Master of Science
Artificial Intelligence and Visual Computing

UNIVERSITY OF WEST ATTICA &

UNIVERSITY OF LIMOGES

"Detection and Classification of Grapevine Varieties using Multispectral Image Analysis and Vegetation Indices"

Olympia Kourounioti

Athens, July 2023

Supervisor: Associate professor Emmanouil Oikonomou

Μέλη Εξεταστικής Επιτροπής συμπεριλαμβανομένου του εισηγητή

Η διπλωματική εργασία εξετάστηκε επιτυχώς από την κάτωθι Εξεταστική Επιτροπή

Α/Α	ΟΝΟΜΑΤΕΠΩΝΥΜΟ	ΒΑΘΜΙΔΑ/ΙΔΙΟΤΗΤΑ	ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ
1	Εμμανουήλ Οικονόμου	Αναπληρωτής Καθηγητής	
2	Αναστάσιος Κεσίδης	Καθηγητής	
3	Λάζαρος Γραμματικόπουλος	Αναπληρωτής Καθηγητής	

Ευχαριστίες

Ολοκληρώνοντας τη διπλωματική μου εργασία θα ήθελα να ευχαριστήσω όσους με βοήθησαν, με καθοδήγησαν και με στήριξαν όλο αυτό το διάστημα.

Αρχικά, θα ήθελα να ευχαριστήσω τον Αναπληρωτή Καθηγητή κ. Εμμανουήλ Οικονόμου για την καθοδήγηση και την επίβλεψη κατά τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας. Ακόμα θα ήθελα να ευχαριστήσω τον Αναπληρωτή Καθηγητή κ. Λάζαρο Γραμματικόπουλο για τις συζητήσεις και τις συμβουλές του το διάστημα αυτό. Επιπλέον θα ήθελα να ευχαριστήσω τον Επίκουρο Καθηγητή κ. Βασίλειο Κρασανάκη καθώς και τον Μωυσή, τον Ανδρέα, τον Παναγιώτη και τον Vishnu για την υποστήριξη και την εμψύχωση όλους αυτούς τους μήνες. Ακόμα θέλω να ευχαριστήσω τον Στέφανο για την υπομονή και τις συζητήσεις σε όλη τη διάρκεια του Μεταπτυχιακού. Τέλος, ένα ευχαριστώ στους φίλους και την οικογένειά μου που με στηρίζουν σε ότι κάνω.

Θα ήθελα ακόμα να ευχαριστήσω το Εργαστήριο Τηλεπισκόπησης που Εθνικού Μετσόβιου Πολυτεχνείου που μου παρείχαν τα δεδομένα εικόνων πάνω στα οποία υλοποιήθηκε η εργασία αυτή καθώς και το Ινστιτούτο Τεχνολογίας Αγροτικών Προϊόντων και συγκεκριμένα τον κ. Δημήτρη Τάσκο που μου έδωσε πρόσβαση στο χώρο των αμπελώνων τους για την πραγματοποίηση επίγειων μετρήσεων στο πλαίσιο της διπλωματικής εργασίας.

Contents

Figures

Tables

Abstract

The present master thesis focuses on investigating and analyzing issues related to the detection of grapevine varieties through the analysis of specific grapevine properties. These properties mostly concern the amount of chlorophyll present in the leaves of the plants, the overall health or the existence of any signs of stress, which affects significantly the process of photosynthesis and the density of the vines' canopy.

In order to provide a comprehensive theoretical approach of the issue, a narrative literature review was conducted, focusing on various aspects related to vines. This review encompassed an exploration of common diseases affecting vines, the methodologies employed for data collection and subsequent analysis, as well as the techniques used for identifying diseases or deficiencies in vine plants. Additionally, a systematic literature review was conducted to examine the identification of different vegetation varieties through image analysis.

To evaluate the health status of the vine, vegetation indicators were employed with the aim of identifying the aforementioned properties. Data collection in the study area included multispectral images captured by other researchers in 2022 through a UAV (Unmanned Aerial Vehicle) flight over the vineyards of the Agricultural Products Technology Institute in Lykovrisi. These images were used to generate an orthophoto of the area, which was subsequently examined. The study area has the unique characteristic of housing wide and diverse vine varieties, comprising over a thousand different types. As part of the research, a sub-vineyard of four acres area was examined, encompassing a total of 112 vine varieties. To accurately detect the precise locations of each of these varieties, ground surveying measurements were taken and a vine map was created.

To isolate the parts depicting vines on the image, supervised classification algorithms were executed. After the classification process, a mask was applied to cover the areas of the image that represented shadows or the ground. In the remaining areas (vine areas), seven vegetation indices were implemented. Among these indices, two pairs that exhibited the least correlation with each other were selected, resulting in a total of four indices. The indices' outcomes were then incorporated by two clustering algorithms: a vector quantization algorithm and a probabilistic model. The primary objective of these algorithms was to group varieties with similar characteristics into classes. Each algorithm was implemented using the four vegetation indices, and their respective outputs were subsequently compared to identify the varieties classified in the same class by both algorithms. This process of identifying varieties with similar characteristics provides valuable insights into vine growth and health. Furthermore, it enables farmers to employ similar approaches on varieties with similar properties for essential works, such as watering, fertilization, and harvesting.

Περίληψη

Στην παρούσα εργασία ερευνήθηκαν και αναλύθηκαν ζητήματα που αφορούν την ανίχνευση ποικιλιών αμπελιών μέσω της ανάλυσης ιδιοτήτων του αμπελιού. Οι ιδιότητες αυτές αφορούν την ποσότητα της χλωροφύλλης στα φύλλα των φυτών, την υγεία ή την ύπαρξη στρες, γεγονός το οποίο συνδέεται με την διαδικασία της φωτοσύνθεσης, καθώς και την πυκνότητα του φυλλώματος των αμπελιών.

Για την θεωρητική προσέγγιση του θέματος πραγματοποιήθηκε αρχικά περιγραφική βιβλιογραφική ανασκόπηση η οποία εστιάζει στα αμπέλια, στις ασθένειες από τις οποίες συνήθως προσβάλλονται, τις μεθόδους με τις οποίες γίνεται η συλλογή δεδομένων, και τις τεχνικές με τις οποίες εντοπίζονται ασθένειες ή ελλείψεις στα αμπέλια. Επιπλέον πραγματοποιήθηκε και συστηματική βιβλιογραφική ανασκόπηση που αφορά τον εντοπισμό διαφορετικών ποικιλιών βλάστησης μέσω εικόνων.

Για την ανάλυση της κατάστασης της υγείας του αμπελιού, εφαρμόστηκαν δείκτες βλάστησης οι οποίοι στοχεύουν στον εντοπισμό των ιδιοτήτων που προαναφέρθηκαν. Σχετικά με την συλλογή δεδομένων στην περιοχή μελέτης, χρησιμοποιήθηκαν πολυφασματικές εικόνες οι οποίες συλλέχθηκαν από άλλους ερευνητές το 2022 μέσω πτήσης UAV (Unmanned aerial vehicle) στους αμπελώνες του Ινστιτούτο Τεχνολογίας Αγροτικών Προϊόντων στην Λυκόβρυση Αττικής. Από τις εικόνες αυτές παράχθηκε ορθοφωτογραφία της περιοχής η οποία έχει το ιδιαίτερο χαρακτηριστικό ότι περιέχει πλήθος διαφορετικών ποικιλιών αμπελιών, και συγκεκριμένα περισσότερες από χίλιες. Στο πλαίσιο της εργασίας μελετήθηκε μία υπό-περιοχή τεσσάρων στρεμμάτων στην οποία βρίσκονται 112 ποικιλίες αμπελιού. Για να εντοπιστεί η ακριβής θέση κάθε ποικιλίας πραγματοποιήθηκαν επίγειες τοπογραφικές μετρήσεις και δημιουργήθηκε χάρτης των ποικιλιών.

Προκειμένου να απομονωθούν τα τμήματα της εικόνας που απεικονίζουν αμπέλια, εκτελέστηκαν αλγόριθμοι επιβλεπόμενης ταξινόμησης. Μετά τη διαδικασία ταξινόμησης τοποθετήθηκε μια μάσκα στις περιοχές τις εικόνας που απεικόνιζαν σκιές ή έδαφος. Στις υπόλοιπες περιοχές (αμπέλια) εφαρμόστηκαν επτά δείκτες βλάστησης. Μεταξύ αυτών των δεικτών επιλέχθηκαν δύο ζεύγη που εμφάνιζαν τη μικρότερη συσχέτιση μεταξύ τους, οπότε συνολικά χρησιμοποιήθηκαν τέσσερις δείκτες. Τα αποτελέσματα από την εφαρμογή των δεικτών χρησιμοποιήθηκαν από δύο αλγορίθμους συσταδοποίησης, έναν αλγόριθμο διανυσματικής κβαντοποίησης και ένα πιθανολογικό μοντέλο με στόχο ποικιλίες με παρόμοια χαρακτηριστικά ώστε να ταξινομηθούν στην ίδια κλάση. Κάθε αλγόριθμος υλοποιήθηκε λαμβάνοντας υπόψη τους τέσσερις δείκτες βλάστησης και στην συνέχεια τα αποτελέσματα τους συγκρίθηκαν προκειμένου να βρεθούν οι ποικιλίες τις οποίες και οι δύο αλγόριθμοι ταξινόμησαν στην ίδια κλάση. Ο εντοπισμός ποικιλιών με παρόμοια χαρακτηριστικά δίνει χρήσιμες πληροφορίες για την ανάπτυξη και την υγεία του αμπελιού και συμβάλει στο ότι οι γεωργοί μπορούν να τις αντιμετωπίζουν με παρόμοιο τρόπο σε διαδικασίες όπως το πότισμα, η λίπανση ή η συγκομιδή.

1. Introduction

Viticulture requires monitoring and care all year, in order to lead to a good quality harvest. Using precision agriculture methods, a danger or a disease could be detected early enough without any crop destruction (Anastasiou et al., 2019). Vineyard's production has much importance for a lot of countries like France, Italy, Greece, Spain, USA etc., due to economic, social and environmental factors (Jones, 2011). Hence, sustainable management of vineyards could be of high importance for countries and companies that own such areas. Overall, the methods for vegetation monitoring can be separated into two categories:

(a) Laboratory methods need the collection of leaves, fruits, soil, etc., for the detailed analysis of the plant's health status. Such a method is costly, time-consuming, and demands the destruction of a part of the plant.

(b) Non-destructive methods like proximal and remote sensing provide important information for the plant's health without any destruction and with less both in cost and time (Kasimati et al., 2022).

Remote sensing is a method that obtains information about the electromagnetic radiation of objects from distance. The objects reflect the sunlight, which travels to the sensor in different lengths of waveform. Depending on the type of sensor the reflection values can be recorded in a big range of wavelength values, such as the visible (RGB), near-infrared (NIR), shortwave infrared (SWIR), thermal infrared (TIR), or microwave spectrum. Those types of sensors constitute passive remote sensing, which means that they only collect the reflected energy of objects. There is also active remote sensing, such as radar or LIDAR devices, which emit their own signal and then collect its reflection (Wójtowicz et al., 2016). Remote sensing platforms can be either satellites or unmanned aerial systems (UAS), while proximal sensing is cameras located on vehicles or mounted on ground systems.

Satellite images cover large areas and can be acquired easily and, therefore, saving considerable time in the whole procedure; however, there are limits in the final spatial resolution for precision agriculture applications. On the other hand, capturing aerial images via UAV systems is an expensive and difficult method as it needs an expert pilot and specific equipment. Nevertheless, it is possible to acquire both high spatial and spectral resolution (from multispectral and/or hyperspectral cameras) and cover a large enough area by adjusting the flight height (Sassu et al., 2021).

Multispectral and hyperspectral images provide a lot of information about plants' health. Vegetation reflection in the visible spectrum gives information about the pigments of the leaves so differences in plant health are distinct. In the range of NIR spectrum, information about the internal leaf structure is possible to be extracted, and also the lack of an element or the presence of a disease can be detected. These

spectral values depend on the variety of the plant, the growth stage, the quantity of content in the plant's tissue, etc. (Akkara, 2022).

Apart from the selection of the proper image acquisition method, another important factor for the data clustering of vineyard varieties (or on the detection of an element, a disease etc.) is the optimal choice of the classification algorithm. Clustering and machine learning methods have been used in literature for diseases or nutrients detection. Also, vegetation indices can provide significant information for a plant's health. Therefore, by combining methods and different data types for the detection of plants conditions, more accurate results can be carried out.

1.1 Problem description

Grapevine health depends on different physical, chemical and topographic factors. Climate, chemical composition of the soil, existence of weeds, pests or diseases they can all influence the productivity of the viticulture (Giovos et al., 2021). Moreover, different varieties of the same plant in the same area can also present differences in the yield, concerning either the growth phases or the grapes or wine quality. Thus, different growth stages need different monitoring.

Consequently, there is a great need for the differentiation between vines encountered at the same area, and also the correlation between varieties and the emergence of some substances is an important factor that must be detected. The results of this correlation will lead to the creation of a vineyard map. For a farmer, knowing of the specific location of each variety may contribute to the avoidance of misplanting, the assurance of each variety's quality, etc. (Karakizi et al., 2016). Thereupon, the importance of detecting different varieties in viticulture seems to have a leading role in the cultivation and the proper management by the farmers.

1.2 Objective

Concerning the need for an efficient management of different vineyard varieties located in the same area, it is purposeful to implement methods for the classification of varieties with common characteristics. The aim of this master thesis is the detection and analysis of several different vineyard varieties through their spectral properties. Multispectral images apart from the visible (RGB), they also include the Red Edge, and the Near-Infrared areas of the spectrum, and they can provide information about the plant's health and state, conditions that are not always visible with bare eyes. To this purpose, vegetation indices will be used. Each index provides details about the health and the condition of an element in vineyards. Therefore, by combining information from vegetation indices, a method for the efficient clustering of vineyards' varieties with same characteristics will be employed. The scope of this method is to provide information to the farmers about the specific characteristics of grapevine varieties; such information can be valuable for the vineyard management, when they have to make decisions about the planting, pruning, fertilizing and harvesting during its life cycle.

1.3 Research questions

The present work endeavors to analyze the spectral values of different vineyards varieties, to find the common properties between them and to cluster them in groups. The specific questions that this research aims to respond are:

- How so far scientific literature has approached the issue of the detection of different varieties and which tools and methods have been preferred more?

- How can spectral properties of multispectral images be used to detect and analyze various vineyard varieties?

- Which vegetation indices are most effective in providing information about the health and condition of grapevine varieties?

- How can the clustering of vineyard varieties with similar characteristics can be undertaken using a combination of vegetation indices?

- For vineyard's varieties, is it possible different clustering/classification methods to find similar outcomes?

- What is the information provided from an area with a lot of different varieties about the topographic or the physical parameters?

1.4 Research methodology

In the initial phase of this study, a comprehensive literature review was conducted. Specifically, a narrative literature review was employed to explore the conventional diseases affecting vineyards and the crucial nutrient elements necessary for vine growth. The analysis also included an examination of common methods for image acquisition, considering parameters such as focus levels and growth stages. Subsequently, a systematic literature review was undertaken to explore the classification of vegetation varieties. A total of 10 relevant documents were selected, using the PRISMA method, key findings from these documents were extracted and the conclusions of the review were presented.

Following the literature review, a collection of aerial multispectral images was used for the creation of an orthophoto image. The vine segments within the orthophoto were then isolated and retained for further analysis. Subsequently, a clustering method was defined to group vineyard varieties with similar characteristics, which was afterwards implemented on the orthophoto. Figure 1 illustrates the sequential steps followed in the present work.
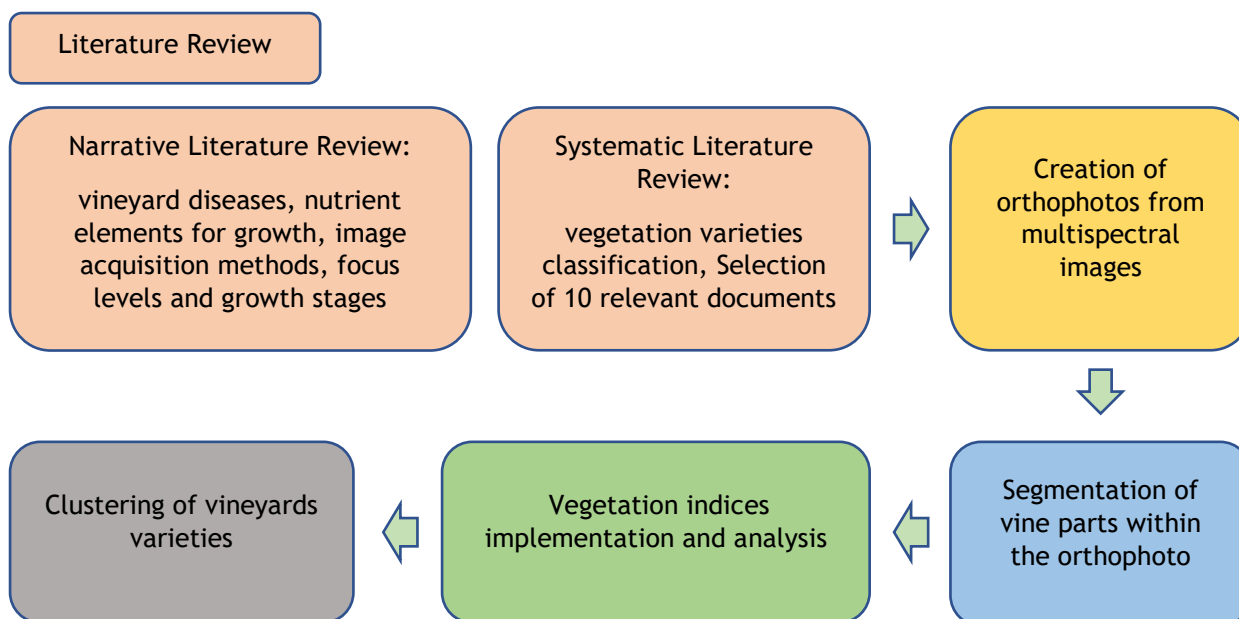
Figure 1: Research methodology diagram.

1.5 Structure

In the second chapter of the master thesis, a narrative literature review is conducted on the vineyard diseases detection methods, vineyard varieties, and classification algorithms. It is important to highlight the frequently used methods for both data acquisition and classification. Subsequently, a systematic literature review was performed, specifically targeting the detection of varieties. The purpose of this review is to identify literature that has not sufficiently examined topics about varieties detection, allowing this master's thesis to provide valuable information for the classification of varieties.

In the third chapter, the conducted experiment is analyzed. Specifically, the characteristics of the study area where the experiment took place are discussed, and the acquired image dataset and its properties are presented. Furthermore, this dataset is processed to generate an orthophotomap, which is then used as the final image for the classification implementation.

The image generated needs to undergo processing to retain only the areas containing vine pixels.

In the fourth chapter, some classification algorithms were implemented to determine the optimal one that provides the best vine's classification. Based on the results of this classification, masks were applied to each variety, enabling them to be treated as distinct polygonal regions. Furthermore, some of the most significant vegetation indices that provide information about the health of the plants were introduced. These indices were applied to the varieties, and a mean value was calculated for each of them. Next, the correlation between the executed indices

was analyzed, and the least correlated indices were selected for the classification of varieties. Continuing, using two classifiers, the varieties were categorized into three different groups based on the mean values obtained from the index implementation. These results were analyzed, and several observations were made.

The fifth chapter consists of the discussion of the findings, the conclusions drawn as well as the contribution of the research along with some proposals for future research.

2. Literature Review

Vineyards are faced with various challenges throughout their lifecycle. Farmers need to monitor several common issues during the growth period of vineyards, including vineyard productivity (Ballesteros et al., 2020; J. Tang et al., 2016),(Arab et al., 2021; Maimaitiyiming et al., 2019)), plant water stress ((Loggenberg et al., 2018; Pôças et al., 2020; Z. Tang et al., 2022)), and nutrient concentrations, such as nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg), and Boron (B) (Chancia et al., 2021; Moghimi et al., 2020; X. Peng et al., 2022). Nutrients have a decisive role in the vineyards growth as nitrogen is important for the productivity of the plant, phosphorus participates in the development of reproductive parts of the plant, while potassium contributes on the movement of substances in the plant. Many research documents focus on the detection of specific diseases, such as **Grapevine vein-clearing virus** (Nguyen et al., 2021), **Mildew disease** (Chen et al., 2020; Hernández et al., 2021; X. Peng et al., 2022; Z. Tang et al., 2022), **Flavescense Dorée grapevine disease** (Silva et al., 2021), and **Esca disease** ((Alessandrini et al., 2021; Kirti et al., 2021)). Another important factor that affects productivity and the monitoring procedure of a disease growth is the vineyard's variety; with more than 10,000 varieties of grapes at a global level, varieties, similar to the detection of diseases, can be identified more easily under specific conditions.

An important factor that facilitates the detection of specific conditions in varieties is the focus level; depending on the specific characteristics to be examined, the focus can be directed towards either a particular part of the plant or encompass the entire field. The capture distance is adjusted accordingly. The most commonly used focus levels include the total vineyard area, canopy, berries, or leaves.

The focus level plays a crucial role in the detection of varieties or diseases; the closer the distance from the plants the earlier specific plant characteristics can be detected. According to Ballesteros et al. (2020) , aerial images, unlike close-up captures, provide the opportunity to obtain not only spectral information but also geometric characteristics of the canopy. Other distinctions between close and long shots depend on factors, such as the size of the area of interest, the availability of collected data, and the timing of data acquisition.

2.1 Narrative Literature Review

2.1.1 Growth stages

An important factor for the vines' health monitoring and production is the growth stage during the data collection. For the determination of the grapevine growth stage, three different systems have been developed until today (COOMBE, 1995). The first one created by Baggiolini (1952), the second one by Eichhorn and Lorenz (1977) and the third one called the BBCH system was developed as a model for the European Union and adapted for the grapevine by Lorenz et al. (1994). The Baggiolini

system consists of 10 stages, but in the newest version another six more stages have been added; thus, resulting in a total of 16 stages (A-P). The Eichhorn and Lorenz (E-L) system includes 22 different stages, whereas the BBCH system consists of 10 macro-stages. Table 1 presents the modified E-L system along with the descriptive name of each of the stages.

Growth stage	Details
Budburst	4 Green tip; first leaf tissue visible
Shoots 10 cm	12 5 leaves separated; shoots about 10 cm long, inflorescence clear
Flowering begins	19 About 16 leaves separated: beginning of flowering (first flower caps loosening)
Full bloom	23 17-20 leaves separated; 50% caps off (= full bloom)
Setting	27 Setting; young berries enlarging (>2 mm diam.), bunch at right angles to stem
Berries pea size	31 Berries pea-size (7 mm diam.)
Veraison	35 Berries begin to colour and enlarge
Harvest	38 Berries harvest-ripe

Table 1: Details about vineyards' growth stages

In most studies, the growth stage of vineyards is considered in the detection of varieties or diseases. In the study of J. Tang et al. (2016), aerial and proximal images were collected during the shoot and veraison stages. The research indicated that for better accuracy, proximal images should be taken earlier in the growth stage. The study of Peng et al. (2022) used aerial images to collect data at various growth stages, including new shoot growth, flowering, fruit expansion, veraison, and maturity. It was found that the quantity of nitrogen gradually decreases as the growth period progresses. Specifically, nitrogen content is higher at the new shoot growth stage compared to the fruit expansion stage, veraison and maturity stage, and nearly zero at the flowering stage. The highest uncertainties for potassium were observed during the fruit expansion stage, and for phosphorus it was during the veraison and maturity stages.

Two articles compared data collected from a UAV (Unmanned Aerial Vehicle) and spectrometer with Sentinel-2 data. The first article (Kasimati et al., 2022) claims that UAV and Spectrosense with GNSS (Global Navigation Satelite Systems) provide better results during the mid-late season with full canopy growth, specifically during the pea-sized berries and veraison growth stages. However, the Sentinel-2 images were less reliable in terms of grape quality. The second paper (Kasimati et al., 2021) argues that UAV + Spectrosense + GNSS provide better results in identifying soluble solids in wine grapes, followed by the CropCircle sensor, and then the Sentinel-2 imagery. The results from the UAV and Spectrosense + GNSS data were obtained during the pea-sized berries and veraison growth stages. In addition to Sentinel-2

images, (Arab et al., 2021) also used images from Landsat-8 during different growth periods: Bud break, Flowering, Fruit Set, Max-Canopy Expansion, Ripening, and Harvest. The results showed that the correlation between grape yield and NDVI (Normalized Difference Vegetation Index), LAI (Leaf Area Index), and NDWI (Normalized Difference Water Index) indices was very low during flowering and harvest periods but high during maximum canopy expansion.

Based on the above, the growth stage at which a vine disease or phenomenon is visible using images or spectroscopy depends on the specific phenomenon being investigated.

2.1.2 Data acquisition methods

In literature, several studies have deployed a UAV platform for acquiring multispectral images. The spatial resolution of the final image depends on the flight height. For instance, in Ballesteros et al., (2020), the resulting image from the multispectral Sequoia camera (Parrot, Paris, France) had a ground resolution of 7 cm at a flight height of 80m, whereas at a flight height of 30m the spatial resolution was 4cm (Romero et al., 2018). Similarly, in Mazzia et al., (2020) and (Kasimati et al., 2021) using the Parrot Sequoia camera at flight heights of 35m and 30m, the ground resolution was 5cm and 3cm, respectively. Padua et al. (2020) deployed the Micro-MCA camera and achieved a spatial resolution of 1.6cm at a flight height of 30m. Z. Tang et al., (2022), used the MicaSense RedEdge camera at a height of 120m above the ground resulting in a final image resolution of 8cm.

The aforementioned methods were implemented in wide areas of vineyards. Therefore, aerial multispectral photography is useful for phenomena occurring across broader fields rather than focusing on individual vines. These research studies were used for productivity measurement, water status or stress detection, and for assessing inclusiveness of nutrients, such as nitrogen, phosphorus, potassium, soluble solids, and more.

Some documents used hyperspectral images for studying specific characteristics of vineyards. Loggenberg et al., (2018) and Nguyen et al., (2021) implemented terrestrial methods for hyperspectral data collection; Silva et al. (2021) and Hernández et al. (2021) took images under laboratory conditions. Only Chancia et al. (2021) integrated a hypespectral camera on a UAS platform. Capturing specific conditions depends on the focus level of the object. In contrast to the predominantly wide-area usage of multispectral images, hyperspectral images were employed to capture vines at the canopy or leaf level, with only one of them addressing a broader variety area.

In addition to multispectral and hyperspectral photography, RGB cameras are commonly used as well. For instance, Chancia et al. (2021) deployed a GoPro camera mounted on a vehicle, Fuentes et al. (2021) used an iPhone11 attached to a selfie stick, Kerkech et al. (2020) and Zhou et al. (2021) employed a UAV and the MAPIR

Survey2 and Parrot Sequoia+ agricultural camera, respectively. Alessandrini et al. (2021) manually captured images using two smartphones and a tablet, while Miranda et al., (2022) utilized a DALSA Genie NanoC2590 camera, and Hernández et al. (2021) conducted image acquisition using the Canon EOS 5D.

Some studies employed alternative methods for image acquisition, such as spectrometer (Maimaitiyiming et al., 2019; Pôças et al., 2020), thermal imagery (Fuentes et al., 2019; Reyes Rojas et al., 2021), satellite images (Sentinel-2, Landsat-8) (Arab et al., 2021; Kisekka et al., 2022), DNA testing (Ampatzidis et al., 2020) or a pre-existing dataset (Cruz et al., 2019; Kaur et al., 2022). These methods can provide satisfactory results in detecting disease or substances in plants.

2.1.3 Data Classification Algorithms

Data classification algorithms have the ability to consider a range of variables related to vineyards, and assess the relation between these variables in near real time, using entire images or segments with common characteristics (Volpi et al., 2021). Chen et al., (2020) suggest that the evaluation of grapevine diseases can be enhanced by incorporating parameters such as weather conditions, location, and the age of the plant. Additionally, other studies implemented weather conditions, such as rainfall, temperature and other climatic variables to develop predictive models for disease detection in vineyards (Rossi & Caffi, 2007; Rouzet & Jacquin, 2003).

Machine learning methods encompass different approaches. Supervised methods involve training and testing datasets, while unsupervised methods employ clustering techniques to analyse data (Dale et al., 2018). In literature, numerous studies have used one or more classifiers or machine learning algorithms to detect diseases or identify natural ingredients. Among the 47 examined documents, the most commonly employed algorithms are Random Forest, Support Vector Machine, Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and various regressors. The following diagram illustrates the distribution of algorithms used for disease detection across the 47 articles.
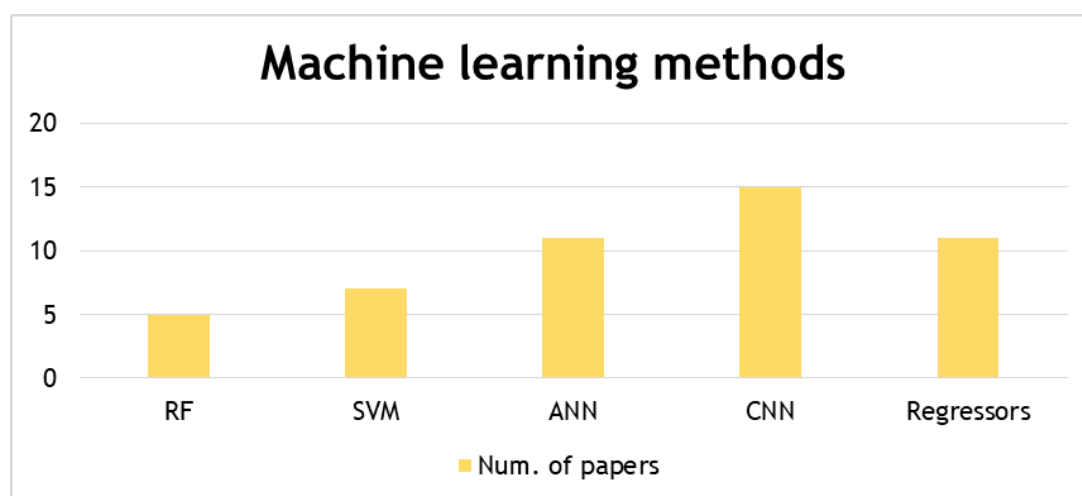
Figure 2: Machine Learning Methods used more in literature

The key points drawn from these articles are the following:

o Many of the studies implemented pre-existed Convolutional Neural Network (CNN) models for classification purposes. The most frequently used models are AlexNet (5 times), SegNet (1 time), LeNet (2 times), ResNet (employed 5 times in various editions, such as ResNet-34, ResNet-50, hybrid ResNet-Jaya model, ResNet-V2, and ResNet-101).

o Some of the algorithms were developed by the authors themselves and results were compared with pre-existing models or methods, like RandomForest, XGBoost, SVM, k-NN and others.

o A common evaluation metric in these publications is model accuracy. Among various methods, ResNet algorithms consistently exhibited higher accuracy compared to other approaches.

o When it comes to disease detection, algorithms have generally achieved a satisfactory level of accuracy in recognizing different types of leaf diseases or distinguishing between healthy and affected leaves.

2.1.4 Vegetation Indices

In addition to machine learning methods, various vegetation indices have been deployed to classify crops components or diseases. These indices are used to identify key characteristics of vineyards, such as productivity, irritation, water status, diseases, chemical elements, or other conditions.

Regarding productivity, J. Tang et al., (2016) emphasized the significance of NDVI-blue, while (Ballesteros et al., 2020) found that the NDVIwiv (well-illuminated vegetation) index yielded important results. (Maimaitiyiming et al., 2019) focused on the water index, whereas (Arab et al., 2021) considered the NDVI and LAI as the most important indices.

For irritation detection, (Ohana-Levi et al., 2019) identified as crucial the CWSI (Crop Water Stress Index). (Romero et al., 2018) suggested the OSAVI index for water stress detection, while Z. Tang et al., (2022) found the NDRE and GRVI indices to be significant.

In the detection of Grapevine vein-clearing virus (GVCV), the FRI1 (physiology indices), WSCT (physiology indices), and AntGitelson (pigment index) were determined as the most significant indices by Nguyen et al., (2021).

For the detection of pests caused by Lycorma delicatula, ARVI, OSAVI, and GNDVI were found to be the most effective indices Zby Zhou et al., (2021).

Finally, Moghimi et al., (2020) selected the NDRE index for the detection of chemical elements such as N, P, K, Ca, Mg, and B.

Detailed documentation of the indices used in this thesis can be found in the sub-section 4.2.

2.2 Vineyard varieties detection

A systematic literature review was conducted to gather and analyze publications related to the detection of varieties in plants. There are two types of literature reviews: narrative and systematic. The former describes and discusses scientific topics from a theoretical standpoint without explaining the methodology and evaluation of the research (Rother, 2007). The latter is a methodological approach that aims to provide an up-to-date summary of an issue (Higgins JPT, 2022). Initially, research questions and keywords are defined, and specific criteria are selected to narrow down the search to a specific area of interest; subsequently, the findings are presented and evaluated (Kitchenham et al., 2009).

A way to visualize the process of a systematic literature review (SLR) is by using the PRISMA diagram, which includes a 27-item checklist and a four-phase flow diagram, ensuring a transparent SLR reporting(Liberati et al., 2009).

As a result, firstly a search was conducted in the Scopus library using specific keywords related to the research topic of vineyard varieties detection. The keywords used were: ("varieties detection" OR "varieties classification") AND ("agriculture" OR "crops" OR "cultivars" OR "vegetation") AND ("UAV" OR "drone" OR "satellite"). The initial search provided a total of 79 documents. Then, the search was refined by limiting the document type to articles and setting the language to English, resulting in 51 documents.

The abstracts of these 51 documents were read, and 12 documents were selected as the most relevant based on their abstracts. Subsequently, the full texts of these selected documents were read and, after thorough evaluation, 8 documents were chosen to meet the requirements of the research.

In addition to these 8 documents, two more relevant documents related to the research topic were included as supporting literature to enhance the evaluation of the topic (Figure 3).
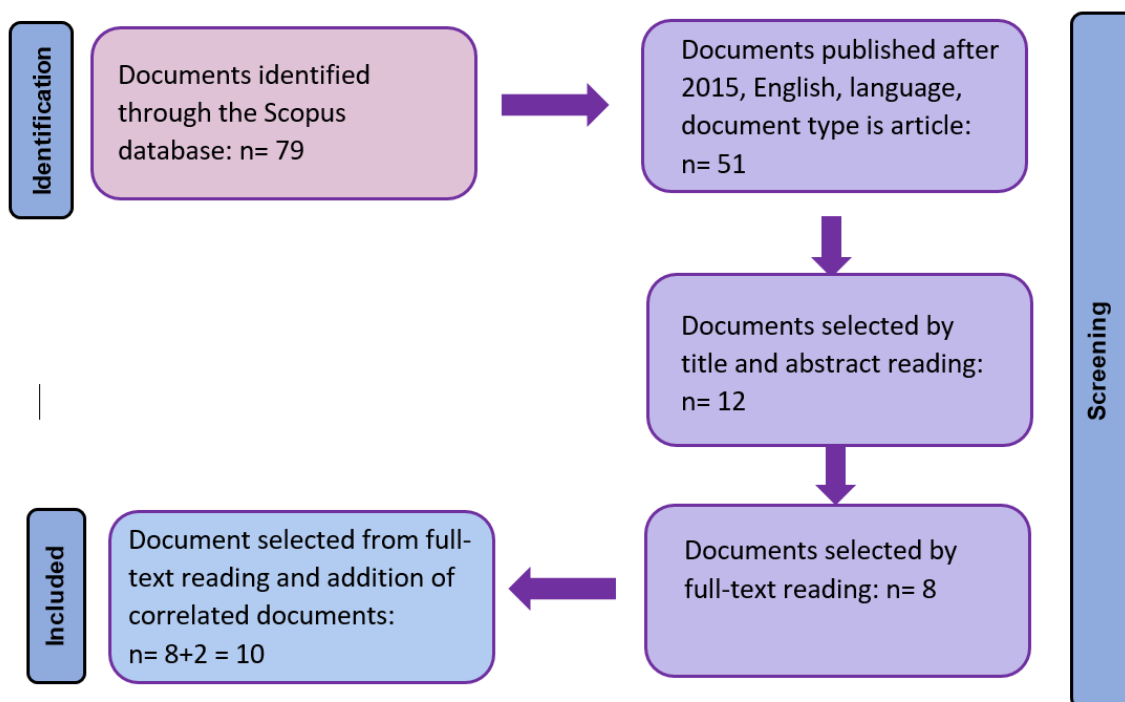
Figure 3: PRISMA diagram steps from the systematic literature review

2.2.1 Selected documents analysis

The selected documents can be divided into four different groups, depending on the distance level at which the image datasets for variety detection were collected. The first group pertains to laboratory photography conditions, where the examination of the variety was conducted at leaf level. This means that leaves are isolated from the plants, photos are captured under laboratory conditions, and then classification is performed on these photos. The second group consists of photos taken at ground level in the field. The third group involves aerial photos captured from a UAV platform, and the fourth group deals with satellite images. Figure 4 illustrates the number of documents corresponding to each group along with Table 2.
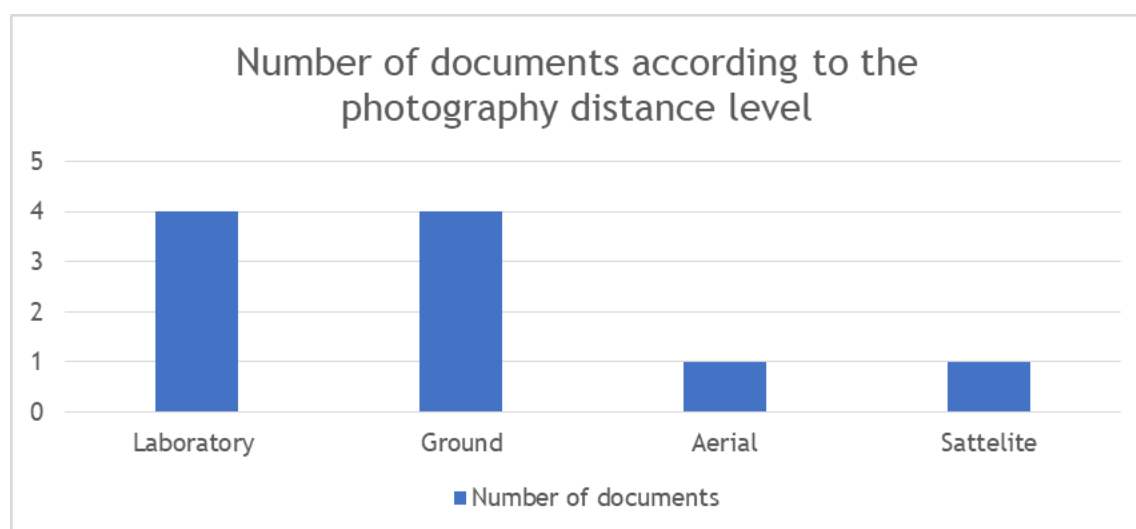
Figure 4: Number of studied documents about distance level image capturing

ID	Crop	Num. of varieties detected	Focus Level	Acquisition method	Spectral range	Best Classifier	References
1	Grapevines	5	Ground	ASD FieldSpec®3 spectroradiometer	350-975, 976–1770, 1771–2500 nm	SVM	(Mirzaei et al., 2019)
2	Grapevines	4	Ground	Portable Spectroradiometer (FieldSpec3, Analytical Spectral Devices, Boulder, CO, USA)	between 350 nm and 2500 nm	SVM and discriminator analysis	(Al-Saddik et al., 2019)
3	Lettuce	3	Ground	RGB camera	RGB	YOLO-VOLO-LS	(Zhang & Li, 2022)
4	Peanuts	4	Laboratory	Image-l "spectral image" series hyperspectral machine	R, G and B were set to 638.7, 551.58 and 442.95 nm	MF-LightGBM-SEL	(Wu et al., 2022)
5	Grapevines	6	Ground	Open dataset of RGB images	RGB	fusion of fc6 (in AlexNet network) and Fc1000 (in ResNet50 network)	(Y. Peng et al., 2021)
6	Cassava	47	Laboratory	Cellular phone or a digital camera	RGB	ANN algorithm	(Unajan & Gerardo, 2019)
7	Grapevines	3	Laboratory	RGB images	RGB	Linear Discriminant Analysis	(Marques et al., 2019)
8	Grapevines	7	Laboratory	OPPO Android phone RGB camera	RGB	Decision Tree (DT)	(Garcia et al., 2022)
9	Maize	25	Aerial	Parrot Sequoia multispectral camera	Green (0.53-0.57 m), Red (0.64–0.68 m), Rededge (0.73-0.74 m) and Near-Infrared (0.77–0.81 m)	Vegetation Indices	(Chivasa et al., 2020)
10	Grapevines	4	Satellite	World View 2 pan-sharpening images (multispectral data)	RGB, NIR, coastal, yellow, red-edge, NIR	SVM	(Karakizi et al., 2016)

Table 2: Documents examined within the systematic literature review process

Each document follows a specific approach for varieties detection based on its objectives.

In the document of (Mirzaei et al., 2019), hyperspectral data was utilized for the detection of vineyard varieties. Before the classification, they attempted to reduce

the dataset dimensions by selecting only the important bands for the final procedure. To achieve this, they implemented the PLSR and ANOVA-PCA methods; the latter was finally chosen due to its higher accuracy results. Subsequently, the SVM and LDA classifiers were compared with the SVM, demonstrating higher accuracy and, thus, being retained for further analysis. They also concluded that the optimal wavelengths and indices for discriminating vineyard varieties from each other are the 695, 752, 1148, 1606 nm and 582, 687, 1154, 1927 nm and R680, WI, SGB and RATIO975_2, DattA, Greenness at leaf and canopy levels, respectively.

Al-Saddik et al., (2019) had as objective to detect the Flavescence dorée disease across four different vineyard varieties. The successive projection algorithm (SPA) and various vegetation indices (VI) were employed for disease detection. These methods were compared against SVM and discriminator analysis (DA), and it was concluded that the SPA method achieved the most accurate disease classification. The experiment was implemented on both healthy and diseased leaves and 12 of the most common vegetation indices were tested as they are related with disease appearance on vegetation.

In the study made by Zhang & Li, (2022), the goal was the detection of five lettuce varieties at seven different growth stages. It was observed that the earlier the growth stage, the more challenging it became to detect varieties or diseases. Initially, the VOLO-D1 classifier was used for varieties detection, but difficultiew were encountered in identifying varieties during the early growth stages. To address that issue, a new method called the YOLO-VOLO-LS classifier was proposed, capable of detecting varieties at an earlier stage. The results of the classification achieve high correlation: 95.961, 93.452, 96.059, 96.014, 96.039 in Val-acc, Test-acc, Recall, Precision, F1-score, respectively Only RGB images were utilized for variety detection, which resulted in limited information compared to multispectral or hyperspectral images.

Wu et al., (2022) attempted the classification of peanut seeds into four different varieties. Unlike others, this research focused on classifying the seeds rather than the plant leaves, and although the method differed, the procedure for variety detection followed the same principles. Hyperspectral sensors were deployed for data collection, providing a wide spectral range to identify differences between seeds. A machine learning algorithm was utilized for classification, which was later compared with four other algorithms. The results indicated that the proposed algorithm, using hyperspectral data, achieved the best performance.

In the document of Y. Peng et al., (2021), CNN and SVM algorithms were applied to an open RGB dataset of ground vineyard image. The dataset insists of five grapevine varieties. These algorithms relied on deep feature detection of different varieties by focusing on various levels of fused data. This approach differed from other methods that primarily focused on the spectral values of the image pixels. The fusion

of fc6 (in AlexNet network) and Fc1000 (in ResNet50 network) deep features obtained the best identification performance.

Unajan & Gerardo (2019) , aimed to detect varieties in images in a manner similar to D3 and D5. Since RGB images alone are not suitable for vegetation detection based only on pixel values, the focus shifted to object detection; 15 leaf features were used to detect 47 varieties, which were detected with an 85.11% accuracy by implementing the ANN algorithm.

In the study of Marques et al., (2019), a method was developed to detect grapevine varieties based on color and shape features. RGB images were utilized, along with the Linear Discriminant Analysis method for image classification, achieving an 87% accuracy for the entire dataset of images.

Garcia et al. (2022) , also used RGB images and 26 features to classify 1,149 images into seven different grapevine varieties. Three supervised algorithms were implemented for classification, with the decision trees algorithm making the best results, achieving an accuracy of 89%.

In the document of Chivasa et al. (2020), multispectral images captured from a UAV, were used to detect 25 maize varieties. Vegetation indices were also employed to calculate the ground MSV infection at three different growth stages and they were classified into resistant, moderately resistant, and susceptible, respectively, with an overall classification accuracy of 77.3%. Six vegetation indices implemented on three different growth stages. Then some variables were tested with GNDVI, CIgreen, CIrededge, and the red band seemed the most important for the classification.

Finally, in the study made by Karakizi et al. (2016), World View 2 satellite images were employed to detect three, four, five, and six different grapevine varieties in four different regions. The method used for variety detection was object-based classification.

2.2.2 Systematic literature review key summaries

Completing the systematic literature review on crop variety detection, it is important to highlight certain conclusions that contribute to the rest of the master thesis.

• Firstly, it is crucial to note that the more spectral values we can acquire from an image dataset, the more information can be extracted about vegetation health, varieties, or specific features. By using only RGB images, the methods applied for the acquisition of effective results are often limited to feature or object detection. These methods typically involve convolution, artificial neural networks, and machine learning algorithms. Multispectral or hyperspectral images have an

advantage over RGB images because they provide more detailed information by analyzing the spectral values of vegetation in the near-infrared or short-wave infrared spectrum. By combining vegetation indices or fusing them with other types of information, even more detailed information can be obtained.

• Secondly, the growth stage during data collection is an important factor for detecting diseases, infections, and varieties. The documents of (Zhang & Li, 2022) and Marques et al. (2019) detected lettuce and grapevine varieties at three and two growth stages, respectively. These documents concluded that the earlier the growth stage, the less information can be extracted from the plant's leaves. It may be valuable to implement variety detection at different growth stages to draw conclusions about the optimal period for variety detection.

• Moreover, when dealing with hyperspectral data, dimension reduction is essential for efficient data utilization. Hyperspectral sensors often capture data across more than 200 spectral bands, but not all bands necessarily provide useful information about plant health or varieties. Algorithms, such as PCA, ANOVA, or PLSR, as used in the document of Mirzaei et al. (2019) , offering the opportunity to decrease the analysis time and extract only the optimal information from the datasets.

• Another crucial aspect is the utility of vegetation indices in providing valuable insights about plants. These indices play a significant role in procedures aimed at detecting phenomena in vegetation areas; for expample Chivasa et al., (2020) tested commonly used vegetation indices, such as NDVI, GNDVI, NDVIrededge, SR, CIgreen, and CIrededge.

• Last but not least, the capturing distance and camera's spectral information are factors that can be chosen according to the available time and the area of interest. Laboratory conditions may offer more detailed results, but they are time and cost consuming, especially for frequent data collection or large areas. On the other hand, satellite images are easily accessible, available every few days, and they can provide coverage of different areas simultaneously. Their spatial resolution, however, is lower compared to aerial or ground images. Aerial images combine the advantages of ground and satellite images, making them a suitable solution for frequent data collection in a specific area with good both spatial and spectral resolution.

3. Material and methods

3.1 Study area

The wider study area of this research is Lykovrisi, in the northern part of Athens Metropolitan Area (AMA) within the Attica region (Figure 5). More specifically, the field work in this research took place at the Institute of Technology of Agricultural Products located in the southwestern part of Lykovrisi. This area is home to over 1,000 different vineyard varieties, making it an intriguing location for implementing the methods described in this thesis. The abundance of diverse varieties presents the opportunity to evaluate various characteristics of the vineyard plants. Figure 5 displays the Institute's location within the Attica region.

(a)

(b)
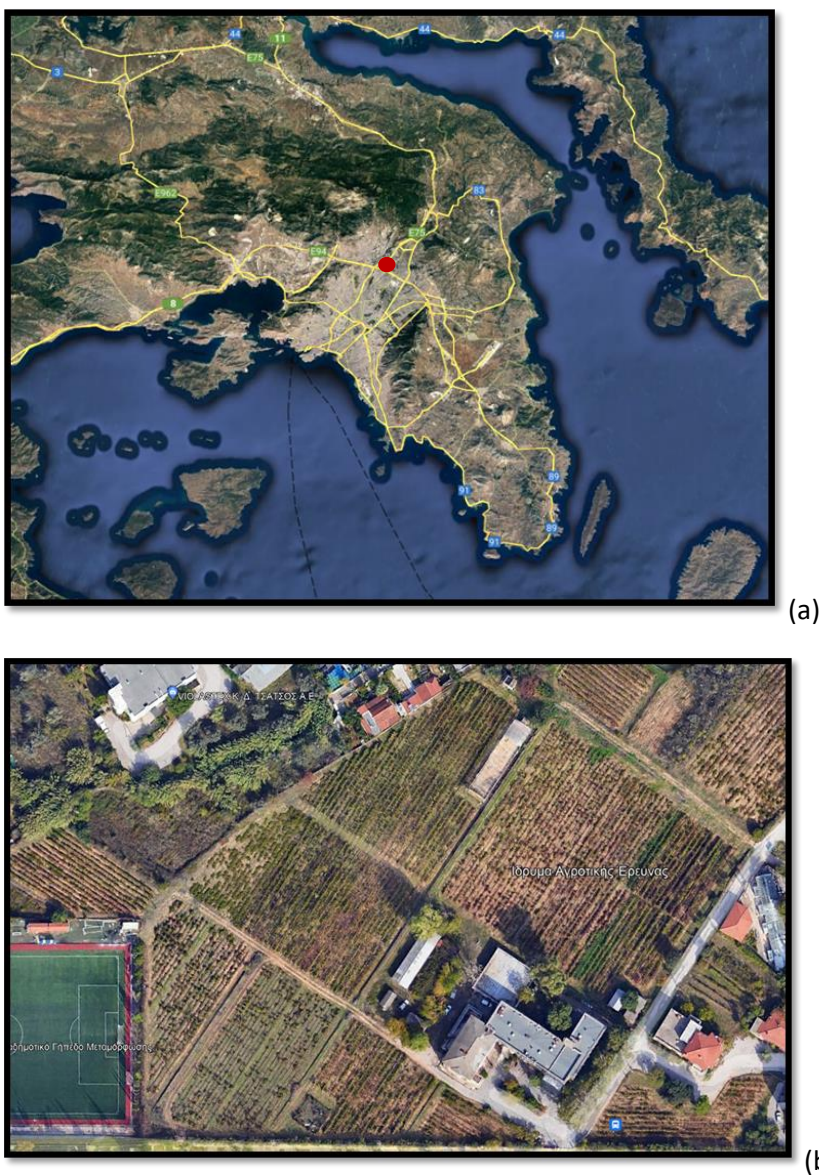
Figure 5: Location of the study area:(a): location of the Institute of Technology of Agricultural Products, (b) its vineyards)

3.2 Equipment

The vineyard area was captured using the multispectral camera Micasense RedEdge-M, which consists of 5 bands (Blue: 475 nm, Green: 550nm, Red: 668nm, RedEdge: 717nm, Near Infrared: 840nm).
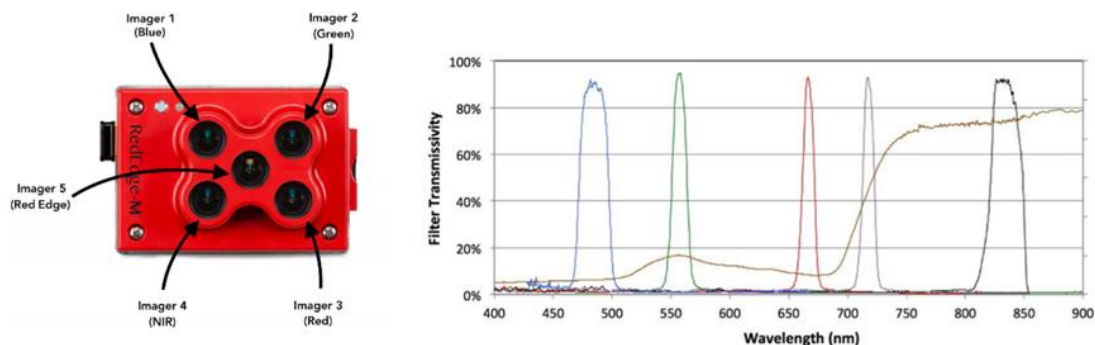
Figure 6: Micasense RedEdge-M multispectral camera

With the camera providing bands in five different spectral values, a wealth of information about the vineyard's condition can be extracted from the images. For instance, vegetation typically exhibits lower values in the blue and red wavelength while higher values are observed in the green and near-infrared spectrum. By combining the values of near infrared spectrum with the red spectrum, differences in vegetation and other land covers, as well as within the same type of vegetation, can be observed.

The images were collected by the laboratory of Remote Sensing of the National and Technical University of Athens. Furthermore, in the given dataset reflectance panel images were not included, and as a result the images used in the methods implementation are not radiometrically calibrated.

The UAV dataset consists of 53 images, with each one containing information from five different spectral bands (Figure 7).
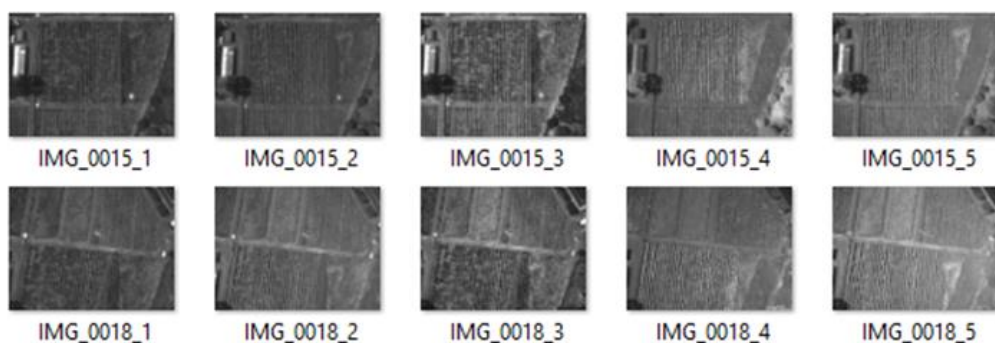
Figure 7: A sample of images captured from the multispectral camera.

3.3 Orthophotomap creation

The production of the orthophotomap was carried out using the commercial software Metashape (Agisoft, 2015). This software uses the algorithms of Structure from Motion (SfM), Multi View Stereo (MVS), and SIFT (Scale-Invariant Feature Transform) (Lowe, 1999), which work with multiple overlapping images. These algorithms identify common points among the images and calculate the camera's position at the capturing time. This process generates a sparse point cloud. The next step involves creating a dense point cloud to reconstruct a 3D model using images with known positions in space. Consequently, the surface of the study area and the Digital Elevation Model (DEM) are generated. Finally, this procedure leads to the creation of the orthophotomap for the area of interest.

The method used for the orthophotomap creation is illustrated in Figure 8.
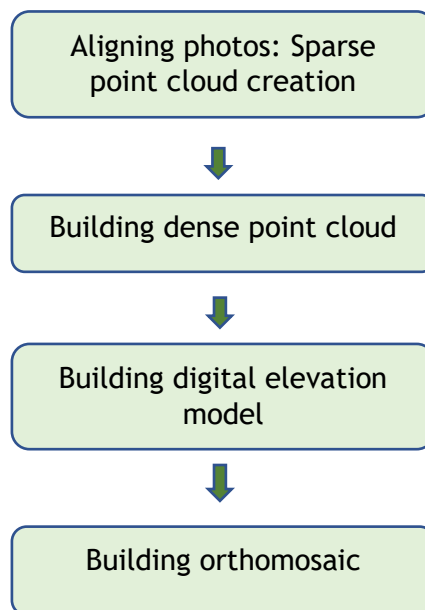
Figure 8: Diagram of steps for the orthophoto creation

3.3.1 Sparse point cloud creation

Firstly, the sparse point cloud was created by aligning common points among the overlapping images with high accuracy. "High accuracy" indicates that the algorithm makes use of the entire image resolution to identify these common points.
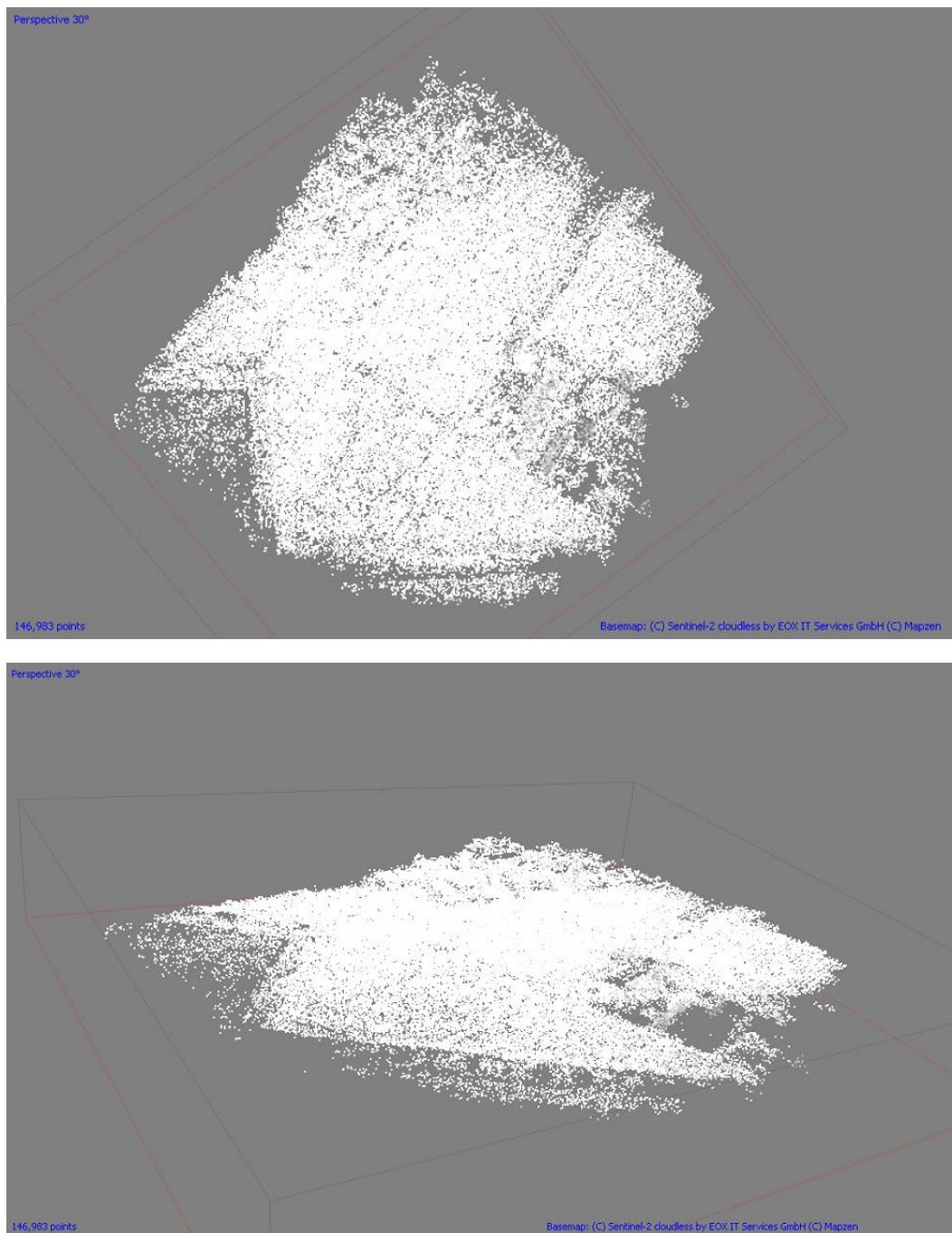
Figure 9: Sparse point cloud of the area of interest

3.3.2 Building dense point cloud

After the generation of the sparse point cloud, additional computational algorithms are applied to densely reconstruct the points of the study area. In this process, the software analyzes the images and calculates the depth information for each pixel by comparing it with neighboring pixels. By using the known camera positions and orientations from the sparse point cloud, the software estimates the 3D coordinates for the rest points of the source images. This results in a much denser point cloud, whose reconstruction algorithm uses various techniques to interpolate the missing depth information and define the point cloud. The goal is to densely capture the geometry of the scene, integrating as many points as possible. The following representations illustrate the dense point cloud that was created based on the sparse point cloud.
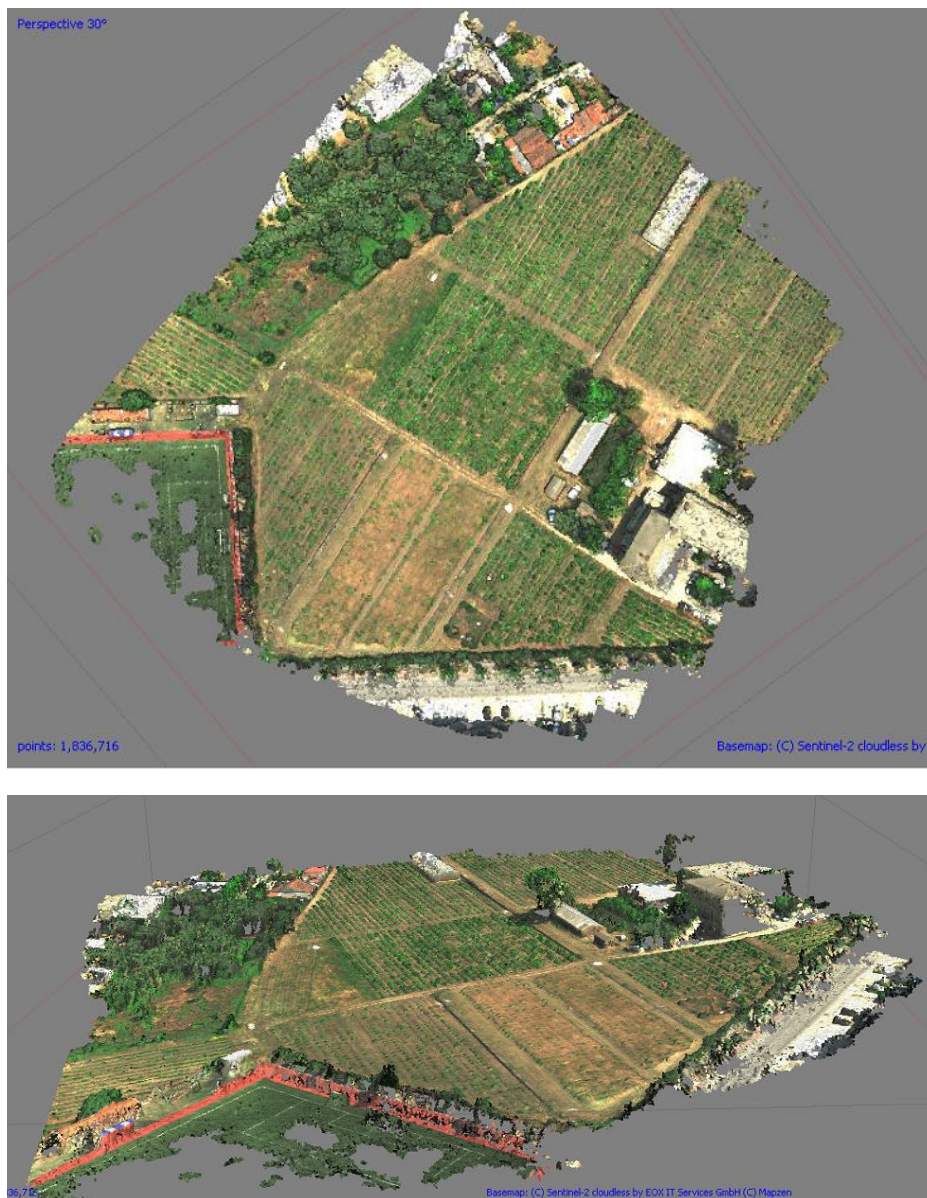
Figure 10: Dense point cloud of the area of interest

3.3.3 Building Digital Elevation Model

The Digital Elevation Model (DEM) was then created from the dense point cloud, including a grid created based on the dense point cloud; on this grid each point represents an elevation value (Figure 11)..
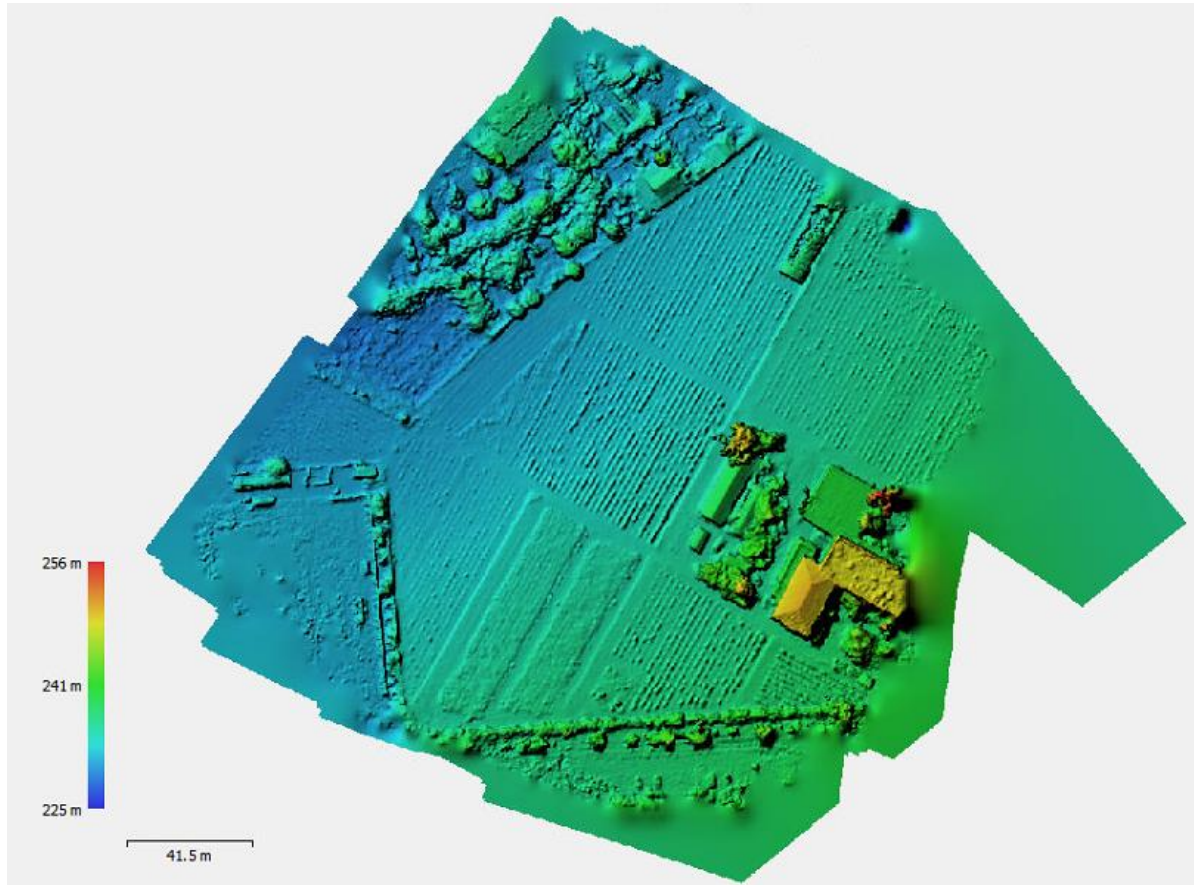
Figure 11: Digital Elevation Model of the area of interest

3.3.4 Building the orthomosaic

An orthomosaic is a high-resolution and orthorectified image created based on source photos. It is an accurate representation of the Earth's surface with minimum distortions and can be used for various purposes, such as measurements, surveying mapping and more. In order to generate an orthomosaic, the source images and a surface model are required. In the current project, the surface model used is the DEM; alternatively, a mesh surface can also be used for the orthomosaic creation. Figure 12 represents the final orthomosaic, while Figure 13 shows the georeferenced image on a map.
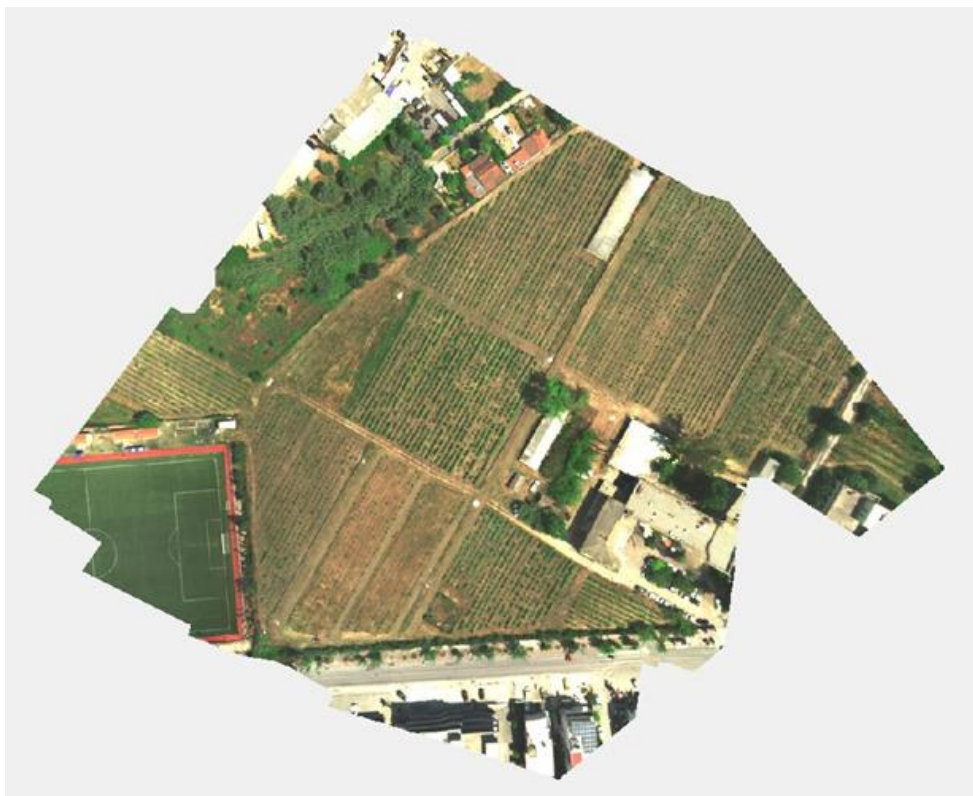
Figure 12: Orthomosaic of the area of interest.

Figure 13: Georeferenced orthophoto of the area of interest.

3.4 Area of Interest: Selection and Analysis

A small part was selected from the entire captured area for further analysis with the final area of interest (AOI) being approximately 3000 m^2.

Figure 14: Selected part of vineyard for analysis

3.4.1 Information about selected AOI

Within the selected AOI, a total of 113 different vineyard varieties are located, with each variety consisting of 2 to 10 vineyard plants, meaning that the number of vines changes among the different varieties. Consequently, certain varieties offer more information regarding spectral values on the plant canopy due to having more plants compared to others. As a result, the accuracy of detecting different varieties may also fluctuate throughout the AOI.
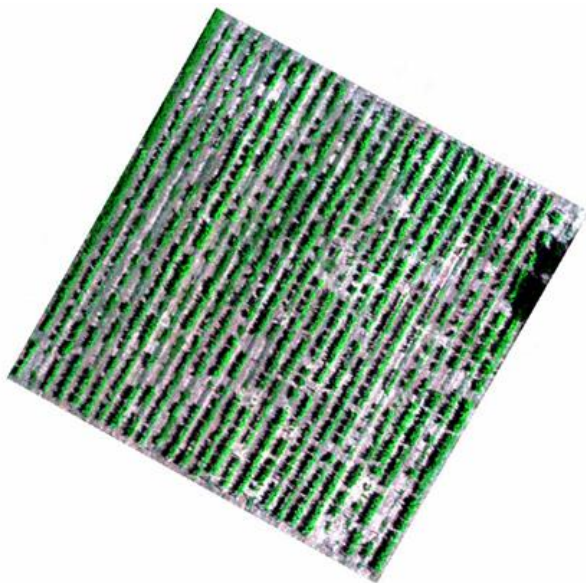
Figure 15: Area of interest

3.4.2 Spectral profile pixel analysis

The image of the AOI contains information from five different spectral bands. In order to extract the maximum from these bands, it is necessary to combine them and identify the areas where each band contributes the most valuable data. By doing so, an analysis about each band with its mean value and standard deviation is initially calculated (Figure 16 and Table 1).
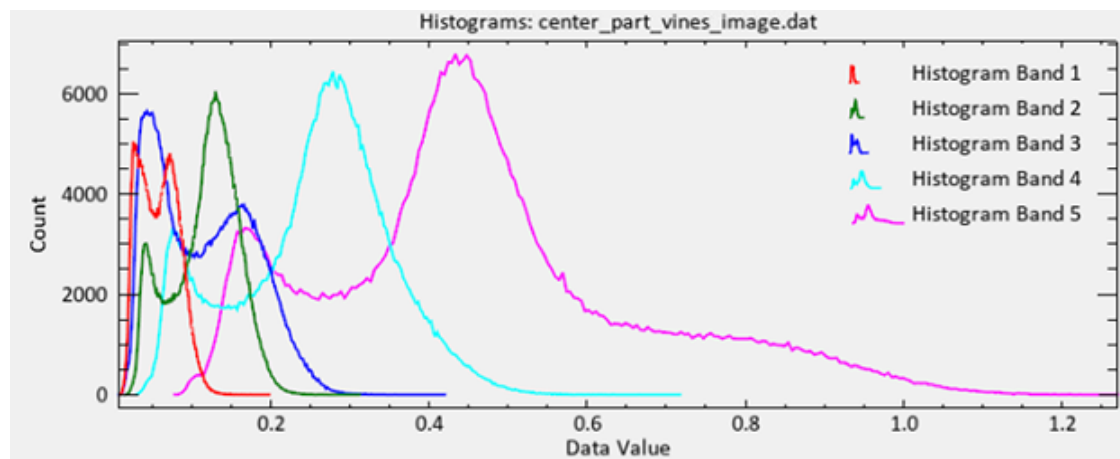
Figure 16: Spectral values of each band of area of interest

Bands	Mean value	Std Dev
Band 1	0.06	0.02
Band 2	0.12	0.04
Band 3	0.12	0.06
Band 4	0.26	0.10
Band 5	0.47	0.2

Table 3: Mean values and Standard deviation of each band

The values of each band were separated into five parts. Moreover, an equalization on values was applied for better visualization results. As each band has a different mean value, it is distinct to provide also different information about images' objects. By focusing on each band it can be observed that the higher values in band 1 and 3 depicts ground pixels, while the higher values in bands 4 and 5 are for vegetation pixels and band 2 values presents both ground and vegetation pixels (Figure 17-21).
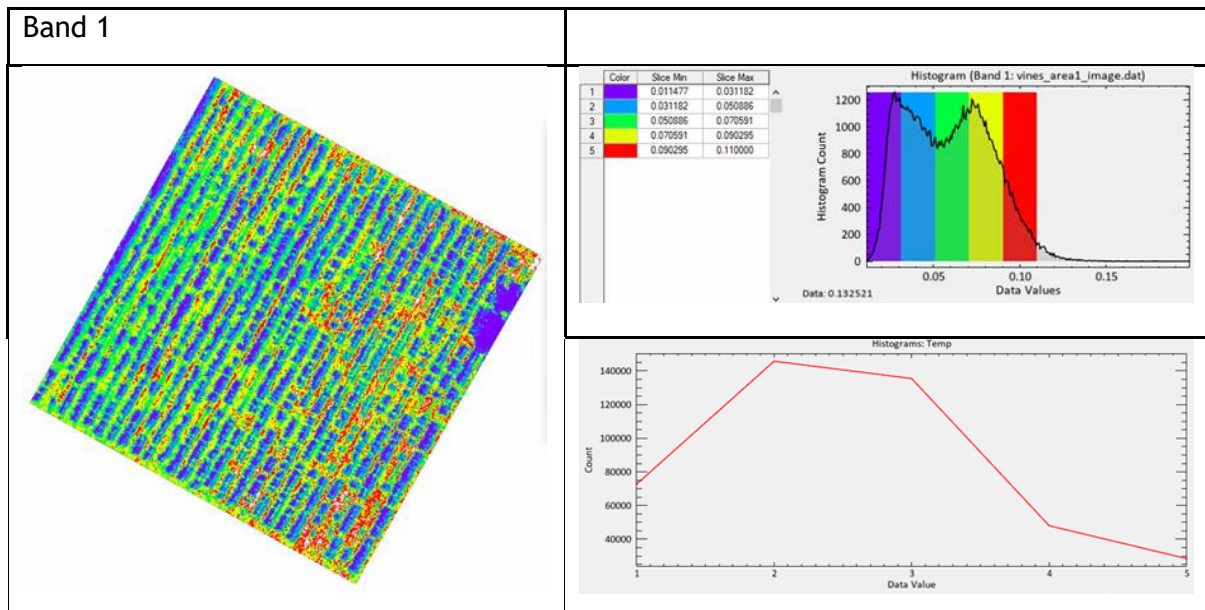
Figure 17: Spectral values of Band 1
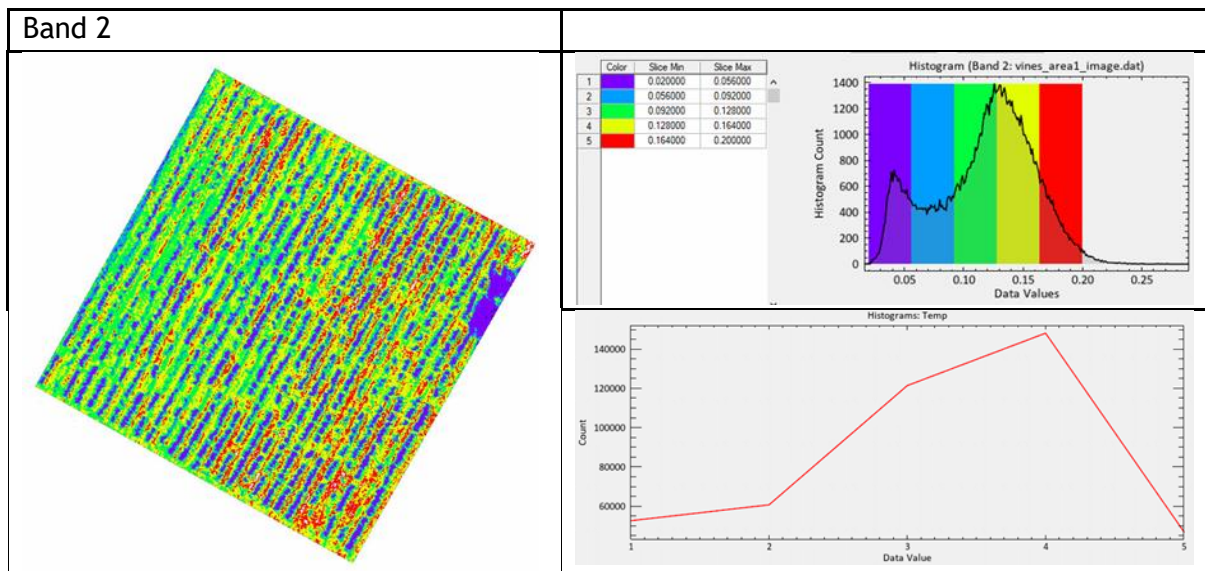
Figure 18: Spectral values of Band 2
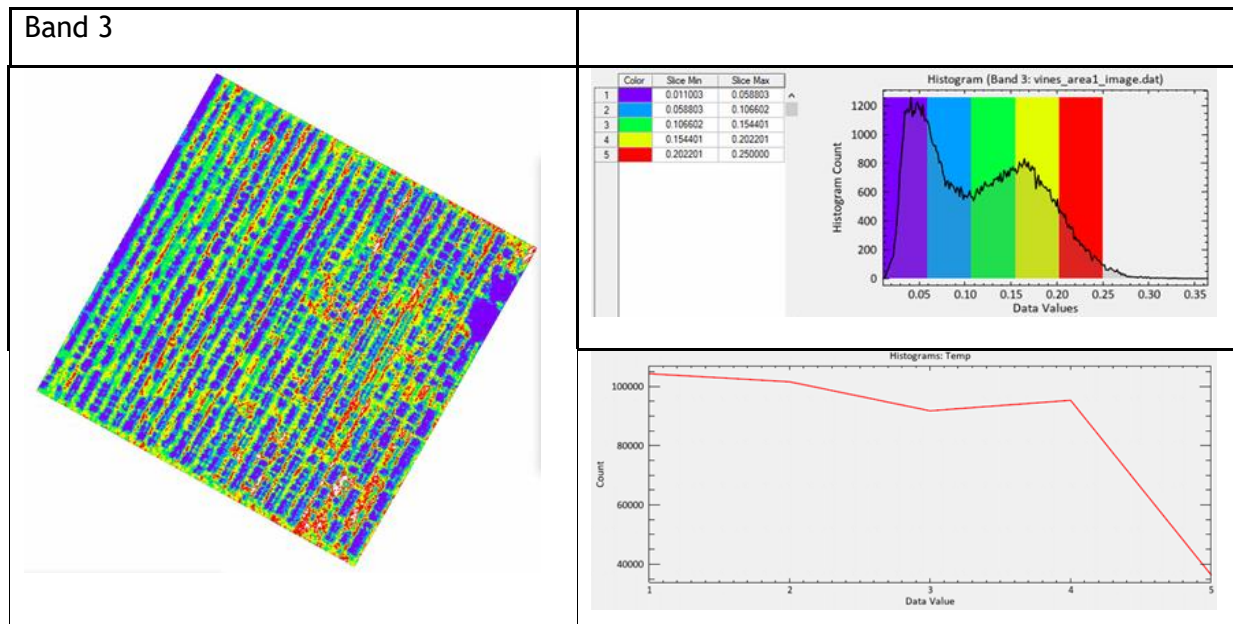
Figure 19: Spectral values of Band 3
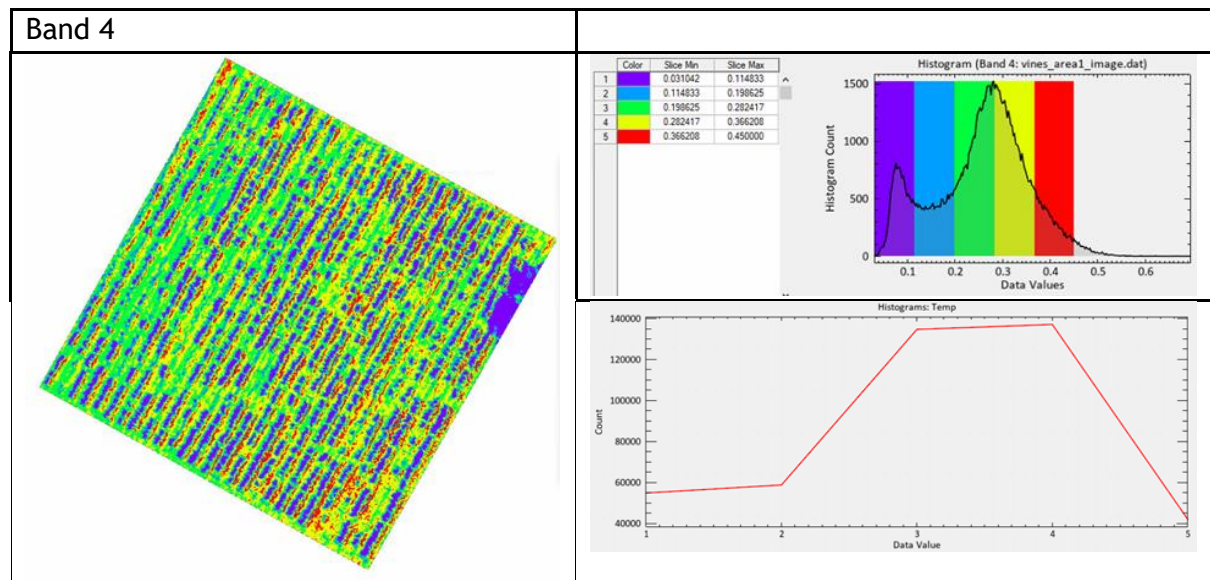
Figure 20: Spectral values of Band 4
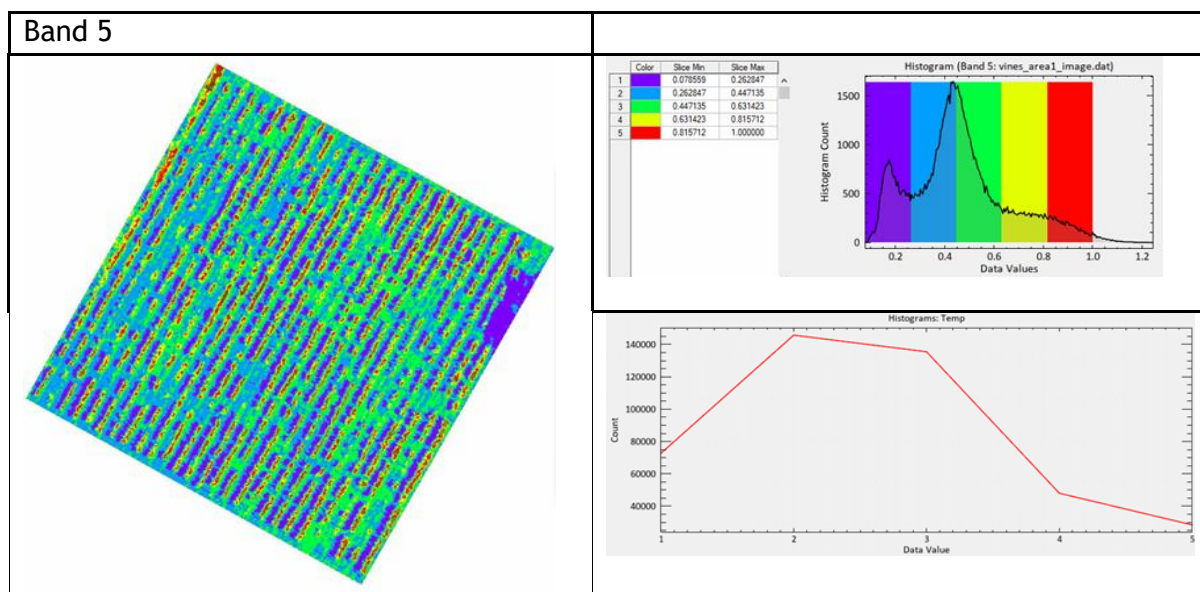
Figure 21: Spectral values of Band 5

3.5 Field measurements of varieties

The purpose of conducting field measurements for variety identification is to accurately detect the exact locations where each variety begins and ends. To map these locations of the varieties, a GNSS GINTEC-M20 receiver was deployed (Figure 22), which is capable of tracking and processing signals from all active satellite systems, including GPS, Glonass, Galileo, and BeiDou. This receiver is capable of achieving horizontal position accuracy of less than 2 cm.

Figure 22: GINTEC-M20 GNSS receiver

In the vineyard fields, labels for each variety are positioned at the starting point of each variety; the locations of these labels were surveyed using the GNSS receiver and in the GGRS87 coordinate reference system, which is the National Greek Coordinate system.

Measuring the precise position of each variety is crucial because it ensures that the ground truth of the variety's dataset is significantly more accurate.

Figure 23. Labels for the position of different vine varieties.

The image below shows the measured points of each distinct variety represented by red dots. It is evident that the first rows contain about 5 varieties each, while the number of different varieties decreases on the last rows.
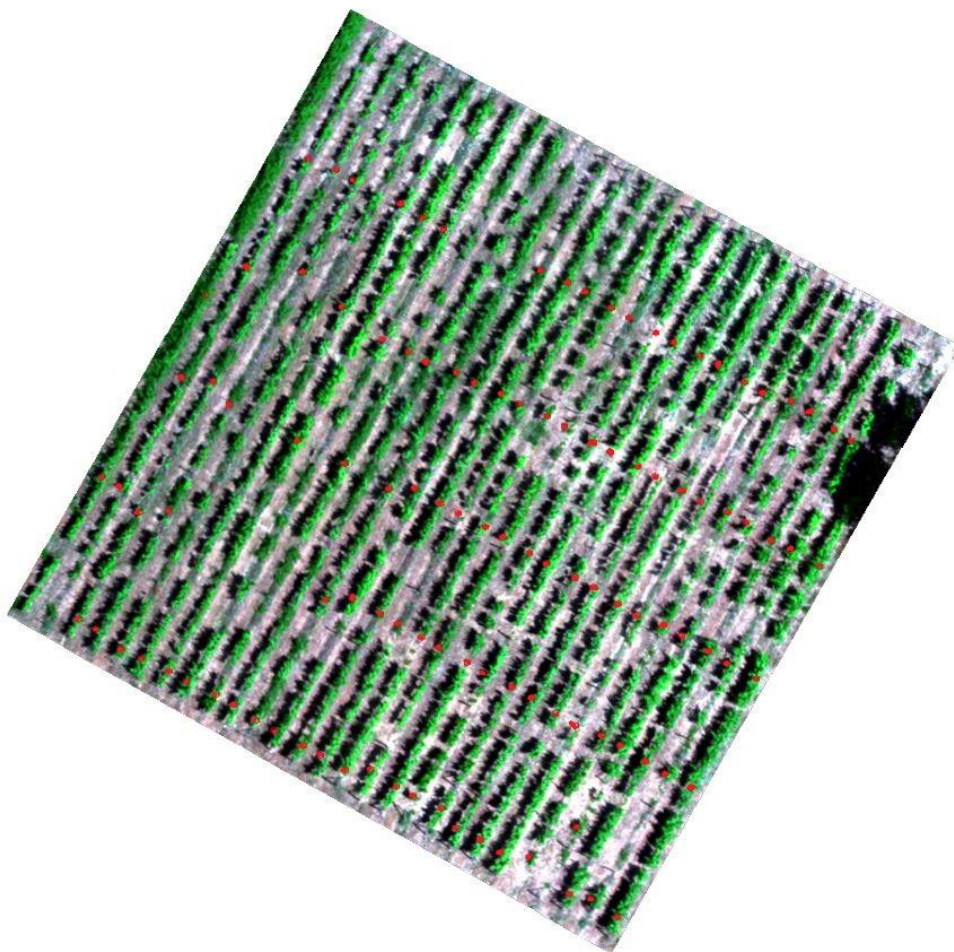
Figure 24: Position of the beginning and the ending of each vine variety.

4. Results

4.1 Mask creation of vineyard region

If to achieve optimal classification between vine, ground, and shadow pixels, aiming to isolate the vine pixels, it is requested to use a mask in the specific area. This process is crucial to focus only on examining the vine canopy area in the subsequent steps. However, apart from vine, ground, and shadow areas, there also weeds; distinguishing this type of vegetation as a separate class poses more challenges, as it consists of low vegetation that may also include ground pixels. Therefore, during the classification process, weed pixels will be categorized within the ground class.

The definition of training and test samples, along with the implementation of various classification algorithms, was conducted. The objective is to determine the most effective classifier for these three regions of interest. Four classifiers were executed, including three supervised algorithms (Maximum Likelihood, Minimum Distance, and Mahalanobis Distance) and one unsupervised algorithm (Iso Data).

4.1.1 Definition of training and testing samples

For the training and testing of these algorithms, 35 samples were used for each region of interest; specifically, 28 samples for training (Figure 25a), while the remaining 7 for algorithm testing (Figure 25b).

(a)   (b)
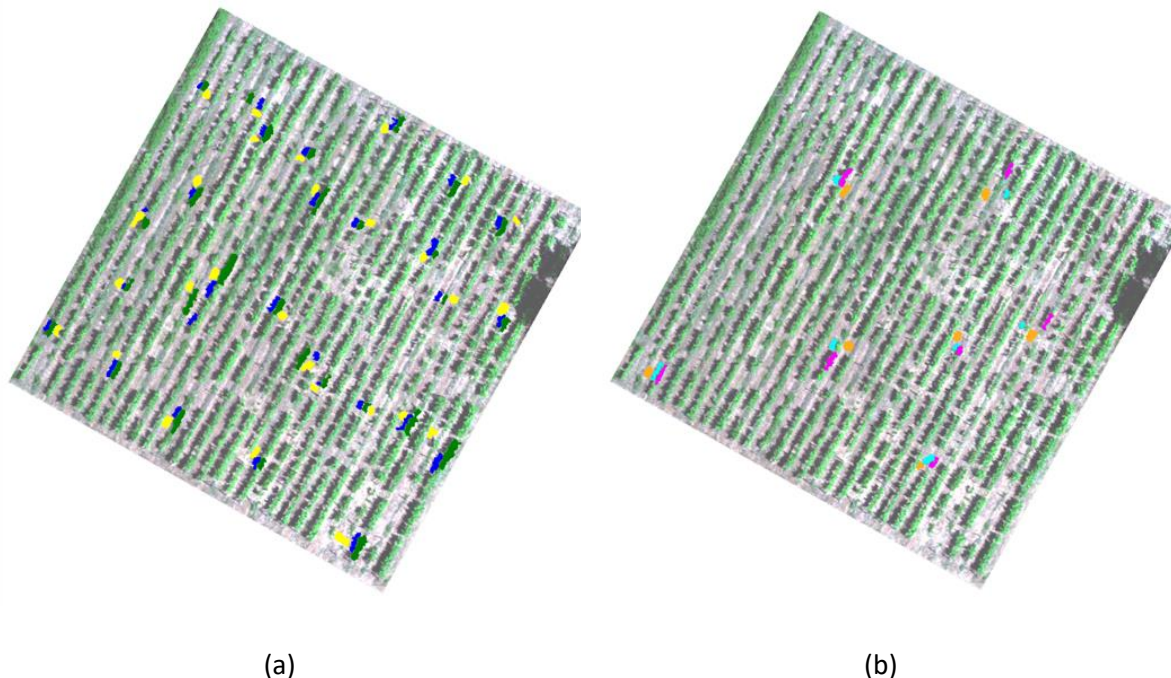
Figure 25: (a) Training and (b) testing samples for the classification

4.1.2 Classification of vineyard, ground and shadow pixels

After defining the training and testing samples, the algorithms were executed, and the results are presented in Table 4 and Figure 26 and Figure 27. Specifically, the following table and diagram illustrate the percentage of area occupied by each class in the classified image for each classification algorithm.

Training	Supervised classification %			Unsupervised classification
	Maximum likelihood	Minimum Distance	Mahalanobis Distance	Iso data
Shadow	21.81	24.14	28.12	25.03
Ground	46.57	46.91	53.7	55.21
Vines	31.62	22.15	25.03	19.75

Table 4:Results of supervised classification
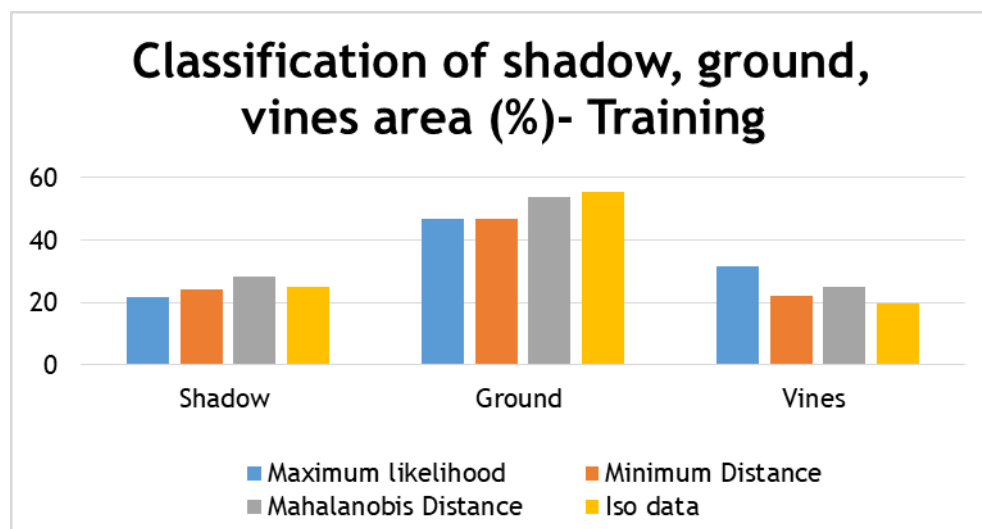
Figure 26: Classification results of each algorithm
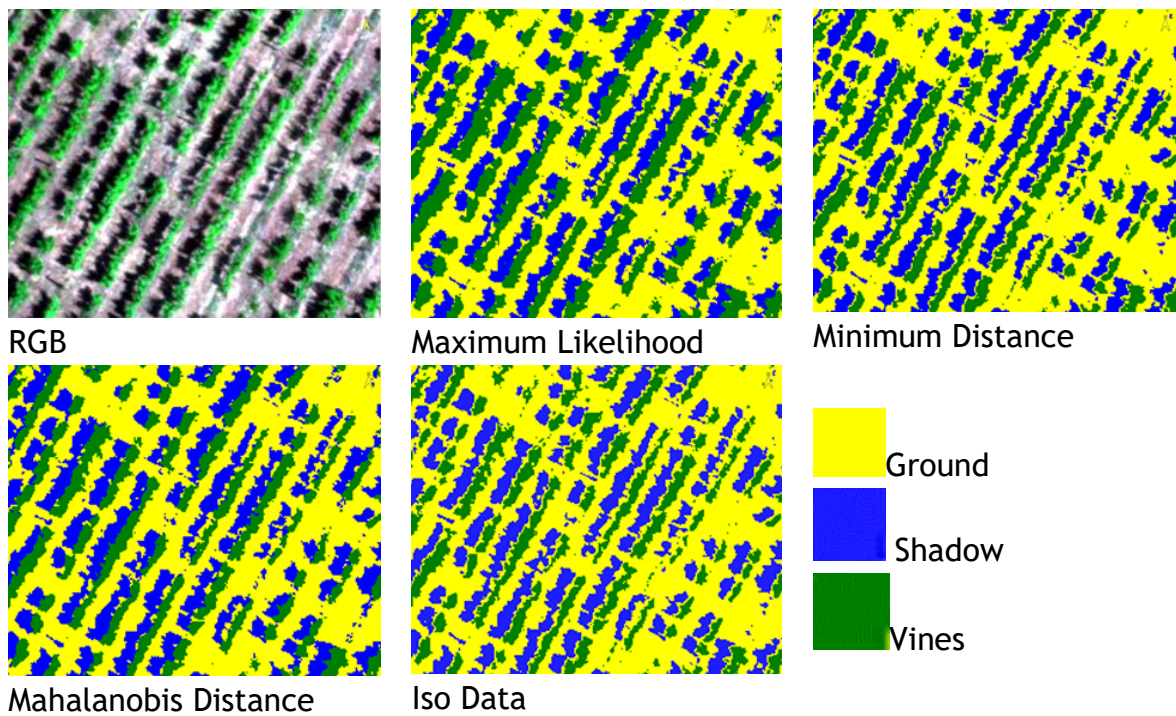
Figure 27: Classification results on a part of the image.

4.1.3 Classifiers' evaluation

For the classifiers' evaluation, confusion matrices were created for both training and testing sets (Table 5)

Maximum Likelihood

Class	test_shadow	test_ground	test_vines	Total
Shadow	**93.97**	0.66	0.91	24.13
Ground	5.69	**97.35**	0.30	38.59
Vines	0.33	1.98	**98.79**	37.28
Total	100	100	100	100

Minimum Distance

Class	test_shadow	test_ground	test_vines	Total
Shadow	**94.31**	0.51	3.11	24.97
Ground	5.69	**97.28**	11.30	42.62
Vines	0	2.20	**85.60**	32.41
Total	100	100	100	100

Mahalanobis Distance

Class	test_shadow	test_ground	test_vines	Total
Shadow	**94.08**	0.66	5.69	25.92
Ground	5.92	**98.46**	1.14	39.37
Vines	0	0.88	**93.18**	34.70
Total	100	100	100	100

Iso Data

Class	test_shadow	test_ground	test_vines	Total
Shadow	**95.65**	0.66	3.18	25.39
Ground	4.35	**97.65**	16.60	44.38
Vines	0	1.69	**80.21**	30.23
Total	100	100	100	100

Table 5: Confusion matrices between training and testing sets

Table 6 and Figure 28 presents the classification results, indicating that the Maximum likelihood algorithm provides the higher percentage of correct classified vineyard pixels.

Testing	Supervised classification			Unsupervised classification
	Maximum likelihood	Minimum Distance	Mahalanobis Distance	Iso Data
Shadow	93.97	94.31	94.08	**95.65**
Ground	97.35	97.28	**98.46**	97.65
Vines	**98.79**	85.60	93.18	80.21

Table 6: Classification results
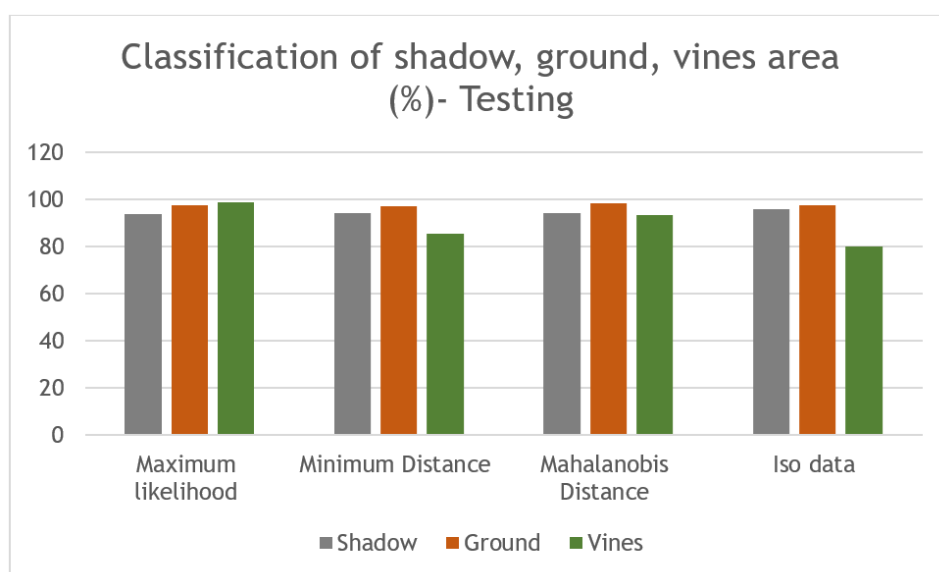
Figure 28: Diagram of classification results

As evident from the Table 6 above, the Iso Data classifier achieved higher accuracy in classifying shadow pixels, the Mahalanobis distance classifier performed better in classifying ground pixels, while the Maximum Likelihood algorithm demonstrated higher accuracy in classifying vine pixels. Since only the vine pixels are required for the subsequent steps, the results that are further analyzed are those obtained from the Maximum Likelihood classification and a mask will be implemented to exclude ground and shadow pixels.

4.1.4 Image noise removal and mask implementation

The classified image that was generated includes pixels that do not belong to the objects in the image; these pixels are considered noise. The objective of this step is to isolate noisy pixels from the image and exclude them from the final result; this procedure is known as "sieve classes". Specifically, the user selects the window size (pixel connectivity=8) for scanning the image, and objects within a class that contain less than 15 pixels are isolated. Consequently, these pixels are labeled as "unclassified".

Notably, the vine class was isolated from the resulting image, and a mask was created (Figure 29) with the vine mask is presented before and after applying the sieve classes procedure; The vine area is reduced from 31.67% of the entire image to 31.14%.
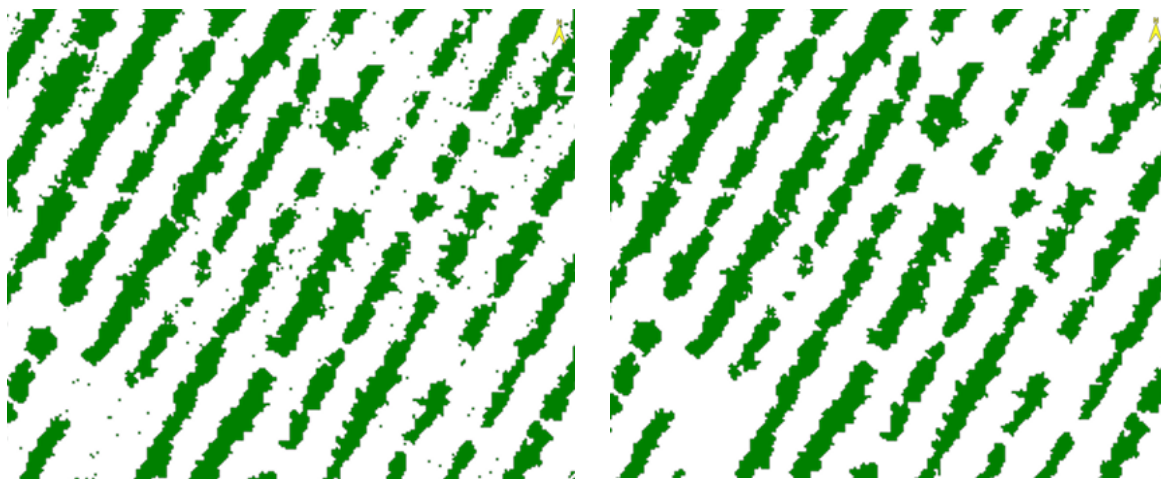
Figure 29: Noise removal results

Taking into account that the focus is only on the vines, it was intentional to apply masks to the ground and shadow areas in the image; thus, the remaining pixels, corresponding to the vines, are considered the regions of interest and were retained, with the final image comprising vine pixels (Figure 30).
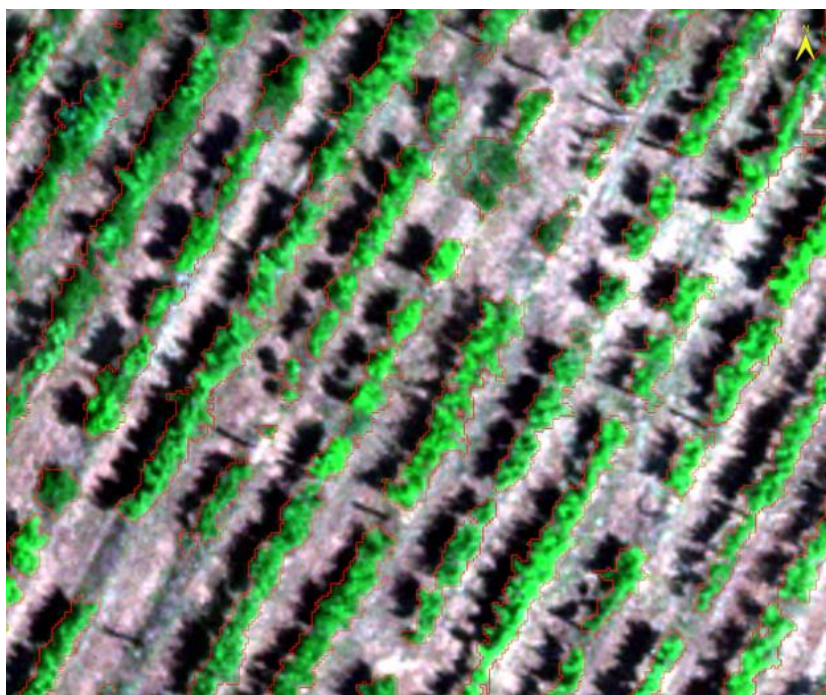
Figure 30: Mask of vines

4.2 Vegetation indices analysis

Since the final region of interest was identified on the image, seven vegetation indices were calculated. The purpose of this step is to identify indices that have lower correlation with each other; this approach will help in clustering varieties with the purpose to detect varieties with common needs or even to identify which of them are on the same growth stage. For instance, if certain varieties are growing at a faster rate than others, their fertilizer or water needs may differ, accordingly. Moreover, variations in ground characteristics across different locations within the same vineyard field can also impact the growth requirements of the vines. The selected indices for implementation and analysis on the area of interest are presented shortly in Table 7.

Index name	Formula	Reference
Chlorophyll Index Green (Clgr)	CLGR= (NIR/GREEN)-1	(Gitelson et al., 2005)
Chlorophyll Index - Red-Edge (Clre)	CLRE= [(NIR/RedEdge)-1]	(Gitelson et al., 2005)
CVI (Chlorophyll Vegetation Index)	CVI= (NIR/GREEN) *(RED/GREEN)	(Vincini et al., 2008)
NDVI (Normalized difference vegetation index)	NDVI= (B5-B3)/(B3+B5)	(Rouse JW, 1974)
GNDVI (Green Normalized difference vegetation index):	GNDVI= (NIR-GREEN)/(NIR+GREEN)	(Gitelson et al., 1996)
EVI 2 (Enhanced vegetation index)	EVI 2= 2.5*(B5-B3)/(B5+2.4*B3+1)	(Clevers, 1999; Kaufman, 1997)
RVI (Ratio Vegetation Index)	RVI= NIR/RED	(Jordan, 1969)

Table 7: Vegetation indices used

4.2.1 Information about selected vegetation indices

(Gitelson et al., 2005) developed two vegetation indices: Chlorophyll Index Green and Chlorophyll Index – Red-Edge (Clre) in order to estimate the chlorophyll content in crops. They conducted experiments in soybean and maize crops but they also proposed the implementation of these indices for other crops. In their study they examined parameters such as LAI, Chl, canopy architecture and leaf structure to estimating these indices. (Gitelson et al., 2005) also proposed the green chlorophyll index (CIgreen), which is a technique based on linear regression of the model against the total chlorophyll content in the canopy; they found a high correlation between CIgreen and canopy chlorophyll content in maize and soybean crops. The CIgreen index was proposed as (NIR/GREEN)-1, where NIR and GREEN represents the reflectance for the Near Infrared and Green bands, respectively.

(Jiang, 2008) considered the findings of (Clevers, 1999)regarding the high correlation among visible bands in agriculture fields, and the research of (Kaufman, 1997) *and* (Karnieli, 2001) who stated that "under clear sky conditions, the SWIR spectral bands are highly correlated with the visible (blue, green and red) spectral bands over various land covers". Based on these insights, Jiang (2008), made the assumption that "visible bands are highly related to each other over agriculture fields". Consequently, they formulated the EVI2 index equation. That index was tested worldwide on 40 sites with diverse ground and crops type (agriculture, grassland, forests) and diverse weather conditions and the implementation of the tests lasted for six years (2000-2005).

(Vincini et al., 2008) proposed the leaf chlorophyll index for canopy-scale in order to estimate the chlorophyll concentration in crops leaves. They conducted experiments to evaluate the effectiveness of this index on sugar beet canopies and they recommended to use this index during periods when crops have open canopies.

(Gitelson et al., 1996) attempted to use the green band in the chlorophyll calculation for crops and they observed that the green index is more sensitive to chlorophyll concentration compared to the red band.

The leaf pigment of plants provides considerable information about the plant's health status. Reflectance values of plants, obtained through specific measurements, it is used in order to measure contents, such as chlorophyll, carotenoids, and other nutrients (Gitelson et al., (2006). Chlorophyll has the property to absorb the light energy and contributes to the photosynthetic process, while carotenoids protect the photosynthetic systems from damage.

4.2.2 Ground truth of vineyards varieties

Next objective is the vine's attribute analysis that will facilitate drawing conclusions about the differences or similarities between various vine varieties. The area of interest consists of 26 vine rows, approximately at 54 m each. Each row consists of 1 to 5 different varieties and, therefore, 112 different varieties are located in the entire study area. Figure 31 illustrates the ground truth of varieties in the area of interest. A, B, C, D and E are the code of each different variety that a row of vines may include. If in a row there are 5 different varieties, the code for each variety in the specific row will be for example 2A, 2B, 2C, 2D and 2E. If a row of vines has only 2 different varieties the codes for that row will be 18A and 18B. Each row depicts each variety in different colour. It is important to note that within the area of interest there are not present the same variety more than once.
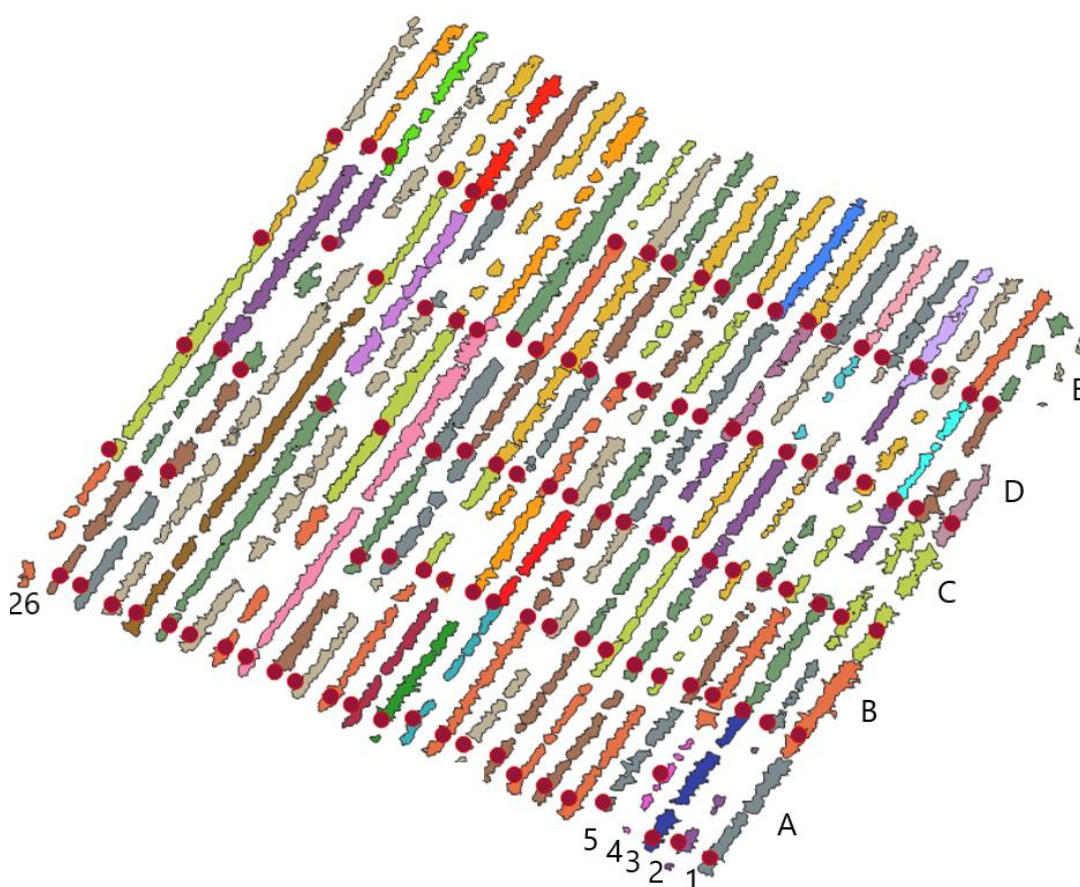
Figure 31: Ground truth of varieties' location

4.2.3 Vegetation indices implementation

To carry out the analysis, the seven different vegetation indices mentioned above were applied to the image; for each variety the average reflectance value for each index was calculated and presented in the following sections of 4.2.3.

4.2.3.1 Chlorophyll Index Green (Clgr)
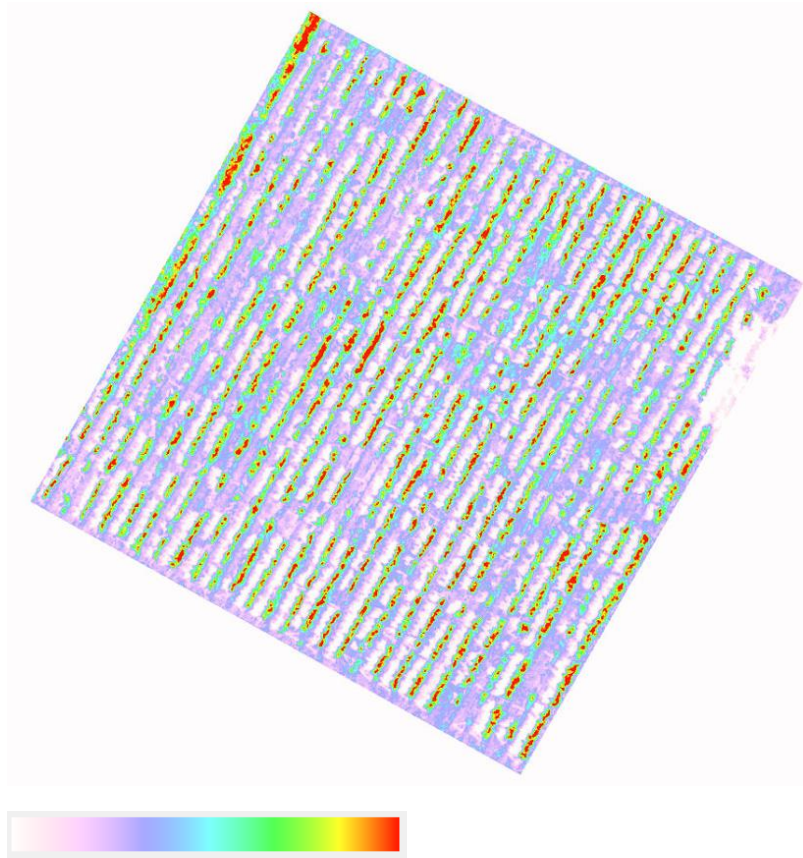
Figure 32: Chlorophyll Index Green (Clgr) implementation.

Mean value of Chlorophyll Index Green					
	A	B	C	D	E
1	-0,39	-0,45	-0,44	-0,50	-0,71
2	-0,47	-0,50	-0,47	-0,56	-0,50
3	-0,46	-0,41	-0,41	-0,47	-0,46
4	-0,47	-0,45	-0,47	-0,51	-0,43
5	-0,44	-0,46	-0,51	-0,48	-0,45
6	-0,43	-0,48	-0,47	-0,50	-0,43
7	-0,42	-0,40	-0,47	-0,48	-0,44
8	-0,40	-0,38	-0,43	-0,45	-0,43
9	-0,44	-0,42	-0,48	-0,45	-0,41
10	-0,45	-0,44	-0,45	-0,46	-0,47
11	-0,44	-0,43	-0,43	-0,48	-0,43
12	-0,41	-0,41	-0,45	-0,54	-0,47
13	-0,41	-0,43	-0,47	-0,44	-0,46
14	-0,46	-0,51	-0,47	-0,45	-0,46
15	-0,46	-0,43	-0,44	-0,42	-0,45
16	-0,46	-0,44	-0,45	-0,43	-0,50

17	-0,44	-0,46	-0,40	-0,44	
18	-0,46	-0,47			
19	-0,46	-0,43	-0,50		
20	-0,47	-0,47	-0,44		
21	-0,47	-0,45	-0,46		
22	-0,46	-0,45	-0,48		
23	-0,50				
24	-0,44	-0,52	-0,53	-0,50	-0,51
25	-0,46	-0,55	-0,48	-0,54	
26	-0,49	-0,45	-0,48	-0,51	

Table 8: Mean value of Chlorophyll Index Green for each variety.

4.2.3.2 Chlorophyll Index - Red-Edge (Clre)
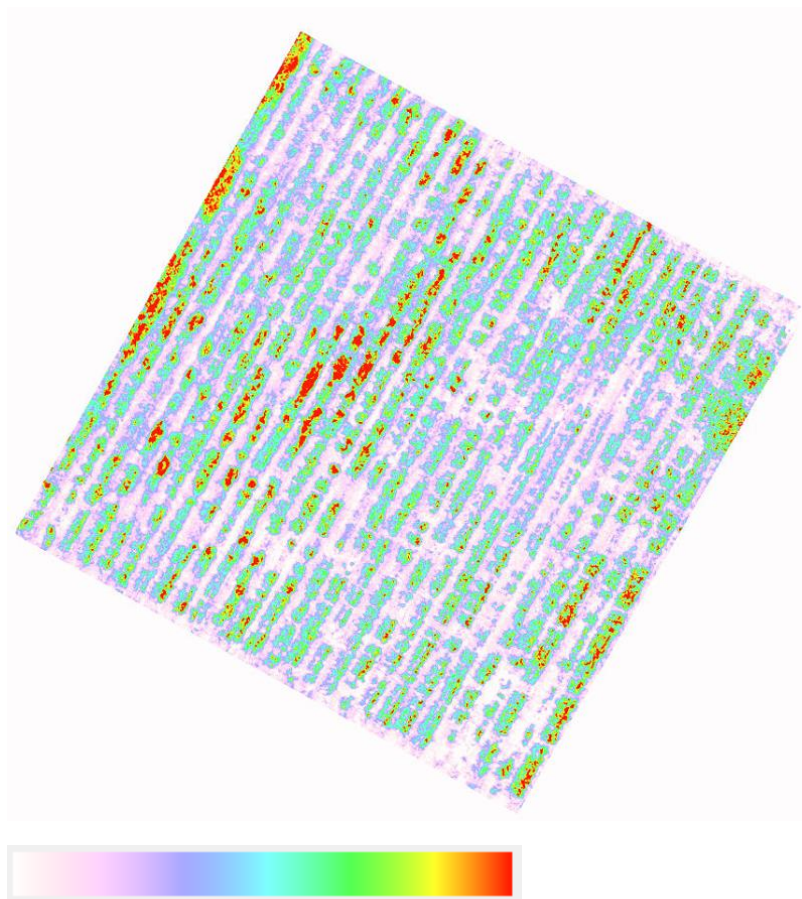
Figure 33: Chlorophyll Index - Red-Edge (Clre) implementation

Mean value of Chlorophyll Index - Red-Edge (CIre)					
	A	B	C	D	E
1	1,33	1,22	1,13	0,97	1,16
2	1,17	1,06	1,02	0,90	1,04
3	1,15	1,10	1,28	1,03	1,08
4	1,06	1,04	1,10	0,87	1,10
5	1,10	1,05	1,01	0,95	1,16
6	1,05	1,00	1,01	0,87	1,19
7	1,10	1,18	0,89	1,04	1,08
8	1,20	1,15	1,00	0,85	1,15
9	1,08	1,13	0,99	0,99	1,27
10	1,17	0,98	1,02	0,95	1,04
11	1,00	1,15	1,09	1,00	1,11
12	1,05	1,10	1,01	0,98	1,18
13	1,09	1,12	1,05	1,13	1,09
14	1,03	1,16	1,10	1,02	1,11
15	1,15	1,21	1,15	1,18	1,13
16	1,18	1,18	1,26	1,34	1,08
17	1,28	1,07	1,40	1,25	
18	1,27	1,10			
19	1,19	1,33	1,15		
20	1,24	1,01	1,05		
21	1,19	1,05	1,05		
22	1,15	1,12	1,04		
23	1,12				
24	1,18	1,25	1,15	1,07	0,97
25	1,20	1,03	1,17	0,97	
26	1,19	1,21	1,07	1,06	

Table 9: Mean value of Chlorophyll Index - Red-Edge (CIre) for each variety

4.2.3.3 CVI (Chlorophyll Vegetation Index)
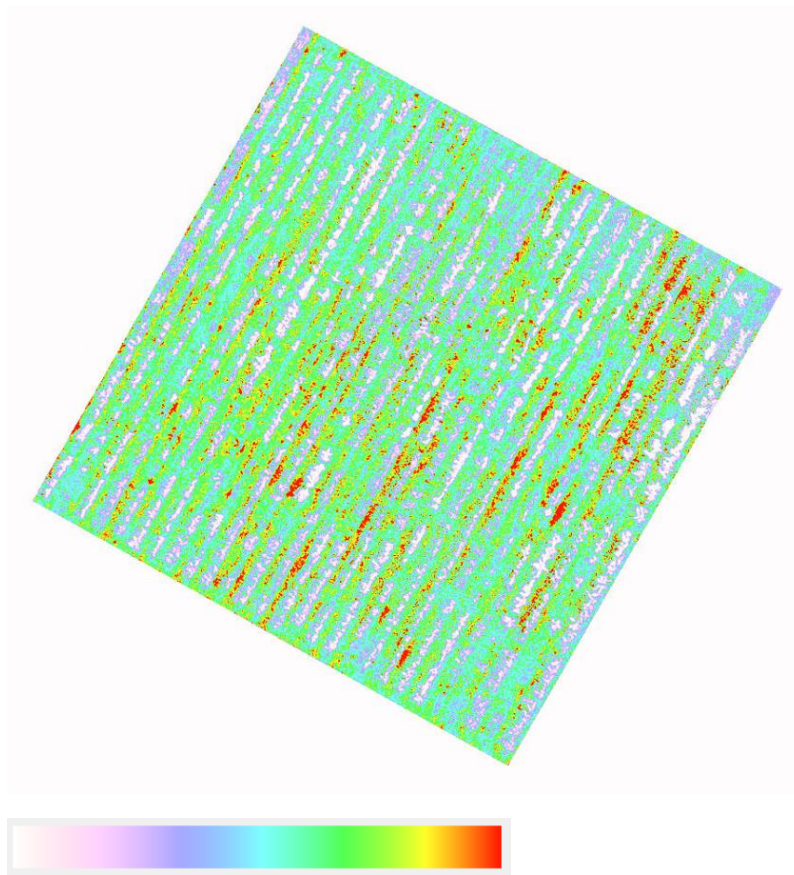
Figure 34: CVI (Chlorophyll Vegetation Index) implementation

Mean value of Chlorophyll Vegetation Index					
	A	B	C	D	E
1	2,84	2,72	2,51	2,21	2,50
2	3,62	3,07	2,84	2,05	2,29
3	2,83	3,01	3,02	2,94	2,64
4	3,91	2,44	2,83	2,46	3,26
5	3,26	3,16	2,86	2,55	3,19
6	2,64	2,88	3,13	2,90	3,06
7	2,69	3,23	2,61	3,05	2,60
8	2,96	2,90	2,94	2,48	2,75
9	2,85	3,04	2,62	2,48	2,99
10	3,06	2,41	2,69	2,72	2,68
11	2,64	3,08	3,15	2,79	2,55
12	2,81	3,34	3,05	3,54	3,09
13	2,80	3,14	3,12	3,08	3,20
14	2,72	3,70	2,95	2,86	3,37
15	2,99	3,15	3,02	2,79	3,05
16	2,92	2,66	3,09	3,21	3,32

17	3,53	3,01	3,67	3,40	
18	3,29	2,94			
19	3,09	3,50	3,21		
20	3,44	2,66	2,71		
21	3,03	2,41	2,74		
22	3,01	2,81	2,65		
23	3,26				
24	3,11	3,48	3,65	2,87	2,69
25	2,81	3,01	3,15	2,94	
26	3,25	3,20	2,74	2,67	

Table 10: Mean value of CVI (Chlorophyll Vegetation Index) for each variety

4.2.3.4 NDVI (Normalized difference vegetation index)
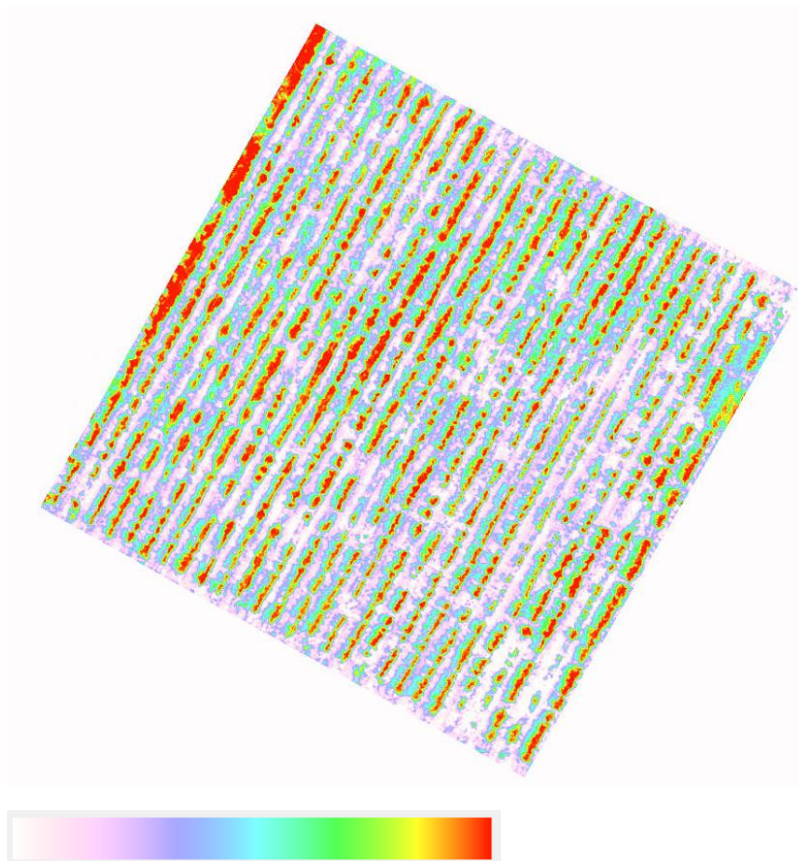
Figure 35: NDVI (Normalized difference vegetation index) implementation

Mean value of NDVI					
	A	B	C	D	E
1	0,82	0,78	0,80	0,80	0,77
2	0,76	0,74	0,77	0,80	0,79
3	0,80	0,79	0,82	0,77	0,79
4	0,74	0,78	0,77	0,76	0,77
5	0,78	0,76	0,75	0,77	0,78
6	0,78	0,74	0,75	0,74	0,80
7	0,78	0,79	0,75	0,76	0,78
8	0,79	0,80	0,77	0,76	0,80
9	0,78	0,79	0,77	0,78	0,80
10	0,79	0,78	0,77	0,77	0,77
11	0,79	0,79	0,77	0,77	0,79
12	0,77	0,81	0,77	0,73	0,78
13	0,77	0,79	0,76	0,79	0,79
14	0,77	0,76	0,78	0,78	0,78
15	0,78	0,81	0,79	0,81	0,79
16	0,79	0,80	0,80	0,82	0,77
17	0,80	0,80	0,82	0,81	
18	0,80	0,79			
19	0,80	0,81	0,79		
20	0,79	0,78	0,78		
21	0,80	0,81	0,79		
22	0,79	0,79	0,76		
23	0,78				
24	0,81	0,78	0,77	0,78	0,75
25	0,81	0,76	0,81	0,74	
26	0,79	0,81	0,79	0,78	

Table 11: Mean value of NDVI (Normalized difference vegetation index) for each variety

4.2.3.5 GNDVI (Green Normalized difference vegetation index)
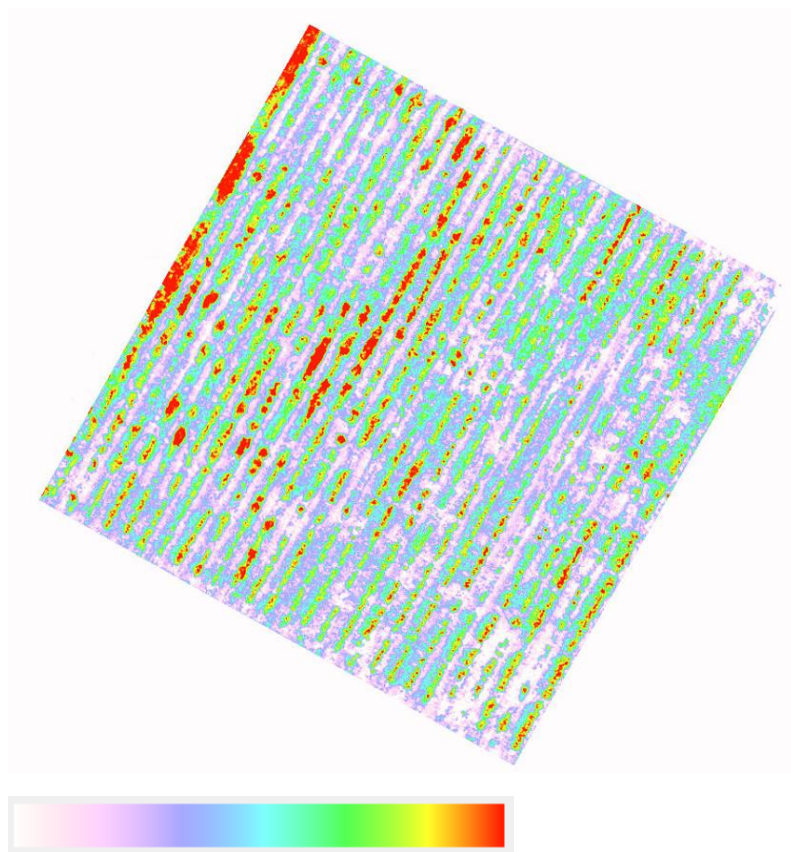
Figure 36: GNDVI (Green Normalized difference vegetation index) implementation

Mean value of GNDVI					
	A	B	C	D	E
1	0,69	0,66	0,66	0,64	0,63
2	0,68	0,64	0,66	0,62	0,64
3	0,68	0,68	0,70	0,66	0,66
4	0,67	0,64	0,65	0,62	0,67
5	0,68	0,66	0,64	0,63	0,68
6	0,65	0,63	0,65	0,63	0,69
7	0,65	0,69	0,62	0,66	0,65
8	0,67	0,68	0,66	0,62	0,67
9	0,66	0,68	0,64	0,64	0,68
10	0,68	0,64	0,65	0,65	0,64
11	0,65	0,68	0,67	0,65	0,66
12	0,65	0,70	0,66	0,65	0,67
13	0,65	0,68	0,65	0,68	0,68
14	0,65	0,68	0,67	0,66	0,68
15	0,67	0,70	0,68	0,68	0,68
16	0,67	0,67	0,69	0,71	0,67

17	0,70	0,68	0,72	0,71	
18	0,69	0,67			
19	0,68	0,71	0,68		
20	0,69	0,65	0,65		
21	0,68	0,66	0,66		
22	0,67	0,67	0,63		
23	0,67				
24	0,69	0,68	0,69	0,66	0,63
25	0,68	0,65	0,69	0,63	
26	0,68	0,70	0,66	0,65	

Table 12: Mean value of GNDVI (Green Normalized difference vegetation index) for each variety

4.2.3.6 EVI 2 (Enhanced vegetation index)
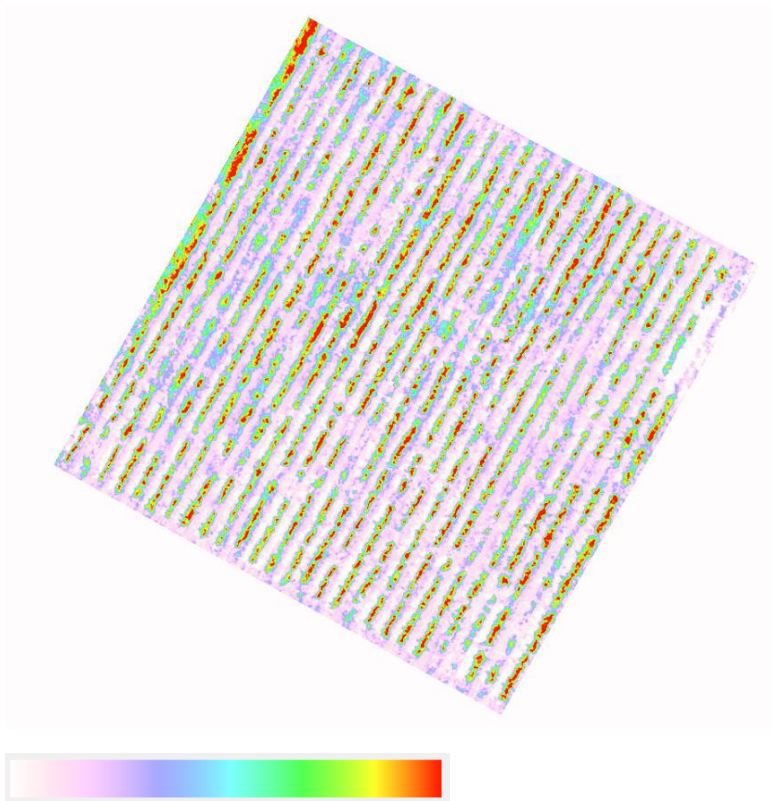
Figure 37: EVI 2 (Enhanced vegetation index) implementation

Mean value of EVI 2					
	A	B	C	D	E
1	0,86	0,80	0,82	0,77	0,53
2	0,75	0,73	0,77	0,71	0,77
3	0,80	0,82	0,85	0,77	0,80
4	0,73	0,80	0,78	0,74	0,80
5	0,79	0,77	0,72	0,77	0,79
6	0,81	0,74	0,75	0,73	0,82
7	0,81	0,83	0,76	0,75	0,81
8	0,83	0,85	0,80	0,79	0,82
9	0,81	0,82	0,76	0,80	0,83
10	0,79	0,81	0,79	0,78	0,77
11	0,80	0,81	0,79	0,77	0,83
12	0,82	0,83	0,78	0,68	0,78
13	0,82	0,81	0,76	0,80	0,78
14	0,78	0,72	0,78	0,80	0,77
15	0,79	0,82	0,81	0,83	0,80
16	0,79	0,82	0,80	0,83	0,74
17	0,80	0,79	0,84	0,81	
18	0,79	0,79			
19	0,79	0,82	0,76		
20	0,77	0,78	0,81		
21	0,78	0,82	0,79		
22	0,79	0,80	0,76		
23	0,75				
24	0,82	0,73	0,71	0,76	0,73
25	0,80	0,70	0,78	0,69	
26	0,76	0,80	0,78	0,75	

Table 13: Mean value of EVI 2 (Enhanced vegetation index) for each variety

4.2.3.7 RVI (Ratio Vegetation Index)
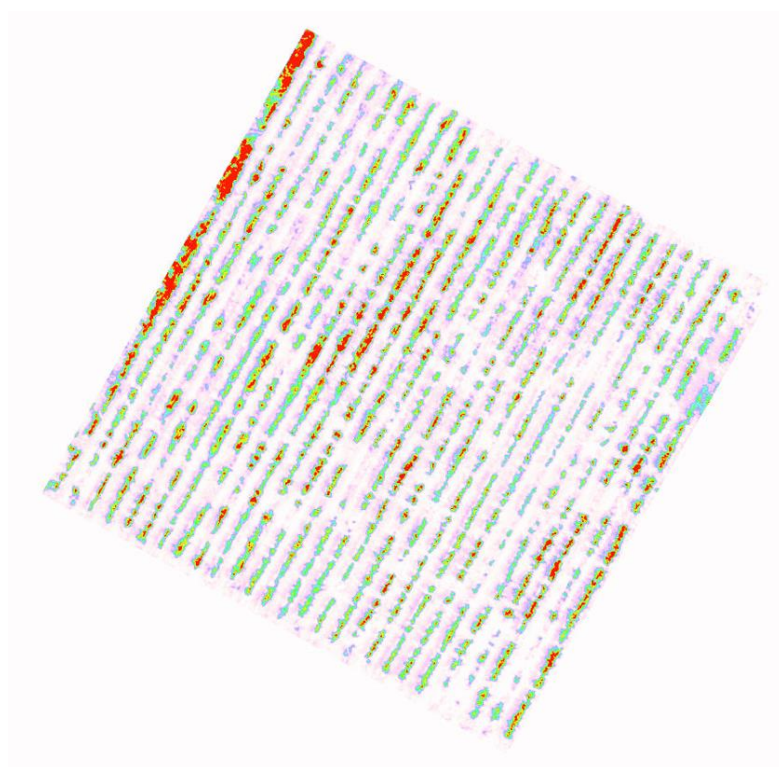
Figure 38: RVI (Ratio Vegetation Index) implementation

Mean value of RVI					
	A	B	C	D	E
1	11,75	10,63	11,34	10,77	8,80
2	8,36	7,99	9,78	10,45	10,27
3	11,09	10,32	12,13	9,04	10,47
4	7,57	10,44	8,90	8,72	9,18
5	9,59	8,60	8,44	8,94	9,77
6	9,59	7,56	7,81	7,38	10,86
7	9,46	10,23	8,18	8,76	9,72
8	9,96	10,31	8,90	8,55	10,79
9	9,96	10,19	9,27	9,99	11,38
10	10,39	9,61	9,37	9,24	9,22
11	9,87	10,26	9,36	9,14	10,51
12	9,29	11,55	9,49	7,08	9,67
13	9,05	10,19	8,71	10,17	9,89
14	9,07	8,83	9,87	9,36	9,36
15	9,73	11,47	10,47	11,20	10,43
16	9,66	11,29	11,10	12,13	8,71

17	10,58	10,36	12,15	11,59	
18	10,77	10,40			
19	10,63	11,46	10,32		
20	9,86	9,45	9,84		
21	10,72	11,42	10,20		
22	10,03	9,98	9,28		
23	9,39				
24	10,87	9,79	8,91	9,60	8,76
25	11,18	8,27	11,09	7,82	
26	10,10	11,12	10,38	9,91	

Table 14: Mean value of RVI (Ratio Vegetation Index) for each variety

4.2.4 Qualitive evaluation of indices

The results of the indices implementation are presented in aggregate in the Figure 39.
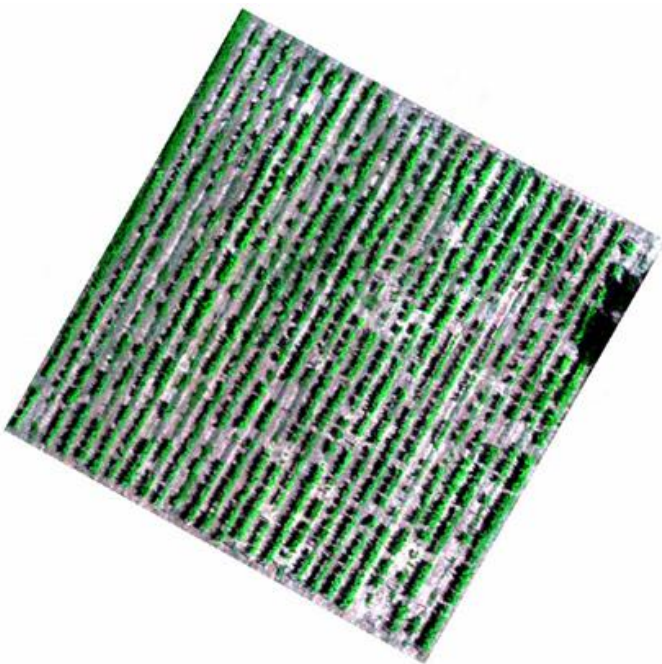
RGB   **CLGR**

CLRE

CVI

NDVI

GNDVI

EVI2

RVI
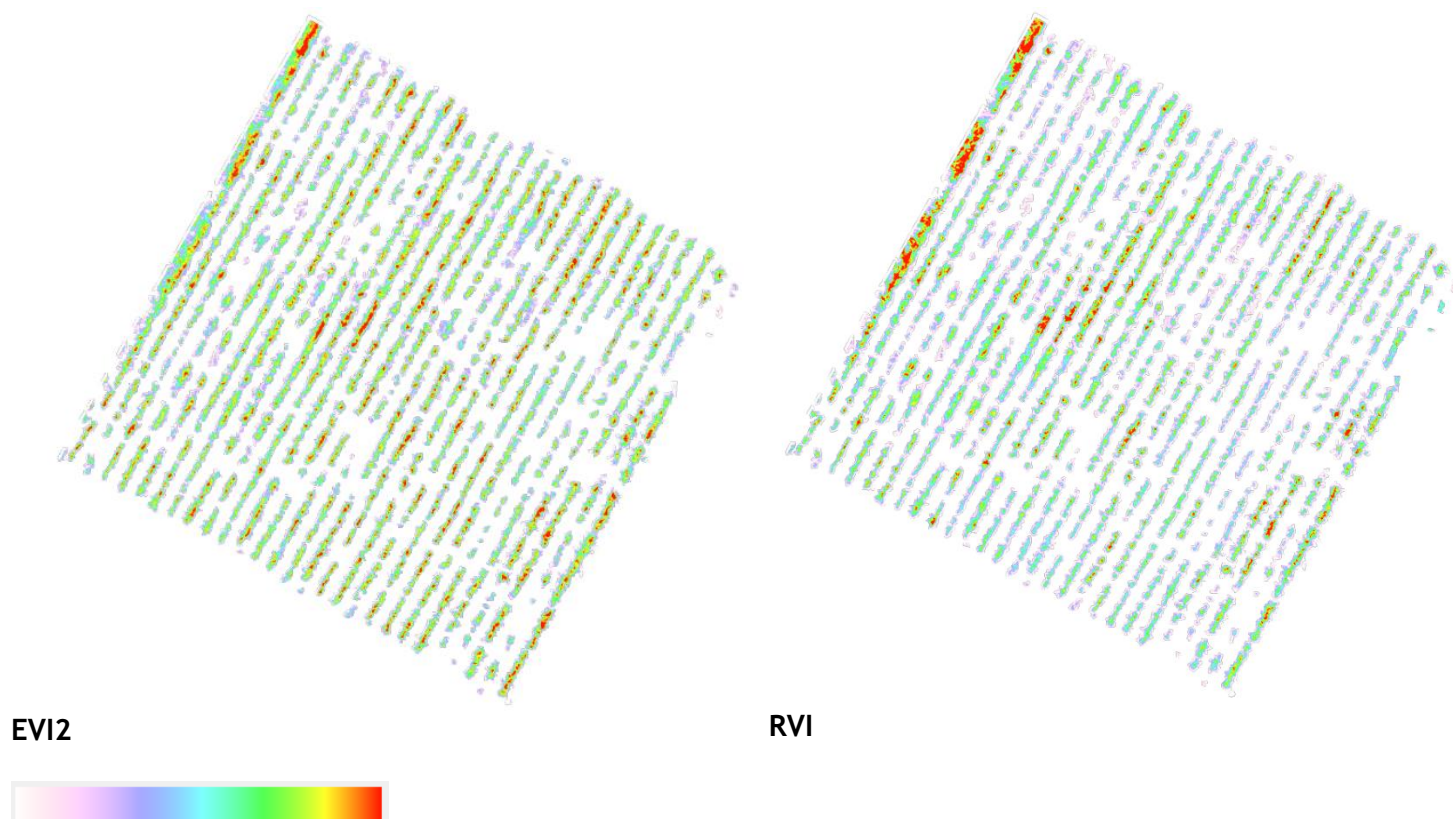
Figure 39: Indices implementation on the image in the masked areas

These vegetation indices have different spectral band combinations and calculation methods, allowing them to capture specific aspects of vegetation health, density, and chlorophyll content. The optical results for the above-presented indices are summarized below:
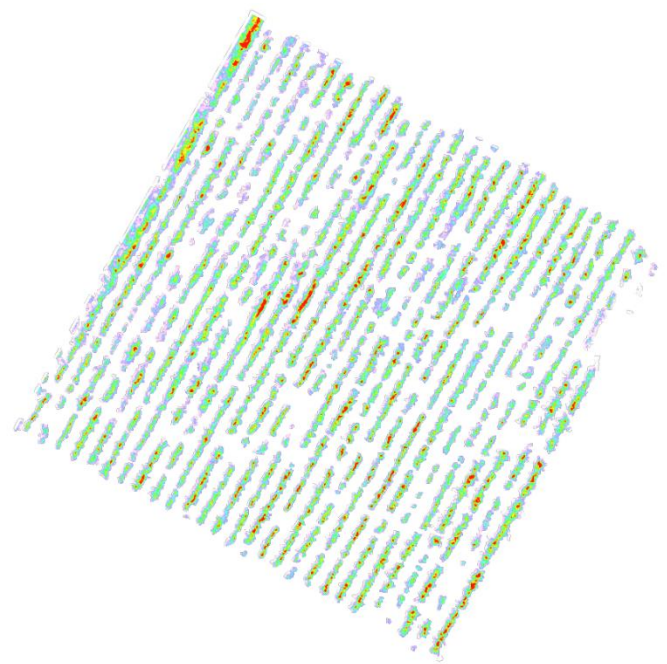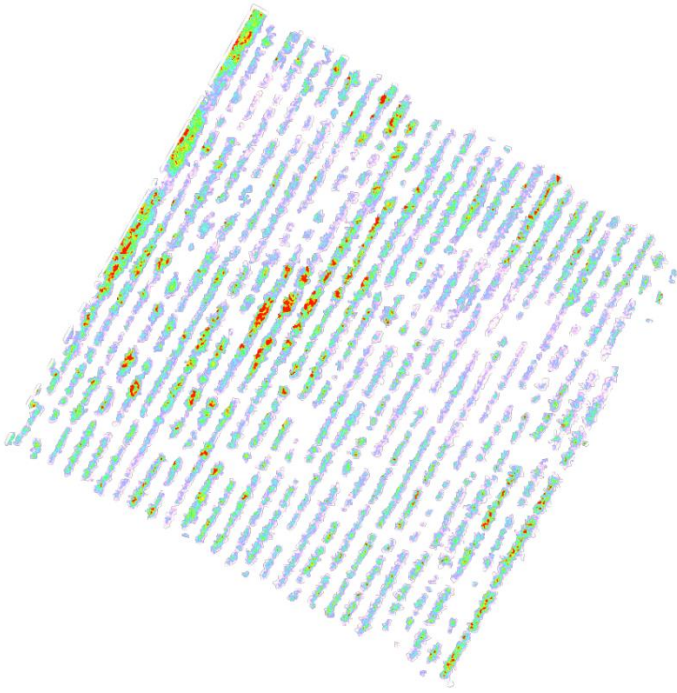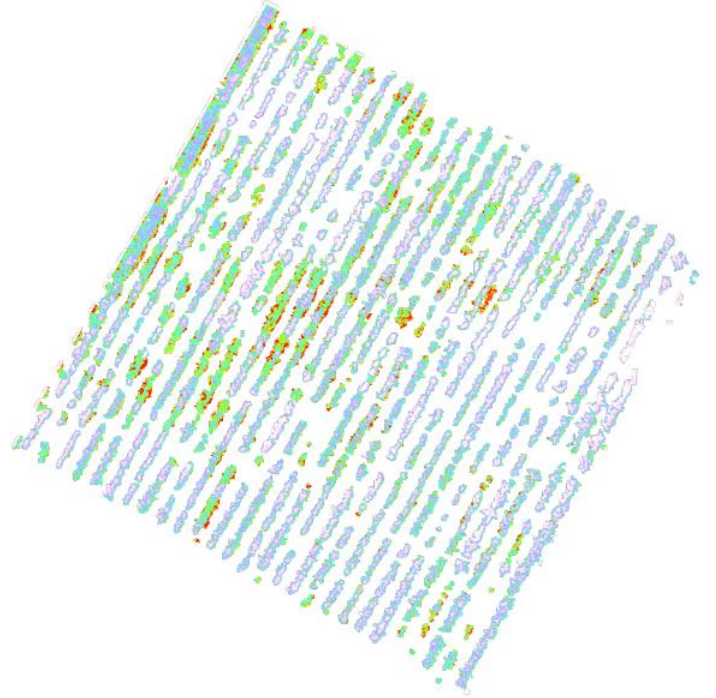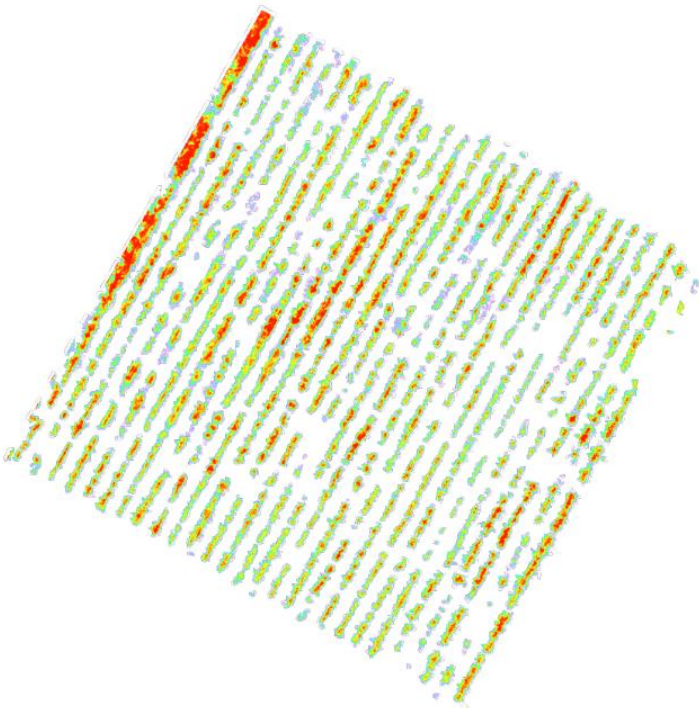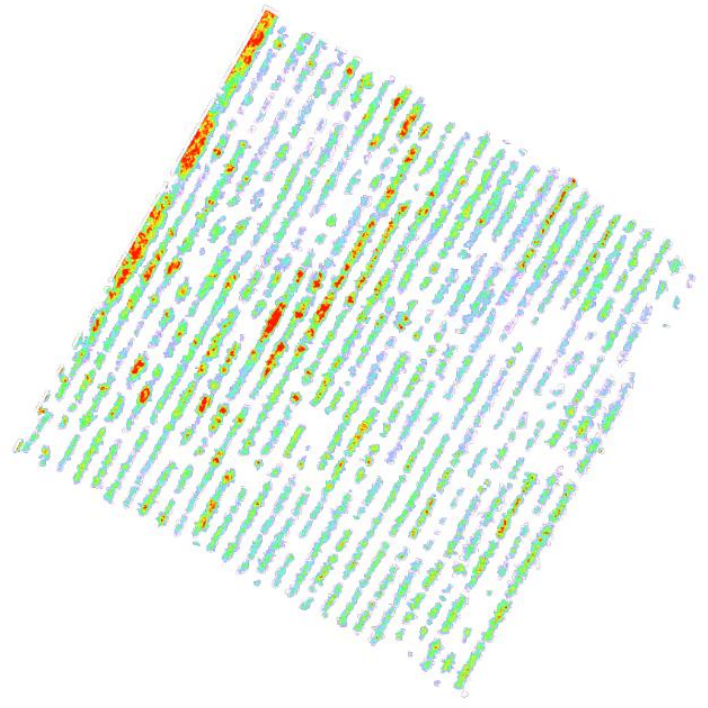
- **Chlorophyll green index - CLGR**: it focuses specifically on the green band and it shows sensitivity to chlorophyll concentration. It is observed that in almost the entire study area the values remain consistently low, while only a few plants exhibit higher values. Also, higher chlorophyll values are presented on the center of each plant as there the foliage is denser.

- **Chlorophyll red-edge index - CLRE**: it aims to calculate chlorophyll concentration in vegetation by using the reflectance in the red-edge region of the electromagnetic spectrum. It is particularly effective in differentiating vegetation types and detecting stress in plants. It is evident that a small area shows higher values, indicating the presence of stress.

- **Chlorophyll Vegetation Index - CVI**: it directly estimates the chlorophyll content in plant leaves and provides insights into leaf health and photosynthetic activity. In the study area high concentration values are missing.

- **Normalized Difference Vegetation Index – NDVI:** is a widely used vegetation index that evaluates the health and vigor of vegetation. It provides information about the density and amount of green vegetation. On the

produced image it seems that all the vines do not provide the same amount of greenness. The difference in this value is related either to the health or the lack of a substance on the leaves or indicates that some plants are younger than others.

- **Green Normalized Difference Vegetation Index – GNDVI:** is similar to NDVI but uses the green band instead of the red band. The main characteristic is that indicates the areas where the vegetation is denser. It is evident that the vines in the center region of the study area show higher values of the GNDVI index.
- **Enhanced Vegetation Index 2 – EVI2:** EVI2 is an improved version of the traditional EVI. It incorporates adjustments to account for atmospheric influences, canopy, and soil. EVI2 provides a more accurate estimation of vegetation conditions, particularly in areas with dense canopies. It seems that only some areas show high values.
- Ratio Vegetation Index - RVI: it evaluates vegetation health and it is sensitive to changes in chlorophyll content and can indicate plant stress. The RVI image showed that only small parts of the entire area appear stress.

4.2.5 Quantitative analyses of vegetation indices

The following steps include a quantitative procedure to evaluate the vegetation indices for clustering varieties into categories with common characteristics. These steps consist of correlating the indices, along with keeping and further analyzing those that indicate lower correlation with each other. From the correlation matrix, two sets of different indices were retained, within a total of four out of the seven indices. Then, two different clustering methods were implemented to categorize varieties into four different groups with similar properties. The results of the clustering methods are illustrated and analyzed in the subsequent steps, and finally, the results of the entire index analysis procedure is presented.

4.2.5.1 Correlation matrix between vegetation indices

Firstly, a matrix was calculated containing the mean values for each vegetation index measured in all the varieties (Table 14).

```
Dataframe is :
     CLGR  CLRE   CVI  NDVI  GNDVI  EVI2    RVI
0   -0.39  1.33  2.84  0.82   0.69  0.86  11.75
1   -0.47  1.17  3.62  0.76   0.68  0.75   8.36
2   -0.46  1.15  2.83  0.80   0.68  0.80  11.09
3   -0.47  1.06  3.91  0.74   0.67  0.73   7.57
4   -0.44  1.10  3.26  0.78   0.68  0.79   9.59
..    ...   ...   ...   ...    ...   ...    ...
107 -0.46  1.09  3.20  0.79   0.68  0.78   9.89
108 -0.46  1.11  3.37  0.78   0.68  0.77   9.36
109 -0.45  1.13  3.05  0.79   0.68  0.80  10.43
110 -0.50  1.08  3.32  0.77   0.67  0.74   8.71
111 -0.51  0.97  2.69  0.75   0.63  0.73   8.76

[112 rows x 7 columns]
```

Table 15: Dataframe with index values.

The correlation matrix consists of 112 rows equal to the total number of different varieties and seven columns, one for each vegetation index. The correlation matrix between the mean values of each index in relation with the others is presented in Figure 40.:
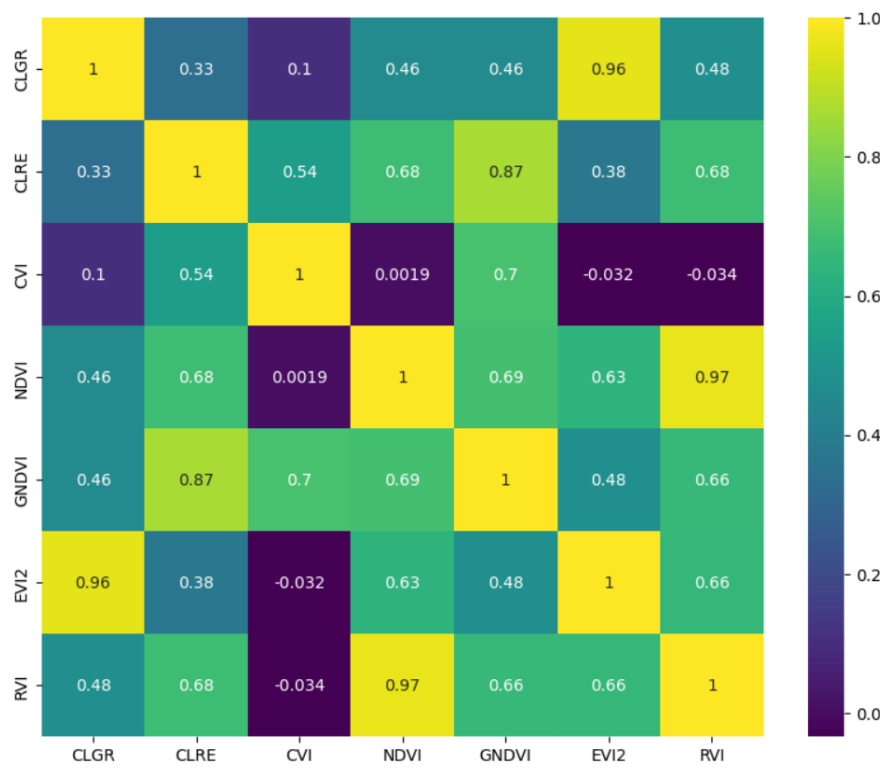
Figure 40: Correlation matrix between indices

Observing the correlation matrix, the following conclusions can be drawn:

- The CVI index appears to have the lowest correlation with most of the other indices. Specifically, it shows a correlation lower than or equal to 0.1 with the RVI, EVI2, NDVI and CLGR indices.
- The CLRE index shows lower correlation values at 0.33 and 0.38 with the CLGR and EVI2 indices, respectively.
- The indices with higher correlation values are NDVI with the RVI index, and CLGR with the EVI2 index. Their correlation value is almost equal to 1, indicating that they have nearly identical results.

For a reliable analysis of the vineyard varieties, it is crucial to retain only the indices that exhibit low correlation with each other. In the subsequent steps of the procedure, two sets comprising of a total of four different vegetation indices will be selected. By considering the lower values in the correlation matrix, we chose the following pairs of indices:

- CVI- RVI
- CLGR- CLRE

4.2.5.2 Linear transformation of indices' values

The value range of each index varies between different scales. It is important though, for better visualization and a more reliable classification, the values range to be adjusted between 0 and 1. To achieve that a linear transformation was executed on the indices value table (Figure 41 and 42)..
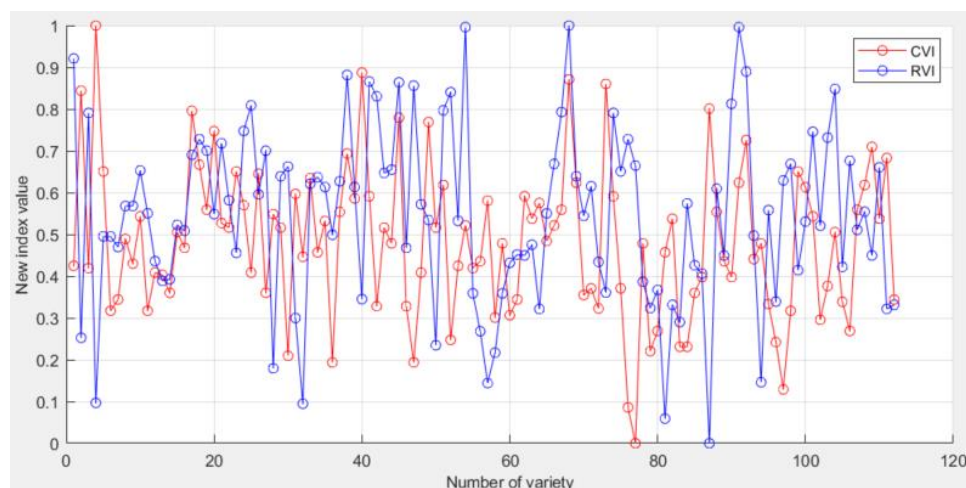
Figure 41: Linear transformed values of CVI and RVI indices.
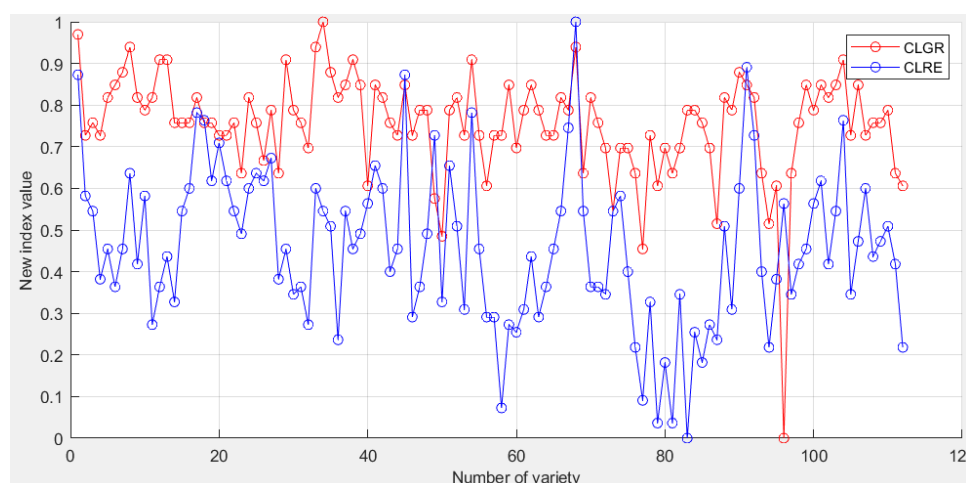
Figure 42: Linear transformed values of CLGR and CLRE indices.

4.3 Clustering of vineyard varieties
4.3.1 Clustering methods

Clustering algorithms fall into two broad categories. The first category is the hard clustering, where each data point belongs exclusively to a single cluster. An example of such an algorithm is the k-means method. The second category is known as soft clustering, where each data point can be assigned to multiple clusters simultaneously. A common algorithm for soft clustering is the Gaussian mixture model (Figure 43) and these are the methods selected for the data clustering.

| k-means clustering | Gaussian mixture model |
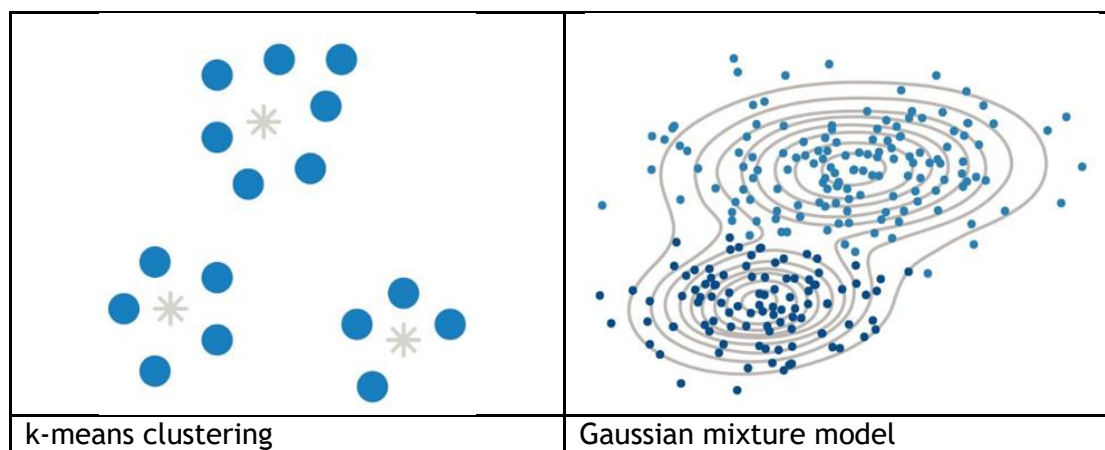
Figure 43: Clustering algorithms presentation

Regarding the k-means method, it is an unsupervised classification which executes an iterative process to cluster data points into groups based on their similarity. The aim is to minimize the sum of distances between data points and their cluster centroids, for an accurate grouping. In the k-means algorithm the Squared Euclidean distance is used, where each centroid represents the mean of the points within that cluster.

Gaussian mixture model, which is the second clustering method applied, it is a probabilistic model, which assumes that all data points are generated from a combination of a finite number of Gaussian distributions with unknown parameters. The method uses the expectation-maximization (EM) algorithm to fit the models. It also has the capability to draw confidence ellipsoids for multivariate models and compute the Bayesian Information Criterion to determine the optimal number of clusters in the data. The Gaussian Mixture offers various options to constrain the covariance of the estimated difference classes, including spherical, diagonal, tied or full covariance.

4.3.2 K-means clustering implementation

The first implemented clustering method is the k-means algorithm, for which the chosen distance metric is the Squared Euclidean distance, and the number of clusters is set to k=3. The classification results are depicted in the Figure 44. To assess the quality of the classification results in terms of cluster compactness, cluster distance, and overlapping, the silhouette coefficient was used. The silhouette coefficient values range between -1 and 1, where a value of 1 indicates that the data points within each cluster are highly compact and the clusters are clearly distinct from each other. Values near 0 suggest the presence of overlapping between data points, while a value of -1 represents the worst-case scenario.

4.3.2.1 K-means clustering on CLGR- CLRE pair indices

Figure 44: Clusters of k-means classification between CLGR and CLRE indices

Based on the above visualizations, it is evident that the values of each variety are evenly distributed across the three classes.
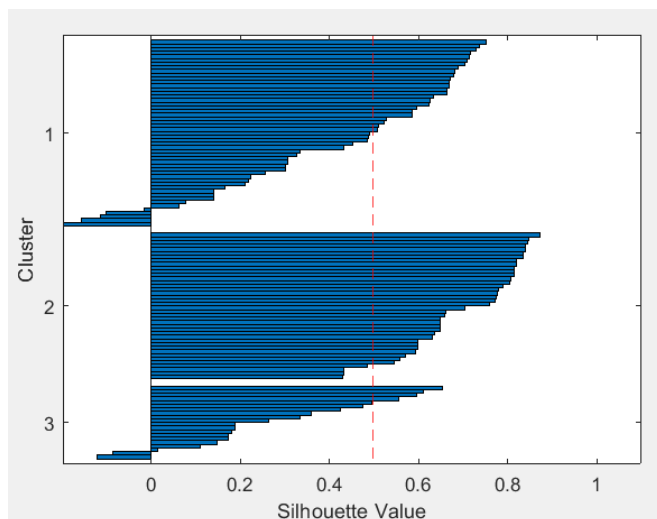
Figure 45: Silhouette Value of CLGR and CLRE clustering

The silhouette coefficient was calculated to be 0.50 (Figure 45). Generally, a silhouette score of 0.5 or higher is considered indicative of good clustering. Therefore, based on the threshold of 0.5, the results can be considered satisfactory, and can be retained.

4.3.2.2 K-means clustering on CVI – RVI pair indices

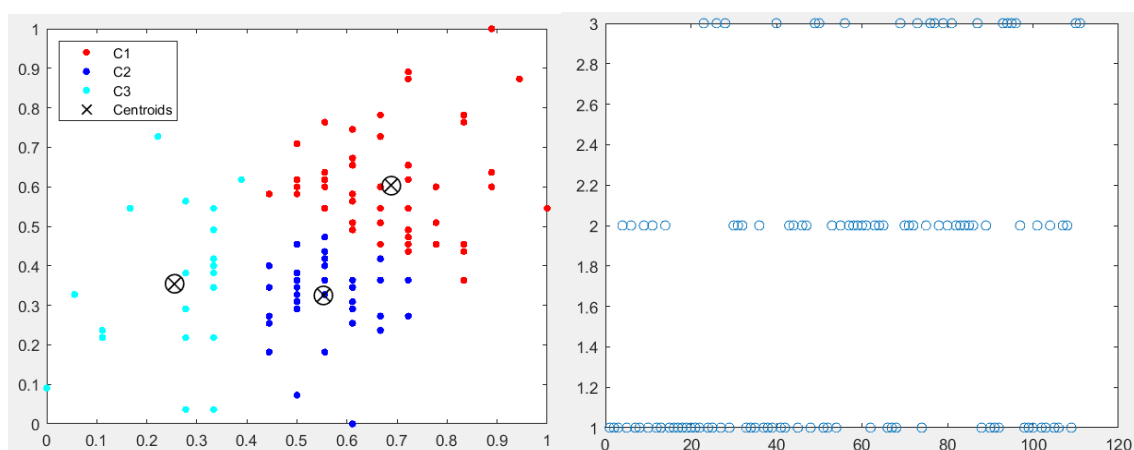The classification results of k-means algorithm for the indices pair CVI – RVI are displayed in Figure 46.
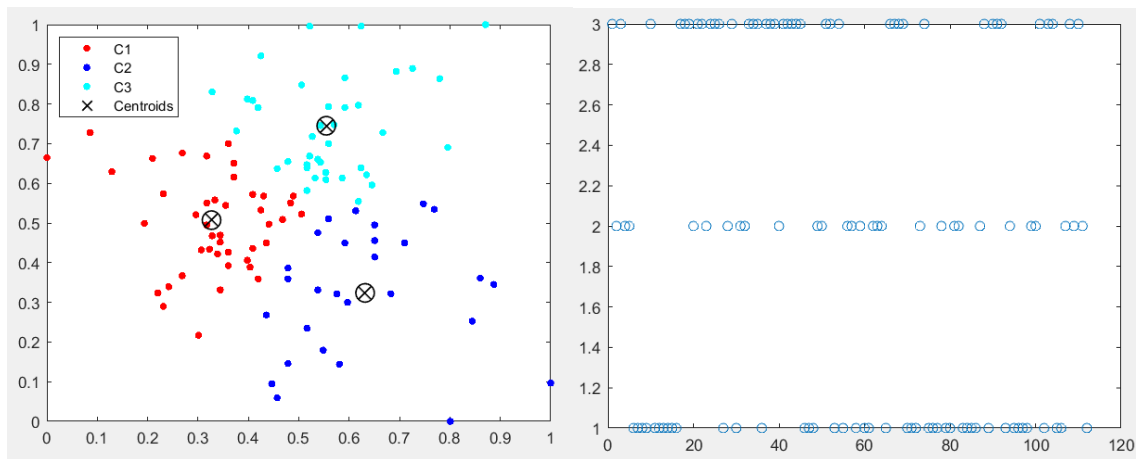
Figure 46: Clusters of k-means classification between CVI and RVI indices

Based on the above visualizations, it is evident that the values of each variety are also evenly distributed across the three classes.
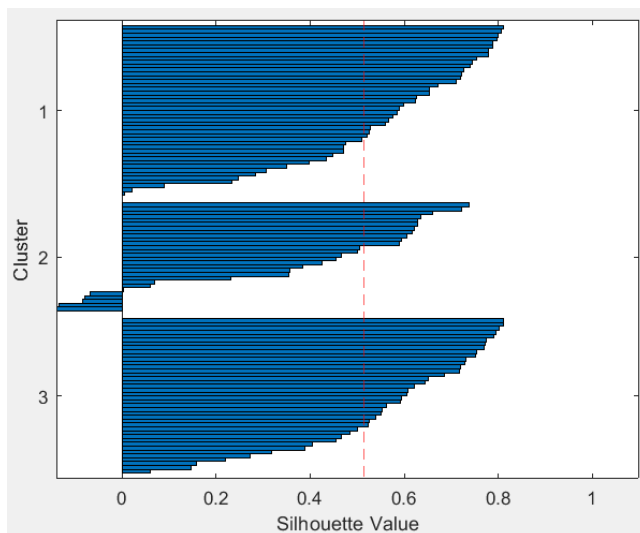
Figure 47: Silhouette Value of CVI and RVI clustering

The silhouette coefficient was calculated to be 0.52. Therefore, based on the threshold of 0.5, the results can be considered satisfactory and they are be retained.

4.3.3 Gaussian mixture model clustering implementation

Continuing, the second algorithm, the Gaussian mixture model, it is used to cluster the same pairs of indices in also three classes.

The machine learning algorithm can be implemented by taking into consideration two parameters: the value of sigma, which can be either diagonal or full, and the shared covariance value, which can be set as true or false.

In diagonal covariance matrices, the predictors are not correlated with each other. The ellipses' axes are either parallel or perpendicular to the x and y axes. This specification increases the number of parameters, but it is more solid compared to using the full covariance.

In the full covariance there are no restrictions on the orientation of the ellipses' axis. However, using full covariance significantly increases the number of parameters, which may lead to a higher probability of overfitting.

In the shared covariance matrices, all components have the same covariance matrix. This means that the size and the orientation of the ellipses are identical. Using shared covariance is more solid compared to unshared covariance, as the number of parameters is increased only for one component.

In the unshared covariance matrices, each component has its own covariance matrix. This allows for differentiation in the size and orientation of the ellipses, with the number of parameters able to increase depending on the size and the orientation of the ellipses, but this approach can effectively capture covariance differences among components.

4.3.3.1 Gaussian mixture model clustering on CLGR- CLRE pair indices

The values distribution for the CLGR- CLRE indices pair is presented in Figure 48.

Figure 48: Values distribution of CLGR – CLRE indices

Based on the algorithm implementation, the results for all combinations of the parameters are shown in Figure 49.
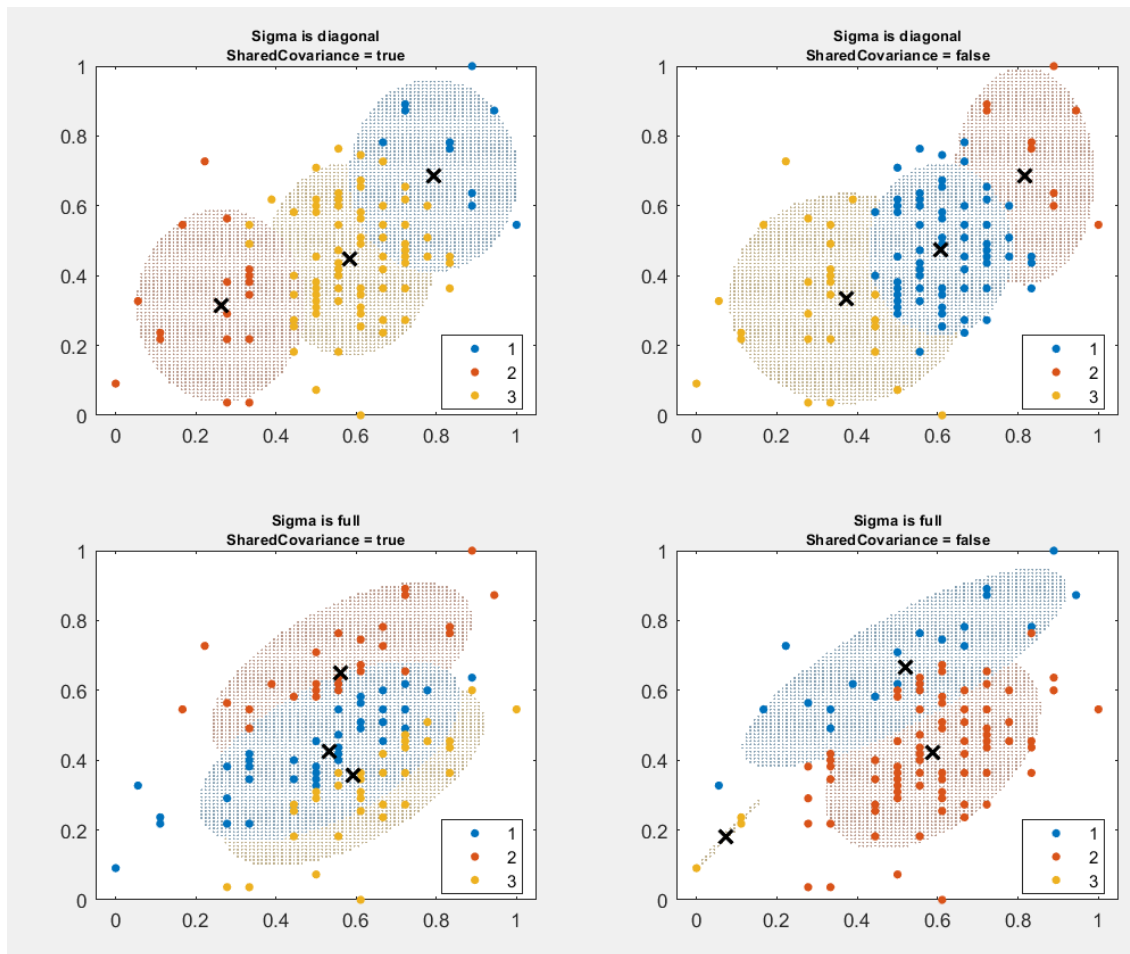
Figure 49: Gaussian mixture model clustering results for CLGR – CLRE indices (where 1,2,3 is the resulted clusters)

Based on the above visualizations, it appears that the combination of sigma-full and shared covariance-true exhibits the most even distribution of values among the three classes.

4.3.3.2 Gaussian mixture model clustering on CVI – RVI pair indices

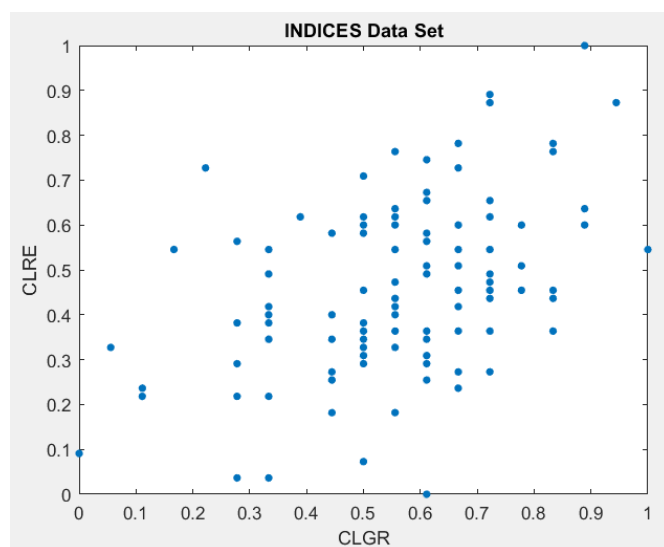Firstly, the values distribution for the CVI-RVI indices pair is shown in Figure 50.
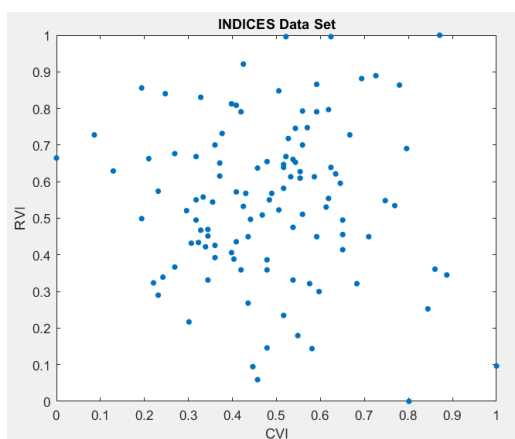
Figure 50: Values distribution of CVI – RVI indices

After the algorithm implementation, the results for all combinations of the parameters are presented in Figure 51.
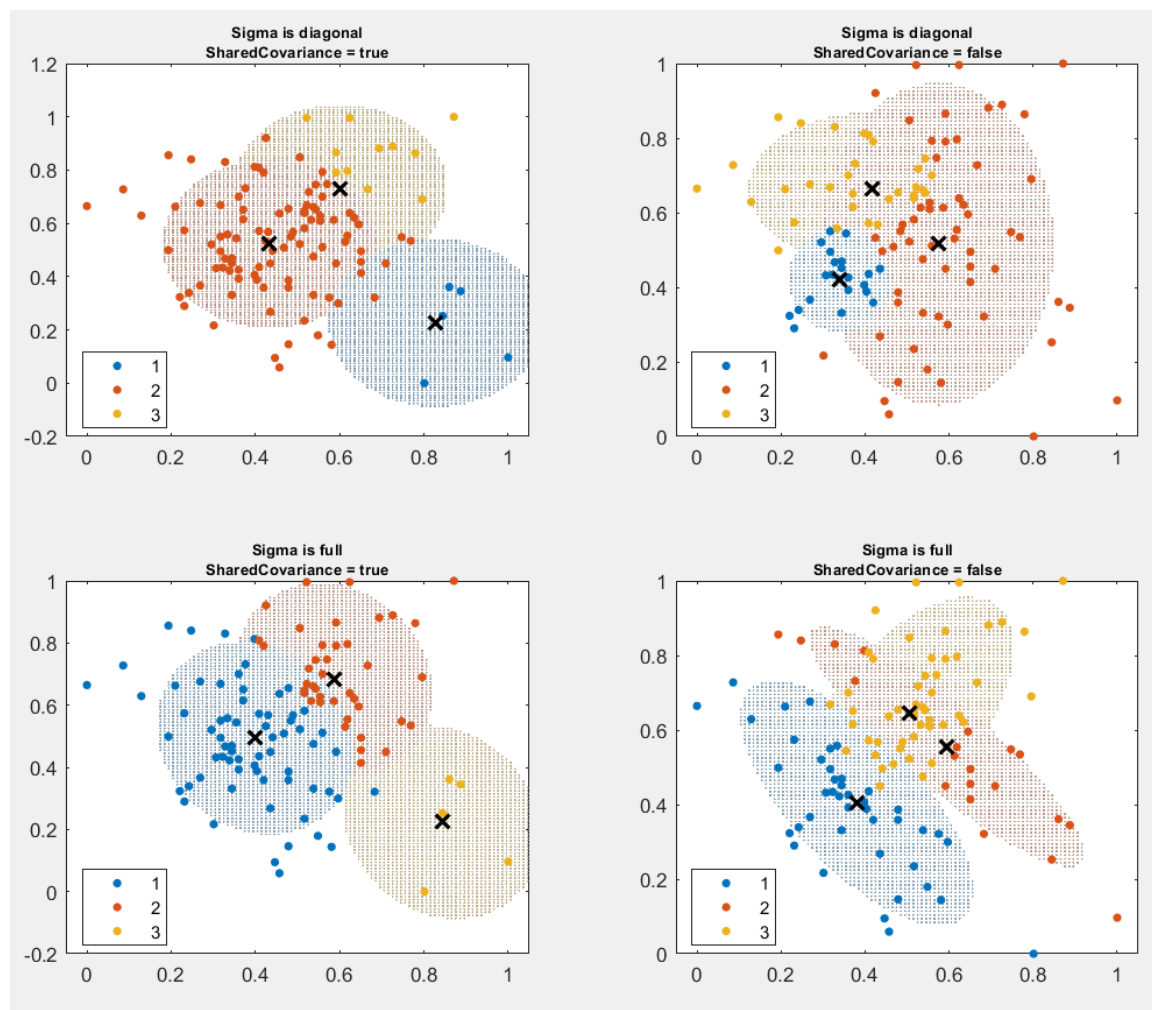
Figure 51: Gaussian mixture model clustering results for CVI- RVI indices (where 1,2,3 is the resulted clusters)

Taking into account the above visualizations, it appears that the combination of sigma-diagonal and shared covariance-false exhibits the most even distribution of values among the three classes.

4.4 Clustering algorithms analysis

After the classification of the varieties' dataset in two different ways and for two combination of pair indices, we evaluated these results. For each algorithm, the results made from the two-index clustering are compared with each other and with the purpose to extract observations related with varieties that both pair of indices classified together, as having common characteristics. For the implementation of this method the results of the first pair of index classification grouped in three classes (based on the classification results), and compared with the same results from the second pair of indices. In Table 15 are presented the variety codes and names that the algorithms cluster in the same class, whereas Figures 52-55 show maps with the location of these varieties inside the vineyard field.

K-means			Gaussian Mixture model		
class1	class2	class3	class1	class2	class3
A6	A23	A1	A6	A1	A3
A9	B2	A3	A7	A2	A10
A11	B14	A10	A11	A16	B16
A14	B24	A17	A12	A17	B17
B4	B25	A18	A13	A18	B18
B10	C5	A19	A14	A20	B22
B20	C24	A21	B20	A23	C1
B21	D6	A22	C9	A26	C15
C2	D12	A24	C10	B14	C26
C4	D25	A25	C20	B15	D1
C7	E16	B3	D4	B19	D2
C9		B7	D5	B24	D15
C10		B8	D8	B26	D26
C14		B9	D10	C3	E2
C20		B11	D11	C16	E3
C21		B12	D14	C17	E6
C22		B13	E7	C19	E8
C26		B15		C24	E15
D5		B16		C25	
D8		B19		D16	
D9		B26		D17	
D10		C1		E9	
D11		C3		E12	
D14		C15			
E3		C16			
E7		C17			
E10		C25			
		D13			
		D15			
		D16			
		D17			
		E6			
		E8			
		E9			
		E15			

Table 16: Final clustering of varieties

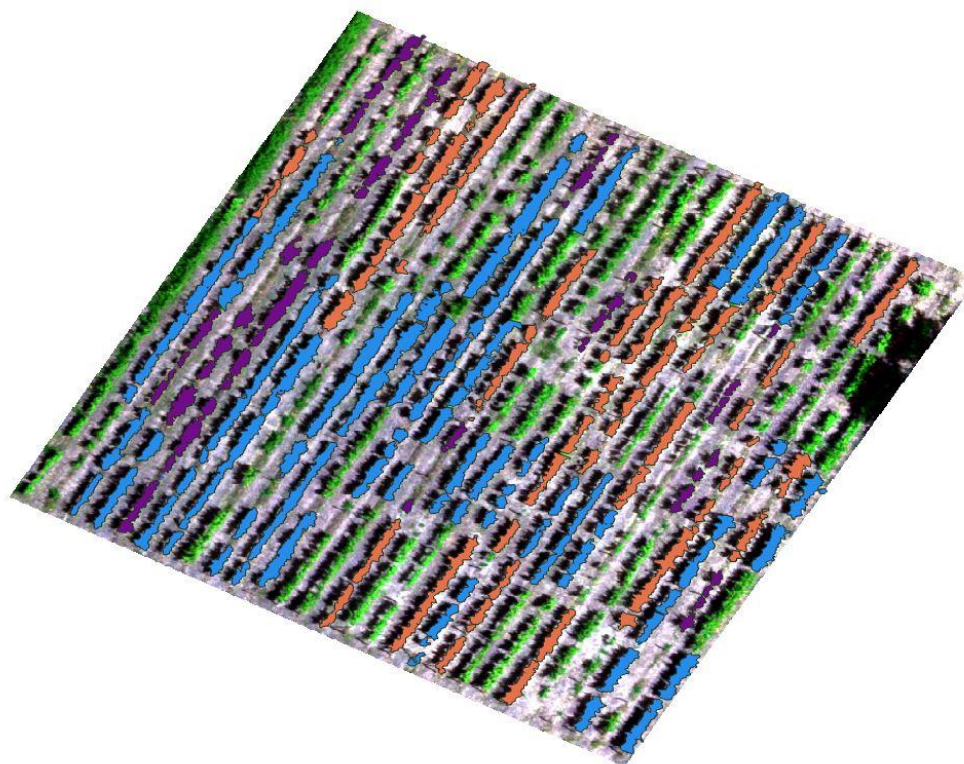Should be noted that the exact names of varieties cannot be revealed as this information is confidential.
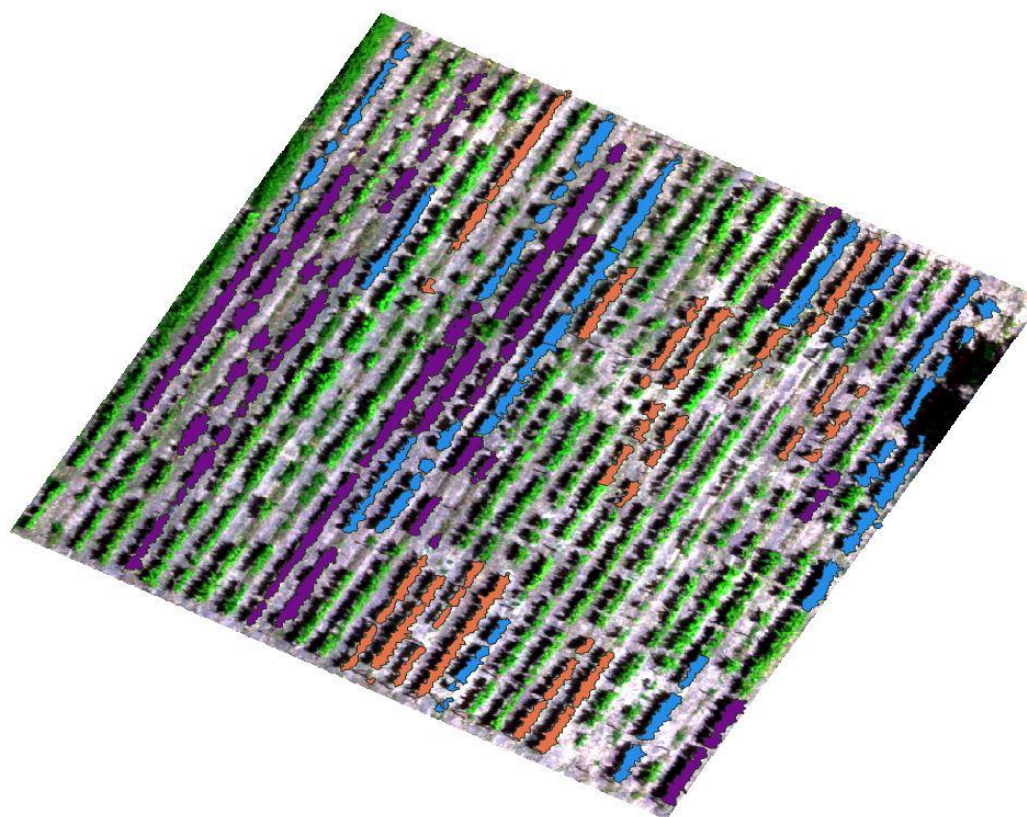
Figure 52: K-means clustering results

Figure 53: Gaussian mixture model clustering results
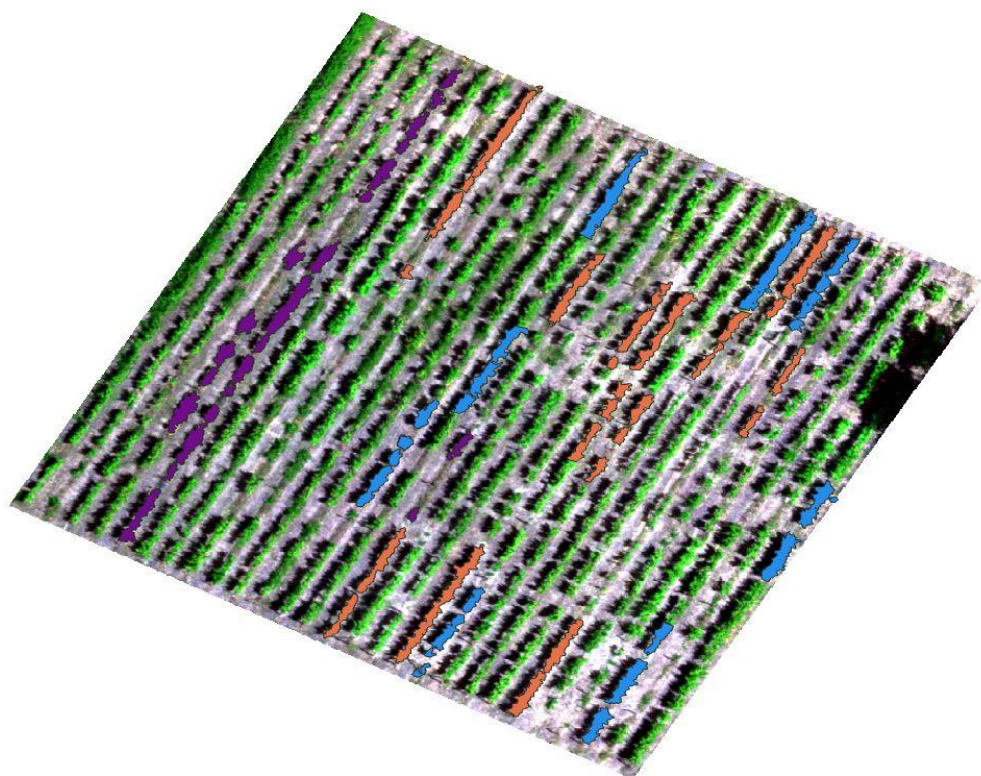
Figure 54: Clustering results by the combination of k-means and Gaussian mixture model algorithms in three classes
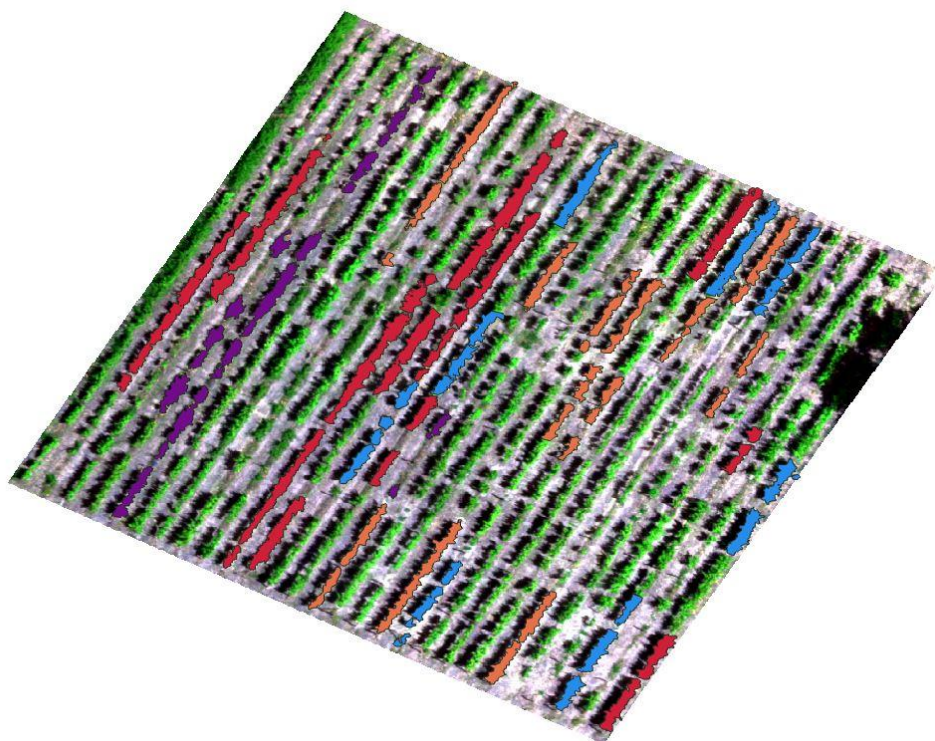
Figure 55: Clustering results by the combination of k-means and Gaussian mixture model algorithms in four classes

4.4.1 Key points from the clustering implementation

Based on the above implementation and the subsequent visualization, the key points found are the following:

- Regarding the k-means algorithm results, and considering the total of the four indices studied (RVI- CVI and the CLGR – CLRE), most of the varieties were classified into the 3 classes. Specifically, 73 of the 112 varieties were classified into the three groups.

- Gaussian mixture model used a stricter classification, as it takes into consideration more parameters and from the four indices analysis classified 58 of the 112 indices into 3 classes. This number is much smaller in contrast with the k-means classification as it represents almost half of the total varieties.

- Many of the varieties were classified in a common class by both clustering algorithms, which ensures the validity of the method.

- The rest of the varieties are characterized as unclassified as the two different pair of indices clustered them in different classes.

- Another important observation is that varieties that belong in the same class, except for the spectral characteristics they may have in common, they have also spatial correlation with each other. This means that neighbor vines, except for spectral characteristics, they also share the same ground conditions; soil characteristics of an area can influence the vine's growth either in a positive or in a negative way. Thus, the ground substances are responsible for the condition inside the plant's tissue and leaves.

- The spectral values of a plant, apart from the plant's health status, depend also on its growth stage. Therefore, it is possible that vines that neighbor may grow simultaneously and this status is related to the portion of water or fertilizers that they have absorbed.

5. Discussion - Conclusion – Further Research

The present study aimed to examine the feasibility of using multispectral images to detect different varieties of a crop, specifically grapevines. The main objective was to find differences between multiple grape varieties using only the information provided by the spectral profiles of the leaf surfaces. The images used in the study were collected through UAV (Unmanned Aerial Vehicle) flights, focusing on entire sections of the selected vineyard (located in Lykovrisi area at the northern part of Attica region) and only on the canopy of them. Extensive literature research revealed that this topic seems to have not been sufficiently examined so far, as usually an area that produces either table grapes or wine would involve one or a small number of varieties. The study area had the characteristic of containing many varieties in a very small area, and the sample from each variety was also relatively small.

Initially, the research showed that it was not feasible to classify and differentiate grape varieties correctly via only using the bands of the multispectral camera. Therefore, it was decided to create categories where the varieties belonging to the same group would show common characteristics that would emerge after applying vegetation indices to the image. Indeed, the varieties were successfully classified into three classes, although some remained unclassified based on the indicated features. Additionally, the classification was performed using two different algorithms to verify the results of one method against the other. It was found that several varieties were classified together by both algorithms. The results indicated also that some of the varieties included in this study did exhibit common characteristics in terms of spectral information. By observing the final map, it emerged that in some cases, the varieties classified in the same class were spatially related. This observation raises the question: if these methods applied in an area containing only one variety, would the values of the indices within clusters of vineyards show such significant differences or would they be very similar, making it more challenging to classify them into different distinctive classes? In other words, the classification derived from this study is influenced by the unique characteristics of the varieties examined, or it was because the vines, due to soil characteristics, water and fertilizer quantities received by the vines, might exhibit different characteristics?

In this study, certain vegetation indices were selected to differentiate grape varieties by identifying their features. The underlying assumption was that each variety possesses distinctive characteristics and, therefore, the growth time and the likelihood of lacking certain substances, leaf density etc. may alter depending on the variety. Thus, it is possible to encounter such differences in the same variety, but this is an issue that should be investigated in future work.

From the maps that were generated after applying the indices, it became apparent that in certain areas, almost all the indices yielded higher values compared to other points in the study area. Considering the small area covered by each variety, it can be concluded that when a region of the vineyard, which includes more than one variety, exhibits high values in certain region, it is maybe related to the

characteristics of that area. This means that either the plants in that specific location grow faster or that they have developed stress due to their proximity, which results in similar spectral information. To validate the results obtained from this study, in addition to the comparisons made using different algorithms and to ensure consistent conclusions, it would be quite essential to have conducted ground measurements, as well on the leaves of the vines in some areas.

As the appearance of diseases is of greater concern to farmers regarding vineyards, it is important to predict or detect diseases or nutrient deficiencies as soon as possible, with the aim to prevent damage or crop loss. Vegetation indices are a way to consider the spectral bands with the most significant information for identifying plant characteristics. The indices used in this study mainly identified the chlorophyll content in vine leaves, determined whether the plant is healthy, indicated areas of stress, or aimed to highlight regions with denser canopy, among other factors. These are some of the basic characteristics that can be detected using multispectral images. However, for deeper analysis and identification of specific substances in plant leaves, the spectral information provided by the five spectral bands of multispectral images is not sufficient. Hyperspectral images are needed to capture values in the near-infrared and shortwave infrared regions of the spectrum at very narrow ranges, in order to identify values that can represent more distinctly the absence or presence of a substance. Therefore, in a continuation of this research, it would be beneficial to differentiate the varieties using the spectral information obtained from a hyperspectral camera. With the large number of spectral bands offered by a hyperspectral camera, it becomes possible to examine very specific ranges of values, and it would be important to identify those values that can differentiate the varieties with each other.

The knowledge of the different varieties' characteristics is very important for the farmers due to various reasons. First, they can optimize the yield of the specific variety. Different grapevines may have varying yields, time of ripening, disease resistance, and other growth characteristics. By accurately detecting and identifying these characteristics, farmers can make specific decisions about planting, pruning, and harvesting, ultimately optimizing their yield and maximizing productivity. Moreover, it is possible for certain varieties to exhibit different levels of resistance specific diseases or pests. By detecting and differentiating these characteristics, farmers can implement the proper disease management strategies, to minimize crop losses and protect their vineyards. Another important issue is that when a farmer plants grapevines in a new field, there is need to select the most suitable variety for their specific growing conditions in order to adapt their cultivation practices accordingly. For these reasons it is crucial to know and differentiate the characteristics of each variety and this information to be used for the best utilization of the vineyard.

In conclusion, to extract information about the growth of grapevines or any other plant based on their variety, several validations and complex analysis need to be conducted, taking into consideration multiple factors. This study was a starting point

for gathering information about grapevines in relation to different varieties. Nevertheless, there are still many potential applications and further scientific research in farming sector that can be carried out in the future, including the introduction of hypespectral cameras, data fusion with in situ and satellite data and imagery, as well as deep and machine learning, time series and multi-dimensional vector analysis along with increasing computational capabilities offered by the continual advancement of technology.

6. References

Agisoft, L. (2015). *AgiSoft PhotoScan User Manual* (Version 1.1. 6 Professional Edition).

Alessandrini, M., Calero Fuentes Rivera, R., Falaschetti, L., Pau, D., Tomaselli, V., & Turchetti, C. (2021). A grapevine leaves dataset for early detection and classification of esca disease in vineyards through machine learning. *Data in Brief*, *35*. https://doi.org/10.1016/j.dib.2021.106809

Al-Saddik, H., Simon, J. C., & Cointault, F. (2019). Assessment of the optimal spectral bands for designing a sensor for vineyard disease detection: the case of 'Flavescence dorée.' *Precision Agriculture*, *20*(2), 398–422. https://doi.org/10.1007/s11119-018-9594-1

Ampatzidis, Y., Cruz, A., Pierro, R., Materazzi, A., Panattoni, A., de Bellis, L., & Luvisi, A. (2020). Vision-based system for detecting grapevine yellow diseases using artificial intelligence. *Acta Horticulturae*, *1279*, 225–230. https://doi.org/10.17660/ActaHortic.2020.1279.33

Anastasiou, E., Castrignanò, A., Arvanitis, K., & Fountas, S. (2019). A multi-source data fusion approach to assess spatial-temporal variability and delineate homogeneous zones: A use case in a table grape vineyard in Greece. *Science of The Total Environment*, *684*, 155–163. https://doi.org/10.1016/j.scitotenv.2019.05.324

Arab, S. T., Noguchi, R., Matsushita, S., & Ahamed, T. (2021). Prediction of grape yields from time-series vegetation indices using satellite remote sensing and a machine-learning approach. *Remote Sensing Applications: Society and Environment*, *22*. https://doi.org/10.1016/j.rsase.2021.100485

Ballesteros, R., Intrigliolo, D. S., Ortega, J. F., Ramírez-Cuesta, J. M., Buesa, I., & Moreno, M. A. (2020). Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques. *Precision Agriculture*, *21*(6), 1242–1262. https://doi.org/10.1007/s11119-020-09717-3

Chancia, R., Bates, T., Heuvel, J. Vanden, & van Aardt, J. (2021). Assessing grapevine nutrient status from unmanned aerial system (UAS) hyperspectral imagery. *Remote Sensing*, *13*(21). https://doi.org/10.3390/rs13214489

Chen, M., Brun, F., Raynal, M., & Makowski, D. (2020). Forecasting severe grape downy mildew attacks using machine learning. *PLoS ONE*, *15*(3). https://doi.org/10.1371/journal.pone.0230254 Clevers

Chivasa, W., Mutanga, O., & Biradar, C. (2020). UAV-based multispectral phenotyping for disease resistance to accelerate crop improvement under changing climate conditions. *Remote Sensing*, *12*(15). https://doi.org/10.3390/RS12152445

Clevers, J. G. P. W. (1999). The use of imaging spectrometry for agricultural applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 299–304.

COOMBE, B. G. (1995). Growth Stages of the Grapevine: Adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research*, *1*(2), 104–110. https://doi.org/10.1111/j.1755-0238.1995.tb00086.x

Cruz, A., Ampatzidis, Y., Pierro, R., Materazzi, A., Panattoni, A., De Bellis, L., & Luvisi, A. (2019). Detection of grapevine yellows symptoms in Vitis vinifera L. with artificial intelligence.

Computers and Electronics in Agriculture, *157*, 63–76. https://doi.org/10.1016/j.compag.2018.12.028

Fuentes, S., Tongson, E. J., De Bei, R., Viejo, C. G., Ristic, R., Tyerman, S., & Wilkinson, K. (2019). Non-invasive tools to detect smoke contamination in grapevine canopies, berries and wine: A remote sensing and machine learning modeling approach. *Sensors (Switzerland)*, *19*(15). https://doi.org/10.3390/s19153335

Fuentes, S., Viejo, C. G., Hall, C., Tang, Y., & Tongson, E. (2021). Berry cell vitality assessment and the effect on wine sensory traits based on chemical fingerprinting, canopy architecture and machine learning modelling. *Sensors*, *21*(21). https://doi.org/10.3390/s21217312

Garcia, L. C., Concepcion, R., Dadios, E., & Dulay, A. E. (2022). Spectro-morphological Feature-based Machine Learning Approach for Grape Leaf Variety Classification. *2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management, HNICEM 2022*. https://doi.org/10.1109/HNICEM57413.2022.10109536

Giovos, R., Tassopoulos, D., Kalivas, D., Lougkos, N., & Priovolou, A. (2021). Remote Sensing Vegetation Indices in Viticulture: A Critical Review. *Agriculture*, *11*(5), 457. https://doi.org/10.3390/agriculture11050457

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, *58*(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A. A., Keydan, G. P., & Merzlyak, M. N. (2006). Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Research Letters*, *33*(11). https://doi.org/10.1029/2006GL026457

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophysical Research Letters*, *32*(8), 1–4. https://doi.org/10.1029/2005GL022688

Hernández, I., Gutiérrez, S., Ceballos, S., Iñíguez, R., Barrio, I., & Tardaguila, J. (2021). Artificial intelligence and novel sensing technologies for assessing downy mildew in grapevine. *Horticulturae*, *7*(5). https://doi.org/10.3390/horticulturae7050103

Higgins J., Thomas J., Chandler J., Cumpston M., Li T., Page M., Welch V. (2022). Cochrane Handbook for Systematic Reviews of Interventions version 6.3.

Jiang, Z.; Huete, A. R.; Didan, K.; Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*, 3833–3845.

Jones, G. V. (2011). SUSTAINABLE VINEYARD DEVELOPMENTS WORLDWIDE. In *WWW.INFOWINE.COM-INTERNET JOURNAL OF ENOLOGY AND VITICULTURE* (Vol. 7, Issue 3).

Jordan, C. F. (1969). Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology*, *50*, 636–666.

Karakizi, C., Oikonomou, M., & Karantzalos, K. (2016). Vineyard Detection and Vine Variety Discrimination from Very High Resolution Satellite Data. *Remote Sensing*, *8*(3), 235. https://doi.org/10.3390/rs8030235

Karnieli, A., Kuafman, Y. J., Remer, L., & Wald, A. (2001). AFRI-aerosol free vegetation index. . *Remote Sensing of Environment*, *77*, 10–21.

Kasimati, A., Espejo-García, B., Darra, N., & Fountas, S. (2022). Predicting Grape Sugar Content under Quality Attributes Using Normalized Difference Vegetation Index Data and Automated Machine Learning. *Sensors*, *22*(9), 3249. https://doi.org/10.3390/s22093249

Kasimati, A., Espejo-Garcia, B., Vali, E., Malounas, I., & Fountas, S. (2021). Investigating a Selection of Methods for the Prediction of Total Soluble Solids Among Wine Grape Quality Characteristics Using Normalized Difference Vegetation Index Data From Proximal and Remote Sensing. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.683078

Kaufman Y J, Wald A E, Remer L A, Gao B, Li R, and Flynn L (1997). The MODIS 2.1-mm band correlation with visible reflectance for use in remote sensing of aerosol. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 1286–1298.

Kaur, P., Harnal, S., Tiwari, R., Upadhyay, S., Bhatia, S., Mashat, A., & Alabdali, A. M. (2022). Recognition of Leaf Disease Using Hybrid Convolutional Neural Network by Applying Feature Reduction. *Sensors*, *22*(2). https://doi.org/10.3390/s22020575

Kerkech, M., Hafiane, A., & Canals, R. (2020). Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Computers and Electronics in Agriculture*, *174*. https://doi.org/10.1016/j.compag.2020.105446

Kirti, K., Rajpal, N., & Yadav, J. (2021). Black measles disease identification in grape plant (vitis vinifera) using deep learning. *Proceedings - IEEE 2021 International Conference on Computing, Communication, and Intelligent Systems, ICCCIS 2021*, 97–101. https://doi.org/10.1109/ICCCIS51004.2021.9397205

Kisekka, I., Peddinti, S. R., Kustas, W. P., McElrone, A. J., Bambach-Ortiz, N., McKee, L., & Bastiaanssen, W. (2022). Spatial–temporal modeling of root zone soil moisture dynamics in a vineyard using machine learning and remote sensing. *Irrigation Science*, *40*(4–5), 761–777. https://doi.org/10.1007/s00271-022-00775-1

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, *51*(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, *62*(10), e1–e34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Loggenberg, K., Strever, A., Greyling, B., & Poona, N. (2018). Modelling water stress in a Shiraz vineyard using hyperspectral imaging and machine learning. *Remote Sensing*, *10*(2). https://doi.org/10.3390/rs10020202

Lowe, D. G. (1999). *Object Recognition from Local Scale-Invariant Features*.

Maimaitiyiming, M., Sagan, V., Sidike, P., & Kwasniewski, M. T. (2019). Dual activation function-based Extreme Learning Machine (ELM) for estimating grapevine berry yield and quality. *Remote Sensing*, *11*(7). https://doi.org/10.3390/rs11070740

Marques, P., Pádua, L., Adão, T., Hruška, J., Sousa, J., Peres, E., Sousa, J. J., Morais, R., & Sousa, A. (2019). Grapevine varieties classification using machine learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11804 LNAI*, 186–199. https://doi.org/10.1007/978-3-030-30241-2_17

Mazzia, V., Comba, L., Khaliq, A., Chiaberge, M., & Gay, P. (2020). UAV and machine learning based refinement of a satellite-driven vegetation index for precision agriculture. *Sensors (Switzerland)*, *20*(9). https://doi.org/10.3390/s20092530

Miranda, M., Zabawa, L., Kicherer, A., Strothmann, L., Rascher, U., & Roscher, R. (2022). Detection of Anomalous Grapevine Berries Using Variational Autoencoders. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.729097

Mirzaei, M., Marofi, S., Abbasi, M., Solgi, E., Karimi, R., & Verrelst, J. (2019). Scenario-based discrimination of common grapevine varieties using in-field hyperspectral data in the western of Iran. *International Journal of Applied Earth Observation and Geoinformation*, *80*, 26–37. https://doi.org/10.1016/j.jag.2019.04.002

Moghimi, A., Pourreza, A., Zuniga-ramirez, G., Williams, L. E., & Fidelibus, M. W. (2020). A novel machine learning approach to estimate grapevine leaf nitrogen concentration using aerial multispectral imagery. *Remote Sensing*, *12*(21), 1–21. https://doi.org/10.3390/rs12213515

Nguyen, C., Sagan, V., Maimaitiyiming, M., Maimaitijiang, M., Bhadra, S., & Kwasniewski, M. T. (2021). Early detection of plant viral disease using hyperspectral imaging and deep learning. *Sensors (Switzerland)*, *21*(3), 1–23. https://doi.org/10.3390/s21030742

Ohana-Levi, N., Bahat, I., Peeters, A., Shtein, A., Netzer, Y., Cohen, Y., & Ben-Gal, A. (2019). A weighted multivariate spatial clustering model to determine irrigation management zones. *Computers and Electronics in Agriculture*, *162*, 719–731. https://doi.org/10.1016/j.compag.2019.05.012

Padua, L., Adao, T., Hruska, J., Guimaraes, N., Marques, P., Peres, E., & Sousa, J. J. (2020). Vineyard Classification Using Machine Learning Techniques Applied to RGB-UAV Imagery. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 6309–6312. https://doi.org/10.1109/IGARSS39084.2020.9324380

Peng, X., Chen, D., Zhou, Z., Zhang, Z., Xu, C., Zha, Q., Wang, F., & Hu, X. (2022). Prediction of the Nitrogen, Phosphorus and Potassium Contents in Grape Leaves at Different Growth Stages Based on UAV Multispectral Remote Sensing. *Remote Sensing*, *14*(11). https://doi.org/10.3390/rs14112659

Peng, Y., Zhao, S., & Liu, J. (2021). Fused deep features-based grape varieties identification using support vector machine. *Agriculture (Switzerland)*, *11*(9). https://doi.org/10.3390/agriculture11090869

Pôças, I., Tosin, R., Gonçalves, I., & Cunha, M. (2020). Toward a generalized predictive model of grapevine water status in Douro region from hyperspectral data. *Agricultural and Forest Meteorology*, *280*. https://doi.org/10.1016/j.agrformet.2019.107793

Reyes Rojas, L. A., Moletto-Lobos, I., Corradini, F., Mattar, C., Fuster, R., & Escobar-Avaria, C. (2021). Determining actual evapotranspiration based on machine learning and sinusoidal approaches applied to thermal high-resolution remote sensing imagery in a semi-arid ecosystem. *Remote Sensing*, *13*(20). https://doi.org/10.3390/rs13204105

Romero, M., Luo, Y., Su, B., & Fuentes, S. (2018). Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. *Computers and Electronics in Agriculture*, *147*, 109–117. https://doi.org/10.1016/j.compag.2018.02.013

Rossi, V., & Caffi, T. (2007). Effect of water on germination of Plasmopara viticola oospores. *Plant Pathology*, *56*(6), 957–966. https://doi.org/10.1111/j.1365-3059.2007.01685.x

Rother, E. T. (2007). Revisão sistemática X revisão narrativa. *Acta Paulista de Enfermagem*, *20*(2), v–vi. https://doi.org/10.1590/S0103-21002007000200001

Rouse JW, H. R. S. J. D. D. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publ*, *351*, 309.

Rouzet, J., & Jacquin, D. (2003). Development of overwintering oospores of *Plasmopara viticola* and severity of primary foci in relation to climate*. *EPPO Bulletin*, *33*(3), 437–442. https://doi.org/10.1111/j.1365-2338.2003.00670.x

Sassu, A., Gambella, F., Ghiani, L., Mercenaro, L., Caria, M., & Pazzona, A. L. (2021). Advances in Unmanned Aerial System Remote Sensing for Precision Viticulture. *Sensors*, *21*(3), 956. https://doi.org/10.3390/s21030956

Silva, D. M., Bernardin, T., Fanton, K., Nepaul, R., Pádua, L., Sousa, J. J., & Cunha, A. (2021). Automatic detection of Flavescense Dorée grapevine disease in hyperspectral images using machine learning. *Procedia Computer Science*, *196*, 125–132. https://doi.org/10.1016/j.procs.2021.11.081

Tang, J., Woods, M., Cossell, S., Liu, S., & Whitty, M. (2016). Non-Productive Vine Canopy Estimation through Proximal and Remote Sensing. *IFAC-PapersOnLine*, *49*(16), 398–403. https://doi.org/10.1016/j.ifacol.2016.10.073

Tang, Z., Jin, Y., Alsina, M. M., McElrone, A. J., Bambach, N., & Kustas, W. P. (2022). Vine water status mapping with multispectral UAV imagery and machine learning. *Irrigation Science*, *40*(4–5), 715–730. https://doi.org/10.1007/s00271-022-00788-w

Unajan, M. C., & Gerardo, B. D. (2019). Recognizing cassava variety using artificial neural network with otsu algorithm for image segmentation. *International Journal of Recent Technology and Engineering*, *8*(2), 131–135. https://doi.org/10.35940/ijrte.A1917.078219

Vincini, M., Frazzi, E., & D'Alessio, P. (2008). A broad-band leaf chlorophyll vegetation index at the canopy scale. *Precision Agriculture*, *9*(5), 303–319. https://doi.org/10.1007/s11119-008-9075-z

Wójtowicz, M., Wójtowicz, A., & Piekarczyk, J. (2016). Application of remote sensing methods in agriculture. In *Communications in Biometry and Crop Science* (Vol. 11).

Wu, Q., Xu, L., Zou, Z., Wang, J., Zeng, Q., Wang, Q., Zhen, J., Wang, Y., Zhao, Y., & Zhou, M. (2022). Rapid nondestructive detection of peanut varieties and peanut mildew based on hyperspectral imaging and stacked machine learning models. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.1047479

Zhang, P., & Li, D. (2022). YOLO-VOLO-LS: A Novel Method for Variety Identification of Early Lettuce Seedlings. *Frontiers in Plant Science*, *13*. https://doi.org/10.3389/fpls.2022.806878

Zhou, X., Yang, L., Wang, W., & Chen, B. (2021). Uav data as an alternative to field sampling to monitor vineyards using machine learning based on uav/sentinel-2 data fusion. *Remote Sensing*, *13*(3), 1–23. https://doi.org/10.3390/rs13030457