OBIA approach for the analysis of medium and high resolution satellite data for environmental

Politecnico di Bari

LIBERATORIA PER L'ARCHIVIAZIONE DELLA TESI DI DOTTORATO

Al Magnifico Rettore
del Politecnico di Bari

Il/la sottoscritto/a Claudio Ladisa nato/a a Manfredonia (FG) il 01/09/1995

residente a Carovigno (BR) in via Parigi 14 e-mail ladisaclaudio@gmail.com

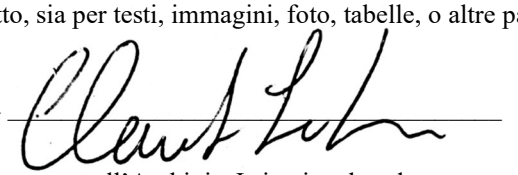
iscritto al 3° anno di Corso di Dottorato di Ricerca in Rischio, sviluppo ambientale, territoriale ed edilizio ciclo 37

ed essendo stato ammesso a sostenere l'esame finale con la prevista discussione della tesi dal titolo:

OBIA approach for the analysis of medium resolution satellite data for environmental monitoring

DICHIARA

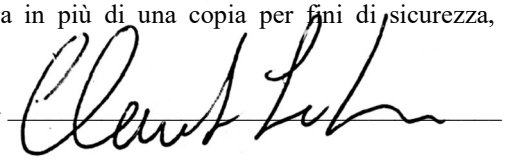
1) di essere consapevole che, ai sensi del D.P.R. n. 445 del 28.12.2000, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici conseguenti al provvedimento emanato sulla base di tali dichiarazioni;
2) di essere iscritto al Corso di Dottorato di ricerca in Rischio, sviluppo ambientale, territoriale ed edilizio ciclo 37 , corso attivato ai sensi del "*Regolamento dei Corsi di Dottorato di ricerca del Politecnico di Bari"*, emanato con D.R. n.286 del 01.07.2013;
3) di essere pienamente a conoscenza delle disposizioni contenute nel predetto Regolamento in merito alla procedura di deposito, pubblicazione e autoarchiviazione della tesi di dottorato nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica;
4) di essere consapevole che attraverso l'autoarchiviazione delle tesi nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica del Politecnico di Bari (IRIS-POLIBA), l'Ateneo archivierà e renderà consultabile in rete (nel rispetto della Policy di Ateneo di cui al D.R. 642 del 13.11.2015) il testo completo della tesi di dottorato, fatta salva la possibilità di sottoscrizione di apposite licenze per le relative condizioni di utilizzo (di cui al sito http://www.creativecommons.it/Licenze), e fatte salve, altresì, le eventuali esigenze di "embargo", legate a strette considerazioni sulla tutelabilità e sfruttamento industriale/commerciale dei contenuti della tesi, da rappresentarsi mediante compilazione e sottoscrizione del modulo in calce (Richiesta di embargo);
5) che la tesi da depositare in IRIS-POLIBA, in formato digitale (PDF/A) sarà del tutto identica a quelle **consegnate**/inviate/da inviarsi ai componenti della commissione per l'esame finale e a qualsiasi altra copia depositata presso gli Uffici del Politecnico di Bari in forma cartacea o digitale, ovvero a quella da discutere in sede di esame finale, a quella da depositare, a cura dell'Ateneo, presso le Biblioteche Nazionali Centrali di Roma e Firenze e presso tutti gli Uffici competenti per legge al momento del deposito stesso, e che di conseguenza va esclusa qualsiasi responsabilità del Politecnico di Bari per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;
6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi, ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto il Politecnico di Bari ed i suoi funzionari sono in ogni caso esenti da responsabilità di qualsivoglia natura: civile, amministrativa e penale e saranno dal sottoscritto tenuti indenni da qualsiasi richiesta o rivendicazione da parte di terzi;
7) che il contenuto della tesi non infrange in alcun modo il diritto d'Autore né gli obblighi connessi alla salvaguardia di diritti morali od economici di altri autori o di altri aventi diritto, sia per testi, immagini, foto, tabelle, o altre parti di cui la tesi è composta.

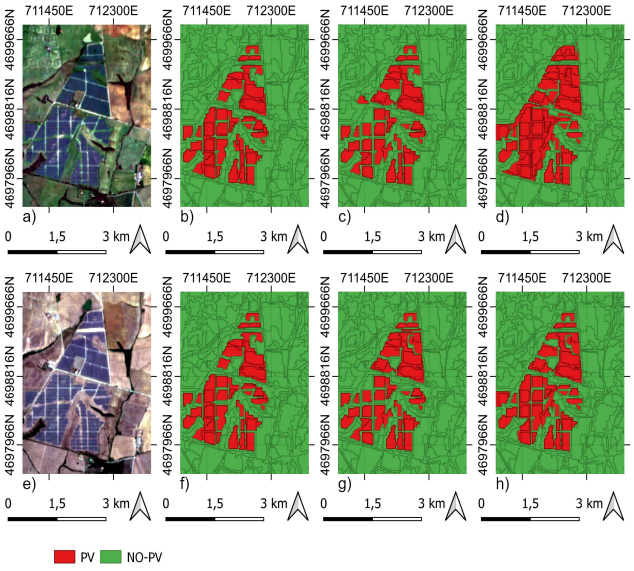
Luogo e data 15/12/2024 Firma 

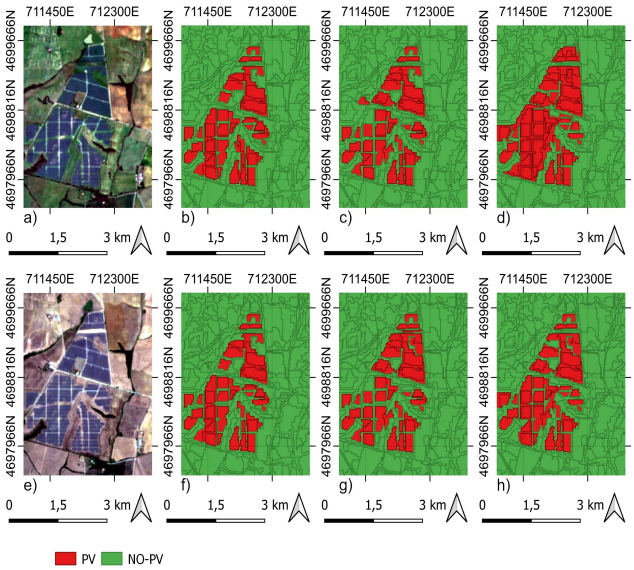
Il/La sottoscritto, con l'autoarchiviazione della propria tesi di dottorato nell'Archivio Istituzionale ad accesso aperto del Politecnico di Bari (POLIBA-IRIS), pur mantenendo su di essa tutti i diritti d'autore, morali ed economici, ai sensi della normativa vigente (Legge 633/1941 e ss.mm.ii.),

CONCEDE

- al Politecnico di Bari il permesso di trasferire l'opera su qualsiasi supporto e di convertirla in qualsiasi formato al fine di una corretta conservazione nel tempo. Il Politecnico di Bari garantisce che non verrà effettuata alcuna modifica al contenuto e alla struttura dell'opera.
- al Politecnico di Bari la possibilità di riprodurre l'opera in più di una copia per fini di sicurezza, back-up e conservazione.

Luogo e data Bari, 15/12/2024 Firma 

D.R.S.A.T.E.	**POLITECNICO DI BARI**	**03**
Doctor in Risk And Environmental, Territorial And Building Development		**2025**
Coordinator: Prof. Vito Iacobellis		
XXXVII CYCLE Curriculum: ICAR/06		
DICATECh Department of Civil, Environmental, Building Engineering and Chemistry		
	Claudio Ladisa	
	OBIA Approach for the Analysis of Medium and High Resolution Satellite Data for Environmental Monitoring: PV Plant Mapping	
	Prof. Eufemia Tarantino DICATECh Politecnico di Bari Dr. Ir. Alessandra Capolupo, Ph.D DICATECh Politecnico di Bari	
		

	D.R.S.A.T.E.	**POLITECNICO DI BARI**	**03**
	Dottorato di Ricera in Rischio e Sviluppo Ambientale, Territoriale ed Edilizio		**2025**
	Coordinatore: Prof. Vito Iacobellis		
	XXXVII Ciclo Curriculum: ICAR/06		
	DICATECh Dipartimento di Ingegneria Civile, Ambientale, Edile e Chimica		
		Claudio Ladisa	
		Approccio OBIA per l'analisi di dati satellitari a media e alta risoluzione per il monitoraggio ambientale: mappatura di impianti fotovoltaici	
		Prof. Eufemia Tarantino DICATECh Politecnico di Bari Dr. Ir. Alessandra Capolupo, Ph.D DICATECh Politecnico di Bari	
			

EXTENDED ABSTRACT *(eng)*

The global expansion of photovoltaic (PV) installations continues to grow in response to rising energy demand and the need for sustainable energy production. However, large-scale solar farms require the use of large areas of land that are diverted from food production, the demand for which has increased dramatically in recent years, owing to the growing world population. To monitor the rapid growth of solar farms, it is essential to collect data on the quantity, distribution, and impact of these systems. Despite the many challenges in mapping PV systems owing to the variety of materials and complexity of layouts, many researchers are trying to improve this process using remote sensing techniques combined with machine and deep learning. This study developed a comprehensive framework for mapping and monitoring PV systems using open-source Sentinel-2 (S2) satellite imagery and remote sensing techniques supported by machine-learning algorithms. The framework was applied to two study areas with different characteristics by adopting an Object-Based Image Analysis (OBIA) approach to improve image segmentation and classification. The preliminary phase included the analysis of the Normalized Difference Vegetation Index (NDVI) index variation within the Ground Truth (GT) polygons, improving the segmentation through the creation of image "composites." During the feature extraction phase, 75 features were obtained from the segments, including spectral indices, such as Enhanced vegetation index (EVI), Normalized Difference Tillage Index (NDTI), and Normalized Difference Built-up Index (NDBI), which turned out to be the most relevant. The Random Forest (RF) models obtained the best performance, with

accuracies above 90% in the test areas and high F1-scores. Support Vector Machine models also showed satisfactory results, whereas K-nearest neighbor models proved to be less effective. The analysis highlighted a trend towards land transformation from agricultural to industrial, with a significant presence of vegetation under many PV arrays, suggesting the development of agrovoltaic plants. The applied methodology represents a potentially valuable predictive tool for monitoring land transformation and territorial planning. The continued availability of high-resolution open-source satellite images, progress in computing capacity, and advancements in image classification algorithms render this methodology easily replicable for both academic and technical design purposes. The results obtained lay the foundation for an approach to PV mapping, which is essential for increasingly intense future monitoring.

keywords

Remote Sensing, OBIA, Photovoltaic, Solar Energy, Machine Learning

EXTENDED ABSTRACT *(ita)*

L'espansione globale degli impianti fotovoltaici continua a crescere in risposta all'aumento della domanda di energia e alla necessità di una produzione energetica sostenibile. Tuttavia, i parchi solari su larga scala richiedono l'utilizzo di ampie aree di terreno, sottraendole alla produzione alimentare, la cui domanda è aumentata notevolmente negli ultimi anni a causa della crescita della popolazione mondiale. Per monitorare la rapida crescita dei parchi solari, è fondamentale raccogliere dati sulla quantità, distribuzione e impatto di questi sistemi. Nonostante le numerose sfide nella mappatura degli impianti fotovoltaici, dovute alla varietà di materiali e alla complessità dei layout, molti ricercatori stanno cercando di migliorare questo processo utilizzando tecniche di telerilevamento combinate con l'apprendimento automatico e l'apprendimento profondo. Questo studio ha sviluppato un quadro metodologico completo per la mappatura e il monitoraggio degli impianti fotovoltaici utilizzando immagini satellitari open-source Sentinel-2 (S2) e tecniche di telerilevamento supportate da algoritmi di machine learning. Il metodo è stato applicato a due aree di studio con caratteristiche diverse, adottando un approccio basato sull'analisi delle immagini per oggetti (OBIA) per migliorare la segmentazione e la classificazione delle immagini. La fase preliminare ha incluso l'analisi della variazione del Normalized Difference Vegetation Index (NDVI) all'interno dei poligoni del Ground Truth (GT), migliorando la segmentazione attraverso la creazione di "compositi" d'immagine. Durante la fase di estrazione delle caratteristiche, sono state ottenute 75 caratteristiche dai segmenti, tra cui indici spettrali come Enhanced vegetation index (EVI), Normalized Difference Tillage Index (NDTI), e Normalized Difference Built-up Index (NDBI), che si

sono rivelati i più rilevanti. I modelli Random Forest hanno ottenuto le migliori prestazioni, con accuratezze superiori al 90% nelle aree di test e alti punteggi F1. Anche i modelli Support Vector Machine hanno mostrato risultati soddisfacenti, mentre i modelli K-nearest neighbor si sono dimostrati meno efficaci. L'analisi ha evidenziato una tendenza alla trasformazione del suolo da agricolo a industriale, con una significativa presenza di vegetazione sotto molti pannelli fotovoltaici, suggerendo lo sviluppo di impianti agrovoltaici. La metodologia applicata rappresenta uno strumento predittivo potenzialmente utile per monitorare la trasformazione del territorio e per la pianificazione territoriale. La continua disponibilità di immagini satellitari open-source ad alta risoluzione, i progressi nella capacità di calcolo e l'avanzamento degli algoritmi di classificazione delle immagini rendono questa metodologia facilmente replicabile in futuro sia per scopi accademici che per la progettazione tecnica. I risultati ottenuti gettano le basi per un approccio alla mappatura degli impianti fotovoltaici, essenziale per un monitoraggio sempre più intenso in futuro.

keywords

Remote Sensing, OBIA, Photovoltaic, Solar Energy, Machine Learning

INDEX

INTRODUCTION

The planet has reached a critical juncture in the transition to sustainable energy owing to the increasing depletion of fossil fuels and growing concerns about climate change (Owusu and Asumadu-Sarkodie, 2016). More than 70% of atmospheric carbon emissions originate from conventional fossil fuel-derived energy sources (Güney, 2022). To ensure that the increase in global temperature is limited to manageable levels of 1.5 °C or less, the international community must fully eliminate these greenhouse pollutants by no later than 2050. According to the findings of the International Renewable Energy Agency (Irena), renewable energy remains the only viable resource if we aspire to meet our extensive electricity demand while limiting further environmental degradation caused by human actions, given the recent significant advances in technological innovation (Wang et al., 2023a).

Solar energy is gaining ground as one of the fastest-growing renewable energy resources. Owing to its exceptional sustainability and quality, the energy generated through PV panels has attracted increasing global interest because of its tremendous prospects in terms of the growth and economics in energy production (Parida et al., 2011). Unlike other sources, such as energy provided by wind power, that from the sun is quiet in the production process and extremely flexible in environmental applications where it is used (Wolniak & Skotnicka-Zasadzień, 2022). As for the global statistics of the sector recorded at the end of the period ending in 2022, the cumulative installed capacity amounted to more than double the number reported only three years earlier, and from approximately 590 GW, we have risen to almost the upper 1000GW currently standing at very high numbers (Kruzel & Helbrych, 2018).

However, the extensive land use for the construction of large-scale PV facilities has an environmental impact that must be considered. Construction of such infrastructure can result in both alterations to perceived landscape values and changes to ecological, hydraulic, and vegetative soil functions (Tawalbeh et

al., 2021) The installation of PV systems exerts a significant influence on photosynthetically active radiation and the radiative budget; reducing surface albedo while altering precipitation patterns (Choi et al., 2020). These changes have a profound effect on local microclimates by affecting phenomena such as evaporation rates, wind speeds, atmospheric temperature, moisture levels within soils, or temperatures within them (Mustafa et al., 2020). Therefore, it is essential to conduct meticulous mapping exercises that highlight the spatial arrangement of the major PV stations.

Geographic data on their distribution play an essential role not only in assessing potential power generation but also in examining the associated environmental impacts from an operational perspective (Dunnett et al., 2020). In addition, exploiting efficient distribution patterns is a valuable resource when planning areas dedicated to landscape design principles, along with the energy demands of urbanization (Ladisa et al., 2024a). Only a few countries have validated databases established through voluntary contributions or collaborative mapping activities (Kasmi et al., 2023).

Geomatics has addressed this topic over the past decade (T. Nguyen & M. Pearce, 2013). For example, various investigations have used drones to map and monitor the spatial distribution and diagnostics of PV cells (Kumar et al., 2018). Thermal photogrammetry using drones has proven to be an interesting technique, particularly for solar panel diagnostics. Low-cost thermal sensors can also be used for this purpose (Zhang et al., 2017a). However, when mapping large-scale solar farms, field surveys can be both time-consuming and expensive (Ladisa et al., 2022a). A more rapid and noninvasive approach to collecting spatial data on such installations is satellite imaging (Ladisa et al., 2022b), which has the added advantage of being able to collect information over larger areas and monitor temporal patterns in the dynamics of expansion among major PV installations (Xia et al., 2023).

1.1. Previous Work on Detection of solar energy from PV by Remote Sensing

In recent years, there has been a notable expansion in large-scale PV installations across multiple regions. Consequently, numerous studies have focused on mapping PV systems using remote sensing imagery. These investigations considered various input data types, image-processing methodologies, and algorithms. Initial research using satellite imagery to detect solar panels primarily relied on high-resolution satellite imaging or Deep Learning (DL) techniques.

(Jiang et al., 2021) recently developed a multi-resolution dataset that segments solar panels in commercial Gaofen-2 and Beijing-2 satellite images with spatial resolutions as low as 0.8m; area images with an even finer resolution down to 0.3 m; alongside drone-based pictures at depths reaching up to just 0.1m. The samples contained diverse installations across different terrains, such as scrublands, grasslands, and agricultural fields. This study utilized a dataset to assess the effectiveness of various deep networks for PV segmentation and explored the viability of incorporating cross-directional samples with varying resolutions. The results revealed that different resolutions were more effective in generating accurate results based on specific distributions within the PV arrays, exhibiting an average metric accuracy exceeding 0.84. Moreover, our findings indicate that integrating cross-direct sample applications does not produce favorable results; thus, precise calibration is crucial when utilizing pretrained networks using target samples. Further research has concentrated on remotely mapping large-scale solar power plants by transitioning from high-resolution commercial imagery to medium-resolution open-source imagery because of the exorbitant costs related to acquiring high-quality commercial images across the extensive areas under investigation.

Recently, the National Aeronautics and Space Administration (NASA) and ESA (European Space Agency) have initiated free-access services providing optical and radar satellite imaging at global coverage but limited resolution levels, resulting in widespread use without any additional expenses incurred (Peters et

al., 2018). Through the EU-led Copernicus program, the ESA has developed global environmental monitoring services based on space technology to ensure autonomous observation capabilities of the Earth in Europe (Jutz & Milagro-Pérez, 2020).

The Copernicus program comprises a spatial component, basic services, and in-situ measurements (Thepaut et al., 2018). The spatial component includes a series of satellites called Sentinels, each with specific missions and instruments for monitoring various aspects of the planet. In particular, multispectral optical data at medium resolution provided by Sentinel-2 (S2) have been widely used in applications such as vegetation monitoring, land-use analysis, and natural disaster mapping (Gascon et al., 2017). These data have proven to be an excellent alternative to NASA's Landsat product due to their higher quality geometric resolution, providing image resolutions up to ten meters compared to Landsat, which is limited to only three times that amount, that is, 30 m.

In a recent study conducted by Costa et al. (2021)), various combinations of DL models were evaluated to classify PV plants using S2 imagery at 10m resolution in the visible near-infrared bands. The analysis included four architectures, U-net, DeepLabv3+, FPN, and PSPNet, along with four backbones, ResNet-50, ResNet101, Eff-b0, and Eff-b7, resulting in 16 potential model combinations. Notably, the performance of all variants was surpassed by U-net using an EfficientNet B7 backbone, achieving an exceptional accuracy above 98%. While highlighting the unique spectral features inherent to its imagery that facilitate detection capability for PV panels, challenges associated with significant computational costs coupled with lengthy processing times complicate this phase of exploration, rendering machine learning (ML) algorithms more favorable compared to DL techniques, considering their minimal usage requirements regarding training set quantities as well as reduced computation demands.

Moreover, Zhang et al. (2022) presented a novel workflow integrating ML-based methodology specifically employing pixel-oriented Random Forest (RF) augmented through visual interpretation practice on Landsat L8 satellite data,

helping create solar power plant datasets covering approximately 2917 km² across China until the FY21-end showed overall accuracies surpassing or equating up-to-the-mark threshold levels exceeding 96%. However, there are noteworthy complexities arising from difficulties encountered in mapping numerous sources depicting such stations, including those noted due to spatial proximities and structural similarities between panel roofs located atop warehouses, which pose hurdles against uninterrupted growth activities, further degrading the soil quality surrounding sites, which inevitably accompanied regional variations, impacting the station layout, and adversely imbibing complexity into tracking growth aspects effectively at scale, requiring tailored solutions to balance desired outcomes comprehensively within a reasonable time parameter upheld.

Wang et al. (2022) used multi-source remote sensing data, including S2 reflectance and Sentinel-1 (S1) polarization, to investigate the transferability of the model between rural and urban landscapes. Using the Google Earth Engine (GEE), an RF classification model was developed based on both the input information from S2 and S1 and spectral indices, such as the Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-Up Index (NDBI), and texture values, such as entropy correlation and contrast. The results showed that a classifier trained in an urban landscape had a higher accuracy with a higher transferability rate of 97.24% compared to a classifier trained in a rural context. The overall maximum accuracy reached 98.90% using samples obtained from the areas of interest in both regions, whereas incorporating VIIRS nighttime light data was found to help improve the predictions by correcting errors in bare rocks or shadowed mountain slopes.

Chen et al. (2022) used Landsat 8 imagery to develop PV identification models by combining spectral features, including the Normalized Difference Tillage Index (NDTI), NDVI, Built-Up Area Index (BUAI), and Normalized Difference Water Index (NDWI), with terrain features and comparing various ML algorithms, including RF, and support vector machine (SVM). The present study introduces a novel approach using spectral curves to discern PV arrays from other land cover

types, thereby enabling the development of a multiclass classification model as opposed to the binary models employed in previous studies. Notably, our investigation revealed that PV arrays are easily distinguishable from water bodies, snow cover, vegetation, and built-up areas but share certain features with barren lands and roads within river valleys. NDTI proved instrumental for detecting PV arrays, whereas additional indices such as Band Intensity (BI), NDVI and Enhanced Vegetation Index (EVI) facilitated their differentiation from urbanization or other forms of vegetation.

Classification of diverse land types using ML algorithms, such as SVM and RF. Current research has significantly improved the capabilities of algorithms, demonstrating high levels of accuracy over larger geographic ranges, even when subjected to limited computational resources. Surprisingly, this was achieved using fewer DL techniques than would normally be required, which is a remarkable advancement.

The uniqueness of this study lies in the sampling phase. As suggested by (Colditz, 2015) for optimal results, the training sample area was approximately 0.25% of the total study area. Furthermore, a combination of high-resolution Google images and multispectral RGB images was used during the sample-labeling process to minimize the temporal disparity between the samples and images. This approach was effective in reducing the classification errors. In this case, the RF classifier was selected because of its fast processing of multi-source data. The results show that the main sources of error include different spectra of the same object and those of different objects. Differences in the installation direction and distance led to significant variations in the spectral and structural information within the PV arrays, hindering complete recognition. For example, panels installed horizontally over long distances have distinct features compared to those on a slope with minimal differentiation between rows. In addition, the spectral information for PV panels located on complicated surfaces, such as agricultural greenhouses, is affected by complex background pixels, resulting in significantly divergent NDVI and NDBI values when comparing these surface-mounted solar

modules. The overall accuracy (OA) achieved in this study was 94.31%. In addition to mapping solar installations, this study examined how PV systems affect changes in local landscapes. Special attention was given to the analysis of vegetation change before and after PV system installation by examining the annual averages of the NDVI values. Different types of PV panels have been observed to have varying effects on the vegetation in different climatic regions. In many cases, these panels tend to reduce the surface green cover; however, their use in temperate zones can promote plant growth.

Pixel-based (PB) image processing is the most commonly used approach in all studies mentioned here, which involves assigning classifications to each individual pixel present in a particular analyzed image. Although it is a simple implementation, this method can produce results with a " salt-and-pepper effect without insufficient consideration of pixel-to-pixel relationships or geometric features of the training sample. Despite this limitation, object-based image analysis (OBIA) has been used alone or in conjunction with other methods, particularly those that rely heavily on semi-spectral classifiers for interpretation. OBIA relies primarily on segmentation algorithms designed to transform covered pixels into homogeneous objects before classification, based on specific spatiotemporal and geometric attributes. These alternative approaches achieve excellent results overall compared to the conventional techniques that preceded them, largely because of the increased efficiency that helps to understand the complex remote sensing procedures related to plant health practices.

For instance, Xia et al. (2022) adopted an OBIA approach by merging segmentation algorithms in the eCognition Developer software with object-based classification algorithms within GEE. Specifically, they used the maximum annual value of the NDBI spectral index obtained from Landsat 8 bands as input and employed a multi-resolution segmentation (MRS) algorithm implemented in eCognition 9.5 (Trimble, Sunnyvale, CA, USA). The MRS algorithm was used to acquire homogenous objects based on three parameters: shape, scale, and compactness. The bottom-up regional fusion technique starts with singular pixel objects and

then merges smaller ones into larger ones iteratively, according to Chirici and Travaglini (2003). For this case study, setting the scale parameter values to approximately 30, while compaction values for both shapes should be approximately one or five, based on interpretation via visual image analysis, were considered acceptable choices. The OBIA approach was used in this study. The RF classifier was trained using the mean and standard deviation of pixels within sample objects from the Red, Near Infrared-1 (NIR1), Near infrared-2 (NIR2), Shortwave Infrared-1 (SWIR1), and Shortwave Infrared-2 (SWIR2) bands over a year, plus two shape features (compactness and landscape shape index). The overall classification accuracies for 2013 and 2019 were 85.8% and 92.8%, respectively, with kappa coefficients of R0.84D and R0RD86. Misclassifications mainly occur between PV power plants and some bare rocks owing to their similar spectral characteristics and segmentation issues related to shapes. The proposed method is computationally efficient, even though it involves time-series analysis, which significantly improves the accuracy levels during testing. The data obtained indicated a significant increase in the area of PV power plants in China between 2007 and 2019, with rapid growth observed from 2013 to 2019.

Another study that favored the OBIA approach over PB was conducted by Plakman et al. (2022). In this study, unlike the previous one, segmentation was carried out using Simple Non-Iterative Clustering (SNIC) within GEE, utilizing both the S2 optical and S1 radar input data. As Mahdianpari et al. (2018) and Achanta & Susstrunk (2017) demonstrated earlier, the SNIC algorithm is highly suitable for segmenting S2 images in terms of processing time compared to other comparable algorithms. Furthermore, the GEE application has an added benefit as it does not require computing resources from users. Statistics such as the mean value, standard deviation, and median were calculated for all resolution bands comprising each segmented object. Objects possessing these characteristics were exported from GEE into the RStudio environment, where they were used for training validation among the various RF models. The present procedure was performed because of RStudio's advantage of allowing the preservation and

application of distinct classification models, a functionality unavailable in GEE. The ability to employ such models on other datasets enhances their potential for study and analysis, which focuses on understanding the knowledge transferability associated with the considered models. In this study, efforts were made to balance the sample quantity per category within the training dataset. An imbalance exists between objects labeled as solar panels and designated non-solar panels, necessitating oversampling and undersampling algorithm implementation possibilities. It should be noted that the vulnerability of machine-learning-based algorithms lies primarily in their tendency to adapt predominantly towards majority classes, which magnifies the significance of our chosen sampling procedures. Hence, various models were executed and compared using the RStudio environment. The training data were manipulated through oversampling or undersampling, or left unaltered to achieve balance. Based on the results of these methods, it is evident that an RF model equipped with both spectral and backscatter properties has tremendous potential for automatically identifying solar farm objects subjected to segmentation. Models augmented via oversampled datasets outperformed their counterparts trained using undersampled versions, demonstrating superior classification accuracy ratings of 99.97% (Kappa = 0,90). Remarkably, although not surprisingly observed during this experimentation process, when retrieving details essential for imparting insights into suitable training samples due to insufficient gathered information following downsizing procedures reflective of undersampling techniques had eliminated said critical observations entirely from consideration, as expected results indicated so thereafter retrospectively analyzed consequences' post-hoc evaluation phase concluded positively supporting the aforementioned speculation describing ramifications associated with lacking necessary foresight requisite proficiency levels appropriate handling analytical tools emerging technologies pass substantial merits promoting allocation of sufficient resources, advancing research developments entrenched in sustainable approaches towards improving global energy ecosystems. In any case, the precision metrics obtained for both the validation and test

data demonstrate that the proposed technique is well-suited for use in transfer learning. Moreover, because the S1 and S2 satellites offer global coverage as open-source platforms, this approach can be easily replicated beyond the defined study area to obtain spectral and backscatter features.

(Ladisa et al., 2022b) research provides further insight into PV panel segmentation by comparatively evaluating two methods: eCognition Developer's multi-resolution algorithm versus Orfeo Toolbox's mean-shift algorithm on extraction from S2 imagery. The distinguishing aspect of their analysis entails determining the optimal parameters using the AssesSeg command-line software, which calculates the Euclidean distance (ED) between segmentations produced by each method. In terms of PV system segmentation purposes, it was observed that although the OTB generated larger total areas collectively within its polygons, optimization utilizing parameter selection techniques suggests that eCognition more accurately captures the ground truth (GT) with greater intersection outputs than OTB. This phenomenon is substantiated by relatively high median values engendered by polygon size output through a mean shift when compared with low standard deviations, affirming that a tendency exists, thus ensuring reproducibility across various sample groups while accounting for limitations associated with currently available algorithms regarding feature input selection problems.

Recursive Feature Elimination (RFE) has proven superior to leave-one-out (LOO) in terms of computational efficiency. RFE works iteratively to select the most relevant features. It begins by training the model on all available data and gradually removing the least significant features in each iteration. This elimination process continued until the desired number of features remained. As shown in Ladisa et al. (2024b), this methodology is not only faster than LOO, but also optimizes the feature set, improving the accuracy of RF classification models.

1.2. Thesis aim

The main objective of this thesis is to propose an OBIA framework for mapping PV parks using remote sensing images, particularly S2, in two different study areas. Solar energy plays a crucial role in renewable energy production and has seen an exponential growth in PV installations over the past 15 years. However, the rapid expansion of solar technology in rural areas poses environmental issues, such as soil loss and habitat modification, leading to changes in landscape characteristics.

Monitoring spatial expansion over time by analyzing its potential environmental impact is essential, and it is crucial to further understand which types of land cover are being transformed through the introduction of solar energy. The adopted OBIA framework comprises three main phases: image segmentation, feature extraction, and object/image classification. However, a literature review highlighted the lack of application of classification models tested on recent and past images to analyze plant progress over time and soil category transformations.

Additionally, integrating OBIA with next-generation ensemble learning approaches is a promising but underexplored approach in land-use analysis. The specific objectives of this thesis are: (i) to study the spectral characteristics of S2 time-series data corresponding to PV plants, (ii) to improve the segmentation phase, (iii) to determine which algorithm is most effective in reducing the feature space during feature selection, (iv) to test different learning algorithms for classification performance and compare results across various study areas, and (v) to analyze how spatial distribution patterns have changed since the first year of availability using S2 data by examining changes in land cover in two selected regions.

1. DATASET AND STUDY AREAS

1.1. Study area

Two study areas, one in Montalto di Castro (VT, Italy) and the other south of Seville (Spain) (Fig.1), covering a rectangular region were chosen to conduct the trials, as indicated in Fig. 2 and 3. Montalto di Castro is 7.5 km from north to south and 5.5 from east to west. km in the region south of Seville is 7.5 km in north-south direction and 5.5 km in the east-west direction. Measurements were taken to ensure that the study areas were as homogeneous as possible in terms of land cover classes. However, it is important to note that PV plants represent only a minority of the total coverage in both areas. These two study areas were selected to compare the same methodology in two different regions with varying vegetation, climate, and PV characteristics.

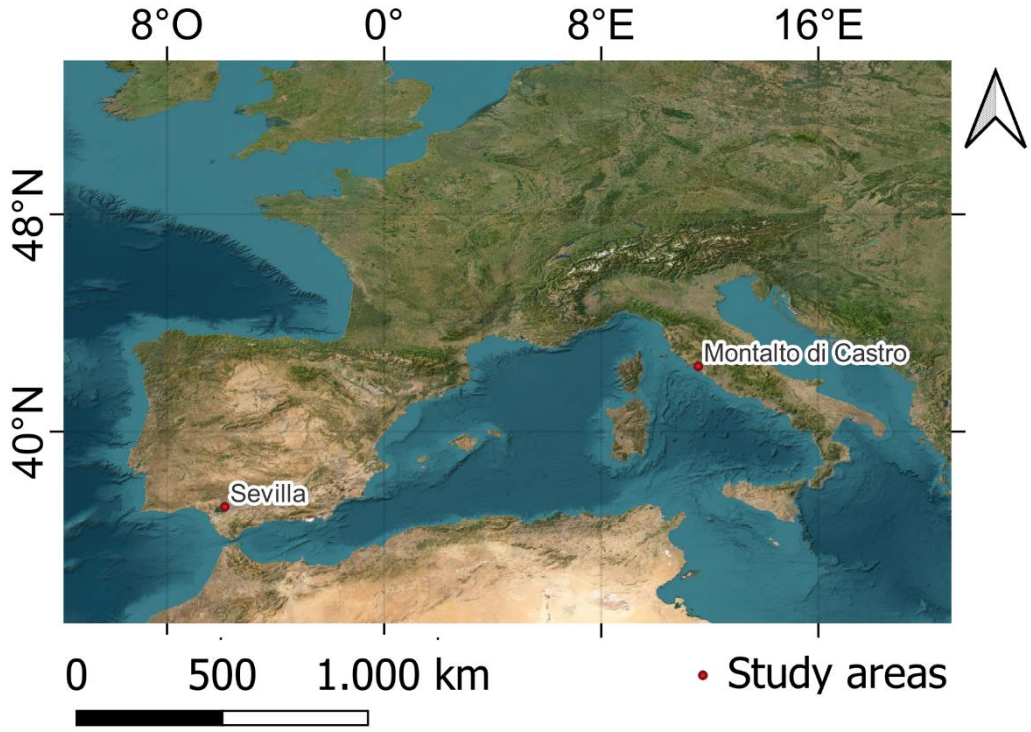


Fig. 1 - Study area location WGS84 reference system (EPSG:4326). Source: ESRI World View.

1.1.1. Montalto di Castro (Viterbo, Italy)

Montalto di Castro is situated in the region of Lazio, specifically within the Province of Viterbo, located centrally within Italy. This area has a multitude of diverse landscapes, including coastal regions, hillsides, and volcanic and mountainous zones (Manrai et al.2017).

In particular regard to Montalto di Castro's location lies its position within Maremma Laziale, an idyllic coastal zone renowned for possessing iron-sand beaches complemented by pristine pine forests that are traversed by not only one but three distinct waterways: The Fiora River and Arrone and Chiarone streams respectively (Schirru, 2019). Sprawling plains are based on abundant sunlight, and the convenient accessibility of Tuscia makes it an optimal zone for the expansion of PV plants. (Chatenoud et al., 2005). Montalto di Castro boasts 78% of all ground-mounted facilities within this region. The inception phase culminated in 2009, with a complete capacity of 24 MegaWatts (MW), followed by a commissioned second stage appraised at 8 MW implemented later that year. (Cristina Volpe Rinonapoli, 2023). Subsequently, both the third and fourth phases were finalized in December of the same year, resulting in an aggregate output potential rating of up to 44 MW. Thus far, the entire solar park has attained 285 MWp of power generation capacity, allocating sufficient energy supply round-the-clock to approximately twenty-four thousand urban dwellers.

In addition, this extensive facility occupies roughly two hundred eighty-three hectares. Plants can be differentiated based on their pattern and arrangement (Andrea Pietrarota, 2023). Close to the Terna substation located in Montalto lies the Cassiopea solar park, which has one of the highest densities among all plant areas. The characteristic fragmented texture comprises PV arrays aligned north-south and organized into multiple groups within this area. Positioned in the northwest direction here is Andromeda's solar park, which presents two distinct compact shaped clusters, unlike its counterpart Cassiopea's layout, producing a total energy output of 51 MW with an average spacing distance between each array of approximately 2 m, giving rise to square forms formed by a

clustered arrangement design (Cristina Volpe Rinonapoli, 2023). Other power-generating parks spread throughout our research site exhibit variations in the interstitial distances separating them, and surface types beneath these structures are seen more prominently during the winter/spring months, creating non-uniform morphological features displaying irregular geometries, where consolidation is not always present.

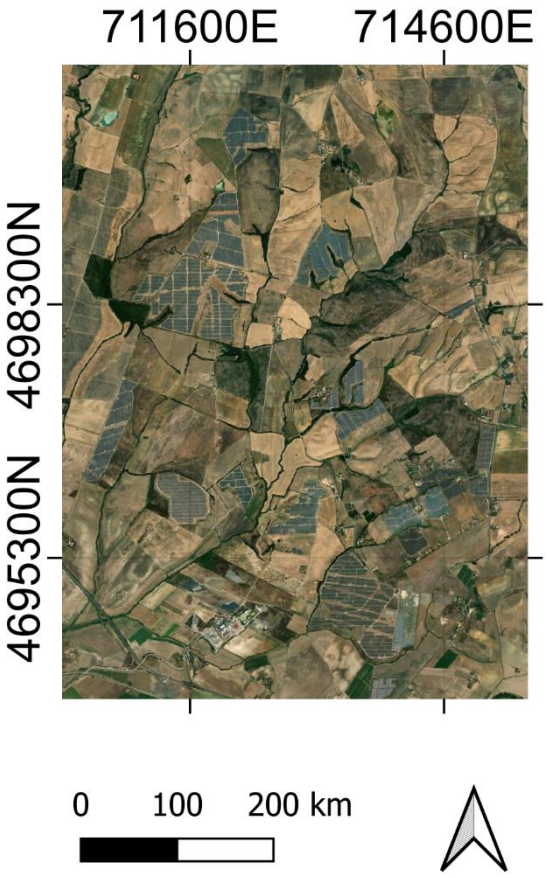


Fig. 2 - Study area location Montalto di Castro. WGS84/ UTM Zone 32N reference system (EPSG:32632). Source: ESRI World View.

1.1.2. Sevilla (Spain)

In the southern area of Seville, between Dos Hermanas and Utrera, a significant cluster of PV plants has been developed in terms of both size and solar energy production. (Ridao et al., 2007). This territory, characterized by one of the highest solar radiation levels in Europe, boasts numerous plants that maximize available solar resources. (Chapman, 2008). Among these, La Isla PV Park is a notable example. Located in the northernmost part of the study area, La Isla covers 520 ha and produces 157.5 MW. (NOVASOL, 2017).

In an area that receives over 1,920 h of sunshine per year, the park's efficiency is optimized owing to a fairly spaced arrangement of arrays, with approximately 4 m of separation between rows. This configuration allows for good air circulation and reduces the shading effect between the panels. Additionally, the perimeter of the plant was compact and there was no vegetation under the solar panels, which facilitated maintenance and reduced fire risks. A little further south, we found Don Rodrigo Solar Park, a historically significant plant and the first large-scale PV plant in Europe. With a total capacity of 175 megawatt-peak and an annual performance of 300 gigawatt-hours (GWh), Don Rodrigo Park can meet the annual energy needs of approximately 93,000 average Spanish households (Global Energy Monitor Wiki 2024).

The plant uses polycrystalline silicon modules mounted on closely spaced horizontal arrays, thereby optimizing the available space. In addition, there was no vegetation under the panels, and the perimeter of the plant closely followed the morphological conformation of the terrain, minimizing the environmental impact. Despite their technical and structural differences, these two PV parks represent symbols of innovation and sustainability in energy production, and make up most of the solar resources in the Seville region. (Ordóñez et al., 2010).

Fig. 3 - Location of the study area in Sevilla. WGS84/ UTM Zone 30N reference system (EPSG:32630). Source: ESRI World View.

1.2. Dataset

This thesis used medium-resolution images from the S2 satellite and ESRI World Satellite Imagery imagery visualization service.

1.2.1. ESRI World Satellite

The ESRI World Imagery provides satellite and aerial imagery of one meter or better quality for most global landmasses, while also providing low-resolution satellite imagery worldwide (Venter et al., 2022). The current sources contributing to this map are as follows. TerraColor imagery at a 15 m resolution is available worldwide for application at small to medium map scales. (ESRI, 2009). Maxar offers several global-scale image-based mapping products. (Smith et al., 2021). The GIS User Community has provided valuable contributions in the form of high-resolution aerial photography. These images span 30 cm to 3 cm in resolution, and collaborative mapping stakeholders are encouraged to share them with Esri via shared contributions for broad dissemination. (Ladisa et al., 2024a). In this thesis, this layer was used to identify the GT features in both study areas.

1.2.2. S2 imagery

The S2 mission involves simultaneous and continuous monitoring of terrestrial and coastal areas through two identical satellites: Sentinel-2A (S2A) and Sentinel-2B (S2B) (Spoto et al., 2012). These satellites employ a Multispectral Instrument (MSI) sensor that captures images from the visible to the shortwave infrared regions of the electromagnetic spectrum. (Wang et al., 2016).

Their temporal resolution is five days owing to their dual-satellite configuration with a swath width of 290 km, capturing images in 13 spectral bands with varying spatial resolutions S2 Products Specification Document, 2021) (Phiri et al., 2020). Additionally, Level-2A (L2A) S2 products provide atmospherically corrected surface reflectance derived from Level-1C (L1C) products or top-of-atmosphere (TOA) radiance reflectance (Spoto et al., 2012).

To generate such data using the Sentinel-Toolbox or the standalone 'Sen-Cor' processor (Main-Knorn et al., 2017), users can first generate L1C data; this processing follows two fundamental phases: scene classification and atmospheric correction, which produces pixel classification maps identifying features

such as clouds/shadows, vegetation, land/desert cover, and water under various surfaces (Aguilar et al., 2020).

Atmospheric correction transforms TOA reflectance into surface reflectance (SR). (Wang et al., 2016). During L2A processing, defective, non-existent, and saturated pixels were excluded. The SR is calculated at the native resolution of the products, which varies depending on the band (10, 20, and 60 m) (Szantoi & Strobl, 2019). The specifications of S2A and S2B are presented in Tab. 1.

Tab.1 – Bands specifications S2 MSI

Band	Band Name	S2A		S2B		Spatial Resolution (m)
		Central λ (nm)	Band Width (nm)	Central λ (nm)	Band Width (nm)	
B2	Blue	496.6	98	492.1	98	10
B3	Green	560.0	45	559.0	46	10
B4	Red	664.5	38	665.0	39	10
B5	Red-edge 1 (RE1)	703.9	19	703.8	20	20
B6	RE2	740.2	18	739.1	18	20
B7	RE3	782.5	28	779.7	28	20
B8	NIR1	835.1	145	833.0	133	10
B8A	NIR2	864.8	33	864.0	32	20
B11	SWIR1	1613.7	143	1610.4	141	20
B12	SWIR2	2202.4	242	2185.7	238	20

This study used 22 S2 L2A (surface reflectance) images without clouds, corresponding to 12 images for Sevilla and the same number for Montalto di Castro.

In both areas, 10 cover an annual period of 2023 (Tab. 2), whereas the other two were published in 2018 (Table. 3). From the bands provided by the S2 products, only those with geometric resolutions of 10m and 20 m were selected.

This is because the 60m bands have a resolution too coarse to be used in this application context.

Tab.2 – S2 Usage Dates Time Series 2023

Study area	**Acquisition date and sensor**
Montalto di Castro	10.01.2023 (S2A) – 14.02.2023 (S2B) – 16.03.2023 (S2B) – 03.04.2023 (S2A) – 27.06.2023 (S2B) – 19.07.2023 (S2A) – 23.08.2023 (S2B) – 30.09.2023 (S2A) – 02.10.2023 (S2B) – 19.11.2023 (S2A)
Sevilla	13.01.2023 (S2A) – 05.02.2023 (S2B) – 02.03.2023 (S2A) – 11.04.2023 (S2A) – 11.05.2023 (S2A) – 25.06.2023 (S2B) – 07.07.2023 (S2A) – 14.08.2023 (S2B) – 13.09.2023 (S2B) – 17.11.2023 (S2A)

Tab.3 – S2 Usage Dates Time Series 2018

Study area	**Acquisition date and sensor**
Montalto di Castro	10.08.2018 (S2A) – 25.12.2028 (S2A)
Sevilla	12.01.2018 (S2A) – 12.08.2018 (S2A)

2. METHODOLOGY

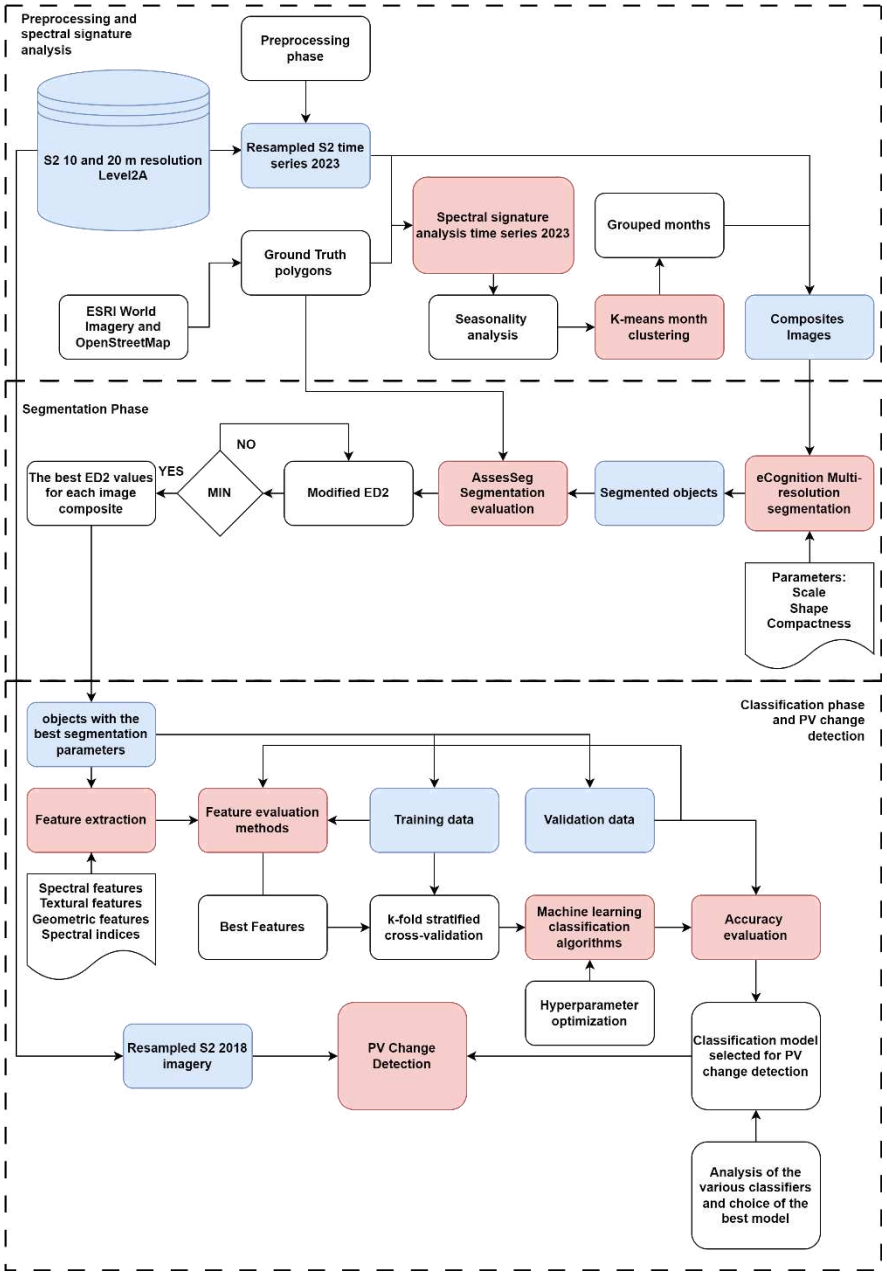


Fig. 4 - Thesis methodology workflow

A methodological flowchart of this study is shown in Fig. 4. Initially, data essential for analytical purposes, as outlined in the preceding section, were collected and subjected to a pre-processing phase. Thereafter, the methodology encompassing the study was divided into several stages.

1. Analysis of spectral signatures and examination of seasonal trends
2. Segmentation and evaluation of segmentation
3. Extraction and computation of features
4. Implementation of the classification models
5. Analysis of PV change detection (CD).

2.1. Pre-processing phase

As outlined in the preceding section, the S2 L2A multispectral images were subjected to atmospheric correction. Following this, a nearest-neighbor approach was utilized to resample all 20-meter resolution bands to retain geometric information at a higher spatial resolution of 10 m. Finally, selection areas for the study were delimited within the confines of pertinent S2 tiles using QGIS software tools.

2.2. Extraction of reference PV polygons

A GT dataset corresponding to the study sites of Montalto di Castro and Sevilla, as shown in Fig. 5 and 6, was extracted from OpenStreetMap by filtering all geometries using only the key:plant: method among the tags. The other geometries that were not updated on OpenStreetMap were manually digitized based on the ESRI WorldImage map. Because of the resolution of the S2 imagery, array clusters at a distance greater than 10m were considered as two distinct reference geometries.

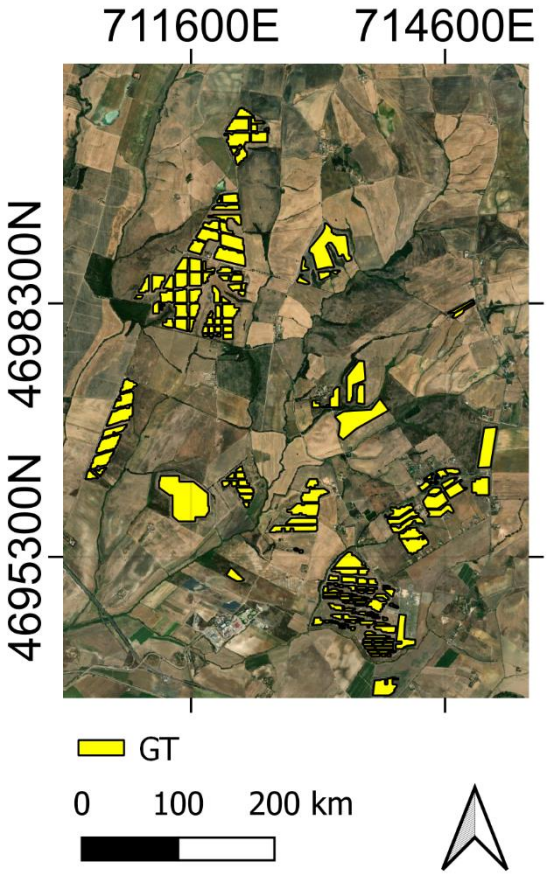


Fig. 5 - PV GT Montalto di Castro WGS84/ UTM Zone 32N reference system (EPSG:32632). Source ESRI World View.

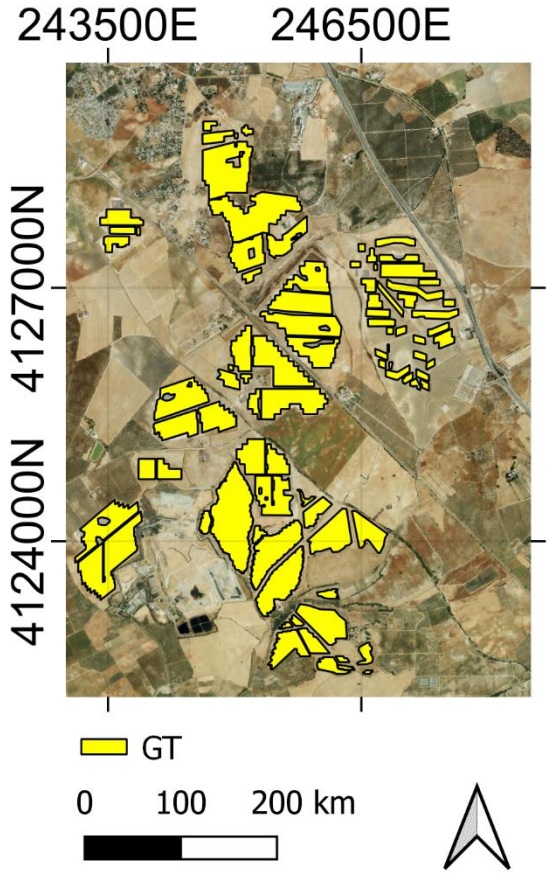


Fig. 6 - PV GT Sevilla WGS84/ UTM Zone 30N reference system (EPSG:32630). Source: ESRI World View.

2.3. Spectral signature analysis

Spectral signature analysis was performed on the pixels within the GT polygons once they had been extracted to obtain a detailed understanding of the multispectral properties of PV installations. Using the 2023 time series, spectral signature analysis was performed for both research locations to determine any seasonality and obtain a signature for each month.

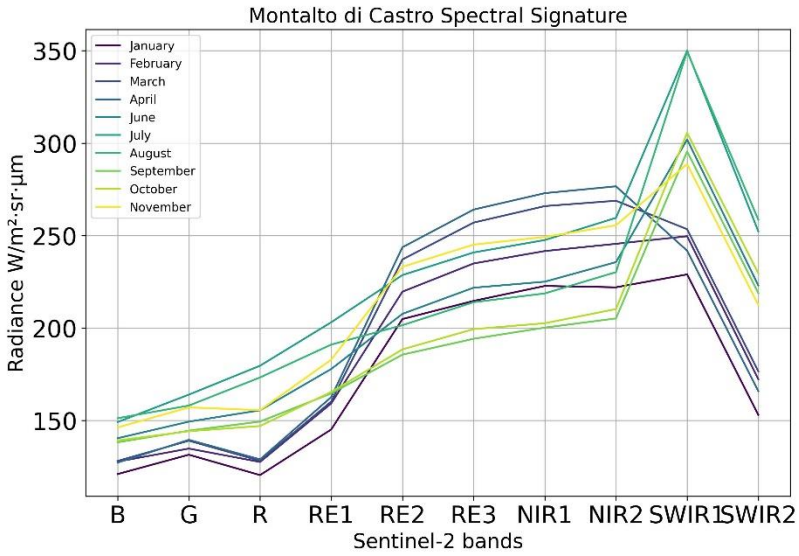


Fig. 7 – Spectral signatures time series 2023 Montalto di Castro.

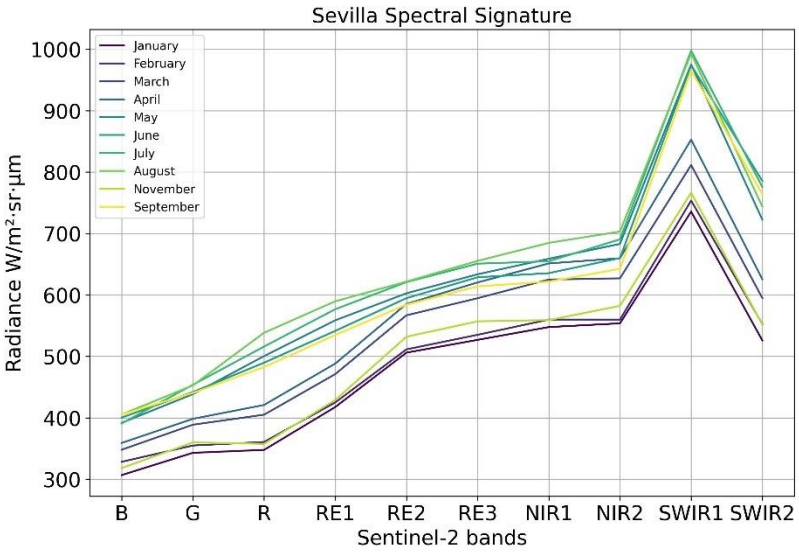


Fig. 8 – Spectral signatures time series 2023 Sevilla.

The spectral characteristics of the time series for the two research regions are shown in Fig. 7 and 8, respectively. It is evident from an analysis of the Montalto di Castro site that spectral signatures follow a seasonal rhythm. The spectral signatures exhibited an essentially continuous increasing trend during the summer and fall, rising from relatively low values in the blue band to higher values in the NIR2. A distinct peak was observed in the SWIR1 band, which was particularly noticeable in July and August. Following this peak, the spectral values in the SWIR1 band drop again, creating a "triangle" pattern that connects the NIR2, SWIR1, and SWIR2 bands. During the winter and spring months, the spectral behavior exhibited markedly different characteristics.

In the visible region up to the RE1 band, the PV signature showed lower values compared to the summer and autumn months, with a slight peak in the green band, akin to the average spectral signature of vegetation (Moroni et al., 2019). Between the RE1 and RE2 bands, a pronounced increase in reflectance values was observed, followed by an almost linear growth towards the NIR2

band. From the NIR2 band to the SWIR1 band, the spectral behavior during the winter and spring months was notably heterogeneous; in certain months, such as January and February, a slight increase was evident, whereas in other months, such as March and April, a decrease was noted. Finally, from the SWIR1 band to the SWIR2 band, a sharp decline in the spectral values was recorded, analogous to that observed in other months. Although less evident, slight seasonality can also be observed in the spectral signatures of the Seville area, as shown in the figure, within the time series. All spectral signatures exhibited a peak in the SWIR1 band, like that observed during the summer and autumn months in the Viterbo study area.

However, in general, Seville shows higher reflectance values than Viterbo owing to the different geographical and climatic characteristics of the area. In the summer months, higher peaks were recorded in SWIR1 compared to the winter months, and the trend of the signatures from B to NIR grew in an almost linear manner. However, in the winter and spring months, the signatures exhibited more variable slopes. From both study areas, it is possible to observe a general spectral signature with rather low values compared with the classic spectral signature curves of other coverages found in the literature and other studies (van Dijk et al., 2021).

In the Seville area and during the summer and autumn months in Montalto di Castro, the spectral curves were characterized by high reflectance in the SWIR1 bands, while other bands (visible and NIR) exhibited very low values. This behavior is attributed to the properties of the materials used in the solar panels, which are designed to absorb most of the solar radiation in the visible and NIR bands to convert solar energy into electricity (Wang et al., 2023b). Consequently, these materials reflect less in these bands, whereas the surfaces of the panels can more effectively reflect the SWIR wavelengths (Deepak et al., 2022). The SWIR bands of S2 are particularly sensitive to material composition, and PV panels, primarily made of silicon or other semiconductor materials, tend to reflect a significant portion of SWIR radiation compared to vegetation or water, which

absorb these wavelengths to a greater extent (Kumar & Thakur, 2021). During the winter and spring months in Montalto di Castro, the change in the shape of the curve occurs because the signatures show higher values in the NIR, RE1, RE2, and RE3 bands, which may indicate the presence of vegetation beneath the arrays. Vegetation has a characteristic spectral reflectance that significantly increases in the NIR bands and RE1, RE2, RE3 bands, reflecting these wavelengths more than the visible bands (Moroni et al., 2019). Suppose that the solar panels are primarily reflected in the SWIR, and the reflectance values of the NIR and red-edge bands increase during winter. In that case, it is plausible that the underlying vegetation influences the spectral signature (van Dijk et al., 2021).

It would be helpful to compute the NDVI for each month of the time series and visually examine the results to better observe this type of occurrence. An indicator that is frequently used to track the density and overall health of vegetation over time. Temporal grouping can benefit from the identification of seasonal patterns using NDVI trend analysis (Pettorelli et al., 2005).

2.4. Seasonal variation analysis of PV NDVI

Through the analysis of the spectral signatures presented in the previous section, it was observed that in both study areas, some installations exhibited variability in spectral values during different seasons of the year. This phenomenon, in addition to the presence of vegetation beneath the arrays, can also be attributed to the excessive distance between the arrays, causing the pixels to contain both parts of the array and soil, thus varying seasonally.

Considering these observations, to better manage and address the temporal variations in the data, it was decided to group the months with similar seasonal characteristics by analyzing the NDVI trends for both study areas. This method was adopted to create image composites that represented homogeneous seasonal periods, which were subsequently used in the segmentation and classification phases. The challenge of this study was to use composites as a starting point for segmentation to improve the image classification process.

Composite images are useful because they typically reduce the noise caused by temporal, meteorological, or atmospheric variations and can provide a more stable and representative image for each seasonal cluster (Yang et al., 2020). This could make classification easier because you will be working with more robust average images rather than single snapshots, reducing intermonthly variability and focusing on clearer seasonal patterns (Kaplan & Avdan, 2018).

The methodology for grouping months to create composites was not based on the simple seasonal division of the calendar (e.g., from June 21 to September 21 for the summer composite), but was determined by the internal pixels of the GT present in both study areas. The NDVI values were then calculated. This process was performed for all images in the 2023 time series, as illustrated in Fig. 9 and 10. For each image in the 2023 time series, the average NDVI values of pixels within the PV installations were calculated. These average NDVI values were subsequently used as inputs for the unsupervised K-means classification algorithm to obtain clusters of months in which the images exhibited spectral similarities.

From Fig. 9, it is evident that there is clear seasonality in the data at Montalto di Castro, attributable to the presence of vegetation beneath some PV installations. The NDVI values were relatively high from January to April, consistently exceeding the threshold of 0.3. During the summer months, the NDVI values decreased significantly, rising again in November. Unfortunately, it was not possible to analyze this trend in May because of the lack of cloud-free S2 images within the study area. In Seville, as shown in the figure, a trend similar to that observed in Viterbo was evident. The index values remained high (greater than 0.2) from January to May with a slight decrease. From May to September, the NDVI values were lower, increasing again in November. Unfortunately, images from September, October, and December were not available because of the dense presence of clouds.

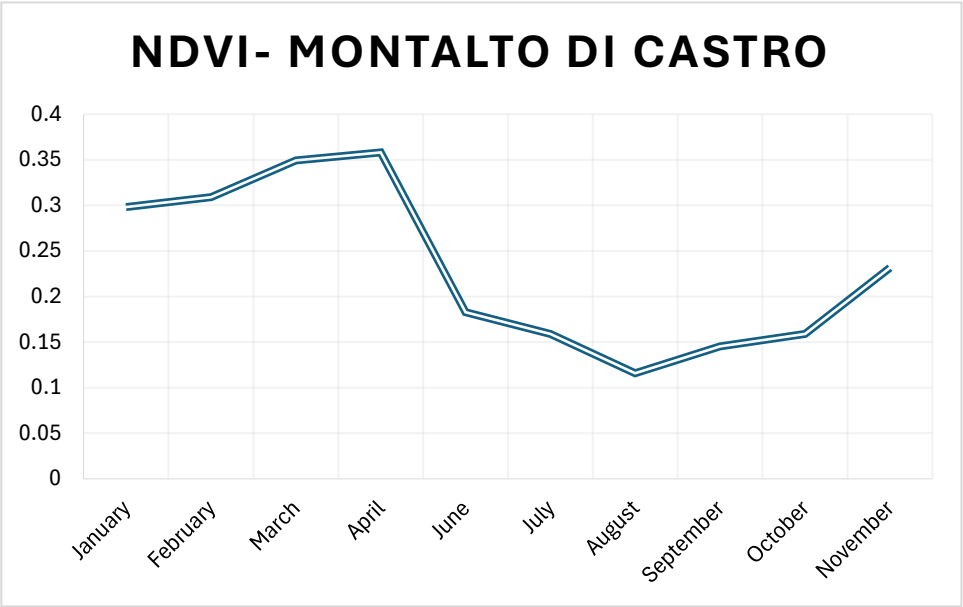


Fig. 9 - NDVI Montalto di Castro 2023 trend values.

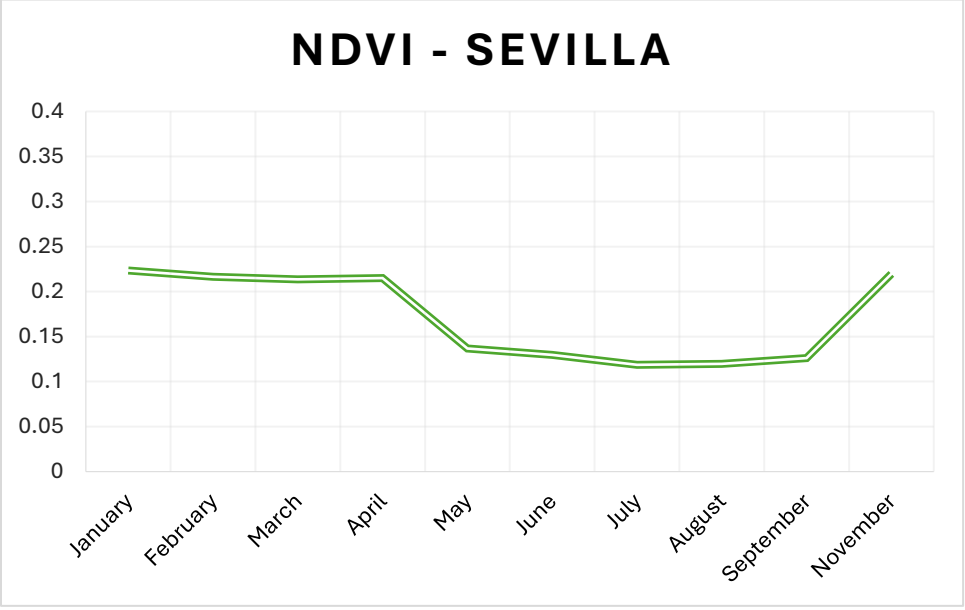


Fig. 10 - NDVI Sevilla 2023 trend values.

2.4.1 K-Means

An essential unsupervised ML technique for data grouping is K-means clustering (Umargono et al., 2020). Using this method, a dataset is divided into "K" clusters and each data point is assigned to the closest centroid cluster. "Mean" refers to the centroids inside each group; in the context of ML, "clusters" are groups of data points that are more similar to each other than to those in other groups (Sammouda & El-Zaart, 2021).

The process of starting a k-means analysis involves selecting a set of centroids at random, from which all other calculations are conducted. These calculations are typically performed using the ED technique, which finds the greatest match between the combined data points as individual records for each cluster (Ahmed et al., 2020). The centroids are then recalculated by iterative cycles until convergence, going over the record set using predetermined repetition settings. This procedure requires the employment of a fresh cluster centroid determination algorithm, which is used automatically until no further iterations within predetermined tolerance levels are possible, yielding a stable result (SAPUTRA et al., 2020). The initial determination of the value of "K" is a crucial factor to consider when applying the K-means clustering algorithm. However, choosing the appropriate number of clusters is not always straightforward, as the optimal number often depends on the dataset and specific task (Sammouda & El-Zaart, 2021).

Various methods can be used to determine the optimal value of (K); among them, the most commonly used method is the “Elbow Method” (Syakur et al., 2018). This method proposes an appropriate number of clusters by tracking the inertia explained by different numbers of clusters and identifying the “elbow” point, where the rate of inertia decreases abruptly and stabilizes (Umargono et al., 2020). In the K-means elbow method, inertia is the sum of the squared distances between each point and the centroid of the cluster to which it belongs (Marutho et al.2018). Low inertia indicates that points are very close to their centroids, suggesting that the clusters are compact and well-defined, whereas high

inertia indicates that points are farther from their centroids, suggesting that the clusters are more dispersed (Nainggolan et al., 2019).

2.4.1 Clustering results and composites creation

Fig. 10 and 11 illustrate the application of the elbow method in both the study areas. It is observed that the slope from 1 to 2 of (k) is very steep and then almost completely stabilizes in both areas. The inertia values of the graph for the Seville area are significantly lower than those of the Montalto di Castro area, thus indicating more compact clusters and closer to the centroids than Montalto di Castro. Therefore, in both study areas, we proceeded with the clustering algorithm using (k = 2), and the results are shown in Tab. 4.

The results of this algorithm suggest the creation of two composite images for each study area. These results were consistent with those of the analysis previously performed on the signatures shown in Fig. 7 and 8. In particular, in Lazio, the results propose the creation of a composite (Composite1M) for January, February, March, April, and November and another composite (Composite2M) for June, July, August, September, and October. In Seville, the situation is similar, as the algorithm proposes to create a composite (Composite1S) for January, February, March, and April and another composite (Composite2S) for June, July, August, and September. Composite images were created in a Python environment using the Rasterio package, which allows the manipulation of geospatial data.

Composites were generated by calculating the median of the image bands within each cluster. The median, rather than the mean, was chosen to avoid the influence of outlier pixels in the final data. Fig. 13, 14, 15, 16, 17 and 18 show Composite1M, Composite2M, Composite1S, Composite2S, CompositeALLM, and CompositeALLS, respectively. The last two are composites containing all months of the year 2023 for the individual study areas and were compared with the other composites mentioned above.

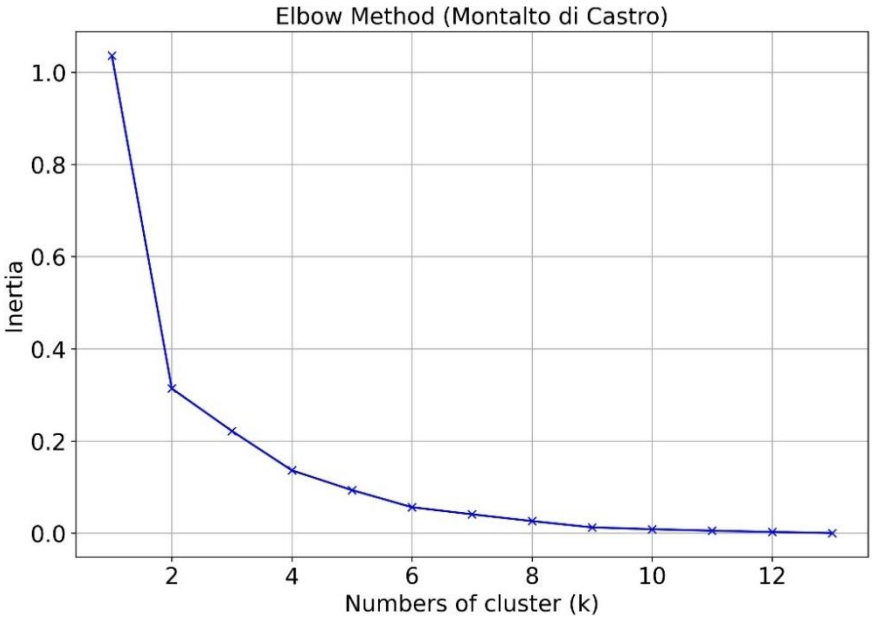


Fig. 11 - Elbow method results in Montalto di Castro.

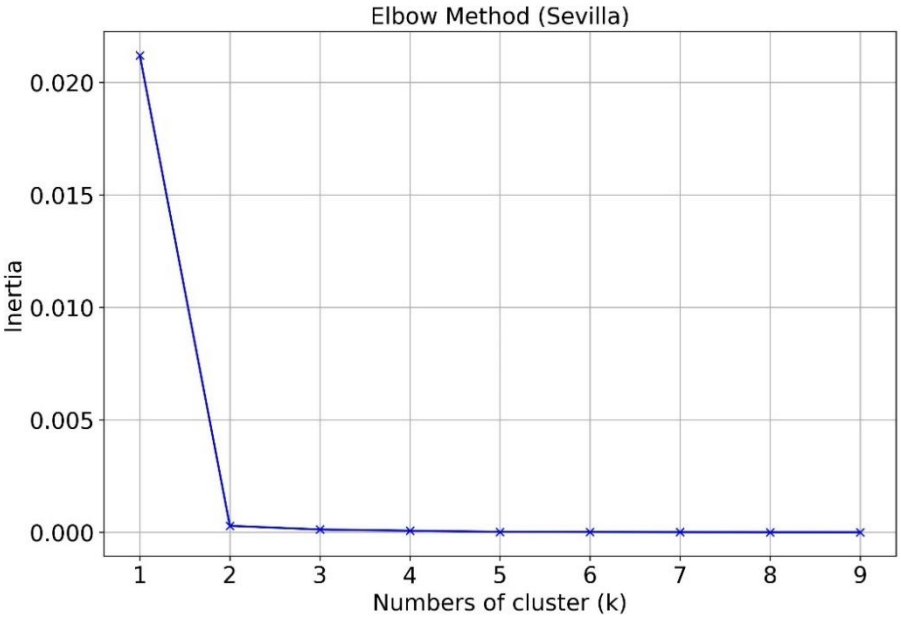


Fig. 12 - Elbow method results in Sevilla.

Tab.4 – K-means results

Montalto di Castro		**Sevilla**	
Month	**Cluster**	**Month**	**Cluster**
January	1	January	1
Febraury	1	February	1
March	1	March	1
April	1	April	1
June	0	May	0
July	0	June	0
August	0	July	0
September	0	August	0
October	0	September	0
November	1	November	1

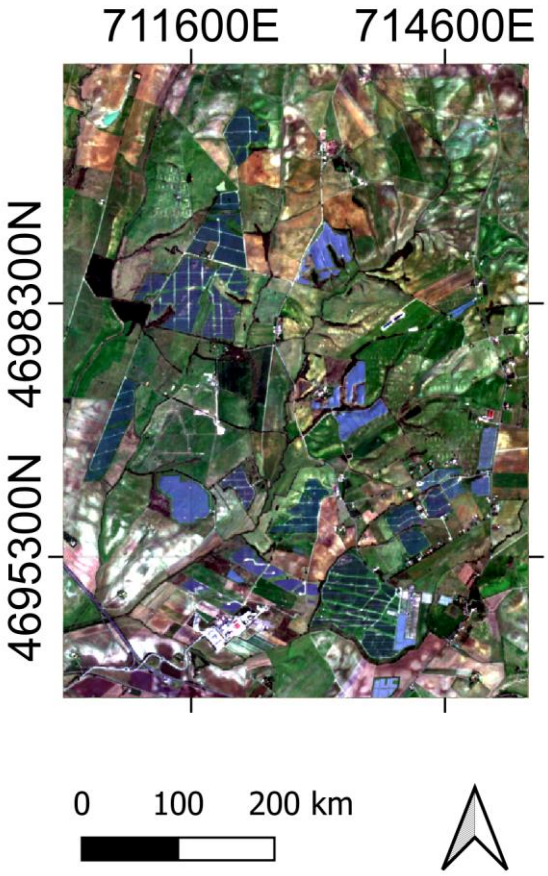


Fig. 13 - S2 Composite1M. WGS84/ UTM Zone 32N reference system (EPSG:32632).

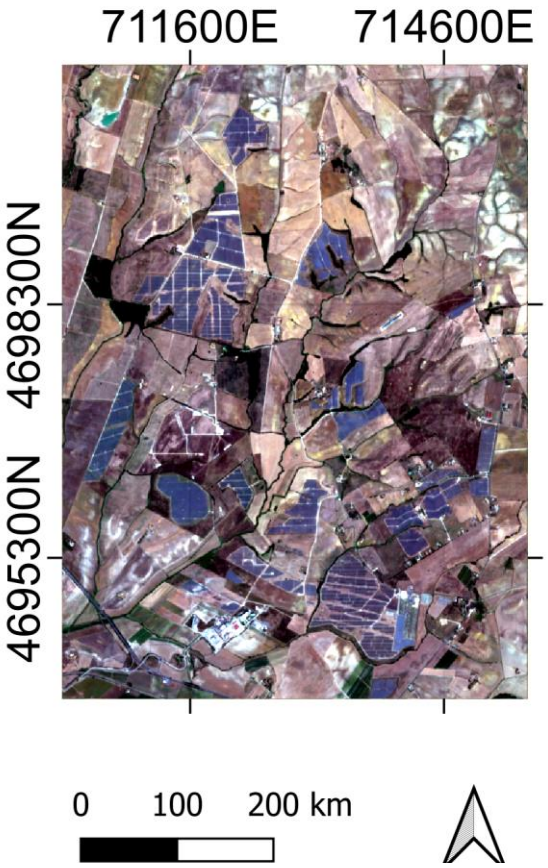


Fig. 14 - S2 Composite2M. WGS84/ UTM Zone 32N reference system (EPSG:32632).

Fig. 15 - S2 Composite1S. WGS84/ UTM Zone 30N reference system (EPSG:32630).

Fig. 16 - S2 Composite2S. WGS84/ UTM Zone 30N reference system (EPSG:32630).

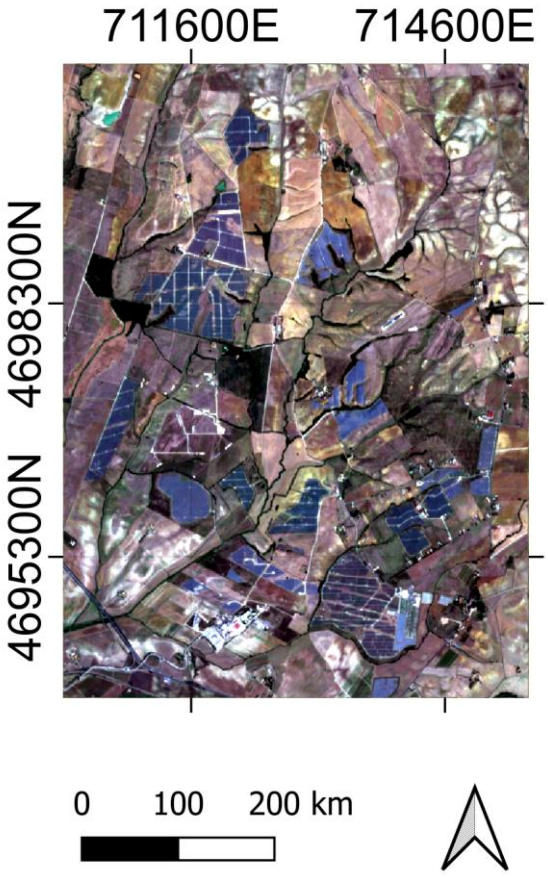


Fig. 17 - S2 CompositeALLM. WGS84/ UTM Zone 32N reference system (EPSG:32632).

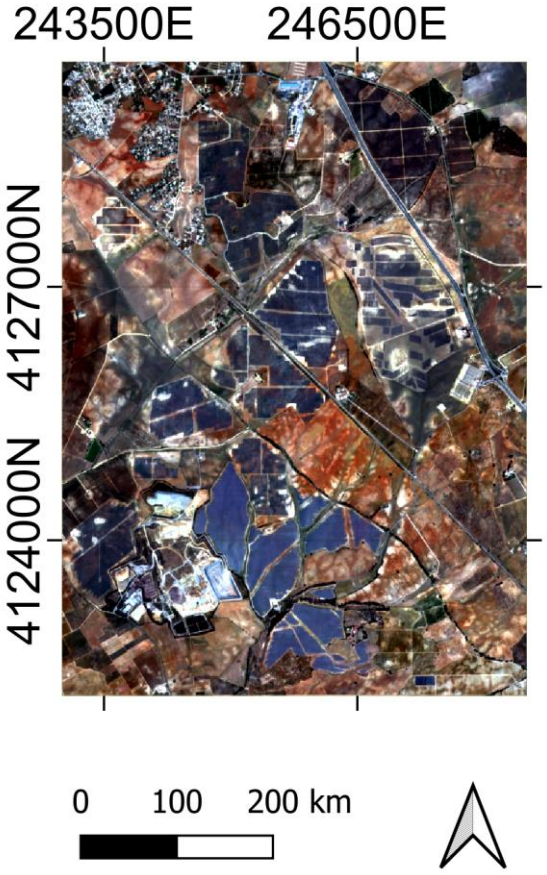


Fig. 18 - S2 CompositeALLS. WGS84/ UTM Zone 32N reference system (EPSG:32630).

2.5. Segmentation

The first crucial stage in the OBIA approach is the segmentation phase (Watkins & van Niekerk, 2019). In this procedure, an image is divided into objects that are groups of uniform and semantically significant pixels (Labib & Harris, 2018). Color, geometric traits, and spectral properties can influence item division.

Despite the presence of numerous image processing segmentation algorithms, the multi-resolution segmentation (MRS) algorithm has been widely used in various remote sensing applications (Tian & Chen, 2007). This particular algorithm is available in the eCognition software (Trimble, Munich, Germany) and has shown excellent results when applied to tasks such as mapping of plastic greenhouses (Aguilar et al., 2016); (Jiménez-Lao et al., 2022). It has also been successfully applied to the mapping of PV plants, demonstrating superior performance compared to other segmentation algorithms used in the same context (Ladisa et al., 2022b); (Ladisa et al., 2022a). In this study, the MRS algorithm was selected for the segmentation phase. The results produced by the algorithm depend on multiple parameters that require optimization by users (Graf et al., 2020). The research objectives influence the choice of these parameters because different terrestrial objects or segments exhibit both spectral and structural variations, implying that the optimal segmentation values may differ between various terrestrial objects (Benz et al., 2004).

However, selecting the best segmentation is a complex task, as blurry edges or image noise can hinder optimal segmentation, even for a single terrestrial object. In such cases, it is necessary to establish a framework through which the quality of the resulting segmentation can be evaluated to ensure generalizability across different images. Therefore, iterative evaluations can determine when the best-fit segmentations have been obtained before implementation. Therefore, the parameters for selecting the MRS algorithm were based on an iterative process for evaluating and improving overall segmentation quality.

2.5.1 MRS

The MRS method represents a bottom-up region-fusion approach in which smaller objects (e.g., pixels) are combined to form larger objects through several iterative steps (Benz et al., 2004). The outcome of the process is influenced by three main components: the shape, scale parameter, and compactness. The scale parameter refers to the maximum heterogeneity allowed per segment, whereas the compactness determines the weight assigned to the smoothness criteria (Baatz et al., 2008). The shape governs the weighting variables of the textural and chromatic features (Tian & Chen, 2007). In addition to these parameters, it is necessary to determine band combinations in advance to ensure the efficiency of the MRS algorithm. In this study, eCognition v10.v1 was specifically used to perform this algorithmic operation.

The required outputs were generated along with the results of the exercises obtained using the same software platform. In theory, the optimization method minimizes the weighted heterogeneity of the image objects throughout the fusion process, resulting in a combination of neighboring image objects. (Jiménez-Lao et al., 2022). The heterogeneity gain, f (Equation 1), is determined by color and shape, with wcolor and wshape contributing to the heterogeneity adjustment (Benz et al., 2004).

$$f = w_{color} \times \Delta h_{color} + w_{shape} \times \Delta h_{shape}, \in [0,1,], \quad (1)$$

$$w_{shape} \in [0,1,], w_{color} + w_{shape} = 1$$

Δhcolor, representing the variations in spectral heterogeneity, allows for multivariate segmentation using the weighting of image channels, wc (Equation 2) (Benz et al., 2004).

$$\Delta h_{color} = \sum_c w_c \times (n_{merge} \times \sigma_{cmerge} - (n_{obj1} \times \sigma_{cobj1c} + n_{obj2} \times \sigma_{cobj2}) \quad (2)$$

Where n is the total number of pixels, σ is the standard deviation within the object, the indices c refer to the channel, obj1 and obj2 refer to the objects to be merged, and merge refers to the merged object. The shape heterogeneity Δhshape (Equation 3) is described by the compactness (Equation 4) and smoothness (Equation 5) of the object's shape, and it refers to the improvement of the shape.

$$\Delta h_{shape} = w_{comp} \times \Delta h_{comp} + w_{smooth} \times \Delta h_{smooth}, \quad w_{comp} + w_{smooth} = 1 \quad (3)$$

$$\Delta h_{comp} = n_{merge} \frac{l_{merge}}{\sqrt{b_{merge}}} - \left(n_{obj1} \frac{l_{obj1}}{\sqrt{b_{obj1}}} + n_{obj2} \frac{l_{obj2}}{\sqrt{b_{obj2}}} \right) \quad (4)$$

$$\Delta h_{smooth} = n_{merge} \frac{l_{merge}}{b_{merge}} - \left(n_{obj1} \frac{l_{obj1}}{b_{obj1}} + n_{obj2} \frac{l_{obj2}}{b_{obj2}} \right) \quad (5)$$

Where l is the perimeter of the object and b is the perimeter of the object's bounding box. The merging process stops if the smallest growth described by f is greater than the limit set by the scale parameter.

2.5.2 Experimental design for the image segmentation phase

The experiments in the segmentation phase were conducted as follows. Once the composites were generated, they were subjected to the iterative process of the MRS algorithm, thereby obtaining different vector layers in the output with varying combinations of shape and scale. Specifically, based on the different sizes of the GTs in the two study areas, it was decided to analyze the scale range from 20 to 80 in Montalto di Castro and from 40 to 180 in Seville. The shape range was analyzed from 0.1 to 0.9, and the compactness was maintained at 0.5, which was analyzed for both study areas. Finally, the various outputs were compared with GT objects to evaluate the segmentation accuracy through metrics. The novelty of this study lies in comparing the best metric results for segmentation accuracy among the different composites, and obtaining the layer with the best combination of MRS parameters to proceed with the subsequent OBIA phases.

2.5.3 Segmentation quality assessment

Quality assessment was conducted using a supervised methodology. The difference between a collection of reference polygons and image segments algorithmically forms the basis of supervised evaluation techniques (D'agostino, 2014). Conversely, unsupervised techniques evaluate the quality of segmentation by examiinter-classclass and inter-class variability and homogeneity (Jozdani & Chen, 2020). Research has shown that supervised assessment techniques that consider geometric boundary correctness are more suitable for artificial target recognition (Chen et al., 2018). Consequently, choosing the optimal segmentation algorithm or determining the segmentation parameter for a certain segmentation method may be performed successfully using supervised evaluation techniques (Zhang et al., 2015).

Additionally, factors such as the variability of soil types beneath PV systems and variations in perimeter configurations affect the intraclass homogeneity within a particular PV segment. Conversinter-classclass heterogeneity can be influenced by regions with high concentrations of PV installations (Ladisa et al.,

2022b). These factors may render unsupervised evaluation techniques unsuitable for assessing the accuracy of PV segmentation (Senel et al. 2023). When the intersection area between a reference polygon and candidate segment exceeds 50% of the total area of the polygon or candidate segment, the candidate segment can be described as a corresponding segment according to the area overlap criterion (D'agostino, 2014).

(Liu et al., 2012) proposed the Euclidean Distance 2 (ED2), based on area overlap criteria, which evaluate both geometric and arithmetic discrepancy criteria. The ED2 measure calculates the quality of segmentation in two-dimensional Euclidean space using the potential segmentation error (PSE) and the number of segments ratio (NSR) (Witharana & Civco, 2014). However, Jozdani and Chen (2020) reported that ED2 does not always correctly identify optimal segmentation results. Yang et al. (2014) proposed advanced segmentation evaluation metrics such as ED3 to address this issue. ED3 integrates the geometric and arithmetic discrepancies (Equation 6).

$$ED3_{ij} = \sqrt{\frac{(OS2_{ij})^2 + (US2_{ij})^2}{2}} \quad ,where$$

$$OS2_{ij} = 1 - \frac{area(x_i \cap y_j)}{area(x_i)}, \tag{6}$$

$$US2_{ij} = 1 - \frac{area(x_i \cap y_j)}{area(y_j)},\ y_j \in Y_{c_i} \cup Y_{d_i}$$

While the range of this metric varies between zero and one, a perfect match is represented by zero (Yang et al., 2015). When the number of reference polygons increases, reference polygons without corresponding segments can be formed because the rule of at least one 50% overlap ratio is not met. In this case, the number of reference polygons used is lower than the original number. Therefore, by altering the ED2 measure, (Novelli et al., 2017) established $PSQnew$ and

$NSRnew$ to prevent biases in the calculation of both PSE and NSR. The modified ED2 (MED2) formula is given in Equation 7.

The number of excluded reference polygons is denoted by n, the maximum segment for a single reference polygon is indicated by max ($|si–rk|$), and the maximum number of corresponding segments in a single reference geometry is indicated by $vmax$. A greater MED2 value implies an arithmetic and/or geometric mismatch, whereas an MED2 value of zero indicates a perfect match between the segmentation and the reference dataset.

$$ED2 = \sqrt{PSE_{\text{new}}^2 + NSR_{\text{new}}^2}$$

$$PSE_{\text{new}} = \frac{\sum | s_i - r_k | + n \times \max(| s_i - r_k |)}{\sum | r_k |} \quad (7)$$

$$NSR_{\text{new}} = \frac{| m - v - n \times v_{\max} |}{m - n}$$

In this study, the MED2 metric was calculated using the open-access command line tool (AssesSeg) developed by Novelli et al. (2017). For the metrics calculation, 50% of the reference polygons were used randomly to avoid making the segmentation result too dependent on the GT, as has already been suggested in previous studies.

2.6. Extraction Features

The multiple features outlined in Tab. 5 were computed within the segmentation output objects to evaluate the effects of various spectral characteristics on PV segmentation. These included spectral features (encompassing mean, minimum, maximum, range, and standard deviation of all S2 bands), texture features from the gray-level co-occurrence matrix (GLCM) (such as angular second moment, contrast, correlation, dissimilarity, entropy, and homogeneity), spectral indices, and geometric object features (such as shape index and pixel width).

To examine the impacts of spectral indices, several indices were employed in the thesis: NDVI (Pettorelli et al., 2011), NDBI (He et al., 2010) ,

Normalized Difference Water Index (NDWI) (Gao et al., 1996), NDTI (Kassouk et al., 2020), Modified Normalized Difference Water Index (MNDWI) (Singh et al., 2015), Enhanced Vegetation Index (EVI) (JIANG et al., 2008), and BUAI (Zha et al., 2003). These indices were selected primarily based on a literature review of PV mapping; their formulations are provided in Tab. 5.

Tab.5 – Extracted features for the image segments.

	Feature	**Description or formula**
Spectral information	Mean	Mean values of bands for each image object
	Minimum	Minimum values of bands for each image object
	Maximum	Maximum values of bands for each image object
	Median	Median values of bands for each image objects
	Standard Deviation	The standard deviation of bands for each image
	Difference	Difference between previously calculated maximum and minimum
Geometric information	Shape index	The smoothness of an image object border
	Width pixel	Width of the object in pixel unit
Texture information	GLCM ASM	The angular second moment of all directions
	GLCM CON	The contrast of all directions
	GLCM COR	Correlation of all directions
	GLCM DIS	Dissimilarity of all directions
	GLCM ENT	The entropy of all directions
	GLCM HOM	Homogeneity of all directions

Spectral indices	Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{(NIR1 - Red)}{(NIR1 + Red)}$
	Normalized Difference Builtup Index (NDBI)	$NDBI = \frac{(SWIR1 - NIR1)}{(SWIR1 + NIR1)}$
	Normalized Difference Water Index (NDWI)	$NDWI = \frac{(Green - NIR1)}{(Green + NIR1)}$
	Normalized Difference Tillage Index (NDTI)	$NDTI = \frac{(SWIR1 - SWIR2)}{(SWIR1 + SWIR1)}$
	Modified Normalized Difference Water Index (MNDWI)	$MNDWI = \frac{(Green - SWIR1)}{(Green + SWIR1)}$
	Enhanced Vegetation Index (EVI)	$EVI = G \times \frac{(NIR1 - Red)}{(NIR1 + C1 \times Red - C2 \times Blue + L)}$
	Built-Up Area Index (BUAI)	$BUAI = NDBI - NDVI$

In the feature-extraction phase, 75 characteristics were evaluated. These included 60 spectral features (derived from 10 bands with six statistics each), 6 textural features, 2 geometric information points, and 7 spectral index characteristics.

2.6.1 Feature Optimization Method

Decreasing the number of features reduces model redundancy, as most features tend to be correlated (Yan & Zhang, 2015). Additionally, the increased diversity of features in OBIA complicates their subjective selection (Granitto et al., 2006). To address this issue, the RFE method was implemented to evaluate and extract the most suitable features for PV recognition. This was accomplished using the scikit-learn library, which is accessible through Python. This algorithm was chosen over other methodologies based on the positive results previously achieved in this field of application.

2.6.2 RFE

RFE aims to select features by gradually reducing the number of considered characteristics, employing an external estimator that assigns weights to each feature (Darst et al., 2018). Initially, the importance of every feature was evaluated by training the estimator on the entire dataset. Subsequently, the least significant features are eliminated. This process was repeated until a final set of selected features was obtained (Bahl et al., 2019). In the thesis, two different methods were employed as external estimators for feature weighting: RF.

2.7. Classification Modeling

The third stage of OBIA involves a classification model that is regularly trained to categorize objects in the images. (Kucharczyk et al., 2020) divided this phase into three subsections: sampling design, object labeling, and classification. The sampling design involves generating training data and verifying sample locations.

An ideal test set should include as many high-quality samples as possible, with an equal number of samples for each class. Labeling involves assigning all objects from the training and test samples, produced or collected during the sampling design phase, to two classes. Reference data were chosen using S2 images and the ESRI World Imagery Service. The final subsection, classification, includes the model training and object classification in the images. There are various classification approaches, such as rule-based or supervised methods. (Ma et al., 2017) noted that supervised object-based classification techniques have been a crucial component of remote sensing research for land cover mapping since 2010. Recently, in addition to traditional ML algorithms such as RF, SVM, or K-nearest neighbor (KNN), newer ensemble learning algorithms have been increasingly used in OBIA. Consequently, this thesis tested several methods, including KNN, SVM, and RF, for PV mapping.

2.7.1 KNN

KNN is a neighbor-based classification algorithm categorized as instance-based or non-generalizing learning (Zhang & Zhou, 2007). This approach retains the training data instances without generating a comprehensive internal model. Classification is determined through a straightforward majority vote among the nearest neighbors of each point, and an unknown object is assigned to the most prevalent data class among its closest neighbors. (Zhang et al., 2017b). The outcomes of this algorithm are primarily affected by parameter k, which is heavily dependent on the data. It has been noted that a higher k value diminishes the impact of noise, but simultaneously blurs the boundaries of classification. (Zhang et al., 2018).

2.7.2 SVM

The SVM is one of the most widely used kernel-based statistical learning algorithms. As a non-parametric method, the SVM is not affected by the underlying data distribution. The primary goal of the SVM algorithm is to determine a hyperplane that divides the dataset into a specific number of classes in line with the training samples. (Patle & Chouhan, 2013). Kernel functions can mitigate the increase in the computational complexity caused by dimensionality. (Yue et al., 2003) The selection of the kernel is vital for SVM applicability, as it directly impacts the classification process outcomes. Various kernel types are employed in SVMs, including linear, polynomial, radial basis function, and sigmoidal kernels (Haifeng Wang & Dejin Hu, 2005). In this thesis, linear and radial basis function kernels, which are commonly used in the remote sensing literature, were tested.

2.7.3 RF

The RF algorithm was one of the earliest successful bagging approaches, combining bagging sampling, random feature selection, and random decision trees (Gislason et al., 2006). Building on the concept of classification and regression trees, RF employs multiple self-learning tree ensembles, referred to as

forests (Chen & Ishwaran, 2012). The classification performance of a single tree is enhanced through the bagging approach, which uses bootstrap aggregation for each tree. Additionally, trees are constructed using a random subset of features at each split node, substantially reducing overfitting and decreasing inter-tree correlation (Belgiu & Drăguţ, 2016). This process effectively eliminates overfitting, leading to improved final predictions for classification problems (Chen & Ishwaran, 2012).

2.7.4 Cross-validation and hyperparameter tuning

This thesis employed the stratified k-fold technique to divide the training and test datasets for model performance evaluation. This approach segments the cross-validation process into stratified folds across n iterations. (Zhang & Liu, 2023). Each k-fold set contains a roughly equal proportion of samples from every class. The study utilized a stratified k-fold with 10 folds repeated 3 times (Morales-Barquero et al., 2019). The stratified k-fold was implemented using the sci-kit-learn library in Python 3.7. To determine the optimal hyperparameters for the algorithms described earlier, the GridSearchCV function from sci-kit-learn was used. This function conducts a comprehensive search for hyperparameter selection, focusing on parameters that are not learned by the classification algorithm.

2.7.5 Accuracy assessment

The confusion matrix (error matrix) serves as the foundation for various descriptive and analytical statistical methods (Lyons et al., 2018). An error matrix was developed to evaluate the accuracy of the thematic maps generated via OBIA and several metrics were derived. To evaluate the effectiveness of the mapping techniques under consideration, error matrices were constructed for each scenario and accuracy evaluation metrics were calculated (Cai et al., 2018). These metrics include OA, producer accuracy (PA), user accuracy (UA), and F1 score (F1). These metrics were calculated using specific formulas in Equations 8, 9, and 10.

$$UA = \frac{TP}{(TP + FP)} \quad (8)$$

$$PA = \frac{TP}{(TP + FN)} \quad (9)$$

$$F1 = 2 \times \frac{(UA \times PA)}{(UA + PA)} \quad (10)$$

In this context, TP, FP, TN, and FN represent true positive, false positive, true negative, and false negative, respectively. TP indicates the sum of correctly identified PV pixels, whereas FP signifies the total number of non-PV pixels mistakenly classified as PV. Additionally, FN represents PV pixels erroneously labeled as non-PV pixels, and FP denotes non-PV pixels incorrectly categorized as PV pixels. These metrics range from 0 to 1, with higher values indicating a superior classification accuracy. Throughout this thesis, the F1 metric was primarily employed to assess the precision of the PV mapping capabilities of the algorithms.

2.8. PV CD

The main objective of this last phase of the thesis is to evaluate the impact of PV plant land consumption from 2018 to 2023 in the two study areas using the best classification models obtained from the previous phase of the thesis. At a practical level, the images of 2018 were segmented with the parameters obtained from the MRS evaluation, and the best-performing classification models classified the objects obtained. It is evident that, after training the models using composite images, each model was applied exclusively to the months included in the time interval of the composite images. For example, the best model developed by Composite1M was applied only to images related to the months included in the composite. This approach allowed the analysis of how PV expansion has varied in recent years in terms of area and to know what changes in the soil have occurred.

2.8.1. Corine Land Cover (CLC)

To analyze land occupation by PV panels classified according to land cover, we employed CLC cartography. CORINE, an acronym for the COoRdination of Information on the Environment, has been an international operational project since 1985 that has addressed various environmental issues. One of its primary features is "a land cover inventory divided into 44 classes, presented as a cartographic product at a scale of 1:100,000, available for most areas of Europe' (SEE, 1995). Specifically, we considered the third level of classification of CLC data updated to 2018, which provides a highly specific characterization. For instance, agricultural areas represent the first level of classification (code 2), permanent crops the second level (code 2.2), while olive groves constitute the third level (code 2.3.3).

3. RESULTS AND OBSERVATIONS

3.1. MRS Results and Evaluation

MRS analysis was conducted for each composite image derived from vegetation variation analysis at two study sites (Montalto di Castro and Seville) using S2 satellite imagery and reflectance storage scales (percentage and 32-bit). The segmentation phase was performed a semi-automatic eCognition rule set, maintaining a compactness of 0.5, following 's approach for plastic greenhouse (Novelli et al., 2017). In this thesis, segmentation was performed using all available bands, as they are considered useful for PV mapping. Notably, the band weights were treated as equal. Fig. 19, 20, and 21 show the results of the segmentation evaluation based on the MED values for Composite1M, Composite2M, and CompositeALLM, respectively.

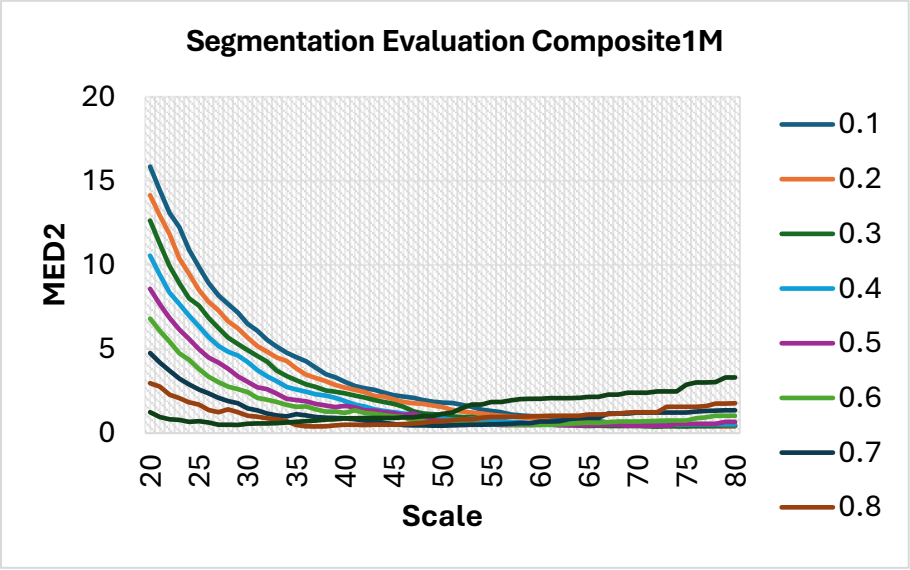


Fig. 19 - Composite 1M segmentation evaluation chart.

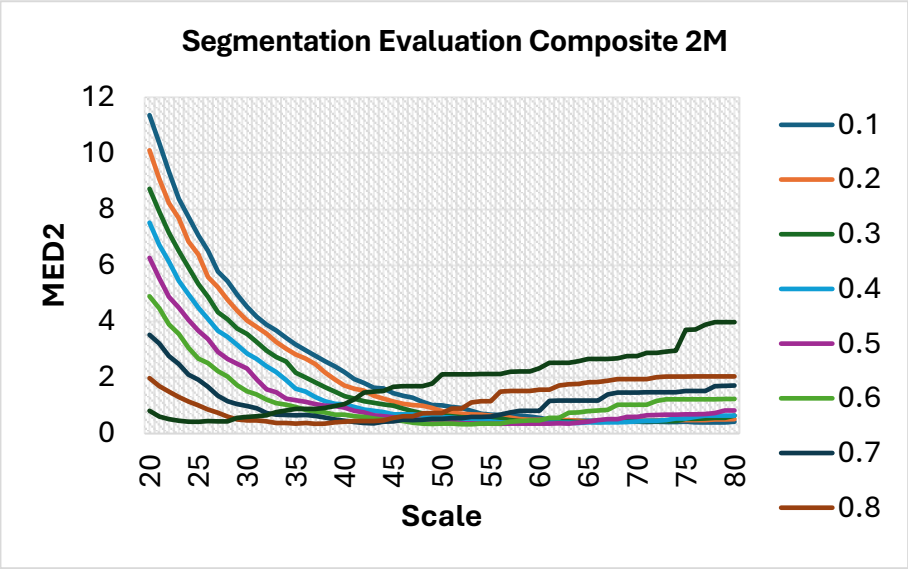


Fig. 20 - Composite 2M segmentation evaluation chart.

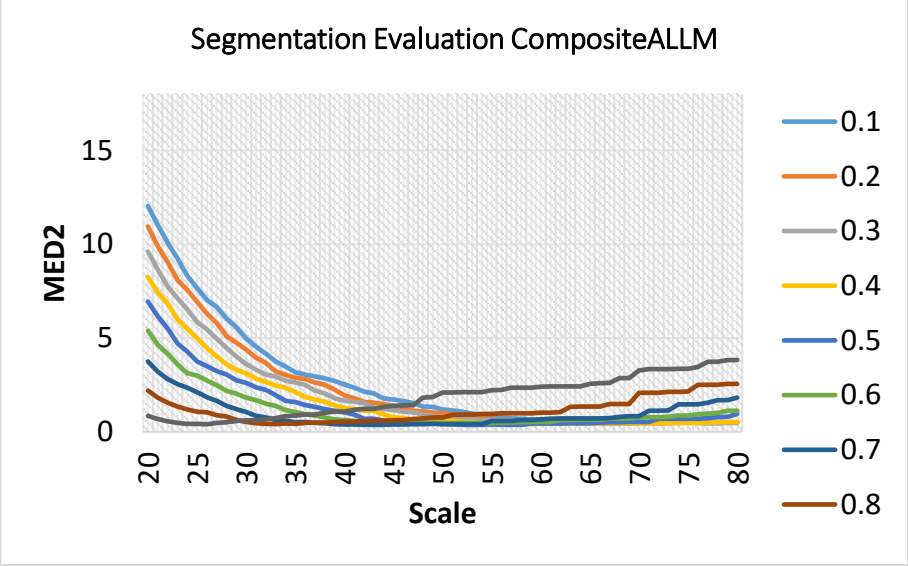


Fig. 21- CompositeALLM segmentation evaluation chart.

In all figures, the x-axis depicts the scale interval, and the y-axis shows the MED2 values. The various curves in the figures represent different shapes. For all three composite images, at lower scale values, MED2 values were notably high for shape values between 0.1 and 0.6, owing to the creation of objects smaller than the reference geometry. As the Scale values increase, the MED2 values gradually decrease for shape curves from 0.1 0.6. The shape curves from 0.7 to 0.9 exhibit an almost opposite trend. Fig. 22, 23, and 24 show the segmentation evaluation results based on the MED values for Composite1S, Composite2S, and CompositeALLS, respectively.

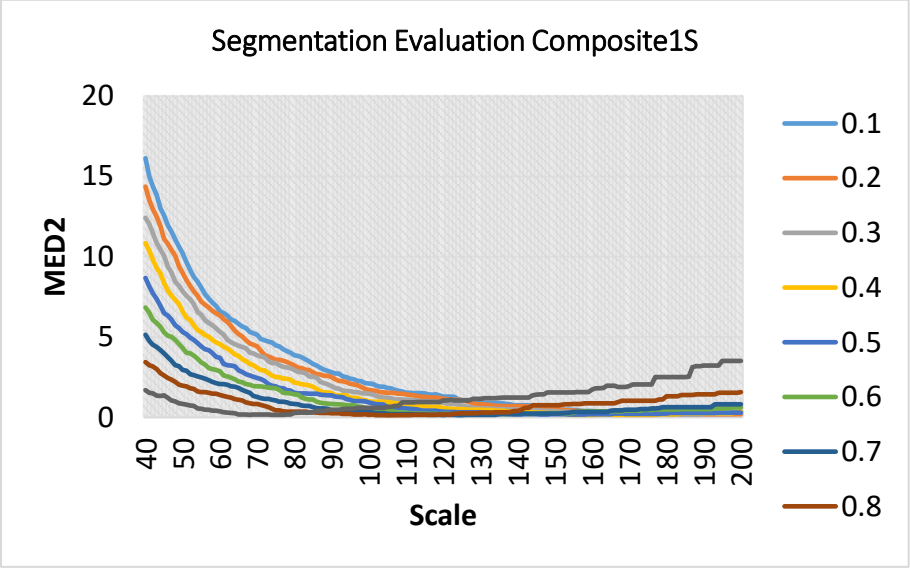


Fig. 22 - Composite1S segmentation evaluation chart.

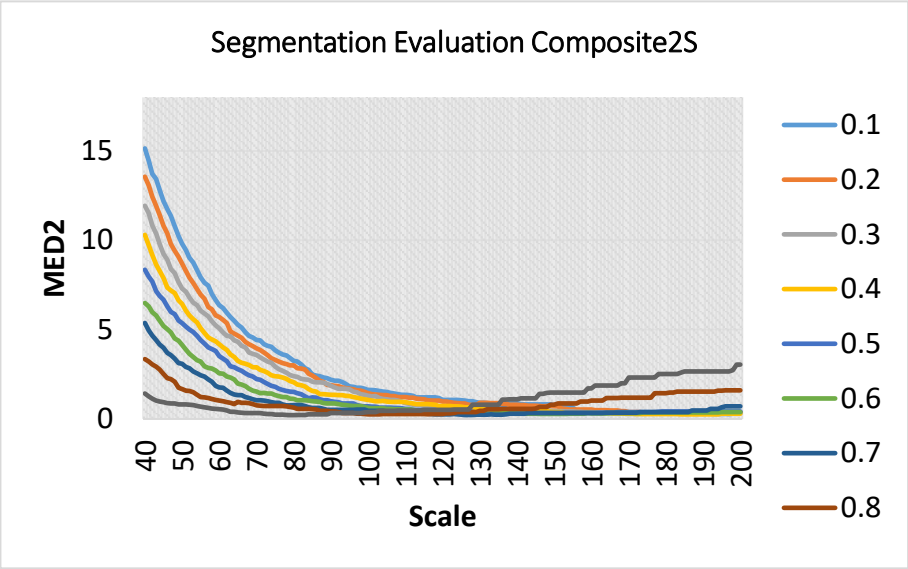


Fig. 23 - Composite2S segmentation evaluation chart.

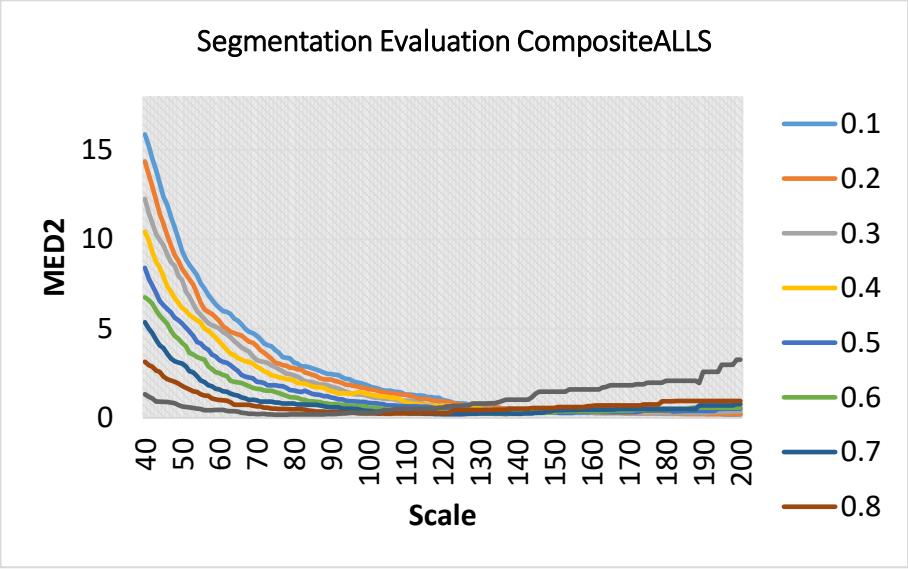


Fig. 24 - CompositeALLS segmentation evaluation chart.

The MED2 trends in Seville's study area resemble those of Montalto di Castro, but with a wider range and higher values owing to the presence ofFig.r areas of plants. Tab. Figure 6 shows the optimal MED2 values for each composite image in both study areas. Despite the previously noted differences in Montalto di Castro, the best MED2 values were relatively consistent across composite images. Nevertheless, Composite2M yielded the most effective segmentation based on the MED2 results. Consequently, objects derived from the MRS with scale parameters of 52, 0.6, and 0.5, extracted from Composite2M, were utilized in subsequent phases for the Montalto di Castro study area. Seville generally exhibited superior MED2 values across all the three composite images. However, unlike Montalto di Castro, Composite1S produced the best MED2 value, favoring winter months. As a result, the subsequent thesis phases for the Seville study area proceeded using objects obtained from the MRS with a scale parameter of 105 and a shape of 0.8.

Tab.6 - MED2 best values.

Images	**Montalto di Castro**		**Sevilla**	
	MED2	**MRS Parameters**	**MED2**	**MRS Parameters**
Composite1	0.389	Scale = 69 Shape = 0.2	**0.132**	Scale = 105 Shape = 0.8
Composite2	**0.332**	Scale = 52 Shape = 0.6	0.213	Scale = 79 Shape = 0.9
CompositeALL	0.364	Scale = 56 Shape = 0.5	0.176	Scale = 77 Shape = 0.9

Fig. 25 and 26 show the segmentation results obtained using the optimized MRS parameters for Montalto di Castro and Seville, respectively. Visual examination corroborated the findings of the MED2 calculation. Seville's segmentation outputs appear more uniform than those of GT, whereas Montalto di

Castro exhibits excessive segmentation and merged geometries encompassing multiple distinct GT objects. This discrepancy can be attributed to the differences in both geometries and distribution patterns, which are characterized by smaller sizes and greater dispersion in Montalto di Castro than in Seville.

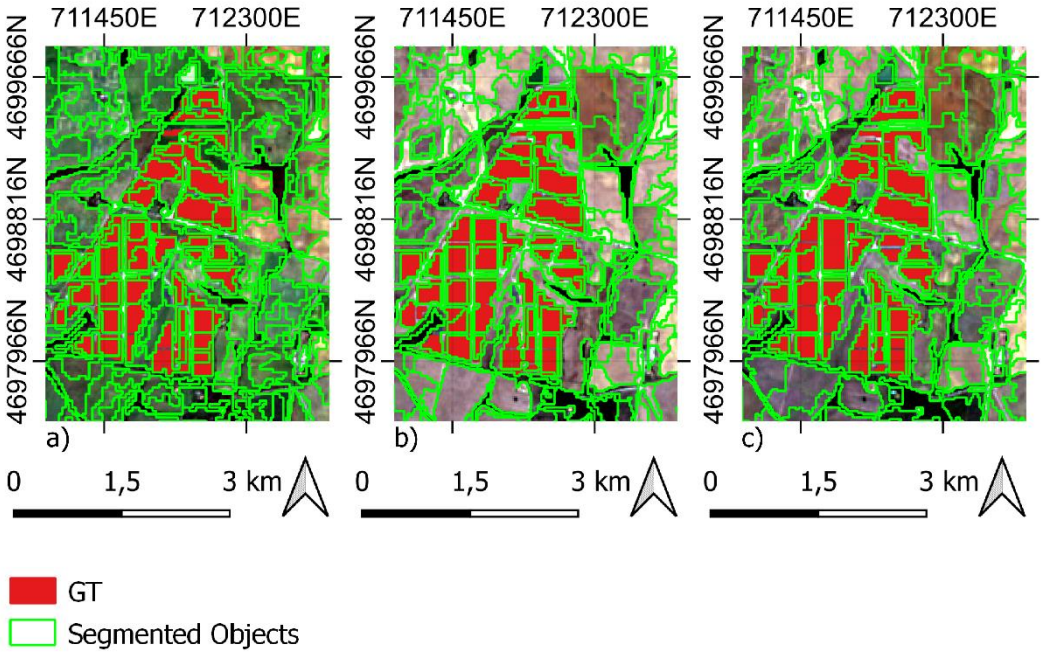


Fig. 25 - Output Segmentation Montalto di Castro: a) Composite1M; b) Composite2M; c) CompositeALLM. WGS84/ UTM Zone 32N reference system (EPSG:32632).

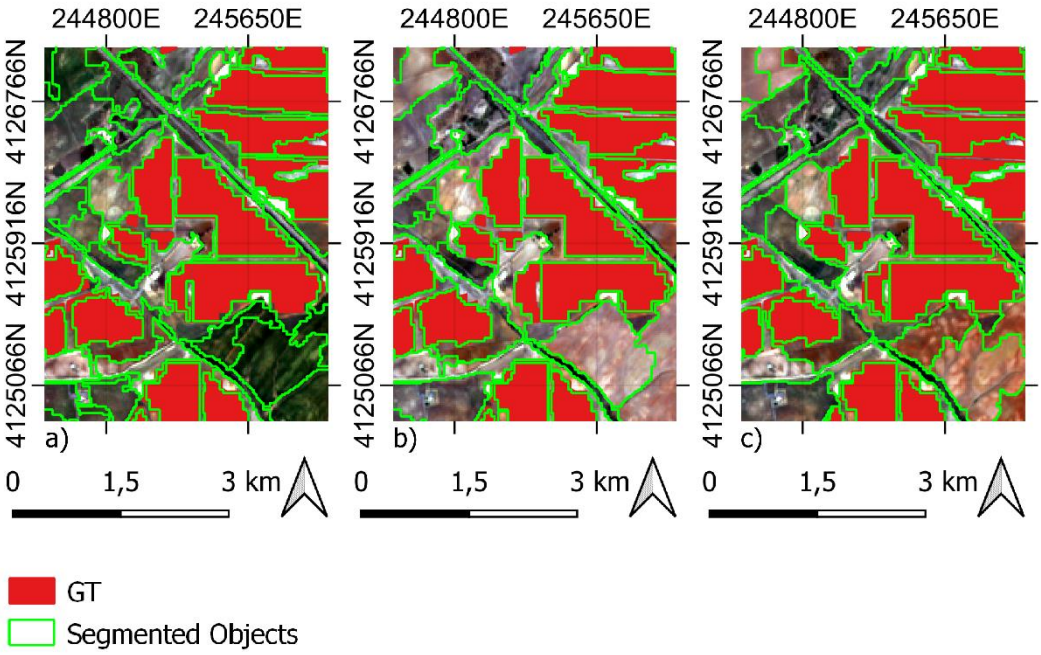


Fig. 26 - Output Segmentation Sevilla: a) Composite1S; b) Composite2S; c) CompositeALLS. WGS84/ UTM Zone 30N reference system (EPSG:32630).

3.2. Feature Extraction

Following the image composite segmentation stage, optimal segmentation outputs were selected based on the initial phase results. The MRS segmentation algorithm was applied with a fixed compactness of 0.5, utilizing all S2 bands. As indicated in Tab. 5, for the Montalto di Castro study area, the chosen segmentation output had a scale of 52 and shape of 0.6, derived from Composite2M (which uses summer months). In contrast, for the Sevilla study area, the selected output had a scale of 105 and a shape of 0.8, originating from Composite1S (which uses winter months).

Using the above procedure, 75 features were extracted from Composite1M, Composite2M, Composite1S, and Composite2S, encompassing spectral, textual, geometric, and spectral indices. CompositeALL features were not extracted because they were unnecessary for developing models specific to the

seasonal periods. The image segments were then manually labeled as PV or NO-PV samples. After training the variables for all segmented objects, the RFE method was employed to evaluate the feature space. Subsequently, PV classification accuracy was assessed by training the RF classifier and calculating its F1 score. Before algorithm testing, hyperparameter optimization was conducted using the GridSearchCV function in Python 3.7.

The model evaluation utilized repeated stratified k-fold cross-validation with ten splits and three repetitions. Tab. 7 presents the optimal hyperparameters for the chosen RF algorithm determined using GridSearchCV.

Tab.7 - The Optimal parameters for the RF algorithm selected within Gridsearch.

Algorithm	Parameter	Optimal value
RF	max_depth	10
	min_samples_split	2
	num_leaves	10

Tab.8 - Top 10 ranked features for Composites S2.

Feature Rank	Montalto di Castro		Sevilla	
	Composite1M	Composite2M	Composite1S	Composite2S
1	EVI	EVI	NDBI	NDBI
2	NDTI	NDBI	NDTI	NDTI
3	Std Dev Blue	NDTI	BUAI	Std Dev NIR1
4	Max SWIR2	Std Dev Blue	EVI	Std Dev RE3
5	Diff SWIR1	Std_Dev NIR2	Std Dev NIR1	BUAI
6	Std Dev RE1	BUAI	Std Dev SWIR2	StdDev
7	NDBI	Std Dev NIR1	Std Dev SWIR1	Std Dev RE2
8	Median Blu	Std Dev RE3	Max RE3	Std Dev NIR2
9	Max RE1	Median Blue	Diff RE3	Diff RE3
10	Diff Blue	Max NIR2	Max SWIR2	Median Blu

Tab. 8 reveals that the EVI index is most significant for Montalto di Castro, whereas NDBI is most relevant for Sevilla. This indicates that vegetation is a key distinguishing factor in Montalto di Castro, whereas built-up areas were more prominent in Sevilla.

Despite this, it is noteworthy that the NDBI index, which ranks highest in Sevilla, also achieves a good importance score in Montalto di Castro, ranking second in Composite2M, and seventh in Composite1M. The NDTI index was the second most important for both locations, highlighting the relationship between the SWIR1 and NIR2 bands typical of PV, as previously observed in the spectral signature study. Another interesting aspect of this result is that the standard deviation of various bands (Blue, NIR, RE) appears frequently in both locations, suggesting that the spectral variability within objects is likely the most important among the statistical values of the bands. Texture and geometric features did not achieve high rankings. Some indices, such as BUAI and Max SWIR2, are present in both locations, but with different priorities, indicating that specific local characteristics influence the importance of these indices.

Fig. 27, 28, 29, and 30 illustrate the variation in accuracy, expressed as the mean F1 scores based on 10-fold cross-validation repeated three times, as a function of the total number of features (ranging from 1 to 75) used to train the RF classifier for the four composites under study Consequently, at the Montalto di Castro study site, the maximum accuracy was achieved using the RFE method with 33 features, yielding an F1 score of 97.89% for Composite1M, and with 54 features, yielding an F1 score of 98.13% for Composite2M. Regarding the Seville study site, the maximum accuracy was attained with 54 features, achieving an F1 score of 97.1% for Composite1S and with 39 features, achieving an F1 score of 99.6% for Composite2S.

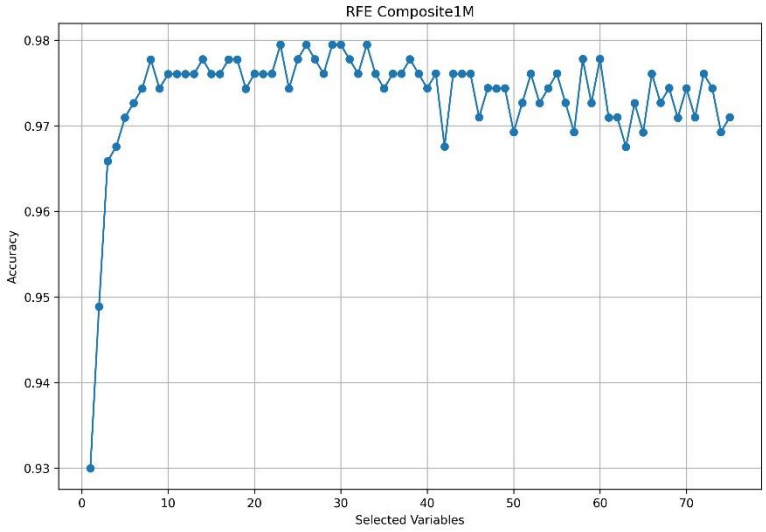


Fig. 27 - RFE results Composite1M.

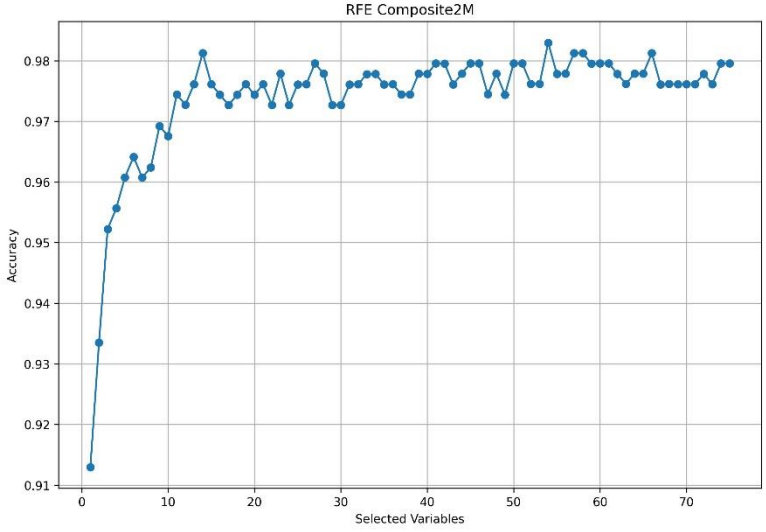


Fig. 28 - RFE results Composite2M.

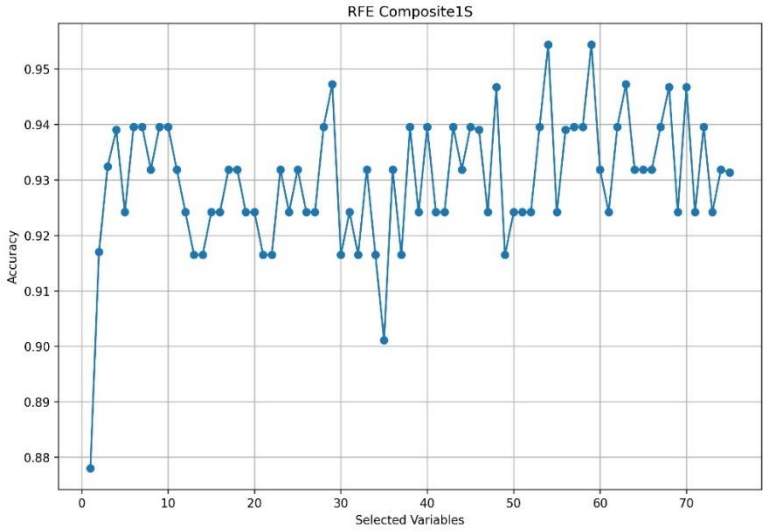


Fig. 29 - RFE results Composite1S.

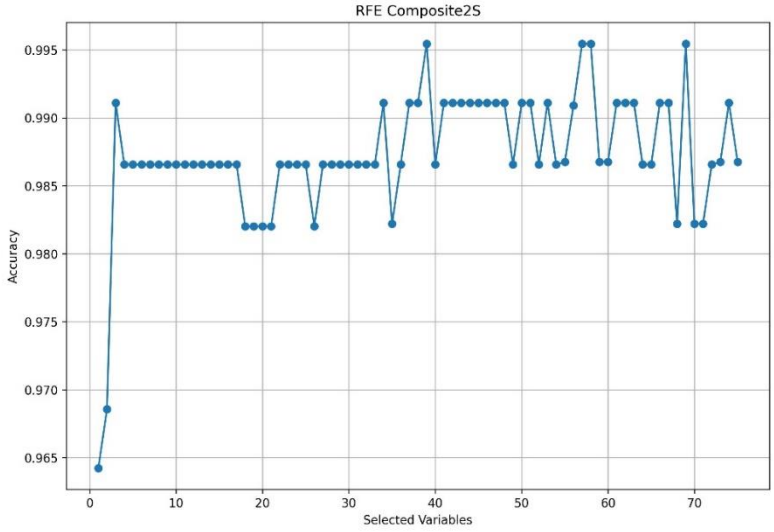


Fig. 30 - RFE results Composite2S.

3.3. Classification model results

Based on the results obtained in the preceding sections, image classification and accuracy evaluation phases were conducted for the S2 composite images. As outlined in the previous section, the image segments obtained in the MRS component and the extracted features based on these objects were used to train the models (KNN, RF, SVM, and GB). Notably, the hyperparameters selected through GridSearchCV, as reported in Tab. 9, were employed to train the algorithms.

Furthermore, the hyperparameters listed in Tab. 7 were utilized for the RF algorithm. Following the training process, the image objects were classified, the accuracy of the models was evaluated by considering the stratified k-fold with 10 subdivisions and three repetitions, and F1 scores were calculated for each subdivision. All experiments were conducted on a computer equipped with an 11th generation Intel® Core™ i7-11370H processor, 16 GB of RAM, a base frequency of 3.30 GHz, and a Windows 10 Pro (64-Bit) operating system. The analysis was performed in Python 3.7 environment.

Tab.9 - Optimal hyperparameters for the algorithm selected within the Gridsearch.

Algorithm	**Parameter**	**Optimal value**			
		Montalto di Castro		**Sevilla**	
		Composite1 M	**Composite2 M**	**Composite1 S**	**Composite2S**
SVM	C	10	10	10	10
	Kernel	Poly	Poly	Linear	Linear
KNN	n_neighbors	5	5	5	5

Accuracy results for the training objects were provided in Tab. 10 for both study sites. The RF performed better in Montalto di Castro than in Sevilla. The F1-Scores were higher for Composite1M and Composite2M than for Composite1S and Composite2S, suggesting that the algorithm is more effective in the Montalto di Castro area. SVM showed a slight variation in performance between the two test sites.

In Montalto di Castro, SVM scored higher for Composite2M than for Composite1M, whereas in Sevilla, Composite1S scored slightly higher than Composite2S. KNN performed slightly better in Montalto di Castro for Composite2M than for Composite1M. In Sevilla, the scores were more balanced between Composite1S and Composite2S, with a slight preference for Composite2S. Overall, RF demonstrated strong classification ability at both test sites. F1-Scores above 90% in all categories indicated high accuracy.

The SVM algorithm achieved good results, with F1-Scores ranging from 87.36% to 93.67%. Although slightly lower than RF, these scores indicate good classification ability. The KNN algorithm achieved the lowest F1-Scores among the three algorithms tested, with values ranging from 86.05% to 90.24%.

Tab.10 – F1 – Score Training data

Algorithm	F1- Score (%)			
	Montalto di Castro		Sevilla	
	Composite1M	Composite2M	Composite1S	Composite2S
RF	96.20	97.44	90.00	94.74
SVM	87. 36	93.67	89.47	92.31
KNN	86.05	90.24	87.18	87.80

Tab. 11 presents the test data results for the various classifiers. RF demonstrated superior performance in Montalto di Castro compared to Sevilla. Composite1M and Composite2M achieved higher F1-Scores than Composite1S and

Composite2S, indicating the algorithm's greater efficacy in the Montalto di Castro area.

The SVM exhibits slight performance variations between the two test regions. In Montalto di Castro, SVM attained a higher score for Composite2M than for Composite1M, whereas in Sevilla, Composite2S marginally outperformed Composite1S. KNN yielded superior results for Composite1S in Sevilla compared to Composite1M in Montalto di Castro. However, KNN achieved a higher score for Composite2M in Montalto di Castro than Composite2S in Sevilla.

The RF test data outcomes closely resembled those of previous results, with minor fluctuations. In Montalto di Castro, Composite1M exhibited a slight decrease, whereas Composite2M improved. Both the composites in Sevilla experienced minor declines. The SVM demonstrated a decreased performance on the test data compared with the earlier results. KNN exhibited a substantial performance reduction in the test data relative to the previous outcomes. These findings suggest that RF is the most robust and effective algorithm for both Montalto di Castro and Sevilla datasets, consistently performing well on both test and training data.

Tab.11 – F1 – Score Test data

Algorithm	F1- Score (%)			
	Montalto di Castro		Sevilla	
	Composite1M	Composite2M	Composite1S	Composite2S
RF	**95.95**	**97.67**	**89.47**	**92.31**
SVM	84.85	89.87	82.67	88.89
KNN	76.47	85.71	82.50	83.33

In Fig. 31 and 32, the visual results of the classification for Montalto di Castro and Sevilla are presented, respectively. Upon visual inspection, it is evident that the results obtained from the accuracy metrics are confirmed in both

study areas; specifically, the RF models demonstrate greater precision in PV classification. However, despite nearly identical accuracy results, the outcomes in Sevilla appear to be more precise and better follow the contours of the plants.

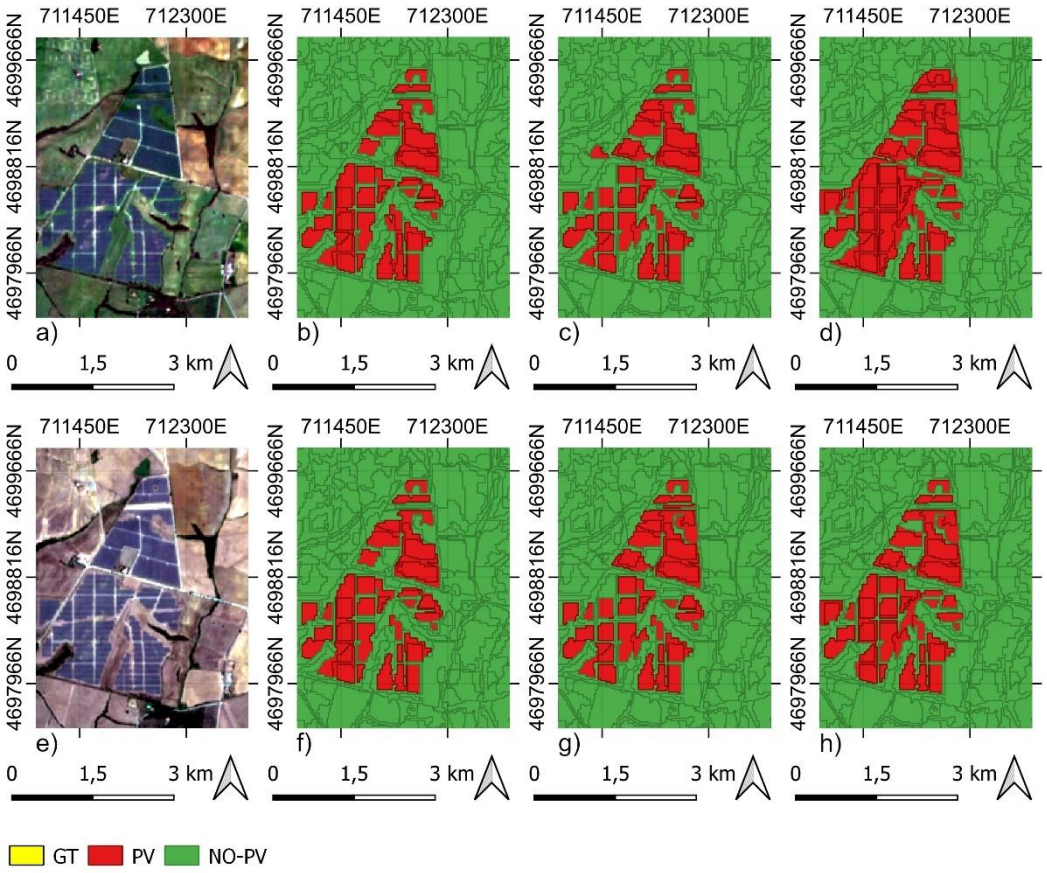


Fig. 31 - Montalto di castro classification model results: a) Composite1M RGB; b) Composite1M RF classification; c) Composite1M SVM classification; d) Composite1M KNN classification; e) Composite2M RGB; f) Composite2M RF classification; g) Composite2M SVM classification; h) Composite2M KNN classification.

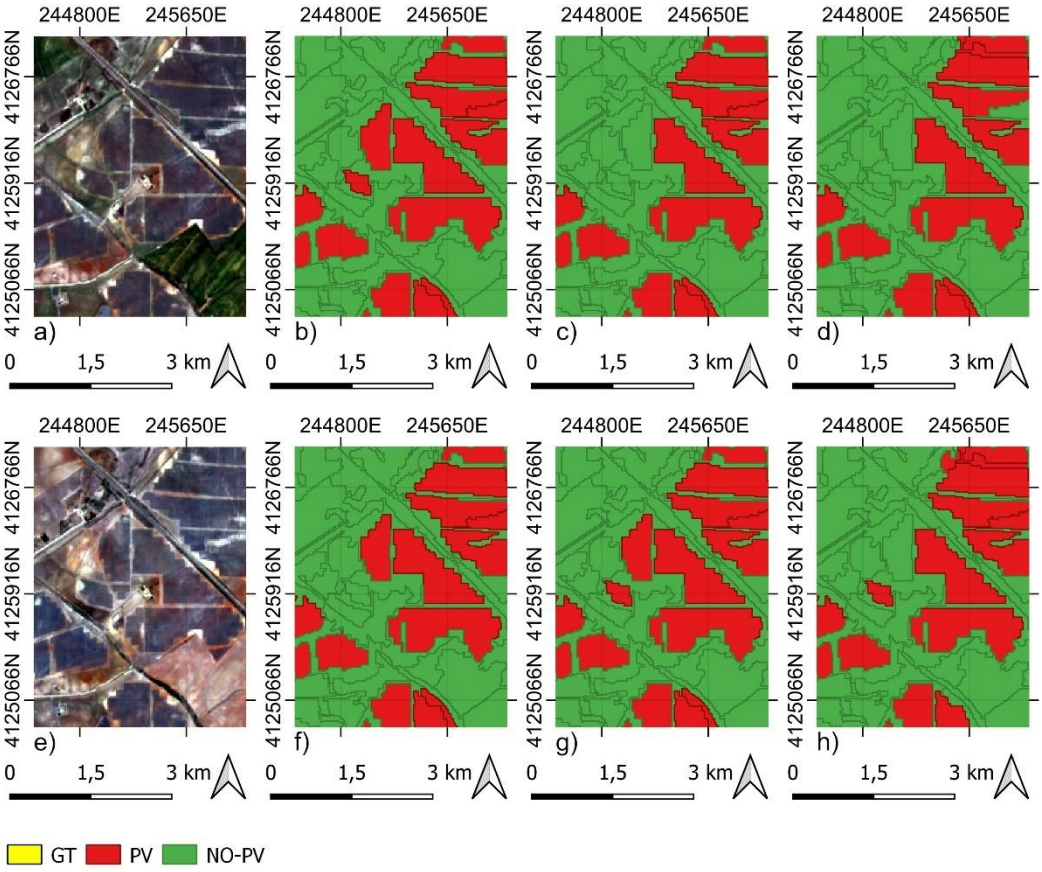


Fig. 32- Sevilla classification model results: a) Composite1S RGB; b) Composite1S RF classification; c) Composite1S SVM classification; d) Composite1S KNN classi-fication; e) Composite2S RGB; f) Composite2S RF classification; g) Compo-site2S SVM classification; h) Composite2S KNN classification.

3.4. PV CD results

In the final section of the thesis, an analysis was conducted on the expansion of areas containing PV installations from 2018 to 2023, as well as the changes in land use types, using the most effective classification models obtained in the

previous step. The RF model achieved the highest accuracy metrics for both study areas, and was therefore chosen for appliFig. to the 2018 data. Tab. Figure 12 shows the F1-score results obtained from the 2018 images. The 2018 images were subjected to the same pre-processing phase as the 2013 time series.

Tab.12 – F1 – Score data 2018 imagery.

Study Area	Imagery	F1 score (%)
Montalto Di Castro	10/08/2018	97,12
	25/12/2018	98,34
Sevilla	12/01/2018	93,22 %
	12/08/2018	86,19 %

For Montalto di Castro, the image acquired on August 10 was classified using the RF classification model trained on the Composite1M image. Conversely, the image from December 25 was analyzed using the RF model trained on the Composite2M image. Regarding Seville, the image from January 12 was classified using the RF model trained on the Composite1S image, whereas the image from August 8 employed the RF model trained on the Composite2S image. The results obtained, measured through the F1-score, highlight the excellent adaptability of the classification models in the respective study areas.

3.4.1 Calculating the area and type of soil changed

The areas encompassing objects classified as PV were calculated using the results of the 2018 image classification. Tab. Fig. 13 shows the area values and their corresponding percentage errors for the actual areas. In Montalto di Castro, the area values were almost identical to the original values in both the images. In Sevilla, however, the area is zero in the January image because there is a PV plant. In the August image, the percentage error was quite high because of the presence of false positives due to the construction sites of PV plants.

Tab.13 – Areas calculation

Study Area	Imagery	Area (Ha)	percentage error (%)
Montalto Di Castro	10/08/2018	259,12	1,64
	25/12/2018	256,3	2,71
Sevilla	12/01/2018	0	0
	12/08/2018	13,1	19,63

Fig. 33 depicts the progression of the PV installation areas from 2018 to 2023. It is evident that in Montalto di Castro, despite the existence of installations prior to 2018, there has not been a substantial growth in new PV facilities in recent years. By contrast, Seville experienced a swift and extensive expansion of new PV installations. At the beginning of 2018, no installations were present in the entire study area, whereas by 2023, the total area covered was slightly over 600 ha, nearly twice the size of the area of Montalto di Castro.

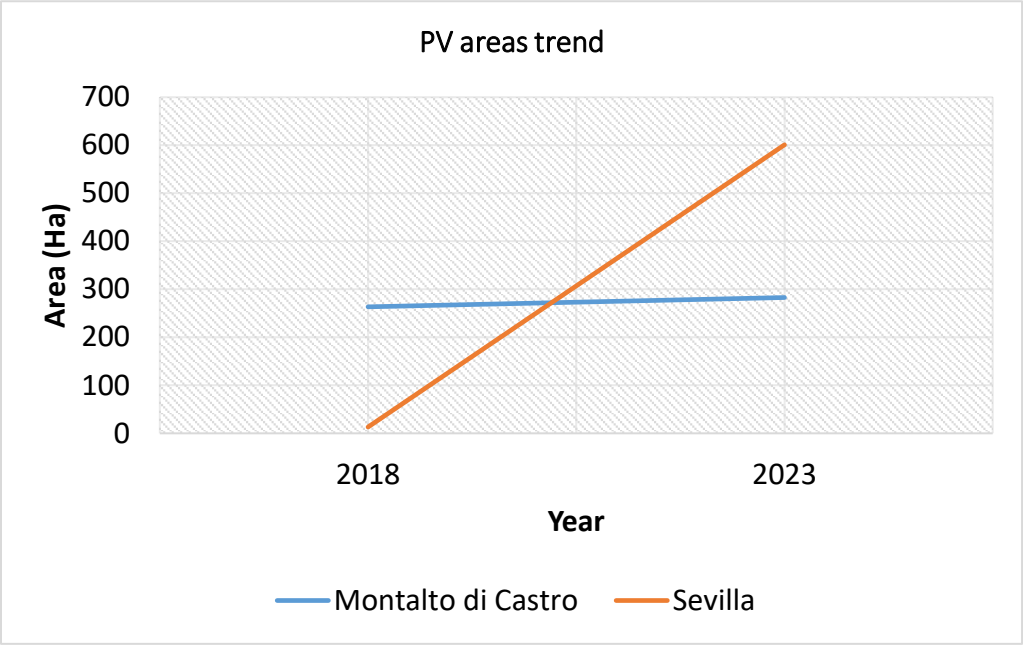


Fig. 33 - PV Areas trend.

3.4.2 Loss area for Land Use Class

The situation in Montalto di Castro and Seville was similar. In both solar parks, the predominant land class before the development of new installations was non-irrigated arable land, which belonged to the Level 1 macro-class of agricultural areas. Therefore, in both cases, the areas designated for PV installations were devoid of specific crops, such as vineyards and olive groves, which are particularly present in the surrounding areas of the study zone. None of the areas where the installations were made were designated as industrial or commercial units; thus, the impact of the installations was directly concentrated in agricultural areas.

4. DISCUSSION OF THE RESULTS

4.1. Spectral analysis and segmentation

The analysis of spectral signatures revealed distinct characteristics of the PV installations in the two study areas. However, the seasonal variability in the spectral signatures necessitated the grouping of months with similar features. The segmentation results, evaluated using the MED2 metric, validated the efficacy of creating image composites based on NDVI values. In Seville, MED2 values were excellent for all compositions, including those encompassing all months. This was primarily attributed to the different geometric characteristics of the installations in the two study areas. Seville's larger and more compact installations are segmented with greater precision. In contrast, segmentation in Montalto di Castro proved to be more challenging owing to its irregular geometry and smaller size. Consequently, optimal segmentation parameters indicate higher scale values in Seville compared to Montalto di Castro. The MRS algorithm demonstrated efficacy for this type of study owing to its ease of use in the eCognition Developer software. The ability to extrapolate segmentation outputs for each scale and shape combination iteratively facilitates an effective segmentation evaluation.

4.2. Features

The features extracted from each segment play a vital role in training the classification algorithm and serve as explanatory variables. In this study, the RFE algorithm was implemented to evaluate the feature space. To examine the effects of various characteristics, including spectral, textural, geometric, and spectral indices, 75 features were extracted and assessed using feature space evaluation methoTab.he findings presented in Table 5 reveal that among the top ten most significant features, EVI and NDTI in the Montalto di Castro study site and NDBI and NDTI in the Seville study site demonstrate considerable importance for the RFE method. Research conducted by Ladisa et al., 2024b, involving a large solar park in India, found that the most relevant features for PV mapping were

NDTI and NDWI, confirming the effectiveness of the NDTI despite differences in installation types, terrain morphology, and latitude. In Montalto di Castro, the strong performance of the EVI index indicates that vegetation beneath PV array installations is a key distinguishing factor. The significance achieved by the NDTI in both study areas highlights the relevance of the relationship between the SWIR1 and NIR2 bands, typical for PV pixels, as demonstrated during the spectral signature analysis. Among the statistics of the individual S2 bands, it was observed that the standard deviation of the Blue, NIR1, and NIR2 bands suggests that spectral variability within PV objects is the most significant among the band statistical values. Texture and geometric features did not perform well in either study area, indicating less relevance compared to other characteristics. These observations demonstrate that while local features can significantly influence mapping accuracy, there are intrinsic characteristics of PV pixels that can be recognized in any study area.

4.3. Classification

The overall results demonstrated that all algorithms achieved an average F1 score of 90% or higher. The RF algorithm exhibited the best overall accuracy, particularly in Montalto di Castro, with F1 scores exceeding 90% for both the image composites. RF's performance confirms its effectiveness in PV mapping, as also highlighted in the studies by Ladisa et al., 2022a and Plakman et al., 2022). The SVM algorithm showed good results, slightly lower than RF but still superior to other studies, with F1 scores ranging from 87.36% to 93.67%. The KNN algorithm, which was used for the first time in PV mapping, obtained the lowest F1 scores among the three algorithms, ranging from 86.05% to 90.24%. Regarding the study area, all algorithms performed better for Montalto di Castro than for Sevilla. However, visual inspection of the results revealed that, despite slightly lower accuracy scores, the installations in the Sevilla area were well classified owing to the improved segmentation quality. Segmentation in Sevilla produced far fewer objects than in Montalto di Castro owing to higher scale values; thus, a

small number of false positives more easily compromised the accuracy values. The use of the RFE algorithm for selecting the best features yielded higher accuracies than (Ladisa et al., 2024b) owing to a more precise segmentation phase and deeper analysis of the seasonal variation in PV pixel values.

4.4. PV CD

The final stage of this research implemented classification models developed in the earlier phases. Notably, the RF classification models applied to the 2018 image data yielded outstanding accuracy, showing remarkable knowledge transfer capability and avoiding overfitting. Regarding the land change analysis results, the two study areas exhibited different new PV installation development patterns from 2018 to 2023. In Montalto di Castro, the PV installation area changed minimally (by 2%) during this period. The 2018 CLC data indicated that most PV installation areas had already been classified, suggesting significant new installation development over a decade ago, followed by a sharp slowdown. Seville presents a contrasting scenario. The CLC data lacked information on PV installations in the study area; the classification results and area calculations revealed ongoing work in 2018. Between 2018 and 2023, substantial new installation developments occurred, with PV installations occupying nearly 600 ha of land. Both study areas shared a common soil type, classified in CLC data as "Non-irrigated arable land." This category encompasses cultivated plots used for non-irrigated agriculture with annual, non-permanent crops, typically in a crop rotation system, including fallow land within this rotation. It also includes fields with occasional sprinkler irrigation using non-permanent devnon-irrigatedort predominantly nonirrigated cultivation.

CONCLUSIONS

The expansion of areas covered by large PV plants continues to grow globally in parallel with the increase in population and energy demand. The growing importance of renewable energy, particularly energy produced by PV plants, and the drastic reduction in the use of fossil fuels, which are responsible for over 70% of carbon emissions, has led the scientific community to focus more on the rapid development of plants and their monitoring. Among the various monitoring methods, such as in situ with drones, this study thoroughly explored the mapping of PV plants using satellite images, making the majority of remote sensing techniques supported by algorithms and the computing power of ML. In particular, this thesis developed a comprehensive fpre-processing starts with the preprocessing phase of satellite images (in this case, open-source S2 images) for the analysis of PV change detection.

The framework was applied to two study areas with distinct characteristics by comparing the differences and similarities in the results. The main feature of this study was the adoption of the OBIA approach rather than the PB approach. The choice to work with an OBIA approach stems from both the good results obtained in the state-of-the-art on this topic and the need to include and test additional features that objects offer compared to normal pixels. The two study areas were selected to test the entire methodology for areas with different characteristics in terms of latitude and land cover type. Moreover, the type of PV plants in the two study areas varied significantly in terms of spatial distribution, size, and type of material. The major novelty of this study is the preliminary phase of analyzing the variation in NDVI within the GT polygons. This type of analysis, in addition to allowing the study of the spectral characteristics of PV pixels and their variations over time, improves the segmentation phase through the creation of image composites that include images from months with similar characteristics. This phase of the methodology was crucial for monitoring the advancement of PV areas between 2018 and 2023, enabling the creation of

classification models capable of operating effectively under all spectral conditions typical of an entire year.

In the image segmentation phase, which is the first phase of the OBIA approach, the freely available command-line tool AssesSeg was used to calculate the MED2 metric, making it easy to access and evaluate the segmentation accuracy. The use of eCognition Developer facilitated the study of segmentation quality owing to the possibility of creating an iterative workflow process that allowed different segmentation outputs for all combinations of scale and shape parameters.

During the feature extraction phase, 75 features were obtained from the segments, including spectral, textural, geometric, and spectral indices. The EVI, NDTI, and NDBI indices were the most relevant in both study areas, according to the RFE method. The use of this feature selection algorithm reduced the dimensionality of the features while improving the accuracy in all case studies.

In the classification model development phase, the RF model achieved the highest F1-score values. In the composites of the summer months (Composite2M and Composite2S), accuracies greater than 94% were achieved in the validation objects and accuracies above 90% for the test areas. These results confirm the excellent performance of the RF in satellite image classification, even in the context of PV mapping. In light of these results, the best-performing RF models in the test phase were used to classify the 2018 images. Despite the lower values, the results of the SVM models were satisfactory, with F1-score values around 89%. KNN instead achieved low values, which proved to be not ideal for PV mapping in this case.

The final phase of this thesis completed the entire development of the classification models by applying them to the 2018 S2 images. This phase of the methodology provides a simple yet effective overview of land consumption trends in both study areas. Although at different speeds and quantities, all areas transformed into PV plants belong to the non-irrigated arable land class, confirming the trend of land use transformation from agricultural to industrial in

recent years. However, the seasonal behavior of the PV pixels addressed in this thesis indicates the presence of vegetation in many PV arrays (especially in the Montalto di Castro area). The presence of vegetation instead of bare soil under the PV arrays indicates a trend towards developing agrivoltaic plants rather than photovoltaic plants, thus maintaining local habitats and reducing land consumption.

Therefore, this methodology represents a potentially valuable predictive tool for planning and decision-making. In general, a large-scale map of photovoltaic parks could offer an opportunity to study land transformation or potential land-use conflicts between land-based economies and existing ecosystem services. This approach could be particularly useful when a rapid assessment is required, such as in the early stages of territorial planning for approximate analyses, or when systematic monitoring and detailed data sources are scarce. This could be the case during rapid photovoltaic expansion or in many developing economies, where unique and reliable data sources are sometimes lacking. In such cases, the absence of a data-sharing culture, rather than a digital divide, could be a more significant factor in undermining the reliability and applicability of this methodology.

Overall, this thesis presents a detailed and innovative PV mapping process owing to its completeness. This study demonstrated excellent applicability. Excluding the segmentation part, the entire process was conducted using open-source data, all within the Python environment. The OBIA technique was of great support for the ML phase in both feature selection and image classification. This thesis lays the foundation for frameworks that are suitable for PV park mapping. The continuous availability of increasingly high-resolution open-source satellite images, progress in computing power of new hardware, and advancement of image classification algorithms will certainly improve this methodology. However, an initial approach to monitoring the advancement of PV areas is fundamental for future monitoring and will undoubtedly be increasingly intense.

ACKNOWLEDGEMENTS

I express my deepest gratitude to my supervisor, Prof. Eufemia Tarantino, for her invaluable support, constant guidance, and wise advice throughout my PhD journey.

Special thanks also go to Prof. Alessandra Capolupo and Prof. Alberico Sonnessa for their encouragement and scientific contribution. I am immensely grateful to all my colleagues at agLab, who shared this journey with passion and dedication.

I would also like to extend my heartfelt thanks to Professors Manuel Angel Aguilar, Fernando Jose Aguilar, and Abderrahim Nemmaoui from the University of Almeria for their support and valuable collaborations. Finally, I wish to thank Prof. Apostolos Papakonstantinou from the Cyprus University of Technology in Limassol for his availability and for the stimulating academic exchanges.

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

Achanta, R. & Susstrunk, S. (2017). Superpixels and Polygons Using Simple Non-iterative Clustering. In: 2017 IEEE Conference on Computer Vi-sion and Pattern Recognition (CVPR). July 2017, IEEE, pp. 4895–4904.

Aguilar, M.A., Aguilar, F.J., García Lorca, A., Guirado, E., Betlej, M., Cichon, P., Nemmaoui, A., Vallario, A. & Paremulti-resolution Assessment of multiresolution segmentation for extracting greenhouses from WorldView-2 imagery. In: International Archives of the Photogrammetry, Remote Sen-sing and Spatial Information Sciences - ISPRS Archives. 2016, International Society for Photogrammetry and Remote Sensing, pp. 145–152.

Aguilar, M.Á., Jiménez-Lao, R., Nemmaoui, A., Aguilar, F.J., Koc-San, D., Tarantino, E. & Chourak, M. (2020). Evaluation of the Consistency of Simultaneously Acquired Sentinel-2 andCovered 8 Imagery on Plastic Co-vered Greenhouses. Remote Sensing. 12 (12). p.p. 2015.

Ahmed, M., Seraj, R. & Islam, S.M.S. (2020). The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. Electronics. 9 (8). p.p. 1295.

Baatz, M., Hoffmann, C. & Willhauck, G. (2008). Progressing from object-based to object-oriented image analysis. In: Object-Based Image Analysis. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 29–42.

Bahl, A., Hellack, B., Balas, M., Dinischiotu, A., Wiemann, M., Brinkmann, J., Luch, A., Renard, B.Y. & Haase, A. (2019). Recursive feature elimination in random forest classification supports nanomaterial grouping. NanoImpact. 15. p.p. 100179.

Belgiu, M. & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogram-metry and Remote Sensing. 114. p.pp. 24–31.

Benz, U.C., Hofmann, P., Willhauck, G., LingenfeMulti-resolutionen, M. (2004). Multiresolution, object-oriented fuzzy analysis of remote sensing data for

GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing. 58 (3–4). p.pp. 239–258.

Cai, L., Shi, W., Miao, Z. & Hao, M. (2018). Accuracy Assessment Measures for Object Extraction from Remote Sensing Images. Remote Sensing. 10 (2). p.p. 303.

Chapman, R. (2008). Producing Inequalities: Regional Sequences in Later Prehistoric Southern Spain. Journal of World Prehistory. 21 (3–4). p.pp. 195–260.

Chatenoud, L., Mosconi, P., Malvezzi, M., Colombo, P., La Vecchia, C. & Apolone, G. (2005). Impact of a major thermoelectric plant on self-perceived health status. Preventive Medicine. 41 (1). p.pp. 328–333.

Chen, X., & Ishwaran, H. (2012). Random forests for genomic data analysis. Genomics. 99 (6). p.pp. 323–329.

Chen, Y., Ming, D., Zhao, L., Lv, B., Zhou, K. & Qing, Y. (2018). Review on High Spatial Resolution Remote Sensing Image Segmentation Evaluation. Photogrammetric Engineering & Remote Sensing. 84 (10). p.pp. 629–646.

Chen, Z., Kang, Y., Sun, Z., Wu, F. & Zhang, Q. (2022). Extraction of Photovoltaic Plants Using Machine Learning Methods: A Case Study of the Pilot Energy City of Golmud, China. Remote Sensing. 14 (11). p.p. 2697.

Chirici, G. & Travaglini, D. (2003). Segmentazione multirisoluzione e classificazione object oriented di immagini telerilevate.

Choi, C.S., Cagle, A.E., Macknick, J., Bloom, D.E., Caplan, J.S. & Ravi, S. (2020). Effects of Revegetation on Soil Physical and Chemical Properties in Solar Photovoltaic Infrastructure. Frontiers in Environmental Science. 8.

Colditz, R. (2015). An Evaluation of Different Training Sample Allo-cation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-Based Algorithms. Remote Sensing. 7 (8). p.p. 9655–9681.

Costa, M.V.C.V. da, Carvalho, O.L.F. de, Orlandi, A.G., Hirata, I., Al-buquerque, A.O. de, Silva, F.V. e, Guimarães, R.F., Gomes, R.A.T. & Júnior, O.A. de C. (2021). Remote Sensing for Monitoring Photovoltaic Solar Plants in Brazil Using Deep Semantic Segmentation. Energies. 14 (10). p.p. 2960.

D'agostino, C. (2014). Investigating Transferability and Goodness of Fit of Two Different Approaches of Segmentation And Model form for estimating safety performance of Motorways. Procedia Engineering. 84. p.pp. 613–623.

Darst, B.F., Malecki, K.C. & Engelman, C.D. (2018). Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC Genetics. 19 (S1). p.p. 65.

Deepak, Srivastava, S., Panda, S. & Malvi, C.S. (2022). Developments in Solar PV Cells, PV Panels, and PVT Systems. In: Solar Thermal Systems: Thermal Analysis and its Application. BENTHAM SCIENCE PUBLI-SHERS, pp. 258–286.

Van Dijk, D., Shoaie, S., van Leeuwen, T. & Veraverbeke, S. (2021). Spectral signature analysis of false positive burned area detection from agricultural harvests using Sentinel-2 data. International Journal of Applied Earth Observation and Geoinformation. 97. p.p. 102296.

Dunnett, S., Sorichetta, A., Taylor, G. & Eigenbrod, F. (2020). Harmonised global datasets of wind and solar farm locations and power. Scientific Data. 7 (1). p.p. 130.

ESRI (2009). arcgis.com. 2009. World Imagery.

Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment. 58 (3). p.pp. 257–266.

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Lan-guille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pes-siot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P. & Fer-nandez, V. (2017). Copernicus Sentinel-2A Calibration and Products Validation Status. Remote Sensing. 9 (6). p.p. 584.

Gislason, P.O., Benediktsson, J.A. & Sveinsson, J.R. (2006). Random Forests for land cover classification. Pattern Recognition Letters. 27 (4). p.pp. 294–300.

Global Energy Monitor Wiki (2024). Global Energy Monitor Wiki. 2024. Don Rodrigo solar farm.

Graf, L., Bach, H. & Tiede, D. (2020). Semantic Segmentation of Sentinel-2 Imagery for Mapping Irrigation Center Pivots. Remote Sensing. 12 (23). p.p. 3937.

Granitto, P.M., Furlanello, C., Biasioli, F. & Gasperi, F. (2006). Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. Chemometrics and Intelligent Laboratory Systems. 83 (2). p.pp. 83–90.

Güney, T. (2022). Solar energy, governance and CO2 emissions. Renewable Energy. 184. p.pp. 791–798. Available from: [Accessed: 25 July 2024].

Haifeng Wang & Dejin Hu (2005). Comparison of SVM and LS-SVM for Regression. In: 2005 International Conference on Neural Networks and Brain. 2005, IEEE, pp. 279–283.

He, C., Shi, P., Xie, D. & Zhao, Y. (2010). Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. Remote Sensing Letters. 1 (4). p.pp. 213–221.

Jiang, H., Yao, L., Lu, N., Qin, J., Liu, T., Liu, Y. & Zhou, C. (2021). Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery. Earth System Science Data. 13 (11). p.pp. 5389–5401.

JIANG, Z., HUETE, A., DIDAN, K. & MIURA, T. (2008). Development of a two-band enhanced vegetation index without a blue band. Remote Sensing of Environment. 112 (10). p.pp. 3833–3845.

Jiménez-Lao, R., Aguilar, M.Á., Ladisa, C., AguilarMulti-resolutionui, A. (2022). Multiresolution Segmentation for extracting plastic greenhouses from Deimos-2 Imagery. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2022.

Jozdani, S. & Chen, D. (2020). On the versatility of popular and frequently proposed supervised evaluation metrics for segmentation quality of remotely sensed images: An experimental case study of building extraction. ISPRS Journal of Photogrammetry and Remote Sensing. 160. p.pp. 275–290.

Jutz, S. & Milagro-Pérez, M.P. (2020). Copernicus: the European Earth Observation programme. Revista de Teledetección. (56).

Kaplan, G. & Avdan, U. (2018). Sentinel-1 and Sentinel-2 data fusion for wetlands mapping: Balikdami, Turkey. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLII–3. p.pp. 729–734.

Kasmi, G., Saint-Drenan, Y.-M., Trebosc, D., Jolivet, R., Leloux, J., Sarr, B. & Dubus, L. (2023). A crowdsourced dataset of aerial images with annotated solar photovoltaic arrays and installation metadata. Scientific Data. 10 (1). p.p. 59.

Kassouk, Z., Zayani, H., Mougenot, B., Mabrouki, R. & Lili-Chabaane, Z. (2020). Detection of tillage areas and periods using high spatial resolution optical image time series. In: 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS). March 2020, IEEE, pp. 334–337.

Kruzel, R. & Helbrych, P. (2018). Analysis of the profitability of a photovoltaic installation in the context of sustainable development of construction. E3S Web of Conferences. 49. p.p. 00061.

Kucharczyk, M., Hay, G.J., Ghaffarian, S. & Hugenholtz, C.H. (2020). Geographic Object-Based Image Analysis: A Primer and Future Directions. Remote Sensing. 12 (12). p.p. 2012.

Kumar, A. & Thakur, A.D. (2021). Impurity photovoltaic and split spectrum for efficiency gain in Cu2ZnSnS4 solar cells. Optik. 238. p.p. 166783.

Kumar, N.M., Sudhakar, K., Samykano, M. & Jayaseelan, V. (2018). On the technologies empowering drones for intelligent monitoring of solar photovoltaic power plants. Procedia Computer Science. 133. p.pp. 585–593. Available from: [Accessed: 27 July 2024].

Labib, S.M. & Harris, A. (2018). The potentials of Sentinel-2 and LandSat-8 data in green infrastructure extraction, using object based ima-ge analysis (OBIA) method. European Journal of Remote Sensing. 51 (1). p.pp. 231–240.

Ladisa, C., Capolupo, A., Ripa, M.N. & Tarantino, E. (2022a). Combining OBIA approach and Machine Learning algorithm to extract photovoltaic panels

from Sentinel-2 images automatically. In: Proceedings of SPIE - The International Society for Optical Engineering. 2022.

Ladisa, C., Capolupo, A., Ripa, M.N. & Tarantino, E. (2022b). Evaluation of eCognition Developer and Orfeo ToolBox Performances for Segmenting Agrophotovoltaic Systems from Sentinel-2 Images.

Ladisa, C., Capolupo, A. & Tarantino, E. (2024a). A Customized JAVA OpenStreetMap Preset to Extract Solar Panel Installations for Humanitarian Purposes.

Ladisa, C., Capolupo, A. & Tarantino, E. (2024b). Optimizing Feature Selection for Solar Park Classification: Approaches with OBIA and Machine Learning. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2024, Springer Science and Business Media Deutschland GmbH, pp. 286–301.

Liu, Y., Bian, L., Meng, Y., Wang, H., Zhang, S., Yang, Y., Shao, X. & Wang, B. (2012). Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. ISPRS Journal of Photogrammetry and Remote Sensing. 68. p.pp. 144–156.

Lyons, M.B., Keith, D.A., Phinn, S.R., Mason, T.J. & Elith, J. (2018). A comparison of resampling methods for remote sensing classification and accuracy assessment. Remote Sensing of Environment. 208. p.pp. 145–153.

Ma, L., Li, M., Ma, X., Cheng, L., Du, P. & Liu, Y. (2017). A review of supervised object-based land-cover image classification. ISPRS Journal of Photogrammetry and Remote Sensing. 130. p.pp. 277–293.

Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Homayouni, S. & Gill, E. (2018). The First Wetland Inventory Map of Newfoundland at a Spatial Resolution of 10 m Using Sentinel-1 and Sentinel-2 Data on the Google Earth Engine Cloud Computing Platform. Remote Sensing. 11 (1). p.p. 43.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U. & Gascon, F. (2017). Sen2Cor for Sentinel-2. In: L. Bruzzone, F. Bovolo, & J. A. Benediktsson (eds.). Image and Signal Processing for Remote Sensing XXIII. 4 October 2017, SPIE, p. 3.

Manrai, L.A., Manrai, A.K. & DeLuca, J. (2017). Twenty Shades of Ita-ly: An Analysis of its Cultural, Natural, and Dual Tourist Attractions with Implications for Global Tourism Marketing. Journal of Global Marketing. 30 (5). p.pp. 297–308.

Marutho, D., Hendra Handaka, S., Wijaya, E. & Muljono (2018). The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News. In: 2018 International Seminar on Application for Technology of Information and Communication. September 2018, IEEE, pp. 533–538.

Morales-Barquero, L., Lyons, M., Phinn, S. & Roelfsema, C. (2019). Trends in Remote Sensing Accuracy Assessment Approaches in the Context of Natural Resources. Remote Sensing. 11 (19). p.p. 2305.

Moroni, M., Porti, M. & Piro, P. (2019). Design of a Remote-Controlled Platform for Green Roof Plants Monitoring via Hyperspectral Sensors. Water. 11 (7). p.p. 1368.

Mustafa, R.J., Gomaa, M.R., Al-Dhaifallah, M. & Rezk, H. (2020). Environmental Impacts on the Performance of Solar Photovoltaic Systems. Sustainability. 12 (2). p.p. 608.

Nainggolan, R., Peranginangin, R., Simarmata, E. & Tarigan, A.F. (2019). Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method. Journal of Physics: Conference Series. 1361 (1). p.p. 012015.

NOVASOL (2017). Proyecto de planta solar fotovoltaica “la isla” 182,5 mwp, subesta-ción “la isla” (30/220kv) y su línea eléctrica de evacuación (220 mva) (alcalá de guadaira, sevilla).

Novelli, A., Aguilar, M.A., Aguilar, F.J., Nemmaoui, A. & Tarantino, E. (2017). AssesSeg A command line tool to quantify image segmentation quality: A test carried out in southern spain from satellite imagery. Remote Sensing. 9 (1).

Ordóñez, J., Jadraque, E., Alegre, J. & Martínez, G. (2010). Analysis of the photovoltaic solar energy capacity of residential rooftops in Andalusia (Spain). Renewable and Sustainable Energy Reviews. 14 (7). p.pp. 2122–2130.

Owusu, P.A. & Asumadu-Sarkodie, S. (2016). A review of renewable energy sources, sustainability issues and climate change mitigation. Co-gent Engineering. 3 (1). p.p. 1167990.

Parida, B., Iniyan, S. & Goic, R. (2011). A review of solar photovoltaic technologies. Renewable and Sustainable Energy Reviews. 15 (3). p.pp. 1625–1636. Available from: [Accessed: 25 July 2024].

Patle, A. & Chouhan, D.S. (2013). SVM kernel functions for classification. In: 2013 International Conference on Advances in Technology and Engineering (ICATE). January 2013, IEEE, pp. 1–9.

Peters, I.M., Liu, H., Reindl, T. & Buonassisi, T. (2018). Global Prediction of Photovoltaic Field Performance Differences Using Open-Source Satellite Data. Joule. 2 (2). p.pp. 307–322.

Pettorelli, N., Ryan, S., Mueller, T., Bunnefeld, N., Jedrzejewska, B., Lima, M. & Kausrud, K. (2011). The Normalized Difference Vegetation In-dex (NDVI): unforeseen successes in animal ecology. Climate Research. 46 (1). p.pp. 15–27.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J. & Stenseth, N.Chr. (2005). Using the satellite-derived NDVI to assess ecologi-cal responses to environmental change. Trends in Ecology & Evolution. 20 (9). p.pp. 503–510.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V., Murayama, Y. & Ranagalage, M. (2020). Sentinel-2 Data Reviewd Cover/Use Mapping: A Re-view. Remote Sensing. 12 (14). p.p. 2291.

Pietrarota, A. (2023). A Montalto di Castro il primo impianto fotovoltaico di Iberdrola in Italia. 2023.

Plakman, V., Rosier, J. & van Vliet, J. (2022). Solar park detection from publicly available satellite imagery. GIScience & Remote Sensing. 59 (1). p.pp. 462–481.

Rinonapoli C. (2023). Montalto di Castro – Mega impianto fotovoltaico entro il 2024. 2023. Lavori affidati al colosso industria-le Sunner Group. Sono 68 ettari di terreno, circa 95 campi da calcio.

Ridao, Á.R., García, E.H., Escobar, B.M. & Toro, M.Z. (2007). Solar energy in Andalusia (Spain): present state and prospects for the future. Renewable and Sustainable Energy Reviews. 11 (1). p.pp. 148–161.

Sammouda, R. & El-Zaart, A. (2021). An Optimized Approach for Prostate Image Segmentation Using K-Means Clustering Algorithm with El-bow Method. Computational Intelligence and Neuroscience. 2021 (1).

Saputra, D.M., Saputra, D. & Oswari, L.D. (2020). Effect of Distance Metrics in Determining K-Value in K-Means Clustering Using Elbow and Silhouette Method. In: Proceedings of the Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019). 2020, Paris, France: Atlantis Press.

Schirru, M.R. (2019). Montalto di Castro - Sustainable Tourism as an Opportunity for Urban and Environmental Regeneration. In: pp. 645–653.

Senel, G., Aguilar, M.A., Aguilar, F.J., Nemmaoui, A. & Goksel, C. (2023). Unraveling Segmentation Quality of Remotely Sensed Images on Plastic-Covered Greenhouses: A Rigorous Experimental Analysis from Supervised Evaluation Metrics. Remote Sensing. 15 (2). p.p. 494.

Singh, K.V., Setia, R., Sahoo, S., Prasad, A. & Pateriya, B. (2015). Eva-luation of NDWI and MNDWI for assessment of waterlogging by integrating digital elevation model and groundwater level. Geocarto Internatio-nal. 30 (6). p.pp. 650–661.

Smith, K.E.L., Terrano, J.F., Pitchford, J.L. & Archer, M.J. (2021). Coastal Wetland Shoreline Change Monitoring: A Comparison of Shorelines from High-Resolution WorldView Satellite Imagery, Aerial Imagery, and Field Surveys. Remote Sensing. 13 (15). p.p. 3030.

Spoto, F., Sy, O., Laberinti, P., Martimort, P., Fernandez, V., Colin, O., Hoersch, B. & Meygret, A. (2012). Overview Of Sentinel-2. In: 2012 IEEE International Geoscience and Remote Sensing Symposium. July 2012, IEEE, pp. 1707–1710.

Syakur, M.A., Khotimah, B.K., Rochman, E.M.S. & Satoto, B.D. (2018). Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster. IOP Conference Series: Materials Science and Engineering. 336. p.p. 012017.

Szantoi, Z. & Strobl, P. (2019). Copernicus Sentinel-2 Calibration and Validation. European Journal of Remote Sensing. 52 (1). p.pp. 253–255.

T. Nguyen, H. & M. Pearce, J. (2013). Automated quantification of solar photovoltaic potential in cities. International Review for Spatial Planning and Sustainable Development. 1 (1). p.pp. 49–60.

Tawalbeh, M., Al-Othman, A., Kafiah, F., Abdelsalam, E., Almomani, F. & Alkasrawi, M. (2021). Environmental impacts of solar photovoltaic systems: A critical review of recent progress and future outlook. Science of The Total Environment. 759. p.p. 143528. Available from: [Accessed: 25 July 2024].

Thepaut, J.-N., Dee, D., Engelen, R. & Pinty, B. (2018). The Copernicus Programme and its Climate Change Service. In: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. July 2018, IEEE, pp. 1591–1593.

Tian, J. & Chen, D. -M. (2007). Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition. International Journal of Remote Sensing. 28 (20). p.pp. 4625–4644.

Umargono, E., Suseno, J.E. & Vincensius Gunawan, S.K. (2020). K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula. In: Proceedings of the 2nd International Seminar on Science and Technology (ISSTEC 2019). 2020, Paris, France: Atlantis Press.

Venter, Z.S., Barton, D.N., Chakraborty, T., Simensen, T. & Singh, G. (2022). Global 10 m Land Use Land Cover Datasets: A Comparison of Dynamic World, World Cover and Esri Land Cover. Remote Sensing. 14 (16). p.p. 4101.

Wang, B., Liu, Q., Wang, L., Chen, Y. & Wang, J. (2023a). A review of the port carbon emission sources and related emission reduction technical measures. Environmental Pollution. 320. p.p. 121000. Wang, J., Liu, J. & Li, L. (2022).

Detecting Photovoltaic Installations in Multi-sourcescapes Using Open Multi-Source Remote Sensing Data. Remote Sensing. 14 (24). p.p. 6296.

Wang, Q., Shi, W., Li, Z. & Atkinson, P.M. (2016). Fusion of Sentinel-2 images. Remote Sensing of Environment. 187. p.pp. 241–252.

Wang, S., Peng, J., Wang, M., Xue, P., Luo, Y., Ma, T. & Zhao, Y. (2023b). Evaluation of the energy conversion performance of different photovoltaic materials with measured solar spectral irradiance. Renewable Energy. 219. p.p. 119431.

Watkins, B. & van Niekerk, A. (2019). A comparison of object-based image analysis approaches for field boundary delineation using multi-temporal Sentinel-2 imagery. Computers and Electronics in Agriculture. 158. p.pp. 294–302.

Witharana, C. & Civco, D.L. (2014). Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure (ED2). ISPRS Journal of Photogrammetry and Remote Sensing. 87. p.pp. 108–121.

Wolniak, R. & Skotnicka-Zasadzień, B. (2022). Development of Photovoltaic Energy in EU Countries as an Alternative to Fossil Fuels. Energies. 15 (2). p.p. 662.

Xia, Z., Li, Y., Chen, R., Sengupta, D., Guo, X., Xiong, B. & Niu, Y. (2022). Mapping the rapid development of photovoltaic power stations in northwestern China using remote sensing. Energy Reports. 8. p.pp. 4117–4127.

Xia, Z., Li, Y., Guo, S., Chen, R., Zhang, W., Zhang, P. & Du, P. (2023). Mapping global water-surface photovoltaics with satellite images. Renewable and Sustainable Energy Reviews. 187. p.p. 113760. Yan, K. & Zhang, D. (2015). Feature selection and analysis on correlated gas sensor data with recursive feature elimination. Sensors and Ac-tuators B: Chemical. 212. p.pp. 353–363.

Yang, J., He, Y., Caspersen, J. & Jones, T. (2015). A discrepancy measure for segmentation evaluation from the perspective of object reco-gnition. ISPRS Journal of Photogrammetry and Remote Sensing. 101. p.pp. 186–192.

Yang, X., Qin, Q., Yésou, H., Ledauphin, T., Koehl, M., Grussen-meyer, P. & Zhu, Z. (2020). Monthly estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data. Remote Sensing of Environment. 244. p.p. 111803.

Yue, S., Li, P. & Hao, P. (2003). SVM classification:Its contents and challenges. Applied Mathematics-A Journal of Chinese Universities. 18 (3). p.pp. 332–342.

Zha, Y., Gao, J. & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. Internatio-nal Journal of Remote Sensing. 24 (3). p.pp. 583–594.

Zhang, H., Tian, P., Zhong, J., Liu, Y. & Li, J. (2023). Mapping Photo-voltaic Panels in Coastal China Using Sentinel-1 and Sentinel-2 Images and Google Earth Engine. Remote Sensing. 15 (15). p.p. 3712.

Zhang, M.-L. & Zhou, Z.-H. (2007). ML-KNN: A lazy learning ap-proach to multi-label learning. Pattern Recognition. 40 (7). p.pp. 2038–2048.

Zhang, P., Zhang, L., Wu, T., Zhang, H. & Sun, X. (2017a). Detection and location of fouling on photovoltaic panels using a drone-mounted in-frared thermography system. Journal of Applied Remote Sensing. 11 (1). p.p. 016026.

Zhang, S., Li, X., Zong, M., Zhu, X. & Cheng, D. (2017b). Learning k for kNN Classification. ACM Transactions on Intelligent Systems and Techno-logy. 8 (3). p.pp. 1–19.

Zhang, S., Li, X., Zong, M., Zhu, X. & Wang, R. (2018). Efficient kNN Classification With Different Numbers of Nearest Neighbors. IEEE Transac-tions on Neural Networks and Learning Systems. 29 (5). p.pp. 1774–1785.

Zhang, X., Feng, X., Xiao, P., He, G. & Zhu, L. (2015). Segmentation quality evaluation using region-based precision and recall measures for remote sensing images. ISPRS Journal of Photogrammetry and Remote Sensing. 102. p.pp. 73–84.

Zhang, X. & Liu, C.-A. (2023). Model averaging prediction by <math altimg="si681.svg" display="inline" id="d1e2762"> <mi>K</mi> </math> -fold cross-validation. Journal of Econometrics. 235 (1). p.pp. 280–301.

Zhang, X., Xu, M., Wang, S., Huang, Y. & Xie, Z. (2022). Mapping photovoltaic power plants in China using Landsat, random forest, and Google Earth Engine. Earth System Science Data. 14 (8). p.pp. 3743–3755.Platone, 360 a.C. *La Repubblica*. edizione a cura di M. Vegetti. Milano: BUR Biblioteca Universale Rizzoli, 2007, p. 1121. ISBN: 978-88-17-01337-6 1183p

CURRICULUM

COGNOME: Ladisa
NOME: Claudio
LUOGO E DATA DI NASCITA:
Manfredonia (FG): 01/09/1995
Mail: claudioladisa@poliba.it cel: +39 3889293738

ESPERIENZE LAVORATIVE:

- Dal 26/03/2024 al 25/06/2024: COASTLINE Secondment. Limassol, Cipro. Aerial and satellite investigations: Modellazione della vulnerabilità e del rischio di erosione costiera nell'area di studio di Paphos (Cipro) mediante l'uso efficace dei dati Copernicus relativi alla variazione della copertura del suolo e agli eventi sismici e climatici.
- Dal 23/05/2021 al 31/10/2021: Permitting Developer Fotovoltaic. Bari. Sviluppo e progettazione di grandi impianti fotovoltaici in aree agricole ed industriali al fine di ottenere le autorizzazioni per la realizzazione degli stessi.

FORMAZIONE

- Dottorato di ricerca in RISCHIO, SVILUPPO AMBIENTALE, TERRITORIALE ED EDILIZIO (Risk and environmental, territorial and building development Ph.D.) IN CORSO.
- Abilitazione all'esercizio della professione di INGEGNERE CIVILE E AMBIENTALE presso il Politecnico di Bari in data 21/12/2021 con la votazione di 77/90.
- Laurea magistrale in Ingegneria per L'Ambiente e il Territorio conseguita presso il Politecnico di Bari in data 23/07/2021 con punti 110/110.

- Laurea triennale in Ingegneria Civile e Ambientale conseguita presso il Politecnico di Bari in data 13/11/2018 con punti 94/110.

MOBILITA' INTERNAZIONALI:

- Dal 01/02/2023 al 02/06/2023: Periodo di Dottorato all'estero presso il "Departamento de Ingeniería. Escuela Superior de Ingeniería"dell'Università di Almeria.
- Dal 21/01/2021 al 21/03/2021 tesi all'estero presso l'Università di Almeria dal titolo "An OBIA approach for mapping plastic greenhouses in the Bari area using Deimos-2 and Sentinel-2 satellite images". Relatori: Eufemia Tarantino e Manuel Angel Aguilar Torres.
- Dal 24/01/2020 al 24/07/2020: Mobilità Erasmus+ presso l'Università di Cantabria.
- Dal 10/12/2019 al 19/12/19: Step Forward - EU for YOUTH organised by Copenaghen Youth Network.

PUBBLICAZIONI SCIENTIFICHE:

- Ladisa, C., Capolupo, A., & Tarantino, E. (2024, July). Optimizing Feature Selection for Solar Park Classification: Approaches with OBIA and Machine Learning. In International Conference on Computational Science and Its Applications (pp. 286-301). Cham: Springer Nature Switzerland.
- Ladisa, C., Capolupo, A., & Tarantino, E. (2023, September). A Customized JAVA OpenStreetMap Preset to Extract Solar Panel Installations for Humanitarian Purposes. In International Conference on Innovation in Urban and Regional Planning (pp. 3-11). Cham: Springer Nature Switzerland.
- Ladisa, C., Capolupo, A., Ripa, M. N., & Tarantino, E. (2022, October). Combining OBIA approach and machine learning algorithm to extract photovoltaic panels from Sentinel-2 images

automatically. In *Remote Sensing for Agriculture, Ecosystems, and Hydrology XXIV* (Vol. 12262, pp. 67-76). SPIE.

- Ladisa, C., Capolupo, A., Ripa, M. N., & Tarantino, E. (2022, July). Evaluation of ecognition developer and orfeo toolbox performances for segmenting agrophotovoltaic systems from sentinel-2 images. In *International Conference on Computational Science and Its Applications* (pp. 466-482). Cham: Springer International Publishing.
- Jiménez-Lao, R., Aguilar, M. A., Ladisa, C., Aguilar, Multi-resolutionui, A. (2022). Multiresolution Segmentation for extracting plastic greenhouses from Deimos-2 Imagery. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2, 251-258.
- Aguilar, M. A., Jiménez-Lao, R., Ladisa, C., Aguilar, F. J., & Tarantino, E. (2022). Comparison of spectral indices extracted from Sentinel-2 images to map plastic covered greenhouses through an object-based approach. GIScience & Remote Sensing, 59(1), 822-842.
- Aloi, A., Alonso, B., Benavente, J., Cordera, R., Echániz, E., González, F., ... & Sañudo, R. (2020). Effects of the COVID-19 lockdown on urban mobility: Empirical evidence from the city of Santander (Spain). *Sustainability*, *12*(9), 3870.