Smart Agricultural Technology 13 (2026) 101911

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

A systematic assessment of remote sensing approaches for agricultural zonation and supporting precision agriculture

Gunay Hasanli [a,b,c,*], Sadra Emamalizadeh [a], Riccardo Mazzoleni [a], Marco Benfenati [a], Gabriele Baroni [a]

[a] Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy
[b] Department of Agricultural and Food Sciences, ADA University, Baku, Azerbaijan
[c] Agro Research Center, The Ministry of Agriculture of Azerbaijan, Baku, Azerbaijan

ARTICLE INFO

Keywords:
Remote sensing
Agricultural zonation
Clustering algorithms
Precision agriculture
Vegetation Indices
Sentinel-2

ABSTRACT

Remote sensing (RS) technologies generate large amounts of data that can be highly valuable for supporting field zonation and precision agriculture. However, numerous processing approaches have been proposed, with no clear guidelines for their application. This study aims to assess the effects of (i) the vegetation index (VI), (ii) the clustering algorithm, and (iii) the timing of image acquisition on the quality and consistency of zonation results. Sentinel-2 images were analyzed over two growing seasons at two experimental sites in Northern Italy. Several ground observations were integrated to define the benchmark. The results indicated that the choice of VI and clustering method had only a minor effect on zonation, whereas image selection had a significant influence. The best results were obtained during the crop growing season and under dry conditions. Overall, this study demonstrates that systematic analyses can provide practical guidance for implementing RS technologies to support precision management practices.

1. Introduction

Agricultural management zonation is relevant for supporting land management and precision agriculture [1]. Delineating homogeneous zones within the field, farmers and technicians can tailor inputs such as water, fertilizers, and pesticides to specific field conditions, thereby optimizing resource use and enhancing agricultural sustainability [2].

In this context, remote sensing (RS) techniques have been considered a promising source of information for land management decisions for several decades [3,4]. The first applications used imagery from NASA's Landsat program, which began with the launch of Landsat 1 in 1972. These early missions provided moderate spatial (80 m) and temporal (16 days) resolution, enabling unprecedented access to large-scale land surface data [5]. With continuous advancements in satellite technology, RS has gained widespread popularity, particularly with the introduction of higher-resolution systems such as ESA's Sentinel-2 and PlanetScope constellation. Sentinel-2, launched in 2015, offers a spatial resolution of up to 10 m across several spectral bands and revisits each location on average every five days, making it highly suitable for vegetation monitoring and agricultural applications [6]. PlanetScope, on the other hand, consists of a dense constellation of CubeSats, each approximately 10 × 10 × 30 cm, that deliver multispectral imagery at 3–4 m spatial resolution with near-daily global coverage [7]. These technological developments have enabled more detailed, timely, and scalable monitoring of agricultural systems.

Despite these technological advances, there remains a notable lack of standardized methodologies for processing RS for agricultural zonation [8,9]. Current approaches vary considerably in their selection of vegetation indices (VI), with many studies defaulting to the Normalized Difference Vegetation Index (NDVI) [10] despite its known limitations in dense vegetation canopies where saturation occurs [11]. Alternative indices have been proposed and applied in different conditions [4]. As examples, the Soil-Adjusted Vegetation Index (SAVI) and the Modified Soil-Adjusted Vegetation Index (MSAVI) [12] have been proposed to reduce the influence of soil background in areas with sparse vegetation cover [3]. The Enhanced Vegetation Index (EVI) has been used in regions with high biomass to minimize canopy background variations and atmospheric influences [13], while the Normalized Difference Water Index (NDWI) has been applied in semi-arid regions to monitor crop water stress and improve irrigation scheduling [14]. But comparative

* Corresponding author.
E-mail address: gunay.hasanli2@unibo.it (G. Hasanli).

https://doi.org/10.1016/j.atech.2026.101911
Received 22 November 2025; Received in revised form 5 February 2026; Accepted 22 February 2026
Available online 24 February 2026

studies across different agro-environmental conditions are limited.

Clustering algorithms have also been identified to play a critical role in translating RS data into meaningful zonation. K-means clustering has been shown to be useful for simple classification of agricultural management zones and it has been widely applied [3,15–17]. But several other algorithms have been proposed [16]. The use of self-organized map (SOM) cluster approach has demonstrated superiority in capturing complex, non-linear interactions in crop productivity and soil variability [9,18,19]. Gaussian Mixture Models (GMMs) have effectively identified uncertainty or overlapping clusters in agricultural landscapes [20,21] and excel at delineating management zones with variable soil properties and crop response patterns [22]. However, also for this processing step, comparative assessments of these methods under varying RS data configurations remain limited.

Finally, temporal image acquisition adds further complexity. Some studies rely on single-date imagery [23], while others employ multi-temporal or time-series approaches using techniques like Principal Component Analysis (PCA) [24] or clustering of temporal profiles [25]. The number of images employed across studies shows considerable variation, ranging from as few as six [26] to fifteen or more [27,28], with limited explanation provided for these choices. Establishing clearer guidelines or justifications for image selection could enhance the consistency and comparability of future research.

In the light of the limitations highlighted above, this study aims to systematically assess how various combinations of vegetation indices, clustering algorithms, and acquisition timing affect agricultural zonation outcomes. The methodology involves a comparative analysis of three vegetation indices (NDVI, SAVI, and MSAVI) derived from high-resolution satellite imagery acquired over two years at different growth stages. We specifically investigate the relative performance of commonly used clustering methods (K-means, SOM, and GMM) across different agricultural fields to identify their strengths and limitations in capturing spatial heterogeneity relevant for agricultural management. Our hypothesis is that while RS provides valuable spatial insights for land-use planning, methodological decisions such as VI selection, clustering techniques and temporal configuration substantially impact the reliability of zonation. The results are used to formulate practical guidelines in using RS for management zonation [29].

2. Materials and methods

2.1. Study area and data collection

Two experimental fields in the Northern part of Italy, Bondeno and Imola, were chosen for this study. Bondeno site (latitude 44°51′13.45″ N, longitude 11°22′40.32″ E) is a walnut orchard (*Juglans regia* L. cv Chandler) managed by Società Agricola Agro Noce S.r.l. The site spans 160 hectares, but a 56-hectare area was investigated in the present study. The orchard was planted in 2019, with an intra-row spacing of 5 m and inter-row spacing of 7 m It is divided into 14 irrigated sectors, each controlled by independent valves and equipped with a micro-sprinkler irrigation (MSI) system. Sprinkler nozzles are suspended 1 m above the soil surface and mounted on polyethylene (PE) tubes spaced 5 m apart. The PE tubes have a flow capacity of 80 m^3/h, and each nozzle delivers water at a rate of 0.078 m^3/h. Imola site (latitude 44°23′5.15″ N, longitude 11°41′50.30″ E) is a 15-hectare vineyard planted with *Vitis vinifera* L. cv Trebbiano Romagnolo and cv Pignoletto. It is managed by Società Agricola CACI in Imola, Bologna (Italy). The vineyard is uniformly irrigated using a sub-surface drip irrigation (SDI) system, installed at a depth of 25 cm along two lines parallel to each row. The SDI system delivers water at a rate of 0.008 m^3/h per emitter. Vine spacing is 1 m within rows and 2 m between rows.

Several ground information and data collected during dedicated field activities have been analyzed for defining the reference zonation that is used for the assessment of the RS approaches. First of all, the regional soil map at the scale of 1:50,000 has been considered (https://ambiente.regione.emilia-romagna.it/geologia/soil/soil-knowledge/cartography). This map primarily describes the main characteristics of the different types of soils. Accordingly, the Bondeno site has a strong soil variability composed of silty, clay, and silty-clay zones, offering moderate to high fertility but indicating a range of variations in the retention of water and nutrients. The Imola site, on the other hand, shows more homogeneity, with a single silty-clay soil type and, thus, high water retention (Fig. 1).

The spatial variability identified by this regional soil map was further assessed through additional ground-based soil texture analysis. Soil samples were taken following the scheme in Fig. 1 at 0–50 cm depth using a manual auger. Particle size distribution (PSD) was determined with a PARIO™ system (METER Group, Pullman, WA, USA), which automatically implements Stokes' Law within the 63–2 μm range to generate continuous PSD curves.

Moreover, soil water content (SWC) field surveys have been conducted over a dense grid as depicted in Fig. 1. The measurements have been collected with the portable time domain reflectometry (TDR) sensor from Spectrum Technologies, Inc. (Aurora, IL, USA). The sensor-enabled mobile, point-scale, repetitive acquisitions of SWC (%), with a vertical spatial resolution of 5 cm depth. The sampling grid is based on positions every 50 m. At each position, 5 observations have been selected for a total of around 500 points for both Bondeno and Imola. While these observations capture only the condition of surface soil layers (i.e., 5 cm soil depth), they have been considered informative for agricultural zonation [30]. The campaigns have been conducted at each location on the 1st and 16th of June 2023 for Bondeno and on the 8th and the 31st of May in 2023 for Imola. Data have been interpolated by means of the Inverse Distance Weighting method [31] to better visualize the spatial variability detected by the measurements. The use of more advanced interpolation techniques is considered beyond the scope of the present study, and it has not been performed.

2.2. Remote sensing data and vegetation indices

High-resolution multispectral imagery, with a temporal frequency of approximately five days at the equator and a spatial resolution of ten meters, provided by the Sentinel-2 mission, was used. Sentinel-2 satellite data are freely available under an open-access policy, enabling large-scale and long-term monitoring. The mission offers suitable spatial and spectral resolution, including key bands essential for vegetation studies [6]. While other missions such as PlanetScope provide higher spatial resolution, their commercial nature and restricted data access limit their suitability for large-scale or long-term research applications [32]. Images for the year 2023 and 2024 were selected to capture seasonal dynamics. Only images with $<$ 20 % cloud cover [33] were retained to minimize the impact of cloud interference on image quality. After filtering, a total of 30 images for Bondeno and 35 images for the Imola site were analyzed.

Three VIs, (NDVI, SAVI, and MSAVI) were computed and compared (Eqs. (1), 2, and 3). The selection of these indices was based on their proven effectiveness in vegetation studies, their complementary strengths, and widespread validation in RS literature [1,10,34] .

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} \tag{1}$$

In Eq. (1), RED (Band 4, ∼ 665 nm) stands for red reflectance, and NIR (Band 8, ∼842 nm) represents near-infrared surface reflectance values. NDVI, which ranges from –1 to 1, estimates vegetation health, with higher values indicating denser and healthier canopies [8].

$$SAVI = \frac{(NIR - RED)(1 + L)}{NIR + RED + L} \tag{2}$$

SAVI incorporates a soil-adjusted factor ($L = 0.5$) [35] to minimize soil background effects in sparsely vegetated areas [36]. Orchard and vineyard surfaces are often heterogeneous, exhibiting a mosaic of dense canopy rows and variable inter-row vegetation or soil exposure. In such

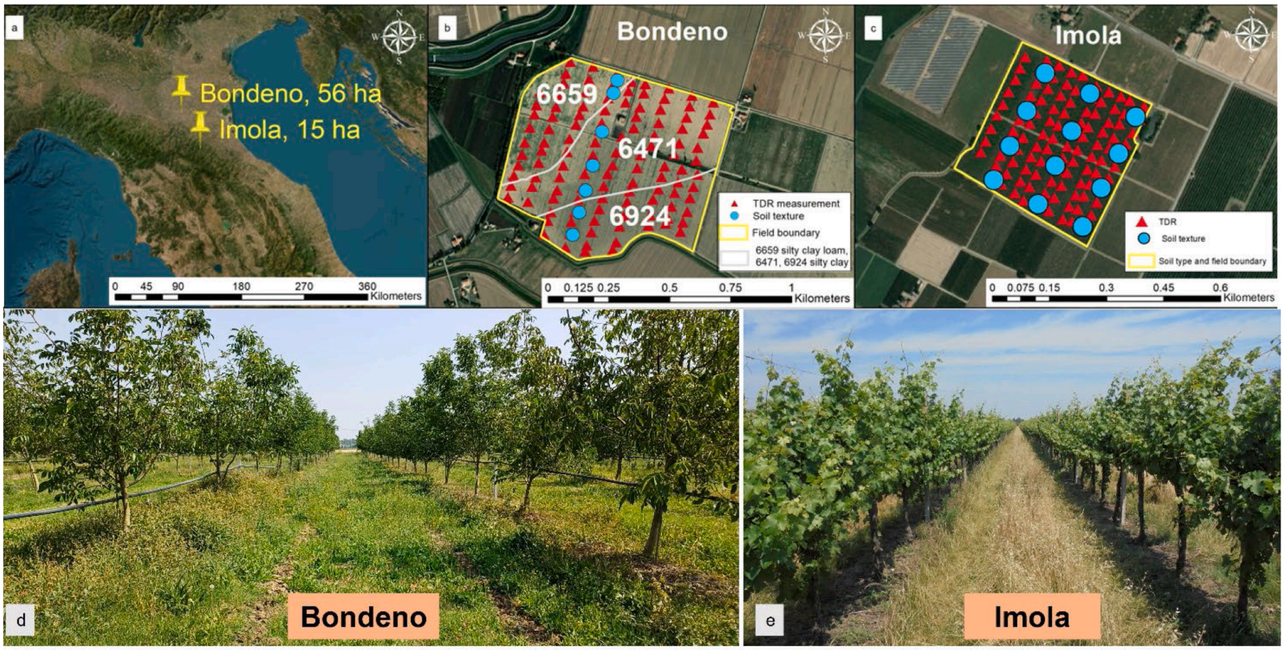

Fig. 1. Geographic locations of the experimental sites in northern Italy (a); Bondeno regional soil map, the soil samples locations, and TDR locations (b); Imola regional soil map, the soil samples, and TDR locations (c); picture of Bondeno on 12/06/2025 (d); picture of Imola on 17/06/2025 (e).

cases, each satellite pixel may capture a mixed signal of vegetation and bare soil. SAVI is particularly effective under these conditions, as it reduces soil influence and improves the accuracy of vegetation monitoring [37].

$$MSAVI = \frac{2\ NIR + 1 - \sqrt{(2NIR + 1)^2 - 8(NIR - RED)}}{2} \tag{3}$$

MSAVI dynamically adjusts the soil correction factor, further reducing soil noise without prior knowledge of vegetation density showing to outperforms SAVI in mixed-pixel scenarios, e.g., partial crop cover [12]. NDVI, SAVI, and MSAVI were selected because together they account for vegetation vigor, soil influence and canopy heterogeneity all of which are characteristic of Bondeno and Imola's perennial cropping systems.

2.3. Clustering algorithms

To explore how different clustering strategies affect agricultural management zonation, we applied three commonly used unsupervised learning techniques: K-means clustering, Gaussian Mixture Models (GMM), and Self-Organizing Maps (SOM). These methods were selected to represent a range of algorithmic approaches, from simpler distance-based clustering to more flexible probabilistic and neural network-based methods. This diversity allows us to assess how each method handles spatial complexity, vegetation heterogeneity, and non-linear patterns present in real agricultural environments [38–40].

K-means clustering is a widely used method due to its simplicity and computational efficiency [41]. It partitions the data into distinct, compact groups by minimizing the distance between individual data points and the center of their assigned cluster. In the context of agricultural zonation, K-means provides a straightforward approach to group areas with similar vegetation characteristics, making it useful for identifying general patterns across a field [42]. However, its assumption of spherical, equally sized clusters can limit its effectiveness in more complex landscapes [43].

GMM offers a more flexible approach by assuming that the data distribution is a combination of multiple overlapping Gaussian components [44]. Unlike K-means, GMM assigns probabilities to each data point's cluster membership, allowing for more nuanced classification in areas with gradual transitions or mixed vegetation conditions. This is particularly relevant in agricultural settings where crop health and soil conditions often change gradually rather than abruptly [45]. GMM is thus better suited for capturing uncertainty and overlapping zones that are common in heterogeneous agricultural fields [46].

Self-Organizing Maps (SOM) are a type of neural network that projects high-dimensional input data onto a two-dimensional grid while preserving topological relationships [47]. This method is especially effective for identifying complex, non-linear interactions in vegetation data, which are often influenced by multiple, interrelated factors such as soil properties, irrigation practices, and crop types [48]. SOM is valuable for exploratory analysis and visualization, helping to reveal subtle patterns that might be missed by more rigid clustering techniques [49].

In this study, all three methods were applied to classify the area into three agricultural zones. This number of classes was guided by agronomic relevance, ensuring the clusters were not only mathematically valid but also feasible for implementing land management. The choice of three zones also aligns with previous studies on crop health assessment [27] and provides a practical framework for targeted agricultural interventions.

2.4. Assessment of the agricultural zonation

To evaluate the quality of clustering outcomes in our zonation study, we used the v-measure, a validation metric particularly well-suited for datasets with imbalanced class distributions, a common feature in real-world agricultural settings [10,50]. Unlike other commonly used clustering metrics, such as the Adjusted Rand Index (ARI) [51], which can be difficult to interpret without a clear class-label mapping, or the Silhouette Coefficient [52] and Davies-Bouldin Index (DBI) [53] which focus more on the compactness of clusters rather than their relevance to external categories, the v-measure directly assesses how well clustering outcomes align with known or interpretable classes in the data.

The v-measure combines two key concepts that are critical in evaluating agricultural zonation results (Eq. (4)).

$$v - measure = \frac{2\ \times\ h\ \times\ c}{h\ +\ c} \tag{4}$$

Homogeneity (h) measures whether each cluster contains only data points from a single reference class. Completeness (c) evaluates whether all data points from a given class are assigned to the same cluster. These two properties reflect how accurately the clustering structure matches meaningful, domain-relevant groupings. The v-measure captures this balance by taking the harmonic mean of homogeneity and completeness. Its values range from 0 to 1, where 1 indicates perfect agreement with the reference classification.

3. Results

3.1. Reference zonation by means of ground observations

The information on vegetation, regional soil maps, soil texture analysis, and field surveys of SWC was compared to delineate a reference zonation for assessing RS approaches. At the Bondeno site, since the orchard was newly planted in 2019, vegetation growth was relatively limited during the two years of experiments and no significant spatial differences were observed. TDR surveys confirmed the spatial variability indicated by the regional soil map (Fig. 2a), and soil texture analysis confirmed the soil classification and the gradient from north to south (Fig. 2b). Based on these results, a reference zonation was established using the regional soil map as the primary source (see Fig. 2b).

In contrast, at the Imola site, vegetation growth showed some differences between the two grape varieties, with denser vegetation in the western section. In addition, soil texture analysis captured some variability, especially in the eastern section (Fig. 2d). This spatial variability was not identified in the regional soil map, but it was confirmed by the SWC surveys with the portable TDR (Fig. 2c). Based on these results, a reference zonation has been defined by combining the two distinct plant varieties and the soil variability detected in the north-south direction by the soil analyses (Fig. 2d).

While it is acknowledged that more sophisticated methods could have identified slightly different zone boundaries [54], this reference zonation was considered a practical perspective that farmers could use to support more precise management decisions. For this reason, this zonation is used as benchmark in the following sections for the assessment of the remote sensing approaches.

3.2. Comparison of zonation by means of vegetation indices

The zonation obtained based on the different RS approaches is compared by means of v-measure, where values close to one identify perfect clustering similarity. The results are shown in Fig. 3a. The results show an average low v-measure (around 0.2), indicating a general disagreement between zonation. The values, however span the entire range [0–1], highlighting a strong effect of the choice of the vegetation indices, the cluster method, and the acquisition time. The results are further aggregated and analysed to identify the role of each processing step, i.e., cluster method, vegetation indices (VI), and acquisition time.

Specifically, Fig. 3b compares the v-measure obtained between the cluster methods used (K-means, GMM, and SOM). Noteworthy, the v-measures between GMM and k-means are on average higher than 0.8, indicating a strong agreement between these two cluster methods. However, the values decrease when compared to SOM. While the values are still higher than 0.6, the results highlight some differences in the classification when using this cluster approach. The results are consistent between the experimental sites. As an example, the different

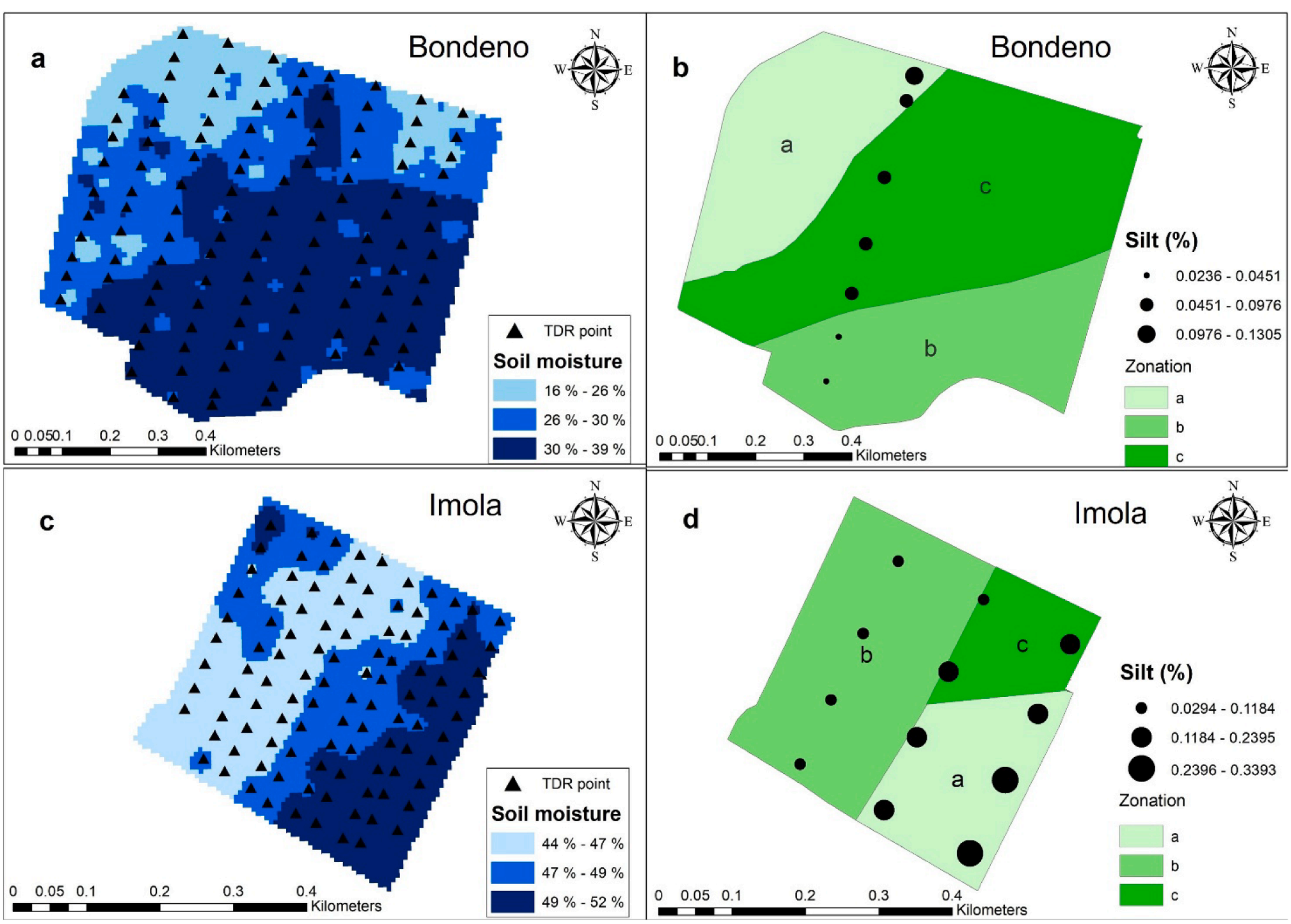

Fig. 2. Interpolated maps of the TDR surveys at Bondeno (a) and at Imola, (c); reference zonation with the results of the soil texture analyses at Bondeno (b) and at Imola (d).

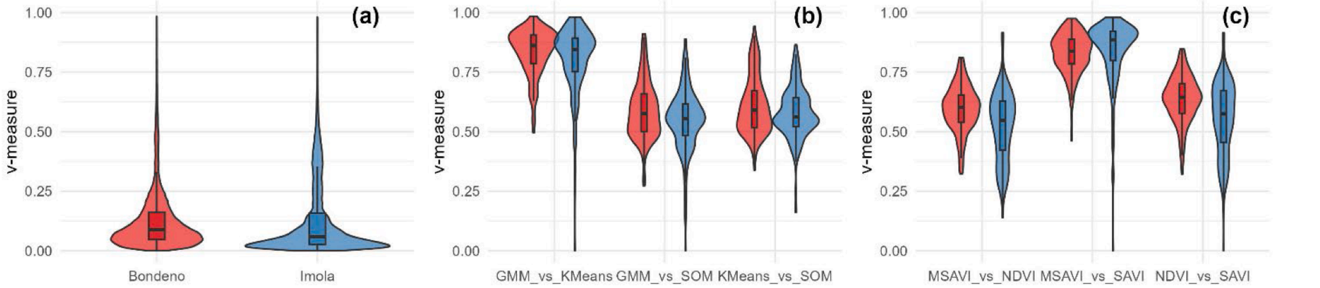

Fig. 3. v-measures between the remote sensing clusters at Bondeno and Imola sites: a) All the indices; b) By comparing the cluster methods; c) By comparing VIs.

zonation is shown in Figure S1 in the supplement material, where it is possible to notice how the clusters identified three zones with good agreement in the locations. The main differences are observed with SOM in the delineation of the southwest area.

The results are further analyzed by looking at the variability of the v-measure due to the vegetation indices (Fig. 3c). A similar behaviour to the cluster choice is noticed. Specifically, a strong agreement between the zonation is detected when comparing SAVI and MSAVI, indicating the similarities of these indices in the zonation. The agreement decreases to around 0.6 when using the NDVI. Also, in this case, the results are consistent between the experimental sites. As an example, the different zonation is shown in Figure S2 in the supplement material, where it is possible to notice very similar zonation with only small areas that are classified differently by using the NDVI.

The results analysed so far show some differences in the zonation by means of cluster methods and vegetation indices, but did not explain all the variability and the low v-measure obtained (Fig. 3a). For this reason, it is noted how these two processing steps did not strongly affect the zonation and could be considered, to some extent, equivalent.

The results were further analysed by looking at the effect of the acquisition time, which is expected to explain most of the differences in the zonation. Specifically, Fig. 4 (a and c) show the v-measure calculated comparing the zonation obtained between images acquired at different times (months) at Bondeno and Imola, respectively. The results show a low agreement between the clusters. Relatively high values (>0.25) are obtained only when the acquisition time is within the month (acquisition time = 0). Otherwise, the v-measures decrease on average to 0.1. An interesting seasonality is, however, detected, especially by looking at the results obtained at Imola (Fig. 4c). The v-measures, in fact, tend to increase when the difference in the acquisition time reaches one year (acquisition time = 12 months). This means that there is a good agreement between the clusters when the images are acquired during different years but the same month.

A final analysis is performed by selecting only the images acquired with a maximum time shift of one month, for which we have obtained the highest v-measure values. In this case, the results are shown in Fig. 4 (b and d) by also distinguishing the acquisition month. The results show a very remarkable seasonality for the Bondeno site (Fig. 4c). High values of v-measure (>0.5) are obtained only during wintertime (January and February) and late Autumn (November and December). In contrast, the v-measure values drop (<0.2 on average) when the images are acquired during the other months. The results indicate that zonation is strongly

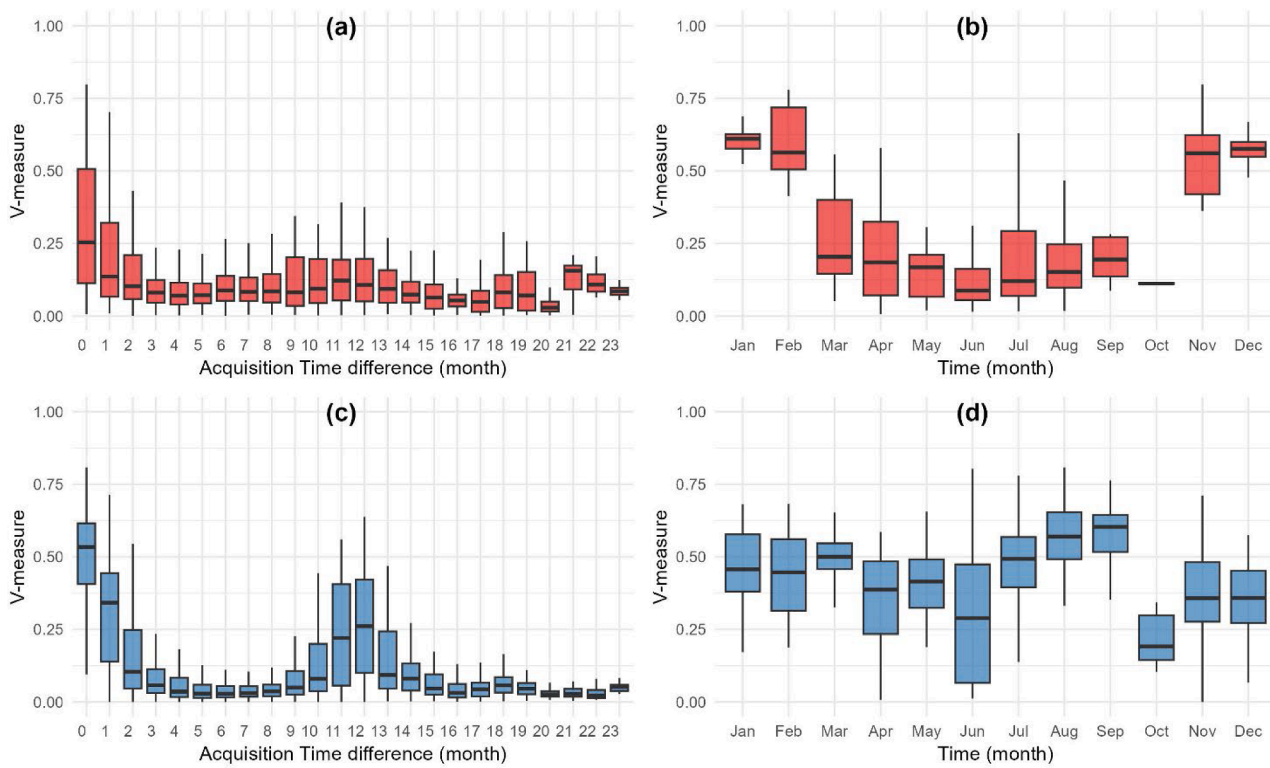

Fig. 4. (a) and (c): v-measures obtained over different acquisition times for Bondeno and Imola, respectively; (b) and (d): v-measure based on images acquired within the same month but over the year.

unstable during the growing season, probably due to plant heterogeneity and crop management. This behaviour is less evident at the Imola site (Fig. 4d), where the v-measures are on average higher (around 0.5) with only a slight decrease during the summer months (e.g., June). Noteworthy, v-measure drops in October. This result highlights the role of the vegetation type and crop management on the zonation.

3.3. Assessment of the zonation by remote sensing approaches

The v-measures calculated by comparing the zonation obtained with ground observations and RS approaches are shown in Fig. 5 (a, b). Two main aspects can be highlighted. On average, the v-measures are relatively low, especially for the Bondeno site. Moreover, a strong temporal variability is detected, especially at the Imola site.

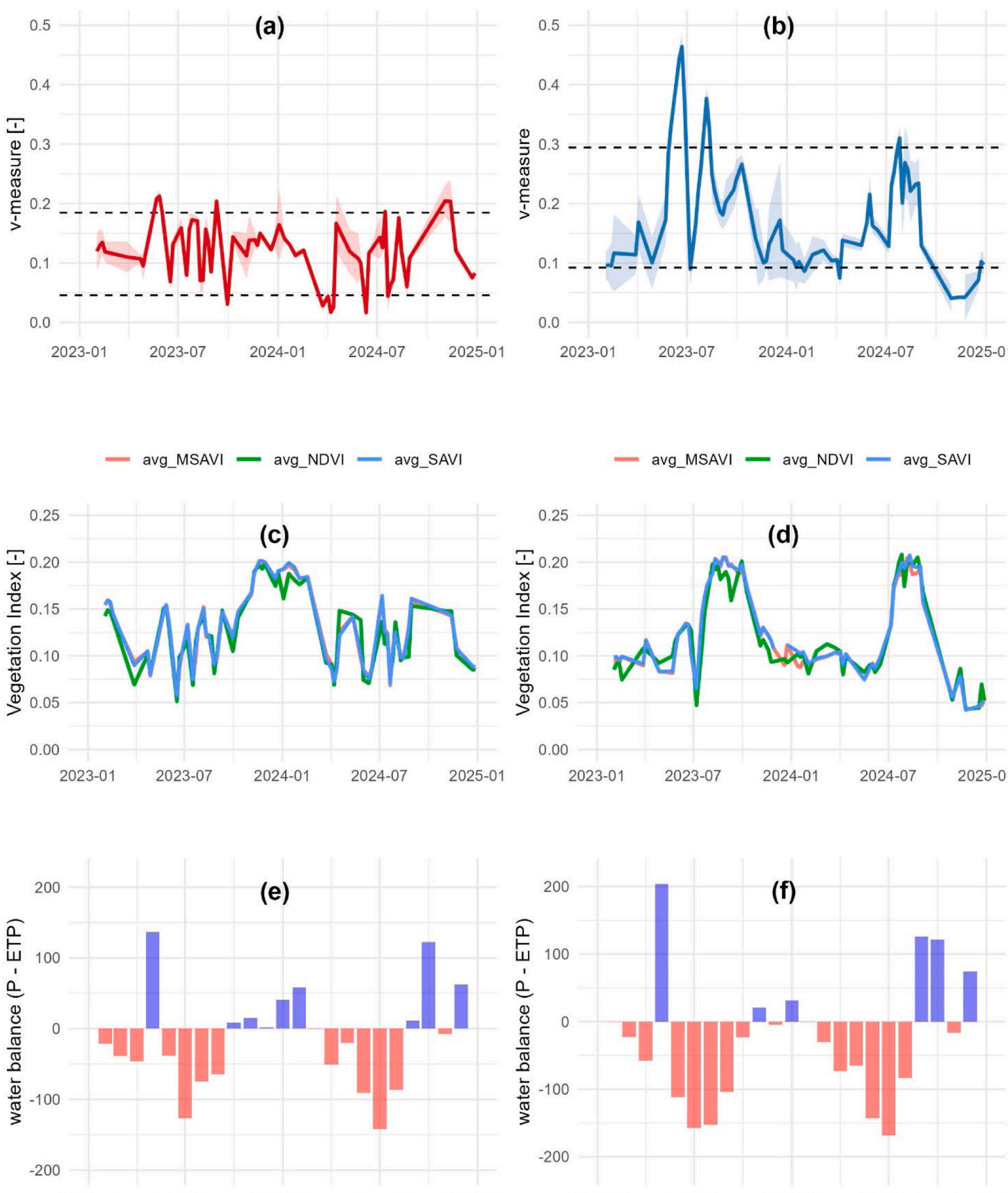

Fig. 5. v-measure between the RS clusters and the reference zonation, at Bondeno (a) and Imola (b), respectively; average VI over time at the two sites (c and d); water balance at the sites calculated based on precipitation (P) – potential evapotranspiration (ETP) (e and f).

Specifically, v-measure at the Bondeno site is on average = 0.2. This can be explained considering that the reference zonation is mainly driven by soil variability (see Section 3.1). Moreover, the walnut orchard was established in 2019, and during the study period, the trees were still in the early stages of development, with limited canopy closure and structural complexity. As a result, the orchard's vegetation cover did not have a strong effect on the spectral response across the field. In contrast, RS indices are mainly affected by the land management, and thus, they do not capture the reference zonation detected with the ground observations.

At the Imola site, the v-measures are, on average, higher than those observed at Bondeno, with a mean value of approximately 0.3. This improved agreement between RS and ground-based zonation can be attributed to crop-related variability. Specifically, the reference zonation is driven by the presence of two distinct grapevine varieties and the long-established nature of the vineyard, which was planted over 20 years ago. Over time, differences in growth dynamics and plant vigor have likely become more pronounced due to the accumulated effects of management practices and environmental conditions. Additionally, cultural operations such as summer pruning, performed occasionally between June and July, introduce further variability in canopy structure. Under these conditions, RS vegetation indices are more effective at capturing the spatial heterogeneity present in the field, leading to a better correspondence with the reference zonation. At both sites, however, the v-measures strongly change over time, highlighting a critical role of the agro-environmental conditions for a proper zonation. In particular, v-measures show quick changes over time at Bondeno, while a more seasonal variation is detected at Imola.

To better understand these behaviours, the temporal dynamic of the v-measure values is compared to the spatial average vegetation indices (Fig. 5c and d). Please note that for a better visualization the indices have been scaled to compare the dynamic rather than the actual absolute values. Moreover, a monthly water balance between precipitation (P) and potential evapotranspiration (ETP) is calculated to assess the hydrological conditions over the area (Fig. 5e and f). Precipitation and weather data have been acquired by the weather stations operated by the regional environmental protection agency [55]. ETP is calculated based on the Penman-Monteith equation. It is possible to note how, on average, the vegetation indices at Bondeno site show sharp changes and do not well capture the growing season (March-August). This confirms how the vegetation indices mainly reproduced the cover crop and the land management, while the young walnut plants are less represented. In contrast, two clear growing seasons are detected at Imola site with the highest vegetation indices values obtained around July. No significant differences are obtained by using the three vegetation indices, confirming the relatively small role of the choice of the index on this type of assessment. By looking at the monthly water balance (P - ET), similar conditions are shown at the two experimental sites. A very dry condition has been experienced during 2023, with only an exceptional intensive precipitation event in May that generated floods and damage over the area [56]. In comparison, a more common weather seasonality has been experienced during 2024, with a relatively wet season followed by a dry summer period.

By comparing these variables, it is noteworthy how the highest v-measures (above the 90 percentiles, dashed line in Fig. 5a and b) are detected when vegetation is actively growing and hydrological conditions are relatively dry. This phenomenon is consistent with findings from previous studies, where vegetation indices and clustering algorithms perform better during periods of strong vegetation signal and reduced soil moisture [57–59]. In contrast, the v-measures are low (below the 10 percentiles, dashed line in Fig. 5a and b) when the value of the vegetation indices is lower and after precipitation events. This behaviour is particularly evident at the Imola site, where v-measure reached values above 0.3 during the growing and dry season in both years. Remarkably, the v-measures drop after the intense precipitation event on May 2023. As an example, the best and worst RS clusters in matching the reference ground-based zonation are shown in figure S3 in the supplement material. It is possible to notice how very different areas have been identified when using different acquisition time.

5. Summary and conclusions

This study systematically compares the effect of different data processing flows to identify homogenous zones and to support precision agriculture based on remote sensing images. The choice of the vegetation indices (VIs), of the cluster approaches and of the acquisition time is assessed. Data have been collected at two experimental sites over two years.

The results show that the choice of the vegetation index and of the cluster approach affects the zonation, but only to some minor extend. The zonation obtained based on NDVI shows some differences (v-measure around 0.6) from the cluster obtained with SAVI and MASVI. Similarly, the zonation obtained based on SOM cluster approach shows some differences (v-measure around 0.6) from the cluster obtained with k-means and GMM. However, these differences are negligible in comparison to the effect of the image acquisition time. In this last case, the classifications are consistent only when the images are acquired within one month (v-measure > 0.5). In contrast, zonation is very different when comparing images acquired over longer time period (> 1 month). Moreover, a seasonality appears over one year period, i.e., the classifications based on images acquired during the same month but at different year are more consistent. The results are explained based on the strong effect of land management and agronomic practices on the (VIs) conducted at the sites. For this reason, the study highlights the need to identify the best acquisition time for zonation. In contrast, the zonation based on the combination of images acquired at different times might be misleading and, for this reason, this approach is not advised.

A comparison with crop growth (average VIs) and agro-environmental conditions (water balance) is conducted to further support the selection of the best image acquisition period. The analysis shows that the greatest performance in zonation (v-measure > 0.5) was realized during the growing seasons (relative high VIs) and during relatively dry conditions. In these conditions, the growing vegetation is expected to be affected by the soil conditions and the VIs better capture the variability of the soil-plant system. In contrast, the quality of zonation decreased when images were acquired when vegetation was not growing actively (winter time) or during wet periods.

These findings have important implications for the development and use of remote sensing approaches to support zonation and precision agriculture. While an increasing number of satellites are in fact providing higher resolutions images that can be valuable in many applications, the results obtained within the present study highlight that the choice of the data processing workflow considerably affect the outcomes. Specifically, the selection of VIs and cluster approaches showed to play a limited role in zonation. In contrast, the definition of the acquisition time showed to be crucial. Although these results may be specific to certain characteristics of the experimental sites (e.g., orchards), similar systematic comparisons across different agro-environmental conditions should be undertaken to identify key research gaps and to develop clear guidelines to producing informative zonation maps to support agronomic practices.

Data processing and codes

A combination of platforms and programming languages was used in this study to process satellite images, perform zonation, conduct the statistical analysis and visualize the results. Most of the data processing was carried out on a standard personal computer (Manufacturer: DELL; Processor: 11th Gen Intel(R), Core(TM) i7-1165G7, 2.80 GHz; Installed RAM 32.0 GB; System type: 64-bit operating system, x64-based processor). Several Python libraries were used: geopandas for vectors data handling, rasterio for raster processing, numpy for numerical computations and MiniSom for SOM clustering [60]. ArcMap was used for spatial visualization and map refinements. Some processing steps were however accelerated using freely available cloud computing resources. Specifically, Google Earth Engine (GEE) was used to acquire and process Sentinel-2 imagery, and to compute vegetation indices (NDVI, SAVI, MSAVI). Classification methods such as K-means, Gaussian Mixture Models (GMM), and Self-Organizing Maps (SOM) has been performed in Google Colab. For reproducibility, a simplified data processing workflow has been integrated in GEE for End-Users [61]. The code is available at the following page: https://github.com/gunaysunmoon/Sentinel-2-NDVI-Zonation-Analysis_RS_GIS.git. Users can select the area of interest, choose Sentinel-2 acquisition time, choose a vegetation index and apply K-means to delineate three zones.

Acknowledgements

The first Author received a PhD scholarship funded by ADA University and the Ministry of Science and Education of Azerbaijan, as part of the State Program for Youth to Study in Prestigious Foreign Higher Education Institutions for 2019–2023. This support is also aligned with the Italian-Azerbaijani University project, a collaborative initiative between the governments of Azerbaijan and Italy. This research has also been supported by the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Program and by the Participating States (funder name, European Partnership on Metrology; funder ID, 10.13039/100,019,599; grant no 21GRD08 SoMMet). We gratefully acknowledge Societa Agricola CACI (Imola, Bologna, Italy) and Andrea Scalabrin from Società Agricola Agro Noce S.r.l. (Bondeno, Ferrara, Italy) for their permission to conduct the trial.

Ethical statement

This research did not involve human participants or animals. Therefore, ethical approval was not required.

CRediT authorship contribution statement

Gunay Hasanli: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Data curation, Conceptualization. **Sadra Emamalizadeh:** Writing – review & editing, Validation, Software, Resources, Formal analysis. **Riccardo Mazzoleni:** Writing – review & editing, Resources, Formal analysis, Data curation. **Marco Benfenati:** Writing – review & editing. **Gabriele Baroni:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2026.101911.

Data availability

I have shared the link.

References

[1] A. Ali, V. Rondelli, R. Martelli, G. Falsone, F. Lupia, L. Barbanti, Management zones delineation through clustering techniques based on soils traits, NDVI data, and multiple year crop yields, Agriculture 12 (2022) 231, https://doi.org/10.3390/agriculture12020231.
[2] R. Gebbers, V.I. Adamchuk, Precision agriculture and food security, Science 327 (2010) 828–831, https://doi.org/10.1126/science.1183899.
[3] X. Song, J. Wang, W. Huang, L. Liu, G. Yan, R. Pu, The delineation of agricultural management zones with high resolution remotely sensed data, Precis. Agric. 10 (2009) 471–487, https://doi.org/10.1007/s11119-009-9108-2.
[4] D.J. Mulla, Twenty five years of remote sensing in precision agriculture: key advances and remaining knowledge gaps, Biosyst. Eng. 114 (2013) 358–371, https://doi.org/10.1016/j.biosystemseng.2012.08.009.
[5] M.A. Wulder, J.G. Masek, W.B. Cohen, T.R. Loveland, C.E. Woodcock, Opening the archive: how free data has enabled the science and monitoring promise of Landsat, Remote Sens Environ. 122 (2012) 2–10, https://doi.org/10.1016/j.rse.2012.01.010.
[6] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, et al., Sentinel-2: ESA's optical high-resolution mission for GMES operational services, Remote Sens Environ. 120 (2012) 25–36, https://doi.org/10.1016/j.rse.2011.11.026.
[7] H. Kimm, K. Guan, C. Jiang, B. Peng, L.F. Gentry, S.C. Wilkin, et al., Deriving high-spatiotemporal-resolution leaf area index for agroecosystems in the U.S. Corn Belt using Planet Labs CubeSat and STAIR fusion data, Remote Sens Environ. 239 (2020) 111615, https://doi.org/10.1016/j.rse.2019.111615.
[8] S. Huang, L. Tang, J.P. Hupy, Y. Wang, G. Shao, A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing, J. Res. 32 (2021) 1–6, https://doi.org/10.1007/s11676-020-01155-1.
[9] A. Bigdeli, A. Maghsoudi, R. Ghezelbash, Application of self-organizing map (SOM) and K-means clustering algorithms for portraying geochemical anomaly patterns in Moalleman district, NE Iran, J. Geochem. Explor. 233 (2022) 106923, https://doi.org/10.1016/j.gexplo.2021.106923.
[10] T. Roßberg, Schmitt M.A Globally, Applicable method for NDVI estimation from Sentinel-1 SAR backscatter using a deep neural network and the SEN12TP dataset, PFG – J. Photogramm. Remote Sens. Geoinf. Sci. 91 (2023), https://doi.org/10.1007/s41064-023-00238-y.
[11] A.R. Huete, H.Q. Liu, K. Batchily, W. van Leeuwen, A comparison of vegetation indices over a global set of TM images for EOS-MODIS, Remote Sens Environ. 59 (1997) 440–451, https://doi.org/10.1016/S0034-4257(96)00112-5.
[12] J. Qi, A. Chehbouni, A. Huete, Y. Kerr, S. Sorooshian, A modified soil adjusted vegetation index, Remote Sens Environ. 48 (1994) 119–126, https://doi.org/10.1016/0034-4257(94)90134-1.
[13] S.A. Shammi, Q. Meng, Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling, Ecol. Indic. 121 (2020), https://doi.org/10.1016/j.ecolind.2020.107124.
[14] S. Beninato, M. Holzman, R. Rivas, Sensitivity of the land surface temperatureand normalized difference water index tosoil moisture on wheat and barleySensibilidad de la temperatura superficial terrestre y del Índice de Diferencia Normalizadadel Agua a la humedad del suelo en trigo y cebada, AgriScientia 42 (2025) 21–33, https://doi.org/10.31047/1668.298x.v42.n1.46537.
[15] Shan Y., Gong A., Su Y., Liu W., Li J., Chen Y., et al. Improvement of soil moisture monitoring using EVI as a key parameter based on TVDI in the north China plain. 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, Melbourne, Australia: IEEE; 2013, p. 3738–41. https://doi.org/10.1109/IGARSS.2013.6723643.
[16] Ilyas F., Priscila S. An optimized clustering quality analysis in K-means cluster using silhouette scores, 2024, p. 49–63. https://doi.org/10.4018/979-8-3693-1355-8.ch004.
[17] G. Fang, C. Wang, T. Dong, Z. Wang, C. Cai, J. Chen, et al., A landscape-clustering zoning strategy to map multi-crops in fragmented cropland regions using Sentinel-2 and Sentinel-1 imagery with feature selection, Agriculture 15 (2025) 186, https://doi.org/10.3390/agriculture15020186.
[18] Filippi A., Dobreva I., Klein A., Jensen J. Self-organizing map-based applications in remote sensing, 2010. https://doi.org/10.5772/9163.
[19] S. Licen, A. Astel, S. Tsakovski, Self-organizing map algorithm for assessing spatial and temporal patterns of pollutants in environmental compartments: a review, Sci. Total Environ. 878 (2023) 163084, https://doi.org/10.1016/j.scitotenv.2023.163084.
[20] Kaliffe K., Santos R. Finding the best tennis serves with K-means and GMM clusters of ball tracking data to interpret serve strategies. 2024. https://doi.org/10.5753/kdmile.2024.244569.
[21] Sharif R., Ghomi M., Taefi M. A novel approach to identifying open star cluster members in Gaia Dr3: integrating mst and GMM techniques. 2025. https://doi.org/10.2139/ssrn.5118409.
[22] P.T. Sorenson, S.J. Shirtliffe, A.K. Bedard-Haughn, Predictive soil mapping using historic bare soil composite imagery and legacy soil survey data, Geoderma 401 (2021) 115316, https://doi.org/10.1016/j.geoderma.2021.115316.
[23] Q. Yu, W. Ji, R. Pu, S. Landry, M. Acheampong, J. O' Neil-Dunne, et al., A preliminary exploration of the cooling effect of tree shade in urban landscapes, Int. J. Appl. Earth Obs. Geoinf. 92 (2020) 102161, https://doi.org/10.1016/j.jag.2020.102161.
[24] V. Sainte Fare Garnot, L. Landrieu, N. Chehata, Multi-modal temporal attention models for crop mapping from satellite time series, ISPRS J. Photogramm. Remote Sens. 187 (2022) 294–305, https://doi.org/10.1016/j.isprsjprs.2022.03.012.
[25] F. Reyes, R. Casa, M. Tolomio, M. Dalponte, N. Mzid, Soil properties zoning of agricultural fields based on a climate-driven spatial clustering of remote sensing time series data, Eur. J. Agron. 150 (2023) 126930, https://doi.org/10.1016/j.eja.2023.126930.

[26] M.L. Villarreal, L.M. Norman, S. Buckley, C.S.A. Wallace, M.A. Coe, Multi-index time series monitoring of drought and fire effects on desert grasslands, Remote Sens Environ. 183 (2016) 186–197, https://doi.org/10.1016/j.rse.2016.05.026.
[27] Q. Zhang, Q. Yuan, C. Zeng, X. Li, Y. Wei, Missing data reconstruction in remote sensing image with a unified spatial–Temporal–Spectral deep convolutional neural network, IEEE Trans. Geosci. Remote Sens. 56 (2018) 4274–4288, https://doi.org/10.1109/TGRS.2018.2810208.
[28] N. Bohn, L. Guanter, T. Kuester, R. Preusker, K. Segl, Coupled retrieval of the three phases of water from spaceborne imaging spectroscopy measurements, Remote Sens Environ. 242 (2020) 111708, https://doi.org/10.1016/j.rse.2020.111708.
[29] G. Hasanli, S. Emamalizadeh, R. Mazzoleni, On the use of remote sensing vegetation indices for agricultural zones' delineation, 2024, https://doi.org/10.1109/MetroAgriFor63043.2024.10948870.
[30] P.R. Rossini, I.A. Ciampitti, T. Hefley, A. Patrignani, A soil moisture-based framework for guiding the number and location of soil moisture sensors in agricultural fields, Vadose Zone j 20 (2021), https://doi.org/10.1002/vzj2.20159.
[31] G.Y. Lu, D.W. Wong, An adaptive inverse-distance weighting spatial interpolation technique, Comput. Geosci. 34 (2008) 1044–1055, https://doi.org/10.1016/j.cageo.2007.07.010.
[32] R. Houborg, M.F. McCabe, A Cubesat enabled Spatio-Temporal Enhancement method (CESTEM) utilizing planet, Landsat and MODIS data, Remote Sens Environ. 209 (2018) 211–226, https://doi.org/10.1016/j.rse.2018.02.067.
[33] K. Tarrio, X. Tang, J.G. Masek, M. Claverie, J. Ju, S. Qiu, et al., Comparison of cloud detection algorithms for Sentinel-2 imagery, Sci. Remote Sens. 2 (2020) 100010, https://doi.org/10.1016/j.srs.2020.100010.
[34] E.P. Glenn, P.L. Nagler, A.R. Huete, Vegetation index methods for estimating evapotranspiration by remote sensing, Surv. Geophys. 31 (2010) 531–555, https://doi.org/10.1007/s10712-010-9102-2.
[35] A.R. Huete, A soil-adjusted vegetation index (SAVI), Remote Sens Environ. 25 (1988) 295–309, https://doi.org/10.1016/0034-4257(88)90106-X.
[36] D. Radocaj, A. Šiljeg, R. Marinović, M. Jurišić, State of major vegetation indices in precision agriculture studies indexed in Web of science: a review, Agriculture 13 (2023) 707, https://doi.org/10.3390/agriculture13030707.
[37] L. Galvão, J. Santos, D. Roberts, F. Breunig, M. Toomey, Y. Moura, On intra-annual EVI variability in the dry season of tropical forest: a case study with MODIS and hyperspectral data, Remote Sens Environ. 115 (2011) 2350–2359, https://doi.org/10.1016/j.rse.2011.04.035.
[38] Thrun M. Approaches to cluster analysis, 2018, p. 21–31. https://doi.org/10.1007/978-3-658-20540-9_3.
[39] B. Everitt (Ed.), *Cluster Analysis*, 5. ed., Wiley, 2011.
[40] Csizér K., Jamieson J. Cluster analysis, 2012. https://doi.org/10.1002/9781405198431.wbeal0138.
[41] J. MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics 5.1, University of California Press, 1967, pp. 281–298.
[42] Mucherino A., Papajorgji P., Pardalos P. Data mining in agriculture. vol. 34. 2009. https://doi.org/10.1007/978-0-387-88615-2.
[43] A.K. Jain, R.C. Dubes, Algorithms For Clustering Data, Prentice-Hall, Inc., USA, 1988.
[44] B. Li, G. Guo, Optimizing electric vehicle clustering through improved MVABC-GMM algorithm based on gaussian mixture model and Monte Carlo prediction, Acad. J. Sci. Technol. 14 (2025) 90–94, https://doi.org/10.54097/g2mn1262.
[45] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.
[46] C. Fraley, A.E. Raftery, Model-based clustering, discriminant analysis, and density estimation, J. Am. Stat. Assoc. 97 (2002) 611–631, https://doi.org/10.1198/016214502760047131.
[47] T. Kohonen, Variants of SOM, in: T. Kohonen (Ed.), Self-Organizing Maps, Springer, Berlin, Heidelberg, 2001, pp. 191–243, https://doi.org/10.1007/978-3-642-56927-2_5.
[48] M. Tobiszewski, S. Tsakovski, V. Simeonov, J. Namieśnik, Surface water quality assessment by the use of combination of multivariate statistical classification and expert information, Chemosphere 80 (2010) 740–746, https://doi.org/10.1016/j.chemosphere.2010.05.024.
[49] J. Vesanto, E. Alhoniemi, Clustering of the self-organizing map, IEEE Trans. Neural. Netw. 11 (2000) 586–600, https://doi.org/10.1109/72.846731.
[50] J. Nowosad, T.F. Stepinski, Spatial association between regionalizations using the information-theoretical V-measure, Int. J. Geogr. Inf. Sci. 32 (2018) 2386–2401, https://doi.org/10.1080/13658816.2018.1511794.
[51] L. Hubert, P. Arabie, Comparing partitions, J. Classif. 2 (1985) 193–218, https://doi.org/10.1007/BF01908075.
[52] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math. 20 (1987) 53–65, https://doi.org/10.1016/0377-0427(87)90125-7.
[53] D.L. Davies, D.W. Bouldin, A cluster separation measure, IEEE Trans. Pattern Anal. Mach. Intell. (PAMI–1) (1979) 224–227, https://doi.org/10.1109/TPAMI.1979.4766909.
[54] S.M. Becker, T.E. Franz, O. Abimbola, D.D. Steele, J.P. Flores, X. Jia, et al., Feasibility assessment on use of proximal geophysical sensors to support precision management, Vadose Zone J. (2022), https://doi.org/10.1002/vzj2.20228.
[55] G. Antolini, L. Auteri, V. Pavan, F. Tomei, R. Tomozeiu, V. Marletto, A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961-2010, Int. J. Clim. 36 (2016) 1970–1986, https://doi.org/10.1002/joc.4473.
[56] C. Arrighi, A. Domeneghetti, Brief communication: on the environmental impacts of the 2023 floods in Emilia-Romagna (Italy), Nat. Hazards Earth Syst. Sci. 24 (2024) 673–679, https://doi.org/10.5194/nhess-24-673-2024.
[57] J. Bendig, A. Bolten, G. Bareth, UAV-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability, Photogramm. - Fernerkund. - Geoinf. 6 (2013) 551–562, https://doi.org/10.1127/1432-8364/2013/0200.
[58] B. Fang, V. Lakshmi, Soil moisture at watershed scale: remote sensing techniques, J. Hydrol. 516 (2014) 258–272, https://doi.org/10.1016/j.jhydrol.2013.12.008.
[59] J.M. Peña-Barragán, M. Ngugi, R.E. Plant, J. Six, Object-based crop identification using multiple vegetation indices, textural features and crop phenology, Remote Sens Environ. 115 (2011) 1301–1316, https://doi.org/10.1016/j.rse.2011.01.009.
[60] Ahmad M., Ali M., Hasan M.R., Mobo F., Rai S. Geospatial Machine Learning and the power of Python programming: libraries, tools, applications, and plugins, 2024, p. 223–53. https://doi.org/10.4018/979-8-3693-6381-2.ch010.
[61] Jayanthi V., Akhil S., Pavan P., Seenu P. Crop area estimation using Sentinel-2 and GEE, 2024, p. 457–67. https://doi.org/10.1007/978-981-97-7467-8_30.