

Comparing Sentinel-1, Sentinel-2, and Landsat-8 data in the early recognition of irrigated areas in central Italy

Lorenzo Vergni,¹ Alessandra Vinci,¹ Francesca Todisco,¹ Francesco Saverio Santaga,² Marco Vizzari¹

¹Department of Agricultural, Food and Environmental Sciences, University of Perugia, Perugia; ²Institute of BioEconomy (IBE), National Research Council (CNR), Florence, Italy

Abstract

This study evaluated the effectiveness of various remote sensing (RS) data (Sentinel-1, Sentinel-2, and Landsat 8) in the early recognition of irrigated areas in a densely cultivated area of central Italy. The study was based on crop data collected on more than 2000 plots in 2016 and 2017, characterized by quite different climatic conditions. The different RS data sources were used both alone and combined and with precipitation to define corresponding random forest (RF) classifiers whose overall accuracy (OA) was assessed by gradually increasing the number of available features from the beginning of the irrigation season. All tested RF

classifiers reach stable OAs (OA 0.9) after 7-8 weeks from the start of the irrigation season. The performance of the radar indexes slightly improves when used in combination with precipitation data, but three weeks of features are required to obtain OA above 80%. The optical indices alone (Sentinel-2 and Landsat 8) reach OA \approx 85% in the first week of observation. However, they are ineffective in cloudy conditions or when rainfed and irrigated fields have similar vigour. The most effective and robust indices are those based on combined sources (radar, optical, and meteorological), allowing OAs of about 92% and 96% at the beginning and in the middle of the irrigation season, respectively.

Correspondence: Lorenzo Vergni, Department of Agricultural, Food and Environmental Sciences, University of Perugia, Borgo XX Giugno 74, 06121 Perugia, Italy.
Tel.: +39.0755856041. E-mail: lorenzo.vergni@unipg.it

Key words: E-OBS grid precipitation; irrigation; maize; random forest; rainfed crops; tobacco.

Acknowledgements: the authors wish to thank Dr. Tommaso Cesaretti for his support in analysing the output of the classification algorithms.

Funding: RICERCA DI BASE 2019 - Università degli Studi di Perugia.

Contributions: LV, AV, FT, FSS, MV, conceptualization and supervision; LV, MV, FSS, software, methodology, formal analysis, data curation; LV, MV, writing original draft; LV, MV, AV, FT, review and editing. All authors have read and agreed to the published version of the manuscript.

Conflict of interest: the authors declare no potential conflict of interest.

Data availability statement: grid precipitation data are openly available from the public E-OBS database through the European Copernicus Climate Change Service website. All satellite data used in the research are openly available in the USGS archives (L8), the Copernicus Open Access Hub (S1, S2), and Google Earth Engine (S1, S2, L8).

Received for publication: 25 August 2021.

Accepted for publication: 21 October 2021.

© Copyright: the Author(s), 2021

Licensee PAGEPress, Italy

Journal of Agricultural Engineering 2021; LII:1265

doi:10.4081/jae.2021.1265

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 4.0) which permits any non-commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Introduction

In many areas of the world, agriculture is strictly dependent on irrigation, ensuring the possibility of increasing and stabilizing productivity even in the driest years (Rosegrant and Cline, 2003). Furthermore, some estimates indicate that irrigation contributes to more than 70% of freshwater withdrawals globally (Shiklomanov, 2000; Cai and Rosegrant, 2002), and in future scenarios, given the growing population and global warming, further increases of the irrigation water use are expected (Rockstrom *et al.*, 2012; Li and Troy, 2018). In this context, it is increasingly important to develop methods for sustainable planning and management of water use in agriculture, an objective that can only be achieved by improving knowledge both on actual irrigation water volumes and on the extent of irrigated areas (Jalilvand *et al.*, 2019; Dari *et al.*, 2020, 2021).

The recent increase of remote sensing (RS) data availability in spectral, spatial, and temporal resolution, combined with the decreasing acquisition and processing costs and the growing availability of machine learning algorithms in many processing environments, have boosted RS research applications. Satellite RS is less costly and time-consuming than traditional surveys and offers excellent potential for mapping or monitoring irrigation practices. Moreover, RS data is more easily integrable with other data sources into modern geographic information systems (GIS) for combined spatial and temporal analyses. Thus, RS spatial information about irrigated lands is very relevant for sustainable resource planning and management. It can provide timing and prioritizing water delivery, assess irrigation performance, quantify the environmental impact of irrigation practices, assess irrigation water use, and identify changes in space and time (Ozdogan *et al.*, 2010).

Studies on irrigation mapping using RS data are still not common considering the relevance of irrigation water use issues. However, synthetic-aperture radar (SAR), optical, and weather data, alone or in combination, have already been applied to irrigation assessment and mapping at the field scale. Most recent studies about this topic rely on RS data and machine learning classifiers

to map irrigated areas. For example, Gao *et al.* (2018) used Sentinel-1 (S1) SAR data and a random forest (RF) classifier on a 400 km² area located in Catalonia, Spain, obtaining an accuracy of around 82%. Using various classification algorithms in the same region, Bazzi *et al.* (2019) obtained accuracies in the range 89-94 % and 89-92% from S1 and Sentinel-2 (S2) data, respectively, with further slight accuracy increments using both sources simultaneously. In the Adour Amont watershed, France, the combined use of radar (S1), optical (S2), and rainfall data improved the accuracy of the classification obtained using the RF classifier on each data separately (overall accuracy, OA≈0.7) (Pageot *et al.*, 2020). S1 and S2 data and a combination of support vector machine (SVM) and decision tree classifiers were used in the semi-arid region of Kairouan plain, in central Tunisia - North Africa (Bousbih *et al.*, 2018). This study obtained the best results with classifications based on soil moisture indices only, with an overall accuracy of 77%. Dari *et al.* (2021), comparing in North-East Spain different RS soil moisture products [*i.e.*, soil moisture and ocean salinity (SMOS) at 1 km, soil moisture active passive (SMAP) at 1 km and 9 km, Sentinel-1 at 1 km, and advanced SCATterometer (ASCAT) at 12.5 km], found that SMAP is the most effective in detecting irrigated areas (OA=78%).

Landsat data were successfully applied in Google Earth Engine (GEE) within a semi-automatic training approach, based on RF, to produce an irrigated croplands map of the United States (US) at 30-m resolution with an overall accuracy of 94% (Xie *et al.*, 2019). Landsat data and the RF classifier were also used in GEE for mapping decades of annual irrigation across the US High Plains Aquifer, obtaining an overall accuracy of 91.4% across years (Deines *et al.*, 2017, 2019).

Other studies have used RS data to model or estimate crop water

requirements, particularly considering RS's recognized potential to produce spatial information about evapotranspiration (Calera *et al.*, 2017). The normalized difference vegetation index (NDVI) derived from S2 was used in a study area of Central Italy to estimate the actual evapotranspiration of irrigated croplands in a simplified water balance predicting the irrigation water (IW) (Maselli *et al.*, 2020b). The results show that over 50% and 70% of the measured IW variance was explained daily and weekly, with mean bias errors below 0.3 mm/day and 2.0 mm/week. S2 data were applied to assess crop evapotranspiration and irrigation water requirement (IWR) in standard conditions in a tomato field in Central Italy (Vanino *et al.*, 2018). The results showed that canopy growth, maximum evapotranspiration, and IWR were accurately estimated from satellite observations following seasonal rainfall and air temperature data. Landsat-8 and RapidEye satellite data, combined with ground-based meteorological data, were successfully used to obtain estimates of evapotranspiration and crop coefficients of table grapes vineyards in the Apulia region (Italy) (Vanino *et al.*, 2015).

However, to our best knowledge, very few studies have compared the effectiveness in the early identification of irrigated fields of SAR and optical data, separately or in their combinations, together with rainfall data. Indeed, early identification of these areas could be very relevant for timely crop irrigation requirements, irrigation monitoring, and sustainable water resources management at the watershed level.

This study analyses the performance of various indices derived from SAR (S1) and optical (S2 and Landsat 8) satellite data, also combined with rainfall data, in the early identification of irrigated fields (*i.e.*, from the start of the irrigation season). Moreover, the variability of the accuracy according to the different crop types was evaluated and discussed.

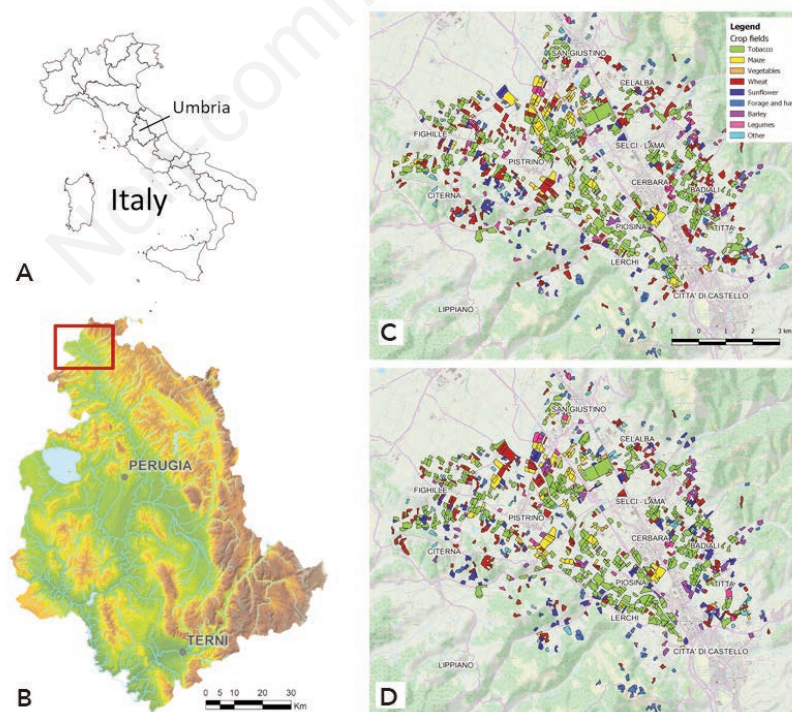


Figure 1. Location of the study area (A, B) and spatial distribution of crop fields considered in the research during years 2016 (C) and 2017 (D).

Materials and methods

Study area

The study area is the Upper Tiber River valley in the Umbria region, central Italy (Figure 1). This region has a typical Mediterranean climate, summers are warm to hot, and winters are generally mild but colder at higher altitudes. The wettest season is autumn, and the driest is summer. In hilly areas, tree crops (olives and vines) and forage crops prevail, while other herbaceous crops (e.g., winter cereals, maize, tobacco, sunflower, grain legumes, and vegetable crops) are widespread in the plains. Most summer crops (e.g., maize, vegetables, and tobacco) require irrigation due to the considerable imbalance between precipitation and evapotranspiration. This work considered the irrigation season between the last week of May (22nd week of the year) and mid-September (38th week of the year). As it is shown in Figure 1, the study area is fragmented as it is characterized by small agricultural plots and small agricultural plots.

Crop data

For the study, we first collected information about the crops grown in the study area. Then, upon request, the producers' associations provided data on the Common Agricultural Policy (CAP) declarations in 2016 and 2017. This made it possible to determine the crops of each cadastral parcel in the two years indicated. To obtain greater accuracy, considering the spatial resolution of RS data used in this research, we selected only the parcels with an area greater than 1 ha and covered by a single crop for at least 90%. Furthermore, we only considered the parcels with annual herbaceous crops. The final database consisted of 2050 fields, 984 in 2016 and 1066 in 2017 (Figure 1).

Given the lack of actual ground truth about these 2050 fields, the presence of irrigation was indirectly defined based on the crop type, assuming the classifications indicated in Figure 2 (which also shows the crop frequency distribution). This classification was determined based on the information collected from the producers'

associations and the Regional Forestry Agency (AFOR), the primary irrigation water provider in the area. This made it possible to exclude the presence of tobacco, maize, and vegetables in non-irrigated conditions in the study area. Furthermore, for tobacco, the production regulation imposes irrigation to reach the required qualitative standards. As for non-irrigated crops, AFOR data indicate that only 2% of the areas of the supply contracts concerns crops such as sunflower, winter cereals, forage, and legumes; therefore, these can be considered non-irrigated, with a minimum margin of error. Of course, the lack of actual ground truth about irrigation introduces possible errors, which are difficult to quantify. However, taking into account that in the area, there is an almost univocal correspondence between crop type and irrigation supply, the errors should be modest and not such as to significantly affect the results of the comparative analysis.

The typical growing season and the sowing and harvesting intervals of the main crops are shown in Table 1.

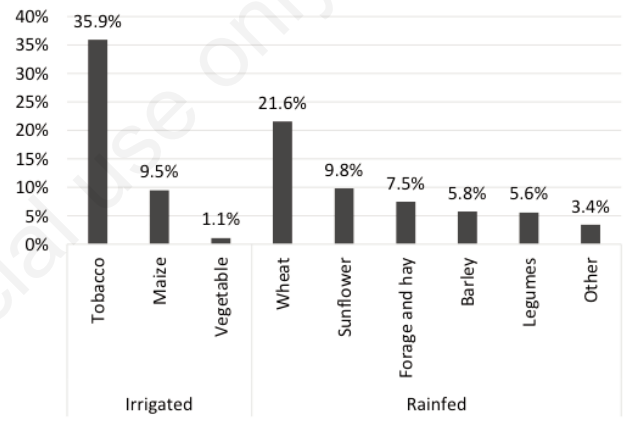


Figure 2. Percentage of the different crops in the study area in the period 2016-2017.

Table 1. Shaded cells indicate the typical growing season of the main crops cultivated in the study area. S and H letters indicate the periods during which seeding or transplanting (S) and harvesting (H) are most likely.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tobacco						S	S					
Maize				S	S			H	H	H		
Vegetables					S	S		H	H	H		
Wheat							H	H			S	S
Sunflower				S	S			H	H			
Forage-hay*					H	H	H	H	H	H		
Barley							H	H			S	S
Legumes						H	H				S	S

Table 2. Optical indices and related formulas used in the research.

Index	Formula	Author
NDVI	$\frac{NIR - RED}{NIR + RED}$	Rouse <i>et al.</i> , 1974
MSAVI	$\frac{(2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - RED)})}{2}$	Qi <i>et al.</i> , 1994a, 1994b

Meteorological data

Spatial information on the weather data of the study area was derived from the E-OBS daily gridded meteorological data (Cornes *et al.*, 2018). To implement this database, time series from European national weather services are collected, processed with extensive quality-check and blending procedures, and finally transformed into a gridded daily dataset containing several atmospheric variables onto a $0.1^\circ \times 0.1^\circ$ regular grid. E-OBS data were downloaded through the European Copernicus Climate Change Service Website. Figure 3 shows the average daily temperature and precipitation for the study area between 01/05/2016 and 30/09/2017, derived from the E-OBS data. Shaded areas indicate the periods corresponding to the irrigation seasons 2016 and 2017. The long-term mean precipitation amount during the irrigation season is about 210 mm. In 2016 it was 228 mm, and in 2017 only 113 mm. For classification purposes, we used only precipitation data aggregated at a weekly time scale and associated with each field considered.

Remote sensing data

We used Sentinel-1 (S1) SAR data and optical data from both Sentinel-2 (S2) and Landsat 8 (L8). We accessed these data and calculated the SAR and spectral indices on a field basis in the GEE platform. The two S1 satellites (A and B), launched by the European Space Agency (ESA), provide data using a dual-polarization C-band SAR sensor with a spatial resolution of 10, 25, or 40 meters and a temporal resolution of 6 days. Within GEE, S1 ground range detected (GRD) scenes are pre-processed using the S1 Toolbox (thermal noise removal, radiometric calibration, terrain correction) to generate a calibrated, ortho-corrected product. We used the single co-polarization, vertical transmit/vertical receive (VV) and the single co-polarization, horizontal transmit/horizontal receive (VH) bands at 10 meters resolution and the VH/VV ratio. VH, VV, and their ratios have been already used for irrigation mapping (e.g., Bousbih *et al.*, 2018; Gao *et al.*, 2018; Pageot *et al.*, 2020). Previous studies confirmed the high explicatory value of the VH/VV ratio, particularly in early phenological stages (Veloso *et al.*, 2017; Vreugdenhil *et al.*, 2018). The ratio partially compensates for the radiometric instability of the sensor and shows higher stability than the single polarization.

The two S2 satellites (A and B) are equipped with a multispectral sensor (MSI), detecting 13 spectral bands, with a spatial resolution ranging of 10, 20, and 60 meters and 4 days of temporal resolution on average. GEE collects S2 data from Copernicus Data Hub in the L1C (orthorectified top-of-atmosphere reflectance), and L2A (orthorectified atmospherically corrected surface reflectance - SR) processing levels. In the L2A level, three quality assessment (QA) bands are included, where one (QA60) is a bitmask band with cloud mask information. We used the L2A images, applying a cloud-filtering, and selecting the B5 (red), B8 (NIR), and B11 (SWIR1) bands.

L8 operational land imager (OLI) provides multispectral images at a 30-meter resolution containing five visible and near-infrared (VNIR) bands and two short-wave infrared (SWIR) bands with a revisit time of 16 days. A 15-meters panchromatic band is available as well. In GEE, L8 data are available in an atmospherically corrected (SR) level which includes a cloud, shadow, water, and snow mask band as well as a per-pixel saturation mask band. In this study, we used the L8 SR images, applying a cloud-filtering, and selecting the B4 (red), B5 (NIR), and B6 (SWIR1) bands. In addition, we used two spectral indices derived from S2 and L8 bands (Table 2).

NDVI is a widely used vegetation index because it directly expresses a measure of vegetation health (Rouse *et al.*, 1974). The combination of its normalized difference formulation and the use of the highest absorption and reflectance regions of chlorophyll make it very explicative over a wide range of conditions. NDVI has already been used in previous studies for irrigation mapping (e.g., Bousbih *et al.*, 2018; Maselli *et al.*, 2020a; Pageot *et al.*, 2020). The modified soil-adjusted vegetation index (MSAVI) (Qi *et al.*, 1994a, 1994b), like its predecessor SAVI (Huete, 1988), seeks to overcome the issues of NDVI related to the soil background effect, especially in areas partially covered by vegetation. This index eliminates the need to find the soil line or even specify the soil brightness correction factor as required for other previous vegetation indices. MSAVI has already been successfully used, with other vegetation indices, for irrigation mapping in olive orchards (Sepulcre-Cantó *et al.*, 2009) and water stress estimation in vineyards (Romero *et al.*, 2018). We chose this index considering the lower sensitivity to the soil background to verify its possible higher reliability than NDVI in irrigation detection, especially during the early crop stages.

Unfortunately, optical data from Sentinel-2 and Landsat 8 in 2016 were characterized by several missing images due to the frequent presence of cloudiness during the irrigation season. Therefore, optical-derived indices were only evaluated for 2017.

Methods

Our methodology relies on a broadly used machine learning algorithm (*i.e.*, RF) applied to various time-series combinations of the above-mentioned SAR and optical indices calculated on a field basis, sometimes combined with rainfall data.

Random forest

RF is a machine learning technique based on many decision trees that work as an ensemble. The starting point for building the RF classifier is a dataset that includes a qualitative response variable (*i.e.*, irrigated or rainfed) and one or more predictor variables (features). The original dataset is randomly subsetted to form the training and the validation datasets. Each decision tree of the forest is obtained from a bootstrap sample of the training dataset and uses only some randomly chosen features during tree growth. The out of bag (OOB) set is the datum not selected in the sampling process of a specific tree. The classification algorithm is based on a majority vote (mode) across decision trees in the classification stage (Breiman, 2001). We applied RF to recognize the irrigated fields

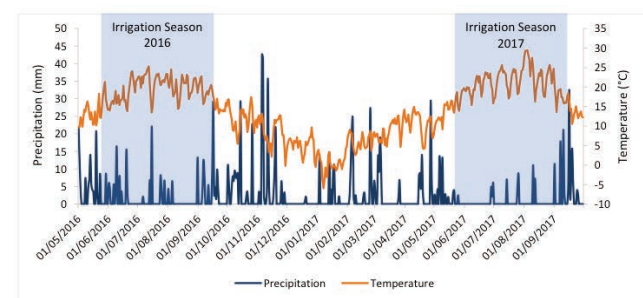


Figure 3. Mean daily precipitation and temperature for the study region in the period 01/05/2016-30/09/2017, including two irrigation seasons (E-OBS meteorological data).

based on different features datasets, as shown in Table 3. The purpose is to assess the capacity of a certain type of dataset to provide early recognition of irrigated fields. We modelled the RF progressively for each dataset, increasing the number of features from the beginning of the irrigation season with weekly steps.

The datasets are of different types: some are single-source (*e.g.*, radar or optical), others have been obtained by combining multiple data sources. In defining the combined datasets, only the best performing single-source datasets were considered.

Each original dataset was randomly split into 50% for training and 50% for validation to improve the robustness of the validation subset. After a preliminary evaluation based on the classification accuracy convergence, the number of trees was set to 5000 for all tested classifiers. Thus, the number of features considered at each split was equal to the square root of the number available features, as usually suggested for classification purposes (Breiman, 2002).

The feature importance was evaluated as the mean decrease in accuracy (MDA) after permutations of each feature. In particular, the MDA is computed averaging (over all trees) the decrease in the OOB accuracy before and after the permutation of each feature. Thus, the features with higher MDA are relatively more important than the others for the accuracy of a specific RF classifier.

The RF analysis was implemented in the R environment, using the functions available in the library *randomForest* (Liaw and Wiener, 2002).

The performance of the RF classifiers, generated on the various training datasets, was evaluated by applying the trained classifiers to the corresponding validation datasets and computing the overall accuracy OA (%), the Producer's accuracy (PA) (%), and the User's accuracy (UA) (%) (Congalton, 1991; Richards and Jia, 2006).

The OA is obtained as:

$$OA = \frac{TC}{N} \quad (1)$$

where TC is the number of truly classified fields, and N is the total number of fields analysed in the validation dataset. The PA, also indicated as precision, is given by:

$$PA = \frac{TC_k}{O_k} \quad (2)$$

where TC_k is the number of fields of category k truly classified and O_k is the number of fields observed in category k. The UA, also indicated as recall, is given by

$$UA = \frac{TC_k}{C_k} \quad (3)$$

where C_k is the number of fields classified in the category k.

Both UA and PA can be quantified for each category (*i.e.*, in this work for both irrigated and rainfed fields). A classifier is considered highly accurate when obtaining a combination of high recall and precision values. In particular, high precision (UA) corresponds to a low representation of commission errors and high recall (PA) to a low value of omission errors (Weaver *et al.*, 2018).

Results

Overall accuracy, user accuracy, producer accuracy

The overall accuracy of the RF classifiers trained on the datasets considered (Table 3) is shown in Figure 4. For completeness of information, the accuracy variations are reported up to the 33rd week. However, it is believed that in the area considered, the recognition of irrigated areas can be defined 'early' if carried out by the end of June (*i.e.*, 27th week).

The analysis in Figure 4 shows a general increase in OA as the number of weeks (*i.e.*, available features) increases. In some cases (*e.g.*, some radar-derived datasets), the increase is gradual and tends to reach an almost steady value around the 28th week. For other datasets (*e.g.*, optical-derived and some combined datasets), the OA is already very high ($OA \geq 85\%$) since the first observation period and rapidly reaches a plateau from as early as the 24th- 25th weeks. For example, the combined datasets 'S2_MSAVI + S1_VV + p' and 'S2_MSAVI + S1_VV' allow obtaining the absolute highest values of OA ($OA \approx 96\%$), followed by 'S2_NDVI' and 'S2_MSAVI', which have OAs close to 95% starting from the 23rd and 24th weeks.

In terms of early evaluation, the best performance is obtained with the combined dataset 'S2_MSAVI + S1_VV + p', which allows an OA of 91.5% since the first week. Other noteworthy results are the 'S2_MSAVI' and 'S2_NDVI' datasets, whose initial

Table 3. List of the datasets (single-source and combined) and the corresponding number of available features (depending on the available years) and observations.

Dataset	Type of data	Years	Number of features (weeks)	Observations (fields)
S1_W	Radar	2016-2017	From 1 to 17	2050
S1_VH	Radar	2016-2017	From 1 to 17	2050
S1_VH/S1_WV	Radar	2016-2017	From 1 to 17	2050
S2_MSAVI	Optical	2017	From 1 to 17	1066
S2_NDVI	Optical	2017	From 1 to 17	1066
LS_MSAVI	Optical	2017	From 1 to 7	1066
LS_NDVI	Optical	2017	From 1 to 7	1066
S1_W + p	Radar and rainfall	2016-2017	From 2 to 34	2050
S2_MSAVI + S1_WV	Optical and radar	2017	From 2 to 34	1066
S1_VH/S1_WV + p	Radar and rainfall	2016-2017	From 2 to 34	2050
S2_MSAVI + S1_WV + p	Optical, radar, and rainfall	2017	From 3 to 51	1066

S1_w, Sentinel-1 data in vertical-vertical polarization; S1_vh, Sentinel-1 data in vertical-horizontal polarization; S2_NDVI, NDVI derived from Sentinel-2 data; MSAVI, derived from Sentinel-2 data; LS_MSAVI, MSAVI derived from Landsat 8 data; LS_NDVI, NDVI derived from Landsat data; p, E-OBS rainfall data.

OA is about 88%. Similar results are also obtained from the Landsat-derived optical data, even if, in this case, the first survey is one week later due to the longer revisit time. The ‘S1_VH’ dataset has the worst performance, with an OA of only 58% in the first week.

The VV-type radar data is slightly more performing than the VH one, especially in the early classification. As expected, the ratio of the two types of backscattering (‘S1_VH / S1_VV’) allows obtaining a significant increase in both the early and medium-late accuracy. All the radar-derived datasets seem to be positively affected by the presence of additional precipitation features. This is clearly demonstrated by comparing the OAs of the datasets ‘S1_VH / S1_VV + p’, ‘S1_VV + p’, and ‘S2_MSAVI + S1_VV + p’ with those of similar datasets not including precipitation.

In general, optical-derived datasets perform better than radar-derived ones, both in terms of early and mid-late classification.

The results of the analysis of the feature importance are shown only for the S2_MSAVI, ‘S1_VH / S1_VV’, and ‘S2_MSAVI + S1_VV + p’ datasets, as they are those that allow obtaining the best results in the respective categories (optical, radar, and combined). Figure 5 shows the MDA of the features involved in those datasets in the early irrigation season (*i.e.*, between the 22nd and the 27th week). For the M2_SAVI dataset, there is an evident predominance of the features detected in the first two weeks (22nd and 23rd). Also, the first two weeks of observation for the radar dataset ‘S1_VH / S1_VV’ are important, but the importance is more evenly distributed than that of the optical dataset. Finally, for the combined dataset, greater importance assumes the optical features (particularly those related to the first two weeks) followed by the radar and rainfall features.

Further details on the performance of the different classifiers are shown in Figure 6, which shows the 1:1 plots of UA and PA for both the rainfed and irrigated categories. For simplicity’s sake, only the values relating to the 22nd and 27th weeks are shown: indeed, the 22nd week corresponds to the minimum number of features (and therefore OA) for any dataset (Figure 4); the 27th week, instead, corresponds to the period in which almost all datasets have

reached a stable OA.

Figure 6 indicates that the differences between UA and PA are always moderate (<10% at the 22nd week and <5% at the 27th week). The only exception appears to be the ‘S1_VV + p’ dataset at the 22nd week for irrigated plots, whose PA is about 20 percentage points higher than the UA. Thus, in general, the UA and PA do not seem to be particularly affected by the irrigation regime.

Finally, to evaluate whether the performances of optical and radar datasets could depend on the different number of data considered (Table 3), we analysed the performance of the RF classifiers based on radar data, using only the data collected in 2017. The results (not shown) were very similar to those obtained with the entire available database.

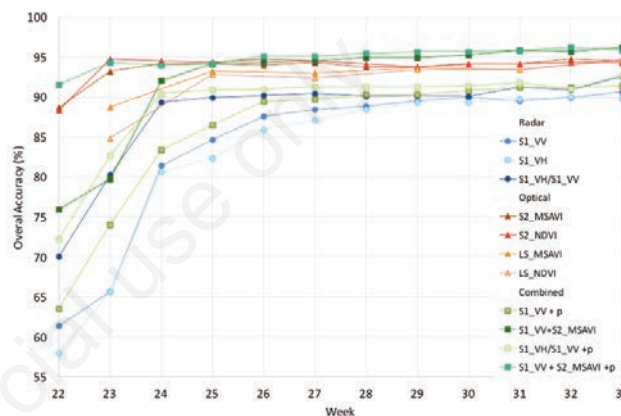


Figure 4. Overall accuracy (%) of the random forest classifiers for the different datasets composed of an increasing number of available features (*i.e.*, weeks). See Table 3 for the description of the dataset codes.

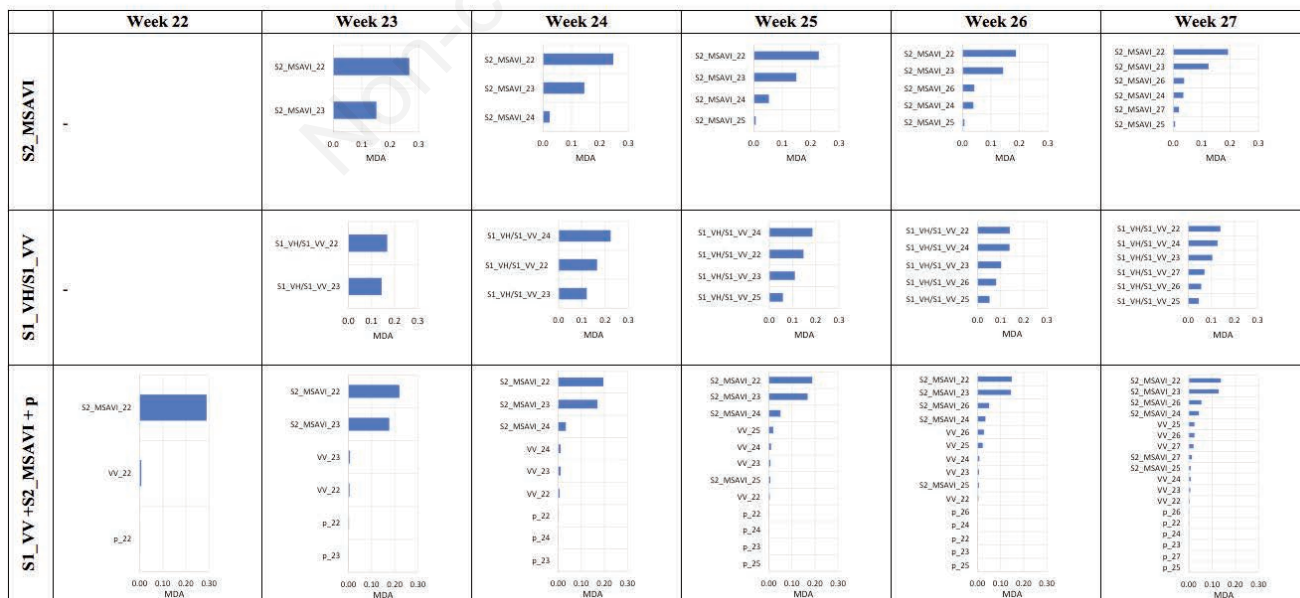


Figure 5. Variable importance plots for the features considered in some relevant classifiers at different weeks of the early irrigation season. The mean decrease accuracy (MDA) of the classifiers based on a single feature is not shown.

Results by crop type

To evaluate the variability of the overall accuracy with the crop type, comparisons were made between the accuracies found in tobacco and corn (the two main irrigated crops, Figure 2) and between winter wheat and sunflower (the two main rainfed crops, Figure 2).

For this particular purpose, Equation (1) was applied using as the denominator the total number of fields of the specific crop and as the numerator the number of plots of the same crop correctly classified as irrigated (for tobacco and maize) or non-irrigated (for wheat and sunflower).

The 1:1 plots of Figure 7 summarize the results of this analysis: on the left (Figure 7A-C), there are the comparisons between

the accuracies obtained for maize and tobacco from the different datasets in three salient phases of the irrigation season (22nd week, 27th week, 33rd week); on the right side (Figure 7D-F), the same information is given for the comparisons sunflower vs. wheat.

Concerning the comparison between tobacco and maize (Figure 7A-C), the analysis shows significant differences in the OAs obtained with the same dataset. In the first week, the datasets, including radar data, alone or combined with precipitation, provide slightly better OAs for maize than tobacco. However, the differences are relatively modest and reach a maximum of 10 percentage points for the ‘S1_VV + p’ dataset. The datasets, including S2 data, are generally much better performing for tobacco, with differences of up to 30 percentage points (e.g., ‘S2_MSAVI’ and ‘S2_MSAVI

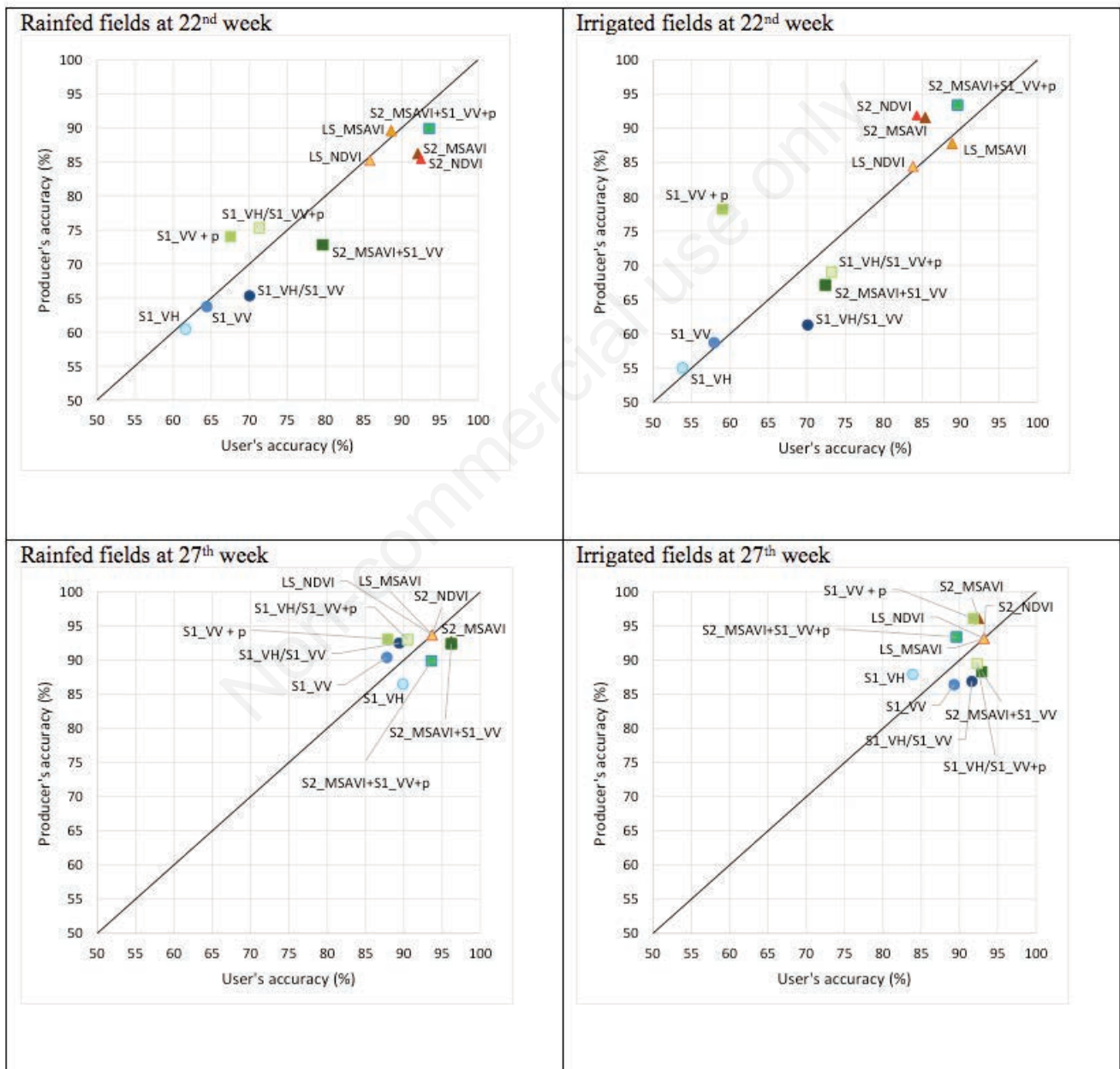


Figure 6. User’s and producer’s accuracies of the models considered for rainfed and irrigated fields based on the features available at the beginning of irrigation season (22nd week) and after 6 weeks (27th week). For Landsat data, the first available week is the 23rd. See Table 3 for the description of the dataset codes.

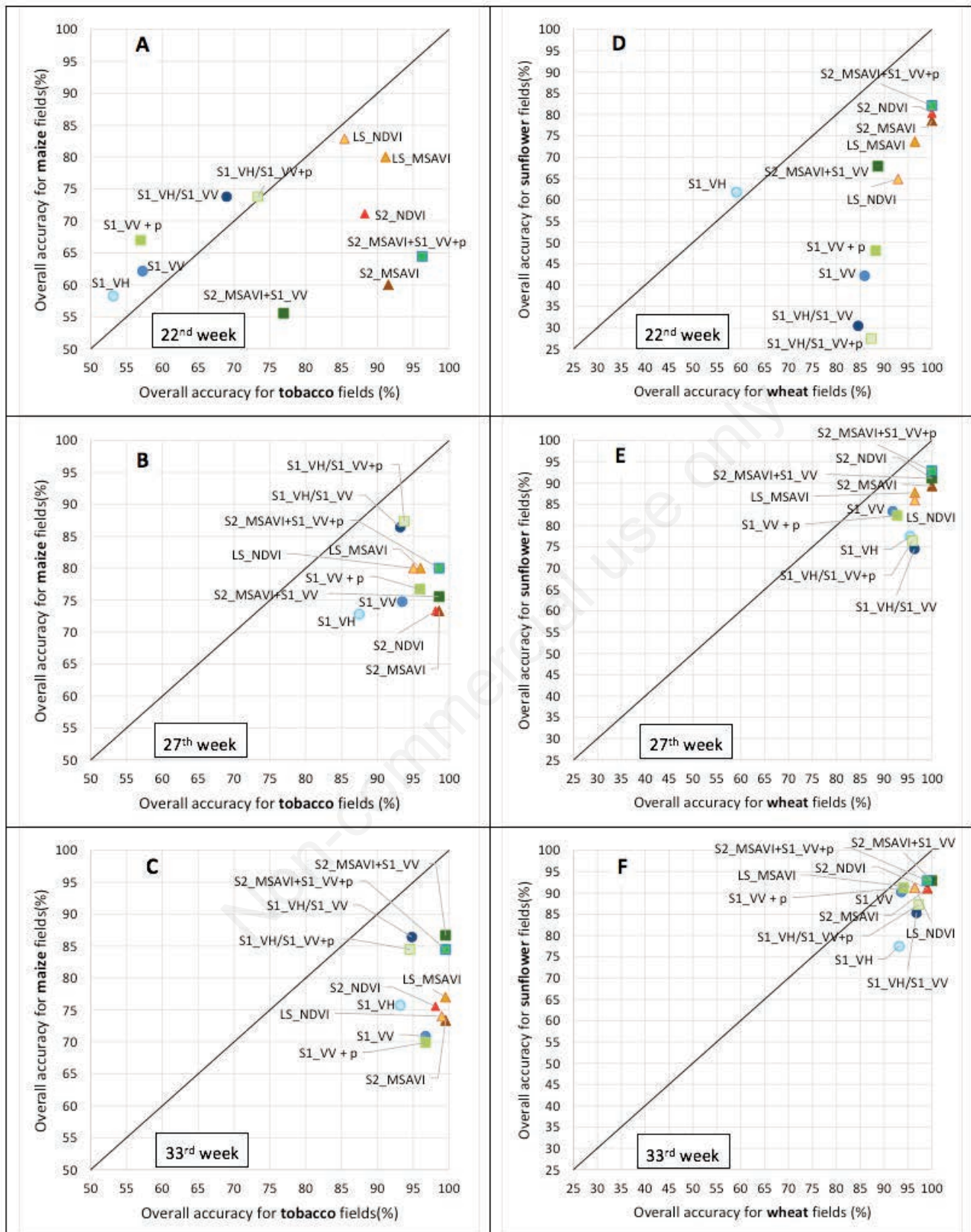


Figure 7. A-F) 1:1 plots showing the comparison of the overall accuracy in recognition of irrigated and rainfed fields for the two main irrigated crops (tobacco and maize) and the two main rainfed crops (wheat and sunflower) in three phases of the irrigation season. For Landsat data, the first available week is the 23rd. See Table 3 for the description of the dataset codes.

+ S1_vv + p'). As the irrigation season progresses and, therefore, with the increase in the number of available features, it is observed that all the datasets provide better performance for tobacco than for maize. For tobacco, many models (e.g., 'S2_MSAVI + S1_VV'; 'S2_MSAVI + S1_VV + p'; 'LS_MSAVI'; 'LS_NDVI'; 'S2_MSAVI'; 'S2_NDVI') reach OAs close to 100%; for maize instead, the best models do not exceed 87% OA.

Concerning the comparison between sunflower and wheat, wheat fields are generally correctly recognized as rainfed more frequently than sunflower fields (Figure 7). The most relevant differences are observed at the beginning of the irrigation season (22nd week, Figure 7D). In this period, most of the radar-derived datasets have OAs lower than 50% for sunflower and higher than 85% for wheat. Relevant differences (between 20 and 30 percentage points) are also observed for the other datasets. The only exception is the S1_VH dataset, for which the OA is around 60% for both wheat and sunflower. With the increase in the number of features (progress of the irrigation season), the differences in accuracy between the two crops become more and more modest, although non-irrigated wheat-type plots continue to be recognized with greater accuracy than those with sunflowers.

Discussion

All the tested datasets reach maximum OA as early as the middle of the irrigation season (i.e., 30th week of the year or second half of July, Figure 4). In this period, the optical-derived datasets have OAs that stand at an average value of 94% against 90% of the radar-derived datasets. Furthermore, all the models are balanced, with modest differences between the UA and PA (Figure 6).

In general, the OAs obtained in this analysis are slightly higher than those found in most previous studies that dealt with similar applications. For example, Pageot *et al.* (2020), based on cumulative data (radar, optical, and weather), indicated OA≈70%; Dari *et al.* (2021) and Gao *et al.* (2018), based only on microwave data, obtained OAs of about 78% and 82%, respectively. Our OAs are somewhat similar to those indicated by Bazzi *et al.* in Catalonia, although in that case, the SAR and optical data accuracies are respectively about 94 and 92%, respectively. Indeed, having considered a relatively small study area and selected agricultural parcels may have increased the overall accuracy.

Landsat optical data prove to be as reliable as Sentinel data, although longer revisit times can make identification less timely and effective. Moreover, the accuracy's difference obtained from two optical indices considered (NDVI and MSAVI) is generally negligible.

The use of combined databases (SAR + rainfall, or SAR + optical, or SAR + optical + rainfall) showed the best performance in all the tests obtaining OAs close to 96% (Figure 4). This result agrees with the previous works found in the literature, highlighting the advantage of the combined use of rainfall and SAR data and SAR and optical data (Bazzi *et al.*, 2019; Pageot *et al.*, 2020). Also, in the case of S1 radar data, following Pageot *et al.* (2020), the advantage of using the ratio of VV and VH polarizations is noted. Furthermore, the combination of SAR and weather information (in this case, precipitation) allows appreciable increases in the OA, particularly in the early irrigation season. In this regard, precipitation features at a spatial resolution finer than the E-OBS grids could further enhance the classification accuracy, but at present, the E-OBS dataset represents the most detailed gridded precipita-

tion product available for the whole Italian Peninsula.

Regarding the early recognition of irrigated areas, there are relevant differences between the datasets considered (Figure 4). In this regard, few previous works have dealt with the analysis of these differences. An example can be found in Bazzi *et al.* (2019), which report increases in OA similar to those shown in Figure 4, underlying, as in our case study, the advantage of combined data sources in early recognition. However, the most evident result is the clear superiority of the datasets based only on optical data compared to data based only on radar data (Figure 4). This was also confirmed by the variable importance analysis (Figure 5), which shows the relatively higher importance of the early optical features. This behaviour can be attributed to the vegetational characteristics of the crops present in the study area in the different phases of the irrigation season. In fact, at the beginning of irrigation season (i.e., 22nd week, end of May), irrigated crops are at the very early growth stage (Table 2), with most of the field surface represented by bare soil.

Conversely, most non-irrigated crops (winter cereals, forage, and hay) are in full development. In this situation, the optical indices allow obtaining high accuracy even from a single feature based on the vegetative vigour (Figure 5). In the late irrigation season, the recognition by the optical indices is still accurate because, in this case, the irrigated crops are at their maximum development. In contrast, the non-irrigated crops are already harvested (e.g., winter cereals) or not very vigorous (e.g., forage). However, this behaviour indicates that recognizing irrigated areas based only on optical indices could be unreliable in areas characterized by crop or cultivation practices different from those present in the Upper Tiber River Valley. To demonstrate this consideration, Figure 8 shows the comparison between the OAs of two exemplary optical and radar datasets ('S2_MSAVI' and 'S1_VV', respectively) using an RF classifier based only on the single feature related to a specific week of the irrigation season.

As shown in Figure 8, while the OA of the radar dataset 'S1_VV' increases and remains relatively stable at around 75%, the OA of the optical dataset is characterized by a wider variability. In particular, the OA is ≤60% in two weeks (27th and 28th). This is a transition period in the study area, during which non-irrigated

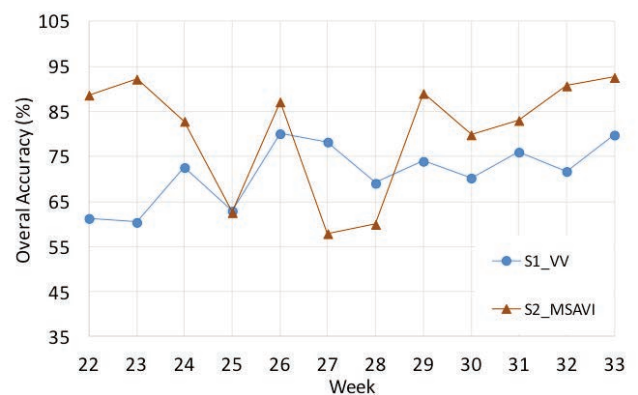


Figure 8. Overall accuracy (%) of the random forest classifiers for the 'S1_VV' and 'S2_MSAVI' datasets was obtained using the single feature of each week. See Table 4 for the description of the dataset codes.

crops begin to be harvested (*i.e.*, vigour decrease) and irrigated crops start to become well developed (Table 2), with a consequent drastic reduction in the accuracy of the classifiers based on a single optical feature. Furthermore, optical data are not always available due to cloud coverage. For example, the optical data in the 2016 irrigation season were not considered due to their high discontinuity. In this situation, even if it would have been possible to evaluate the effectiveness of recognition in the late irrigation season, an early evaluation, similar to that obtained in 2017, would not have been possible. On the other hand, the radar data, although it takes 2-3 weeks of observations to reach satisfactory accuracy values (Figure 4), proved to be effective in both dry and wet years.

The study has also shown that the crop type significantly affects the identification accuracy of irrigated and rainfed fields (Figure 7). Among the irrigated plots, those with tobacco are recognized with greater accuracy than those with maize. Among rainfed fields, those with wheat are better recognized than those with sunflowers. In both cases, a possible explanation is that crops associated with lower recognition accuracies (maize and sunflower) are less represented in the study area (Figure 2). We carried out further simulations imposing balanced numbers of fields with the crops mentioned above to verify this effect. In the case of irrigated crops, the use of datasets with a similar number of tobacco and maize plots does not lead to significant changes in the accuracies indicated in Figure 7A-C.

On the other hand, maize is a crop with a slightly earlier growth cycle than tobacco (Table 1); therefore, in the initial phases of the irrigation season, this crop may already be sufficiently developed to produce an optical response similar to some rainfed crops, thus reducing the related recognition accuracy (Figure 7A). This problem does not affect radar indices which perform similarly for tobacco and maize (Figure 7A). As the season progresses (Figure 7B and C), the OA differences between tobacco and maize tend to reduce, but the better recognition accuracy of tobacco fields is still evident. The explanation may lie in the fact that the irrigation management reserved for tobacco (a high-income crop) is usually more careful than that applied to a staple crop like maize. Moreover, maize's high crop density could negatively affect radar indices, which tends to suppress information on surface soil moisture (Gao *et al.*, 2017).

In the case of rainfed crops (sunflower and wheat), we found that more balanced databases allow to significantly improve the recognition accuracy for sunflower as a rainfed crop of all datasets. The poor accuracy obtained by many datasets in the first period of the irrigation season (Figure 7D) is justifiable since the sunflower, although not irrigated, has a growth cycle very similar to that of the main irrigated crops. This factor can make recognition based on single features and single sources unreliable in the initial stages of the irrigation season (Figure 7D). However, even in this challenging case, multiple sources datasets (*e.g.*, 'S2_MSAVI + S1_VV + p') allow obtaining satisfactory accuracy from the beginning of the season.

Conclusions

The main objectives of the work were to evaluate the effectiveness of different RS data in the early recognition of irrigated and non-irrigated plots and evaluate how the crop type can influence the recognition.

The analysis of the results has shown that the classifications based on the indices derived from S1 SAR data are reliable in most

situations. The index based on the VH and VV polarizations ratio has the best performance with accuracies close to 90% with only three features (*i.e.*, three weeks). All the radar-type indices benefit from additional information on precipitation, although the contribution of this feature becomes less and less important as the irrigation season progresses.

Under conditions of limited cloud coverage, optical indices like NDVI and MSAVI (from Landsat 8 and Sentinel-2) have shown excellent performance in the early recognition of irrigated areas. The overall accuracy is close to 90% already from the beginning of the irrigation season. However, it was also underlined how the recognition with optical indices only could be critical when rainfed and irrigated crops have similar vegetative characteristics.

In conclusion, we recommend using combined datasets because, in general, they allow obtaining satisfactory accuracy with a single observation period and are robust to recognition problems deriving from the low representativeness of some crop typologies or their peculiar characteristics.

References

- Bazzi H., Baghdadi N., Ienco D., El Hajj M., Zribi M., Belhouchette H., Escorihuela M.J., Demarez V. 2019. Mapping irrigated areas using Sentinel-1 time series in Catalonia, Spain. *Remote Sens.* 11:1836.
- Bousbih S., Zribi M., El Hajj M., Baghdadi N., Lili-Chabaane Z., Gao Q., Fanise P. 2018. Soil moisture and irrigation mapping in a semi-arid region, based on the synergetic use of Sentinel-1 and Sentinel-2 Data. *Remote Sens.* 10:1953.
- Breiman L. 2001. Random forests. *Mach. Learn.* 45:5-32.
- Breiman L. 2002. Manual on setting up, using, and understanding random forests v3. 1. Tech. Report, Stat. Dep. Univ. Calif. Berkeley. Available from: <http://oz.berkeley.edu/users/breiman>
- Cai X., Rosegrant M.W. 2002. Global water demand and supply projections: part 1. A modeling approach. *Water Int.* 27:159-69.
- Calera A., Campos I., Osann A., D'Urso G., Menenti M. 2017. Remote sensing for crop water management: from ET modelling to services for the end users. *Sensors.* 17:1104.
- Congalton R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 37:35-46.
- Cornes R.C., van der Schrier G., van den Besselaar E.J.M., Jones P.D. 2018. An ensemble version of the E-OBS temperature and precipitation data sets. *J. Geophys. Res. Atmos.* 123:9391-409.
- Dari J., Quintana-Seguí P., Escorihuela M.J., Stefan V., Brocca L., Morbidelli R. 2021. Detecting and mapping irrigated areas in a Mediterranean environment by using remote sensing soil moisture and a land surface model. *J. Hydrol.* 596:126129.
- Dari J., Brocca L., Quintana-Seguí P., Escorihuela M.J., Stefan V., Morbidelli R., 2020. Exploiting high-resolution remote sensing soil moisture to estimate irrigation water amounts over a Mediterranean region. *Remote Sens.* 12:2593.
- Deines J.M., Kendall A.D., Hyndman D.W. 2019. Annual irrigation dynamics in the U.S. Northern high plains derived from landsat satellite data. *Geophys. Res. Lett.* 44:9350-60.
- Deines J.M., Kendall A.D., Crowley M.A., Rapp J., Cardille J.A., Hyndman D.W. 2019. Mapping three decades of annual irrigation across the US High Plains Aquifer using Landsat and Google Earth Engine. *Remote Sens. Environ.* 233:111400.
- Gao Q., Zribi M., Escorihuela M.J., Baghdadi N. 2017. Synergetic

- use of sentinel-1 and sentinel-2 data for soil moisture mapping at 100 m resolution. *Sensors (Switzerland)*. 17:1966.
- Gao Q., Zribi M., Escorihuela M., Baghdadi N., Segui P. 2018. Irrigation mapping using Sentinel-1 time series at field scale. *Remote Sens.* 10:1495.
- Huete A.R. 1988. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* 25:295-309.
- Jalilvand E., Tajrishy M., Hashemi S.A.G., Brocca L. 2019. Quantification of irrigation water using remote sensing of soil moisture in a semi-arid region. *Remote Sens. Environ.* 231:111226.
- Li X., Troy T.J. 2018. Changes in rainfed and irrigated crop yield response to climate in the western US. *Environ. Res. Lett.* 13:064031.
- Liaw A., Wiener M. 2002. Classification and regression by random forest. *R News*. 2-3:18-22.
- Maselli F., Battista P., Chiesi M., Rapi B., Angeli L., Fibbi L., Magno R., Gozzini B. 2020a. Use of Sentinel-2 MSI data to monitor crop irrigation in Mediterranean areas. *Int. J. Appl. Earth Obs. Geoinf.* 93:102216.
- Maselli F., Chiesi M., Angeli L., Fibbi L., Rapi B., Romani M., Sabatini F., Battista P. 2020b. An improved NDVI-based method to predict actual evapotranspiration of irrigated grasses and crops. *Agric. Water Manag.* 233:106077.
- Ozdogan M., Yang Y., Allez G., Cervante C. 2010. Remote sensing of irrigated agriculture: opportunities and challenges. *Remote Sens.* 2:2274-304.
- Pageot Y., Baup F., Inglada J., Baghdadi N., Demarez V. 2020. Detection of irrigated and rainfed crops in temperate areas using Sentinel-1 and Sentinel-2 time series. *Remote Sens.* 12:3044.
- Qi J., Kerr Y., Chehbouni A. 1994a. External factor consideration in vegetation index development. pp 723-730 in *Proc. 6th Int. Symp. Phys. Meas. Signatures Remote Sens*, France.
- Qi J., Chehbouni A., Huete A.R., Kerr Y.H., Sorooshian S. 1994b. A modified soil adjusted vegetation index. *Remote Sens. Environ.* 48:119-26.
- Richards J.A., Jia X. 2006. *Remote sensing digital image analysis: An introduction*. Springer, Berlin, Germany.
- Rockström J., Falkenmark M., Lannerstad M., Karlberg L. 2012. The planetary water drama: dual task of feeding humanity and curbing climate change. *Geophys. Res. Lett.* 39.
- Romero M., Luo Y., Su B., Fuentes S. 2018. Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. *Comput. Electron. Agric.* 147:109-17.
- Rosegrant M.W., Cline S.A. 2003. Global food security: challenges and policies. *Science* 302:1917-9.
- Rouse W., Haas R.H., Deering D.W. 1974. Monitoring vegetation systems in the Great Plains with ERTS, NASA SP-351. *Third ERTS-1 Symp.* Vol. 1.
- Sepulcre-Cantó G., Zarco-Tejada P.J., Sobrino J.A., Berni J.A.J., Jiménez-Muñoz J.C., Gastellu-Etchegorry J.P. 2009. Discriminating irrigated and rainfed olive orchards with thermal ASTER imagery and DART 3D simulation. *Agric. For. Meteorol.* 149:962-75.
- Shiklomanov I.A. 2000. Appraisal and assessment of world water resources. *Water Int.* 25:11-32.
- Vanino S., Pulighe G., Nino P., De Michele C., Falanga Bolognesi S., D'Urso G. 2015. Estimation of evapotranspiration and crop coefficients of tendone vineyards using multi-sensor remote sensing data in a Mediterranean environment. *Remote Sens.* 7:14708-730.
- Vanino S., Nino P., De Michele C., Falanga Bolognesi S., D'Urso G., Di Bene C., Pennelli B., Vuolo F., Farina R., Pulighe G. 2018. Capability of Sentinel-2 data for estimating maximum evapotranspiration and irrigation requirements for tomato crop in Central Italy. *Remote Sens. Environ.* 215:452-70.
- Veloso A., Mermoz S., Bouvet A., Le Toan T., Planells M., Dejoux J.F., Ceschia E. 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sens. Environ.* 199:415-26.
- Vreugdenhil M., Wagner W., Bauer-Marschallinger B., Pfeil I., Teubner I., Rüdiger C., Strauss P. 2018. Sensitivity of Sentinel-1 backscatter to vegetation dynamics: an Austrian case study. *Remote Sens.* 10:1396.
- Xie Y., Lark T.J., Brown J.F., Gibbs H.K. 2019. Mapping irrigated cropland extent across the conterminous United States at 30 m resolution using a semi-automatic training approach on Google Earth Engine. *ISPRS J. Photogramm. Remote Sensing* 155:136-49.
- Weaver J., Moore B., Reith A., McKee J., Lunga D. 2018. A comparison of machine learning techniques to extract human settlements from high resolution imagery. in *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*.