of Table 5 the lowest RMSE of the different spectral indices were achieved by SEN2COR and iCOR, while MAJA performed slightly worse.

Table 10. RMSE of spectral indices from different AC methods for the six plots and four dates.

Best method for each index highlighted in bold

RMSE	Level-1C	SEN2COR	6S	MAJA	iCOR
NDVI	0.0316	0.0307	0.0280	0.0307	0.0255
NDII	0.0599	0.0304	0.0365	0.0368	0.0370
NBR	0.1006	0.0298	0.0350	0.0494	0.0539

Level-1C products showed an overestimation of NDII and NBR indices, which derived in greater RMSE values. These values were clearly reduced by corrected BOA reflectances, in line with the results of Figures 8 and 9. On the contrary, the performance of each AC method in the estimation of NDVI varied for each land cover. This varying performance affected the RMSE

values of Table 10, with only minor improvements of RMSE for NDVI index. These results suggest that the selection of the suitable method is application dependent.

4. Conclusions

A rigorous evaluation of the accuracy of Level-2A BOA reflectance products obtained with different AC methods requires their comparison with ground measurements obtained with spectroradiometers. This comparison is necessary to quantify the performance of each AC method both band wise and per land cover. The coefficient of determination or the root mean square error can be used to quantitatively measure this performance, and therefore the different AC methods are evaluated.

The results obtained by the RMSE and R^2 suggest a varying performance of AC methods depending on the land cover and spectral bands. Although minor differences were observed in most cases, MAJA and iCOR showed a better performance according to the RMSE in the analysis per land cover. Regarding the band wise analysis, MAJA ranked first in all spectral bands. A detailed inspection of the scatterplots, showed a common trend of overestimation of low reflectance values, together with a clear underestimation of high BOA reflectance values in the Level-1C product. This overestimation was more pronounced in band B8A, so the three spectral indices derived from it showed also clear signals of overestimation. Among the AC methods, SEN2COR performed the best in the estimation of two spectral indices (NDII and NBR), while iCOR showed a slightly superior performance in NDVI.

Summing up, the results obtained on this inter-comparison of AC methods showed a clear enhancement of the accuracy of BOA reflectance estimation compared to the performance of uncorrected Level-1C product. This proves the need of an accurate AC, essential preprocessing step for a large number of Remote Sensing applications, especially multitemporal studies. Nevertheless, the comparison of four AC methods showed minor differences between them.

On the one hand, the statistical analysis of spectral signatures showed a better performance of MAJA, as seen in the figures of merit. On the other hand, the assessment of spectral indices derived from BOA reflectances showed a slightly superior performance of SEN2COR. Consequently, the selection of an AC method needs to be aligned with the study purposes. Yet, this study is based on a reduced amount of comparison points and further research is required to assess the performance of each method on a higher variety of land covers and atmospheric conditions. Besides, the quality of additional products such as geophysical and cloud masks must be assessed to decide which method is the best to correct the S2 imagery.

Acknowledgments: This work has been supported by the SynerTGE project (CGL2015-69095-R), funded by the Spanish Ministry of Economy and Competitiveness; the HyZCP project (2015-17), funded by the Centro Universitario de la Defensa de Zaragoza, and projects CGL2016-75217-R (MINECO/FEDER, EU) and PyrenEOS EFA 048/15, which has been 65% cofinanced by the European Regional Development through the Interreg V-A Spain-France-Andorra programme (POCTEFA 2014-2020).

Conflicts of Interest: The authors declare no conflict of interest

References

- Abdou, W. A.; Conel, J. E.; Pilorz, S. H.; Helmlinger, M. C.; Bruegge, C. J.; Gaitley, B. J.; Ledebor, W.C.; Martonchik, J. V. 2001. Vicarious calibration a reflectance-based experiment with airMISR. *Remote Sensing of Environment* 77(3), 338–353.
- Anyamba, A.; Tucker, C.J. 2005. Analysis of sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. *J. Arid Environ.* 63, 596–614.
- ASD inc. 2017. Fieldspec®4 user manual. Boulder (Colorado). Available online: <http://support.asdi.com/Products/FileGet.aspx?f=600979-g%20FieldSpec4UG.pdf> (Accessed on 1 October 2017).
- Astrium services. 2013. SPOT 6 & SPOT 7 imagery user guide. Available online: https://www.spaceoffice.nl/blobs/Dataportaal/User_Guide_SPOT6_V1.0.pdf (Accessed on 21 January 2018)
- Baldrige, A.; Hook, S.; Grove, C.; Rivera, G. 2009. The Aster spectral library version 2.0. *Remote Sensing of Environment*, 113(4), 711–715.
- Berk, A.; Anderson, G.P.; Acharya, P.K.; Bernstein, L.S.; Muratov, L.; Lee, J.; Fox, M.; Adler-Golden, S.M.; Chetwynd, J.H.; Hoke, M.L.; 2006. MODTRAN5: 2006 update, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII. *International Society for Optics and Photonics*, p. 62331F.
- Borràs, J.; Delegido, J.; Pezzola, A.; Pereira, M.; Morassi, G.; Camps-Valls, G. 2017. Clasificación de usos del suelo a partir de imágenes Sentinel-2, *Revista de Teledetección*, 48, 55–66.

- Bru, D.; Lubac, B.; Normandin, C.; Robinet, A.; Leconte, M.; Hagolle, O.; Martiny, N.; Jamet, C., 2017. Atmospheric Correction of Multi-Spectral Littoral Images Using a PHOTONS/AERONET-Based Regional Aerosol Model. *Remote Sens.* 9, 814.
- Carlson, T.N.; Ripley, D.A. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sens. Environ.* 62(3), 241-252.
- Clevers, J.; Kooistra, L.; van den Brande, M. 2017. Using Sentinel-2 data for retrieving LAI and leaf and canopy chlorophyll content of a potato crop. *Remote Sens.* 9, 405.
- Colkesen, I.; Kavzoglu, T. 2017. Ensemble-based canonical correlation forest (CCF) for land use and land cover classification using Sentinel-2 and Landsat OLI imagery. *Remote Sens. Lett.* 8, 1082-1091.
- Chrysafis, I.; Mallinis, G.; Siachalou, S.; Patias, P. 2017. Assessing the relationships between growing stock volume and Sentinel-2 imagery in a mediterranean forest ecosystem. *Remote Sens. Lett.* 8, 508-517.
- Chuvieco, E. 2010. Teledetección espacial: La observación de la tierra desde el espacio. Madrid. ISBN: 978-84-344-3498-1.
- Danner, M.; Locherer, M.; Hank, T.; Richter, K. 2015. Spectral Sampling with the ASD FieldSpec 4 – Theory, Measurement, Problems, Interpretation. EnMAP Field Guides Technical Report, GFZ Data Services. <http://doi.org/10.2312/enmap.2015.008>.
- Davidson, A.; Wang, S.; Wilmschurst, J. 2006. Remote sensing of grassland–shrubland vegetation water content in the shortwave domain. *Int. J. Appl. Earth Obs. Geoinf.* 8, 225-236.
- Doxani, G.; Vermote, E.; Roger, J.-C.; Gascon, F.; Adriaensen, S.; Frantz, D.; Hagolle, O.; Hollstein, A.; Kirches, G.; Li, F.; Louis, J.; Mangin, A.; Pahlevan, N.; Pflug, B.; Vanhellemont, Q., 2018. Atmospheric Correction Inter-Comparison Exercise. *Remote Sens.* 10, 352.
- Dörnhöfer, K.; Göritz, A.; Gege, P.; Pflug, B.; Oppelt, N. 2016. Water constituents and water depth retrieval from Sentinel-2A—A first evaluation in an oligotrophic lake. *Remote Sens.* 8, 941.
- Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P., et al., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES operational services. *Remote Sens. Environ.* 120, 25-36.
- ESA, 2015. Sentinel-2 User Handbook, in: Agency, E.S. (Ed.).
- Fraser, R.; van der Sluijs, J.; Hall, R. 2017. Calibrating satellite-based indices of burn severity from UAV-derived metrics of a burned boreal forest in NWT, Canada. *Remote Sens.* 9, 279.
- Gascon, F.; Bouzinac, C.; Thépaut, O.; Jung, M.; Francesconi, B.; Louis, J.; Lonjou, V.; Lafrance, B.; Massera, S.; Gaudel-Vacaresse, A., et al. 2017. Copernicus Sentinel-2A calibration and products validation status. *Remote Sens.* 9, 584.
- Guanter Palomar, L., 2007. New algorithms for atmospheric correction and retrieval of biophysical parameters in Earth Observation. Application to ENVISAT/MERIS data. Universitat de València.
- Hagolle, O.; Huc, M.; Villa Pascual, D.; Dedieu, G., 2015a. A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENμS and Sentinel-2 Images. *Remote Sens.* 7, 2668.
- Hagolle, O.; Sylvander, S.; Huc, M.; Claverie, M.; Clesse, D.; Dechoz, C.; Lonjou, V.; Poulain, V. 2015b. Spot-4 (take 5): Simulation of Sentinel-2 time series on 45 large sites. *Remote Sens.* 7, 12242.
- Immitzer, M.; Vuolo, F.; Atzberger, C. 2016. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* 8, 166.
- Korhonen, L.; Hadi; Packalen, P.; Rautiainen, M. 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sens. Environ.* 195, 259-274.

- Kukawska, E.; Lewinski, S.; Krupinski, M.; Malinowski, R.; Nowakowski, A.; Rybicki, M.; Kotarba, A. 2017. Multitemporal Sentinel-2 data - remarks and observations. In 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images.
- Lange, M., Dechant, B., Rebmann, C., Vohland, M., Cuntz, M., Doktor, D., 2017. Validating MODIS and Sentinel-2 NDVI products at a temperate deciduous forest site using two independent ground-based sensors. *Sensors*, 17, 1855.
- Lantzanakis, G.; Mitraka, Z.; Chrysoulakis, N. 2017. Comparison of physically and image based atmospheric correction methods for Sentinel-2 satellite imagery. In *Perspectives on atmospheric sciences*, Springer, 255-261.
- Le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailier, J. Y., Bréda, N., Genet, H., Davi, H., Dufrêne, E. 2008. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sensing of Environment*, 112(10), 3846–3864.
- Lebourgeois, V.; Dupuy, S.; Vintrou, É.; Ameline, M.; Butler, S.; Bégué, A. 2017. A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). *Remote Sens.* 9, 259.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. Sentinel-2 Sen2Cor: L2A Processor for Users, *Proceedings Living Planet Symposium 2016*. Spacebooks Online, pp. 1-8.
- Lefebvre, A.; Sannier, C.; Corpetti, T. 2016. Monitoring urban areas with Sentinel-2A data: Application to the update of the Copernicus high resolution layer imperviousness degree. *Remote Sens.* 8, 606.
- Lillesand, T.; Kiefer, R.W.; Chipman, J. 2014. *Remote sensing and image interpretation*. John Wiley & Sons. ISBN: 978-1118343289
- Lonjou, V.; Desjardins, C.; Hagolle, O.; Petrucci, B.; Tremas, T.; Dejus, M.; Makarau, A.; Auer, S. 2016. MACCS-ATCOR joint algorithm (MAJA), In *Proceedings of SPIE - The International Society for Optical Engineering*.
- Mallinis, G.; Mitsopoulos, I.; Chrysafi, I. 2017. Evaluating and comparing Sentinel 2A and Landsat-8 Operational Land Imager (OLI) spectral indices for estimating fire severity in a Mediterranean pine ecosystem of Greece. *GIScience & Remote Sens.* 55, 1-18.
- McCoy, R.M. 2005. Field spectroscopy. In *Field methods in remote sensing*, Guilford Press, New York, 42-58. ISBN: 978-1593850791
- Marcello, J.; Eugenio, F.; Perdomo, U.; Medina, A. 2016. Assessment of atmospheric algorithms to retrieve vegetation in natural protected areas using multispectral high resolution imagery. *Sensors*, 16, 1624.
- Martins, V.S.; Barbosa, C.C.F.; de Carvalho, L.A.S.; Jorge, D.S.F.; Lobo, F.L.; de Moraes Novo, E.M.L. 2017. Assessment of atmospheric correction methods for Sentinel-2 MSI images applied to amazon floodplain lakes. *Remote Sens.* 9, 322.
- Melendo-Vega, J.R., Martín, M.P., Vilar del Hoyo, L., Pacheco-Labrador, J., Echavarría, P., Martínez-Vega, J. 2017. Estimation of grassland biophysical parameters in a “dehesa” ecosystem from field spectroscopy and airborne hyperspectral imagery. *Revista de Teledetección* 48, 13-28.
- Mendiguren, G., Martín, P., Nieto, H., Pacheco-Labrador, J., Jurdao, S. 2015. Seasonal variation in grass water content estimated from proximal sensing and MODIS time series in a Mediterranean Fluxnet site. *Biogeosciences* 12 (18), 5523-5535.
- Milton, E.J.; Schaepman, M.E.; Anderson, K.; Kneubühler, M.; Fox, N. 2009. Progress in field spectroscopy. *Remote Sens. Environ.* 113, S92-S109.
- Müller-Wilm, U. 2017. Sentinel-2 MSI – Level-2A prototype processor installation and user manual. Available online: <http://step.esa.int/thirdparties/sen2cor/2.2.1/S2PAD-VEGA-SUM-0001-2.2.pdf> (Accessed on 28 November 2017).

- Munyati, C. 2017. The potential for integrating Sentinel 2 MSI with Spot 5 HRG and Landsat 8 OLI imagery for monitoring semi-arid savannah woody cover. *Int. J. Remote Sens.* 38, 4888-4913.
- Murphy, S. 2017. Github repository of Sam Murphy. Available online: <https://github.com/samsammurphy> (Accessed on 1 of December 2017).
- Navarro, G.; Caballero, I.; Silva, G.; Parra, P.-C.; Vázquez, Á.; Caldeira, R. 2017. Evaluation of forest fire on Madeira island using Sentinel-2A MSI imagery. *Int. J. Appl. Earth Obs. Geoinf.* 58, 97-106.
- Nazeer, M., Nichol, J.E., Yung, Y.-K., 2014. Evaluation of atmospheric correction models and Landsat surface reflectance product in an urban coastal environment. *Int J Remote Sens* 35, 6271-6291.
- Novelli, A.; Aguilar, M.A.; Nemmaoui, A.; Aguilar, F.J.; Tarantino, E. 2016. Performance evaluation of object based greenhouse detection from Sentinel-2 MSI and Landsat 8 OLI data: A case study from Almería (Spain). *Int. J. Appl. Earth Obs. Geoinf.* 52, 403-411.
- Paul, F.; Winsvold, S.; Kääb, A.; Nagler, T.; Schwaizer, G. 2016. Glacier remote sensing using Sentinel-2. Part II: Mapping glacier extents and surface facies, and comparison to Landsat 8. *Remote Sens.* 8, 575.
- Pacheco-Labrador, J., Martín, M. P., Riaño, D., Hilker, T., Carrara, A. 2016. New approaches in multi-angular proximal sensing of vegetation: Accounting for spatial heterogeneity and diffuse radiation in directional reflectance distribution models. *Remote Sensing of Environment* 187, 447-457.
- Patel, P. N., Bhatt, H., Mathur, A. K., Prajapati, R. P., Tyagi, G. 2016. Reflectance-based vicarious calibration of INSAT-3D using high-reflectance ground target. *Remote Sensing Applications: Society and Environment* 3, 20-35.
- Pesaresi, M.; Corbane, C.; Julea, A.; Florczyk, A.; Syrris, V.; Soille, P. 2016. Assessment of the added-value of Sentinel-2 for detecting built-up areas. *Remote Sens.* 8, 299.
- Pons, X.; Pesquer, L.; Cristóbal, J.; González-Guerrero, O. 2014. Automatic and improved radiometric correction of Landsat imagery using reference values from MODIS surface reflectance images. *Int. J. Appl. Earth Obs. Geoinf.* 33, 243-254.
- Quintano, C.; Fernández-Manso, A.; Fernández-Manso, O. 2016. Sentinel-2A red-edge spectral indices suitability for discriminating burn severity. *Int. J. Appl. Earth Obs. Geoinf.* 50, 170-175.
- Richter, R.; Louis, J.; Berthelot, B. 2017. Sentinel-2 MSI – Level 2A products algorithm theoretical basis document. Available online: https://earth.esa.int/c/document_library/get_file?folderId=349490&name=DLFE-4518.pdf (Accessed on 28 November 2017).
- Richter, R.; Schlöpfer, D. 2015. ATCOR-2/3 user guide, version 9.0.0. DLR report DLR-IB, 565-01.
- Rouquié, B., Hagolle, O., Bréon, F.-M., Boucher, O., Desjardins, C., Rémy, S., 2017. Using Copernicus Atmosphere Monitoring Service Products to Constrain the Aerosol Type in the Atmospheric Correction Processor MAJA. *Remote Sens.* 9, 1230.
- Roy, D., Li, Z., Zhang, H., 2017a. Adjustment of Sentinel-2 Multi-Spectral Instrument (MSI) Red-Edge Band Reflectance to Nadir BRDF Adjusted Reflectance (NBAR) and Quantification of Red-Edge Band BRDF Effects. *Remote Sensing* 9, 1325.
- Roy, D.P., Li, J., Zhang, H.K., Yan, L., Huang, H., Li, Z., 2017b. Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. a School of Surveying, University of Otago. PO Box 56, Dunedin, New Zealand 199, 25-38.
- Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., Kovalsky, V., 2016. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. a School of Surveying, University of Otago. PO Box 56, Dunedin, New Zealand 176, 255-271.

- Schaepman-Strub, G.; Schaepman, M.E.; Painter, T.H.; Dangel, S.; Martonchik, J.V. 2006. Reflectance quantities in optical remote sensing—definitions and case studies. *Remote Sens. Environ.* 103(1), 27-42.
- Shoko, C.; Mutanga, O. 2017, Examining the strength of the newly-launched Sentinel 2 MSI sensor in detecting and discriminating subtle differences between C3 and C4 grass species. *ISPRS J. Photogramm. Remote Sens.* 129, 32-40.
- Sibanda, M.; Mutanga, O.; Rouget, M. 2016. Comparing the spectral settings of the new generation broad and narrow band sensors in estimating biomass of native grasses grown under different management practices. *GIScience & Remote Sens.* 53, 614-633.
- Sterckx, S., Knaeps, E., Adriaensen, S., Reusen, I., De Keukelaere, L., Hunter, P., Giardino, C., Odermatt, D., 2015a. OPERA: An atmospheric correction for land and water, *Proceedings of the ESA Sentinel-3 for Science Workshop, Venice, Italy*, pp. 2-5.
- Sterckx, S., Knaeps, S., Kratzer, S., Ruddick, K., 2015b. SIMilarity Environment Correction (SIMEC) applied to MERIS data over inland and coastal waters. *Remote Sens Environ* 157, 96-110.
- Toming, K.; Kutser, T.; Laas, A.; Sepp, M.; Paavel, B.; Nõges, T. 2016. First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery. *Remote Sens.* 8, 640.
- Traganos, D.; Reinartz, P. 2017. Mapping Mediterranean seagrasses with Sentinel-2 imagery. *Marine Pollution Bulletin*.
- Van der Meer, F.D., van der Werff, H.M.A., van Ruitenbeek, F.J.A., 2014. Potential of ESA's Sentinel-2 for geological applications. *Remote Sens. Environ.* 148, 124-133.
- Verhegghen, A.; Eva, H.; Ceccherini, G.; Achard, F.; Gond, V.; Gurllet-Fleury, S.; Cerutti, P. 2016. The potential of Sentinel satellites for burnt area mapping and monitoring in the Congo basin forests. *Remote Sens.* 8, 986.
- Veraverbeke, S.; Stavros, E.N.; Hook, S.J. 2014. Assessing fire severity using imaging spectroscopy data from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and comparison with multispectral capabilities. *Remote Sens. Environ.* 154, 153-163.
- Vermote, E.F., Kotchenova, S., 2008. Atmospheric correction for the monitoring of land surfaces. *Journal of Geophysical Research: Atmospheres* 113.
- Vermote, E.F., Tanré, D., Deuze, J.L., Herman, M., Morcette, J.-J., 1997. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Trans. Geosci. Remote Sens.* 35, 675-686.
- VITO Remote Sensing. 2017. Available online: https://blog.vito.be/remotesensing/icor_available (Accessed on 27 November 2017).
- Vuolo, F.; Żółtak, M.; Pipitone, C.; Zappa, L.; Wenng, H.; Immitzer, M.; Weiss, M.; Baret, F.; Atzberger, C. 2016. Data service platform for Sentinel-2 surface reflectance and value-added products: System use and examples. *Remote Sens.* 8, 938.
- Wilson, R.T., 2013. Py6S: A Python interface to the 6S radiative transfer model. *Computers and Geosciences* 51, 166-171.
- Yang, X.; Chen, L. 2017. Evaluation of automated urban surface water extraction from Sentinel-2A imagery using different water indices. *J. Appl. Remote Sens.* 11, 026016
- Yengoh, G.T.; Dent, D.; Olsson, L.; Tengberg, A.E.; Tucker III, C.J. 2015. Use of the normalized difference vegetation index (NDVI) to assess land degradation at multiple scales: Current status, future trends, and practical considerations. Springer. ISBN: 978-3-319-24112-8.
- Yilmaz, M.T.; Hunt, E.R.; Jackson, T.J. 2008. Remote sensing of vegetation water content from equivalent water thickness using satellite imagery. *Remote Sens. Environ.* 112, 2514-2522.
- Zanter, K. 2015. Landsat 8 (L8) data users handbook. Department of the Interior US Geological Survey.