HAROKOPIO UNIVERSITY

SCHOOL OF ENVIRONMENT,
GEOGRAPHY AND APPLIED ECONOMICS
DEPARTMENT OF GEOGRAPHY

POSTGRADUATE PROGRAMME:
APPLIED GEOGRAPHY AND SPATIAL PLANNING
COURSE OF GEOINFORMATICS

Exploring the synergy of EnMAP hyperspectral satellite data with machine learning for LULC mapping in a Mediterranean setting

Christina Lekka

Master's Thesis

Athens, 2023

ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΠΕΡΙΒΑΛΛΟΝΤΟΣ, ΓΕΩΓΡΑΦΙΑΣ ΚΑΙ
ΕΦΑΡΜΟΣΜΕΝΩΝ ΟΙΚΟΝΟΜΙΚΩΝ
ΤΜΗΜΑ ΓΕΩΓΡΑΦΙΑΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ:
ΕΦΑΡΜΟΣΜΕΝΗ ΓΕΩΓΡΑΦΙΑ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗ ΤΟΥ ΧΩΡΟΥ
ΚΑΤΕΥΘΥΝΣΗ ΄ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ΄

Διερεύνηση της συνέργειας των υπερφασματικών δορυφορικών δεδομένων EnMAP με τη μηχανική μάθηση για χαρτογράφηση της χρήσης και κάλυψης γης σε μεσογειακό περιβάλλον

Χριστίνα Λέκκα

Διπλωματική Εργασία

Αθήνα, 2023

HAROKOPIO UNIVERSITY

SCHOOL OF ENVIRONMENT,
GEOGRAPHY AND APPLIED ECONOMICS
DEPARTMENT OF GEOGRAPHY

POSTGRADUATE PROGRAMME:
APPLIED GEOGRAPHY AND SPATIAL PLANNING
COURSE OF GEOINFORMATICS

Examination committee

Petropoulos, George P. (Supervisor)
Assistant Professor of Geoinformation,
Department of Geography, Harokopio University

Faka, Antigoni
Assistant Professor of Geography,
Department of Geography, Harokopio University

Ferentinos, Konstantinos
Senior Researcher
Department of Agriculture Engineering
Soil & Water Resources Institute-HAO

Ethics and Copyright Statement

hereby declare that:

1) I am the owner of the intellectual rights of this original work and to the best of my knowledge, my work does not insult persons, nor does it offend the intellectual rights of third parties.

2) I accept that Library and Information Centre of Harokopio University may, without changing the content of my work, make it available in electronic form through its Digital Library, copy it in any medium and / or any format and hold more than one copy for maintenance and safety purposes.

3) I have obtained, where necessary, permission from the copyright owners to use any third-party copyright material reproduced in the master thesis while the corresponding material is visible in the submitted work.

Acknowledgements

I would like to express my sincere gratitude to Dr. George P. Petropoulos, for his excellent supervision, continued support and scientific guidance throughout every stage of my thesis. His expertise and valuable feedback have greatly influenced the outcome of this research.

Special thanks also to Spyridon Detsikas, doctoral student of Geography department, for the valuable insights and exceptional support at any technical issue, which have greatly contributed to the success of this research.

Many thanks to all my professors throughout the Master's program in Geoinformatics, in the Department of Geography at Harokopio University, from whom I gained invaluable knowledge throughout the courses of my studies.

Table of Contents

Abstract in Greek

Η υπερφασματική τηλεπισκόπηση ή φασματοσκοπία παρέχει υψηλής ποιότητας φασματικές πληροφορίες για τις ιδιότητες της επιφάνειας της γήινης επιφάνειας και τα οικολογικά συστήματα παγκοσμίως. Η φασματοσκοπία απεικόνισης βελτιώνει την αναγνώριση των χαρακτηριστικών της γης ευνοώντας πλήθος εφαρμογών, συμπεριλαμβανομένης καταγραφής και χαρτογράφησης της χρήσης και κάλυψης γης. Ακριβείς πληροφορίες σε τακτικά χρονικά διαστήματα για τις χρήσεις και τις αλλαγές της επιφάνειας της γης είναι κρίσιμες για την παγκόσμια παρακολούθηση και διαχείριση των οικοσυστημάτων. Τα επιχειρησιακά προϊόντα θεματικών χαρτών κάλυψης γης, επί του παρόντος είναι διαθέσιμα σε παγκόσμια ή εθνική κλίμακα και προέρχονται από τη χρήση πολλαπλών παθητικών οπτικών αισθητήρων. Ωστόσο, η χαμηλή χωρική ανάλυση που εξακολουθούν να παρέχουν αυτά τα προϊόντα τα καθιστά ακατάλληλα για πολυάριθμες εφαρμογές σε κλίμακα περιφερειακής και τοπικής ενότητας.

Η διαθεσιμότητα δεδομένων ΕΟ από σύγχρονες δορυφορικές αποστολές προσφέρει μοναδικές ευκαιρίες για την αντιμετώπιση αυτού του περιορισμού. Το Πρόγραμμα Περιβαλλοντικής Χαρτογράφησης και Ανάλυσης (EnMAP) είναι μια δορυφορική υπερφασματική αποστολή της Γερμανίας με στόχο την παρακολούθηση και τον χαρακτηρισμό του γήινου περιβάλλοντος σε παγκόσμια κλίμακα. Το πρόγραμμα EnMAP προορίζεται να γεφυρώσει το χάσμα παρέχοντας πλούσιες λεπτομερείς φασματικές πληροφορίες στο VNIR και SWIR εύρος στην περιοχή μεγάλης κλίμακας με ευρεία χρονική κάλυψη και υψηλή χωρική ανάλυση. Αξιοποιώντας τα υψηλής ποιότητας δεδομένα και ανοικτής πρόσβασης στην επιστημονική κοινότητα, αποκαλύπτεται ένα μεγάλο δυναμικό σε ένα ευρύ φάσμα οικολογικών και περιβαλλοντικών εφαρμογών, όπως οι ακριβείς και ενημερωμένοι θεματικοί χάρτες κάλυψης-χρήσης γης.

Στο πλαίσιο των παραπάνω, η παρούσα μελέτη αποτελεί μία από τις πρώτες που διερευνούν τις δυνατότητες του δορυφόρου ENMAP στο πλαίσιο της χαρτογράφησης των χρήσεων και της κάλυψης γης, με σκοπό να διερευνήσει τα πλεονεκτήματα του. Στόχος αποτελεί η διερεύνηση της αποτελεσματικότητας των πιο δημοφιλών αλγορίθμων μηχανικής μάθησης επιβλεπόμενης ταξινόμησης, όπως ο αλγόριθμος Support Vector Machines και Random Forest, χρησιμοποιώντας ένα σύνολο υπερφασματικών δεδομένων του δορυφορικού προγράμματος EnMAP. Ως περίπτωση μελέτης χρησιμοποιείται ένας τυπικό Μεσογειακό τοπίο.

Οι θεματικοί χάρτες που προέκυψαν χρησιμοποιώντας τους αλγόριθμους μηχανικής μάθησης SVM και RF αξιολογήθηκαν ως προς την απόδοσή τους χρησιμοποιώντας βασικές στατιστικές αξιολόγησης. Επιπλέον, διεξήχθη συγκριτική ανάλυση των αποτελεσμάτων χρησιμοποιώντας τον στατιστικό έλεγχο σημαντικότητας McNemar's. Τα αποτελέσματα ανέδειξαν την υπεροχή της μεθόδου SVM έναντι RF, όπου προέκυψαν υψηλότερη τιμή της συνολικής ακρίβειας, με 90,5% έναντι 87,5%, αντίστοιχα. Η υπεροχή της ακρίβειας για την SVM επιβεβαιώθηκε περαιτέρω από το στατιστικό έλεγχο McNemar's.

Τα ευρήματα της παρούσας μελέτης ανέδειξαν ότι τα υπερφασματικά δεδομένα του EnMAP διαθέτουν μεγάλο δυναμικό στον τομέα της χαρτογράφησης χρήσης/κάλυψης γης και αναμένεται να παρέχουν πολύτιμες πληροφορίες για την περαιτέρω αξιολόγηση των συνόλων δεδομένων του EnMAP σε σχετικές εφαρμογές.

Λέξεις-κλειδιά: EnMAP, χρήσεις γης και κάλυψη γης, μηχανική μάθηση, υπερφασματική τηλεπισκόπηση, φασματοσκοπία

Abstract

Hyperspectral remote sensing or imaging spectroscopy provides high-quality spectral information on terrestrial surface properties and ecology systems worldwide. Imaging spectroscopy enhances the identification of characteristics of derivative surface features advantaging various applications, including land use and land cover (LULC) mapping. Accurate and repetitive information on land use and its changes are crucial in monitoring and managing ecosystems globally. LULC operational products are currently available at global or national scales utilizing imaging data acquired from multiple optical sensors. Yet, the coarse spatial resolution these products are still provided makes them not suitable for numerous applications on regional and local scales.

The availability of EO data from contemporary satellite missions offers unique opportunities towards addressing this limitation. The Environmental Mapping and Analysis Program (EnMAP), is a German hyperspectral satellite mission aiming at monitoring and characterizing the Earth's environment on a global scale. EnMAP program is intended to bridge this gap by providing abundant detailed spectral information in visible - near infrared (VNIR) and shortwave infrared (SWIR) ranges within a large-scale area in wide temporal coverage and high spatial resolution. Taking advantage of high-quality and open-access data for the scientific community, great potential is revealed in a wide range of environmental applications, such as, i.e., accurate and up-to-date LULC thematic maps.

In the purview of the above, this thesis evaluates – to the authors' knowledge for the first time - the use of recently launched ENMAP in the context of LULC mapping, aiming at exploring the advantages of the hyperspectral EnMAP datasets. The overall objective of this study is to demonstrate the effectiveness of some of the most popular machine learning (ML) pixel-based algorithms, i.e., Support Vector Machines (SVMs) and Random Forest (RF), for image classification on EnMAP hyperspectral dataset. As a case study a typical Mediterranean landscape is used.

The thematic maps obtained using fine-tuned SVM and RF algorithms, provided comparable accuracy assessed using standard classification accuracy metrics. Further, a comparative analysis of results was conducted using McNemar's chi-square statistical significance testing. Results indicated that SVMs exhibited over RF showed higher overall accuracy, of 90.5% and 87.5%, respectively. The superiority of SVMs was further supported by McNemar's statistic test.

Findings of present study have shown that EnMAP datasets hold great potential in the field of land-cover and use mapping and are expected to provide valuable input for further evaluation of EnMAP datasets in relevant applications.

Keywords: Hyperspectral remote sensing, imaging spectroscopy, EnMAP, land cover and land use, machine learning

List of Figures

List of Tables

List of Abbreviations

FAO	Food and Agriculture Organization of the United Nations
EEA	European Environment Agency
EO	Earth Observation
RS	Remote Sensing
EnMAP	Environmental Monitoring and Analysis Program
DESIS	DLR Earth Sensing Imaging Spectrometer
LULC	Land use and land cover
SAC	Special Areas of Conservation
SPA	Special Protection Areas
ML	Machine Learning
DL	Deep Learning
GIS	Geographic Information System
RGB	Red Green Blue
NIR	Near infrared
VNIR	Visible – Near Infrared
SWIR	Shortwave infrared
VIS	Visible
TIR	Thermal Infrared
MSI	Multispectral Instrument
EM	Electromagnetic spectrum
LiDAR	Light Detection and Ranging
HSI	Hyperspectral Imagery
NASA	National Aeronautics and Space Administration
JPL	Jet Propulsion Laboratory
ISRO	Indian Space Research Organization
UAV	Unmanned Aerial Vehicle
GSD	Ground sample distance
SNR	Signal-to-noise ratio
ESA	European Space Agency
CLC	Corine Land Cover
LDA	Linear discriminant analysis
SVM	Support Vector Machines
RBF	Radial Basis Function
RF	Random Forest
ANN	Artificial Neural Network
KNN	K-Nearest Neighbours
DT	Decision Trees
SAM	Spectral Angle Mapper
CNN	Convolutional Neural Network
EGP	EOWEB Geoportal
DSDA	German Satellite Data Archive
IPS	Instrument Planning Subsystem
UTM	Universal Transverse Mercator
FTPS	File Transfer Protocol Secure
QL	Quality Level
BIL	Band Interleaved by Line
BIP	Band Interleaved Pixel
BSQ	Band Sequential
OA	Overall Accuracy
PA	Producer Accuracy
UA	User Accuracy
K_c	Cohen's Kappa

1. Introduction

Chapter 1

This chapter aims at providing a general overview to the background of this thesis, the research statement as well as the aim and objectives. Furthermore, the chapter closes presenting the overall thesis structure.

1.1. A general overview

The Earth's terrestrial surface and ecological systems are currently facing a significant pressure from a variety of sources, both natural causes and human factor activities (FAO, 2022). The rate of change in land use and land cover has accelerated significantly due to unregulated population growth and the increase of economic and industrial activities, such as expansion of urban areas, deforestation and mining operations, particularly in developing countries (Talukdar et al., 2020). Therefore, there is an urgent need for accurate, consistent and up-to-date information on global land cover and land use alterations which in turn serves as a critical asset in enabling effective and sustainable management of natural resources (FAO, 2022). Such information plays a crucial role in informing policy-makers for guiding land use practices and strategic planning efforts aimed at achieving sustainable development and environmental sustainability goals (Estoque, 2020; Giuliani et al., 2020; Andries et al., 2022).

The rapidly growing advances in the field of Earth Observation, entailed a vast range of remote sensing platforms equipped with passive sensors (multispectral or hyperspectral) and scanning sensors as well as microwave radiometers and active systems, to capture electromagnetic radiation reflected or emitted by the earth's surface (Jensen, 2015; Toth & Jóźków et al., 2016; Fu et al., 2019). For passive sensors, multispectral or hyperspectral instruments mounted on airborne or spaceborne satellite platforms and function by recording electromagnetic radiation reflected or emitted from earths terrain. Multispectral systems are capable of capturing energy across several distinct bands of the electromagnetic spectrum, while hyperspectral instruments are capable of recording energy across hundreds of narrow, continuous bands within the electromagnetic spectrum (Zhu et al., 2018). By employing various digital image processing techniques in remotely-sensed data, detailed information can be obtained about the Earth's surface features, such as topography, vegetation structure and land use/cover, supporting earth monitoring at different scales and wavelengths. Applications of Earth Observation (EO) in remote sensing field provides a global perspective for the material of the Earth’s surface (Winkler et al., 2021; Zhao et al., 2022).

Advanced techniques in imaging spectroscopy or hyperspectral imaging (HSI) provide a wide range of image processing techniques which enable understanding better the Earths’ surface features and their spatial distribution. This information collected through the use of remote sensing airborne and spaceborne platforms, corresponds to data with numerous contiguous bands in the electromagnetic spectrum (Buckingham & Staenz et al., 2008; Rast et al., 2019). The availability of such powerful data provides detailed information about the reflectance characteristics of the features at multiple wavelengths. This information can be used to identify the unique spectral signatures of different feature properties, such as natural vegetation and forests, different type of crops, built-up areas and water bodies (Rasti et al., 2018). In recent years, advances in hyperspectral remote sensing technology and increasing computing power have made it possible to classify large areas of land cover and land use and perform fine-grained discrimination with high accuracy (Pandey et al., 2020; Akar & Gormus, 2021; Moharram et al., 2023). This has important applications in fields such as agriculture, urban planning and environmental monitoring where detailed information about the earth's surface benefits with various applications within the scientific community and policy makers (Singh et al., 2020; Weiss et al., 2020; Wellmann et al., 2020).

Imaging spectroscopy enhances the identification of characteristics of derivative surface features advantaging various applications, including detailed land use and land cover (LULC) mapping. Accurate and repetitive information on land cover and its changes are crucial for many post-analysis tasks in monitoring and managing ecosystems globally. Measuring and predicting changes in land use and land cover through quantitative analysis is the most effective way to manage and comprehend landscape transformations and natural resources (Verrelst et al., 2018; Lamine et al., 2018). There are several approaches that can be used for hyperspectral LULC classification, including traditional statistical and geospatial techniques and state-of-the-art machine learning algorithms and deep learning techniques (Lv & Wang et al., 2020; Mughees & Tao, 2021; Datta et al., 2022). The advancements in hyperspectral remote sensing, improved the image enhancement in terms of spectral and spatial information and better discrimination of surface materials in data acquired from hyperspectral sensors.

A wide range of methodological approaches can be employed to develop accurate and up-to-date LULC maps utilizing algorithms and statistical models to analyse large datasets, identify patterns, and predict land cover changes. The advantages of hyperspectral imaging coupled with technological advancements in the field of ML algorithms have boosted classification techniques and increased efficiency of derived classified thematic maps (Pandey et al., 2019). In recent years, ML-based techniques have made significant progress in the field of land-use and cover mapping, overcoming the limitations over conventional algorithms and emerging as a promising area of research. ML in supervised classification, such as pixel-based, and various ensemble methods, have demonstrated robust capabilities of algorithms and achieved exceptional accuracy in hyperspectral LULC applications.

1.2. Research statement

The spaceborne imaging spectroscopy satellite mission *Environmental Mapping and Analysis Program* (EnMAP) designed from a consortium of German institutions headed by the German Aerospace Center (DLR). The mission aiming to monitor and characterize the Earth’s surface and environment on a global scale by providing high spectral resolution hyperspectral imagery data with large spatial coverage, enabling a comprehensive view of the Earth's surface (Guanter et al., 2015). EnMAP mission endorses a wide range of EO and environmental applications, including identifying land cover/use, vegetation analysis and natural resources monitoring (Kaufmann et al., 2008; 2015). In contrast with previously launched hyperspectral spaceborne missions, acquiring data at medium or coarse resolution, the high quality in spatial and spectral characteristics of EnMAP derived hyperspectral data allows for a better discrimination of diversity in landscapes and improved characterization of the environment. Such information can be utilized in EO applications to elaborate LULC maps, enhance the performance and boost the classification accuracy (Yokoya et al., 2016; Rosentreter et al., 2017; Marcinkowska-Ochtyra et al., 2017). Overall, the design of EnMAP mission can support gap filling in space-based imaging spectrometer data for environmental and Earth observation applications, providing high-quality hyperspectral data that can be used to support a wide range of studies and decision-making processes (Chabrillat et al., 2022).

The launch of EnMAP mission was a significant step in comprehensive hyperspectral data coverage on a large-scale (Keller et al., 2017). The deployment of the EnMAP hyperspectral sensor into its operational phase has prompted the scientific community to investigate the potential benefits and capabilities for employing various advanced machine learning methods on high-quality, derived hyperspectral datacubes. However, the extraction of information about different land-cover attributes with similar characteristics in HSI image classification is a challenging task due to highly complexity of hyperspectral data, as well as the variability and diversity of the land cover classes. To tackle this issue, advanced ML algorithms have been employed in the analysis of these data. These algorithms offer powerful capabilities for extracting valuable information from high-dimensional and complex hyperspectral datasets, particularly in distinguishing land-cover attributes with similar characteristics. These advanced ML algorithms leverage the significance of EnMAP data and provide sophisticated techniques for classification and mapping. By harnessing the capabilities of these algorithms, researchers and practitioners can effectively overcome the variability of land cover classes in hyperspectral image classification. In this regard, the recent availability of EnMAP HSI (since April 2022) expands the potential for enhancing and improving operational EO land-cover products. The introduction of high spectral, spatial, and temporal resolution open-access data could contribute to enhanced mapping of global ecosystems and land change dynamics. These advancements open up new possibilities for accurate and detailed updates of land cover mapping, supporting various applications in discipline fields.

1.3. Aim & Objectives

In purview of the above, this thesis aims at exploring the potential of ML algorithms for LULC mapping applied on novel hyperspectral data, in order to identify high-performing algorithms that can overcome the limitations of traditional classification methods. The overall goal is to acesse the effectiveness of most popular ML pixel-based algorithms in EnMAP hyperspectral data through cost-effective tools and approaches for supporting accurate and efficient LULC mapping. More specifically, this thesis objectives are to: (i) assess the efficiency of widely-used ML algorithms i.e., Support Vector Machines (SVMs) and Random Forest (RF) for hyperspectral image classification, (ii) compare the performance of fine-tuned algorithms and evaluate their potential for practical applications in land cover mapping using EnMAP hyperspectral data as well as (iii) compare the results with operational EO LULC products, i.e., (ESAs' WorldCover). As a case study is used a typical Mediterranean setting over a heterogeneous landscape located in Greece. The methodology to be implemented is based on the synergistic use of ML techniques coupled with ENMAP imagery and ancillary data. The study was carried out using EnMAP Box-3, within open-source QGIS software and open-access remotely-sensed data sources, allowed for a cost-effective implementation of the analysis. To the best of authors' knowledge, the present thesis is one of the first of its kind attempting to explore the advantages of recently launched EnMAP hyperspectral satellite mission in the context of LULC mapping. This study emphasizes the strengths and capabilities of EnMAP datasets in combination with ML.

1.4. Thesis outline

This thesis is structured into seven chapters, as follows: *Chapter 1* functions as a preface and aims at providing a general overview to the background of this thesis, the research statement and objectives. *Chapter 2*, provides an introduction on fundamentals in hyperspectral remote sensing along with a representative review of existing literature on hyperspectral image classification methods and identify the key research gaps and challenges in the field. *Chapter 3* depicts the experimental site, including the case study and datasets employed in the study. In *Chapter 4*, the detailed methodological framework is presented. *Chapter 5* presents the results of the analysis while in *Chapter 6* the key findings and limitations are discussed. *Chapter 7* presents the concluding remarks and outlines potential future research directions.

2. Literature Review

Chapter 2

This chapter serves as a review of the existing literature related to the research topic. It will provide an introduction on principals of hyperspectral remote sensing along with a comprehensive overview of applied hyperspectral image classification methods as well as in identifying the key research gaps and challenges in the field linked to the present study objectives.

2.1. Fundamentals of imaging spectroscopy

EO is experiencing significant growth, with an expanding range of remote sensing platforms incorporating passive sensors such as optical cameras, scanning sensors, and microwave radiometers for capturing electromagnetic energy emitted or reflected by the Earth's surface (Zhao et al., 2022). The adequacy of technical specifications of a passive remote sensor in terms of spectral and spatial characteristics directly influences the quality and quantity of the derived data. Multispectral sensors collect reflected or emitted energy of target area in a limited number of discrete spectral bands, typically for tens, in a certain spectral range within the EM spectrum (Curran, 2016; Panuju et al., 2020). Multispectral instruments (MSI) mounted on remote sensors provide a cost-effective trade-off between spatial resolution and operational costs and a suitable balanced between spatial, spectral and temporal resolution, which can make them a practical choice for certain applications. Although, MSI are limited in their ability to capture detailed spectral information across multiple spectral bands, which in turn restrict their capability to cover a broader range of the EM (Jameel et al., 2020). Sensor characteristics, such as spectral resolution, can influence the ability of remote sensing data to detect features in complex environments and identify different vegetation species (Priyadarshini et al., 2019).

Conversely to multispectral, hyperspectral sensors are able to provide advanced image data up to hundreds of spectral narrow bands with higher spectral resolution, which enhance the quality and increase the quantity of retrieved information. In environmental and vegetation analysis low-resolution sensors may not be optimal to detect heterogeneity between different vegetated areas, while high-resolution sensors can capture detailed information in order to sufficiently discriminate different plant species (Marcello et al., 2018; Lu et al., 2019). The ability to differentiate between features that possess similar physical properties relies on the enhanced spectral resolution and the continuous wavelength range acquired the image. Higher spectral resolution enhances the capability to retrieve information on the physicochemical characteristics in soil and vegetation with greater precision (Thenkabail et al., 2016; 2018; Srivastava et al., 2020). Therefore, data acquired through hyperspectral sensors can be beneficial for a variety of tasks that require more detailed spectral information in field of agroforestry and relevant ecological applications. Yet, the operational cost remains higher than multispectral sensors and they can be challenging and time-consuming due to the high volume of data that must be processed (Pandey et al., 2020; Dubovik et al., 2021). For further insights into the principles of remote sensing and characteristics of different sensors, one can refer to relevant literature (e.g. Lillesand et al., 2014; Richards & Jia, 2013; Jensen, 2015; Curran, 2016; Richards, 2022).

Hyperspectral remote sensing, also known as imaging spectroscopy, assign to the technology that integrates imaging and spectroscopy techniques to capture spatial and spectral information in high resolution within. Spectroscopy refers to the study of the interaction between light and matter based on its wavelength, whether the light is emitted, reflected or scattered by a material i.e., solid, liquid, or gas (Rast et al., 2019). The imaging systems capture the spatial information while the spectroscopy functions by measuring the spectral reflectance or emission of light using numerous of narrow continuous spectral bands in a wide range of wavelengths ranging from visible – near infrared (NIR) to shortwave infrared (SWIR) in the EM (Srivastava et al., 2020). The amount of the spectral insights provided within hundreds of adjacent near spectral bands, is capable to detect and retrieve detailed information about subtle variations in the absorption or emission features of different attributes even if these attributes exhibit a degree of in physical properties similarity. In the domain of vegetation analysis, imaging spectroscopy enabled species discrimination and has played a crucial role in estimating chlorophyll content, detecting water stress, identifying nutrient deficiencies, and diagnosing disease symptoms (Maschler et al., 2018; Lassalle et al., 2021; Srivastava et al., 2021; Sethy et al., 2022) as well as for vegetation extraction in urban related studies (Petropoulos et al., 2015). Regarding the agricultural production, it has been utilized to identify crop species, detect nutrient deficiencies, and monitor overall crop health throughout various stages of growth (Thenkabail et al., 2018; Yao et al., 2018; Verrelst et al., 2019; Sanchez et al., 2020; Aneece & Thenkabail, 2021; 2022). Studies of soil chemical and physical properties also relies in imaging spectroscopy (Ben-Dor et al., 2009; Goswami et al., 2020; Milewski et al., 2022). Furthermore, in the field of mineral exploitation, imaging spectroscopy has proven valuable in identifying mineral types and mapping their distributions within mining sites (Gupta & Venkatesan, 2020; Kumar et al., 2020; Peyghambari & Zhang, 2021) as well as in geochemical analysis (Sun et al., 2019). HSI has also found significance in material identification (Chisense et al., 2012; Pandey & Tiwari 2020), as well as in applications related soil and vegetation metal contamination (Lamine et al., 2019). Overall, imaging spectroscopy has demonstrated its value as a powerful tool for scientific research and practical applications across multiple disciplines.

2.1.1. Hyperspectral remote sensors characteristics

The imaging spectroscopy systems developed since 1980s in both airborne and spaceborne platforms. Airborne sensors are typically mounted on aircraft that flown at low altitudes, providing high spatial resolution but limited coverage. Spaceborne sensors, on the other hand, orbit the Earth at high altitudes, providing wide-area coverage but at lower spatial resolution. Imaging spectroscopy systems operating in four primarily modes: pushbroom, whiskbroom, snapshot, and tunable filter, with pushbroom and whiskbroom to be the most commonly used modes. Additionally, there is a dispersive mode and a compressive sensing mode (Coulter et al., 2007). In pushbroom mode, a hyperspectral sensor uses a two-dimensional (2D) detector array to continuously capture a narrow, contiguous strip of the Earth's surface as the sensor platform moves forward. This mode works by scanning the Earth with a line of detectors that collect data in a range of wavelengths simultaneously. In this mode, the spatial resolution of the image is determined by the size of the individual detector of pixels, and the spectral resolution is determined by the number of spectral bands captured by the sensor. Pushbroom mode has the advantage of being able to capture wide swaths of imageries with a high spatial and spectral resolution. (Fig.2.1) However, it requires precise synchronization between the motion of the sensor platform and the scanning of the detectors and is therefore sensitive to motion and vibration disturbances. In whiskbroom mode, a hyperspectral sensor uses a one-dimensional (1D) linear detector array to scan across the surface in a series of narrow parallel lines. This mode works by sweeping the detector array across the surface of the Earth in a series of rows, collecting data in a range of wavelengths as it goes. In this mode, the spatial resolution of the image is determined by the size of the detector array and the number of rows scanned, and the spectral resolution is determined by the number of spectral bands captured by the sensor (Fig.2.2).

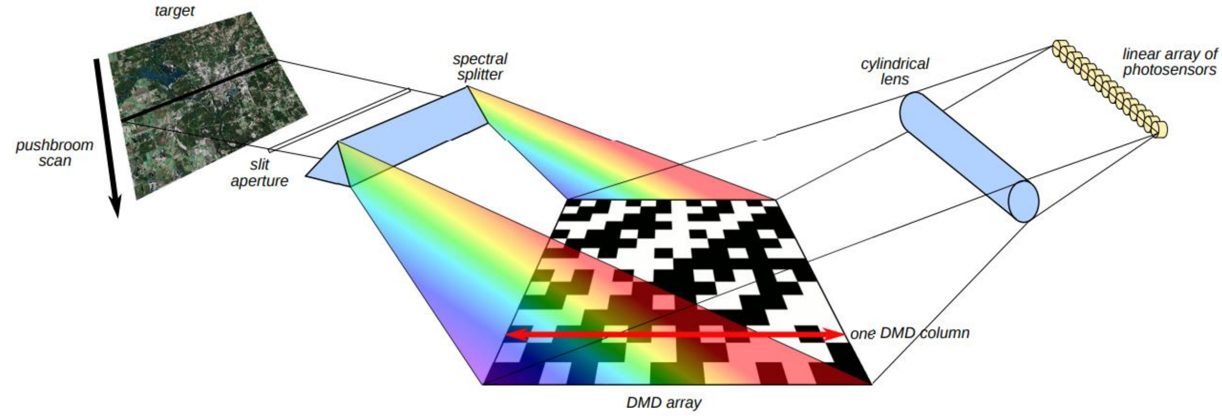


Figure 2.1. Comprehensive overview of hyperspectral sensor operating in pushbroom mode (Source: Fowler, 2014)

A well-known example of a hyperspectral airborne sensor operating in whiskbroom mode is the airborne AVIRIS with a whiskbroom-scan architecture, while there are currently no spaceborne hyperspectral imagers that operate in whiskbroom mode. In pushbroom mode, a 2D detector array captures a continuous strip of the Earth's surface as the sensor platform moves forward, while in whiskbroom mode, a 1D linear detector array scans across the surface in parallel lines. Pushbroom mode captures wider swaths of imagery with high spatial and spectral resolution but is sensitive to motion disturbances. Whiskbroom mode is less sensitive to motion disturbances but has a narrower swath width and limited spatial resolution.

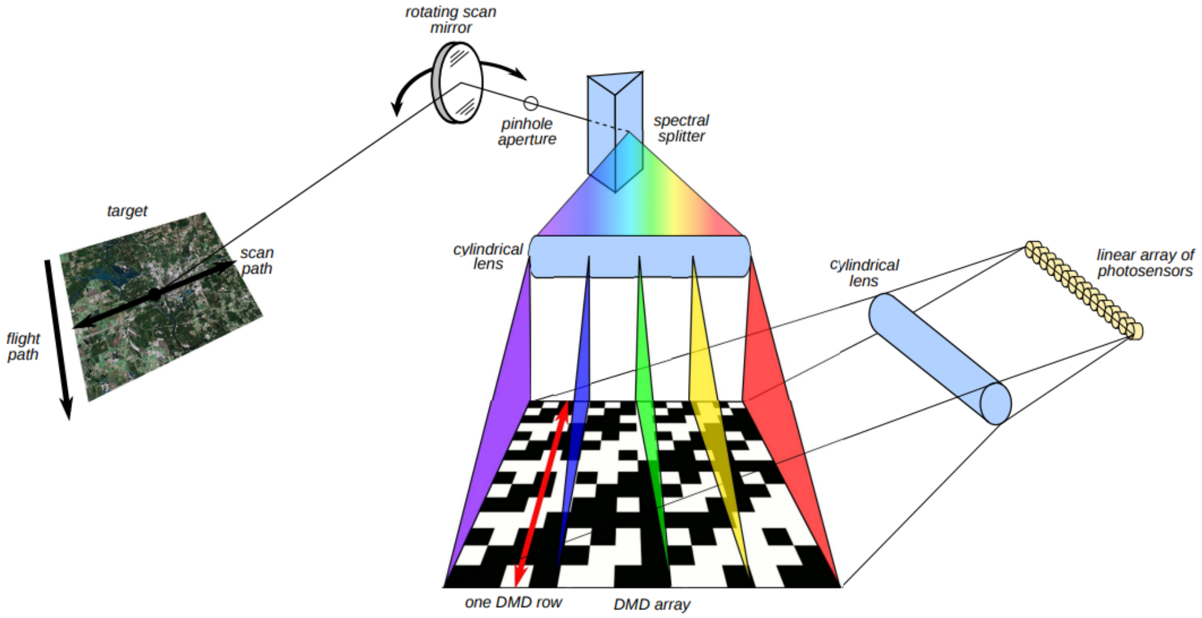


Figure 2.2. Comprehensive overview of hyperspectral sensor operating in whiskbroom mode (Source: Fowler, 2014)

The majority of hyperspectral sensors, including airborne and spaceborne, employ 2D detector arrays and functions with pushbroom scanning, i.e., the airborne CASI and FLI hyperspectral platforms, the spaceborne sensors Hyperion/EO-1, MERIS onboard ESA's ENVISAT, the Italian PRISMA, HISUI onboard Japan Experiment Module (JEM) on ISS, German DLR DESIS and EnMAP as well as future launched missions such as FLORIS for ESA's Fluorescence Explorer (FLEX). This configuration enables high spectral and spatial resolution imagery, as well as large coverage of the target area. A briefly overview of most well-known past, current, and future airborne and spaceborne hyperspectral missions is presented in Table

2.1. Major airborne and spaceborne missions are further discussed above (Coulter et al., 2007; Buckingham & Staenz, 2008; Pandey et al., 2020; Qian, 2021).

Table 2.1. Overview of major past, current and future airborne and spaceborne hyperspectral missions

Sensor	Platform	Spectral range (nm)	Spectral bands	Ground Sampling distance (m)	Swath Width (km)	Launch	Phase	Organization
APEX [a]	Dornier DO-228	380 – 2500	Up to 334 (default 114)	1 – 2	2 – 2.5	2011	✓	ESA, RSL, VITO
AVIRIS [a]	Aircraft ER-2	400 – 2500	224	20	11	1987	✕	NASA, JPL
CHIME [a]	CHIME	400 – 2500	210	20 – 30	120	2029 (planned)	✕	ESA, COM
CASI [a]	Aircraft flexible	418 – 826	288	2 – 5	1 – 5	1989	✕	ITRES, Canada
CHRIS [b]	PROBA-1	400 – 1050	62 (MODE 1) 18 (MODE 2-4) 37 (MODE 5)	34 (MODE 1) 17 (MODE 2-5) 5 viewing angles	13	2001	✓	ESA-UK
DESIS [b]	ISS	400 – 1000	235	30	30	2018	✓	LR Germany/GFZ
EnMAP [b]	German HS	420 – 2450	244	30	30	2022	✓	LR Germany/GFZ
FLORIS [b]	FLEX	500 –780	420	300	150	2024 (planned)	✕	ESA
HISUI [b]	JEM/ISS	400 – 2500	185	30	20	2019	✓	Japan
HICO [b]	ISS	350 – 1080	128	90	51	2009	✕	ONR/USA
Hyperion [b]	EO-1	400 – 2500	242	30	7.7	2000	✕	NASA
HyperScout [b]	GomX-4b	400 – 1000	45	50	200	2018	✓	ESA
HySI [b]	IMS-1	400 – 950	64	500	130	2008	✕	ISRO
HysIS [b]	IMS-2	400 – 2400	256	30	30	2018	✓	ISRO
HyspIRI [a]	HyspIRI	400 – 2500	220	60	150	2024 (planned)	✕	NASA-JPL
MERIS [b]	ENVISAT -1	390–1040	520 (transmit 15)	300	1150	2002	✕	ESA
PRISMA [b]	AVIO Italian launcher	400 – 2500	237	30	30	2019	✓	Italy's ASI
SHALOM [b]	MBT Space	400 – 2500	275	10	30	2025 (planned)	✕	ASI, ISA

a. Airborne
b. Spaceborne.
✓ Active
✕ Non-active

Airborne Hyperspectral Imaging missions

The first operational airborne spectrometer Airborne Visible/Infrared Imaging Spectrometer (**AVIRIS**), was designed by National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL) in California in the early ‘80s. It was set to flown aboard a NASA ER-2 aircraft in 1986, and was fully operated since 1989, under NASA contract. AVIRIS has a whiskbroom-scan architecture and uses a one-dimensional (1D) linear detector array to scan across the Earth's surface in a series of narrow parallel lines. It covers the entire spectrum from 0.4 nm to 2.45 nm with high spectral resolution providing data 224 spectral contiguous narrow bands and a high spatial resolution from a 20 km attitude over a swath of 12 km. Later in 90’s the AVIRIS was replaced by the Airborne Visible Infrared Imaging Spectrometer – Next Generation (**AVIRIS NG**) which was built upon AVIRIS with improvements in radiometric performance and spectral sampling interval of 5 nm higher. Later, the deployment of the AVIRIS induced further advancements in airborne hyperspectral mission with the Canadian hyperspectral missions **FLI** (1983), **CASI**, and **SFSI** (1992) in 1980s and 1990s, the Australian **HyMAP** (Barnsley et al., 1998) by the HyVista Corporation. Over the following years, the development of imaging spectroscopy led to the innovation of other well-known today airborne imaging mission such as the Airborne Prism Experiment (**APEX**) by the University of Zurich in Switzerland and VITO in Belgium (Schaepman et al., 2006). In the last decade, there has been a rise in the employment of hyperspectral cameras mounted on of Unmanned Airborne Vehicles (UAVs) and it is expected that they expect to play a key role in airborne imaging spectroscopy (Yan et al., 2019). This is due to the fact that the instruments used in imaging spectroscopy are becoming more compact and lightweight (Jia et al., 2022; Sousa et al., 2022).

Spaceborne Hyperspectral Imaging mission

Spaceborne hyperspectral imagers have been around since the early 1980s. The launched of the first generation spaceborne imaging spectrometers for scientific and operational purposes was made with the **Hyperion** aboard NASA’s Earth Observing-1 (**EO-1**) launched in November 2000 (https://data.nasa.gov/dataset/EO-1-Hyperion/). Hyperion was built by TRW within a fast-track schedule of development. With an initial estimate of 1-year lifespan, Hyperion provided calibrated spaceborne hyperspectral data for 16 years (decommissioned on 2017). Hyperion is a bushbroom hyperspectral imager, operating in a range of 400-2500 nm with a spectral resolution of 10 nm in VNIR and SWIR. The camera had a wide field of view, with a 7.65 km wide swath and a spatial resolution of 30 meters (Ungar et al., 2003). Over the past three decades, there have been numerous space missions that employed true dispersive element imaging spectrometers. These missions include some of the major well-established hyperspectral sensors, such as those described briefly above:

The Compact High-Resolution Imaging Spectrometer (**CHRIS**) is a hyperspectral imager developed by the UK's Sira Electro-Optics Ltd. and launched in 2001 as part of ESA's Project for On-Board Autonomy (PROBA) platform (https://earth.esa.int/eogateway/missions/proba-1/data). PROBA's primary objective was testing of innovations in spacecraft platform design, including autonomous operation. CHRIS has five operating modes, each with a different nominal number of bands, wavelength range, spectral bandwidth, and nominal ground sampling distance (GSD). The GSD decreases as the spectral bandwidth increases. The hyperspectral imager is using the push-broom technique that operates in the VNIR range (400-1050 nm). With a lifespan of over 20 years (as of June 2023), CHRIS simultaneously collects spectral data at five viewing angles and at two spatial resolutions of 17 m and 34 m, providing 19 and 62 spectral bands, respectively, over a typical image area of 13 km^2 (Cutter, 2008).

Within the same decade, the Medium-Resolution Imaging Spectrometer (**MERIS**), developed by other prominent sensor systems and the European Space Agency (ESA), was launched in 2002 aboard the ENVISAT satellite (https://meris-ds.eo.esa.int/oads/access/). It operated successfully for ten years before its termination and was designed by incorporating the technical specifications of prior hyperspectral airborne systems, i.e., Canadian and NASA. Employing a push-broom mode, MERIS provided hyperspectral images in the VNIR region. The instrument was intended to capture 520 spectral bands within the wavelength range of 390 to 1040 nm with a native instrument SSI of 1.25 nm. However, the downlink capability of the MERIS was limited and it only transmitted 15 channels, where each channel represented an average of eight to ten native spectral elements of the detector arrays. The GSD varied between 260 m at nadir and 390 m at swath extremities in the cross-track direction (Verstraete et al., 1999).

Lately, the DLR Earth Sensing Imaging Spectrometer (**DESIS**) has been developed as a push-broom hyperspectral imager that operates within the VNIR region, covering wavelengths between 400 and 1000 nm with a minimum SSI of 2.55 nm (https://geoservice.dlr.de/data-assets/). The instrument offers a ground swath width of 30 km and a GSD of 30 m. DESIS is mounted on the multi-user system for earth sensing, which was launched on June (29), 2018, as part of the SpaceX CRS-15 logistics flight to the ISS and installed on the exterior of the ISS on August (27), 2018. DESIS can operate in either static mode, which enables acquiring hyperspectral data to produce BRDF products or stereo images with 3° angle steps, or dynamic mode, which permits continuous observations of the same targets with ground motion compensation to enhance the signal-to-noise ratio (SNR) of the acquired hyperspectral data, with up to 1.5° change in viewing direction per second (Heiden et al., 2022).

Furthermore, the Italian is hyperspectral satellite **PRISMA**, was launched in 2019 with the aim of testing and developing technology for environmental monitoring and risk assesment (https://prismauserregistration.asi.it/). It consists of a hyperspectral imager and a panchromatic camera operating in pushbroom mode and covers spectral bands ranging from 400 to 1010 nm (VNIR) and from 920 to 2505 nm (SWIR). The hyperspectral imager provides images with 30 m GDS, 30 km swath width and spectral bands at an SSI of 12 nm. The PAN camera provides images at a 5 m spatial resolution, which are co-registered with the hyperspectral images to allow for image fusion and sharper spatial resolution (Shaik et al., 2023).

Recently, the Environmental Mapping and Analysis Program (**EnMAP**), a spaceborne imaging spectroscopy satellite mission designed from a consortium of German institutions headed by the German Aerospace Center (DLR), launched on April 1, 2022. EnMAP aims at monitoring and characterizing the Earth's surface and environment on a global scale by providing high spectral resolution hyperspectral imagery data in a large spatial coverage. The EnMAP mission will endorse a range of earth observation and environmental applications, such as mapping land use and cover and species discrimination. A detailed overview of the technical characteristics and specifications of the EnMAP hyperspectral satellite mission will be covered in section 3.2.1. of Chapter 3.

2.2. Remote Sensing image classification in land cover mapping

There is a growing awareness of the alterations of the world's natural resources which have caused significant implications for ecosystems, biodiversity and human society in a global scale (FAO, 2022). Accurate information on land use and land cover is essential for understanding changes in the way land is used (i.e., conversion of natural forested land for agricultural exploitation, urban expansion) and changes in the physical and biological characteristics of the land cover (i.e., deforestation, reforestation). Knowledge of changes in physical and biological properties due to human activities and natural processes allows for better understanding of LULC dynamics and landscape transformation (Gudo et al., 2022). The utilization remotely sensing data in image classification presents an effective solution to numerous challenges encountered in traditional methods for retrieving information for mapping purposes on land-cover and land-use. Terrestrial surveying or other common mapping methods; these traditional approaches tend to be time-consuming and costly, especially in large-scale areas (Al-Doski et al., 2020). Remotely sensed data offer the advantage of swiftly acquiring LULC information at diverse spatial and temporal scales and with a significantly reduced operational cost, covering large spatial extents in efficient time and easily storage. Quantitative assessment of spatiotemporal patterns and dynamics in landscapes is a prevalent research field within Earth observation, that have been widely used for analysis and decision-making processes (Navin & Agilandeeswari, 2020); in urban planning and industrial development (Heiden et al., 2012; Nisha & Anitha, 2022), and for management of natural resources (Chander et al., 2020).

There are several LULC products available on a global scale that have been developed using EO (Table 2.2). These products provide valuable information for management and planning purposes, environmental monitoring, and sustainable development. Among the most widely used LULC products, is GlobCover, a global land cover map produced by the European Space Agency (ESA) and the European Commission (EC). It utilizes data from the 300m MERIS sensor on board the ENVISAT satellite mission, covering the period from December 2004 to June 2006 (http://due.esrin.esa.int/page_globcover.php). Another notable product is ESA WorldCover, which provides global land cover products at 10 m resolution for the years 2020 and 2021 (https://esa-worldcover.org/en/data-access). These products are developed and validated in near-real time using data from Sentinel-1 and Sentinel-2 satellites (Zanaga et al., 2021; 2022). Additionally, the MODIS Land Cover Type/Dynamics product operating at a coarser resolution, incorporates five different land cover classification schemes and provides land cover information at yearly intervals (2001-2018) to study vegetation dynamics and seasonal cycles (https://lpdaac.usgs.gov/products/mcd12q1v006/). Last but not least, the European Corine Land Cover (CLC) provides consistent land cover information across EU. The initial establishment of the CLC took place between 1986 and 1998. New versions are released every six years, resulting in a total of five implemented versions (so far), i.e., CLC 1990, CLC 2000, CLC 2006, CLC 2012, CLC 2018, employed enhanced quality data and refined classification techniques (https://land.copernicus.eu/pan-european/corine-land-cover). These widely-used operational products that corresponding to the EU coverage, have been extensively employed by the scientific community for time series analysis and monitoring of land cover dynamics. Their open access policy further enhances their accessibility and usability for researchers in the relevant field.

Table 2.2. Characteristics of widely used land cover datasets operating across pan-European regions (Manakos & Braun, 2014)

Dataset	Spatial Resolution	Date	Input data	Land-cover classes	Classification Method	Organization
GlobCover	300 m	Since 2009	MERIS FR	22	(Un)supervised methods, spatiotemporal clustering	ESA
MODIS	500 m	Since 2001	MODIS	17	Supervised decision tree	NASA
ESA WorldCover	10 m	Since 2021	Sentinel-1 Sentinel-2	11	(Un)supervised methods	ESA
Corine Land Cover (CLC)	250 m 1:100 000	Since 1990	Landsat, SPOT, LISS III, RapidEye, Sentinel	Hierarchical (3 levels, 44 classes)	Computer assisted photo-interpretation	EEA

Principles of image classification

The accurate assignment of land-cover classes in remotely sensing data can be achieved through the utilization of suitable classification techniques that align with the nature of the research and the characteristics of the datasets (Abburu & Golla, 2015; Ghorbani et al., 2016). The primary approaches involve the use of either a hard classifier, where each pixel is allocated to a unique class, or a soft classifier, which provides a degree of similarity for each class. Soft classifiers, such as spectral mixture analysis, offer enhanced precision and detailed information, particularly in scenarios involving mixed pixels or coarse spatial resolution data that may introduce errors. Classifiers can be further classified into parametric and non-parametric methods depending on their assumptions about the data distribution. Parametric classifiers, including Maximum Likelihood and linear discriminant analysis (LDA), which rely on specific assumptions about the data distribution i.e., the assumptions of the Gaussian distribution (Al-doski et al., 2013). Conversely, non-parametric algorithms, such as Support Vector Machine (SVM), and Random Forest (RF), and Artificial Neural Networks (ANN), which do not rely on statistical parameters and can handle diverse data distributions without making assumptions about distribution (Maxwell et al., 2018; Dhingra & Kumar, 2019)

The most common classification methods typically involve two main categories: pixel-based classification and object-based classification; which in turn pixel-based classification can be broadly divided into two main sub-types: supervised, and unsupervised techniques (Pal, 2005; Pal and Mather, 2005). Object-based classification process is based upon spectrum information, geometry as well as colours of objects. In object-based classification pixels are grouping into segments and treated like objects, the creation of a segmented image is later used for the classification process (Ma et al., 2017). In contrast to object-based, pixel-based classification utilized the reflectance statistics and relies only on the spectral information of pixel, i.e., the value that correspond to each pixel, is used for the classification process (Jog & Dixit, 2016). Several deep learning network models have been efficiently employed for land-cover mapping in the field of remote sensing image processing (Li et al., 2019; Wang et al., 2021).

Pixel-based classification assigns pixels to represent land use/cover classes, whether training samples are used to define land cover classes, the classification process of a pixel is further categorized into supervised or unsupervised. In unsupervised classification, pixels are grouped together based on their reflectance properties using unsupervised clustering-based algorithms. This process utilizes the statistical information derived from the spectral image without prior definition of classes. Yet, in this type of classification it is necessary a prior knowledge of the characteristics of the region being classified. The analyst is responsible for managing and labelling each class in specific category after the classification process, to ensure that obtained results are meaningful. Examples of commonly used techniques for unsupervised classification include ISODATA, and the K-means clustering algorithm. In case of supervised classification, ancillary reference data are utilized to create a training set. This training set comprises spectral signatures associated with each class. The classification process is based on these spectral signatures which are used to train the classifier to assign pixels to specific classes during the classification process. Commonly used classifiers in supervised classification include SVM, Mahalanobis Distance, Naive Bayes, RF, Spectral Angle Mapper (SAM), Decision Tree (DT), Maximum Likelihood classifier (Al-doski et al., 2020).

These algorithms are commonly used and proved to be highly effective techniques for both supervised and unsupervised classification tasks. For further insights in methods and taxonomy of classification algorithms employed in satellite imageries the reader can refer to (Lu & Weng, 2007; Abburu & Golla, 2015; Dhingra & Kumar, 2019; Macarringue et al., 2022).

2.3. A review on image hyperspectral classification

The amount of that information in LULC features depends on sensors' characteristics. In terms of the quality of information varies from high spectral and spatial resolution context to lower resolutions. Note there is always a trade-off between spectral, spatial and temporal resolution; typically, the higher the spatial resolution, the lower the spectral and temporal resolution, and the higher the temporal resolution, the lower the spatial and spectral resolutions (Warner et al., 2009). Multispectral images, which are commonly superior in spatial resolution compared to hyperspectral, face challenges when it comes to accurately identifying distinct features within similar groups (Kumar et al., 2015; Marcello et al., 2018). As a consequence, they do not offer comprehensive LULC mapping and classification across various algorithms due to their limited spectral resolution. This limitation hampers precise identification of diverse species and the accurate classification of land cover types (Pandey et al., 2021). In hyperspectral remote sensing, the high amount of spectral information given by hundreds of continued narrow bands EM makes hyperspectral datasets capable in discrimination of land-cover types due to the distinctive spectral signatures observed in different objects (Lv & Wang, 2020). The spectral response of different materials in earth's surface are attributed to their unique chemical and physical properties; i.e., in soil or vegetation analysis, varies are caused by differences in pigments, structural characteristics, and water content (Thenkabail et al., 2016; 2018). These advances in spectral and spatial resolution enabled for obtaining better discrimination and accuracy in land-use and cover mapping and multi-temporal change analysis.

Continued research and development of various state-of-the-art techniques, coupled with the availability of high-quality derived remotely sensed data, hold great potential for advancing current mapping techniques (Pandey et al., 2019; Wang et al., 2022). Currently, among the most commonly and widely-used hyperspectral sensors in land-cover mapping are EO-1/Hyperion (Khosravi & Jouybari-Moghaddam, 2019; Dou et al., 2020; Pal et al., 2020; Roy et al., 2021), HySpex (Schmidt et al., 2017; Sabat-Tomala et al., 2020; Wang et al., 2022; Constans et al., 2022), as well as airborne campaigns datasets such as Pavia University and Center of Pavia from ROSIS, and the Salinas and Indian Pines datasets, both provided by the AVIRIS imaging spectrometer (Wang et al., 2020; Khan et al., 2022; Pandey & Tiwari, 2022). Additionally, simulated data from HyspIRI and EnMAP (Clark, 2017; Marcinkowska-Ochtyra et al., 2017; Rosentreter et al., 2017), as well as hyperspectral cameras mounted on UAVs (Yan et al., 2019; Liu et al., 2020; Matese et al., 2021; Sousa et al., 2022), have been widely utilized for land-cover mapping. Recently launched hyperspectral platforms such as PRISMA and DESIS have also demonstrated great potential for improve performance in applications of land-cover mapping in heterogeneous landscapes (Vangi et al., 2021; Aneece & Thenkabail, 2022; Asam et al., 2022; Farmonov et al., 2023; Kalantar et al., 2022). Other use cases involve utilizing multiple datasets from sensors such as HyMAP, APEX, and CASI (Marcinkowska-Ochtyra & Zagajewski et al., 2017; Raczko & Zagajewski, 2017; Amini et al., 2018).

There are several techniques utilized for land-cover classification, with various cutting-edge methods and approaches (Pandey et al., 2019; Wang et al., 2022). Some of the remarkable state-of-the-art techniques include ML algorithms combined with advanced hyperparameter tuning techniques (Yang et al., 2020; Aneece & Thenkabail, 2021), deep learning approaches, such as CNN-based architectures such as 3D-CNNs or 2D-CNNs with recurrent connections (Paoletti et al., 2018; Swalpa et al., 2021), deep belief networks (DBNs) (Lie et al., 2019; Chintada et al., 2021), and recurrent neural networks (RNNs) (Wu & Prasad, 2017; Mou et al., 2018). Other approaches focus on transfer and active learning techniques (Xie et al., 2021; Thoreau et al., 2022) as well as ensemble methods to enhance robustness and mitigate the impact of misclassifications (Jafarzadeh et al., 2021; Manian et al., 2022; Colkesen & Ozturk, 2022).

These advanced methods employ cutting-edge technologies and algorithms to achieve accurate and detailed classification results (Singh et al., 2020; Moharram et al., 2023).

2.3.1. Machine learning in supervised image classification

Among the mentioned approaches, supervised classification techniques hold prominence in hyperspectral image LULC classification due to their ability to produce accurate and reliable results. Extensive research has been conducted to investigate the robustness and performance of different ML classifiers in the context of hyperspectral remotely sensed data (Petropoulos et al., 2012; Carranza-García et al., 2019; Akar & Gormus, 2022). These studies aim to provide valuable insights regarding different levels of accuracy and determine the most suitable classifier to achieve most accurate results in image classification (Kaul and Raina, 2022). In the last decade, non-parametric algorithms as well as ensemble methods such us SVMs and RF have shown to be most commonly employed in hyperspectral land-cover classification, due to their simplicity, high performance, and ability to handle high dimensional data, and derive accurate results with small amount of training dataset (Petropoulos et al., 2015; Kale et al., 2017; Alcolea et al., 2020; Kaul and Raina, 2022). The unique strengths of SVM lie in its ability to handle complex decision boundaries and capture non-linear relationships through kernel functions. The polynomial and radial basis function (RBF) kernels have been used most frequently in remote sensing, however SVM is the most popular methodology and provides superior accuracy than the other conventional methods for LULC classification (Sheykhmousa et al., 2020; Talukdar et al., 2020). On the other hand, RF's strength lies in its ensemble learning approach, which combines multiple decision trees to achieve accurate and reliable classifications. Numerous studies have demonstrated the superior performance and accuracy of SVM and RF in identifying and categorizing land-cover properties with exceptional precision (Raczko & Zagajewski, 2017; Sabat-Tomala et al., 2020; Aneece & Thenkabail, 2021).

For example, Chen and Cheng (2018) performed image classification using RF and SVM classifiers in order to compare its performance. The research focuses on a Hymap experimental data with a 3.5 m resolution image of Berlin, to classify into five land-cover and use classes. Results demonstrated that the RF yielded an overall classification accuracy of 92.6% with a K_c coefficient of 0.902, and slightly outperformed SVM algorithm, which achieved 91.2% with K_c of 0.884, respectively. These findings indicated the effectiveness of both algorithms which derived high accuracy results but also highlight the suitability and effectiveness of the RF method for high-resolution remotely-sensed classification.

In another study conducted by Christovam et al. (2019), the performance of various pixel-based classifiers, including SVMs and RF, was assessed using the HyRANK dataset across 14 land-use and cover classes. Results of this study revealed that RF perform slightly better compared to SVMs in terms of OA and K_c, although both classifiers achieved comparative high accuracies (91% and 0.89 for RF, and 88% and 0.86 for SVM, respectively).

Nhaila et al. (2019), performed classification HSI using four supervised learning algorithms, namely SVM, RF, K-Nearest Neighbors (KNN), and LDA. The experiments were conducted on three real hyperspectral datasets obtained from NASA's AVIRIS i.e., Indian Pines and Salinas, and ROSIS i.e., University of Pavia. The dimensionality of the datasets was reduced using mutual information. The results showed that the SVM classifier with RBF kernel and RF classifier produced statistically better classification accuracies compared to the other algorithms. The results demonstrate that the SVM classifier with RBF kernel achieves the highest classification accuracy among the tested methods, with OA of 93.2% for the Indian Pines dataset, 93.4% for the Salinas dataset, and 91.9% for the Pavia University dataset, slightly better that RF. Thus, SVM with a radial kernel and RF were shown to be more effective in supervised classification.

More recently, Kokal et al. (2022), performed SVMs on PRISMA HSI with a 30 meter spatial resolution for land-use and cover classification, in a diverse landscape in Turkey. In this study, classification performed across 9 land-cover classes, and reported a remarkably high overall accuracy of 92.3% and a K_c coefficient of 0.91.

Based on literature most case studies suggest that SVMs algorithm is considered to be more popular in hyperspectral data analysis due to its ability to achieve higher accuracy compared to other conventional methods. Yet, according to findings RF, demonstrates higher efficiency compared to the SVM when dealing with datasets containing a larger number of target classes (Li et al., 2016). Sheykhmousa et al. (2020) conducted a systematic review of recently published articles (251) based on supervised remotely-sensed image classification and revealed that medium and high spatial resolution images are most common image types for SVM and RF, respectively. When it comes to lower spatial resolution images, the RF method consistently outperforms SVM, despite the fact that more studies have been conducted using SVM-based approaches on low spatial resolution imageries. A comparison within the review of the maximum average accuracies of RF and SVM methods suggests the superiority of the SVM method while classifying data with many more features. Regarding the utility and performance based on hyperspectral datasets the same study reported that the percentages of the HSI remotely sensed data for SVM and RF, are 21% and 10%, respectively. Consequently, among pixel-based classifiers SVM receives the most attention in working hyperspectral datasets. For SVM and RF, the mean classification accuracy based on hyperspectral datasets remains highest at 91.5% for SVM and 79.59% for RF, respectively (Sheykhmousa et al., 2020).

These studies highlighted the robustness and efficacy of SVM and RF as prominent approaches in the relevant fields of land-cover classification. Overall, the majority of studies based on ML algorithms employ advanced optimization techniques for hyperparameter tuning to enhance the performance of classifiers and overall accuracy (Yang & Shami, 2020; Sahithi et al., 2022). However, several studies highlighted that regarding the ML algorithm and spatial resolution of datasets used, increasing the land-cover classes will in turn decrease the overall accuracy. The number of classes identified is a crucial factor which affects the accuracy of land cover classification; as the number of classes increases, spectral differences between classes become less distinct, decreasing classification accuracy (Van Thinh et al., 2019; Zeferino et al., 2020; Sheykhmousa et al. 2020; Dabija et al., 2021).

2.4. Final Remarks

The approaches based on ML algorithms for LULC classification have proven highly effective in extracting valuable information regarding land use properties and alteration processes from images acquired from HSI sensors, as capable to handle large-scale datasets, consider complex relationships between spectral and spatial features and adapt to various land-cover types and environmental conditions (Gupta et al., 2021; Jia et al., 2021; Wang et al., 2022).

Deep learning and its complex architectures, such as CNN-based approaches, demonstrated significant advancements over other conventional methods. Nevertheless, these methods face challenges that necessitate expert knowledge and specialized hardware, primarily due to the high computational requirements during process (Kaul and Raina, 2022). According to relevant research, SVM-based and RF algorithms are among most widely used methods in ML that can enhance classification performance and provide competitive results with CNNs. The advantage of these methods is that they are relatively simple to implement, are robust against overfitting and can be trained with a small amount of data (Sheykhmousa et al., 2020).

However, challenges remain such as computational and operational cost, as well as selection of suitable features in order to handle the variability and diversity of the land cover classes. The computational costs can be substantial due to dimensionality issues of hyperspectral data as well as the complexity of machine learning algorithms. The processing and analysis of large-scale datasets require significant computational resources and time. It is important to note that there are challenges that need to be addressed during the pre-processing steps of HSI data before the analysis, regarding the complexity and volume issues in HSI datasets, i.e., curse of dimensionality. Additionally, operational costs associated with processing procedures, such as storage and analysis of hyperspectral data, can pose significant constraints on the practicality of implementing these advanced methods on large-scale.

The accuracy and precision of the output results for LULC classification primarily rely on sensors characteristics and related factors to image acquisition. All in all, the selection of an appropriate technique that aligns with the data characteristics, training data, and algorithm fine-tuning play an important role in ensuring accurate and precise mapping outcomes. Addressing these challenges involves overcome computational and operational costs, selecting suitable features to handle the variability and diversity of land-cover classes, and optimizing algorithms. The computational costs associated with processing related procedures to large-scale datasets and the storage requirements for hyperspectral data.

This research aims at addressing the limitations related to efficient feature discrimination in diverse and heterogeneous Mediterranean settings by exploring the performance of fine-tuned ML algorithms, and leveraging the enhanced spectral information provided by the EnMAP HSI. By doing so, this study contributes to the advancement of the field by evaluating the potential of well-established ML algorithms on EnMAP hyperspectral data and addressing the specific challenges associated with land-cover classification in Mediterranean environments.

Chapter 3

3. Experimental set up

This chapter highlights the significance of the chosen data for the research as a thorough overview of the study area and datasets used in this thesis. It provides an overview of the geographical location and spatial extent of the research area, along with sources and types of data used in the analysis and for quality assessment.

3.1. Experimental sites

The selected study sites featuring a diverse range of land cover forms, varying from a mixture of mountainous terrain and agroforestry areas, density urban areas, surrounding suburban regions, agricultural fields, including fragments of large-scale farming. The selected scenes located in the northern part of Greece, near the border with Macedonia and it is composed by three imageries across the sensors path, each covering an area of 30 × 30 km, encompasses a tile of total 90 × 90 km swath width (Fig.3.1). The selection of areas was primarily based on the availability of EnMAP hyperspectral data within a Mediterranean region as well as on the presence of diverse forms of land cover types.

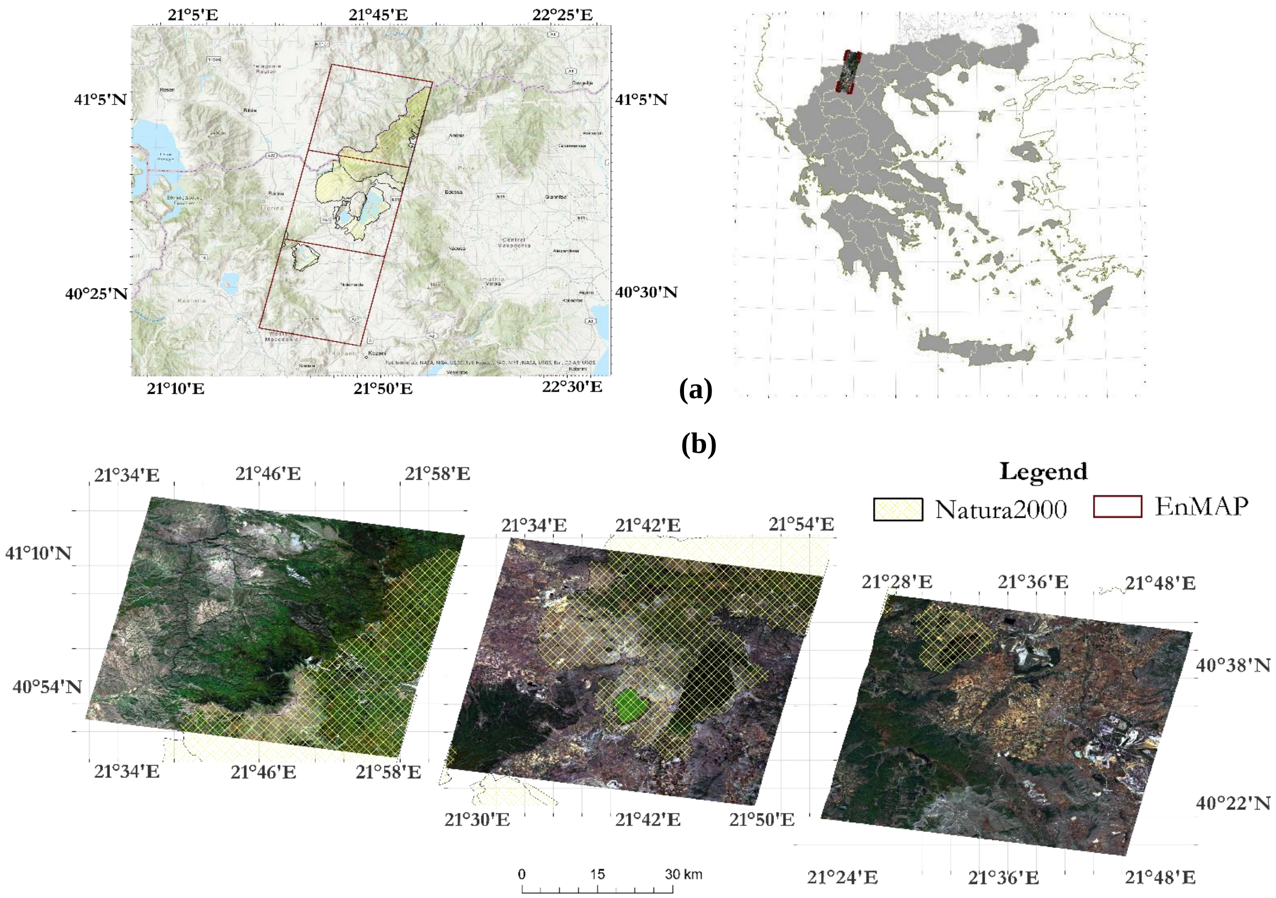


Figure 3.1. Overview of the research areas: **(a)** Research areas are located in Northern Greece close to the borders with north Macedonia (north-western part Greece and central-eastern part of North Macedonia), and **(b)** EnMAP hyperspectral tiles with the yellow polygons correspond regions declared as Natura 2000 within the study area

The scene in the first tile (1) is located in the southern part of the Balkan Peninsula, in the country of North Macedonia above the borders with Greece. The region is mainly mountainous, with elevations ranging from approximately 1867 meters to over 2525 meters above sea level. There are also several peaks and valleys visible in the area. In the north and northwest of the region, there are several smaller mountain ranges, while to the south and southeast, the terrain appears to be flatter and more open. The vegetation in the region would depend on the altitude and the amount of precipitation, but it is likely to include forests, grasslands, and shrublands.

The scene in the second tile (2) is found in the western part of North Macedonia, near the border with Greece. It is located near Amyntaio, a municipality in the regional unit of Florina of West Macedonia, Greece. The terrain in this region is mainly mountainous, and the area is primarily characterized by forests and woodland, with some areas of shrubs and natural grasslands. There agricultural land use in the lower elevations, particularly along the valleys. Mineral resources in the region include deposits of lead, zinc, and copper, as well as marble and other types.

The scene within the third tile (3) corresponds to a region in Greece, located in the north-eastern part of the country. The landscape in this region is diverse and includes both agricultural and forested areas, as well as some mineral deposits. Agriculture is an important economic activity in the study region, with the major crops including wheat, corn, and tobacco. The lowlands are mostly used for agriculture, while the mountainous areas are used for grazing livestock. In terms of mineral resources, the region contains deposits of lignite coal, which is used for electricity generation.

Natura 2000 is composed of two types of designated areas, namely Special Areas of Conservation (SACs) and Special Protection Areas (SPAs). The Habitats Directive specifies the requirement for Sites of Community Importance (SCIs) to be designated, which can then become Special Areas of Conservation (SACs) upon approval by the European Commission (EEA, 2020). These areas are designated to protect species other than birds and specific habitat types, such as forests, grasslands, wetlands, and more. They are not subject to strict protection in terms of human activities and usage; many of these sites are utilized for farming or forestry. Regarding Natura sites in extend of study area, the mountain hill located in the northwest of the study site, known as Oros Voras (SPA) in Pella, encompasses a total area of 79,454 ha. The central part of the region, in Florina, is characterized by significant water bodies. These include Lakes Vegoritida – Petron (SCI), spanning an area of 12,569 hectares, Lake Petron (SPA) covering 6,696 ha, and Lakes Cheimaditida – Zazari (SPA) in the southeast, with a combined area of 5,193 ha.

The landscape is a diverse and dynamic area with a mixture of natural and human-made features. Overall, the region is important for cultivation, agroforestry and mining activities, with a diverse array of natural vegetation. In the higher elevation areas to the north, there are forests dominated by deciduous and conifer trees. In lower elevations, the forests are more characterized as scrubland and grasslands. The lowland areas to the south are primarily used for agriculture, mostly small-scale family-owned farms. In terms of mining, there are several active mines in the study area that extract various minerals, including lead, zinc, copper, gold, and silver. The most notable is a large open-pit lignite mine located in the central part of the study area; is one of the largest open-pit mines in Greece, with a production capacity~10 million mt/yr. There are also several smaller mines in the surrounding hills. Activities in mineral deposits in the area caused environmental degradation, highlighting the need for sustainable resource management practices. The lignite mines in the area have had a significant impact on the environment, as they have altered the landscape and caused pollution of air and water resources in the surrounding areas. Activities in mineral deposits in the area caused environmental degradation, highlighting the need for sustainable resource management practices. In addition, the agricultural practices in the area, including the use of fertilizers and pesticides, poses a threat for soils' health as can lead to soil erosion and water pollution if not properly managed. In terms of CLC (2018) of the European Environment Agency (EEA) within the region are mainly formed Mediterranean forests, woodlands-scrub and temperate broadleaf and mixed forests (Fig.3.2).

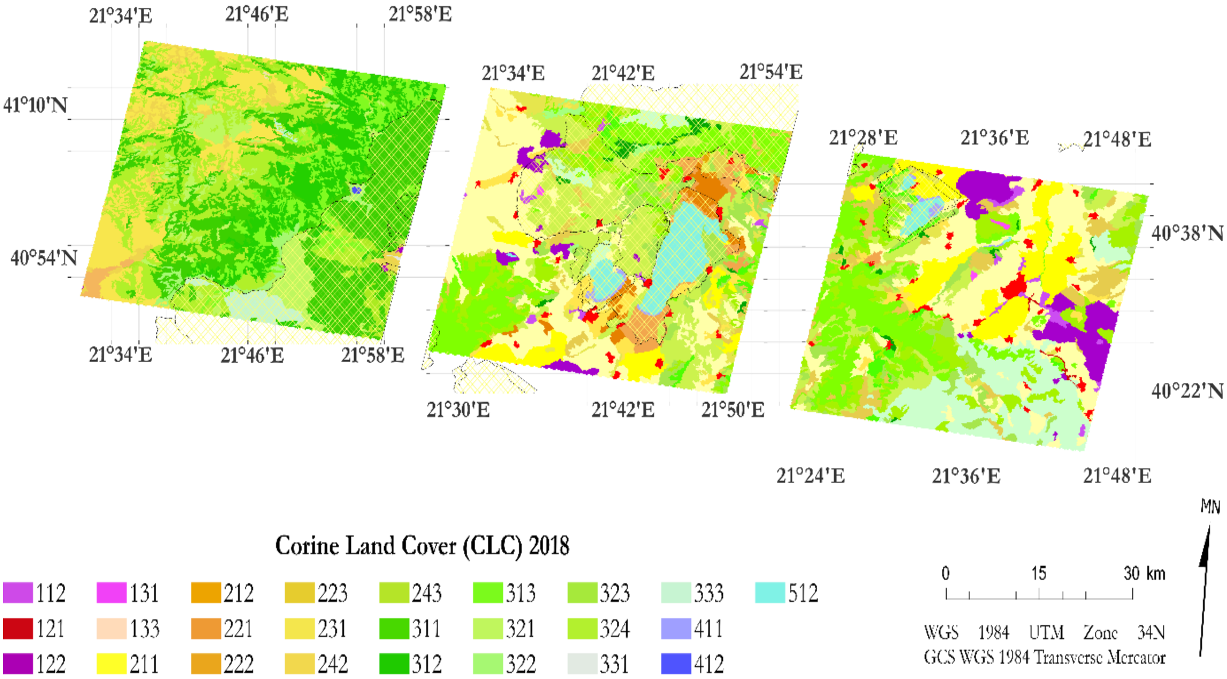


Figure 3.2. The Corine Land Cover maps obtaining from the official Corine Land Cover for 2018 of the EEA for each scene. Codes: **112**: discontinuous urban fabric; **121**: industrial or commercial units; **122**: Road and rail networks and associated land; **131**: mineral extraction sites; **133**: construction sites; **211**: non-irrigated arable land; **212**: permanently irrigated land; **221**: vineyards; **222**: fruit trees and berry plantations; **223**: olive groves; **231**: pastures; **242**: complex cultivation patterns; **243**: land principally occupied by agriculture, with significant areas of natural vegetation; **311**: broad-leaved forest; **312**: coniferous forest; **313**: mixed forest; **321**: natural grasslands; **322**: moors and heathland; **323**: sclerophyllous vegetation; **324**: transitional woodland/shrub; **333**: sparsely vegetated areas; **411**: inland marshes; **412**: peatbogs; **512**: water bodies

3.2. Datasets

3.2.1. Environmental Mapping and Analysis Program (ENMAP)

EnMAP is a spaceborne hyperspectral mission designed from a consortium of German institutions headed by the German Aerospace Center (DLR). EnMAP platform launched in April 2022, aiming to monitor the environment and characterize the Earth's surface on a global scale by providing high spectral resolution hyperspectral imagery data in a large spatial coverage, which allows for a comprehensive view of the Earth's surface. The high quality of the derived data, allows for a better discrimination of different materials which will improved characterization of the environment.

3.2.1.1. Technical specifications

The mission is characterized by a specialized hyperspectral sensor with a pushbroom design that covers a spectral range of 420 nm to 1000 nm (VNIR) and 900 nm to 2450 nm (SWIR) with high accuracy and stability in both ranges. The sensor provides a 30 km wide image with a 30 m x 30 m spatial resolution and an off-nadir pointing of 30° for rapid target revisits. The onboard memory capacity will enable the acquisition of 1,000 km of image data per orbit and a total of 5,000 km per day. The EnMAP expedition is fitted with a strict information handling mechanism that involves advanced radiometric and atmospheric adjustment algorithms. These mechanisms guarantee data's quality and precision, facilitating the production of precisely-corrected data products, that can be used for a variety of applications (Table 3.1). The design of EnMAP can support gap filling in space-based imaging spectrometer data for environmental and earth observation applications, providing high-quality hyperspectral data that can be used to support a wide range of studies and decision-making processes (Chabrillat et al., 2022).

Table 3.1. EnMAP HSI Instrument technical specifications

EnMAP HSI Instrument Specifications	
Image system	Pushbroom-prism
Spectral range	VNIR: 420-1000 / SWIR: 900-2450
Spectral sampling distance:	6.5 nm (420 nm - 1000 nm; VNIR)/10 nm (900 nm - 2450 nm; SWIR)
Signal-to-Noise ratio:	> 500 (at 495 nm; VNIR), > 150 (at 2200 nm; SWIR)
Number of bands	VNIR: up to 99 bands / SWIR: up to 163 bands
Processing types	L0, L1B, L1C, L2A
Ground sampling distance	30 m × 30 m
Swath width	30 km
Geometric co-registration:	< 0.2 pixel (at Level 1C)
Spectral sampling distance:	6.5 nm (420 nm - 1000 nm; VNIR) 10 nm (900 nm - 2450 nm; SWIR)
Spectral accuracy / stability:	0.5 nm / 0.5 nm (VNIR) / 1.0 nm / 0.5 nm (SWIR)
Smile and keystone:	< 0.2 pixel
Orbit repeat cycle:	398 revolutions in 27 days
Orbit altitude:	653 km (7021.8 km semi-major axis)
Inclination angle:	97.96° (polar, sun-synchronous)
Orbital period:	5856 s
Local time descending node:	11:00 h ± 18 min.
Revisit time:	4 days (±30° off-nadir tilt) / 21 days (±5° off-nadir tilt)

3.2.2. EnMAP Dataset

The EnMAP dataset made publicly available for download at no cost from EnMAP Ground Segment Instrument Planning Subsystem (IPS) Portal within EOWEB Geoportal (EGP) (https://eoweb.dlr.de/egp/). EOWEB is the interface to the German Satellite Data Archive (DSDA). EOWEB Geoportal (EGP) is a multi-mission Earth observation data portal of the DLR. Through traditional map- and filter-based, and ordering functions, Earth observation data from the German Satellite Data Archive can be requested using a user account (Table 3.2).

Table 3.2. Basic characteristics of the EnMAP datasets that used in this study

Characteristics of acquired dataset		
	Product level	L2A
Landscape	Tile **1**: Tile **2**: Tile **3**:	Temperate Broadleaf and Mixed Forests Mediterranean Forests, Woodlands and Scrub Mediterranean Forests, Woodlands and Scrub
Acquisition settings	**Correction type**	Combined (land and water)
	Terrain correction	Yes
	Image resampling	Bilinear Interpolation
	Cirrus and haze	Removal
	Water type	Clear
Lat/long (center frame):	Tile **1**: Tile **2**: Tile **3**:	41.037228825 / 21.778811225 40.7682671 / 21.701071525 40.49930575 / 21.62320955
Projected coordinates	UTM	UTM 34N – Greece
Date and time (UTC)	Tile **1**: Tile **2**: Tile **3**:	2022-10-07 09:54:38 2022-10-07 09:54:42 2022-10-07 09:54:47

3.2.3. Ancillary data

The Corine Land Cover 2018 from EEA and Planet Scope image dataset were utilized for visual interpretation and validation purposes, respectively. To focus on specific research areas, both the Corine Land Cover data and the Planet Scope image data were clipped accordingly. The projected coordinate system used was the Universal Transverse Mercator (UTM) with UTM zone 34N, Greece. The CLC data for year 2018 containing the Corine Land Cover (CLC) data for the year 2018 was acquired in vector format (500 × 500 m) through the Copernicus Land Monitoring Service (https://land.copernicus.eu/) and was utilized to identify and extract comprehensive information on land-cover categories with the highest extent. For the purpose of validation, the PlanetScope dataset was acquired for date 07/10/2022 and incorporated into the study. PlanetScope, an innovative satellite data platform developed by Planet Labs, a leading private Earth imaging company based in the United States, offers a distinctive constellation of small CubeSat satellites equipped with high-resolution imaging cameras.

4. Methodology

This chapter serves as a comprehensive guide to the methodology used in this thesis. It provides an overview of the research design, including the methodological approach and validation metrics that were used. More specifically, it presents the data acquisition, pre-processing, analysis, and validation steps undertaken to access the objectives of the research. In addition, it highlights the techniques, tools, and software utilized for data analysis, along with their respective potentials and limitations.

4.1. Methodological framework

The extraction of thematic information from satellite data is mainly performed through the identification of spectral and spatial patterns associated with different types of land-use /land-coverage. Satellite data classification algorithms use the spectral response values for each pixel in the image to determine the category to which it belongs according to the characteristics of features on the earth's surface (Lv, & Wang 2020; Zhang, & Chen, 2020). For pre-processing and implementation of the analysis the classification workflow was used the EnMAP Box-3 (EnMAP-Box Developers, 2019; van der Linden et al., 2021). The overall methodology implemented for land-cover classification is illustrated in Figure 4.1.

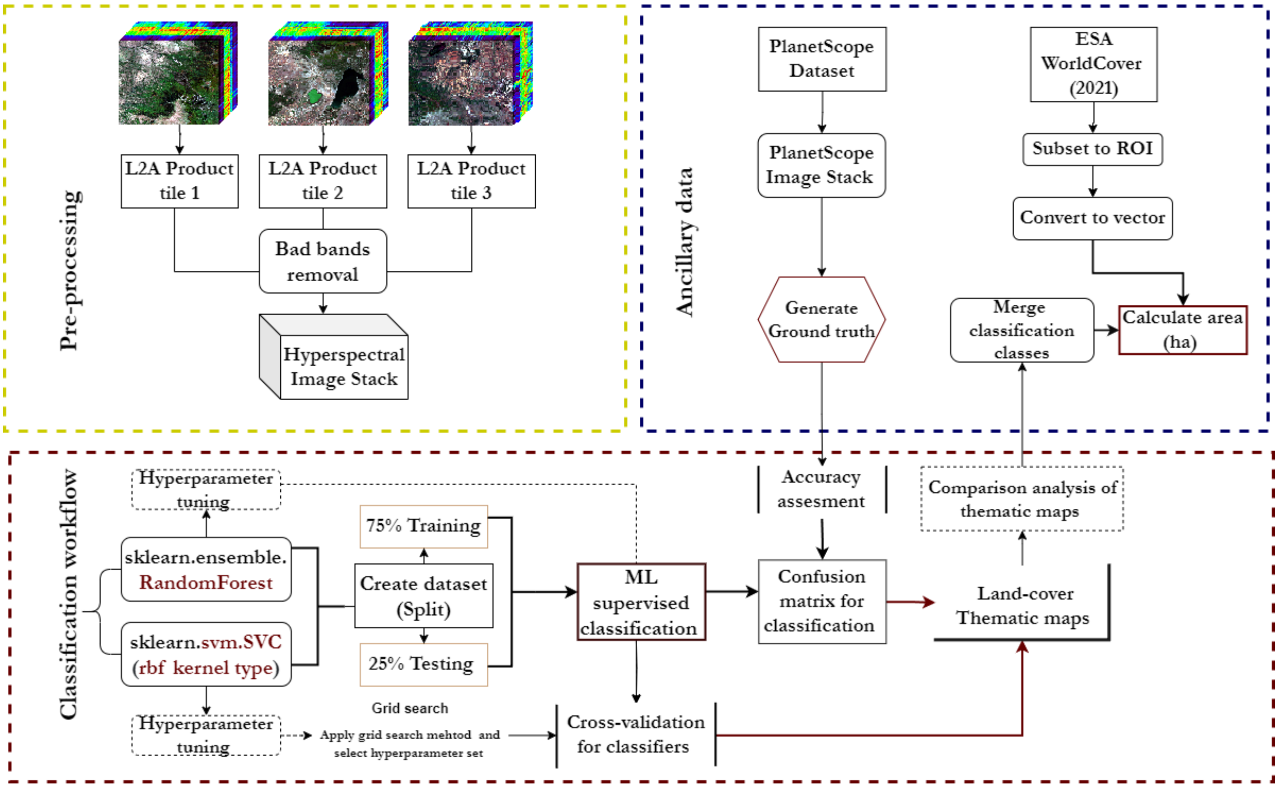


Figure 4.1. Methodological framework of the land-cover classification analysis

4.1.1. Pre-processing

The dataset comprises three tiles along the sensor's track, acquired from the EOWEB Geoportal (EGP) at L2A processing level. These products have undergone atmospheric, radiometric, and geometric corrections, ensuring data's quality and accuracy. The EOWEB Geoportal interface offers additional pre-processing options, allowing users to specify atmospheric correction types based on landscape characteristics and land-cover types. Furthermore, the interface provides various image analysis options to further enhance the data analysis process. Regarding to the study sites used for evaluating EnMAPs' potential in land-cover mapping, datasets obtained in selected GeoTIFF format with Metadata in Universal Transverse Mercator (UTM) coordinate system; other available options include BIL, BSQ, BIP or JPEG2000 formats, provided also with Metadata.

During ordering phase, configurations for processing options were determined. As such, terrain correction was selected to be applied and atmospheric processing type was set combined (which include land and water correction) for each tile due to the presence of water bodies in every scene. The bilinear interpolation method was chosen for image resampling, and no additional band interpolation was applied. Further processing involved selecting the water reflectance type and water type, which were set to normalized and clear, respectively. Cirrus and haze removal were performed, while no cloud removal was necessary as the obtained dataset was free of clouds and cloud shadows. Dataset are delivered after request using the File Transfer Protocol Secure (FTPS) and accessed through the FileZilla Server 1.7.0, using host connection (download.dsda.dlr.de) and ID of user's personal account to EOWEB Geoportal.

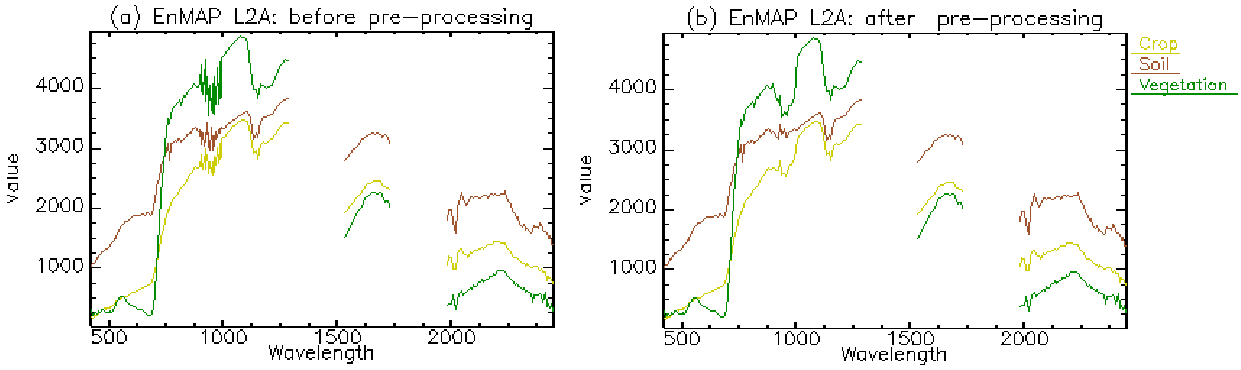


Figure 4.2. Spectral response of vegetation (green), soil (brown) and crops (yellow) of EnMAP hyperspectral data in Level-2A **(a)** before pre-processing and **(b)** after pre-processing applied

For processing and post-processing, EnMAP Box-3 plugin was used within open-source QGIS 3.30.3 software. The L2A product contain initially 212 bands from 420-1000 nm (VNIR) to 900-2450 nm (SWIR). The metadata were imported as an EnMAP L2A product and then exported as spectral images after undergoing additional processing steps. Firstly, a moving average filter with a 3x3 kernel window was used to detect overlapping bands in the VNIR and SWIR regions (Fig.4.2). Further, bad bands occurring between 1358 nm and 1453 nm, as well as between 1814 nm and 1961 nm, due to water absorption, have been removed to avoid impacts at any further analysis. In this case, bad bands ranging from band 115 (1295.28 nm) – band 129 (1519.22 nm) and band 149 (1738.93 nm) – band 155 (1967.95 nm) identified and removed. Each tile exported as spectral image containing 192 bands. Subsequently, the tiles were merged into a single raster, encompasses a scene of total 90 km swath width. Following the merging process, a scale factor of 0.0001 was applied to the TOA reflectance of the mosaic image. Lastly, an essential step involves the application of Quality Level (QL) test flags files, which are provided as part of the data. This step serves to effectively detect and eliminate line noise that may be present in the data. Line noise and white spots resulting from sensor saturation, can adversely impact the accuracy of subsequent data analysis and classification processes. Therefore, it is crucial to identify and remove these pixels prior to the classification analysis (Storch et al., 2013). The image then projected in WGS 1984 UTM zone 34N, which corresponds to Greece.

4.2. Classification of ENMAP imagery

In this research, supervised classification approaches are employed in order to identify and categorize various land cover classes using advanced machine learning fine-tuning algorithms, i.e., SVM and ensemble learning algorithm RF. The SVM with a radial kernel was implemented due to superiority in performance and accurate results (Talukdar et al., 2020). RF is widely used with hyperspectral (HSI) data due to its ability to provide good classification results, without relying on any underlying probability distribution for the input data (Jafarzadeh et al., 2021). These advanced ML algorithms have demonstrated exceptional performance in LULC classification tasks, leveraging their robustness, and capability in handling and processing complex and noise datasets (Kale et al., 2017). In this research, the classification scheme was based on the Corine Land Cover from the European Environment Agency (EEA), and implied to classify 11 distinct classes, after modifications based on sub-classes at level 3.

4.2.1. Support Vector Machines (SVM)

The SVM was introduced in the late 1970s by Vapnik (Vapnik, 1995), and since then, it has become the most widely employed kernel-based algorithm for various classification tasks. In the realm of classification, it holds a great advantage over other statistical methods as a non-parametric algorithm in remote sensing image classification, mainly because it does not pose any constraints based on the distribution of data.

SVM is a supervised machine learning algorithm that aims to find an optimal hyperplane that maximally separates data points of different classes. A hyperplane is a decision boundary that functions through separating different classes in the feature space with aim to find hyperplane with the largest margin. A commonly used technique for finding optimal combinations of hyperparameters is Grid Search, which functions by exhaustively searching through a specified range of parameter values. The margin in SVM refers to the distance between the decision boundary and the closest support vectors; SVM aims to maximize this margin to achieve better generalization and improve classification performance. The C parameter in SVM controls the trade-off between maximizing the margin and minimizing classification errors. Support Vectors are the data points that are closest to the decision boundary, i.e., hyperplane. The kernel trick is a technique used to transform data into a higher-dimensional space, enabling SVMs to effectively handle complex relationships between features. Based on the type of kernels utilized to establish decision boundaries, SVMs can be categorized as linear, which employs a linear kernel, and non-linear, which utilizes non-linear kernels such as polynomial or radial basis functions. Given that a comprehensive description of SVMs can be found in existing literature, such as Foody and Mathur (2004) and Pal and Mather (2005), a detailed explanation of SVMs it is not provided in this context. The kernel-based techniques that are extensively employed in the utilization of SVMs are briefly presented above (Scholkopf et al., 1999):

Liner kernel:

$$k(x_i, x_j) = x_i \cdot x_j \tag{4.1}$$

Polynomial kernel:

$$k(x_i, x_j) = (\gamma x_i \cdot x_j + r)^d, \gamma > 0 \quad 4.2$$

Radial basis function (RBF) kernel:

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad 4.3$$

Sigmoid kernel:

$$k(x_i, x_j) = \tanh(\gamma x_i \cdot x_j + r) \quad 4.4$$

The linear kernel (*Equation 4.1*) computes the dot product in feature space. The polynomial kernel (*Equation 4.2*) uses a user-defined constant, denoted as d, to determine the kernel order. The RBF kernel (*Equation 4.3*) incorporates a weight term, c, to represent another kernel type. The sigmoid kernel (*Equation 4.4*) illustrates a two-layer sigmoid neural network that acts as a similarity indicator between xi and xj, using the dot product term (x x i j) in each kernel. User-defined parameters, such as γ, d, and r, significantly impact the accuracy of SVMs' solutions in different kernel functions.

4.2.2. Random Forest

Random Forest is an ensemble learning algorithm that combines multiple decision trees in order to create a powerful model in image classification tasks. Ensemble learning refers to the technique of combining multiple models to improve overall performance. A random selection of characteristics and data is used to individually train each decision tree, reducing overfitting and increasing diversity among the trees. The approach employs bootstrap aggregating, a method known as bagging, to train various bootstrap samples of the training data into diverse decision trees (Breiman, 2001). The final prediction, is based on different tree projections are combined during prediction through voting or average. The significance of each characteristic in the classification task is shown by RF's measures of variable importance. Additionally, it avoids the requirement for a separate validation set by estimating model performance using out-of-bag (OOB) error. In the Random Forest algorithm, once the forest is grown, each tree contributes to the decision by voting for a particular class, and the final label is determined by the majority vote. The key benefits of the RF algorithm include its ability to handle noise effectively, requiring fewer parameters for tuning, and offering a computationally efficient solution. RF method can handle large datasets with several features or channels since it is scalable. Due to the randomization of the feature, RF is naturally resistant to overfitting (Christovam et al., 2019).

4.3. Classification scheme design

A classification scheme was developed by utilizing the archival data of Corine Land Cover (CLC) from 2018. The scheme focused on classifying land cover into 11 distinct classes based on a moderate version of sub-classes in Level-3. To determine these classes, the extent covered by each CLC class in the study area was calculated and classes with the highest coverage were identified. This approach allowed for the identification of land-cover types that represented the dominant land cover categories in study area and specified the number of categories that will be further used for the classification analysis. The classification scheme is shown in Table 5, which provides the names of categories at Level-1, and the classification scheme after modifications based on sub-classes (Level 3).

Table 4.1. Classification scheme generated with reference the Corine Land Cover from 2018 (Level 1) and classification scheme used in the analysis, after modifications based on sub-classes (Level 3). Each class is assigned to its corresponding Level-1 section

Land-use and cover Classification Scheme		
Level 1 (CLC 18)	**Classification scheme**	**Class description**
1. Artificial surfaces	**1.** Urban fabric **2.** Mineral extraction sites	-Discontinuous urban fabric, construction sites, including road and rail networks and associated land -Mining sites of lignite
2. Agricultural areas	**3.** Non-irrigated arable land **4.** Permanently irrigated arable land **5.** Pastures	-Non-irrigated arable land - Permanently or periodically irrigated arable land, - Herbaceous vegetation and grass cover, abandoned arable land, permanent grasslands under grazing by domestic animals
3. Forest and semi-natural areas	**6.** Broadleaved forests **7.** Coniferous forest **8.** Natural grasslands **9.** Mixed forests – Shrub	-Broadleaved forests, -Coniferous forests - Natural grasslands, under minimum human influence, -Mixed forest transitional woodlands and shrublands
4. Wetlands	**10.** Inland marshes	-Flooded vegetation, flowering aquatic plants such as water lily
5.Water bodies	**11.** Water bodies	-Water bodies, Lakes, clear water

4.3.1. Training points selection

The approach used for the generation of the classification scheme was based on the CLC (2018) land-cover map in order to identify the dominant land cover classes within the study area and define the number of classes. The dominant land-cover classes in the study area and those that could not be dismissed were included in the scheme. These classes comprised artificial surfaces (discontinuous urban fabric), broad-leaved and coniferous forest, non-irrigated arable land and permanently irrigated arable land, mixed forest and transitional woodland-shrub which represents the class mixed forest-shrubs, natural grasslands and pastures, mineral extraction sites, water bodies, and inland marshes corresponding to flooded vegetation such as water lilies (Table 4.1). Mixed forests and shrublands were merged into a single category due to spectral similarities. Similarly, areas associated with mineral extraction and construction activities were identified and labelled as mining sites due the spatial limited extent within the study area.

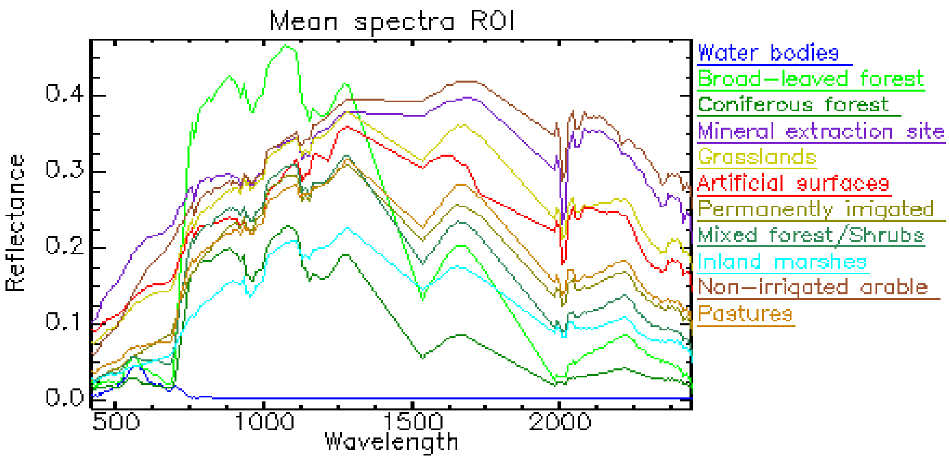


Figure 4.3. Mean spectral signatures of land-use and cover categories used for EnMAP image classification analysis

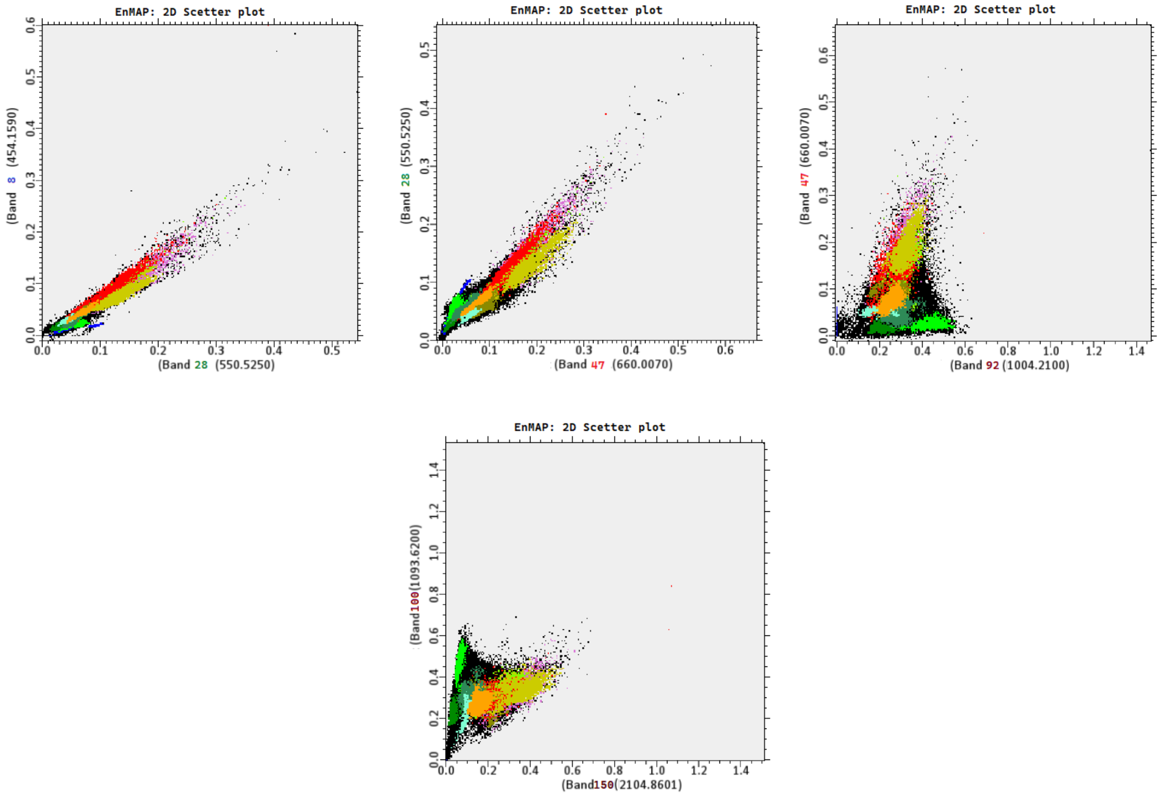


Figure 4.4. The 2-D scatter plot of EnMAP spectral bands imported with ROIs in **(a) Blue** (B8) vs. **Green** (B28); **(b) Green** (B28) vs. **Red** (B47); **(c) Red** (B47) vs. **NIR** (B92); **(d) NIR** (B100) vs. **SWIR** (B150)

Approximately 3500 training points were very carefully selected for each class, a total dataset consist of ~30,000 pixels, used to train each classifier (Fig.4.3). The Jeffries-Matusita and the Transformed Divergence statistical measures performed using ENVI Classic 5.3 (64 bit). To address the high correlation among nearby classes, particularly in VNIR bands, for the analysis of Jeffries-Matusita and Transformed Divergence statistics, the half bands of EnMAP imagery were used. This was primarily motivated by the narrow spectral sampling distance between bands (6.5 nm VNIR and 10 nm in the SWIR). Specifically, a 2-band step was performed, selecting bands at intervals of two bands. By selecting bands at intervals of two bands, the spectral sampling distance was effectively increased between adjacent bands. Figure 4.4 displays a 2-d scatter plot depicting the EnMAP spectral bands imported with ROIs across different wavelengths. The mean spectra of the collected ROI pixels derived after applying pre-processing steps of overlapping detection and of bad bands removal, as shown in Figure 4.5.

The obtained metrics for ROIs separability provide insights into the distinctiveness and discriminability of different land cover classes. The computed spectral separability values indicated a satisfactory discriminative ability, with values ranging from 1.9 to 2. Values higher than 1.9 indicated sufficient separability. The analysis revealed that forested areas exhibited high separability from water bodies, and artificial surfaces while the separability between agricultural areas comparatively lower. Notably, the separability ranging from 1.99 to 2.00 indicating a high and satisfactory degree of separability. The lowest separability was observed between the pastures, grasslands and mixed forests-shrubs as well as artificial surfaces and mineral extraction sites. The separability values of the selected ROIs for training set is shown in Table 4.2.

Table 4.2. Spectral separability of training dataset (ROIs) according to Jeffries-Matusita, Transformed Divergence statistics

	Mining sites	Coniferous forest	Grassland	Pasture	Mixed forest-Shrubs	Inland marsh	Non-irrigated arable land	Broad-leaved	Permanently irrigated	Water bodies
Artificial surfaces	1.98	1.99	1.97	1.99	1.99	1.99	1.99	2	1.99	2
Mineral extraction sites		2	1.99	2	2	2	1.99	2	2	2
Coniferous forest			2	1.99	1.99	2	2	1.99	2	2
Grasslands				1.99	1.99	1.99	1.99	2	2	2
Pastures					1.90	1.99	1.99	2	2	2
Mixed forest-Shrubs						1.99	1.99	1.99	1.99	2
Inland marshes							1.99	2	1.99	2
Non-irrigated arable land								2	1.99	2
Broad-leaved forest									2	2
Permanently irrigated arable										2

4.4. Classification implementation

The classification workflow employed in this study incorporated widely-used ML algorithms, Random Forest (RF) and Support Vector Machine (SVM), for image classification within the advanced classification workflow of EnMAP Box-3 plugin (der Linden et al., 2015) using the *scikit-learn* library (Pedregosa et al., 2011). Performance of each model evaluated using overall accuracy and 'f1 macro' score metric.

For SVM classifier The RBF kernel type was chosen for this study due to well-established popularity and superior accuracy in land cover classification tasks. RF supervised machine learning algorithm is known to provide enhanced classification performance with high accuracy and robustness to noise, and ability to capture non-linear relationships and outlier detection (Talukdar et al., 2020). For RF, the power of ensemble learning, which takes advantage of the diversity among multiple models to enhance the overall performance and achieve accurate results even without hyperparameters tuning; while the fine-tuning process using grid search for the SVM model enabled the identification of optimal parameters and refined further the model and its classification capabilities.

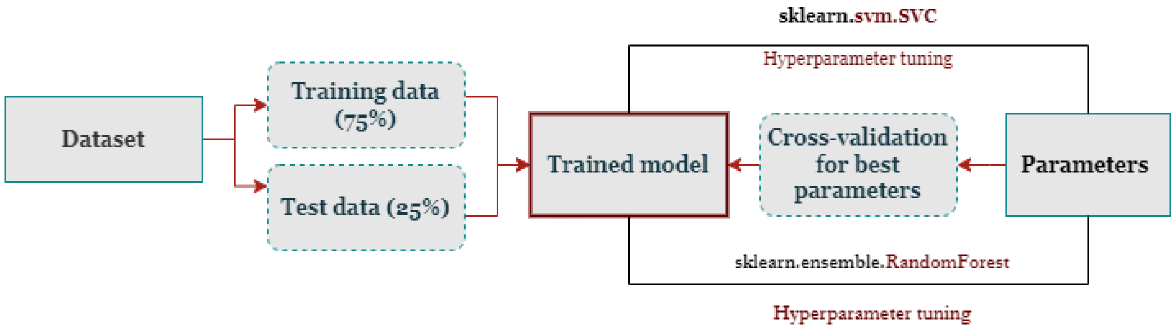


Figure 4.5. Hyperparameter tuning workflow

The training and classification processes for the RF and SVM algorithms were conducted in the EnMAP Box-3 toolbox using the *Random Forest Classifie*r and *Support Vector Classifier (RBF kernel)* packages from the *scikit-learn* library, for RF and SVM respectively (Pedregosa et al., 2011). For the k-fold cross-validation procedure on current dataset, the train/test split procedure was used to evaluate the model. The dataset was randomly split into subsets of train (75%) and test (25%), and same training dataset (pickle file) was used for fitting and evaluating the performance of each classifier as indicated in Figure 4.5.

For the RF classifier, parameters were tuned manually, using python coding through the scikit-learn API within the classification workflow. In order for best estimators and parameters to be found and to improve the algorithm's performance the *GridSearchCV* was used, with cv = 5. The criterion was set to gini and decision trees values set to [10,20,30,50,100,200,500, 1000]. For max features values changed from default auto to (1, 'sqrt', 'log2') and max depths from default None to [2,3,5,10,20], to optimize feature selection and mitigate overfitting. For the min samples leaf values and min samples split, the ranging [5,10,20,50,100,200] and [2,5,10,20] was used, respectively. The adjustments were made based on recommendations within the sklearn’s documentation and multiple experimental iterations using different configurations of parameters. After the fine-tuning process, the optimal combination with gini criterion was found to 1000 estimators.

In case of SVMs with an RBF kernel, the hyperparameter optimization process was conducted in order to find the optimal values for each parameter for the algorithm, taking into account the input dataset. By systematically exploring different parameter combinations, the optimization process sought to maximize the performance of the algorithms and improve their accuracy in classifying the data. A hyperparameter tuning for cost C and gamma γ value was performed using an automated process called *GridSearchCV*, using a k-fold cross-validation, with k = 5. A range of values were considered for cross-product of C and γ values ranging in [10^{-3},10^{-2}, 10^{-1}, 1, 10^{1} ,10^{2} ,10^{3}]. After the fine-tuning process, the optimal combination of regularization parameter C and gamma γ was set at 1000 and 0.01, respectively.

4.4.1. Accuracy assessment approach

The accuracy assessment of the thematic land-cover maps involved the utilization of various statistical metrics, as evaluated through the confusion matrix. These metrics included the overall accuracy Overall Accuracy (OA), User's Accuracy (UA), Producer's Accuracy (PA), Kappa coefficient (Kc). OA provides an estimate of the overall classification accuracy, representing the percentage of correctly classified pixels in the output map. K_c measures the level of agreement between the reference data and the chosen classifier, compared to the agreement expected by chance. PA indicates the percentage of correctly classified pixels within a specific class and reflects the omission error, which represents the pixels omitted from their reference class. Similarly, UA represents the percentage of pixels classified as a specific category that actually belong to other ground truth classes, indicating the commission error. Mathematically, these parameters can be expressed as shown (Congalton, 2001).

$$OA = \frac{1}{N}\sum_{i=1}^{r} n_{ii} \tag{4.5}$$

$$UA = \frac{n_{ii}}{n_{irow}} \tag{4.6}$$

$$PA = \frac{n_{ii}}{n_{icol}} \quad 4.7$$

$$K_c = N\sum_{i=1}^{r} n_{ii} - \sum_{i=1}^{r}\frac{n_{icol}n_{irow}}{N^2} - \sum_{i=1}^{r} n_{icol}n_{irow} \quad 4.8$$

where the symbol n_{ii} represents the count of pixels that are accurately classified within a particular category; the variable N refers to the total number of pixels present in the confusion matrix, while r denotes the number of rows; n_{icol} represents the sum of columns in the reference data; n_{irow} signifies the sum of rows in the predicted classes.

For the computation of the statistical metrics, the validation set was independently selected from each categorization class to ensure accurate classification through random selection. The multispectral imaging obtained in eight spectral bands in VNIR, namely coastal blue, green I, green, yellow, red, red edge, NIR with a spatial resolution of 3 m (https://www.planet.com/). Approximately 9,000 pixels were generated using the PlanetScope multispectral image as reference, with around 800 pixels allocated for each class (which accounted for 25% of the training set). The process of selection validation points (pixels) was primarily guided by the PlanetScope satellite imagery (pixel size ~3 meters) used as reference, which acquired closely in time to the EnMAP datasets (Fig.4.6). In order to prevent overlap between the training data and validation sites and to ensure that the pixels used for training did not coincide with those used for validation, the validation sites were chosen at locations separate from where the random points were collected. Moreover, validation pixels were collected in regions exhibiting homogeneity to mitigate pixel mixing effects regarding the high resolution PlanetScope imagery in conjunction with EnMAP HSI.

To determine OA as well as UA and PA from classified maps the *estimating area and map accuracy for simple random sampling* tool within the EnMAP Box-3 software was utilized (Stehman, 2014). Subsequently, these metrics were used to assess the Cohen's Kappa (K_c), which was calculated using Microsoft Excel. The statistic F1 scores was also produced to gain insights regarding misclassifications within the classification results. These statistical metrics provide valuable information regarding the performance and reliability of the classification process, allows for a comprehensive evaluation of the accuracy and robustness of the results.

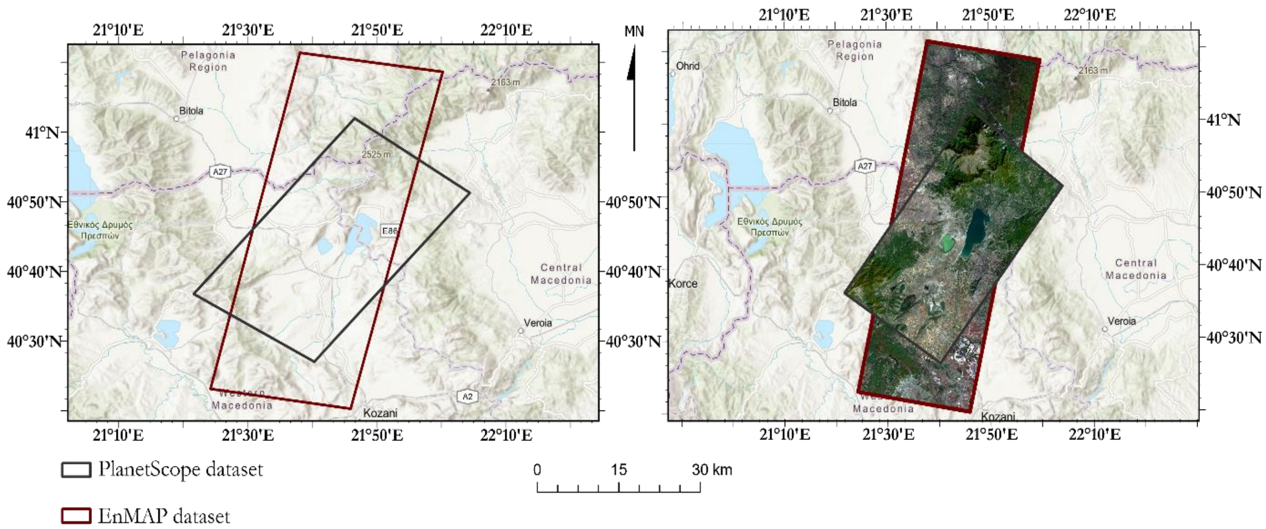


Figure 4.6. PlanetScope dataset overlap in EnMAP dataset used for the collection of validation pixels. Dataset are presented in true colour RGB composite i.e., for PlanetScope R:6, G:3, B:2 and for EnMAP R:47, G:32, B:14

4.5. Comparative analysis of thematic land-cover maps

The same set of validation data was used to evaluate the classification accuracy of implemented algorithms RF and SVM based on EnMAP imagery, for the production of land-use maps. To assess the statistical significance and determine the superiority of one classifier over another, McNemar's chi-squared (χ^2) test was employed (McNemar, 1947), which is a parametric test known for its simplicity and accuracy (De Leeuw et al., 2006). Existing literature suggests that it is more precise and sensitive than the Kappa z-test (Manandhar et al., 2009). McNamar utilizes two error matrices to calculate a chi-square (χ^2), which compares the misclassifications between the two classifiers.

More specifically, it considers the cases where one classifier incorrectly classifies a sample that the other classifier classifies correctly and vice versa. This statistical test, which based on the chi-square value, provides insights into the comparative performance of the classifiers and represents a straightforward approach for comparing the predictions of different algorithms on a per-class basis. The computation of McNemar's test is as follows:

$$\chi^2 = \frac{(f_{12}-f_{21})^2}{f_{12}+f_{21}} \quad 4.9$$

where f_{12} represents the number of cases that are misclassified by classifier one but correctly classified by classifier two, while f_{21} represents the number of cases that are correctly classified by classifier one but misclassified by classifier two.

5. Results

Chapter 5

This chapter provides a detailed account of the results obtained in this thesis. It presents an accuracy assessment of the classification results, along with a detailed comparative analysis based on different machine learning algorithms. Furthermore, the chapter offers a comparison of the thematic maps obtained from the analysis and current operational products.

5.1. Accuracy assessment of classification analysis

The land-use and cover maps produced using tuned SVMs and RF algorithms. The selected best estimators for each classifier used for the analysis, are shown in Table 5.1. Thematic maps illustrating the land-cover and land-use categories derived from the implementation of each classifier based on EnMAP HSI, are presented in Figure 5.1. Table 5.2 summarize the results of the classification assessment i.e., OA, Cohen's Kappa (K_c), PA and UA accuracy metrics. In Tables 5.3 and 5.4 the confusion matrix of each classifier is shown, which represent the pixel misclassifications and pixels that classified correctly for each classifier. The visual examination of both the thematic maps and the statistical analysis indicates that both classifiers performed well in accurately representing the spatial distribution of different land-cover and use types across the study area.

For RF model parameters were tuned manually using python coding, and best estimators were found using *GridSearchCV*, with *cv = 5*. For RF with 'gini' criterion, the optimal value of estimators (trees) was found to be 100 and the minimum number of samples required to split an internal node 2. For SVMs RBF kernel, the best values for hyperparameters were automatically identified using the grid search tuning technique. The optimal values using grid search method yielded values of the regularization parameter C at 1000 and γ parameter at 0.01, which found to be most effective resulting in a significant OA of 99%. The grid search technique allows for a systematic exploration of different hyperparameter combinations to identify the optimal values. By employing grid search technique, fine-tuning of SVM and RF model achieved significant performance of fitted classifier (Table 5.1).

Table 5.1. Best values of hyperparameters tuned for each model using 5-fold cross-validation and OA of two classifiers

Algorithm	Parameters	Value	OA
	criterion	gini	
	n_estimators	1000	
RF	max_depth	20	95%
	max features	sqrt	
	min_samples_split	2	
	min_samples_leaf	5	
	C	1000	99 %
SVM (RBF)	γ	0.01	

The SVMs classifier achieved an average UA range of 71.3% to 99.1% and an average PA ranging from 73.5% to 99.5% across eleven different land-cover classes. The Random Forest classifier showed an average UA range of 55% to 98.6% and an average PA range of 55% to

99.5%. The SVM (RBF) classifier demonstrated promising results across land-cover classes used. For artificial surfaces, the UA and PA were 94% and 98.2%, respectively, with an F1-score of 96%. Broad-leaved forest achieved high accuracy with a UA of 97.5% and a PA of 98.2%, resulting in an F1-score of 97.9%. Similarly, the coniferous forest exhibited excellent classification performance with a UA of 99.1% and a PA of 98.3%, yielding an F1-score of 98.7%. The SVM classifier performed well in identifying mineral extraction sites, with a UA of 87.7% and a PA of 96.9%, resulting in an F1-score of 92.1%. The classification of permanently irrigated arable land achieved a UA of 92.9% and a PA of 95.2%, with an F1-score of 94%. For non-irrigated arable land, the SVM (RBF) classifier demonstrated a UA of 98.78% and a PA of 83.9%, resulting in an F1-score of 90.79%. The classifier showed promising performance for transitional woodland-shrub, with a UA of 88.6% and a PA of 95.8%, yielding an F1-score of 92%. Water bodies were accurately classified, with a UA and PA of 94.9% and 95.3%, respectively, resulting in an F1-score of 95.1%. Inland marshes achieved a UA of 95.7% and a PA of 84.5%, indicating a relatively accurate classification with an F1-score of 89.7%. However, grasslands, achieved a relatively lower UA of 83.48% and a PA of 73.5%, resulting in an F1-score of 78.2%.

However, the grasslands achieved a relatively lower user accuracy (UA) of 83.4% and producer accuracy (PA) of 73.5%, resulting in an F1-score of 78.2%. Similarly, the pastures exhibited comparatively lower accuracy, with a UA of 71.39% and a PA of 82.6%, resulting in an F1-score of 76.6%. This discrepancy can be attributed to the homogeneity between the two land-cover classes, i.e., grasslands and pastures. Overall, SVMs indicating a high level of agreement between the predicted and reference land-cover classes, with OA of 90.5% and K_c of 0.89.

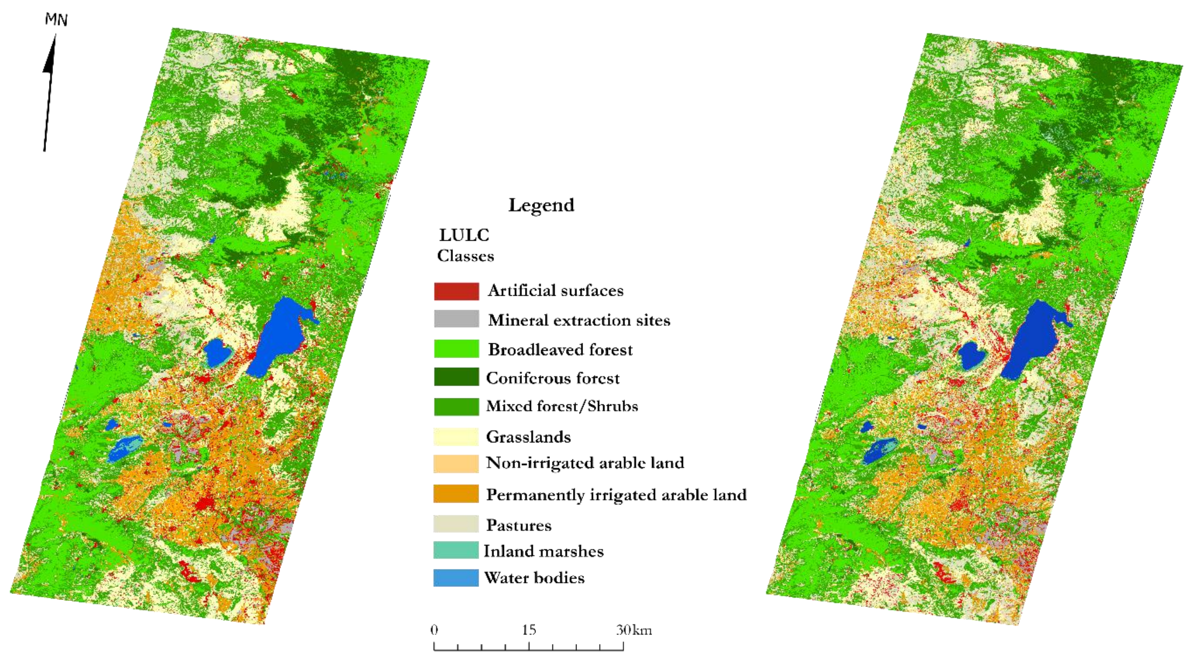


Figure 5.1. Land-cover and use thematic maps obtained using ML pixel-based algorithms: **(a)** SVMS and **(b)** RF classifiers based on EnMAP hyperspectral imagery

Table 5.2. Summary of results in accuracy assessment for land-cover thematic maps using SVMs and RF based on EnMAP hyperspectral satellite imagery

	SVM (RBF)			Random forest		
Land cover classes	**UA (%)**	**PA (%)**	**F1-score**	**UA (%)**	**PA (%)**	**F1-score**
Artificial surfaces	0.940	0.982	0.96	0.851	0.927	0.887
Broad-leaved forest	0.975	0.982	0.979	0.994	0.979	0.986
Coniferous forest	0.991	0.983	0.987	0.986	0.995	0.991
Grasslands	0.834	0.735	0.782	0.805	0.550	0.654
Inland marshes	0.957	0.845	0.897	0.958	0.827	0.887
Mineral extraction sites	0.877	0.969	0.921	0.819	0.912	0.863
Non-irrigated arable	0.987	0.839	0.907	0.965	0.804	0.877
Pastures	0.713	0.826	0.766	0.674	0.866	0.758
Permanently irrigated arable land	0.929	0.952	0.940	0.868	0.971	0.917
Transitional woodland-shrub	0.886	0.958	0.920	0.876	0.955	0.914
Water bodies	0.949	0.953	0.951	0.949	0.952	0.950
Overall accuracy (%)			**OA** = 90.5	**OA** = 87.5		
Kappa coefficient			**K_c** = 0.897	**K_c** =0.862		

Table 5.3. Adjusted confusion matrix of SVMs classifier: predicted (rows) vs. observed (columns). The diagonal elements of the matrix represent the pixels that were classified correctly for each class

Support Vector Machines (RBF)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	Total
Artificial surfaces (1)	**943**	0	1	2	28	13	1	2	3	0	10	1003
Broad-leaved forest (2)	0	**851**	12	0	0	0	0	0	0	9	0	872
Coniferous forest (3)	0	1	**819**	0	3	0	0	0	0	3	0	826
Grasslands (4)	4	0	0	**859**	22	12	1	107	19	5	0	1029
Inland marshes (5)	0	0	0	0	**611**	0	0	0	0	0	27	638
Mineral extraction sites (6)	10	0	0	0	0	**822**	104	0	0	0	1	937
Non-irrigated arable (7)	0	0	0	0	0	1	**892**	1	8	1	0	903
Pastures (8)	0	0	0	298	0	0	2	**811**	3	18	4	1136
Permanently irrigated arable land (9)	0	0	0	1	0	0	62	3	**877**	0	1	944
Mixed forest-Shrublands (10)	3	14	1	8	12	0	0	57	11	**824**	0	930
Water bodies (11)	0	0	0	0	47	0	0	0	0	0	**876**	923
Total	960	866	833	1168	723	848	1062	981	921	860	919	10141

Table 5.4. Adjusted confusion matrix of RF classifier: predicted (rows) vs. observed (columns). The diagonal elements of the matrix represent the pixels that were classified correctly for each class

Random Forest	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	Total
Artificial surfaces (1)	**890**	0	1	67	27	40	15	2	1	1	1	1045
Broad-leaved forest (2)	0	**848**	1	0	0	0	0	0	0	4	0	853
Coniferous forest (3)	0	4	**829**	0	4	0	0	0	0	0	3	840
Grasslands (4)	20	0	0	**643**	8	9	10	86	9	13	0	798
Inland marshes (5)	0	0	0	0	**598**	0	0	0	0	2	24	624
Mineral extraction sites (6)	37	0	0	10	0	**774**	123	0	0	0	0	944
Non-irrigated arable (7)	5	0	0	1	0	25	**854**	0	0	0	0	885
Pastures (8)	6	0	0	352	16	0	0	**850**	7	17	13	1261
Permanently irrigated arable land (9)	1	0	0	64	4	0	60	4	**895**	1	2	1031
Mixed forest-Shrublands (10)	1	14	2	31	19	0	0	39	9	**822**	1	938
Water bodies (11)	0	0	0	0	47	0	0	0	0	0	**875**	922
Total	960	866	833	1168	723	848	1062	981	921	860	919	10141

The Random Forest classifier also demonstrated high accuracy in obtained results. For artificial surfaces, the UA and PA were 85.1% and 92.7%, respectively, with an F1-score of 88.78% and performed well in identifying mineral extraction sites, with a UA of 81.9% and a PA of 91.2%, resulting in an F1-score of 86.3%. Broad-leaved forest achieved exceptional accuracy, with a UA of 99.4% and a PA of 97.9%, resulting in an F1-score of 98.6%. Similarly, the coniferous forest exhibited high precision with a UA of 98.6% and a PA of 99.5%, yielding an F1-score of 99.10%. For non-irrigated arable land, the Random Forest classifier demonstrated a UA of 96.5% and a PA of 80.4%, resulting in an F1-score of 87.7%. In case of permanently irrigated arable land achieved a UA of 86.8% and a PA of 97.1%, with an F1-score of 91.7%. The Random Forest classifier showed promising performance for transitional woodland-shrub, with a UA of 87.6% and a PA of 95.5%, yielding an F1-score of 91.43%. Water bodies were accurately classified, with a UA and PA of 94.9% and 95.2%, respectively, resulting in an F1-score of 95.06%. Inland marshes achieved a UA of 95.8% and a PA of 82.7%, indicating a relatively accurate classification with an F1 score of 88.79%.

Similar to SVMs, grasslands and pastures posed a challenge also for RF classifier, exhibited comparatively lower accuracy resulting in an F1-score of 65.4% and 75.8%, respectively. RF classifier achieved an OA of 87.5% and K_c of 0.86, indicating also a strong agreement between the predicted and reference land-cover classes. However, despite the lower accuracy reported for the grassland and pastures categories in comparison with the other land-cover classes, both classifiers demonstrated satisfactory discrimination capabilities with high F1 scores (Table 5.2).

5.1.1. Comparative analysis of thematic maps

The findings of this study indicate that both pixel-based algorithms examined show significant potential and achieved high accuracy in distinguishing land cover classes based on EnMAP HSI, according to OA and K_c metrics.

To further evaluate if there is a significance among the classifiers performance, McNemar's test was employed on a 2×2 contingency matrix created for the correctly and incorrectly classified pixels (Abdi, 2020). In the accuracy assessment of a total of 10140 pixels, the misclassified pixels after SVM classification accounted for 956 pixels, while the RF method yielded 1263 misclassified pixels. On the other hand, SVM correctly classified 9185 pixels, and the RF approach correctly classified 8878 pixels (Table 5.3 and 5.4, respectively). The resulting chi-square test statistic value was 42.20, which exceeded the critical chi-square value of 29.59 at a significance level of 0.001. This indicates a significant difference and leads to the rejection of the null hypothesis (Table 5.5).

Table 5.5. The results of McNemar's test for land-use and cover classification accuracy using pairwise algorithms, i.e., SVMs and RF, for f12 and f21 are represents the number of cases that were correctly classified in classifier one but wrongly classified in classifier two and vice versa

Pairwise algorithms	Support Vector Machines vs. Random Forest				
	f12	*f21*	*df.*	Chi-square	p-value
SVM vs. RF	956	1263	10	42.20	0.001

5.1.2. Thematic maps against operational EO land-cover product

To evaluate the performance of each classifier in comparison to operational EO land-cover and use products, the ESA's WorldCover (2021) was used. The latest version of ESA's WorldCover (2021) was selected for this study due to its close temporal proximity to the EnMAP dataset acquired in 2022 and its high spectral resolution. ESA's WorldCover, which utilizes Sentinel-1 and Sentinel-2 data at a 10-meter resolution, emerges as the optimal choice for the current analysis. It not only delivers up-to-date information but also provides superior resolution and a classification scheme that satisfactorily aligns with the characteristics of the thematic maps obtained from the EnMAP dataset. Within the study area, ESA's WorldCover consists of seven land-cover classes: tree cover, grassland, cropland, built-up areas, herbaceous wetlands, permanent water bodies, and bare/sparse vegetation. To ensure compatibility with ESA's WorldCover, a separate dataset was created by merging sub-classes within the primary land-cover categories derived from the classification results. This merging process was performed to align the number of classes with those present in ESA's WorldCover. Specifically, forests were merged to represent tree cover, natural grasslands and pastures were merged to represent grasslands, permanently irrigated and non-irrigated arable lands were merged to represent cropland, artificial surfaces remained as built-up areas, mining sites were renamed as bare/sparse vegetation, representing areas of exposed soil in the study area. For herbaceous wetlands and permanent water bodies, the categories inland marshes and water bodies were renamed, accordingly (Table 5.6).

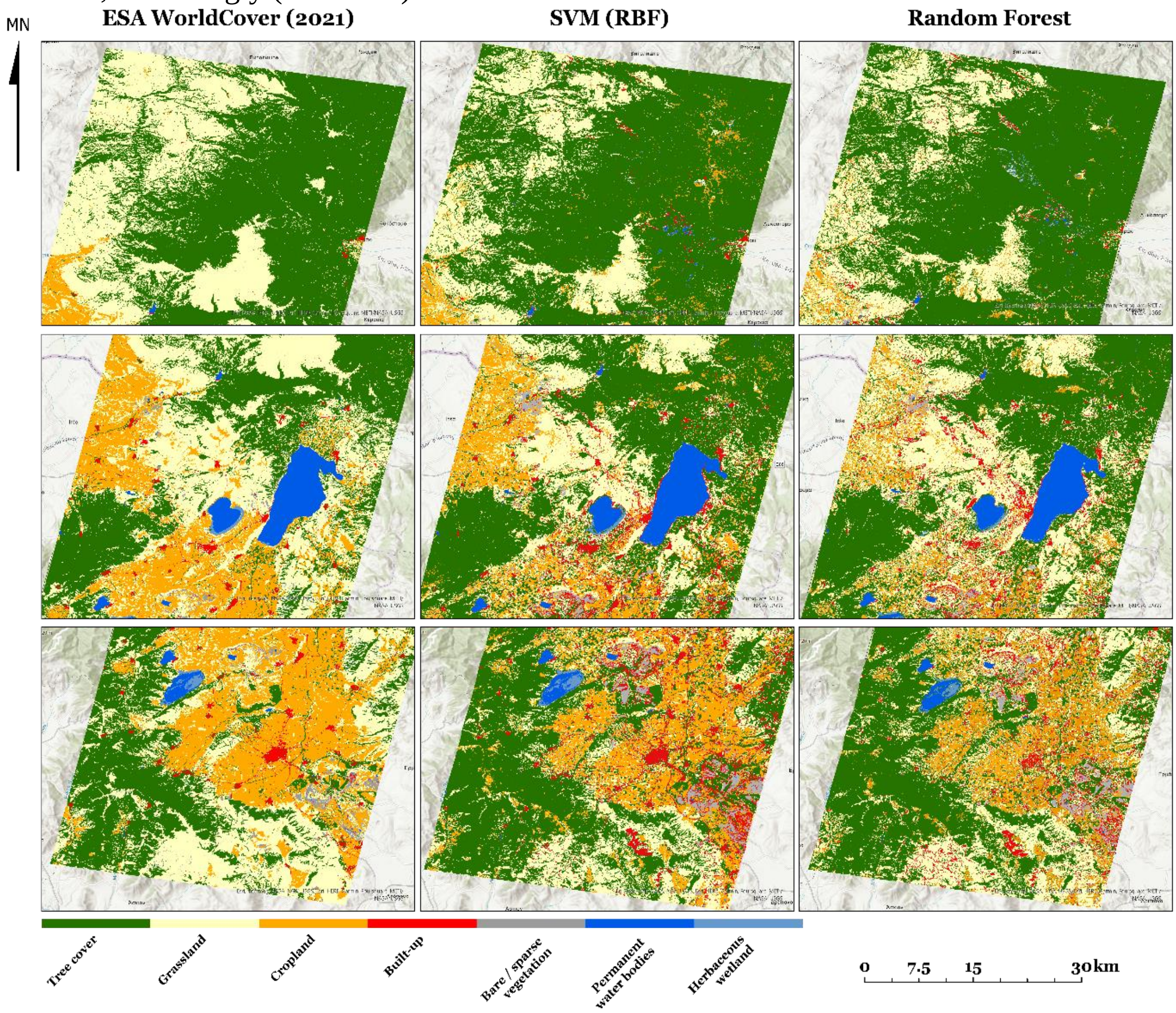


Figure 5.2. Overview of ESAs WorldCover (2021) reference map and thematic land-cover maps obtained from SVMs and RF algorithms, after adjustments in classes of classification of the latter

Table 5.6. Adjustments in classification scheme used in analysis in accordance to ESA's WorldCover

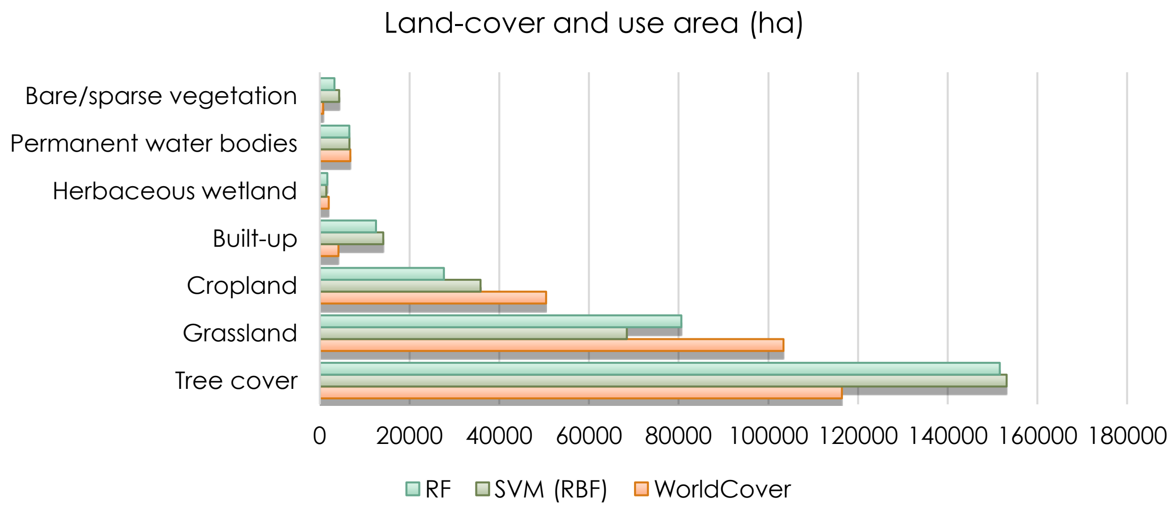
Classes	WorldCover ESA Classes	EnMAP classification scheme Adjustment in classes
1	Tree cover	Broad-leaved forest Coniferous forest Mixed forest, Shrubland
2	Grassland	Natural grasslands Pastures
3	Cropland	Permanently irrigated arable land Non-irrigated arable land
4	Built-up	Artificial surfaces
5	Herbaceous wetland	Inland marshes
6	Permanent water bodies	Water bodies
7	Bare/sparse vegetation	Mining sites

Geometry attributes of the WorldCover dataset were used to calculate the predicted area of each land cover class in hectares, allowing for a spatial comparison of the outcomes thematic maps, in terms of land-cover and land-use predictions. Comparing the classification maps obtained from the analysis with the ESAs WorldCover dataset, it becomes apparent that the former provides a higher level of detail, particularly in regions characterized by greater heterogeneity, despite no significant differences observed in the distribution of land-cover classes (Fig.5.2 & 5.3). With respect to the year difference between the datasets i.e., ESAs WorldCover (2021) and EnMAP datasets (obtained for 2022), significant variations in land-cover and land-use classes, such as tree cover or grasslands, can be attributed to temporal changes in land-use/-cover patterns. Seasonal variations should also be taken into consideration, as grasslands and tree cover can display different phenological patterns and seasonal dynamics. More likely, differences in the adopted classification methodologies and classification scheme for the LULC product. In this study, the classification scheme generated contained classes such as natural grasslands and pastures, but pixels corresponding to transitional woodland and shrubs (which are neither pastures nor natural grasslands) were treated as mixed forest and merged into forest categories i.e., tree cover (Table 5.7). On the other hand, associated land may be correspond to grasslands regarding to ESAs WorldCover. These factors contribute to an underestimation in grasslands land cover class and subsequently an overestimation in tree cover land cover class between ESAs WorldCover and classification outputs.

Table 5.7. Quantification of land-cover classes in ESA's WorldCover and outcome thematic maps after adjustments

	ESAs WorldCover		**SVM (RBF)**		**RF**	
Class	**Area (ha)**	**Cover (%)**	**Area (ha)**	**Cover (%)**	**Area (ha)**	**Cover (%)**
Tree cover	116410.4	41%	153144.87	54%	151616.51	53%
Grassland	103386.92	36%	68456.05	24%	80611.24	28%
Cropland	50459.82	18%	35838.37	13%	27668.33	10%
Built-up	4157.42	1%	14148.79	5%	12531.46	4%
Herbaceous wetland	1996.96	1%	1443.44	1%	1680.2	1%
Permanent water bodies	6807.46	2%	6605.07	2%	6585.54	2%
Bare/sparse vegetation	740.24	0%	4322.62	2%	3265.94	1%
Total (ha)	**283959.25**	**100%**	**283959.25**	**100%**	**283959.25**	**100%**

Figure 5.3. Area-based comparison (in hectares) of ESAs' WorldCover LULC product and thematic maps obtained using EnMAP hyperspectral dataset



Concerning the built-up area an ~3% discrepancy occurred, which may be attributed to the fact that the classification scheme generated in this study, in which urban fabric, construction sites, roads, and associated land were all classified as artificial surfaces. Regarding bare/sparse vegetation, referring to areas with little or no vegetation, classification scheme in this study classifies and identified them as mining sites with bare soil. However, the ESAs WorldCover dataset did not accurately identify the extent of bare soil areas, resulting in only a few hectares being classified, despite the presence of several remarkable mineral extraction sites within the study area (Table 5.7). In this case, it is believed that the classification outputs from the EnMAP dataset achieved a better discrimination of bare areas where in case of ESAs WorldCover, are corresponds to grasslands. A satisfactory agreement between both SVMs and RF classification outputs and the ESAs WorldCover in herbaceous wetland and permanent water bodies, as expected, indicating the efficiency of the obtained results. However, with respect to the temporal difference between two datasets, both classification outputs showed good overall agreement with the ESAs WorldCover, indicating the potential of classification outcomes.

6. Discussion

Chapter 6

This chapter provides a critical evaluation and interpretation of the key findings and limitations of the analysis. It also offers insights into the practical implications of the research for the field of remote sensing and land use/land cover analysis.

6.1. Discussion of classifiers performance

The application of machine learning algorithms for land-cover mapping has gained significant attention over the field of remotely sensed imagery. Broad range techniques based on ML are employed to leverage the capabilities of hyperspectral data (Pandey et al., 2019; Ahmad et al., 2022). Many researchers have been conducted to evaluate the performance of different ML algorithms in LULC classification using hyperspectral sensors (Hasan et al., 2019; Talukdar et al., 2020; Gupta et al., 2021). These studies have been conducted across various scales and settings, achieving different levels of performance and accuracy (Colkesen & Ozturk, 2022). According to Sheykhmousa et al. (2020), there an increase has been noticed in the use of RF and SVM across the globe in a wide range of applications, including vegetation and urban mapping and particularly LULC applications, which had the highest average accuracy of any application.

This research evaluated the efficiency of EnMAP hyperspectral imagery in conjunction with well-established ML algorithms, SVMs with RBF kernel type, and ensemble RF, for LULC mapping in a typical Mediterranean environment. Both classifiers demonstrated precise and accurate results upon eleven land-cover and use classes i.e., two forested and two semi-forest (broadleaved forest, coniferous forest and mixed-shrubs, natural grasslands, respectively), three agricultural classes (permanently irrigated arable, non-irrigated arable land and pastures), two classes representing artificial surfaces (mining sites and total artificial areas), and two classes represents water bodies and wetlands (water bodies and inland marshes, respectively). With respect to the total land cover coverage across the study area, vegetation cover is the dominant land cover class in the region. Both algorithms classified about half of the total land as forested (considered all forested classes represents forests), followed by grasslands (considered that pastures and natural grasslands are both attributed to grasslands) and arable land (considered irrigated and non-irrigated arable land in total as croplands). The area-based comparison of spatial extent in land-cover classes among ESA's WorldCover and LULC thematic maps derived from EnMAP imagery revealed both areas of agreement and discrepancy. Cropland shows a similar trend, with ESAs WorldCover having a larger area than the other classifiers. Built-up areas show relatively high agreement and remarkable agreement is shown for herbaceous wetland and permanent water bodies, while discrepancies between the classifiers regarding bare/sparse vegetation. In vegetation classes referring to tree cover and grasslands exhibits some variation, with ESAs WorldCover reporting a lower and higher coverage percentage compared to SVM and RF, respectively, may be attributed to differences in the classification methodologies and schemes, canopy density due to seasonal variations the temporal difference, as well as different sensors (SAR & multispectral vs. hyperspectral) used for generating LULC thematic maps (Fig.5.3). Overall, with respect to spectral, spatial and temporal differences between datasets, both SVMs and RF classification outputs have shown good agreement the ESAs WorldCover EO land-use and cover product.

The SVM classifier with a radial kernel achieved higher accuracy regarding OA and K_c performance metrics compared to the RF classifier, indicating a stronger agreement with the reference data. Since K_c statistic of two classifiers obtained relatively close, the significance of statistical differences between the accuracy of the models was further evaluated using the McNemar's test. The McNemar's test is well suited for testing significance of result among two algorithms found statistically significant (p-value < 0.001). Consequently, the null hypothesis suggesting equal performance between the SVMs and RF was rejected, suggesting the relatively superior performance of the SVMs approach over RF, in case of current study. However, RF classifier also exhibited notable accuracy for all land-cover classes. In case of SVMs (RBF kernel), best performance achieved for artificial surfaces, broad-leaved forest, and coniferous forest, with UAs ranging from 94% to 99.1%. Overall, the SVM classifier achieved an OA of 90.5% and a K_c of 0.89. Regarding RF, best performance also obtained for artificial surfaces, broad-leaved forest, and coniferous forest, with UAs ranging from 85.1% to 99.4%. The Random Forest classifier achieved an OA of 87.5% and a K_c of 0.86, indicating good agreement between predicted and reference land-cover classes. Both classifiers have also accurately identified mineral extraction sites and arable land, including non-irrigated and irrigated arable land, mixed forest-shrub, water bodies, and inland marshes, while grasslands and pastures exhibited relatively lower levels of accuracy. For both classifiers, grasslands and pastures exhibited comparatively lower accuracies, compared to other land-cover classes. This may be attributed to the inherent complexity and variability of grassland environments, especially when existing with pastures, making them more difficult to accurately classify (Wu et al., 2023). Misclassifications between grasslands and pastures can occur due to the mixing of spectral signatures, particularly given the 30 m spatial resolution of EnMAP's imagery, which makes it difficult to discriminate between areas with similar vegetation characteristics.

Based on the authors' current knowledge, this study is the first attempt to evaluate the potential of recently launched EnMAP hyperspectral satellite imagery in land-use and cover mapping using ML algorithms. Therefore, results obtained from this study could not be directly compared to previous studies in terms of directly comparison regarding the sensors' effects on classification accuracy in a Mediterranean setting. However, in terms of classifiers accuracy the results are comparable to those of previous studies that evaluated the performance of various ML algorithms for land-use and cover classification using hyperspectral datasets (Petropoulos et al., 2012; 2015; Hasan et al., 2019; Gopinath et al., 2020). For example, Lamine et al. (2018), employed SVMs to assess the spatiotemporal changes in land-use and cover using Hyperion hyperspectral data, in a typical Mediterranean setting, using Hyperion hyperspectral data with a 30 m spatial resolution. The studies reported a maximum OA and K_c of 91% and 0.90, respectively, which aligns with the results obtained from SVM classifiers using also a RBF kernel type. Similar in case of results obtained using RF, have shown similar levels of accuracy in land-use and cover classification (Clark, 2017; Colkesen & Ozturk, 2022). The results obtained are also similar to previous studies which indicated that SMVs achieved high accuracy and outperformed compare to other pixel-based classifiers, i.e., RF, employed on different hyperspectral datasets (Puletti et al., 2016; Ghamisi et al., 2017; Raczko & Zagajewski 2017; Cheng & Cheng, 2018; Nhaila et al., 2019). For example, Alcolea et al. (2020) conducted a survey, based on state-of-the-art supervised ML classifiers employed for LULC classification, including SVMs and RF, using various land cover classes, applied to various hyperspectral datasets derived from different airborne campaigns, with diverse spatial and spectral resolutions and different context, i.e., the Indian Pines (IP) from AVIRIS, University of Pavia (UP) from ROSIS, Salinas Valley (SV) from AVIRIS, Kennedy Space Center (KSC) also from AVIRIS and University of Houston (UH) from CASI, Results from this study demonstrated that RF algorithm consistently yields the lowest accuracy values across all datasets, while kernel-based methods, i.e., SVMs, outperformed the other methods (Table 6.1).

Table 6.1. An overview of accuracies performance in LULC mapping with up to 10 land-cover categories, employed on widely-used hyperspectral airborne campaigns

Author Publication	Classes (No.)	Dataset and resolution	Algorithm	OA	KA
Alcolea et al., 2020	16	Indian Pines (20 m)	RF	75.32 ± 0.44	71.42 ± 0.53
			SVM	83.46 ± 0.35	81.08 ± 0.41
	13	Kennedy Space Center (18 m)	RF	88.88 ± 0.43	87.61 ± 0.48
			SVM	90.51 ± 0.56	89.43 ± 0.62
	16	Salinas Valley (3.7 m)	RF	90.08 ± 0.17	88.93 ± 0.19
			SVM	93.2 ± 0.17	92.42 ± 0.19
	15	University of Houston (2.5 m)	RF	73.0 ± 0.07	70.99 ± 0.07
			SVM	76.96 ± 0.0	75.21 ± 0.0
	9	University of Pavia (1.3 m)	RF	86.8 ± 0.25	81.98 ± 0.35
			SVM	93.98 ± 0.15	91.99 0.2

However, the algorithm to be used must be selected according to the size of input datasets, as well as the number of target classes. In addition to the selection of the appropriate classifier, the parameters of each algorithm must be tested under different parameter settings in order for best values of parameters to be found and set, which in turn will increase the performance of the ML algorithm. A wide range of studies have indicated the efficiency in tuning of hyperparameters to increase predictive power of ML models. Comparisons among results obtained based on these studies using SVMs and RF, demonstrated that after hyperparameter tuning, better results are produced than before tuning process, regardless of the datasets used, indicating a direct effect on the performance of the model (Yang & Shami, 2020; Chava et al., 2021). Regarding our study, the high performance achieved by both algorithms can be primarily attributed to the hyperparameter tuning process. For the classification analysis, an automated sklearn's GridSearchCV process employed to optimize the hyperparameters via Scikit-Learn Estimators in EnMAP-Box 3.12.1 using a task-oriented interface. Best estimators achieved exceptional scores for each model (i.e., 99% and 95%, with *cv=5*, for SVMs and RF, respectively). Although, fine-tuned SVMs has been found to produce more accurate results compare to conventional algorithms in supervised classification, SVM tuning process requires selecting a suitable kernel and optimizing kernel parameters, which can involve relatively complex computations. While RF does not require extensive parameter tuning, is more computationally efficient and has been proven to perform well even with default parameter settings (Adugna et al., 2022).

All in all, both classifiers demonstrate robustness to outliers and noise in the data. SVMs use a margin-based approach, in order to find the optimal hyperplane that maximally separates different classes, which helps in mitigating the influence of outliers. On the other hand, RF constructs an ensemble of decision trees and combines their predictions, which reduces the impact of individual noisy or outlier data points. This robustness makes both SVMs and RF suitable algorithms in LULC classification when coupled with HSI. In terms of the potential of SVMs compared to RF regarding the analysis, the high overall accuracy of outcome thematic maps, as reported also in case of other studies, may be linked to the algorithms capability to find optimally separating hyperplanes for classes in contrast to other pixel-based techniques (Petropoulos et al., 2012). For example, for a dataset that contains a lot of noise or outliers, the separating hyperplane cannot accurately approximate the true data distribution and as a result, the provided separating margin, performs unevenly in terms of generalization (Ireland et al., 2015).

6.1.1. Challenges and Limitations

Despite the remarkable potential and advancements in classification of hyperspectral remotely-sensed imageries, there are still challenges regarding that type of data, that must be taken into account. One of the key challenges remain, regarding the complexity and volume issues, refer to the computational and operational cost associated with processing procedures i.e., storage and analysis, that can pose significant constraints on the practicality of these data. The high dimensionality of HSI containing hundreds of narrow, redundant spectral bands make HSI processing extremely complex and time-consuming, requiring advanced hardware and software (Moharram et al., 2023). In that sense, numerous of challenges can arise during the pre-classification procedure with collecting the labelled training set, including the massive time consumption, and difficulty gathering the labelled training samples. The limited number of labelled training samples poses a significant threat for hyperspectral land-cover classification performance and the selection of suitable features in order to handle the variability and diversity of the land cover classes. In the context of this study, a total of 192 spectral bands of hyperspectral EnMAP were used for the classification process. To address the high-volume issues a substantial number of training samples was collected, which in turn introduce an increase in computational costs during the analysis process. In this case, the classification process using RF exhibited a notably faster computational efficiency in contrast with SVM which required a significant amount of time to complete. Regarding to the analysis outcomes. In addition, spatial resolution poses challenges as it may limit the distinction between land-cover classes with similar spectral characteristics, such as grasslands and pastures (Wu et al., 2023). A major drawback of both classifiers is that they don't operate at the sub-pixel level, which could theoretically greatly reduce potential mixture issues resulting from a sensor's coarse spatial resolution and, as a result, misclassification outcomes (Petropoulos et al., 2012).

6.1.2. Key findings

The current study represents the first in the field attempt to assess the capabilities of EnMAP hyperspectral satellite imagery, using a total of 192 spectral bands, for land-use and cover mapping based on a comparative analysis of machine learning (ML) pixel-based algorithms. The application of ML algorithms, SVMs and RF, showcased their effectiveness in mapping eleven land-cover and use classes in a typical Mediterranean environment. The overall pre-processing and processing analysis implemented within the freely available python plug-in EnMAP-Box 3.12.1 in open-source QGIS software. The hyperparameter tuning process using played a crucial role in achieving exceptional scores for both SVMs and RF classifiers. The hyperparameter optimization process performed using sklearn's GridSearchCV automated approach, within a task-oriented and user-friendly interface provided by Scikit-Learn library. The SVM classifier with a radial kernel outperformed RF in terms of overall accuracy metrics, indicating a stronger agreement with the reference data. McNemar's statistic test supported the relatively superior performance of the SVMs over RF. Current finding aligns with previous studies that demonstrated the superiority of SVM against RF for LULC classification using hyperspectral data. In order to further evaluate the performance of ML classifiers in compare with EnMAP data, an area-based assessment performed among obtained land-cover thematic maps from EnMAP dataset and ESAs' WorldCover EO product. The comparison of predicted area in land-cover classes indicated a generally satisfactory agreement beetween most of the LULC categories among datasets. However, notable discrepancies were observed specifically for the grasslands and tree cover classes. The generated thematic maps exhibited a tendency to underestimate the extent of grasslands while overestimating the extent of tree cover. Overall, both SVMs and RF classification outputs have demonstrated good agreement with the ESA's WorldCover EO land-use and cover product, taking into account spectral, spatial, and temporal differences between datasets.

Chapter 7

7. Conclusions & Future Work

This chapter summarizes the objective of the study, which is to evaluate the efficiency of EnMAP in land-use and cover mapping. It highlights the early assessment of EnMAP's effectiveness through the evaluation of commonly used classification algorithms. Moreover, research aims to provide recommendations for future research and directions.

7.1. Concluding remarks

The objective of this study was to assess the combined effectiveness of EnMAP hyperspectral imagery and the performance well-known ML algorithms, namely SVM along with RBF kernel type and ensemble RF for producing land-use/cover thematic maps in a typical Mediterranean environment.

To achieve the research objectives and compare the performance of algorithms, several statistical analyses were performed. These included comparing the overall accuracy of all algorithms and individual classes accuracy, evaluate the statistical significance between algorithms, as well as perform an area-based comparison of outcome thematic maps against ESAs Worldcover EO land-use and cover products. Results of the analysis have shown that EnMAP hyperspectral data demonstrated a highly promising potential in the field of land-cover and use mapping. The key study findings are summarised as follow:

- Comparing the classification results of SVMs and the RF approaches, based on EnMAP HSI, SVMs exhibited showing higher OA and K_c accuracy values. For SVMs an OA and K_c of 90.5% and 0.89 obtained, respectively. While RF produced also high OA and K_c accuracy of 87.5% and 0.86 respectively. The superiority of SVMs was further supported by the McNemar’s chi-square statistic. However, it is important to note that the RF classifier also exhibited notable OA and accuracies for individual land-cover classes. Thus, both classifiers can be considered to be appropriate for land-cover and use mapping using EnMAP HSI.
- Further, an area-based comparison of outcomes performed against ESAs WorldCover operational EO product. Considering the differences in methodological approach and the temporal disparity between the two datasets, both thematic maps i.e., EnMAP HSI outcomes, demonstrated a satisfactory agreement in land-cover classes covered area with ESAs WorldCover. This suggests that further investigation and exploration should be conducted regarding the promising potential of the latter for operational applications. Such studies as the current one can be linked to ongoing efforts aimed at enhancing and improving operational products.
- In regard to the classification analysis, a total of 192 spectral bands of L2A EnMAP HSI were used. To tackle the challenges related to the large volume of data, a significant number of training samples (~30,000) were collected. However, this increase in the amount of data introduces additional computational costs during the analysis process. It should be taken into account the large and high-volume size of input datasets, since all high-quality bands were used, and no reduction or optimal band selection techniques were undertaken.

- The hyperparameter tuning process using played a crucial role in achieving exceptional scores for both SVMs and RF algorithms (99% and 95%, respectively). Regarding the parameter settings and computational time, parameters for RF are easiest to be set in contrast with SVM (selection of appropriate kernel type and parameter values), in addition, the computational time for RF classification process is significant lesser form SMVs. Thus, in terms of computation cost showed to be RF is more computationally efficient.
- The overall analysis implemented in open-source QGIS 3.30.3 software using EnMAP-Box 3.12.1 within python Scikit-Learn (sklearn) library, allowed for a cost-effective implementation of the analysis. The combination of an open-source and user-friendly interface, and the availability of open-access hyperspectral data resources coupled with advancement in ML provides a valuable tool for researchers working in the field of hyperspectral data analysis and disciplines.

To the best of the authors' knowledge, the findings of this study contribute to one of the initial assessments of the potential of EnMAP hyperspectral satellite imagery in LULC mapping and to provide a comparison among most commonly used ML pixel-based algorithms in the field land-cover classification.

The findings of this study highlight the significance of evaluating the newly deployed EnMAP hyperspectral satellite data in combination with advanced ML classification approaches. The study not only evaluates EnMAP's potential but also compares its performance using well-established methods, thus contributing to the existing knowledge and understanding in the field. This research contributes to the field by addressing the unique challenges associated with land-cover classification in Mediterranean environments. Findings can serve as valuable insights and can be further used for future comparison studies, regarding the performance of EnMAP hyperspectral data for applications related to land-use and cover in research areas with similar characteristics. By assessing the ability of these methods to be applied in diverse environmental contexts or landscapes with similar characteristics, their broader applicability and reliability can be evaluated, allowing for potential future utilization.

7.2. Future research directions

Further research needs to be conducted for the exploitation of the possibilities of this new satellite by employing various state-of-the-art spectral-spatial approaches in multi-seasonal datasets. It is noting that findings obtained from this study should be viewed as exploratory since were based exclusively on a single analysis of EnMAP HSI. The analysis was based on images from a single acquisition and the classification was performed using the spectral response of features during that specific time period, without supplementary time-series data. By incorporating relevant multi-temporal satellite data and considering the seasonal dynamics and patterns of vegetation classes, allow to capture seasonal variations, facilitating the discrimination among categories that require time-series information for precise detection, such as irrigated and non-irrigated land, and various vegetation classes. Regarding the vegetation classes, such as coniferous forests, broad-leaved forests, grasslands, and croplands, undergo noticeable changes throughout different seasons such as growing or peak season. Considering the seasonal dynamics and phenological patterns of vegetation classes is crucial for achieving precise land-cover classification results.

The integration of data from various sensors, such as multispectral and radar, can effectively overcome these challenges. Additionally, incorporating ancillary data, such as topographical information, can improve the precision and reliability of the outcomes. EnMAP's high spectral and temporal resolution could also contribute to improved mapping and monitoring of large-scale agricultural areas, facilitating early identification of vegetation stress and monitoring of crop growth dynamics. Overall, EnMAP is a valuable tool for boosting precision agriculture, improving crop health and ultimately leading to increased productivity. All in all, EnMAP offers advantages in discriminating land-cover types, enabling efficient and cost-effective mapping of land use and land cover over large, inaccessible regions.

Last but not least, the framework of the study aims to contribute to ongoing efforts focused on enhancing and refining existing operational products. The findings indicate that EnMAP holds promise for improving existing EO-based LULC operational products in various fields, including land-use planning, environmental monitoring, and natural resource management. These findings emphasize the potential of integrating EnMAP's data into existing frameworks.

References

Abburu, S., & Golla, S. B. (2015). Satellite Image Classification Methods and Techniques: A Review. *International Journal of Computer Applications*, *119*(8), 20–25. https://doi.org/10.5120/21088-3779

Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, *57*(1), 1–20. https://doi.org/10.1080/15481603.2019.1650447

Adugna, T., Xu, W., & Fan, J. (2022). Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote Sensing*, *14*(3), 574. https://doi.org/10.3390/rs14030574

Ahmad, M., Shabbir, S., Roy, S. K., Hong, D., Wu, X., Yao, J., Khan, A. M., Mazzara, M., Distefano, S., & Chanussot, J. (2022). Hyperspectral Image Classification—Traditional to Deep Models: A Survey for Future Prospects. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 968–999. https://doi.org/10.1109/jstars.2021.3133021

Alcolea, A. A., Paoletti, M. E., Haut, J. M., Resano, J., & Plaza, A. (2020). Inference in Super-vised Spectral Classifiers for On-Board Hyperspectral Imaging: An Overview. *Remote Sensing*, *12*(3), 534. https://doi.org/10.3390/rs12030534

Al-Doski, J., Mansor, S. B., Ng, H., San, P., & Khuzaimah, Z. (2020). Land cover mapping using remote sensing data. *Am. J. Geogr. Inf. Syst*, *2020*, 33-45. 10.5923/j.ajgis.20200901.04

Al-Doski, J., Mansorl, S. B., & Shafri, H. Z. M. (2013). Image classification in remote sensing. Department of Civil Engineering, Faculty of Engineering, University Putra, Malaysia, 3(10). ISSN 2225-0948

Amini, S., Homayouni, S., Safari, A., & Darvishsefat, A. A. (2018). Object-based classification of hyperspectral data using Random Forest algorithm. *Geo-spatial Information Science*, *21*(2), 127–138. https://doi.org/10.1080/10095020.2017.1399674

Aneece, I. P., & Thenkabail, P. S. (2021). Classifying Crop Types Using Two Generations of Hyperspectral Sensors (Hyperion and DESIS) with Machine Learning on the Cloud. *Remote Sensing*, *13*(22), 4704. https://doi.org/10.3390/rs13224704

Aneece, I., & Thenkabail, P. S. (2022). New Generation Hyperspectral Sensors DESIS and PRISMA Provide Improved Agricultural Crop Classifications. *Photogrammetric Engineering and Remote Sensing*, *88*(11), 715–729. https://doi.org/10.14358/pers.22-00039r2

Asam, S., Schwarzenbacher, F., Marshall, D., & Bachmann, M. (2022). Detection of Grassland Degradation in Azerbaijan By Combining Multi-Decadal NDVI Time Series And Fractional Cover Estimates Based On DESIS Data. https://elib.dlr.de/187095/

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *Isprs Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Ben-Dor, E., Chabrillat, S., Demattê, J. a. M., Taylor, G. N., Hill, J. C., Whiting, M., & Sommer, S. S. (2009). Using Imaging Spectroscopy to study soil properties. *Remote Sensing of Environment*, *113*, S38–S55. https://doi.org/10.1016/j.rse.2008.09.019

Breiman, L. (2001). Random Forest. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/a:1010933404324

Buckingham, R., & Staenz, K. (2008). Review of current and planned civilian space hyperspectral sensors for EO. *Canadian Journal of Remote Sensing*, *34*(sup1), S187–S197. https://doi.org/10.5589/m08-002

Carranza-García, M., García-Gutiérrez, J., & Riquelme, J. C. (2019). A Framework for Evaluating Land Use and Land Cover Classification Using Convolutional Neural Networks. *Remote Sensing*, *11*(3), 274. https://doi.org/10.3390/rs11030274

Chabrillat et al., "EnMAP Pre-Launch and Start Phase: Mission Update," IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022, pp. 5000-5003, doi: 10.1109/IGARSS46834.2022.9884773.

Chander, S., Gujrati, A., Krishna, A. S., Sahay, A., & Singh, R. P. (2020). Remote sensing of inland water quality: a hyperspectral perspective. In *Elsevier eBooks* (pp. 197–219). https://doi.org/10.1016/b978-0-08-102894-0.00017-6

Chandra, M. A., & Bedi, S. (2018). Survey on SVM and their application in image classification. *International Journal of Information Technology*, *13*(5), 1–11. https://doi.org/10.1007/s41870-017-0080-1

Cheng, G., Han, J., & Lu, X. (2017). Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proceedings of the IEEE*, *105*(10), 1865–1883. https://doi.org/10.1109/jproc.2017.2675998

Chintada, K.R., Yalla S.P. and Uriti, A. (2021). A Deep Belief Network Based Land Cover Classification. Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, pp. 1-5, doi: 10.1109/i-PACT52855.2021.9696524.

Chisense, C. (2012). Classification of Roof Materials using Hyperspectral Data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B7*, 103–107. https://doi.org/10.5194/isprsarchives-xxxix-b7-103-2012

Clark, M. M. (2017). Comparison of simulated hyperspectral HyspIRI and multispectral Landsat 8 and Sentinel-2 imagery for multi-seasonal, regional land-cover mapping. *Remote Sensing of Environment*, *200*, 311–325. https://doi.org/10.1016/j.rse.2017.08.028

Colkesen, I., & Ozturk, M. B. (2022). A comparative evaluation of state-of-the-art ensemble learning algorithms for land cover classification using WorldView-2, Sentinel-2 and ROSIS imagery. *Arabian Journal of Geosciences*, *15*(10). https://doi.org/10.1007/s12517-022-10243-x

Congalton, R. G. (2001). Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, *10*(4), 321. https://doi.org/10.1071/wf01031

Constans, Y., Fabre, S., Seymour, M. H., Crombez, V., Deville, Y., & Briottet, X. (2021). Hyperspectral Pansharpening in the Reflective Domain with a Second Panchromatic Channel in the SWIR II Spectral Domain. *Remote Sensing*, *14*(1), 113. https://doi.org/10.3390/rs14010113

Coulter, D., Hauff, P. L., & Kerby, W. L. (2007). Airborne hyperspectral remote sensing. In *Proceedings of Exploration* (Vol. 7, pp. 375-386).

Curran, P. J. (2016). Principles of remote sensing. Routledge.

Dabija, A., Kluczek, M., Zagajewski, B., Raczko, E., Kycko, M., Al-Sulttani, A. H., Tardà, A., Pineda, L., & Corbera, J. (2021). Comparison of Support Vector Machines and Random Forests for Corine Land Cover Mapping. Remote Sensing, 13(4), 777. https://doi.org/10.3390/rs13040777

Datta, D., Mallick, P. K., Bhoi, A. K., Ijaz, M., Shafi, J., & Choi, J. (2022). Hyperspectral Image Classification: Potentials, Challenges, and Future Directions. *Computational Intelligence and Neuroscience*, *2022*, 1–36. https://doi.org/10.1155/2022/3854635

De Leeuw, J., Jia, H. Y., Yang, L., Liu, X., Schmidt, K., & Skidmore, A. K. (2006). Comparing accuracy assessments to infer superiority of image classification methods. *International*

Journal of Remote Sensing, *27*(1), 223–232. https://doi.org/10.1080/01431160500275762

Dhingra, S., & Kumar, D. (2019). A review of remotely sensed satellite image classification. *International Journal of Electrical and Computer Engineering*, *9*(3), 1720. https://doi.org/10.11591/ijece.v9i3.pp1720-1731

Dou, P., & Zeng, C. (2020). Hyperspectral Image Classification Using Feature Relations Map Learning. *Remote Sensing*, *12*(18), 2956. https://doi.org/10.3390/rs12182956

Dubovik, O., Schuster, G. L., Xu, F., Hu, Y., Boesch, H., Landgraf, J., & Li, Z. (2021). Grand Challenges in Satellite Remote Sensing. *Front. Remote Sens.*, *2*. https://doi.org/10.3389/frsen.2021.619818

EnMAP-Box Developers (2019): EnMAP-Box 3 - A QGIS Plugin to process and visualize hyperspectral remote sensing data. https://enmap-box.readthedocs.io/en/latest

European Environment Agency (EEA). 2020. State of nature in the EU. Results from reporting under the nature directives 2013-2018. Report 10/2020. *EEA*, Luxembourg

FAO (2022). Framework for Environmental and Social Management. Rome. https://doi.org/10.4060/cb9870en

FAO (2022). The future of food and agriculture: Drivers and triggers for transformation – Summary version. Rome. https://doi.org/10.4060/cc1024en

FAO (2022). The State of the World's Land and Water Resources for Food and Agriculture – Systems at breaking point. Main report. Rome. https://doi.org/10.4060/cb9910en

Farmonov, N., Amankulova, K., Szatmári, J., Sharifi, A., Abbasi-Moghadam, D., Nejad, S. M. M., & Mucsi, L. (2023). Crop Type Classification by DESIS Hyperspectral Imagery and Machine Learning Algorithms. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 1576–1588. https://doi.org/10.1109/jstars.2023.3239756

Foody, G. M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(6), 1335–1343. https://doi.org/10.1109/tgrs.2004.827257

Fowler, J. H. (2014). Compressive pushbroom and whiskbroom sensing for hyperspectral remote-sensing imaging. https://doi.org/10.1109/icip.2014.7025137

Ghamisi, P., Plaza, J., Chen, Y., Li, J., & Plaza, A. (2017). Advanced Spectral Classifiers for Hyperspectral Images: A review. *IEEE Geoscience and Remote Sensing Magazine*, *5*(1), 8–32. https://doi.org/10.1109/mgrs.2016.2616418

Gómez, C. A., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *Isprs Journal of Photogrammetry and Remote Sensing*, *116*, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Gopinath, G., Sasidharan, N., & Surendran, U. (2020). Landuse classification of hyperspectral data by spectral angle mapper and support vector machine in humid tropical region of India. *Earth Science Informatics*, *13*(3), 633–640. https://doi.org/10.1007/s12145-019-00438-4

Goswami, C., Singh, N. K., & Handique, B. K. (2020). Hyperspectral Spectroscopic Study of Soil Properties- A Review. *International Journal of Plant and Soil Science*, 14–25. https://doi.org/10.9734/ijpss/2020/v32i730301

Green, R. C., Eastwood, M., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J., Pavri, B., Chovit, C. J., Solis, M. I. V., Olah, M. R., & Williams, O. (1998). Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVI-RIS). *Remote Sensing of Environment*, *65*(3), 227–248. https://doi.org/10.1016/s0034-4257(98)00064-9

Guanter, L., Brell, M., Chan, J. C., Giardino, C., Gomez-Dans, J., Mielke, C., Morsdorf, F., Segl, K., & Zhu, X. X. (2018). Synergies of Spaceborne Imaging Spectroscopy with Other Remote Sensing Approaches. *Surveys in Geophysics*, *40*(3), 657–687. https://doi.org/10.1007/s10712-018-9485-z

Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., Kuester, T., Hollstein, A., Rossner, G., Chlebek, C., Straif, C. J., Fischer, S., Schrader, S., Storch, T., Heiden, U., Mueller, A., Bachmann, M. F., Mühle, H., Müller, R., . . . Sang, B. (2015). The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sensing*, *7*(7), 8830–8857. https://doi.org/10.3390/rs70708830

Gudo, A. J. A., Deng, J., & Qureshi, A. S. (2022). Analysis of Spatiotemporal Dynamics of Land Use/Cover Changes in Jubek State, South Sudan. *Sustainability*, *14*(17), 10753. https://doi.org/10.3390/su141710753

Gupta, D. B., Sethi, D., & Bathija, R. (2021). A Review on Land-use and Land-change with Machine Learning Algorithm. *IOP Conference Series*. https://doi.org/10.1088/1757-899x/1119/1/012006

Gupta, P., & Venkatesan, M. (2020). Mineral Identification Using Unsupervised Classification from Hyperspectral Data. In *Springer eBooks* (pp. 259–268). https://doi.org/10.1007/978-981-15-0135-7_25

Guryanov, M. A., Гурьянов, M., Борзов, C. M., & Borzov, S. M. (2017). Spectral-Spatial Classification of Vegetative Cover Types Using Hyperspectral Data. *Vestnik Novosibirskogo Gosudarstvennogo Universiteta*, *15*(4), 14–21. https://doi.org/10.25205/1818-7900-2017-15-4-14-21

Hasan, H., Shafri, H. Z. M., & Habshi, M. (2019). A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models For Hyperspectral Image Classification. *IOP Conference Series*, *357*(1), 012035. https://doi.org/10.1088/1755-1315/357/1/012035

Heiden, U., Heldens, W., Roessner, S., Segl, K., Esch, T., & Mueller, A. (2012). Urban structure type characterization using hyperspectral remote sensing and height information. *Landscape and Urban Planning*, *105*(4), 361–375. https://doi.org/10.1016/j.landurbplan.2012.01.001

Heiden, Uta and Alonso, Kevin and Bachmann, Martin and Burch, Kara and Carmona, Emiliano and Cerra, Daniele and Dietrich, Daniele and Knodt, Uwe and Krutz, David and Lester, Heath and Marshall, David and Müller, Rupert and de los Reyes, Raquel and Reinartz, Peter and Tegler, Mirco (2022). The Spaceborne Imaging Spectrometer DESIS: Mission summary and potential for scientific developments. In: 12th EARSeL Workshop on Imaging Spectroscopy. 12th EARSeL Workshop on Imaging Spectroscopy, 22.-24. June 2022, Potsdam, Germany. http://is.earsel.org/workshop/12-IS-Potsdam2022/wp-content/uploads/2022/06/programm-booklet_long_complete_v2.pdf

Jafarzadeh, H., Mahdianpari, M., Gill, E. W., Mohammadimanesh, F., & Homayouni, S. (2021). Bagging and Boosting Ensemble Classifiers for Classification of Multispectral, Hyperspectral and PolSAR Data: A Comparative Evaluation. *Remote Sensing*, *13*(21), 4405. https://doi.org/10.3390/rs13214405

Jensen, J. R. (2015). *Introductory Digital Image Processing: A Remote Sensing Perspective*. Prentice Hall.

Jia, J., Chen, X., Chen, J., Guo, R., Shu, R., & Wang, J. (2020). Status and application of advanced airborne hyperspectral imaging technology: A review. *Infrared Physics & Technology*, *104*, 103115. https://doi.org/10.1016/j.infrared.2019.103115

Jia, S., Jiang, S., Lin, Z., Zhang, G., Xu, M., & Yu, S. (2021). A survey: Deep learning for hyperspectral image classification with few labeled samples. *Neurocomputing*, *448*, 179–204. https://doi.org/10.1016/j.neucom.2021.03.035

Jog, S., & Dixit, M. (2016). *Supervised classification of satellite images*. https://doi.org/10.1109/casp.2016.7746144

Kalantar, B., Seydi, S. T., Ueda, N., Saeidi, V., Halin, A. A., & Shabani, F. (2022). Deep Ensemble Learning for Land Cover Classification Based on Hyperspectral Prisma Image. In *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*. https://doi.org/10.1109/igarss46834.2022.9884203

Kale, K. V., Solankar, M. M., Nalawade, D. B., Dhumal, R. K., & Gite, H. R. (2017). A Research Review on Hyperspectral Data Processing and Analysis Algorithms. *Proceedings of the National Academy of Sciences, India. Section a, Physical Sciences*, *87*(4), 541–555. https://doi.org/10.1007/s40010-017-0433-y

Kaufmann, H. H., Segl, K., Guanter, L., Hofer, S. O., Foerster, K., Stuffler, T., Mueller, A., Richter, R. P., Bach, H., Hostert, P., & Chlebek, C. (2008). *Environmental Mapping and Analysis Program (EnMAP) - Recent Advances and Status*. https://doi.org/10.1109/igarss.2008.4779668

Kaufmann, H., Segl, K., Itzerott, S., Bach, H., Wagner, A., Hill, J. C., Heim, B., Oppermann, K., Heldens, W., Stein, E., Müller, A., Van Der Linden, S., Leitão, P. J., Rabe, A., & Hostert, P. (2010). Hyperspectral Algorithms: Report in the frame of EnMAP preparation activities. *Potsdam: Deutsches GeoForschungsZentrum GFZ*, 1–268. https://doi.org/10.2312/gfz.b103-10089

Kaul, A, Raina, S. Support vector machine versus convolutional neural network for hyperspectral image classification: A systematic review. Concurrency Computat Pract Exper. 2022; 34(15): e6945. doi:10.1002/cpe.6945

Khan, S. S., Khan, M. A., Haider, S., & Damaševičius, R. (2022). Hyperspectral image classification using NRS with different distance measurement techniques. *Multimedia Tools and Applications*, *81*(17), 24869–24885. https://doi.org/10.1007/s11042-022-12263-x

Khosravi, I., & Jouybari-Moghaddam, Y. (2019). Hyperspectral Imbalanced Datasets Classification Using Filter-Based Forest Methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(12), 4766–4772. https://doi.org/10.1109/jstars.2019.2914668

Kokal, A. T., İsmailoğlu, İ. İ., & Musaoglu, N. (2022). COMPARISON OF LANDSAT-9 AND PRISMA SATELLITE DATA FOR LAND USE / LAND COVER CLASSIFICATION. *The International Archives of the Photogrammetry, Remote Sensing and Spatial*

Information Sciences, *XLVI-M-2–2022*, 197–201. https://doi.org/10.5194/isprs-archives-xlvi-m-2-2022-197-2022

Kumar, C., Chatterjee, S., & Oommen, T. (2020). Mapping hydrothermal alteration minerals using high-resolution AVIRIS-NG hyperspectral data in the Hutti-Maski gold deposit area, India. *International Journal of Remote Sensing*, *41*(2), 794–812. https://doi.org/10.1080/01431161.2019.1648906

Kumar, P., Pandey, P. C., Kumar, V., Singh, B. K., Tomar, V., & Rani, M. (2015). Efficient Recognition of Forest Species Biodiversity by Inventory-Based Geospatial Approach Using LISS IV Sensor. *IEEE Sensors Journal*, *15*(3), 1884–1891. https://doi.org/10.1109/jsen.2014.2368394

Lamine, S., Petropoulos, G. P., Brewer, P., Bachari, N. E. I., Srivastava, P. K., Manevski, K., Kalaitzidis, C., & Macklin, M. G. (2019). Heavy Metal Soil Contamination Detection Using Combined Geochemistry and Field Spectroradiometry in the United Kingdom. *Sensors*, *19*(4), 762. https://doi.org/10.3390/s19040762

Lamine, S., Petropoulos, G. P., Singh, S. K., Szabó, S., Bachari, N. E. I., Srivastava, P. K., & Suman, S. (2018). Quantifying land use/land cover spatio-temporal landscape pattern dynamics from Hyperion using SVMs classifier and FRAGSTATS®. *Geocarto International*, *33*(8), 862–878. https://doi.org/10.1080/10106049.2017.1307460

Lassalle, G., Fabre, S., Credoz, A., Hédacq, R., Dubucq, D., & Elger, A. (2021). Mapping leaf metal content over industrial brownfields using airborne hyperspectral imaging and optimized vegetation indices. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-020-79439-z

Li, C., Wang, Y., Zhang, X., Gao, H., Yang, Y., & Wang, J. (2019). Deep Belief Network for Spectral–Spatial Classification of Hyperspectral Remote Sensor Data. *Sensors*, *19*(1), 204. https://doi.org/10.3390/s19010204

Li, S., Song, W., Fang, L., Chen, Y., Ghamisi, P., & Benediktsson, J. A. (2019). Deep Learning for Hyperspectral Image Classification: An Overview. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(9), 6690–6709. https://doi.org/10.1109/tgrs.2019.2907932

Li, X., Chen, W., Cheng, X., & Wang, L. (2016). A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery. Remote Sensing, 8(6), 514-540. https://doi.org/10.3390/rs8060514

Lianjun Chen and Xinwen Cheng, 2016. Classification of High-resolution Remotely Sensed Images Based on Random Forests. *Journal of Software Engineering*, 10: 318-327. 10.3923/jse.2016.318.327

Liu, M., Yu, T., Gu, X., Sun, Z., Yang, J., Zhang, Z., Mi, X., Cao, W., & Li, J. (2020). The Impact of Spatial Resolution on the Classification of Vegetation Types in Highly Fragmented Planting Areas Based on Unmanned Aerial Vehicle Hyperspectral Images. *Remote Sensing*, *12*(1), 146. https://doi.org/10.3390/rs12010146

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870. https://doi.org/10.1080/01431160600746456

Lv, Q., Dou, Y., Niu, X., Xu, J., & Li, B. (2014). *Classification of land cover based on deep belief networks using polarimetric RADARSAT-2 data*. https://doi.org/10.1109/igarss.2014.6947537

Lv, W., & Wang, X. (2020). Overview of Hyperspectral Image Classification. *Journal of Sensors*, *2020*, 1–13. https://doi.org/10.1155/2020/4817234

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. F. G. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *Isprs Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

Macarringue, L. S., Bolfe, É. L., & Pereira, P. a. A. (2022). Developments in Land Use and Land Cover Classification Techniques in Remote Sensing: A Review. *Journal of Geographic Information System*, *14*(01), 1–28. https://doi.org/10.4236/jgis.2022.141001

Manakos, I., & Braun, M. (2014). *Land Use and Land Cover Mapping in Europe*: Practices & Trends. Springer.

Manandhar, R., Odeh, I. O. A., & Ancev, T. (2009). Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement. *Remote Sensing*, *1*(3), 330–344. https://doi.org/10.3390/rs1030330.

Manian, V., Alfaro-Mejía, E., & Tokars, R. (2022). Hyperspectral Image Labeling and Classification Using an Ensemble Semi-Supervised Machine Learning Approach. *Sensors*, *22*(4), 1623. https://doi.org/10.3390/s22041623

Marcello, J., Rodríguez-Esparragón, D., & Moreno, D. (2018). *Comparison of Land Cover Maps Using High Resolution Multispectral and Hyperspectral Imagery*. https://doi.org/10.1109/igarss.2018.8517878

Marcinkowska-Ochtyra, A., Zagajewski, B., Ochtyra, A., Jarocińska, A., Wojtuń, B., Rogass, C., Mielke, C., & Lavender, S. (2017). Subalpine and alpine vegetation classification based on hyperspectral APEX and simulated EnMAP images. *International Journal of Remote Sensing*, *38*(7), 1839–1864. https://doi.org/10.1080/01431161.2016.1274447

Maschler, J., Atzberger, C., & Immitzer, M. (2018). Individual Tree Crown Segmentation and Classification of 13 Tree Species Using Airborne Hyperspectral Data. *Remote Sensing*, *10*(8), 1218. https://doi.org/10.3390/rs10081218

Matese, A., Berton, A., Di Gennaro, S.F., Gatti, M., Squeri, C., Poni, S. 2021. Testing performance of UAV-based hyperspectral imagery in viticulture. In Proceedings of the Precision Agriculture; Wageningen Academic Publishers: Wageningen, The Netherlands; Budapest, Hungary, 2021; pp. 509–516. 10.3920/978-90-8686-916-9_61

Mather, P. J. (1987). Computer processing of remotely-sensed images — An introduction. *Geocarto International*, *2*(4), 64. https://doi.org/10.1080/10106048709354125

Maulik, U., & Chakraborty, D. (2017). Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques. *IEEE Geoscience and Remote Sensing Magazine*, *5*(1), 33–52. https://doi.org/10.1109/mgrs.2016.2641240

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157. https://doi.org/10.1007/bf02295996

Milewski, R., Schmid, T., Chabrillat, S., Jiménez, M., Escribano, P., Pelayo, M., & Ben-Dor, E. (2022). Analyses of the Impact of Soil Conditions and Soil Degradation on Vegetation Vitality and Crop Productivity Based on Airborne Hyperspectral VNIR–SWIR–TIR Data in a Semi-Arid Rainfed Agricultural Area (Camarena, Central Spain). *Remote Sensing*, *14*(20), 5131. https://doi.org/10.3390/rs14205131

Moharram, M. A., & Sundaram, D. M. (2023). Land use and land cover classification with hyperspectral data: A comprehensive review of methods, challenges and future directions. *Neurocomputing*, *536*, 90–113. https://doi.org/10.1016/j.neucom.2023.03.025

Mou, L., Bruzzone, L., & Zhu, X. X. (2018). Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(2), 924–935. https://doi.org/10.1109/tgrs.2018.2863224

Mughees, A., & Tao, L. (2021). Deep Learning for Hyperspectral Image Analysis and Classification. In *Engineering Applications of Computational Methods*. Springer Nature. https://doi.org/10.1007/978-981-33-4420-4

Navin, M. S., & Agilandeeswari, L. (2020). Multispectral and hyperspectral images-based land use / land cover change prediction analysis: an extensive review. *Multimedia Tools and Applications*, *79*(39–40), 29751–29774. https://doi.org/10.1007/s11042-020-09531-z

Nhaila, H., Elmaizi, A., Sarhrouni, E., & Hammouch, A. (2019). Supervised classification methods applied to airborne hyperspectral images: comparative study using mutual information. *Procedia Computer Science*, *148*, 97–106. https://doi.org/10.1016/j.procs.2019.01.013

Nisha, A, & Anitha, A, (2022). Current Advances in Hyperspectral Remote Sensing in Urban Planning. In *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*. https://doi.org/10.1109/icicict54557.2022.9917771

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217–222. https://doi.org/10.1080/01431160412331269698

Pal, M., & Mather, P. M. (2005). Support vector machines for classification in remote sensing. *International Journal of Remote Sensing*, *26*(5), 1007–1011. https://doi.org/10.1080/01431160512331314083

Pal, M., Rasmussen, T. M., & Porwal, A. (2020). Optimized Lithological Mapping from Multispectral and Hyperspectral Remote Sensing Images Using Fused Multi-Classifiers. *Remote Sensing*, *12*(1), 177. https://doi.org/10.3390/rs12010177

Pandey, D. N., & Tiwari, K. N. (2020). Feature identification and extraction of urban built-up surfaces and materials in AVIRIS-NG hyperspectral imagery. *Geocarto International*. https://doi.org/10.1080/10106049.2020.1797187

Pandey, P. C., Koutsias, N., Petropoulos, G. P., Srivastava, P. K., & Ben-Dor, E. (2019). Land use/land cover in view of earth observation: data sources, input dimensions, and classifiers—a review of the state of the art. *Geocarto International*, *36*(9), 957–988. https://doi.org/10.1080/10106049.2019.1629647

Pandey, P. C., Srivastava, P. K., Balzter, H., Bhattacharya, B., & Petropoulos, G. P. (2020). *Hyperspectral Remote Sensing: Theory and Applications*. Elsevier.

Paoletti, M. E., Haut, J. M., Plaza, J., & Plaza, A. (2018). Deep&Dense Convolutional Neural Network for Hyperspectral Image Classification. *Remote Sensing*, *10*(9), 1454. https://doi.org/10.3390/rs10091454

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. (2011). Scikit-learn: Machine Learning in {P}ython. Journal of Machine Learning Research, 12, 2825--2830}.

Petropoulos, G. P., Arvanitis, K., & Sigrimis, N. (2012). Hyperion hyperspectral imagery analysis combined with machine learning classifiers for land use/cover mapping. *Expert Systems with Applications*, *39*(3), 3800–3809. https://doi.org/10.1016/j.eswa.2011.09.083

Petropoulos, G. P., Kalivas, D., Georgopoulou, I. A., & Srivastava, P. K. (2015). Urban vegetation cover extraction from hyperspectral imagery and geographic information system spatial analysis techniques: case of Athens, Greece. *Journal of Applied Remote Sensing*, *9*(1), 096088. https://doi.org/10.1117/1.jrs.9.096088

Petropoulos, G. P., Vadrevu, K. P., & Kalaitzidis, C. (2013). Spectral angle mapper and object-based classification combined with hyperspectral remote sensing imagery for obtaining land use/cover mapping in a Mediterranean region. *Geocarto International*, *28*(2), 114–129. https://doi.org/10.1080/10106049.2012.668950

Peyghambari, S., & Zhang, Y. (2021). Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review. *Journal of Applied Remote Sensing*, *15*(03). https://doi.org/10.1117/1.jrs.15.031501

Puletti, N., Camarretta, N., & Corona, P. (2016). Evaluating EO1-Hyperion capability for mapping conifer and broadleaved forests. *European Journal of Remote Sensing*, *49*(1), 157–169. https://doi.org/10.5721/eujrs20164909

Qian, S. (2021). Hyperspectral Satellites, Evolution, and Development History. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 7032–7056. https://doi.org/10.1109/jstars.2021.3090256

Rabe, A., Jakimow, B., Thiel, F., Hostert, P., & Van Der Linden, S. (2018). *EnMAP-Box 3 a free and open source Python plug-in for QGIS*. https://doi.org/10.1109/igarss.2018.8518854

Raczko, E., & Zagajewski, B. (2017). Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral

APEX images. *European Journal of Remote Sensing*, *50*(1), 144–154. https://doi.org/10.1080/22797254.2017.1299557

Rast, M., & Painter, T. H. (2019). Earth Observation Imaging Spectroscopy for Terrestrial Systems: An Overview of Its History, Techniques, and Applications of Its Missions. *Surveys in Geophysics*, *40*(3), 303–331. https://doi.org/10.1007/s10712-019-09517-z

Rasti, B., Scheunders, P., Ghamisi, P., Licciardi, G., & Chanussot, J. (2018). Noise Reduction in Hyperspectral Imagery: Overview and Application. *Remote Sensing*, *10*(3), 482. https://doi.org/10.3390/rs10030482

Richards, J. A. (2022). *Remote Sensing Digital Image Analysis*. Springer Nature.

Richards, J. A., & Jia, X. (2006). *Remote Sensing Digital Image Analysis: An Introduction*. Springer Science & Business Media.

Richards, J. A., & Jia, X. (2013). *Remote Sensing Digital Image Analysis: An Introduction*. Springer Science & Business Media.

Rodriguez-Galiano, V., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P. M., & Jeganathan, C. (2012). Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment*, *121*, 93–107. https://doi.org/10.1016/j.rse.2011.12.003

Rodriguez-Galiano, V., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sánchez, J. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *Isprs Journal of Photogrammetry and Remote Sensing*, *67*, 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Rosentreter, J., Hagensieker, R., Okujeni, A., Roscher, R., Wagner, P. E., & Waske, B. (2017). Subpixel Mapping of Urban Areas Using EnMAP Data and Multioutput Support Vector Regression. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(5), 1938–1948. https://doi.org/10.1109/jstars.2017.2652726

Roy, S. K., Mondal, R. K., Paoletti, M. E., Haut, J. M., & Plaza, A. (2021). Morphological Convolutional Neural Networks for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 8689–8702. https://doi.org/10.1109/jstars.2021.3088228

Rwanga, S., & Ndambuki, J. M. (2017). Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS. *International Journal of Geosciences*, *08*(04), 611–622. https://doi.org/10.4236/ijg.2017.84033

Sabat-Tomala, A., Raczko, E., & Zagajewski, B. (2020). Comparison of Support Vector Machine and Random Forest Algorithms for Invasive and Expansive Species Classification Using Airborne Hyperspectral Data. *Remote Sensing*, *12*(3), 516. https://doi.org/10.3390/rs12030516

Sahithi, V. S. S., Krishna, I. V. M., & Giridhar, M. V. S. S. (2022). Analysing the Sensitivity of SVM Kernels on Hyperspectral Imagery for Land Use Land Cover Classification. Journal of Image Processing and Artificial Intelligence, 8(2), 15–23. https://doi.org/10.46610/joipai.2022.v08i02.003

Sanchez, L., Ermolenkov, A., Biswas, S., Septiningsih, E. M., & Kurouski, D. (2020). Raman Spectroscopy Enables Non-invasive and Confirmatory Diagnostics of Salinity Stresses, Nitrogen, Phosphorus, and Potassium Deficiencies in Rice. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.573321

Schaepman, M. E., Jehle, M., Hueni, A., D'Odorico, P., Damm, A., Weyermann, J., Schneider, F., Laurent, V., Popp, C., Seidel, F. C., Lenhard, K., Gege, P., Küchler, C., Brazile, J., Kohler, P. J., De Vos, L., Meuleman, K., Meynart, R., Schläpfer, D., . . . Itten, K. I. (2015). Advanced radiometry measurements and Earth science applications with the Airborne Prism Experiment (APEX). *Remote Sensing of Environment*, *158*, 207–219. https://doi.org/10.1016/j.rse.2014.11.014

Schmidt, J., Fassnacht, F. E., Lausch, A., & Schmidtlein, S. (2017). Assessing the functional signature of heathland landscapes via hyperspectral remote sensing. Ecological indicators, 73, 505-512. 10.1016/j.ecolind.2016.10.017

Schölkopf, B., Burges, C. J. C., & Smola, A. J. (1999). Advances in kernel methods: support vector learning. In *International Conference on Neural Information Processing*. https://doi.org/10.5555/299094

Segl, K., Guanter, L., Rogass, C., Kuester, T., Roessner, S., Kaufmann, H. H., Sang, B., Mogulsky, V., & Hofer, S. O. (2012). EeteS—The EnMAP End-to-End Simulation Tool. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(2), 522–530. https://doi.org/10.1109/jstars.2012.2188994

Sethy, P. K., Pandey, C., Sahu, Y., & Behera, S. K. (2022). Hyperspectral imagery applications for precision agriculture - a systemic survey. *Multimedia Tools and Applications*, *81*(2), 3005–3038. https://doi.org/10.1007/s11042-021-11729-8

Shaik, R. U., Periasamy, S., & Zeng, W. (2023). Potential Assessment of PRISMA Hyperspectral Imagery for Remote Sensing Applications. *Remote Sensing*, *15*(5), 1378. https://doi.org/10.3390/rs15051378

Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghamisi, P., & Homayouni, S. (2020). Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 6308–6325. https://doi.org/10.1109/jstars.2020.3026724

Singh, P., Pandey, P. C., Petropoulos, G. P., Pavlides, A., Srivastava, P. K., Koutsias, N., Deng, K. a. K., & Bao, Y. (2020). Hyperspectral remote sensing in precision agriculture: present status, challenges, and future trends. In *Elsevier eBooks* (pp. 121–146). https://doi.org/10.1016/b978-0-08-102894-0.00009-7

Sousa JJ, Toscano P, Matese A, Di Gennaro SF, Berton A, Gatti M, Poni S, Pádua L, Hruška J, Morais R, Peres E. UAV-Based Hyperspectral Monitoring Using Push-Broom and Snapshot Sensors: A Multisite Assessment for Precision Viticulture Applications. Sensors (Basel). 2022 Aug 31;22(17):6574. doi: 10.3390/s22176574. PMID: 36081033; PMCID: PMC9460142.

Srivastava, P. K., Gupta, M., Singh, U. K., Prasad, R., Pandey, P. C., Raghubanshi, A. S., & Petropoulos, G. P. (2021). Sensitivity analysis of artificial neural network for chlorophyll prediction using hyperspectral data. *Environment, Development and Sustainability*, *23*(4), 5504–5519. https://doi.org/10.1007/s10668-020-00827-6

Stehman, S. V. (2014). Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *International Journal of Remote Sensing*, *35*(13), 4923–4939. https://doi.org/10.1080/01431161.2014.930207

Storch, T., Bachmann, M. F., Eberle, S., Habermeyer, M., Makasy, C., De Miguel, A., Mühle, H., & Müller, R. (2013). EnMAP Ground Segment Design: An Overview and its Hyperspectral Image Processing Chain. In *Lecture notes in geoinformation and cartography*. Springer Nature. https://doi.org/10.1007/978-3-642-32714-8_4

Sun, L., Khan, S. D., & Shabestari, P. (2019). Integrated Hyperspectral and Geochemical Study of Sediment-Hosted Disseminated Gold at the Goldstrike District, Utah. *Remote Sensing*, *11*(17), 1987. https://doi.org/10.3390/rs11171987

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y.-A., & Rahman, A. (2020). Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. Remote Sensing, 12(7), 1135. https://doi.org/10.3390/rs12071135

Thenkabail, P. S. (Ed.). (2019). Hyperspectral remote sensing of vegetation. CRC Press.

Thenkabail, P. S., Lyon, J. G., & Huete, A. (2016). Hyperspectral Remote Sensing of Vegetation. In *CRC Press eBooks*. https://doi.org/10.1201/b11222

Thenkabail, P. S., Lyon, J. G., & Huete, A. (Eds.). (2018). *Hyperspectral Indices and Image Classifications for Agriculture and Vegetation*. CRC press.

Thoreau, R., Achard, V., Risser, L., Berthelot, B., & Briottet, X. (2022). Active Learning for Hyperspectral Image Classification: A comparative review. *IEEE Geoscience and Remote Sensing Magazine*, *10*(3), 256–278. https://doi.org/10.1109/mgrs.2022.3169947

Usery, E. L. (2015). Remote Sensing and Image Interpretation. 7th Edition. *Photogrammetric Engineering and Remote Sensing*, *81*(8), 615–616. https://doi.org/10.14358/pers.81.8.615

Van Der Linden, S., Rabe, A., Held, M., Jakimow, B., Leitão, P. J., Okujeni, A., Schwieder, M., Suess, S., & Hostert, P. (2015). The EnMAP-Box—A Toolbox and Application Programming Interface for EnMAP Data Processing. *Remote Sensing*, *7*(9), 11249–11266. https://doi.org/10.3390/rs70911249

Van Der Linden, S., Rabe, A., Jakimow, B., Thiel, F., Cooper, S., Okujeni, A., and Hostert, P., "Integrating Imaging Spectroscopy and GIS – Free and Open Source Image Analysis in QGIS with the EnMAP-Box 3," in *OSA Optical Sensors and Sensing Congress 2021 (AIS, FTS, HISE, SENSORS, ES),* S. Buckley, F. Vanier, S. Shi, K. Walker, I. Coddington, S. Paine, K. Lok Chan, W. Moses, S. Qian, P. Pellegrino, F. Vollmer, G. , J. Jágerská, R. Menzies, L. Emmenegger, and J. Westberg, eds., OSA Technical Digest (Optica Publishing Group, 2021), paper HF4E.2. https://opg.optica.org/abstract.cfm?URI=HISE-2021-HF4E.2

Vangi, E., D'Amico, G., Francini, S., Giannetti, F., Lasserre, B., Marchetti, M., & Chirici, G. (2021). The New Hyperspectral Satellite PRISMA: Imagery for Forest Types Discrimination. *Sensors*, *21*(4), 1182. https://doi.org/10.3390/s21041182

Vapnik, V. (1995). The Nature of Statistical Learning Theory. In *Springer eBooks*. https://doi.org/10.1007/978-1-4757-2440-0

Verrelst, J., Malenovský, Z., Van Der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J., Lewis, P., North, P., & Moreno, J. M. (2018). Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surveys in Geophysics*, *40*(3), 589–629. https://doi.org/10.1007/s10712-018-9478-y

Verstraete, M. M., Pinty, B., & Curran, P. J. (1999). MERIS potential for land applications. *International Journal of Remote Sensing*, *20*(9), 1747–1756. https://doi.org/10.1080/014311699212452

Wang, J., Bretz, M., Dewan, M. a. A., & Delavar, M. A. (2022). Machine learning in modelling land-use and land cover-change (LULCC): Current status, challenges and prospects. *Science of the Total Environment*, *822*, 153559. https://doi.org/10.1016/j.scitotenv.2022.153559

Wang, S., Zhou, K., Wang, J., & Zhao, J. (2022). Identifying and Mapping Alteration Minerals Using HySpex Airborne Hyperspectral Data and Random Forest Algorithm. *Frontiers in Earth Science*, *10*. https://doi.org/10.3389/feart.2022.871529

Wang, Y., Liang, B., Ding, M., & Li, J. (2020). Dual-branch dense residual network for hyperspectral imagery classification. International Journal of Remote Sensing, 41(7), 2581-2602. https://doi.org/10.1080/01431161.2019.1693078

Winkler, K., Fuchs, R., Rounsevell, M., & Herold, M. (2021). Global land use changes are four times greater than previously estimated. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22702-2

Wu, H., & Prasad, S. (2017). Convolutional Recurrent Neural Networks forHyperspectral Data Classification. *Remote Sensing*, *9*(3), 298. https://doi.org/10.3390/rs9030298

Wu, N., Crusiol, L. G. T., Liu, G., Wuyun, D., & Han, G. (2023). Comparing Machine Learning Algorithms for Pixel/Object-Based Classifications of Semi-Arid Grassland in Northern China Using Multisource Medium Resolution Imageries. *Remote Sensing*, *15*(3), 750. https://doi.org/10.3390/rs15030750

Xie, F., Gao, Q., Jin, C., & Zhao, F. (2021). Hyperspectral Image Classification Based on Superpixel Pooling Convolutional Neural Network with Transfer Learning. *Remote Sensing*, *13*(5), 930. https://doi.org/10.3390/rs13050930

Yan, Y., Deng, L., Liu, X., & Zhu, L. (2019). Application of UAV-Based Multi-Angle Hyperspectral Remote Sensing in Fine Vegetation Classification. *Remote Sensing*, *11*(23), 2753. https://doi.org/10.3390/rs11232753

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061

Yokoya N, Chan JC-W, Segl K. Potential of Resolution-Enhanced Hyperspectral Data for Mineral Mapping Using Simulated EnMAP and Sentinel-2 Images. Remote Sensing. 2016; 8(3):172. https://doi.org/10.3390/rs8030172

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221

Zhang, J., & Chen, Y. (2020). Overview of hyperspectral image classification methods. Nanjing Xinxi Gongcheng Daxue Xuebao, 12(1), 89-100. https://doi.org/10.1155/2020/4817234

Zhao, Q., Yu, L., Du, Z., Peng, D., Hao, P., Zhang, Y., & Gong, P. (2022). An Overview of the Applications of Earth Observation Satellite Data: Impacts and Future Trends. *Remote Sensing*, *14*(8), 1863. https://doi.org/10.3390/rs14081863

Zhong, P., Gong, Z., Li, S., & Schönlieb, C. (2017). Learning to Diversify Deep Belief Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(6), 3516–3530. https://doi.org/10.1109/tgrs.2017.2675902

Zhu, L., Suomalainen, J., Liu, J., Hyyppä, J., Kaartinen, H., & Haggrén, H. (2018). A Review: Remote Sensing Sensors. In *InTech eBooks*. https://doi.org/10.5772/intechopen.71049