remote sensing

MDPI

Article

Early-Season Crop Mapping by PRISMA Images Using Machine/Deep Learning Approaches: Italy and Iran Test Cases

Saham Mirzaei [1], Simone Pascucci [1,*], Maria Francesca Carfora [2], Raffaele Casa [3], Francesco Rossi [4], Federico Santini [1], Angelo Palombo [1], Giovanni Laneve [4] and Stefano Pignatti [1]

[1] Institute of Methodologies for Environmental Analysis (IMAA), Italian National Research Council (CNR), C/da S. Loja, 85050 Tito Scalo, Italy; sahammirzaei@cnr.it (S.M.); federico.santini@cnr.it (F.S.); angelo.palombo@cnr.it (A.P.); stefano.pignatti@cnr.it (S.P.)

[2] Istituto per le Applicazioni del Calcolo "Mauro Picone" (IAC), Italian National Research Council (CNR), Via Pietro Castellino 111, 80131 Napoli, Italy; f.carfora@na.iac.cnr.it

[3] Department of Agriculture and Forestry Sciences (DAFNE), University of Tuscia, Via San Camillo de Lellis, 01100 Viterbo, Italy; rcasa@unitus.it

[4] Scuola Ingegneria Aerospaziale (SIA), University of Rome "La Sapienza", Via Salaria 851, 00138 Roma, Italy; francesco.rossi@uniroma1.it (F.R.); giovanni.laneve@uniroma1.it (G.L.)

* Correspondence: simone.pascucci@cnr.it; Tel.: +39-06-4548-8685

Abstract: Despite its high importance for crop yield prediction and monitoring, early-season crop mapping is severely hampered by the absence of timely ground truth. To cope with this issue, this study aims at evaluating the capability of PRISMA hyperspectral satellite images compared with Sentinel-2 multispectral imagery to produce early- and in-season crop maps using consolidated machine and deep learning algorithms. Results show that the accuracy of crop type classification using Sentinel-2 images is meaningfully poor compared with PRISMA (14% in overall accuracy (OA)). The 1D-CNN algorithm, with 89%, 91%, and 92% OA for winter, summer, and perennial cultivations, respectively, shows for the PRISMA images the highest accuracy in the in-season crop mapping and the fastest algorithm that achieves acceptable accuracy (OA 80%) for the winter, summer, and perennial cultivations early-season mapping using PRISMA images. Moreover, the 1D-CNN algorithm shows a limited reduction (6%) in performance, appearing to be the best algorithm for crop mapping within operational use in cross-farm applications. Machine/deep learning classification algorithms applied on the test fields cross-scene demonstrate that PRISMA hyperspectral time series images can provide good results for early- and in-season crop mapping.

Keywords: PRISMA; Sentinel-2; early-season crop mapping; machine learning; deep learning

check for updates

Citation: Mirzaei, S.; Pascucci, S.; Carfora, M.F.; Casa, R.; Rossi, F.; Santini, F.; Palombo, A.; Laneve, G.; Pignatti, S. Early-Season Crop Mapping by PRISMA Images Using Machine/Deep Learning Approaches: Italy and Iran Test Cases. *Remote Sens.* **2024**, *16*, 2431. https://doi.org/10.3390/rs16132431

Academic Editor: Yuxin Miao

Received: 6 May 2024
Revised: 19 June 2024
Accepted: 30 June 2024
Published: 2 July 2024

1. Introduction

Early-season crop mapping is valuable for estimating the area under cultivation and for yield prediction [1], which is vital information for marketing and decision making involved in food security. Moreover, accurate and detailed in terms of number of crop species, crop mapping is important for agricultural management, economic development planning, and agroecosystem conservation. Precise crop maps could be used for analyzing the implementation and effects of agri-environmental policies [2] and planning of optimal and sustainable agronomic management based on crop rotations [3]. National and international institutions require an early estimation of the planted area for food security issues [4]. Parallelly, there is increasing interest from crop insurance companies in precise early crop mapping and weather-based crop health indicators to be included in their area-yield crop insurance schemes, especially in the early season when the vegetation cover is low and fields do not show a dense homogeneous canopy.

Time series-based analysis of multispectral images, e.g., Landsat and Sentinel, are the most common data sources that are used for crop mapping [5–7]. The time series method is more accurate when data covering the whole growing season are used [8]. Interference from clouds and aerosols is the main drawback of optical remote sensing data, making time series analysis more challenging [9].

Concerning hyperspectral remote sensing, UAV hyperspectral cameras [10–13] and airborne hyperspectral data [14,15] have been extensively used for crop mapping. Spaceborne hyperspectral images have been used for crop mapping [16–18]. However, moving from proximal/airborne to spaceborne sensors, new challenging issues, related to low signal-to-noise ratio (SNR), spectral mixing issues related to the 30 m/pixel supported by the present missions, and atmosphere attenuations, must be considered. The two hyperspectral technology demonstration missions, i.e., the NASA Hyperion onboard the Earth Observing-1 (EO-1) spacecraft [19] and the Compact High Resolution Imaging Spectrometer (CHRIS) on ESA's Proba-1 microsatellite [20] did not show suitable temporal resolution and did not receive much attention from geoscientists and agronomists for crop-type mapping. Conversely, significant improvements in this context can be obtained nowadays, following the operation of the Italian Space Agency (ASI) PRISMA (Hyperspectral Precursor of the Application Mission) [21] hyperspectral satellite mission since 2019, the German EnMAP (Environmental Mapping and Analysis Program) since 2022 [22], the JPL-NASA EMIT (Earth Surface Mineral Dust Source Investigation) [23] since 2022, and the German DLR Earth Sensing Imaging Spectrometer (DESIS) [24], which provide hyperspectral images to the community worldwide. Furthermore, potential added value would be explored by forthcoming hyperspectral missions such as the ESA CHIME [25], ASI-PRISMA Second Generation (https://space.leonardo.com/en/news-and-stories-detail/-/detail/prisma-accordo-progetto-seconda-generazione, accessed on 5 May 2024), and NASA SBG [26]. Despite improvements in the suitable swath width and revisit time, the ongoing hyperspectral satellites are still not fully suitable for early crop mapping and monitoring [27] unless applied in a tandem configuration on specific selected test sites. However, it is foreseen that the higher temporal resolution of the hyperspectral imagery will be able to provide, in the following years, denser time series suitable to the development of new high value products covering the crop growing season [28]. One of the main challenges posed by spaceborne hyperspectral imagery is the availability of relevant information in high-dimensional data containing highly correlated spectral information. As a branch of artificial intelligence, machine learning (ML) refers to algorithms that can explore and derive meaningful information from data and use them within a self-learning approach to construct algorithms/models for accurate classification or prediction. ML has gradually gained popularity due to its accuracy and reliability. ML algorithms, e.g., artificial neural networks (ANNs), support vector machines (SVMs), and random forest (RF) have been widely applied in the last decade in crop type identification [16,17,29–32].

Furthermore, the use of deep learning (DL) algorithms has allowed to explore several levels of distributed representations from the input training dataset. Among them, convolutional neural network (CNN) is the most common DL method applied to crop mapping [33]. This is because CNNs reduce the need for building handcrafted feature extractors, by demonstrating an exceptional ability to learn complex representations directly from image bands. Larger and more diverse training data lead to improved functionality of CNNs. A long running time, especially when working with big data, has been a challenging issue that is expected to be solved by improved hardware and software components of machine vision systems [34,35]. In comparison with CNNs, the architectures with more layers and parameters, like Visual Geometry Group (VGG-16), AlexNet, and Res-Net, need more training data to fit the model parameters [36].

Classical ML algorithms process each pixel vector based on spectral features without considering spatial contextual information [37]. Keeping the spatial information could help crop type mapping because neighboring pixels have higher probabilities of belonging to the same crop class [38]. Hence, 2D-CNN can calculate spatial features in both directions, but it does not have spectral features. 2D-CNN has been used for crop mapping [39] to identify the phenology in Sentinel-2 images [40] and leaf age from a single image [41].

3D-CNN, even if at a higher computational cost, is able to exploit high-dimensional spectral features [42]. The main issue in 3D methods is the increased dimensionality of samples that may affect the classification accuracy and efficiency [43]. Spectral/spatial pixel-wise hyperspectral image classification can be achieved by integrating spatial features into spectral information. Therefore, different feature extraction methods have been used to overcome this issue [18,44–46].

Early-season crop mapping using supervised methods could be severely hampered by the absence of contemporary Earth observation (EO) data and ground truth acquisitions. Ideally, ground truth should be obtained from surveys that collect targeted first-hand information, but it is often impeded by the substantial cost of time and labor [47]. Most classifiers work well through training using similar ground truth data, while the spatial transferability of these methods is still a challenging issue. In addition, environmental conditions (rainfall, temperature, etc.) could change over the years, severely impacting crop growing stages. To overcome this limitation, several early-season mapping methodologies, e.g., sample migration [48] and temporal encoding [49], have been developed to effectively classify crop types, even in situations where there is limited ground truth data available for that specific growing season. These approaches leverage historical data to generate crop labels and prior information for early-season crop identification. Moreover, a transfer learning (TL) scheme based on DL approaches helps to gain crop mapping results in transfer sites.

This study presents the application of the 1D-CNN and 3D-CNN deep learning algorithms and an ensemble of learning techniques like RF, SVM, K-nearest neighbor (KNN), and multiclass naive bayes (MNB) algorithms to a first available time series of new hyperspectral imagery belonging to four farm test cases. At this scope, the PRISMA mission provides an ideal benchmark for demonstrating the advancement offered by hyperspectral time series for the application in the agriculture context for crop mapping. This brings us to clarify the objectives of our study: (1) to develop a cross-farm ML/DL training and accuracy assessment when there are no in situ data available for training and (2) to test fine early-season crop mapping capabilities of hyperspectral time series data in comparison with multispectral imagery with a fine temporal resolution. The novelty of this research lies in (1) the supply of contemporaneous ground truth prepared by analyzing the PRISMA and Sentinel-2 optical images for training and validation; (2) demonstrating by ML and DL algorithms the importance of the use of hyperspectral satellite time series data for the accurate early- and in-season crop mapping. Section 2 describes the study areas used for the field data collection and analysis, and the methodology for achieving this study's objectives. Sections 3 and 4 report and discuss the results.

2. Materials and Methods

2.1. Study Areas

Three agricultural test sites, Maccarese, Grosseto, and Jolanda di Savoia, in Italy (Figure 1a), and a site in Iran (Figure 1b) were selected as case studies for crop mapping.

The Jolanda di Savoia farm is in northeast Italy (latitude 44°52′59″N, longitude 11°58′48″E), with an altitude of about 1 m.a.s.l. and mean annual precipitation of 691 mm. This region has a warm and temperate climate with an annual mean temperature of 13.6 °C [50]. The soils mostly show clayey and silty textures. The area belongs to the Bonifiche Ferraresi S.p.A., the largest farm in Italy, encompassing about 3850 ha (Figure 1c). The sowing time of winter cereals varies between October and December, while for summer crops it varies between April and June

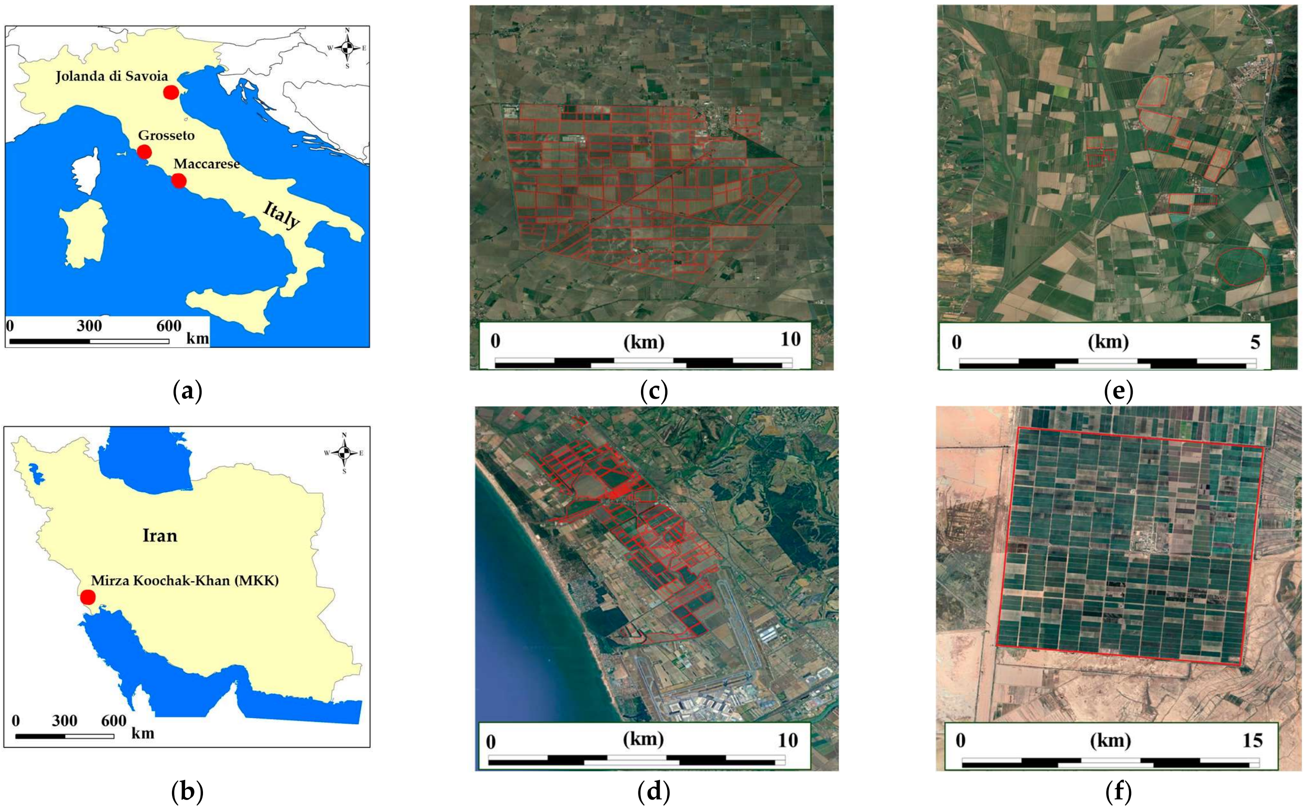


Figure 1. Location of study areas in (**a**) Italy and (**b**) Iran, and field boundary of (**c**) Jolanda di Savoia, northeast Italy, (**d**) Maccarese, central Italy, (**e**) Grosseto, central Italy, and (**f**) MKK southwest Iran.

The Maccarese S.p.A. farm (latitude 41°52′18″N, longitude 12°14′05″E), with an altitude of 8 m.a.s.l. and a long-term average annual precipitation of 812.9 mm, is in central Italy near Rome (Figure 1d). The farming area has soil texture ranging from sandy to clay loam [50]. The area has a typical coastal Mediterranean climate (hot summer Mediterranean climate, Csa in Köppen classification), with an average minimum temperature in the winter of 5.0 °C and an average maximum temperature in summer of 27.4 °C. The rainiest seasons are autumn and winter, and the sowing time for winter cereals varies between October and December, while for summer crops it varies between April and June, according to farm management needs and precipitation patterns [50].

The Grosseto test area (Figure 1e) is in central Italy (latitude 42°50′19″N, longitude 11°02′18″E, altitude 10 m.a.s.l.) and has a long-term average annual precipitation of 655 mm. The Grosseto area has a Mediterranean climate with very mild wet winters and very hot dry summers.

The Mirza Koochak-Khan (MKK) farming and industrial lands (Figure 1f) are in the Khuzestan province in southwest Iran (30°55′15″N; 48°15′35″E) with a mean altitude of 5 m.a.s.l., and a long-term average annual precipitation of 266 mm. The MKK soil is mostly characterized by silty clay loam and clay loam textures.

2.2. Overview of the Implemented Crop Mapping Procedure

The block diagram in Figure 2 shows the flowchart designed to map crop types using PRISMA and Sentinel-2 images. Data gathering, including ground truth data collection and satellite image acquisition, is described in Section 2.3. The pre-processing and processing of data is described in Sections 2.4.1 and 2.4.2, respectively. The classification algorithms and the TR/ACC (TRaining and ACCuracy assessment) procedure are described in Sections 2.5 and 2.6.

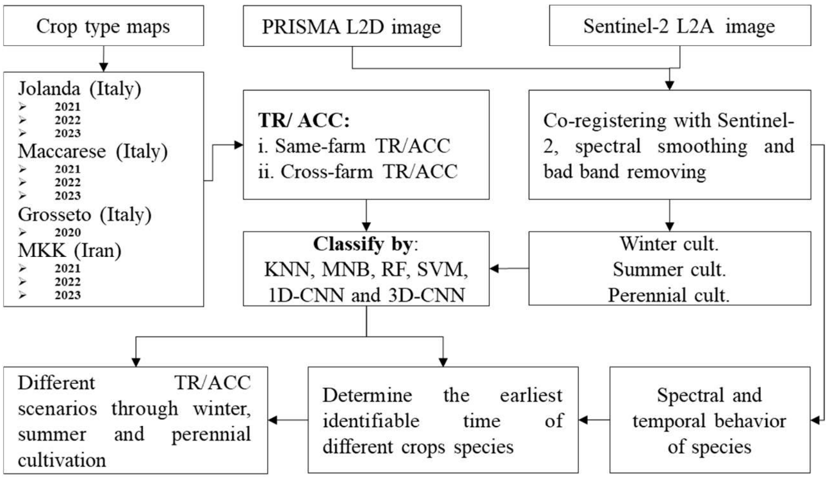


Figure 2. Flowchart of crop mapping by Sentinel-2 and PRISMA satellite images via ML/DL method.

2.3. Data Collection

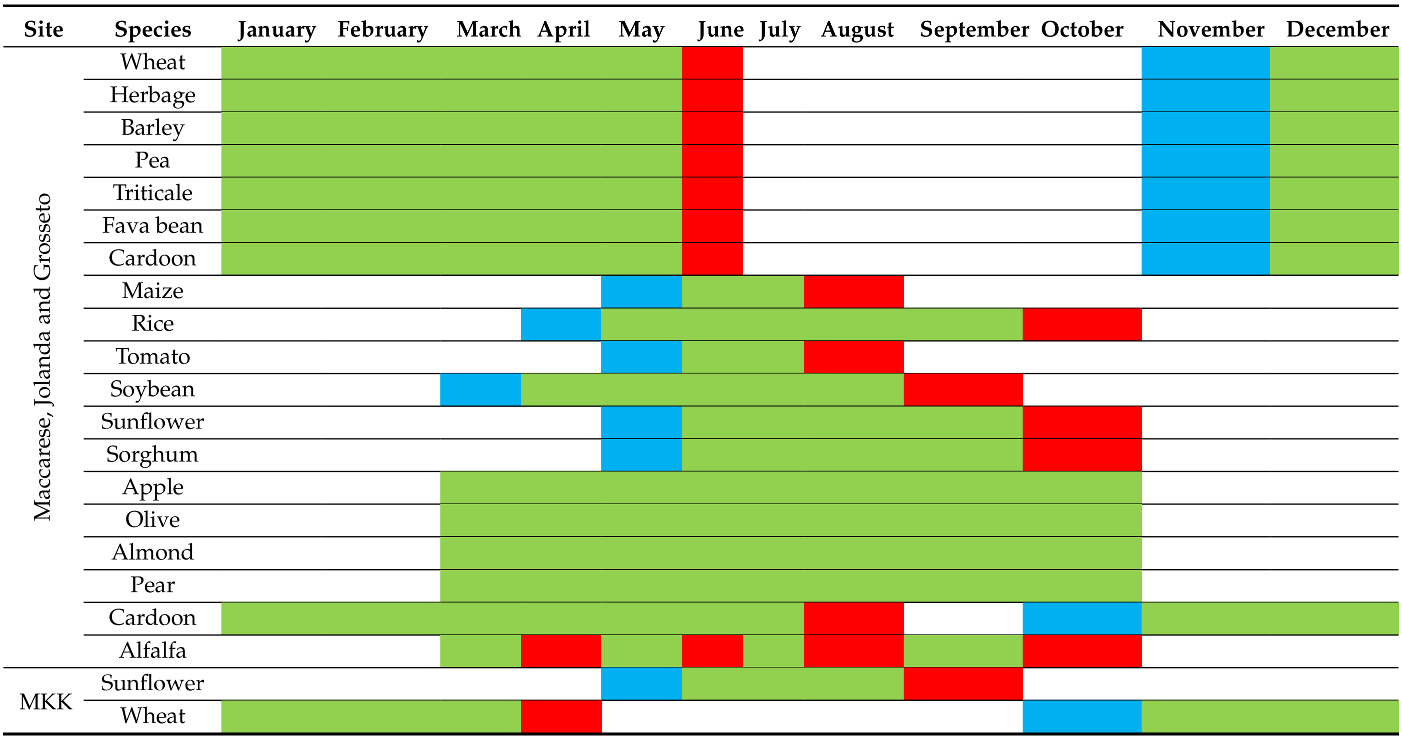
2.3.1. Ground Reference Data

The ground truth data, including the type of crop grown in the farms, were provided directly by field campaigns, while the 2020 crop map of the Grosseto site was extracted from the last published research by Spiller et al. [16]. MKK site crop maps for 2021–2023 were provided by Mirza Koochak-Khan farming and industrial company. The common species of this site among the other sites were wheat and sunflower. Some field photos of field campaigns in the Maccarese and Jolanda di Savoia farms of wheat and maize at different phenological stages from leaf development to ripening are shown in Figure 3.

Figure 3. Field photos for maize and wheat at different phenological stages in the Maccarese and Jolanda di Savoia farms of Italy.

The crop calendar of the investigated crops in the three sites located in Italy (Maccarese, Jolanda, and Grosseto) and in the MKK site is shown in Table 1 to briefly provide an outlook of the development stage of each crop in correspondence with the PRISMA and Sentinel-2 acquisitions.

Table 1. Crop calendar of different crops in Maccarese, Grosseto, Jolanda di Savoia, and MKK sites. The planting, growth, and harvesting stages are shown in blue, green, and red colors, respectively.



In Italy, durum (*Triticum durum* Desf.) and winter wheat (*Triticum aestivum* L.), herbage, barley (*Hordeum vulgare* L.), pea (*Pisum sativum* L.), triticale (x *Triticosecale Wittmack*), and fava bean (*Vicia faba* var. equina L.) are sown in October/November and grow during winter and spring with harvest between May and July depending on the species. These species were considered here as winter crops. Maize (*Zea mays* L.), rice (*Oryza sativa* L.), tomato (*Solanum lycopersicum* L.), soybean (*Glycine max* L.), sunflower (*Helianthus annus* L.), and sorghum (*Sorghum bicolor* Moench) are planted from March to June and harvested from August to October depending on the species. These species were considered as summer crops. Apple, almond, and pear are perennial trees actively vegetating from March until November or December, while olive is an evergreen species. Cardoon (*Cynara carduncculus* L. var. *Altilis*) is a perennial crop, starting to grow in October, and is harvested in August. Alfalfa (*Medicago sativa* L.) starts to grow in March, and it is harvested approximately every 45 days (4 times per each growing season). These species are considered perennial crops.

Table 2 reports the list of the ground truth crop data for Maccarese, Grosseto, Jolanda di Savoia, and MKK sites for 2020–2023 winter, summer, and perennial cultivations that are available along the time frame covered by PRISMA and Sentinel-2 acquisitions. Different test sites and different crops have been used within two training and validation scenarios: (i) same-farm TR/ACC and (ii) cross-farms TR/ACC. For the cross-farm TR/ACC, the dataset was divided into two groups: Group A (MKK, Grosseto, and Maccarese farms, shown with green color bar) and Group B (Jolanda di Savoia farm, shown with red color bar). For the cross-farm TR/ACC scenario, in order to cover all relevant species, the 2021 peas, sorghum, pear, and apple species growing in Jolanda di Savoia (Group A) and the 2022 triticale, olive, almond, and cardoon species growing on the farms in Maccarese (Group B), not being cultivated annually in the other group, were assumed to be cultivated in the other group. These data are shown by diagonal borders ($\times$) in Table 2.

Table 2. Availability of ground truth crop data of MKK, Grosseto, Maccarese, and Jolanda di Savoia sites for 2021–2023 winter, summer, and perennial cultivations. Group A is shown in green color, Group B is shown in red color. Species not being cultivated annually in the other group are shown by diagonal borders (×).

Cult.	Species		Grosseto	MKK			Maccarese			Jolanda		
			2020	2021	2022	2023	2021	2022	2023	2021	2022	2023
Winter	Wheat	D *										
		C *										
	Herbage											
	Barley											
	Pea									×		
	Triticale							×				
	Fava bean											
Summer	Maize											
	Rice											
	Tomato											
	Soybean											
	Sunflower											
	Sorghum									×		
Perennial	Apple									×		
	Olive							×				
	Almond							×				
	Pear									×		
	Cardoon							×				
	Alfalfa											

* D—durum, C—common.

2.3.2. Satellite Imagery

PRISMA is a satellite hyperspectral push-broom sensor with a GSD of 30 m, a swath of 30 km, and a spectral resolution better than 12 nm in the spectral range from 400 nm to 2500 nm. Sentinel-2 is an operational wide-swath (290 km), high-resolution (10 m), and multispectral (13 bands) imaging satellite constellation with a 5-day revisit frequency. Based on the crop calendar (Table 1) and crop ground truth occurrence maps (Table 2), the dataset used for this study was divided into three cultivations consisting of winter, summer, and perennial crops. This study applied PRISMA hyperspectral images for all the sites, and 12 images from Jolanda di Savoia, 11 images from Maccarese, and 15 images from MKK were available. Moreover, the closest (±6 days to each PRISMA acquisition) Sentinel-2 cloudy free images (11 bands resampled to the 30 m spatial resolution) were used for reference and for classification comparison. A summary of all the imagery used is reported in Table 3, where the PRISMA images used for Group A and Group B are shown in green and red color, respectively. For the cross-farm TR/ACC scenario, to cover all the relevant species cultivated exclusively in Group A or Group B, the images highlighted in bold were assumed as if they were acquired in the other group, as similarly indicated with × in Table 2.