Cloud Detection in Hyperspectral Images: Application to PRISMA Images

Maria Francesca Carfora, Italia De Feis, and Diana Carolina Fonnegra Mora

***Abstract*—Hyperspectral sensors provide researchers and governmental authorities with a wealth of information due to their fine spectral resolution, numerous bands, and wide spectral range. These sensors are used in various fields, including agriculture, environmental and forestry monitoring, geology, biology, medicine, and food quality assessment, among others. Generally, they measure across the visible and infrared parts of the electromagnetic spectrum, but they cannot penetrate thick cloud layers, which makes observations unusable under cloudy conditions. Also, the presence of thin and very thin clouds is a problem for the accurate retrieval of surface and atmospheric parameters. The PRecursore IperSpettrale della Missione Applicativa (PRISMA) is a medium-resolution hyperspectral imaging satellite, developed, owned, and operated by Agenzia Spaziale Italiana, launched in orbit on the 22 March 2019. PRISMA carries two sensor instruments, the HYC Hyperspectral Camera module and the panchromatic camera module. In this article, we present the results we obtained by testing some machine learning techniques for cloud detection on Top of Atmosphere (TOA) reflectance data. In particular, we focused on k-nearest neighbors, random forest, and extreme gradient boosting trained on a dataset of manually annotated images by the authors, after transforming the L1 TOA radiance in reflectance data. We also provide numerical comparison with the Cloud detection in hyperspectral images with atmospheric column water vapor method.**

***Index Terms*—Cloud detection, hyperspectral, machine learning, PRecursore IperSpettrale della Missione Applicativa (PRISMA), remote sensing.**

I. Introduction

HYPERSPECTRAL sensors offer unique opportunities for monitoring agriculture and forestry, identifying crops, mapping, disaster management, assessing urban growth, and resource mining because they provide researchers and governmental authorities with a wealth of information due to their fine spectral resolution, numerous bands, and wide spectral range. Their development has significantly advanced our ability to analyze and interpret complex scenes, providing valuable insights across various scientific and industrial domains.

Received 17 July 2025; revised 20 November 2025; accepted 6 January 2026. Date of publication 12 January 2026; date of current version 2 February 2026. This work was supported by the Italian Space Agency (ASI) through SAPP4VU Project under Grant 2022-13-U.0 CUP N. F83C2200055005 and in part by Italian Space Agency through Project "SAPP4VU: Sviluppo di Algoritmi Prototipali Prisma per la Stima del Danno Ambientale e della Vulnerabilità alla Land Degradation" under Grant 2022-13-U.O. and Grant BANDO ASI DC-UOT-2019-061. *(Maria Francesca Carfora, Italia De Feis, and Diana Carolina Fonnegra Mora contributed equally to this work.) (Corresponding author: Diana Carolina Fonnegra Mora.)*

The authors are with the Istituto per le Applicazioni del Calcolo "Mauro Picone," CNR, 00131 Napoli, Italy (e-mail: dianacarolinafonnegramora@cnr.it).

Digital Object Identifier 10.1109/JSTARS.2026.3653197

Hyperspectral images comprise hundreds of narrow contiguous wavelength bands over a wide range of the electromagnetic (EM) spectrum. These hypercubes hence possess a high spectral resolution and a vast amount of embedded information compared to multispectral images. Processing such data enables algorithms to detect and identify targets of interest (e.g., land cover classes) in a hyperspectral scene by revealing subtle differences in their spectral signatures that are not perceived by multispectral sensors. The rich spectral information of hyperspectral data can yield improved accuracy and robustness in characterization, identification, and classification of land covers. However, hyperspectral data also present unique challenges. The fine spectral resolution implies narrow and very contiguous channels, which introduces more noise and high correlation of adjacent bands. Such redundancy may result in duplicated information. Moreover, the spectral signature can exhibit significant spatial variability. When classifying hyperspectral data, the high number of spectral channels and the generally limited availability of labeled training samples lead to the well-known curse of dimensionality and, consequently, increase the risk of overfitting.

The precursor to these hyperspectral instruments is the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [1],[1] which has been capturing spectral images for scientific research and applications every year since 1986. AVIRIS has been the first imaging spectrometer to measure the reflected solar spectrum from 400 to 2500 nm through 224 contiguous spectral channels at 10 nm intervals. AVIRIS was succeeded in 2013 by the Airborne Visible InfraRed Imaging Spectrometer - Next Generation (AVIRIS-NG[2]), which doubled the number of spectral bands to 425 by halving the spectral resolution to 5 nm, covering the range from 380 to 2510 nm.

The first hyperspectral sensor to be deployed in space was Hyperion, on board NASA's Earth Observing-1 (EO-1) satellite.[3] Launched on 21 November 2000, as part of NASA's New Millennium Program, the mission was initially planned as a one-year technology demonstration. Upon successful completion of its objectives, NASA extended the mission, allowing EO-1 to operate until March 30, 2017. Hyperion collected 220 unique spectral channels spanning wavelengths from 400 to 2500 nm, with a 10-nm bandwidth. The instrument achieved a spatial resolution of 30 m across all bands. Each standard scene covered

[1] [Online]. Available: https://aviris.jpl.nasa.gov/aviris/index.html
[2] [Online]. Available: https://avirisng.jpl.nasa.gov/aviris-ng.html
[3] [Online]. Available: https://eospso.nasa.gov/missions/earth-observing-1

an area of 7.7 km × 42 km, with the capability to extend the scene length up to 185 km. Hyperion was succeeded by several other hyperspectral sensors. On 9 May 2018, China launched the Gaofen-5 satellite into orbit, equipped with an Advanced Hyperspectral Imager (AHSI) [2]. The AHSI operates across both the visible–near infrared (VNIR) and shortwave infrared (SWIR) spectral ranges, from 390 to 2510 nm, with a spectral resolution of 5 nm in VNIR and 10 nm in SWIR, totaling 330 bands. It has a spatial resolution of 30 m and a swath width of 60 k. Subsequently, India launched the Hyperspectral Imaging Satellite (HySIS)[4] into orbit on 29 November 2018. HySIS covers a spectral range between 400 and 2400 nm, with 60 bands in the VNIR and 256 bands in the SWIR, a spatial resolution of 30 m, and a swath width of 30 km. These missions were followed by the Italian mission PRecursore IperSpettrale della Missione Applicativa (PRISMA),[5] launched on 22 March 2019. The hyperspectral sensor on PRISMA observes 240 spectral bands across the 400–2500 nm range, with a spectral resolution of less than 12 nm. It has a spatial resolution of 30 m, and each standard scene covers an area of 30 km × 30 km. Similarly, the German Environmental Mapping and Analysis Program (EnMAP)[6] was launched into orbit on 1 April 2022. EnMAP captures 242 spectral bands from 420 to 2450 nm with a spatial resolution of 30 meters, a swath width of 30 kilometers, and a swath length of up to 5550 kilometers per day.

Generally, hyperspectral sensors, which measure across the visible and infrared parts of the EM spectrum, cannot penetrate cloud layers, rendering observations unusable under cloudy conditions. Consequently, cloud detection becomes a fundamental step to ensure the reliability of subsequent analyses and to fully leverage their high spectral and spatial resolution. Moreover, these characteristics offer unique opportunities to enhance cloud detection accuracy and reduce false positives (FP) and negatives. Despite extensive literature on the topic, cloud detection remains challenging due to several factors. Physically, the presence of thin clouds and poor contrast against underlying surfaces complicates identification. Furthermore, algorithms must be adapted to varying sensor hardware—specifically across different spatial, spectral, and temporal resolutions—while accounting for high dimensionality and noise sensitivity in the data [4], [5], [6], [7], [8].

Over the years, various cloud detection methods have been developed, leveraging insights gained from multispectral data. Some methods use a time series of images from the same area to calculate the cloud mask by comparing them with a cloud-free reference image, as done in [7]. In that work, the authors proposed a principal component analysis (PCA)-based fast spatial-spectral random forest (FSSRF) method for cloud removal. FSSRF first applies PCA to known hyperspectral images to produce a series of principal components, and then uses a few of these components to reconstruct the missing data in each cloudy band of the contaminated image. However, this approach requires that each image under study has a temporally close, cloud-free counterpart captured under similar conditions, and assumes that the surface has remained unchanged (e.g., no seasonal variations), which is not always feasible.

In contrast, threshold-based methods are notable for their simplicity and operational efficiency. This is the reason why most of the hyperspectral sensor have their own native cloud/land use mask based on thresholding combinations of bands (single bands, band differences, band ratios, or other combinations) and histograms.

Pioneer of these methods is function of mask (Fmask) [4], [5] proposed first for Landsat imagery and subsequently extended to Sentinel 2. Fmask uses top of atmosphere (TOA) reflectance and brightness temperature as inputs and considers a set of rules based on cloud physical properties to separate potential cloud pixels (PCPs) and clear-sky pixels. Next, a normalized temperature probability, spectral variability probability, and brightness probability are combined to produce a probability mask for clouds over land and water separately. Then, the PCPs and the cloud probability mask are used together to derive the potential cloud layer. Moreover, by using an object-based cloud and cloud shadow matching algorithm, it is capable of providing cloud, cloud shadow, and snow masks for each image. Fmask could also be considered for hyperspectral sensors.

In 2017, Sun et al. [9] presented a cloud detection algorithm-generating (CDAG) method for remote sensing data obtained from AVIRIS. The authors propose to simulate multispectral TOA reflectance data by combining hyperspectral data into broader multispectral pixel datasets using a band combination technique based on the spectral response function of sensors and other parameters. The cloud detection algorithms for the multispectral remote sensing sensor are generated automatically in the next step based on the spectral difference between the clouds and underlying surfaces, and the synthetic cloud detection result is then obtained according to the threshold and accuracy.

Recently, Alakian [10] proposed the Cloud detection in Hyperspectral Images With Atmospheric column WAter vapor (CHIWAWA) method. It is a threshold algorithm that uses criteria based on reflectance to compute a mask of PCPs and refine it by exploiting the water vapor (WV) map. The method differs from other threshold algorithms because it uses ground reflectance data, which, once shooting conditions and atmospheric effects are corrected, is more suitable than TOA reflectance for characterizing surface materials. The use of the WV map is central to the method and plays a role similar to that of thermal bands that are very useful for cloud detection but are absent in the reflective spectral range. Moreover, the author presents two specific procedures to detect snow and ice pixels and reduce the altitude dependence of the WV map in mountain images too. CHIWAWA performance is assessed on PRISMA and AVIRIS-NG and compared with Fmask, CDAG, and native cloud optical thickness (COT) map provided by L2C PRISMA data. The COT map is estimated from a lookup table covering a large number of possible scenarios [11].

Although threshold methods are simple, fast, and easy to implement, they have some limitations. They may be less effective at detecting thin or optically thin clouds and often require manual tuning of thresholds. For these reasons, it may be beneficial to

[4][Online]. Available: https://www.isro.gov.in/HysIS.html [3]
[5][Online]. Available: http://www.prisma-i.it/index.php/en/
[6][Online]. Available: https://www.enmap.org/

consider more advanced methods based on machine learning, deep neural networks (DNNs), and convolutional neural networks (CNNs). Machine learning methods have demonstrated their potential in the Cloud Masking Intercomparison eXercise (CMIX) on Landsat 8 (NASA/USGS) and Sentinel-2 (ESA) missions [12], where ten algorithms, developed by nine teams from 14 different organizations—including universities, research centers, industry, and space agencies (CNES, ESA, DLR, and NASA)—were evaluated across five benchmark datasets. Among the ten algorithms, three were machine learning-based, and in particular, s2cloudless, a gradient boosting algorithm, showed very good performance.

Deep learning approaches, particularly CNNs, have also shown strong performance on a variety of multispectral sensors, including QuickBird, Sentinel-2, Landsats 7 and 8, Gaofen-1 and 2, and ZY-3. For instance, Xie et al. [13] proposed a CNN model for cloud detection, Yang et al. [14] tested architectures from fully convolutional networks to encoder–decoder models, and recent studies, such as [15] and [16], introduced attention-based networks and hybrid CNN-Transformer architectures, demonstrating their effectiveness in capturing both spatial and spectral patterns for accurate cloud masking.

More recent works have explored the use of deep neural networks for cloud detection in hyperspectral images, also considering their potential for on-board processing [8], [17], [18], [19], [20]. Advancements in lightweight deep learning models, together with dedicated hardware, now make it feasible to perform cloud masking directly on satellites. In particular, the study presented in [17] focused on HyperScout-2, which acquires 50 bands in the VNIR range (400–1000 nm) and four thermal bands in the 8–14 μm range, and studies in [8] and [19] were applied to HYPSO-1, with 120 bands spanning 430–800 nm. These works represent initial efforts to adapt deep learning approaches, previously developed for multispectral sensors, to the high-dimensional and spectrally rich data provided by hyperspectral instruments.

Building on initial hyperspectral efforts, such as HyperScout-2 and HYPSO-1, which demonstrated the potential of deep neural networks for cloud detection in high-dimensional spectral data, similar approaches have been explored for Tropospheric Monitoring Instrument (TROPOMI) on Sentinel-5 Precursor. TROPOMI, while not a traditional imager, provides very high spectral resolution that enables advanced cloud detection. Indeed, its spectral richness permits advanced cloud retrieval and characterization, and recent studies have leveraged neural-network-based methods for this purpose, including the operational TROPOMI algorithms [21] and the CRANN method [22], [23], which incorporates a deep neural network model with radiative transfer model to retrieve cloud fraction and cloud-top pressure from the oxygen–oxygen collision-induced (O_4) absorption band. These examples illustrate how deep learning approaches developed for hyperspectral sensors can be adapted to instruments with thousands of spectral channels, extending the potential for accurate cloud analysis across diverse platforms.

This study focuses on cloud detection in PRISMA TOA reflectance data, presenting and summarizing the results obtained by the authors within the SAPP4VU project, funded by the Italian Space Agency [Agenzia Spaziale Italiana (ASI)]. The project aimed to develop prototype algorithms that leverage PRISMA hyperspectral data to estimate environmental damage and vulnerability to land degradation. The authors contributed by characterizing PRISMA noise [24], [25] and adapting a robust cloud detection procedure specifically for PRISMA images. The Python code implementing these algorithms has been delivered to ASI and is also available on GitHub upon request.

Several assessment studies [26], [27], [28] have evaluated the performance of the PRISMA hyperspectral instrument across various land cover types—such as rural areas, inland and coastal waters, and urban environments—confirming that the availability of contiguous spectral information is particularly promising. PRISMA data have already contributed to a range of applications, including the retrieval of multiple vegetation traits [29], [30], mapping of water quality parameters [31], [32], and characterization of surface material composition across different landscapes [33]. Other notable applications include fire severity assessment [34], early-season crop mapping [35], and estimation of atmospheric CO_2 content [36].

However, to the best of the authors' knowledge, this is the first study—alongside the recent work by Alakian [10]—to address the problem of cloud detection specifically for PRISMA data. Moreover, machine learning methods have not yet been applied or systematically evaluated in this context, primarily due to the absence of a benchmark, manually annotated dataset.

In this work, we present a manually annotated dataset comprising samples from various spectral categories, including urban areas, wetlands, forests, bare soil, water bodies, agricultural land, snow, and clouds. The samples were collected from multiple sites across Europe and northern Africa—particularly the Sahara Desert—to ensure comprehensive representation of all categories. We evaluate the performance of three machine learning algorithms: k-nearest neighbors (kNN), random forest (RF), and extreme gradient boosting (XGBoost). In addition, we provide a numerical comparison with the CHIWAWA cloud detection method and the native PRISMA land use mask provided by ASI alongside the Level-1 PRISMA images [11].

The rest of this article is organized as as follows. Section II presents the dataset, the preprocessing steps, the construction of the training and validation sets, and the band selection process. Sections III and IV describe the methods used, along with the evaluation metrics. Section V is devoted to the numerical experiments. Section VI discusses the results, and finally, Section VII concludes this article.

II. PRISMA Data

PRISMA is an Earth observation satellite mission funded by the Italian Space Agency (ASI), aiming at acquiring and delivering data for multiple remote sensing applications, such as resource management, environmental monitoring, and disaster management [37]. PRISMA payload includes a high-spectral resolution imaging spectrometer operating in the VNIR, and SWIR, and a medium-resolution panchromatic (PAN) camera. The satellite is flying on a sun-synchronous low Earth orbit at an altitude of 615 km; the nominal orbit repeat cycle is 29

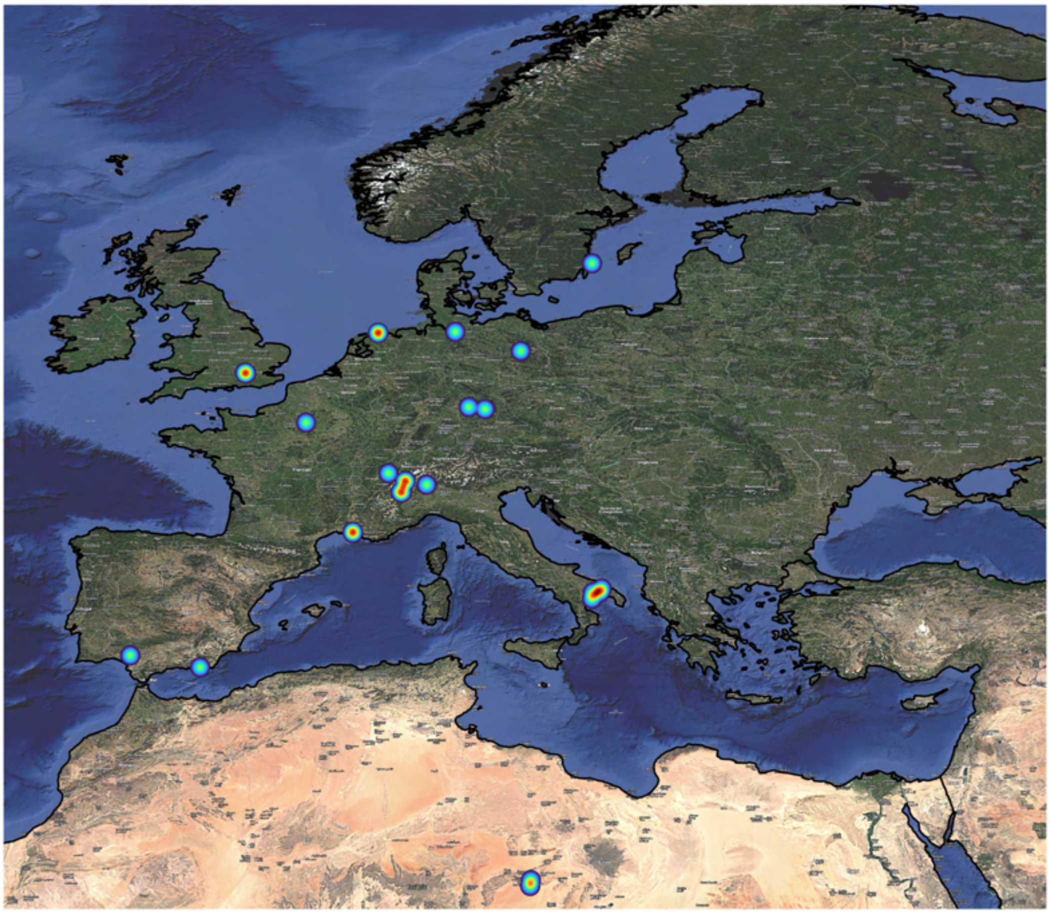

Fig. 1. PRISMA data coverage area. Blue spots indicate the acquisition of a single image, while red spots indicate multiple images of the same site.

days (from nadir) with a revisit capability of seven days with a roll maneuver. The geographic coverage is nominally between 70°S - 70°N latitude and 180°W - 180°E longitude, but imagery can be acquired at higher latitudes with the only limitation related to typically low solar zenith angle (SZA). The standard size of a single image is 30 km × 30 km with a ground sampling distance of 30 m (VNIR-SWIR) and 5 m (PAN). The imaging spectrometer covers the nominal 400–2500 nm spectral range with two separated instruments: the VNIR spectrometer features 66 spectral bands from about 400–1010 nm, while the SWIR detector provides 174 spectral bands between 920 and 2500 nm; both instruments have a bandwidth lower than 15 nm. The PAN camera provides coregistered 5 m spatial resolution imagery, useful for a better interpretation of the imaging spectrometer data.

The PRISMA ground processor systematically generates the following data levels.

1) *Level-0 (L0):* Raw data stream, including metadata, ancillary data, and formatting information.
2) *Level-1 (L1):* TOA radiometrically and geometrically calibrated radiance images.
3) *Level-2B (L2B):* Coregistered at-surface radiance in sensor geometry, with an attached geocoding model.
4) *Level-2 C (L2C):* Coregistered surface reflectance in sensor geometry, with an attached geocoding model.
5) *Level-2D (L2D):* Geolocated surface reflectance.

This study explores Level-1 (L1) PRISMA imagery. The hyperspectral L1 data consist of calibrated and coregistered hypercubes of TOA radiance in sensor geometry. This product is generated on demand and delivered in Hierarchical Data Format [11].

In addition, the product includes an error mask that identifies unreliable or less reliable pixels—such as those that are defective, saturated, radiometrically inaccurate, or contain NaN or Inf values at the end of processing. It also provides three thematic data masks that indicate pixels affected by clouds, sunglint, and a Level-1 land cover classification of the image, respectively.

We selected 26 L1 PRISMA images with varying cloud cover conditions—ranging from heavily clouded to cloud-free—to build the training and validation datasets. The dataset includes samples from various spectral categories, such as urban areas, wetlands, forests, bare soil, waterbodies, agricultural land, snow, and clouds. These samples were collected from multiple sites across Europe and northern Africa, particularly the Sahara Desert, to ensure comprehensive representation of all categories.

Fig. 1 illustrates the spatial distribution of the selected areas: blue markers indicate locations where a single image was acquired, while red markers denote sites with multiple image acquisitions. The selected images span the late spring and summer seasons from 2021 to 2023.

TABLE I
PRISMA Images Used for Training: Id, Filename, Location, and Coordinates of the Image Center

Id	File	Place	Center
T1	PRS_L1_STD_OFFL_20210722104421_20210722104425	France	**Lat** 43°28′30.63” **Lon** 4°54′33.62”
T2	PRS_L1_STD_OFFL_20210730094853_20210730094857	Lybia	**Lat** 26°7′33.09” **Lon** 13°48′20.45”
T3	PRS_L1_STD_OFFL_20210730094857_20210730094902	Lybia	**Lat** 25°51′40.23” **Lon** 13°44′35.95”
T4	PRS_L1_STD_OFFL_20210808103703_20210808103707	Switzerland	**Lat** 46°4′3.72” **Lon** 7°28′10.45”
T5	PRS_L1_STD_OFFL_20210816093833_20210816093837	Italy	**Lat** 40°30′5.99” **Lon** 17°12′46.61”
T6	PRS_L1_STD_OFFL_20210819102715_20210819102719	Switzerland	**Lat** 46°3′15.09” **Lon** 7°30′55.92”
T7	PRS_L1_STD_OFFL_20210819102723_20210819102727	Italy	**Lat** 45°31′41.76” **Lon** 7°21′8.59”
T8	PRS_L1_STD_OFFL_20210827110636_20210827110640	Spain	**Lat** 36°45′44.47” **Lon** −2°45′23.01”
T9	PRS_L1_STD_OFFL_20220623103830_20220623103834	Netherlands	**Lat** 53°29′16.81” **Lon** 6°10′7.36”
T10	PRS_L1_STD_OFFL_20220703101206_20220703101211	Germany	**Lat** 52°34′25.40” **Lon** 13°17′2.68”
T11	PRS_L1_STD_OFFL_20220716103518_20220716103522	Germany	**Lat** 53°32′54.06” **Lon** 9°59′45.49”
T12	PRS_L1_STD_OFFL_20220722104028_20220722104032	Switzerland	**Lat** 46°29′46.12” **Lon** 6°40′7.70”
T13	PRS_L1_STD_OFFL_20220803104755_20220803104800	France	**Lat** 43°33′53.86” **Lon** 4°53′31.82”
T14	PRS_L1_STD_OFFL_20220811112632_20220811112637	Spain	**Lat** 37°22′26.61” **Lon** −6°14′28.12”
T15	PRS_L1_STD_OFFL_20230707111230_20230707111234	U.K	**Lat** 51°29′4.40” **Lon** −0°29′25.19”
T16	PRS_L1_STD_OFFL_20230725094837_20230725094841	Italy	**Lat** 40°30′23.30” **Lon** 16°54′23.32”
T17	PRS_L1_STD_OFFL_20230820103350_20230820103354	Italy	**Lat** 45°54′30.47” **Lon** 8°34′34.44”
T18	PRS_L1_STD_OFFL_20230827105158_20230827105202	Netherlands	**Lat** 53°29′13.01” **Lon** 6°10′3.52”
T19	PRS_L1_STD_OFFL_20230831102255_20230831102259	Germany	**Lat** 49°40′12.68” **Lon** 11°29′14.76”
T20	PRS_L1_STD_OFFL_20230902105636_20230902105640	France	**Lat** 49°1′3.14” **Lon** 2°32′17”
T21	PRS_L1_STD_OFFL_20230906102611_20230906102616	Germany	**Lat** 49°47′1.23” **Lon** 10°4′48.90”

The dataset was divided into two subsets—one for training (see Table I) and one for model validation (see Table II).

A. Preprocessing

L1 data are released as digital number (DN) values. After excluding erroneous and defective pixels—using the error and sunglint masks provided as auxiliary data with each image, we converted the DN values to radiance according to the method described in [11]. This conversion was performed separately for each spectral region (VNIR and SWIR) using the following equation:

$$L_\lambda = \frac{\mathrm{DN}_\lambda}{\mathrm{scalefactor}} - \mathrm{offset} \tag{1}$$

where L_λ is the spectral radiance at the sensor aperture, expressed in $\mathrm{mW}\cdot\mathrm{cm}^{-2}\cdot\mathrm{sr}^{-1}\cdot\mu\mathrm{m}^{-1}$, DN_λ is the raw DN for each band, and scalefactor and offset are constants provided in the image metadata for each spectral region.

Next, the radiance values were converted to reflectance as described in [11]

$$\rho_\lambda = \frac{\pi L_\lambda d^2}{E_{\mathrm{SUN}} \cos\theta_s} \tag{2}$$

where ρ_λ is the reflectance for each band, d is the Earth–Sun distance in astronomical units (with $d^2 \approx 0.9663$ for the considered period), E_{SUN} is the spectral extra-atmospheric solar irradiance in mW $\cdot\mathrm{cm}^{-2}\cdot\mu\mathrm{m}^{-1}$, and θ_s is the SZA in degrees. The SZA θ_s is calculated from the latitude and longitude of each pixel and the UTC timestamp of the corresponding frame. The value

TABLE II
PRISMA IMAGES USED FOR VALIDATION: ID, FILENAME, LOCATION, AND COORDINATES OF THE IMAGE CENTER

Id	File	Place	Center
V1	PRS_L1_STD_OFFL_20210707094808_20210707094812	Italy	**Lat** $40°42'31.15$” **Lon** $17°19'57.18$”
V2	PRS_L1_STD_OFFL_20220627100725_20220627100729	Sweden	**Lat** $56°58'31.67$” **Lon** $16°50'30.95$”
V3	PRS_L1_STD_OFFL_20210808103711_20210808103716	Italy/France	**Lat** $45°32'43.67$” **Lon** $7°16'32.51$”
V4	PRS_L1_STD_OFFL_20230817094518_20230817094523	Italy	**Lat** $40°11'50.69$” **Lon** $16°49'52.27$”
V5	PRS_L1_STD_OFFL_20230817111911_20230817111916	U.K	**Lat** $51°28'28.84$” **Lon** $-0°27'9.86$”

TABLE III
CLASSES DISTRIBUTION, CLOUDY/NOT CLOUDY, AND LAND COVER

Multiclass Label	Binary Label	Class	Unbalanced db (px)	Balanced db (px)
1	1	Cloud	149 0621	978 537
2	0	Bare soil	795 044	139 791
3	0	Urban	232 382	139 791
4	0	Wetlands	153 868	139 791
5	0	Agriculture	412 488	139 791
6	0	Forest	214 722	139 791
7	0	Waterbodies	447 770	139 791
8	0	Snow	139 791	139 791

of E_{SUN} is derived from the Thuillier solar irradiance model, following the methodology described in [11], [28].

B. Hyperspectral Training Dataset

To build reliable information for training the classification algorithms, we performed manual annotation on the selected images. This process resulted in a vector file associated with each image in the training dataset; see Table I.

For each image, we identified and labeled the land cover classes proposed in [11], see Table III, by drawing polygons corresponding to the various classes. The annotation was carried out through visual inspection and manual sampling. Despite being time-consuming, manual sampling remains the most effective method for accurate land cover labeling [38].

We gave special attention to the construction of the dataset to avoid factors that could influence cloud recognition and, consequently, affect algorithm performance. To this end, our pixel-based dataset incorporated not only spectral information but also spatial variability within the same class. For example, the bare soil class included samples from diverse environments such as the Sahara Desert, coastal zones in Ameland, and mountainous areas with snow in Bagnes. Similarly, anthropogenically modified regions—such as greenhouse areas in Spain—were also considered to enhance class representation and improve model generalization.

We used QGIS software to visually interpret land cover types. For this task, we extracted a subset of bands from each training image to create false-color composites. These composites helped us highlight the radiometric characteristics of the target classes and accurately delineate them using polygons. Figs. 2 and 3 show a sample of the selected areas, grouped into eight main categories: Agriculture, Bare soil, Forest, Urban, Snow, Waterbodies, Wetlands, and Clouds.

We observed marked heterogeneity in the spectral responses of several classes. For instance, the Clouds category showed significant variability in spectral signatures depending on cloud thickness, shape, and the underlying surface [9]. This is shown in Figs. 2 and 3, where image B—a fragmented cloud—shows a lower reflectance than case A—a thick cloud with a higher reflectance.

Another interesting case is the Urban category: particularly in Image C, representing greenhouses, the spectral signature shows a higher reflectance very similar to clouds and a marked spectral difference with pixels extracted from other urbanized areas. Also in the Wetlands, the saline (Image C) shows a higher reflectance in the visible than its counterparts.

These findings highlight the importance of carefully representing and balancing the spectral variability within each class during the construction of the training dataset, especially for classes with high internal variability or spectral similarity to clouds.

We overlaid the associated vector file onto each image in the training dataset and extracted pixels based on the labeled polygons. We assigned labels according to the corresponding class values and stored all extracted points in a single table to build a preliminary multiclass database, where each column represents a feature (e.g., spectral value for a given band) or the associated class. To ensure balanced representation of land cover types under clear-sky conditions, we adjusted the dataset by equalizing the number of pixels across classes 2–8, all corresponding to the binary label noncloudy: 0. We then selected an equal number of pixels for class 1, labeled as cloudy: 1. Table III presents the final classes distribution.

C. Band Selection

We reduced the number of bands by selecting only the channels whose central wavelengths were closest to those listed in Table IV. We used these channels as a reference because they match those employed by PRISMA’s *land-use mask generator*. In addition, we included two WV absorption bands—at 942 and

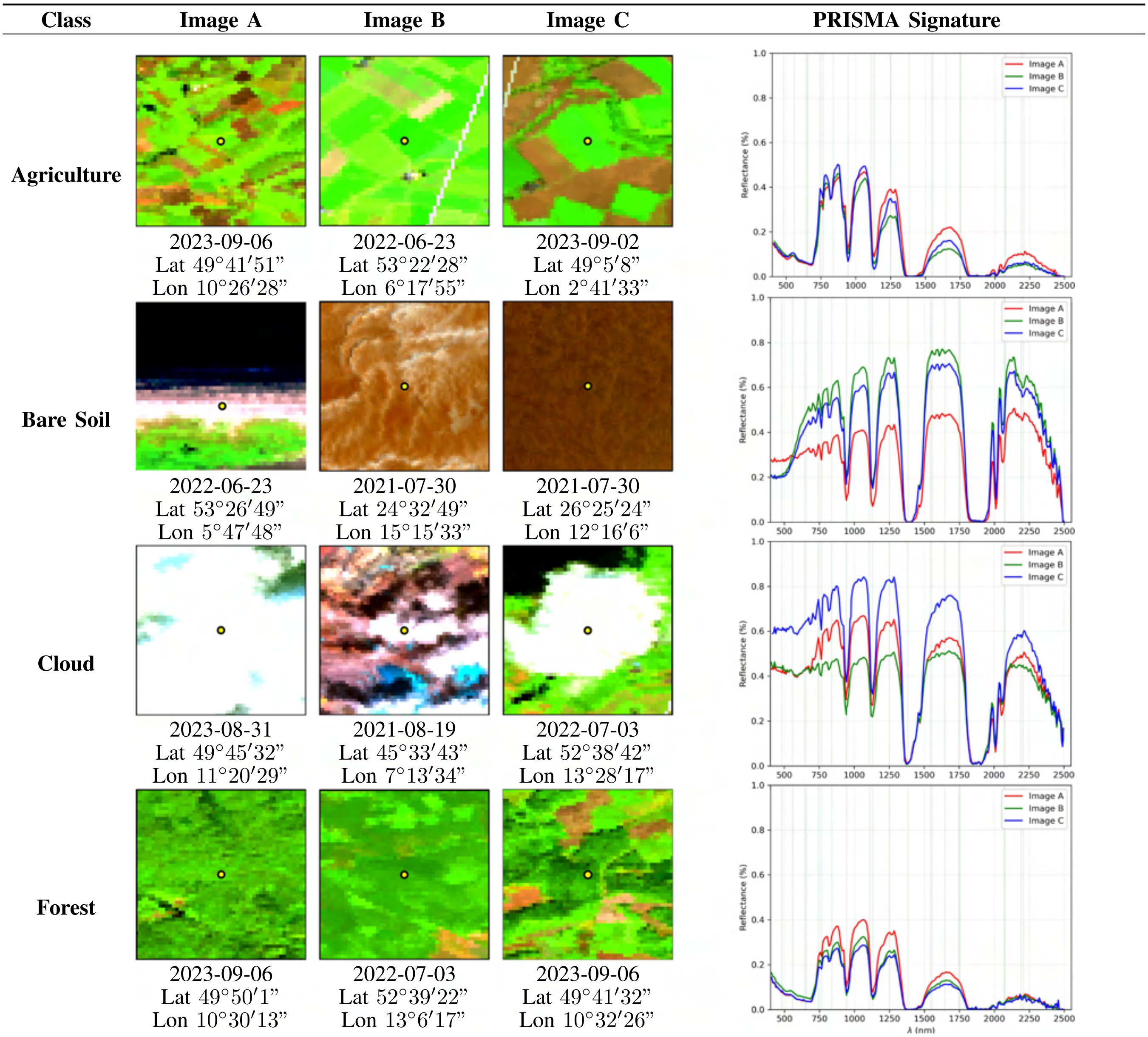

Fig. 2. Image samples from the training dataset, grouped by land cover classes (left panel) and the corresponding spectral signature for the central pixel of each image (right panel). Selected channels for false color composition on PRISMA: *Red:* 1647 nm, *Green:* 838 nm, and *Blue:* 482 nm.

TABLE IV
SELECTED BANDS

Region	Bands [nm]
VNIR	480, 559, 650, 660, 742, 762, 840, 942
SWIR	1114, 1131, 1250, 1386, 1548, 1558, 1651, 1749, 2072, 2082, 2193, 2203

1131 nm—due to their well-established importance in detecting atmospheric water content [39].

We also explored an alternative strategy for band selection by applying PCA to two versions of the training dataset: the full dataset containing all 234 spectral bands, and a filtered version with 202 bands, where we removed those located in atmospheric absorption regions. In both cases, we retained the first ten principal components and applied the algorithms described in Section III. However, the results on the validation dataset were unsatisfactory, showing a high number of false negatives (FN) and FP. Due to this poor performance, we opted to use the more reliable, physics-based band selection approach described in Table IV.

D. Validation Dataset

We built the validation dataset (see Table II) by selecting images with characteristics similar to those in the training set. Specifically, the images varied in cloud-cover percentage, acquisition time (between 2021 and 2023), and terrain morphology. We also prioritized scenes with a high number of pixels belonging to spectrally bright classes that typically challenge

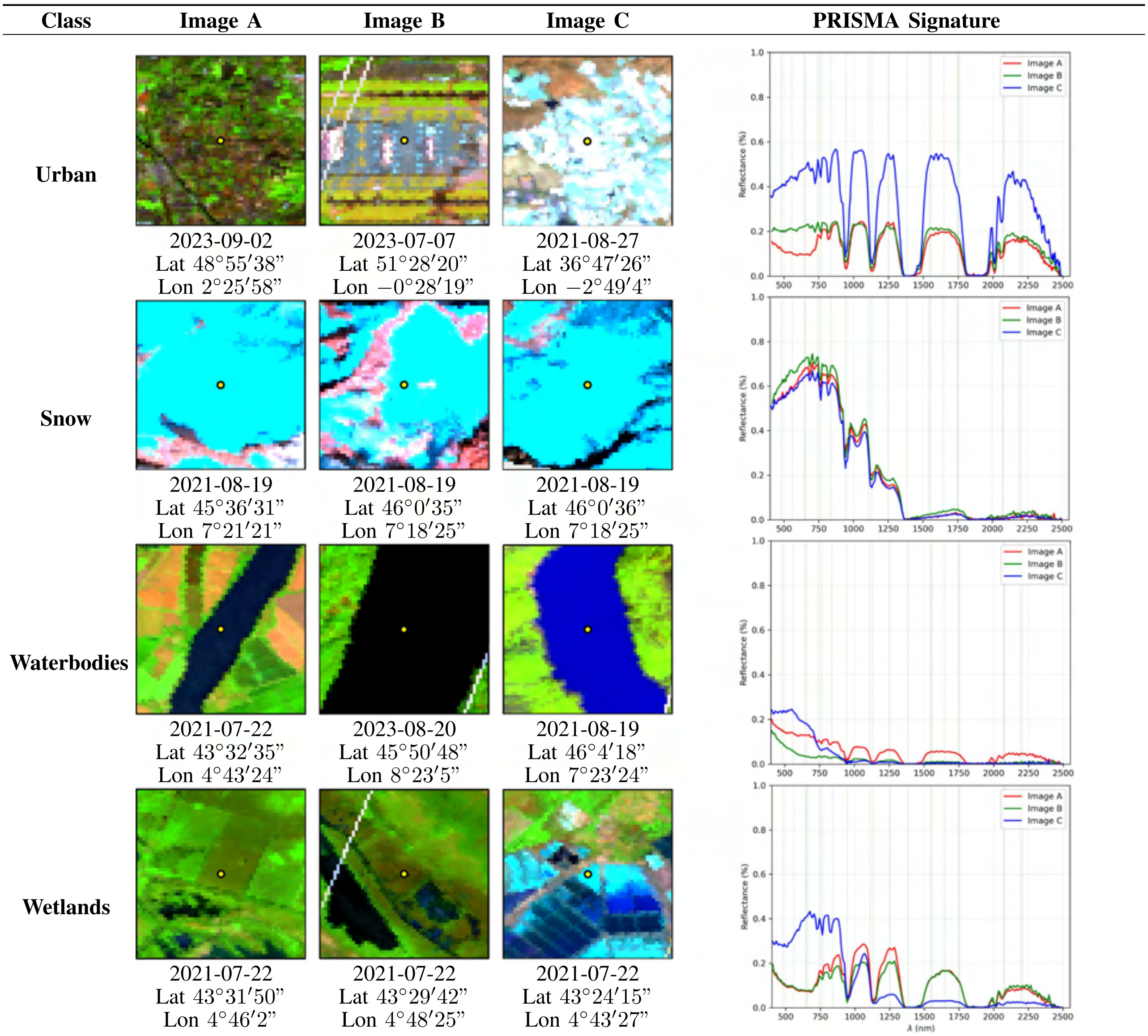

Fig. 3. Image samples from the training dataset, grouped by land cover classes (left panel) and the corresponding spectral signature for the central pixel of each image (right panel). Selected channels for false color composition on PRISMA: *Red:* 1647 nm, *Green:* 838 nm, and *Blue:* 482 nm.

cloud-detection algorithms, such as snow-covered areas, bare soil, and urban zones. Each image was manually annotated to construct the dataset. As with the training database, we compiled the extracted pixels into a single table. In contrast to the training dataset, this set contains only two labels: *Cloud* and *No cloud*. It is important to mention that cloud boundaries are difficult to represent, and this has an important role in the performance metrics [10].

The land cover annotations for both the training and validation datasets have been made available in a public repository[7] while the related PRISMA images can be requested and freely downloaded through the official PRISMA portal after registration.

[7][Online]. Available: https://github.com/argennof/CloudDetectionPRISMA

III. Methods

We implemented three different algorithms for cloud detection on PRISMA data: the classical kNN and two ensemble learning methods as RF and XGBoost.

Let us indicate the training dataset as $(\mathbf{X_i}, Y_i)$, $i = 1, \ldots, N$, with $\mathbf{X_i} \in \mathbb{R}^d$ vector of features and Y_i its associate label. Indicating with $C_1, \ldots, C_K$ the K classes, Y_i assumes values in $\{1, \ldots, K\}$, i.e., $Y(\mathbf{X}_i) = j$ if $\mathbf{X}_i \in C_i$.

- kNN is a supervised machine learning technique, which classifies each given point by assigning it the label associated with the majority of its closest k neighbors. This method relies on the chosen distance function between samples, and common choices are the Euclidean, Minkowsky, or Manhattan distance. The algorithm was introduced in [40] and later expanded by [41]. Given an unlabeled point $\mathbf{X} \in \mathbb{R}^d$, we look at the k training points $\mathbf{X_r}, r = 1, ..., k$, that are closest in distance to $\mathbf{X}$. Then, we assign to $\mathbf{X}$ the class label $\hat{Y}(\mathbf{X})$ that is most often represented among these k neighbors (the *mode* of this discrete distribution)

$$\hat{Y}(\mathbf{X}) = \text{mode}(Y_1, \ldots, Y_k). \tag{3}$$

kNN is very simple, intuitive, and has no training time. It is often successful in some classification problems, when the decision boundary is very irregular and the dataset is small. But it is also very sensitive to the number of features d, because the more dimensions we add, the larger the search volume needs to be to capture k neighbors (curse of dimensionality), and then, these neighbors may not be particularly similar to $\mathbf{X}$ anymore.

Ensemble methods combine multiple simple models, or *weak learners*, to obtain a single, powerful model. RF [42] and gradient-boosted trees (GBT) [43] are both ensemble methods; they combine multiple decision trees, such as decision trees or classification and regression trees, to improve predictive performance.

- In RF, a number of trees are created through bootstrap aggregation, or *bagging* [42] and random feature selection. Bagging is a procedure for reducing the variance of a statistical learning method: by averaging many noisy but approximately unbiased models, we obtain a single low-variance prediction model. Trees are ideal candidates for bagging, since they can capture complex interaction structures in the data, and if grown sufficiently deep, have relatively low bias. Each tree is independently trained on a bootstrapped subset of the training data, with a series of splitting rules applied to build the tree, using a specific impurity measure (such as Gini impurity or entropy) to evaluate the split. More precisely, in a node m representing a splitting region R_m with N_m observations, let

$$\hat{p}_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(Y_i = k)$$

the proportion of class k observations in node m, for $k = 1, \ldots, K$. We classify the observations in node m to class $k(m) = \arg\max_k \hat{p}_{mk}$, the majority class in node m. To evaluate the goodness of the split, different measures are possible and include the following:

$$\text{Gini index} : \sum_1^K \hat{p}_{mk}(1 - \hat{p}_{mk})$$

$$\text{Cross-entropy} : \sum_1^K \hat{p}_{mk} \log \hat{p}_{mk}.$$

Now, RF improves over bagged trees by considering, instead of all the d features, just a random subset of size p while building each split in a tree. This random feature selection enhances decorrelation between trees. After training, predictions for unseen samples $\mathbf{X}$ are made by aggregating the predictions from all individual trees in a majority vote. See [44] (chapter 15 and Algorithm 15.1) for more details.

The bagging process can be summarized as follows.

1) For $b = 1, \ldots, B$ (where B is the total number of trees), the following holds.
   a) Sample, with replacement, k training examples from $(\mathbf{X_i}, Y_i)$ $i = 1, \ldots, n$; call these $(\mathbf{X}_b, \mathbf{Y}_b)$, where $\mathbf{X}_b \in \mathbb{R}^{k \times d}$ and $\mathbf{Y}_b \in \mathbb{R}^{k \times 1}$.
   b) Train a classification tree T_b on $(\mathbf{X}_b, \mathbf{Y}_b)$ by recursively repeating the following steps for each terminal node of the tree, until the minimum node size is reached.
      i. Select p variables at random from the d variables.
      ii. Pick the best variable/split-point among the p.
      iii. Split the node into two daughter nodes.
   c) Output the ensemble of trees $\{T_b\}_1^B$.
2) After training, predictions for an unseen sample $\mathbf{X}$ is made by aggregating the predictions from all individual trees, i.e., let $\hat{C}_b(\mathbf{X})$ be the class prediction of the bth RF tree. Then

$$\hat{C}_{RF}^B(\mathbf{X}) = \text{majority vote}\left\{\hat{C}_b(\mathbf{X})\right\}_1^B.$$

By aggregating M decorrelated trees, RF effectively reduces variance and produces a robust final estimate.

- GBT, on the other hand, uses boosting [43] to aggregate trees. The key difference is that in bagging, each tree is built independently, whereas in boosting, each new tree is added sequentially, specifically designed to correct the errors made by previous trees. In this case, the loss function is the multinomial deviance

$$L(Y, p(\mathbf{X})) = -\sum_{k=1}^K \mathbb{I}(Y = k) \log p_k(\mathbf{X})$$

where $\mathbb{I}(\cdot)$ is the indicator function, $\mathbf{X} \in \mathbb{R}^{d \times 1}$ is the generic feature vector, and $p_k(\mathbf{X})$ denotes the posterior class conditional probability

$$p_k(\mathbf{X}) = \Pr(Y = k|\mathbf{X}) = \Pr(C_k|\mathbf{X}) = \frac{e^{f_k(\mathbf{X})}}{\sum_{l=1}^K e^{f_l(\mathbf{X})}}. \tag{4}$$

Once the estimates $\hat{f}_k, k = 1, \ldots, K$, are obtained from a sample of N data, for a new sample $\mathbf{X}$, we can either evaluate the probability according to (4) or classify according to Bayes' Rule for Classification

$$\text{pr}(\text{C}_\text{k}|\mathbf{X}) = \max_{j=1,\ldots,K} \Pr(C_j|\mathbf{X})$$

where we classify to the most probable class using the estimated posterior conditional distribution of the classes. The original GBT algorithm proposed by Friedman [43], as described in [44], can be written as follows.

1) Initialize $f_0(\mathbf{X}) = \arg\min_\gamma \sum_{i=1}^n L(Y_i, \gamma(\mathbf{X}_i))$, where L is the loss function.
2) For $m = 1$ to M (for each boosting iteration), the following holds.
   – For $k = 1$ to K (for each class), the following holds.

a) For $i = 1, \ldots, n$, compute the pseudoresiduals

$$g_{ik}^{(m)} = -\frac{\partial L(Y_i, f_{m-1}(\mathbf{X}_i))}{\partial f_{m-1}(\mathbf{X}_i, k)} = \mathbb{I}(Y_i = k) - p_k(\mathbf{X}_i)^{(m-1)}$$

where $p_k(\mathbf{X}_i)^{(m-1)}$ is the predicted probability of class k for the ith observation in the previous iteration.

b) Fit a least square tree to the pseudoresiduals $g_{ik}^{(m)}$, yielding terminal regions $R_{jk}^{(m)}, j = 1, \ldots, J_k^{(m)}$, where $J_k^{(m)}$ is the number of terminal nodes (leafs) for tree k at iteration m.

c) For each terminal region $R_{jk}^{(m)}$, compute the best step size $\gamma_{jk}^{(m)}$ to minimize the loss

$$\gamma_{jk}^{(m)} = \arg\min_{\gamma} \sum_{\mathbf{X}_i \in R_{jk}^{(m)}} \times L\left(Y_i, f_{m-1}(\mathbf{X}_i, k) + \gamma\right).$$

This step adjusts the value of f within each region $R_{jk}^{(m)}$ to minimize the loss function.

d) Update the model by adding the contribution of the new tree

$$f_m(\mathbf{X}, k) = f_{m-1}(\mathbf{X}, k) + \eta \sum_{j=1}^{J_k^{(m)}} \gamma_{jk}^{(m)} \mathbb{I}\left(\mathbf{X} \in R_{jk}^{(m)}\right)$$

where η is the learning rate, and $\mathbb{I}(\mathbf{X} \in R_{jk}^{(m)})$ is an indicator function that is 1 if $\mathbf{X}$ falls in region $R_{jk}^{(m)}$, and 0, otherwise.

3) Output the final model: $\hat{f}_M(\mathbf{X}, k), \quad k = 1, \ldots, K$, where K is the number of classes, and M is the number of boosting iterations.

This iterative approach enables GBT to focus on improving the predictions where prior models underperformed, often resulting in enhanced accuracy.

XGBoost is a scalable, distributed gradient-boosted decision tree machine learning library developed by the authors in [45] and [46]. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems. Unlike gradient boosting, which works as gradient descent in function space, XGBoost works as Newton–Raphson: a second-order Taylor approximation is used in the loss function.

All three methods have a number of parameters that need to be tuned. Optuna [47] is an automatic hyperparameter optimization open-source framework specifically designed for machine learning. This framework offers high modularity, allowing users to dynamically build the search space for hyperparameters. It combines efficient search and pruning algorithms to enhance the cost-effectiveness of optimization. Our experiment includes a Bayesian optimization technique as a search criteria strategy, fivefold cross-validation, and Successive Halving Pruner as a pruner.

Fig. 4. Color scale for the classification mask.

IV. Evaluation Metrics

We applied both qualitative visual inspection and quantitative criteria to evaluate the performance of the considered methods. Regarding the quantitative metrics, and assuming the hypothesis

$$H_0 : \text{pixel cloudy against the alternative} H_1 : \text{pixel not cloudy}$$

we defined the quantities as follows:

- *true positives (TP):* pixels correctly identified by the method as cloudy;
- *true negatives (TN):* pixels correctly labeled by the method as noncloudy;
- *FP:* noncloudy pixels in the ground-truth data but labeled as cloudy by the method;
- *FN:* cloudy pixels in the ground-truth data but labeled by the method as clear sky.

We then considered the following indicators for each image:

$$\text{OA} = 100 \times \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (5)$$

$$\text{P} = 100 \times \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Sen} = 100 \times \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

$$\text{TNR} = 100 \times \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (8)$$

$$\text{BOA} = 0.5 \times (\text{Sen} + \text{TNR}) \quad (9)$$

where OA refers to the overall accuracy, P is the precision/positive predicted value, Sen is the sensitivity/recall/TP rate, TNR is the TN rate, and BOA the balanced overall accuracy. Fig. 4 shows the color scale used to represent TP, TN, FP, and FN values in the classification mask. Moreover, high Sen values show that cloud pixels have been effectively masked out, while high P values indicate that the algorithm is cautious, excluding noncloudy pixels. High BOA values are related to a good balance of FP and FN [48].

Following [10], we also considered the Cloud Cover percentage CC%, defined as the ratio between the number of pixels detected as cloudy and the total number of pixels in the image expressed in percentage. This allows us to construct a metric that quantifies the difference (ΔCC) between the estimated cloud cover (CC%_est) and the ground truth cloud cover (CC%_ref), i.e.,

$$\text{CC\%\_ref} = 100 \times \frac{(\text{TP} + \text{FN})}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (10)$$

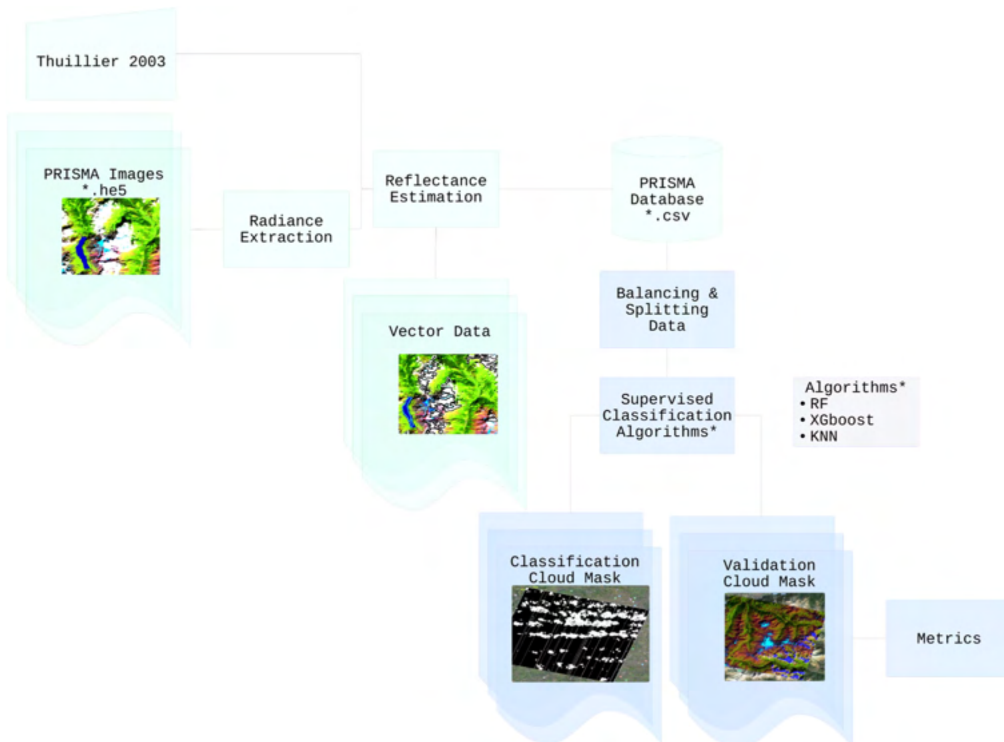

Fig. 5. Workflow of the proposed methodology.

$$\mathrm{CC\%\_est} = 100 \times \frac{(\mathrm{TP}+\mathrm{FP})}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{11}$$

$$\begin{aligned}\Delta\mathrm{CC} &= |\mathrm{CC\%\_est} - \mathrm{CC\%\_ref}| \\ &= \frac{|\mathrm{FP}-\mathrm{FN}|}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}.\end{aligned} \tag{12}$$

Fig. 5 summarizes the flowchart of the methodology we proposed.

V. Experimental Results

In this section, we show the performance of kNN, RF, and XGBoost, whose hyperparameters were tuned using Optuna. Table V presents the best selected hyperparameters.

We compared the performance of the three machine learning methods with the native PRISMA land use mask provided by ASI along with L1 images [11] and with the results obtained by the CHIWAWA algorithm [10], as provided by Alakian. The native PRISMA land use mask is a hierarchical threshold classifier based on surface reflectance data. It classifies the signal into certain spectral categories, i.e., cloud, sunglint, water, snow, not-vegetated land, urban, vegetated land (crop and rangeland, forests), wetlands, using 7 VNIR bands and 12 SWIR bands and their combination. CHIWAWA [10] is an unsupervised cloud detection algorithm specifically tailored to hyperspectral images. It combines the complementary information available in the ground reflectance image and the atmospheric column WV map. The method uses threshold criteria based on cloud physical properties to compute a mask of PCPs and refine it by exploiting the WV information. For computations, we used a

TABLE V
Hyperparameters Selected by OPTUNA, Search Criteria Strategy: Bayesian Optimization and Fivefold Cross-Validation

Classifier	Params and Intervals	Selected value
kNN	n_neighbors: [5, 10]	5
	weights: [uniform, distance]	Distance
	algorithm [ball_tree]	ball_tree
RF	criterion [gini]	gini
	n_estimators: [8, 60]	43
	max_depth: [23, 27]	24
	max_features: [3, 7]	3
	min_samples_split: [4, 8]	4
XGBOOST	objetive [binary:logistic]	binary:logistic
	tree_method: [hist, approx]	approx
	max_depth: [3, 12]	11
	min_child_weight: [1,2 50]	1
	Learning_rate [0.3]	0.3
	subsample: [0.7, 1]	0.87
	colsample_bytree: [0.5, 1]	0.57
	reg_lambda [0.01, 10]	0.80

Microsoft Windows 11 system with 12th Gen Intel Core i9-12900H, 2900 MHz, 14 Core(s), 20 Logical Processors CPU, 16 GB RAM and NVIDIA GeForce RTX 4060 graphics card. Regarding CHIWAWA, its running time depends on cloud content and the number of clouds present. For PRISMA images, it varies from 40 to 60 s (CC below 5%) to 6–8 min (CC around 60%). Table VI shows the CC%_est, CC%_ref, and ΔCC metrics, along with the processing time for each validation image. We observe that cloud-free scenes were mostly correctly labeled as noncloudy. However, as indicated by the ΔCC for

TABLE VI
ESTIMATED CLOUD COVERAGE AND PROCESSING TIME

Criteria	Id	V1	V2	V3	V4	V5
	CC%_ref	0	0	6.4	0	24.7
CC%_est	PRISMA	0	0	5.5	0	18.8
	CHIWAWA	0	0	2.63	0.03	27.87
	kNN	0.02	0	5.71	0	20.13
	RF	0.01	0	3.79	0	17.82
	XGBOOST	0.01	0	4.13	0	18.43
Δ CC	PRISMA	0	0	-0.9	0	-5.9
	CHIWAWA	0	0	-3.77	+0.03	+3.17
	kNN	+0.02	0	-0.69	0	-4.57
	RF	+0.01	0	-2.61	0	-6.88
	XGBOOST	+0.01	0	-2.27	0	-6.27
Time (s)	kNN	634	439	1008	650	958
	RF	1.19	1.52	1.79	1.27	1.80
	XGBOOST	1.19	1.36	1.38	1.10	1.67

Cloud coverage is evaluated for all methods, whereas processing time is reported only for kNN, RF, and XGBoost.

TABLE VII
OA, P, SEN, AND BOA METRICS FOR PRISMA, CHIWAWA, KNN, RF, AND XGBOOST FOR IMAGES ID V3 AND V5

Id	Metrics	PRISMA	CHIWAWA	kNN	RF	XGBOOST
V3	OA	95.75	95.51	**98.12**	97.30	97.57
	P	69.14	85.50	89.23	**98.39**	97.69
	Sen	60.10	35.38	**80.12**	58.54	63.36
	BOA	79.14	67.49	**89.73**	79.24	81.63
V5	OA	92.35	94.48	**94.50**	92.78	93.30
	P	95.32	84.43	97.75	**99.08**	98.85
	Sen	72.63	**95.22**	79.59	71.45	73.73
	BOA	85.73	**94.73**	89.50	85.62	86.72

images Ids V1 and V4, the proposed methods produced some false alarms in isolated pixels.

Table VII reports the results for the OA, P, Sen, and BOA metrics on images with IDs V3 and V5, which we considered as reference. We observe a slight improvement for the proposed methods compared to PRISMA and CHIWAWA. The results suggest that kNN achieves the highest OA performance. However, in terms of execution time (see Table VI), this algorithm is significantly outperformed by RF and XGBoost. Regarding P, RF and XGBoost achieved the best performance, with an average accuracy of over 98%. The average BOA results indicate better performance for kNN, XGBoost, and RF, with scores of 89.15%, 84.17%, and 82.43%, respectively. In terms of Sen, CHIWAWA achieved the best result for image Id V5 with 95.22%, in contrast to its poor performance on image Id V3, where it achieved only 35.38%. Generally speaking, the machine learning methods demonstrate better control over FPs, exhibiting lower P values across images ID v3 to v5, as well as good control over FNs in mountainous regions, as seen in image ID v3. In contrast, CHIWAWA performs well in controlling FNs, showing lower Sen values, except in image ID v3, where it struggles in mountainous areas due to the presence of bright pixels and thin clouds.

This quantitative evaluation is complemented by a qualitative visual inspection, shown in Figs. 7 and 8, which display the full extended masks and selected regions of interest in false color using CHIWAWA, kNN, RF, and XGBoost for the validation dataset. Regarding Fig. 7, for image Id V1, we observe isolated pixels erroneously classified as FP by kNN, RF, and XGBoost, but not by CHIWAWA. In image Id V3, RF and XGBoost primarily underestimate the presence of clouds, while CHIWAWA and kNN both underestimate and overestimate cloud presence in different areas, often misclassifying transition zones between clouds and bare soil. In image Id V4, CHIWAWA shows isolated pixels incorrectly classified as cloudy. In image Id V5, CHIWAWA and kNN tend to overestimate cloud presence, whereas RF and XGBoost behave more conservative. Fig. 8 shows a detailed analysis of the regions highlighted by the yellow boxes in the full images. These areas include productive agricultural zones, as seen in images with Ids V2, V4, and V5. Image Id V1 contains a mixture of agricultural fields and urban areas, while image Id V3 corresponds to a mountainous region affected by both snow and cloud cover.

In the zoomed-in regions of images Id V1 and Id V4, we identify productive cloud-free areas that include some bright pixels.

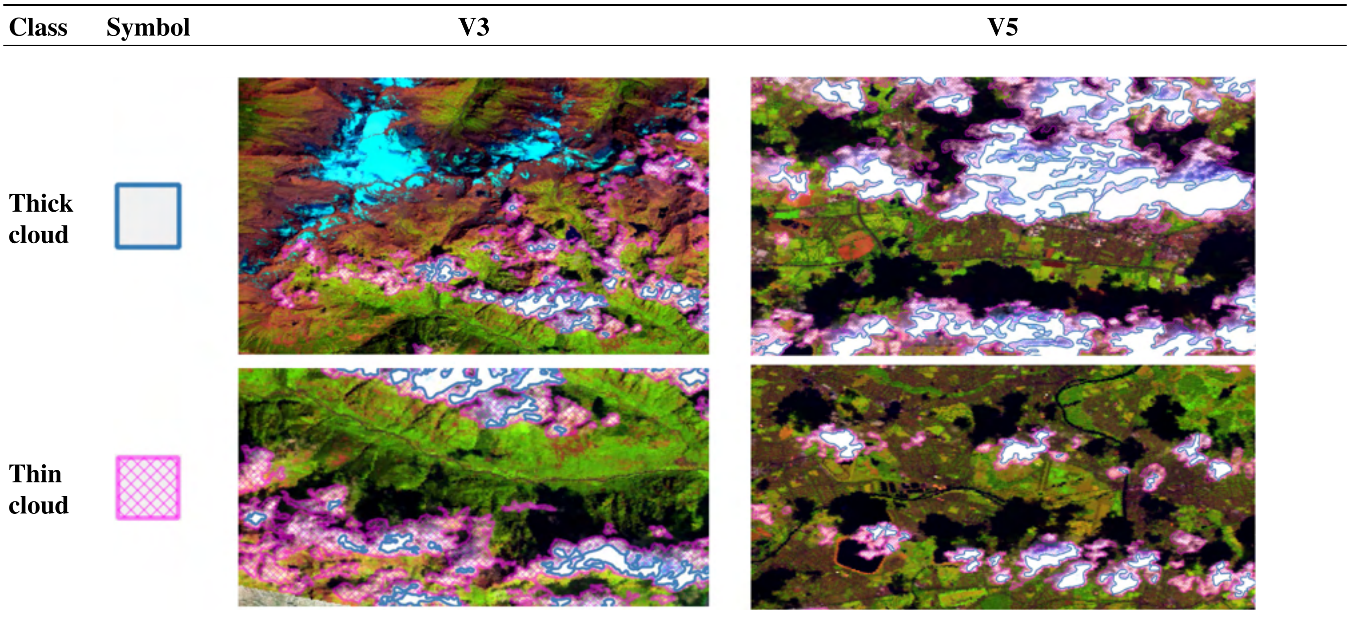

Fig. 6. Example illustrating the differentiation between thin and thick clouds (represented by hashed pink and blue polygons, respectively) for images V3 and V5. The background images are displayed in false color: *Red:* 1647 nm, *Green:* 838 nm, and *Blue:* 482 nm.

TABLE VIII
ESTIMATED CLOUD-COVER SENSITIVITY BY CLOUD-THICKNESS CATEGORY FOR ALL METHODS: PRISMA, KNN, RF, CHIWAWA, AND XGBOOST

Criteria	Id	Thin Clouds		Thick Clouds	
		V3	V5	V3	V5
	Total (px)	50340	155653	10925	81215
Sen	**PRISMA**	52.90	60.76	93.25	95.38
	CHIWAWA	21.54	**92.73**	99.14	**99.99**
	kNN	**75.81**	68.98	**99.98**	99.93
	RF	49.57	56.63	99.88	99.86
	XGBOOST	55.43	60.08	99.89	99.88

Most methods correctly classify these areas as noncloud, except by kNN and CHIWAWA, which misclassify some isolated pixels as clouds. In contrast, the zoomed-in region of image Id V2 shows that all crop areas are correctly labeled as noncloud by every method.

Image Id V3 depicts a mountainous region characterized by high-brightness surfaces, such as snow and bare soil. In this context, CHIWAWA tends to overestimate cloud presence by misclassifying bare soils as clouds. It also underestimates cloud presence, incorrectly labeling thin cloud areas as clear blue sky. This limitation has also been documented by [10], who reported that an abundance of bright pixels increases the false alarm rate. Conversely, RF and XGBoost tend to produce a higher number of FN. Although kNN shows comparatively better performance in identifying TP, all methods exhibit notable challenges in accurately detecting thin clouds.

Image Id V5 shows a region with clouds over a flat area. In this set of masks, kNN, RF, and XGBoost tend to underestimate cloud edges, whereas CHIWAWA tends to overestimate them.

Finally, to further investigate the performance of the considered methods in detecting thin clouds, we manually annotated the cloudy pixels in images V3 and V5 as either "thin" or "thick," as shown in Fig. 6. Then, we evaluated the sensitivity metric (Sen) of all methods for Thin Clouds and Thick Clouds separately for the validation images. The results are reported in Table VIII. Overall, all methods were consistent in detecting thick clouds, achieving accuracies above 99.8% for the evaluated images. This contrasts with the native PRISMA cloud mask, which reached 93.25% and 95.38% for images with Ids V3 and V5, respectively.

For areas affected by thin clouds, the sensitivity shows greater variability. In image Id V3, kNN achieved the highest sensitivity (75.81%), whereas CHIWAWA obtained the lowest (21.54%). Conversely, in image Id V5, CHIWAWA had the highest value (92.73%) and RF had the lowest (56.63%). In this last case, it is interesting to underline the kNN's performance, which achieved an overall sensitivity of 68.98%. The kNN results for both images suggest that leveraging information from neighboring pixels may be beneficial for classification, especially in this

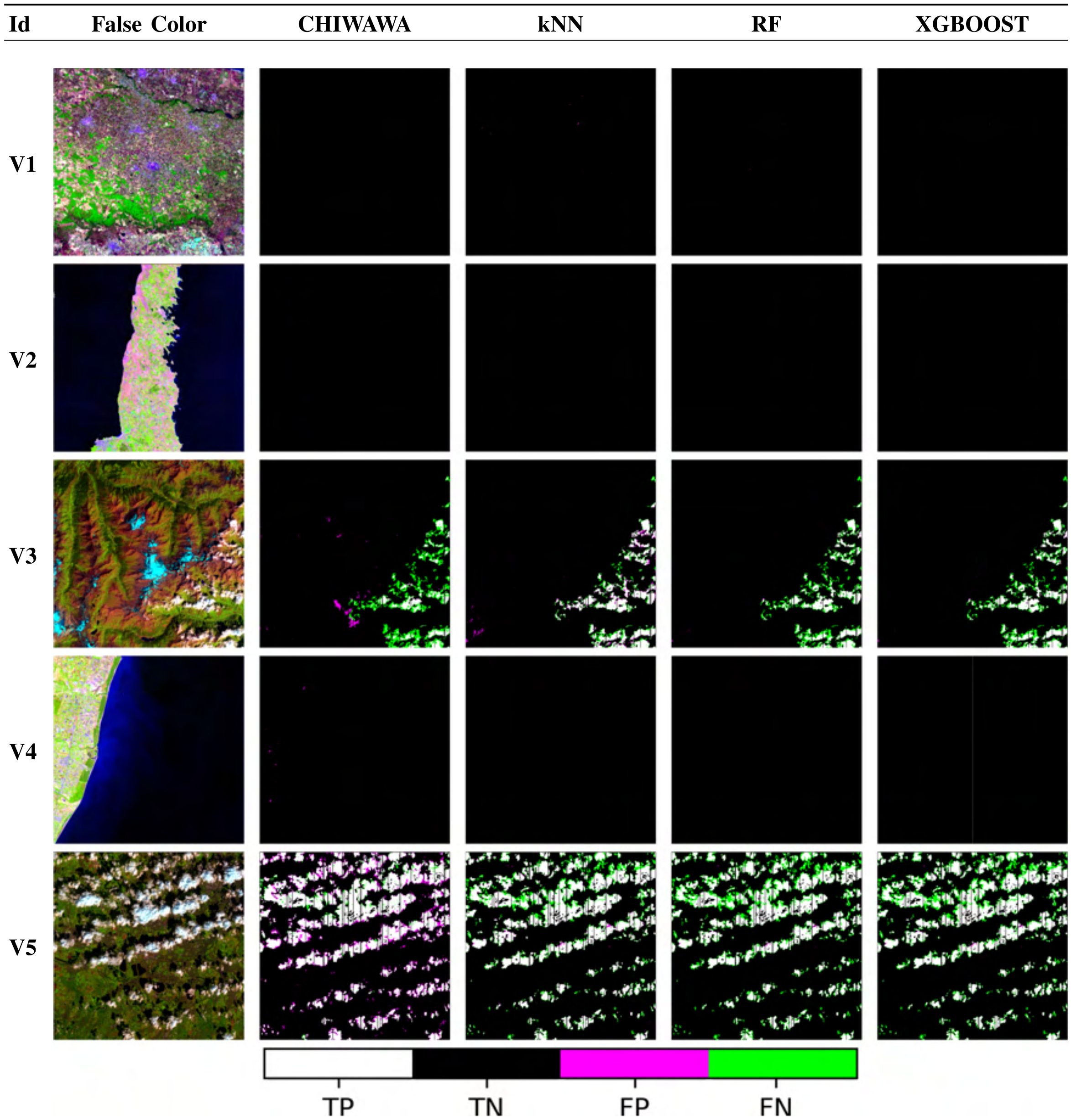

Fig. 7. Visual comparison of CHIWAWA, kNN, RF, and XGBOOST over the validation dataset. False color: *Red:* 1647 nm, *Green:* 838 nm, and *Blue:* 482 nm.

context where all the tested machine learning methods were trained only on a generic "cloud" class, without explicit distinction for "thick/thin" classes. However, these disparate results collectively highlight that accurately identifying thin clouds remains a significant challenge for all the evaluated methods.

VI. Discussion

Cloud detection using multiple spectral bands is a well-studied topic, as shown by the CMIX, the first comprehensive international evaluation of cloud detection algorithms tailored for moderate-resolution Earth observation satellites, specifically Landsat 8 and Sentinel 2 [12].[8] The participants were 10 cloud masking algorithms developed by nine international teams, representing universities, research labs, industry, and space agencies—including CNES, ESA, DLR, and NASA. In particular, one of the methods was based on deep neural network, precisely a modified U-Net architecture (CD-FCNN), another one on GBT (S2cloudless), and the last one is a multitemporal extension of the s2cloudless algorithm (InterSSIM). In contrast, cloud detection for hyperspectral sensors is indeed an evolving field, and while it draws heavily from techniques developed for multispectral data, there are key differences and growing specialization due to the unique characteristics of hyperspectral data [4], [5], [7], [9], [10]. This article is the first attempt to systematically assess the performance of three machine learning methods for cloud detection in PRISMA hyperspectral data. It presents a prototype software implementation of kNN, RF,

[8][Online]. Available: https://calvalportal.ceos.org/cmix

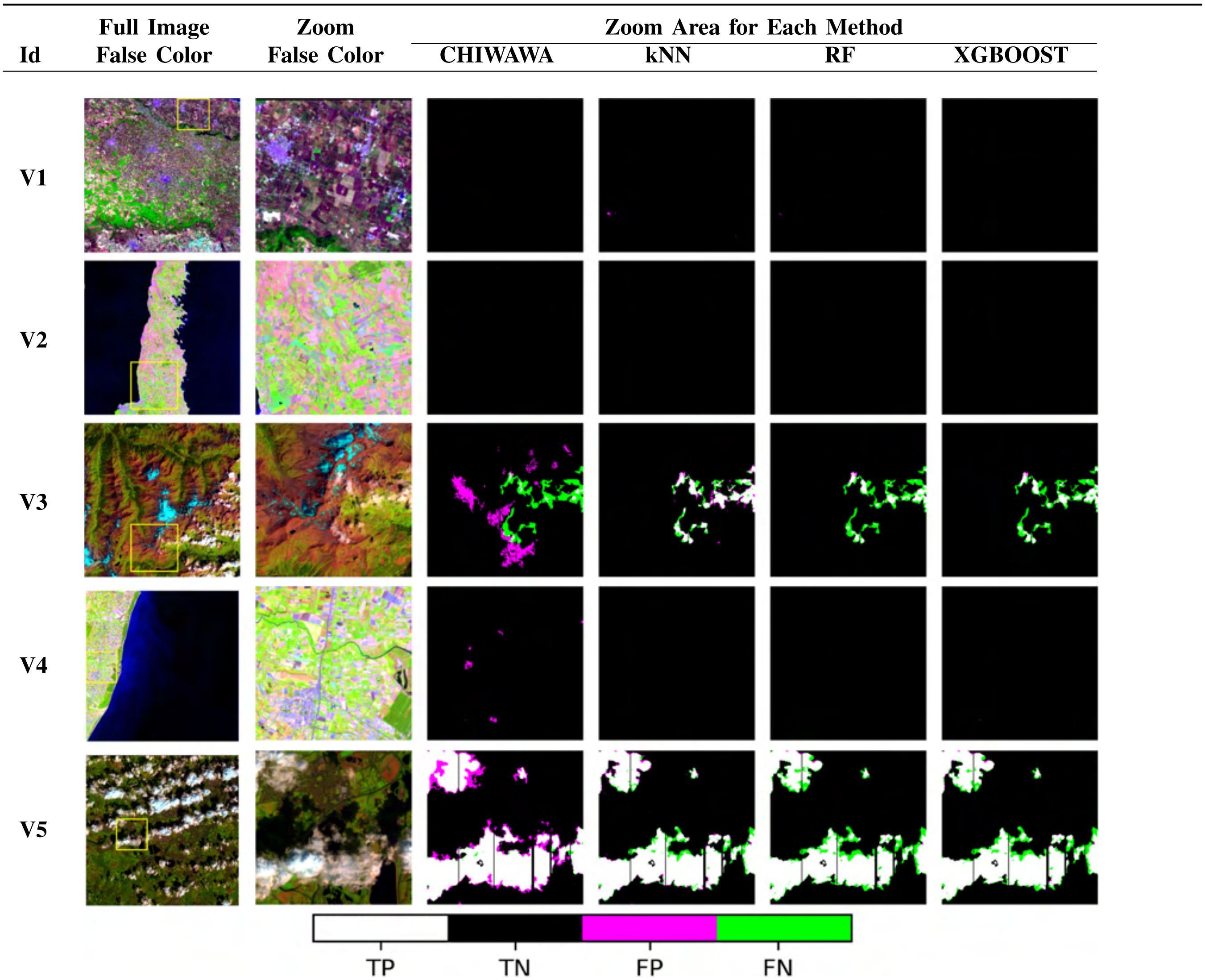

Fig. 8. Visual comparison of CHIWAWA, kNN, RF, and XGBoost. Yellow areas represent a zoom on the Cloud mask results. *Red:* 1647 nm, *Green:* 838 nm, and *Blue:* 482 nm.

and XGBoost, featuring automatic hyperparameter tuning and a comprehensive dataset of 26 *manually annotated* PRISMA images. The proposed techniques are compared against the native PRISMA land use mask and the CHIWAWA algorithm, which was specifically designed for both PRISMA and AVIRIS-NG data. The results demonstrated the strong performance of the methods, achieving an accuracy of 98.12% in mountainous regions—areas where both the native PRISMA mask and CHIWAWA encounter some difficulties. This is noteworthy considering that CHIWAWA incorporates a WV map, which serves a function similar to thermal bands—known to be highly effective for cloud detection but absent in the reflective spectral range. Furthermore, all three methods showed excellent control over FP, with precision reaching 98.39% in mountainous regions and 99.08% in flat areas, and good control over FN in mountainous regions with a sensitivity of 80.12%, less over flat regions. However, all the methods generally exhibit limitations in accurately detecting thin clouds and edge clouds. Indeed, a critical aspect underlying the performance of the proposed machine learning methods is the quality of the preprocessing steps and the manually curated training dataset. Although manual annotation remains the most reliable approach in the absence of a benchmark cloud mask for PRISMA, by-eye classification inevitably introduces uncertainty—particularly for thin clouds, cloud edges, and bright surfaces whose spectral signatures overlap with those of optically thin cloud layers. First, the conversion from DN values to TOA reflectance involves band-dependent calibration parameters (offsets, scale factors, E_{Sun} values, and solar geometry). Small inaccuracies in this step—especially within bands affected by atmospheric absorption or low SNR—may influence the separability between classes. Because our cloud detection models operate on a reduced set of 20 VNIR–SWIR channels, any reflectance instability in these selected bands could affect the decision boundary. This is particularly relevant for the WV absorption bands (942 and 1131 nm), which are known to be more sensitive to atmospheric variability and may behave differently over bright surfaces or elevated terrain. Second, while thick clouds show a distinctive spectral signature and high brightness, thin clouds and cloud edges often merge smoothly into the background reflectance of land or snow, producing transition zones where the cloud–clear boundary is ambiguous even upon visual inspection. In these regions, the manual polygons inevitably contain mixed pixels. As a consequence, some pixels labeled "cloud" may include partial contributions from the underlying surface, whereas some pixels labeled "clear" may contain residual cloud contamination. These effects can introduce label noise in the training set, which in turn may bias the learning process toward either underdetecting thin clouds (if many thin clouds were mistakenly labeled as clear) or producing false alarms (if some bright surfaces were inadvertently included in cloud polygons). Third, the heterogeneity of bright noncloud classes contributes additional ambiguity. Snow, high-altitude bare soil, and certain urban materials—in particular greenhouse areas—exhibit high reflectance across the VNIR and SWIR, sometimes mimicking the spectral distribution of thin clouds. During annotation, distinguishing these materials from semi-transparent clouds is nontrivial unless texture or contextual information is taken into account. Since our models are trained on single-pixel spectra without spatial context, any mislabeled boundaries or ambiguous samples may train the algorithm to associate high reflectance with the cloud class even in the absence of atmospheric features. This effect partly explains why RF and XGBoost tend to generate FP in isolated pixels, especially near the edges of bright surfaces.

A further aspect to be considered concerns the feasibility of applying the proposed machine learning methods in an operational context, where large volumes of hyperspectral data must be processed routinely. The kNN algorithm, while conceptually simple and competitive in terms of classification accuracy—especially in the detection of thin clouds—does not scale efficiently with data size due to its high computational cost at inference time. Since predictions require computing distances between each test pixel and the entire training set, the complexity grows linearly with the number of training samples and the spectral dimensionality. In this regard, several optimization methods could be explored, such as approximate nearest neighbor search, KD-tree or ball-tree acceleration, locality-sensitive hashing, and prototype-reduction techniques that condense the training set. These strategies would significantly reduce query time while preserving most of the classification accuracy, thereby making kNN more suitable for large-scale applications. In contrast, both RF and XGBoost are considerably more attractive for operational scenarios. After the initial training phase—which is computationally heavier but performed only once—RF and XGBoost provide very fast inference, with prediction times independent of the size of the training set and governed instead by the number and depth of the trees. This characteristic makes them scalable and cost-effective for near-real-time cloud detection over large image archives or continuous data streams. Furthermore, these tree-based models can be parallelized efficiently across multiple CPU cores or GPUs and integrated into distributed processing frameworks commonly used in remote sensing (e.g., Dask, Spark, and cloud-based pipelines). Their relatively small memory footprint and robustness to missing or noisy spectral bands also contribute to their suitability for operational PRISMA data processing. Overall, although kNN benefits from simplicity and interpretability, its direct application to large-scale hyperspectral datasets would require algorithmic optimization or approximate search strategies. By contrast, RF and XGBoost exhibit stronger operational scalability and are inherently more compatible with real-time or high-throughput processing environments. Future work should focus on expanding the *manually annotated dataset* both spatially and temporally, in order to fully leverage the capabilities of transformer-based architectures [15], CNNs [14], and hybrid models [16], whose performance on image data is well established provided that a sufficiently large training dataset is given. In this perspective, the approach proposed in some lightweight models for real-time hyperspectral cloud detection [8], [17], [18], [19], [20] would also suggest optimization strategies for the applicability of RF/XGBoost, as well as these deep learning approaches, in operational contexts.

VII. Conclusion

We leveraged the capabilities of three machine learning methods—kNN, RF, and XGBoost—for cloud detection in PRISMA hyperspectral images. Their performance is compared against two threshold-based techniques: the native PRISMA land-use mask provided by ASI alongside L1 images and the CHIWAWA algorithm.

To support this analysis, we built a benchmark dataset by manually annotating 26 PRISMA images through visual inspection. These images represent a range of cloud-cover conditions—from heavily clouded to cloud-free. The dataset includes samples from various spectral categories, such as urban areas, wetlands, forests, bare soil, water bodies, agricultural land, snow, and clouds. These samples were collected from multiple sites across Europe and Northern Africa, including the Sahara Desert, to ensure a comprehensive representation of all categories. The selected images span the late spring and summer seasons from 2021 to 2023.

The results indicate that machine learning methods yield strong and promising performance for cloud detection, particularly in mountainous regions where threshold-based approaches require careful tuning to mitigate FP and FN caused by bright surfaces and thin clouds. A common challenge for all methods lies in the estimation of cloud edges and the detection of thin clouds. Specifically, kNN, RF, and XGBoost tend to underestimate cloud boundaries, while CHIWAWA tends to overestimate them. Experimental findings also highlight the potential benefits of incorporating WV bands for improved cloud differentiation.

Acknowledgment

PRISMA data were available from the Italian Space Agency (ASI) website. The authors would like to thank A. Alakian for providing cloud masks results for the CHIWAWA algorithm on the selected PRISMA images. This work is dedicated to Umberto Amato, scientist, mentor, who taught us the value of questioning and the courage to look beyond the first answer.

References

[1] R. O. Green et al., "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425798000649

[2] Y.-N. Liu, D.-X. Sun, X.-N. Hu, S.-F. Liu, and K.-Q. Cao, "AHSI: The hyperspectral imager on China's GaoFen-5 satellite," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 509, no. 1, 2020, Art. no. 012033. [Online]. Available: https://iopscience.iop.org/article/10.1088/1755-1315/509/1/012033

[3] S. Mahalingam et al., "Reflectance based vicarious calibration of HySIS sensors and spectral stability study over pseudo-invariant sites," in *Proc. 2019 IEEE Recent Adv. Geosci. Remote Sens. : Technol., Standards Appl.*, 2019, pp. 132–136.

[4] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425711003853

[5] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images," *Remote Sens. Environ.*, vol. 159, pp. 269–277, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425714005069

[6] S. Mahajan and B. Fataniya, "Cloud detection methodologies: Variants and development—A review," *Complex Intell. Syst.*, vol. 6, no. 2, pp. 251–261, 2020.

[7] L. Wang and Q. Wang, "Fast spatial-spectral random forests for thick cloud removal of hyperspectral images," *Int. J. Appl. Earth Observation Geoinf.*, vol. 112, 2022, Art. no. 102916. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1569843222001170

[8] M. Ali et al., "Lightweight cloud masking models for on-board inference in hyperspectral imaging," 2025, *arXiv:2507.08052*.

[9] L. Sun et al., "A cloud detection algorithm-generating method for remote sensing data at visible to short-wave infrared wavelengths," *ISPRS J. Photogrammetry Remote Sens.*, vol. 124, pp. 70–88, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271616306189

[10] A. Alakian, "Cloud detection in hyperspectral images with atmospheric column water vapor: Application to PRISMA and AVIRIS-NG images," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5510824.

[11] A. S. I. ASI, "PRISMA Algorithm Theoretical Basis Document (ATBD)," ASI—Italian Space Agency, Tech. Rep. 1, 2021.

[12] S. Skakun et al., "Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2," *Remote Sens. Environ.*, vol. 274, 2022, Art. no. 112990. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425722001043

[13] F. Xie, M. Shi, Z. Shi, J. Yin, and D. Zhao, "Multilevel cloud detection in remote sensing images based on deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3631–3640, Aug. 2017.

[14] J. Yang, J. Guo, H. Yue, Z. Liu, H. Hu, and K. Li, "CDNet: CNN-based cloud detection for remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6195–6211, Aug. 2019.

[15] B. Zhang, Y. Zhang, Y. Li, Y. Wan, and Y. Yao, "CloudViT: A lightweight vision transformer network for remote sensing cloud detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 5000405.

[16] N. Ma, L. Sun, Y. He, C. Zhou, and C. Dong, "CNN-TransNet: A hybrid CNN-Transformer network with differential feature enhancement for cloud detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 1001705.

[17] G. Giuffrida et al., "CloudScout: A deep neural network for on-board cloud detection on hyperspectral images," *Remote Sens.*, vol. 12, no. 14, 2020, Art. no. 2205. [Online]. Available: https://www.mdpi.com/2072-4292/12/14/2205

[18] R. Pitoňák, J. Mucha, L. Dobis, M. Javorka, and M. Marusin, "CloudSatNet-1: FPGA-Based hardware-accelerated quantized CNN for satellite on-board cloud coverage classification," *Remote Sens.*, vol. 14, no. 13, 2022, Art. no. 3180. [Online]. Available: https://www.mdpi.com/2072-4292/14/13/3180

[19] J. A. Justo et al., "Semantic segmentation in satellite hyperspectral imagery by deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 273–293, 2025.

[20] J. Kněžık, J. Herec, and R. Pitoňák, "Fast-SEnSeI: Lightweight sensor-independent cloud masking for on-board multispectral sensors," 2025, *arXiv:2509.20991*.

[21] D. G. Loyola et al., "The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor," *Atmospheric Meas. Techn.*, vol. 11, no. 1, pp. 409–427, 2018. [Online]. Available: https://amt.copernicus.org/articles/11/409/2018/

[22] W. Wang et al., "A novel physics-based cloud retrieval algorithm based on neural networks (CRANN) from hyperspectral measurements in the O_2-O_2 band," *Remote Sens. Environ.*, vol. 311, 2024, Art. no. 114267. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0034425724002852

[23] W. Wang et al., "Development of an algorithm for the simultaneous retrieval of cloud-top height and cloud optical thickness combining radiative transfer and multisource satellite information from O_4 hyperspectral measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 4104111.

[24] N. Acito, M. F. Carfora, M. Diani, G. Corsini, S. Pascucci, and S. Pignatti, "Noise coefficients retrieval in PRISMA hyperspectral data," in *Proc. IGARSS 2023-2023 IEEE Int. Geosci. Remote Sens. Symp.*, 2023, pp. 1493–1496.

[25] S. Pignatti et al., "PRISMA L1 and L2 performances within the PRISCAV project: The Pignola test site in Southern Italy," *Remote Sens.*, vol. 14, no. 9, 2022, Art. no. 1985. [Online]. Available: https://www.mdpi.com/2072-4292/14/9/1985

[26] S. Cogliati et al., "The PRISMA imaging spectroscopy mission: Overview and first performance analysis," *Remote Sens. Environ.*, vol. 262, 2021, Art. no. 112499. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0034425721002170

[27] N. Acito, M. Diani, and G. Corsini, "PRISMA spatial resolution enhancement by fusion with Sentinel-2 data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 62–79, 2022.

[28] F. Braga et al., "Assessment of PRISMA water reflectance using autonomous hyperspectral radiometry," *ISPRS J. Photogrammetry Remote Sens.*, vol. 192, pp. 99–114, 2022.

[29] J. Verrelst et al., "Mapping landscape canopy nitrogen content from space using PRISMA data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 178, pp. 382–395, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271621001751

[30] G. Tagliabue et al., "Hybrid retrieval of crop traits from multi-temporal PRISMA hyperspectral imagery," *ISPRS J. Photogrammetry Remote Sens.*, vol. 187, pp. 362–377, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271622000879

[31] M. Niroumand-Jadidi, F. Bovolo, and L. Bruzzone, "Water quality retrieval from PRISMA hyperspectral images: First experience in a turbid lake and comparison with Sentinel-2," *Remote Sens.*, vol. 12, no. 23, 2020, Art. no. 3984. [Online]. Available: https://www.mdpi.com/2072-4292/12/23/3984

[32] F. N. Begliomini et al., "Machine learning for cyanobacteria mapping on tropical urban reservoirs using PRISMA hyperspectral data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 204, pp. 378–396, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271623002617

[33] G. D. Luca, J. L. Pancorbo, F. Carotenuto, B. Gioli, G. Modica, and L. Genesio, "PRISMA imaging for land covers and surface materials composition in urban and rural areas adopting multiple endmember spectral mixture analysis (MESMA)," *ISPRS J. Photogrammetry Remote Sens.*, vol. 225, pp. 196–220, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0924271625001807

[34] C. Quintano, L. Calvo, A. Fernández-Manso, S. Suárez-Seoane, P. M. Fernandes, and J. M. Fernández-Guisuraga, "First evaluation of fire severity retrieval from PRISMA hyperspectral data," *Remote Sens. Environ.*, vol. 295, 2023, Art. no. 113670.

[35] S. Mirzaei et al., "Early-season crop mapping by PRISMA images using machine/deep learning approaches: Italy and Iran test cases," *Remote Sens.*, vol. 16, no. 13, 2024, Art. no. 2431. [Online]. Available: https://www.mdpi.com/2072-4292/16/13/2431

[36] N. Acito, M. Diani, M. Alibani, and G. Corsini, "Matched filter based on the radiative transfer model for CO_2 estimation from PRISMA hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5529413.

[37] R. Guarini et al., "PRISMA hyperspectral mission products," in *Proc. IGARSS 2018-2018 IEEE Int. Geosci. Remote Sens. Symp.*, 2018, pp. 179–182.

[38] C. Aybar et al., "CloudSEN12 : The largest dataset of expert-labeled pixels for cloud and cloud shadow detection in Sentinel-2," *Data Brief*, vol. 56, 2024, Art. no. 110852.

[39] B.-C. Gao and A. F. Goetz, "Column atmospheric water vapor and vegetation liquid water retrievals from airborne imaging spectrometer data," *J. Geophys. Res.: Atmospheres*, vol. 95, no. D4, pp. 3549–3564, 1990.

[40] E. Fix and J. L. Hodges, "Discriminatory analysis. Nonparametric discrimination; consistency properties," USAF Sch. Aviation Med., Randolph Field, TX, USA, Tech. Rep. 4,1951.

[41] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.

[42] L. Breiman, “Bagging predictors,” *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996. [Online]. Available: https://doi.org/10.1007/BF00058655

[43] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.

[44] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, 2nd ed., ser. Springer Series in Statistics. New York, NY, USA: Springer, 2009.

[45] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. data mining*, 2016, pp. 785–794.

[46] O. Sagi and L. Rokach, “Approximating XGBoost with an interpretable decision tree,” *Inf. Sci.*, vol. 572, pp. 522–542, 2021.

[47] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework,” in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2019, pp. 2623–2631.

[48] C. Aybar et al., “CloudSEN12, a global dataset for semantic understanding of cloud and cloud shadow in Sentinel-2,” *Sci. Data*, vol. 9, no. 1, 2022, Art. no. 782.

Maria Francesca Carfora received the M.S. degree in mathematics and the Ph.D. degree in computational mathematics from the University of Naples “Federico II,” Italy, in 1988 and 2014, respectively.

In 1996, she joined the Italian National Research Council, where she is currently a Senior Researcher. Since 1997, she has been PI or Partner in several research projects, funded by National and International Organizations. She also acts as a Research Director for young fellows and participate in Organizing Committees. She has authored and coauthored more than 50 publications in peer-reviewed academic journals. Her research interests include the development and implementation of numerical methods for ODE, PDE, and SDE with applications to medical and biological images, remote sensing, biological systems; other fields of interest are data and image analysis, statistical modeling and parameter estimation.

Italia De Feis received the Ph.D. degree in applied mathematics and computer science from the University of Naples “Federico II,” Italy, in 1999.

She is currently a Senior Researcher with the Istituto per le Applicazioni del Calcolo “M. Picone,” Italian National Research Council (CNR), Naples. She has served as a Principal Investigator (PI) and Co-Investigator in several projects regarding next-generation infrared sounders. These projects include the Infrared Atmospheric Sounding Interferometer (IASI), the IASI – NextGeneration (IASI-NG), the InfraRed Sounder of Meteosat Third Generation (MTG-IRS) and the Hyperspectral Precursor of the Application Mission (PRISMA). Her research interests include inverse problems (radiative transfer equation in the atmosphere), statistical machine learning, and data integration (L3 products and set up of water stress indexes).

Diana Carolina Fonnegra Mora received the M.S. degree in remote sensing and a postgraduation diploma in data science from the National University of Córdoba, Crdoba, Argentina, in 2018 and 2022, respectively. She is currently working toward the Ph.D. degree in engineering for innovation and sustainable development with the University of Basilicata, Potenza, Italy.

Since 2012, she has been a Topographical Engineer with the Universidad del Valle, Cali, Colombia. Since 2023, she has collaborated with the Istituto per le Applicazioni del Calcolo “M. Picone’ (CNR-IAC), Naples, Italy. Her research interests include satellite data processing and developing statistical machine learning models for hyperspectral data integration.