Masters Program in Geospatial Technologies

Spatiotemporal Analysis of Forest Cover Changes and Climate Dynamics in Castellón Province

Fahima Begum

Dissertation submitted in partial fulfilment of the requirements for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS Information Management School
Universitat Jaume I
Westfälische Wilhelms-Universität Münster

Spatiotemporal Analysis of Forest Cover Changes and Climate Dynamics in Castellón Province

Dissertation supervised by

Prof. Filiberto Pla Bañón, PhD
Institute of New Imaging Technologies, Universitat Jaume I
Castellón de la Plana, Spain

Dissertation co-supervised by

Ana Cristina Costa, PhD
Associate Professor, NOVA IMS
Universidade NOVA de Lisboa, Portugal

Prof. José Salvador Sánchez Garreta, PhD
Computer Languages and Systems
Universitat Jaume I, Spain

February 20, 2025

Acknowledgement

At first, I express my profound gratitude to the Almighty Allah, the most gracious and most benevolent.

I would like to express my sincere gratitude to my supervisor Prof. Filiberto Pla Bañón (Universitat Jaume I, Spain), for his inspiring advice, constructive guidance and continuous follow up in every working stage. Without his proper guidance, insightful suggestions, constructive criticism, continuous support and kind supervision this study could not have been successfully conducted. I am grateful to him for his kind and friendly attitude that motivated me to complete this research. I am thankful to my co-supervisors Prof. José Salvador Sánchez Garreta ((Universitat Jaume I, Spain) and Ana Cristina Costa (Associate Professor, NOVA IMS, Portugal) for providing the insightful comments. These comments improve the quality of my thesis significantly. Without their consecutive comments and detailed suggestions, the thesis could not be in structured shape and standard quality. I want to thank all my lecturers in Geospatial Technologies for providing me with multidisciplinary knowledge and skills that have significantly facilitated my professional development in this remote sensing and GIS field.

I want to sincerely thank the European Commission for providing me the Erasmus Mundus Scholarship which has been an incredible educational and memorable experience for me. I want to thank everyone related to the program for their continuous support for making our paths easier while studying in this master's program.

I would like to express my heartfelt gratitude to my classmates for their help and support throughout the whole master's program. I would like to express indebtedness to Farhan Ahmed for his invaluable support throughout the whole journey. I am deeply grateful to Sourav Dey, Pramila Paul, Ramiz Muktadir, Sourav Karmakar, Maliha Momtaz, Ashfak Mahmud, Masuma Chowdhury Mishma, Srijon dutta, Kanak Biswas, Tabassum Saba, Fahim Tasnim, Ayesha Tariq for their continuous mental support. I convey my regards to all the persons who supported and helped me during this thesis work.

I convey my gratitude to my mother Mrs. Monoara Begum for her continuous support and inspiration that she always provided. I am grateful to my father, Late Hussain Ahmed, who remains the constant inspiration of my life. Without the support of my parents, I would not be here today. I express my heartfelt gratitude to my sister Farhana Begum for her inspiration that kept me motivated during the entire period of this thesis. I would like to express my sincere gratitude to my beloved husband Mohammad Mobasher Hossain for his continuous mental support, motivation throughout this journey. With his continuous support, the journey became easy for me.

Finally, I want to thank the open-source data community, whose resources have greatly contributed to the success of this research.

Abstract

Understanding forest cover change dynamics and its association with climatic conditions is crucial for sustainable environmental management. This research examined forest cover change in Castellón Province by using satellite-based remote sensing and machine learning classification during 2010-2024. The supervised classification methods—Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Tree (CART)—were applied on Landsat images using Google Earth Engine environment. In this study, spatiotemporal trends integrate Google Earth Engine (GEE), Python-based statistical analysis, and GIS mapping to identify the best classifier as well as reliable model. RF outperformed other classifiers. It proved maximum accuracy using different performance and evaluation metrics. Classification results indicated a significant forest cover loss, especially between 2015 and 2020 due to urbanization, land use change, and potential climatic disturbances. The dynamics of forest conversion have been explored using transition matrices. It helped in detecting overall loss of areas of forest along with appreciable changes to non-forest categories.

The areas of significant deforestation along with afforestation trends were also recognized using spatial maps. This puts focus on finding out the factors behind the degradation like climate change, with anthropogenic effects on land use changes. The time series analysis of NDVI trends and Climate variables (Temperature-2m, Land Surface Temperature (LST) and Precipitation) shows alarming trends in the study area. The climatic variables showed possible strong correlation between vegetation health and climate fluctuations. Pearson correlation analysis confirmed the positive relationship between NDVI and precipitation while demonstrating the negative relationship between temperature and land surface temperatures (LST). The integration of the use of machine learning, remote sensing, along with statistical analysis identifies key insight into data-driven reasoning and detailed analysis of the forest cover changes within a Mediterranean ecosystem. The findings can contribute to sustainable forest management approaches and climate adaptation strategies by providing valuable insights into forest loss patterns including climatic drivers. It can also contribute to regional-level programs about forest conservation, climate adaptive measures, along with sustainable land-use planning and land management using scientific evidence and data-driven knowledge of the province.

Keywords:
Forest Cover change, Remote Sensing, Machine Learning, RF, SVM, CART, Transition Matrices, Climate variables, NDVI, Castellón Province, NDVI-Climatic interaction, Sustainable mediterranean ecosystem.

Contents

List of Figures

List of Tables

List of Abbreviation

AIRS- Atmospheric Infrared Sounder

CART- Classification and Regression Trees

CNN- Convolutional Neural Network

FAO- Food and Agriculture Organization

FRA- Forest Resource Assessment

FP- False Positive

FN- False Negative

GAUL- Global Administrative Unit Layer

GIS- Geographic Information System

GEE- Google Earth Engine

GFW- Global Forest Watch

LST- Land Surface Temperatures

LULC- Land Use Land Cover

ML- Machine Learning

MODIS- Moderate Resolution Imaging Spectroradiometer

MLC- Maximum Likelihood Classification

NNC- Neural Network Classifiers

NDVI- Normalized Difference Vegetation Index

NIR- Near-Infrared Spectroscopy

OLI- Operational Land Imager

OA- Overall Accuracy

PCC- Pearson Correlation Coefficient

RF- Random Forest

SLC- Scan Line Corrector

SVM- Support Vector Machine

SWIR- Short-Wave Infrared

SQ- Square Kilometre

TM-Thematic Mapper

TP- True Positive

TN- True Negative

WGS 84- World Geodetic System84

Chapter 1. Introduction

1.1 Background and Motivation

Forests are essential for maintaining biodiversity, regulating the climate, sinking carbon, maintaining ecological balance, and providing essential ecosystem services (Psistaki et al., 2024). Numerous species find habitats in forests, which helps to preserve biodiversity. Also, forest conservation is crucial for carbon sequestration, which indirectly sustains biodiversity in coastal habitats (Raman et al., 2024). Forests provide diverse ecosystem services that are essential for human well-being. Forests and natural areas play a major role in sustainable development, even in urban areas. It has been observed by Raman et al., (2024) that Sustainable Development Goal 11 which refers to sustainable cities and communities are supported by urban forests by cooling cities, enhancing air quality, and creating recreational areas.

A shift in forest regions over time due to both natural and man-made forces is called "forest cover change". It is a complex phenomenon that affects equally natural and human-induced processes, leading to significant changes in forest ecosystems over time. Different studies also proved the same. Xiao et al., (2022) found that both natural and anthropogenic phenomena such as extreme climatic events, urbanization, and agricultural activities contributed to shifts in the forest cover. Researchers emphasise how important it is for conservation efforts to understand changes in forest cover. So, for the betterment of life it is crucial to identify forest cover changes for taking proper mitigative actions towards forest degradation and all. Changes in forest cover have significant impact on wetlands, biodiversity, carbon cycles, and global climate regulation. Therefore, knowledge of forest dynamics is essential for sustainable environmental management and policy making.

According to Global Forest Watch (GFW), globally 488 million hectares of tree cover decreased between 2001 and 2023, which is 12% compared to what has been there since 2000, which is alarming. Furthermore, Food and Agriculture Organization (FAO) Global Forest Resources Assessment highlighted that deforestation has caused the loss of 420 million hectares of forest globally since 1990 (FAO, 2022). The forest dynamics are different in each region. Mediterranean forests such as those in the province of Castellón in Spain face unique challenges due to their susceptibility to shifting weather patterns, frequent droughts, wildfires, and land use changes. Martínez-Vega et al., (2021) investigated LULC differences in Mediterranean agricultural ecosystems with a focus on southern Spain. These alterations were found to be influenced by human activity, water scarcity, and climate change.

The FAO emphasized the cost-effectiveness of remote sensing in minimizing the need for considerable field work in their article on national forest monitoring systems. It claims that remote sensing enhances fieldwork performance and is especially helpful in determining specific characteristics' boundaries (Holmgren & Marklund, 2007). Farmonaut, (2024) report states that satellite imagery can detect deforestation with up to 95% accuracy. They also

mentioned the ability of modern satellites to capture images with resolutions as fine as 30 cm per pixel allows them to conduct comprehensive analysis of the health of individual trees and the structure of the forest canopy.

Sentinel, MODIS, and Landsat satellites data can provide detailed observations regarding forest dynamics, degradation, and long-term change. Over years in cloud geospatial platforms and improvement in machine learning (ML), Google Earth Engine (GEE) and such tools have advanced a lot in improving mapping efficiency, transition classification, and analysis in forests. High-dimensional, complex datasets in remote sensing have achieved advancement in application of a variety of machine classifiers including Random Forest (RF), Support Vector Machine (SVM), and Classification and Regression Trees (CART) in different land cover classification.

The province of Castellón holds a remarkably big forest area (68%), which is much larger than that of Spain (55%), and the Valencian Region (57%) (Delgado-Artés et al., 2022a)**.** Castellón province is also a good representation of the Mediterranean ecosystem which is an ecologically diverse yet fragile landscape with mixed forest ecosystems, agricultural lands, urban regions and protected natural areas which makes it ideal for the study. Notable changes in forest cover, including deforestation, land abandonment, and afforestation attempts, have occurred in the region as a result of urbanization, climate variability, and anthropogenic issues collectively. Additional factors that have led to forest degradation and vegetation stress include wildfires, extended droughts, and rising trend of land surface temperatures (LST).

Analysing spatiotemporal cover trends and its relationship with key meteorologic factors including temperature, precipitation, and LST become critical in consideration of these environment-related issues. There have been limited studies particularly focused on Castellón that combine remote sensing techniques with machine learning classification approaches. Most past studies have analysed land use change in Spain and Valencia at a broader scale. However detailed studies at local levels in Castellón province remained scarce.

1.2 Research Questions and Objectives

Considering the research gaps previously discussed, the main research questions of this study are as follows:

- RQ1: What are the spatio-temporal patterns and trends of forest cover change in Castellón province, Spain, over the study period (2010-2024)?

- RQ2: Are the changes in forest cover related to the climatic variables (Precipitation, Tropospheric Temperature and Land Surface Temperature) over different time periods in Castellón province, Spain?

Based on the research questions, the primary objective of this study is to analyse the spatio-temporal trends in forest cover changes using satellite-based remote sensing

techniques and to examine the relationship between forest cover change and climate variables in Castellón province, Spain. Other specific objectives as follows:

1. To compare machine learning based supervised classification techniques for forest cover change analysis and find out the most suitable method for accurate forest cover mapping.

2. To assess Forest Cover Change Analysis during a 14-year period (2010, 2015, 2020, and 2024) using multi-temporal Landsat imagery.

3. To identify general and recent trends of Normalized Difference Vegetation Index (NDVI) and the climatic variables like: precipitation, temperature-2m, and LST.

4. To evaluate the NDVI variation over the study period and investigate its correlation with the climatic variables like- precipitation, tropospheric temperature, and LST.

5. To analyse the overall correlation between forest cover changes and climatic variables, integrating the classified forest and non-forest cover changes with the three climatic variables.

1.3 Overall Approach to Achieve Each Objective

In this study, to systematically address the research objectives- a combination of satellite-based remote sensing, machine learning supervised classification methods and statistical analysis was used. Landsat data (from 2010-2024) was used for forest cover classification. Using supervised classification (RF, SVM, and CART) in GEE provided accurate mapping of forest cover changes. Following that, 10-fold-cross-validation and confusion matrices were used to verify accuracy. Based on the selection of best classifiers the mapping and analysis was done using ArcGIS pro.

Then, MODIS-derived NDVI and climate data from (CHIRPs precipitation, ERA5 temperature and MODIS LST) were utilized to investigate the correlation between NDVI (microlevel) and Forest (macro level) with climatic variables. Pearson correlation analysis was used to quantify seasonal (monthly) and annual patterns as well as the correlation between NDVI and climate parameters. Finally, climatic variability and forest cover change were combined using GIS-based spatial analysis. ArcGIS Pro, Python and Excel were used for the integration. By using these overall approaches, a complete comprehension view of forest cover changes and climate interactions in Castellón province was obtained.

1.4 Expected Contributions of the Study

The study contributes to scientific knowledge and practical applications in the field of forest monitoring and climate impact assessment in different aspects like:

- Provides a long-term perspective of spatio-temporal analysis of forest cover changes in Castellón province.
- Evaluates the effectiveness of machine learning classification techniques (RF, SVM, CART) for forest cover change analysis.
- It gives us an insightful understanding of Mediterranean forest dynamics.
- It evaluates the NDVI's response to precipitation, temperature and LST, which offers insights into climate related vegetation changes.
- Integrates the multi-sourced climate data with NDVI, enhancing the capability to monitor vegetation-climate correlations using remote sensing techniques.

1.5 Potential Real-World Applications

The findings of the study can be applied to environmental management practices, policy making and future research in following aspects:

- **Forest conservation and land management:**
  Provides information about forest cover changes around the province along with the forest cover change maps which will be assisting local governments and environmental agencies in forest conservation efforts. The high-resolution forest cover change maps can assist local and regional authorities of the province in monitoring deforestation, afforestation and degradation trends and provides a data-driven scenario about forest loss and recovery areas. Understanding the NDVI or forest-climate correlations can help in developing sustainable forest management strategies in these mediterranean ecosystems.

- **Climate change mitigation and adaptation strategies:**
  Understanding the correlation between NDVI and climatic variables enables policy makers to design effective adaptation measures for forest ecosystems under changing climatic conditions. The study highlights the role of temperature and other climatic variables like precipitation in influencing vegetation health, supporting decision-making for water resource management in fire prone forest areas or in areas like drought prone mediterranean forests.

- **Advancing remote sensing and ML-based Land monitoring:**
  The comparative analysis of RF, SVM, and CART classifiers improves the accuracy and reliability of ML-driven land cover classification techniques. The study demonstrates the effectiveness of Google Earth Engine (GEE) and machine learning in large-scale forest monitoring, providing a model for future application. The method can be adapted to other Mediterranean or semi-arid regions which are similar to our study area in terms of land cover and climate dynamics.

- **Supports for future studies on climate change impact on forest:**
  The study provides baseline data for long-term forest monitoring, enabling researchers to track the climate change impact over time. The results of the study can be applied in large-scale studies examining forest cover change at regional, national and global levels. The study enables the validation of forest-climate or vegetation-climate interaction models, supporting interdisciplinary environmental science and remote sensing studies.

  So, overall, in this study it integrated the forest cover changes with climatic variables, which is a helpful reference for researchers, conservationists, and decision-makers aiming in climate resilience and sustainable forest management.

1.6 Outline of This Study

This thesis is structured into five chapters and each focusing on a distinct aspect of the study.

Chapter 1 provides the basic context of this study where background and motivation are discussed first, followed by a clear description of research questions and relevant objectives. It outlines the overall approach to achieve the objectives. This chapter also highlights the expected contributions and potential real-world applications of this research.

Chapter 2 presents a comprehensive literature review covering key aspects of remote sensing-based forest cover change detection, machine learning-based classification approaches for the classification of forest and accuracy assessment techniques. Additionally, it discusses the role of climate variables in forest cover dynamics in the study area. Finally, research gaps are identified in this chapter which leads to the justification of this study.

Chapter 3 details the study area and data collection from satellite and climate datasets, along with the data preprocessing steps. The methodology is introduced, including supervised classification approach, algorithms for machine learning techniques, accuracy assessment and correlation analysis and transition matrices computations. Overall, it provides a structured approach to achieving the research objectives.

Chapter 4 presents the results and relevant discussion of the results. It starts with the comparative analysis of the classifier performance and their validation metrics. It then explores spatio-temporal trends in forest cover change including area statistics and transition matrix analysis. Additionally, the relationship between NDVI and climatic factors, and correlation of vegetation coverage and climate factors are discussed, concluding with an integrated assessment of forest cover change to climatic variables.

Chapter 5 summarizes the key findings, contributions, and limitations of the study. Recommendations for future research on remote sensing and machine learning in forest monitoring have been included in this chapter.

Chapter 2. Literature Review

2.1 Satellite-Based Remote Sensing for Detecting Forest Cover Change

Remotely sensed techniques help in mapping processes in forests at long-term seasonal scales. It is a powerful tool for tracking and studying changing cover in forests over time, providing high resolution imagery. It also provides multi-temporal datasets for analysis of contribution of factors such as land use and climatic influence forests. Various satellite platforms, such as Landsat, MODIS, and Sentinel-2 significantly used for this purpose. Studies such as Savastru et al., (2019) highlights the importance of satellite observations for tracking disturbances in forests under factors of anthropogenic and climate change.

Landsat imagery has monitored changes in the forest for decades since they provide reliable data based on its high spectral and spatial resolution. Because of its 30 m resolution and long temporal record Landsat is suitable for monitoring deforestation, damage to forests, and reforestation. A comparative study of Landsat and sentinel datasets for Mediterranean forest monitoring by De Luca et al., (2022) found that while sentinel-2 offers higher spatial resolution, Landsat gave similar details but remains significant for its historical consistency. It is to be noted that mediterranean regions have unique environmental characteristics and differ in the environment. Jiménez-Olivencia et al., (2021) reviewed 53 case studies from 6 countries (including Spain) based on extensive local scaled study to acquire an extensive and systematic evaluation of land use change. These six counties are located on the northern shore of the median and studied to gain a comprehensive understanding of Mediterranean regions of Europe. They showed that landscape changes are truly evident in Euro-Mediterranean mountains. They describe some reasons like deforestation, reforestation, urbanisation etc. Also, the Mediterranean region has a risk of fire events. Landsat's applicability is also noteworthy here. Mallinis et al., (2016) used Landsat imagery for monitoring land degradation using dynamics because of the fire events. Such analyses demonstrate the utility of Landsat in evaluating forest resilience under increasing climatic stressors.

Altogether these studies validated the purpose of satellite remote sensing in understanding forest dynamics, climate interactions, and long-term ecological trends.

2.2 Machine Learning Approaches for Forest Classification and Accuracy Assessment

Machine learning (ML) techniques have revolutionized forest classification by enabling researchers to reliably differentiate between various forest types. Supervised classification techniques such as CART, SVM, and RF have gained widespread use in remote sensing for mapping land cover.

In some cases, RF is more accurate in superior feature selection and error minimization capabilities. Several studies reveal compared to CART, RF is more stable, as CART is sensitive to training data. Jog & Dixit, (2016) showed in their analysis that compared to CART and SVM, RF outperformed by showing robustness, stability, particularly in handling mixed and complex land cover types. Another study done by Li et al., (2014) showed that, compared to SVM, RF is more efficient for large datasets. This study also refers to RF as a significant supervised algorithm as they can give high classification accuracies under some relevant conditions. The study done by Praticò et al., (2021b) explored the machine learning classification algorithms (RF, SVM, CART) for Mediterranean forest habitat using GEE and Sentinel-2 data. This study reports that RF performed best with overall accuracy of 88%. Though the study was focused on using Sentinel-2 data, it aligns with our methodological approach. As their work reports that RF continuously showed greater accuracy in complicated first environments, this reference strongly supports the use of RF as the main classification method in our work.

Another key aspect of ML-based algorithms is accuracy assessment which ensures reliability in land cover maps. For selecting best classifiers, studies like Reddy et al., (2013) highlights the importance of accuracy assessment in classification based on remote sensing for detecting forest change. They also emphasise on how it can be utilized using overall accuracy and kappa coefficient to evaluate model performance or classification results. The integration of ML classifiers with GEE has further stimulated the significance of accuracy assessment in remote sensing application for forest change.

2.3 Climate Variables and Their Influence on Forest/Vegetation Dynamics

Forest ecosystems are highly sensitive to climatic variables. Climatic variables influence vegetation growth, species distribution and ecosystem stability. Precipitation, tropospheric temperature, and LST are often considered important variables in climate analysis due to their interdependent roles in the Earth's climate system. These interdependent variables provide important insights into various aspects of the climate and weather patterns. They can offer a comprehensive view of local climate, surface conditions, atmospheric processes and water cycles etc.

Precipitation is a crucial component of the hydrological cycle which directly affects water availability and soil moisture. Precipitation is inherently linked with both LST and tropospheric 2m temperature. The precipitation changes can affect soil moisture, vegetation cover etc, which can eventually affect LST and near-surface air temperatures (Bindajam et al., 2020).

Temperature-2m represents near-surface air temperature, which is critical for understanding atmospheric conditions. It usually represents the tropospheric temperature measured at 2m above the ground. This variable is critical for understanding the immediate environment experienced by humans and most terrestrial ecosystems. 2m air temperature is comparable to

LST but not identical. In a comparative study conducted in Romania, the 2m air temperature observations were compared with LST derived from Landsat data. They observed significant differences between these two variables, especially in urban areas (Mihăilă et al., 2024). Differences between LST and 2m air temperature can vary with land cover classes and impervious surface density.

LST is a key variable which represents the temperature in the land surface-atmosphere interface. LST represents the thermal response of the land surface. It is widely used in studies of surface evapotranspiration estimation, urban heat island modelling, drought monitoring, and ecological assessments (Tang et al., 2024).

These variables together provide a more detailed view of the climate system. The interaction between these variables is crucial for measuring and comprehending climate change impacts. For instance, temperature increased in Canada at about double the global average rate and the mean annual temperature increased about 1.7°C (likely range 1.1°C-2.3°C) from 1948-2016 (Bush & Lemmen, 2019; Zhang et al., 2019). Satellite LST data obtained from sources like Landsat, MODIS etc. can offer valuable high-resolution information on spatial and temporal variations in temperature, especially in poor terrain like mountains (Gök et al., 2024). For their effectiveness these variables are often used together in remote sensing studies, allowing for large-scale analysis of climate patterns and changes over time. Satellite observations and reanalysis data can be used to estimate LST on a large scale. For example- J. Liu et al., (2020) investigated the long-term trends of LST from 2003 to 2017 on a per-pixel basis using three distinct datasets including- the Atmospheric Infrared Sounder (AIRS), the Moderate Resolution Imaging Spectroradiometer (MODIS), the recently released ERA5-Land reanalysis data. The spatiotemporal aspects of the data were found to agree rather well. So, aggregation of some of these datasets using remote sensing studies can provide us a holistic view of climate dynamics, improving climatic models to better understand the impacts of climate change on various ecosystems.

Also, climate variables have impacts on forest cover change. Yan et al., (2020) proved that changes in temperature and precipitation patterns can affect the composition, growth, and distribution of forests. Another aspect is that these variables can influence extreme events like drought and eventually can lead to the forest diebacks and increased wildfire risks (I. A. Shah et al., 2024). Different ecosystems including the mediterranean climate can be vulnerable to changes in temperature and precipitation because of its warm, rainy winters and hot, dry summers (J. Liu et al., 2020). Wildfires are also common in Spain. So along with the different forest type, the changes in precipitation patterns in Spain can have been linked to shifts in forest composition and increased vulnerability to wildfires as in one particular study researchers mentioned about forest resilience (Peris-Llopis et al., 2024).

Using global reanalysis dataset and remote sensing dataset for climate variable analysis is a widely used practice. One study conducted in semi-arid region of east Africa conforms to our work in leveraging information in MODIS, a reliable source for spatio-temporal analysis of

vegetative cover and its association with climate, in analysis of vegetative-climate relations over long-term (Measho et al., 2019).

The Pearson correlation method is widely applied in studies of climate and vegetation to evaluate the statistical relationship between vegetative indices like NDVI and climatic factors including temperature, precipitation etc. Pearson correlation coefficient (PCC) was calculated in an endeavour to evaluate spatio-temporal relationships between vegetative indices and climatic factors over an 18-year period. During a study conducted in semi-arid region,Eritrea, the significant part of the study region showed a positive association between NDVI and precipitation, and a strong negative association between NDVI and temperature, alternatively, reveals that high temperatures could have a role towards vegetative degradation and stress (Measho et al., 2019). Climatic variables have their influence on vegetation dynamics is proved by many studies. Precipitation is seen to have significant impact over vegetative cover growth, particularly in arid and semi-arid environments. A study revealed in Central Asia, precipitation is a controlling variable in governing productivity in terms of NDVI (Jiang et al., 2017). Temperature and LST have reverse impacts over vegetative cover. Temperature fluctuations in studies over the Tianshan Mountains in China have been a dominant reason for interannual variation in NDVI (Q. Liu et al., 2016).

Most of the studies are examining NDVI with precipitation and temperature. But the study of (Tran et al., 2017) motivates this study to take LST in consideration also. As in their study they emphasised on the impact of LST on land use and Land Cover (LULC) and they confirmed that LST is strongly associated with land cover changes. NDVI sensitivities differ with changing seasons regarding climatic factors. For instance, NDVI showed high sensitivities towards precipitation during monsoon, but temperature played a controlling role during drier months in Nepal (Baniya et al., 2018). Similar in the case of the Tibetan Plateau, vegetative greenness showed heightened sensitivities towards precipitation in wet regions, but temperature in arid regions played a controlling role (Zhong et al., 2010).

Also, there are some exceptions seen in the studies also. For example, one study revealed that sometimes vegetative cover behaves with regards to climatic factors varies with region in many cases (Jiang et al., 2017). So, understanding changes in forest cover and how they relate to climatic factors has been significantly aided by much prior research. But we found some gaps that may justify the necessity of the study.

2.4 Research Gaps and Justification for the Study

In recent decades, forest cover changes globally and in Europe have been at the centre stage of intense research activity due to their importance in reducing climate change, conserving biodiversity, and delivering vital ecosystem services. Forest changes have been analysed by researchers by applying sophisticated algorithms like RF. A LULC analysis conducted over the Valencian community by (Sobrino et al., 2024) but Castellón province specifically lacks

comprehensive, long term studies that incorporate remote sensing methods for forest monitoring.

Delgado-Artés et al., (2022) analyses long-term forest evolution in Castellón province where the research developed an increase in dense forest cover (including shrubland and woodlands) from 17% to 28%. But there was no quantification of only forest cover changes in this province or experimenting with climate vs forest/vegetation interaction using spatio-temporal correlation tools or techniques. This study addresses this gap by quantifying forest transitions through pixel-based classification and spatial analysis. This current work will enable a detailed understanding of forest expansion, degradation, and stability over a 14-year period (2010-2024).

A recent study conducted in 2024 analysed the effects of afforestation on European climates from 1986 to 2015 using high resolution climate model simulations. In that study Breil et al., (2024) showed forest cover in Europe has generally increased and afforested forest cover has effects on regional climate systems both locally and non-locally. This study highlighted the complex interaction between forest cover changes and regional climate, emphasising the need for detailed, long-term analysis to understand these dynamics fully. Studies have examined NDVI vs other climatic relationships in various global regions. For example, Santoro et al., (2024) did it for EU level 2000 natural cites but there are very few assessments that are particularly related to mediterranean regions like Castellón province which combined NDVI's sensitivity to different climatic variables.

Also, previous study done by Jiménez-Olivencia et al., (2021) focused on the human induced deforestation and land use change but not investigated the climate induced forest transitions of mountain regions of mediterranean Europe. Their work did not integrate the long-term climate trends as potential drivers for forest cover change. Given the increasing climatic variability specifically in Mediterranean regions of Spain, it is critical to quantify the role of climatic factors of forest cover change in the province.

So, we find it crucial to do specific and more targeted analysis is required to capture localized vegetation responses because of the province's varied geography and Mediterranean climate. Our study minimises those gaps by finding a model for identification of forest cover change and by performing Pearson correlation analysis between NDVI and climatic parameters (precipitation, temperature, and LST). The overall comprehensive result will be providing useful information for climate adaptation plans, sustainable forest management, and conservation efforts in Castellón province, aiding in data-driven policy making for regional environmental sustainability.

Chapter 3. Data and Methods

3.1 Study Area

Castellón province, located on Spain's east coast. It is situated in eastern Spain, within the Valencian community. It is one of the three provinces that constitute the Valencian Community. It is bordered by the Mediterranean Sea to the east, with coastline known as the Costa del Azahar. Geographically, Castellón is split into three distinct regions: the northwestern mountainous Maestrazgo, the valleys of the Míjares and Palancia rivers. They are separated by the Sierra de Espadán, and the Mediterranean coastal plains.

It is the second most mountainous province in Spain. Also, it has a diverse geography, which includes both coastal and inland mountain ranges. It is characterized by a mediterranean climate with arid and warm summers and mild and wet winters. Castellón province is geographically diverse with a combination of forests and wooded hills, agricultural plots, and urban settlements. Castellón province ecologically plays an important role in supporting a range of types of flora, including forests and shrubby semi-arid regions. The diverse landscapes of Castellón, ranging from the beaches to mountainous interiors, make it a unique and significant area for environmental study, particularly that which examines the interaction between climate and forest cover. Castellón province, with its changing environment and years of use, is an ideal location for studying variation in forests and its association with climatic factors. The following figure 2 shows the geographical location of the study area.

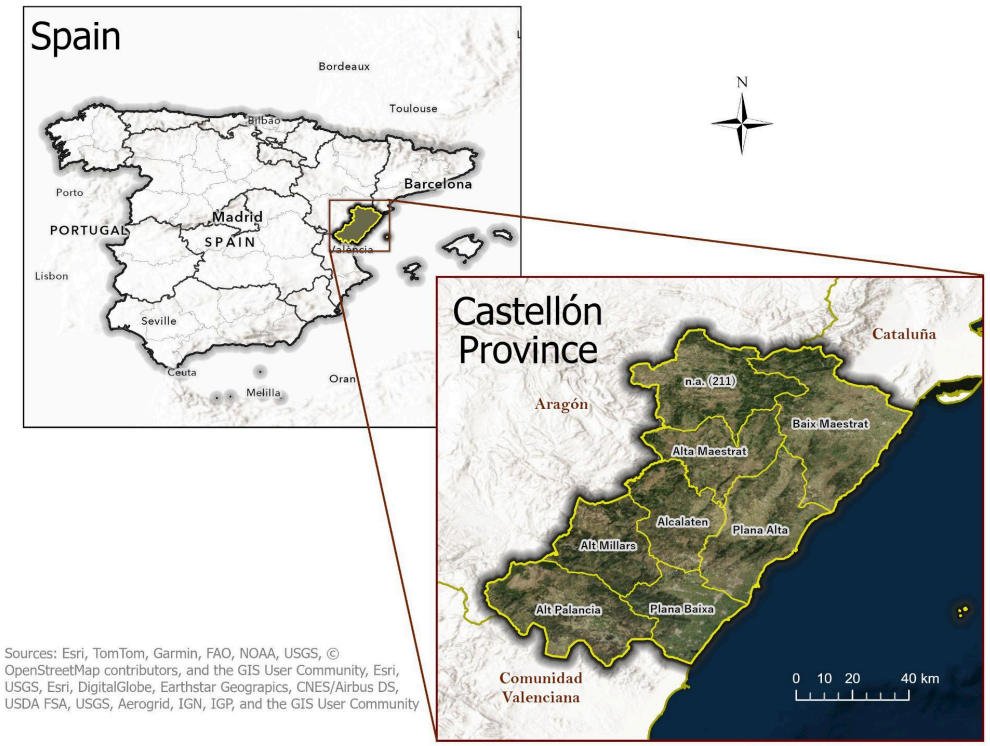


Figure 1: Map showing the geographical location of the study area

The area of Castellón is 6632 sq km. and it is situated between the longitudes 0°50′W and 0°35′ E and the latitudes 40°50′ N and 39°40′ N. For spatial boundary delimitation, The FAO/Gaul level 2 dataset was used to extract the Castellón administrative boundary. As a static dataset, it represents a fixed boundary, maintaining spatial accuracy throughout the study. Also, the reference year for this dataset is 2015, which aligns with the temporal framework of the research. All geospatial processing conducted in the WGS 84 coordinate system (EPSG: 4326) for uniformity in all datasets.

3.2 Data Collection

The study utilized satellite-derived remotely sensed and climatic datasets acquired from various sources for spatiotemporal variation in forest cover changes and its relationship with climate variables. Data collection and preprocessing were conducted by using Google Earth Engine (GEE) and global reanalysis dataset. The spatial and temporal conformity are ensured in every variable. To provide consistent and efficient handling of data acquisition and processing, Google Earth Engine (GEE) and python-based Google Colab were used for processing satellite imagery and climate datasets, respectively.

3.2.1 Satellite Data

Landsat imagery in GEE format was utilized for change mapping and forest classification and detection. Due to problems faced with the Scan Line Corrector (SLC) failure of Landsat 7, Landsat 5 Thematic Mapper (TM) was selected for the analysis of 2010. For the other years of 2015, 2020, and 2024 Landsat 8 Operational Land Imager (OLI) was used. The selection was based on availability of data, spatial resolution, and spectral properties required for accurate forest cover classification. In chapter 2 of literature review scientific ground for using Landsat imagery was observed.

Now, for study period duration this research found some scientific baseline. Choosing study periods in a 5-year interval is a common practice in forest cover change analysis. This approach is supported by scientific literature, as forest does not change rapidly.

In a recent published study, Zang et al., (2025) examined forest regrowth in relation to deforestation regions globally in a 5-year interval. In this study the researchers looked at forest structural factors at 5-year intervals. Also, the Food and Agricultural Organization (FAO) collects and analyses forest resource data through its Global Forest Resources Assessment (FRA) program every 5 to 10 years. This approach makes it possible to thoroughly analyse long-term trends in the shifting forest cover (D'Annunzio et al., 2014). So, these studies demonstrated that using a 5-year interval in forest cover change analysis is a scientifically sound approach.

3.2.2 Preprocessing Step for the Landsat Data in GEE

Remote sensing studies often use images with low cloud cover and combine multiple images using a median composite method to reduce atmospheric distortions and ensure data quality. This practice is supported by scientific literature.

Gao, (2009) discusses the need for atmospheric correction of remote sensing, particularly in water bodies, where accurate corrections are necessary to derive useful water-leaving reflectance. They emphasize the need for appropriate image selection and processing to minimize atmospheric effects. Ajayi & Ojima, (2022) stated in their study that cloud screening is a significant preprocessing step in almost all optical and infrared remote sensing applications. This finding is consistent with the use of imagery with less than 10% cloud cover. It reduces atmospheric distortions from the dataset. In this study, Landsat images were acquired and processed in GEE. The Google Earth Engine is an open-access cloud computing environment that processes and retrieves remotely sensed data globally. It is built on the base of Google's cloud and the JavaScript programming language (Praticò et al., 2021). During the study cloud filtering was applied. Only images with <10% cloud cover were utilized to reduce the atmospheric distortions. Multiple images were aggregated using a median composite approach which is shown below in Table 2. Scaled reflectance values were applied using band-specific coefficients as per USGS collection 2 level 2 guidance. This ensures consistency in spectral band scaling. For optical bands 30m resolution was used which was default. But the thermal bands in Landsat 5 (120m) and Landsat 8 (100m) were resampled to 30m. The resampling was done in GEE to align with the optical bands and maintain accuracy in classifications. All datasets were projected in WGS 84 (EPSG: 4326) to maintain consistency across spatial analysis.

The Landsat data utilized in the study, along with their spatial and spectral properties, are presented in Table 1.

Dataset	Sensor/product	Time frame	Resolution	Resolution used	Bands used	Projection	Unit (in dataset), Unit used	Path/ row
Landsat 5	Thematic mapper (TM)	2010	30 m (optical bands) 120 m (Thermal bands)	30m (Thermal bands resampled in median composite)	SR_B1-(Blue) SR_B2-(Green), SR_B3-Red), SR_B4 (NIR), SR_B5 (SWIR-1), SR_B7-(SWIR-2), ST_B6 (Thermal)	WGS 84 (EPSG: 4326)	Reflectance, Scaled Reflectance	198,199 /32,33
Landsat 8	Operational land imager (OLI)	2015, 2020, 2024	30 m (optical bands), 100m (Thermal bands)	30m (Thermal bands resampled in median composite)	SR_B1- (Ultra blue), SR_B2-(Blue), SR_B3 (Green), SR_B4-(Red), SR_B5-(NIR), SR_B6 (SWIR-1), SR_B7 (SWIR-2), ST_B10 (Thermal)	WGS 84 (EPSG: 4326)	Reflectance, Scaled Reflectance	198,199 /32,33

Table 1: Specification of Landsat remote sensing data used in the study

3.2.3 Temporal Aggregation and Image Composition

The median composite technique is widely used in remote sensing studies to reduce atmospheric distortion and ensure data quality. (Flood, 2013) introduced the medoid compositing method for Landsat imager, which is an approximation of median composite approach. This research explained that this approach is also resistant to extreme values and seems to be better at representing the scenario of the real time period. Qiu et al., (2023) evaluated various Landsat image composting algorithms, including median composite approach for selecting the most appropriate method for different applications. They highlighted the effectiveness of the median composite technique in mitigating cloud impacts and atmospheric distortions.

The details of Landsat image acquisition, including the number of images used per year for median composite generation, are summarized in Table 2.

Dataset and Sensors	Year	Date of image collection (DD/MM/YYYY)	Data aggregation	Total number of images used for median composite
Landsat 5 Thematic mapper (TM)	2010	10/02/2010, 26/02/2010, 14/03/2010, 30/03/2010, 17/05/2010, 05/08/2010, 21/08/2010, 06/09/2010, 24/10/2010, 11/12/2010, 10/02/2010, 30/03/2010, 17/05/2010, 18/06/2010, 20/07/2010, 21/08/2010, 06/09/2010, 24/10/2010, 27/12/2020, 24/05/2010, 11/07/2010, 27/07/2010, 13/09/2010, 16/11/2010	Median composite	24
Landsat 8 Operational land imager (OLI)	2015	24/02/2015, 12/03/2015, 28/03/2015, 13/04/2015, 15/05/2015, 31/05/2015, 03/08/2015, 20/09/2015, 22/10/2015, 08/02/2015, 24/02/2015, 28/03/2015, 15/05/2015, 31/05/2015, 03/08/2015, 20/09/2015, 22/10/2015, 07/11/2015, 14/01/2015, 20/04/2015, 06/05/2015, 07/06/2015, 108/08/2015, 26/08/2015, 14/11/2015, 30/11/2015	Median composite	26
Landsat 8 Operational land imager (OLI)	2020	05/01/2020, 06/02/2020, 22/02/2020, 28/05/2020, 13/06/2020, 31/07/2020, 16/08/2020, 06/12/2020, 22/12/2020, 05/01/2020, 22/02/2020, 28/05/2020, 13/06/2020, 15/07/2020, 31/07/2020, 16/08/2020, 03/10/2020, 19/10/2020, 06/12/2020, 22/12/2020, 12/01/2020, 03/05/2020, 19/05/2020, 06/07/2020, 07/08/2020, 11/11/2020, 13/12/2020	Median composite	27
Landsat 8 Operational land imager (OLI)	2024	01/02/2024, 07/05/2024, 10/07/2024, 26/07/2024, 11/08/2024, 27/08/2024, 28/09/2024, 30/10/2024, 07/05/2024, 26/07/2024, 11/08/2024, 27/08/2024, 28/09/2024, 30/10/2024, 12/04/2024, 30/05/2024, 02/08/2024, 21/10/2024	Median composite	18
Landsat 8 Operational land imager (OLI)	May-July (2024)	05/07/2024, 10/07/2024, 26/07/2024, 07/05/2024, 26/07/2024, 30/05/2024	Median composite	6

Table 2: Overview of Landsat image composition and aggregation

Thus, prior studies also used these composite techniques in different ways related to their purpose. For example: (Zang et al., 2025) examined the track of forest fragmentation under urban expansion using long-term Landsat observations and spatial analysis. In their study they used composite techniques by selecting two optical images. Here in our study, we used a median composite approach for aggregating multiple cloud filtered images. These references collectively support the median composite approach as an effective method. So, in this study median composite was used to reduce noise and minimize the effects of the outliers in each yearly dataset.

3.2.4 Climate Data

In this study three variables are considered for the analysis. These three variables are: precipitation, temperature (2m), and land surface temperature (LST). Precipitation data derived from CHIRPS PENTAD that provides high-resolution precipitation estimates. Temperature 2m data extracted from ERA5 hourly dataset. LST data derived from MODIS MOD11A2 to assess surface thermal variations. These datasets were selected because of their continuous availability between 2010-2024, enabling a long-term climate-vegetation interaction analysis.

The use of these kinds of datasets are validated by different studies. For example, J. Liu et al., (2020) showed that combining remote sensing data (Satellite-based) and reanalysis data (generated by numerical models using historical observation) enables us to more thoroughly describe the spatiotemporal dynamics of LST and gain a better understanding of the main causes of LST variations across various locations. They used datasets like MODIS, ERA5 etc. for their analysis. It can be applicable for the other variables too. So, for or examining NDVI'S sensitivity to climatic variables the relevant data have been extracted from satellite-based datasets (CHIRPS PENTAD and MODIS) and global reanalysis dataset (ERA5 hourly) in this study. These datasets are efficient for their continuous and free availability of data.

3.2.5 Climate Data Preprocessing Steps in Python (Google Colab)

Precipitation, temperature and LST datasets were processed using python in google colab. For NDVI while using the MODIS MOD13A3 dataset, the values were scaled to reflect an NDVI range between -1 to +1. For the NDVI trend, the dataset was processed to derive annual means across Castellon province. CHIRPS precipitation was summed annually to obtain total rainfall per year. ERA5 temperature data was converted from kelvin to Celsius for meaningful interpretation. LST values from MODIS MOD11A2 were similarly converted from kelvin to Celsius using standard MODIS scaling factor. And for temporal aggregation precipitation and temperature datasets were processed to obtain consistent monthly and yearly means to ensure consistency with NDVI and forest classification trends.

These values were later correlated with NDVI trends in Pearson correlation analysis to assess the relationship among them. One thing to note is that different datasets had varying resolutions (shown in Table 3). To maintain consistency, climate data was aggregated over the study area using yearly means with slight exception for precipitation. Following Table 3 provides an overview of the climate datasets, including their resolution, projection, and temporal aggregation.

Dataset	Sensor/ product	Resolution	Projection	Unit (dataset)	Unit used	Data aggregation	Time frame
NDVI	MODIS MOD13A3	1km	WGS 84 (EPSG: 4326)	NDVI index (-1 to +1)	NDVI index	Monthly, Yearly mean	2010-2024
Precipitation	CHIRPS PENTAD	5km	WGS 84 (EPSG: 4326)	mm	mm	Monthly, Yearly mean	2010-2024
Temperature-2m	ERA5 Land Hourly	9km	WGS 84 (EPSG: 4326)	Kelvin	Celsius (°C)	Monthly, Yearly mean	2010-2024
LST	MODIS MOD11A2	1km	WGS 84 (EPSG: 4326)	Kelvin	Celsius (°C)	Monthly, Yearly mean	2010-2024

Table 3: Specification of climate dataset used in the study

3.3 Methods

The methodological workflow of this research follows a systematically integrated remote sensing methods, machine learning based- classification, and interrelation analysis of climate with vegetation. This process involves retrieval of data, data pre-processing, supervised classification, accuracy evaluation, forest cover change detection, and correlation analysis.

The methodology includes a chain of various analyses serving the individual objectives. At first the methodology adopted for the forest cover classification in a structural approach, including satellite image data collection and preprocessing (this is described well in the previous section 3.2), machine learning based-classification, post-classification refinement, and accuracy assessment to achieve reliable results.

The following Figure. 2 illustrates a comprehensive visual representation of all the steps of this study from data collection to final evaluation.

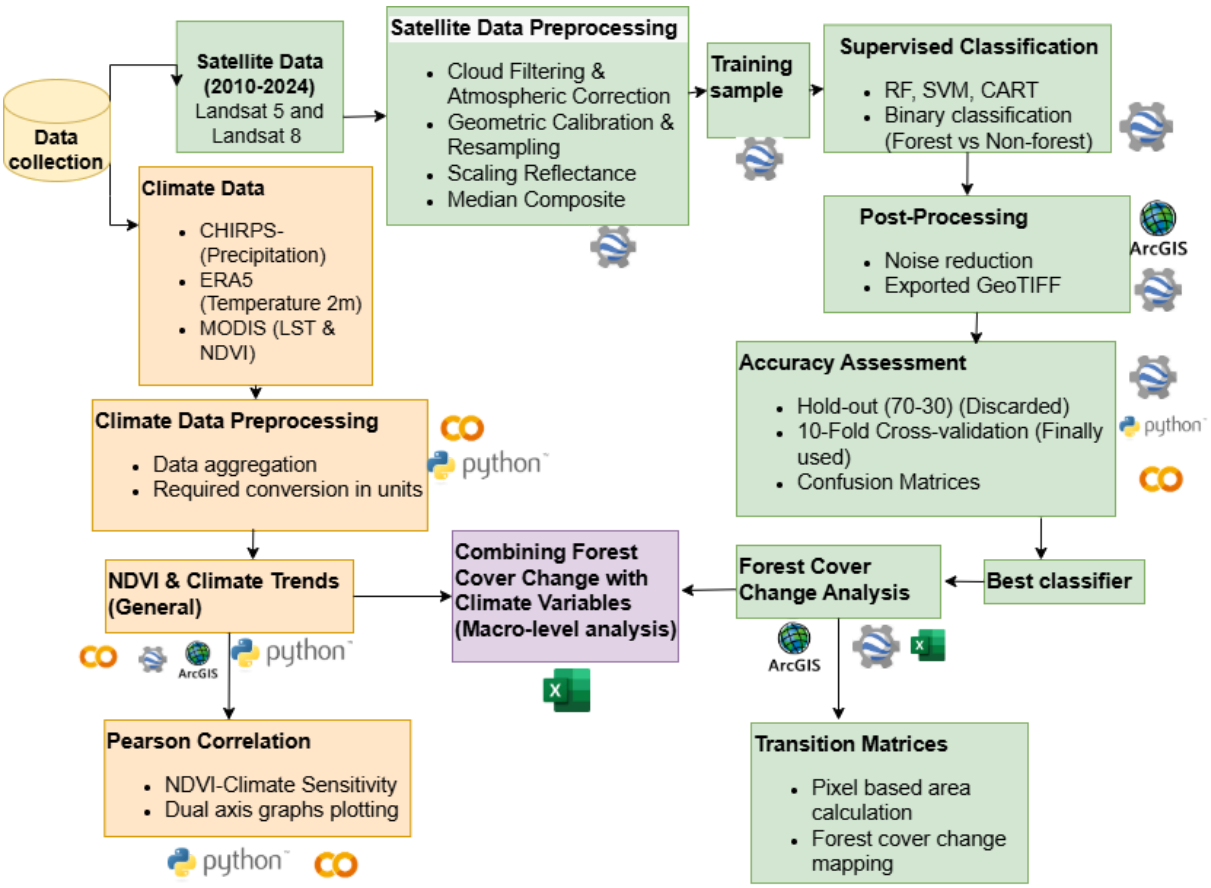


Figure 2: Methodological framework for this study

3.3.1 Supervised Classification Approach

To classify forest and non-forest areas, a machine-learning-based supervised classification approach was applied using Landsat satellite imagery. Supervised classification in land cover classification relies on user-labelled training data to train a model, which then classifies unknown land cover types. In this technique, users give training samples to the model and based on that the algorithm categorises pixels in an image to predefined classes. The users labelled a representative sample of each class from the satellite image, which are used to train the classification algorithm. After processing the image's pixel spectral properties, the algorithm assigns the most similar class to each pixel based on the training data (Díaz-García & Regos, 2024). In a recent study, Khudhur et al., (2024) evaluated the performance of four supervised classification techniques for land cover classification: Support Vector Machine (SVM), Artificial Neural Network (ANN), Maximum Likelihood (ML), and Mahalanobis Distance (MLD). They used Sentinel-2 satellite imagery with 10m spatial resolution and reported very good results by SVM and ANN with overall accuracy rates of 90% and 87%, respectively. So, the technique is commonly applied and has been used in scientific research.

In this study, three machine learning classifiers (RF, SVM, CART) were implemented. And based on the studies objective a binary classification system was defined:

- Forest (1)
- and non-Forest (2)

This binary classification approach was chosen to maintain specificity or clarity in detecting forest cover changes over time. It simplifies change detection and ensures classification accuracy.

3.3.2 Machine Learning Classification Algorithms for Forest Classification (Level 1)

As mentioned before, the study employed three machine learning algorithms. These three algorithms have significant popularity in land cover classification due to their ability to handle complex datasets and produce accurate results.

1. **Random Forest (RF)**

RF is a widely-used machine learning algorithm which creates multiple decision trees and combines their outputs to get more accurate predictions (Breiman, 2001). It uses one of the most recognized methods called bagging, also known as bootstrap aggregation. This method creates diverse training sets for each tree to reduce overfitting and improve generalization (Rani & Inayathulla, 2024). RF is a popular method as it can handle both regression and classification tasks with a high degree of accuracy. RF is also effective in land cover classification with high accuracy.

2. **Classification and Regression Tree (CART)**

CART is a single tree decision classifier like RF (Breiman, 2017). It is a predictive machine learning algorithm which divides the data into subsets at each node of the tree based on an attribute's threshold value. The attribute with higher value is considered for final decision making (Bishop & Nasrabadi, 2006). In land cover classification CART has shown strong performance. For instance, a case study in Northern Mayanmar's gold mining area revealed that CART achieved accuracy of 91.05% with a kappa value of 0.84 and the performance was better than Maximum Likelihood classification but slightly lower than RF (Oo et al., 2022).

3. **Support Vector Machine (SVM)**

SVM is a supervised non-parametric machine learning algorithm based on kernel function and used for classification and regression (Cortes & Vapnik, 1995; Vapnik, 1998). It aims to find a decision boundary called optimal hyperplane to distinguish different classes within high dimensional feature space. SVM has represented high performance in land cover classification in comparison to the traditional methods such as Maximum Likelihood Classification (MLC) and Neural Network Classifiers (NNC). SVM learns the boundary between trainers of different classes to train the algorithm. It helps to project data into a multidimensional space and identify one or more hyperplanes to maximize the training dataset separation among a predefined number of classes (Huang et al., 2002). Effectiveness of SVM is supported by many studies. For example, (Rudrapal & Subhedar, 2015) showed more than 90% accuracy for land cover classification by using SVM.

These machine learning algorithms have demonstrated high potential in optimizing accuracy as well as efficiency of land cover classification. Their ability of high-dimensional data handling and learning complex patterns from limited training samples makes them well-suited for this remote sensing application.

3.3.3 Experimental Setup for Training and Validation

In total 500 samples were selected, 250 for forest and 250 for non-forest based on random sampling across different vegetation types compared with Landsat imagery. This sample collection was done before classification and used to train and validate the classifiers. While sampling, random point sampling was applied to ensure representation across diverse terrain types, different vegetation types and to cover the whole study area. A study by (Millard & Richardson, 2015) demonstrates the use of random point sampling in forest cover classification. They used a stratified random sampling approach to generate training data for their Random Forest classifier.

To ensure accuracy in classification across different years, a separate set of training and validation samples were collected for each year (2010, 2015, 2020, 2024). Landsat imagery for each respective year was examined to identify the major forested and non-forested areas The logic behind this is: land cover conditions significantly vary over time, using a single database for all years can cause mistakes. The sampling process was conducted in GEE. The outcome is manually cross-referenced with Global Forest Watch datasets. The spatial distribution of training samples used for classification, as visualized in Google Earth Engine (GEE) screenshot, is illustrated in Figure 3.

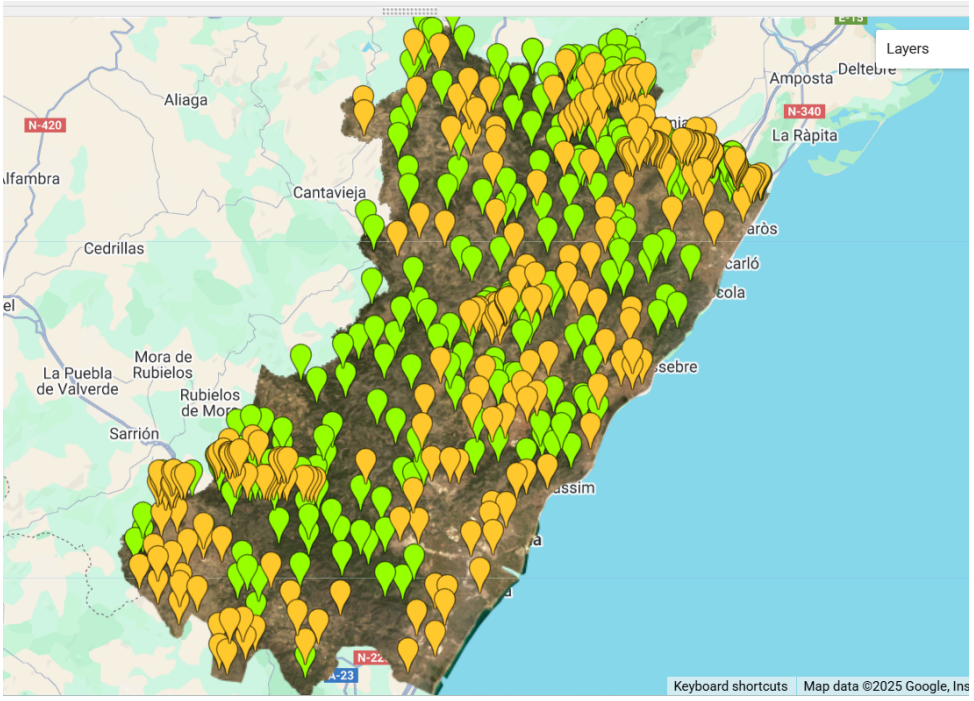


Figure 3: Training sample distribution in GEE (Green colour: Forest, Orange colour: non-Forest)

3.3.4 Post Processing of Classified Image

Before conducting the final image classification, neighbourhood analysis a 3×3 focal mode smoothing filter was applied to reduce noise and enhance spatial consistency in the classified results. These post processing steps correct small misclassifications and ensure that final maps represent continuous land cover patches rather than isolated misclassified pixels. Classified images were exported as GeoTIFFs for further GIS analysis. The visualization and further analysis were done in ArcGIS pro. Combined Mapping done in ArcGIS pro which gave a detailed and comprehensive view of forest cover change in the study area.

3.3.5 Accuracy Assessment & Model Performance

Two validation techniques were used to evaluate classifier performance.

1. **Hold-out validation (70%-30%)**
   Initially the dataset conducted an overall classification accuracy assessment. For this the dataset was divided into 70% training and 30% validation split. Specifically, out of a total 500 samples, 350 samples (70%) were used for training, and 150 samples (30%) were used for validation. These hold out 70-30 at first evaluated. Then accuracy was measured in GEE. And based on confusion metrics derived from GEE, the evaluation matrix (Overall Accuracy, Kappa Coefficient, Precision, Recall, and F1-score) were calculated in python-based Google collab.

2. **10-fold-cross-validation**
   10-fold cross-validation has been performed using Python (Scikit-learn). The dataset was split into 10 subsets and classification was repeated 10 times using different subset for validation each time. The accuracy in terms of Overall Accuracy, Kappa Coefficient, Precision, Recall, and F1-score have been measured in python-based Google collab. The final accuracy was computed as the average of the 10 iterations.

The classifications results were evaluated using five key metrics: Overall Accuracy, Kappa Coefficient, Precision, Recall, and F1-score based on derived confusion metrics. A confusion metric is a table used to quantify the performance of the classification models between predicted and actual values. The matrix is usually good in dealing with binary or multiclass classification problems. In binary classification, usually the matrix contains four entries: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

According to Qiu et al., (2023) confusion metrics prove the overall performance of a model's description in relation to actual class labels. Confusion matrix allows for the calculation of various performance metrics (Sokolova & Lapalme, 2009). Accuracy alone is often insufficient for a comprehensive evaluation of a classification model. When there is a class imbalance, metrics like precision, recall, and F1-score provide more insightful assessments of the model's performance than accuracy alone (Ruslan & Arbaiy, 2024). Precision and recall

are particularly useful and common techniques when there is a need to minimize false positives or false negatives.

The definitions and equations for precision, recall, F-score, accuracy, and Cohen's kappa are given here for a proper understanding:

Accuracy

Accuracy or Overall Accuracy (OA) represents the proportion of correctly classified pixels relative to all reference pixels**.**

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Where,

TP = True Positives (forest pixels which are correctly classified)

TN = True Negatives (non-forest pixels which are correctly classified)

FP = False Positives (non-forest pixels which are misclassified)

FN = False Negatives (forest pixels which are misclassified)

(Maxwell et al., 2021)

Kappa Coefficient (κ)

Kappa measures the agreement between the classification results and reference data while taking into account the agreement occurring by chance.

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (2)$$

Where,

P_o = Observed agreement

P_e = Expected agreement by chance

(Pandya & Gontia, 2025)

Precision (User's Accuracy)

Precision is the ratio of correctly classified pixels out of all predicted pixels for a given class.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

(Maxwell et al., 2021)

Recall (Producer's Accuracy)
Recall is the ratio of correctly classified pixels out of all actual pixels in that class.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

(Maxwell et al., 2021)

F1-score
The F-score is the weighted harmonic mean of precision and recall, providing a single measure that balances both metrics.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

(Maxwell et al., 2021)

3.3.6 Qualitative Evaluation in ArcGIS Pro

To validate and interpret the results beyond all numerical area comparison among the classifiers and l accuracy assessments, all four years (2010, 2015, 2020, and 2024) classified maps were visually inspected in ArcGIS Pro. In this section the composing four classified maps were used in side-by-side visual interpretation. This manual visual inspection examined the classification consistencies, and observed forests aligned with the expected patterns of forest cover dynamics in this province.

3.3.7 Summarization of Level 1

Above mentioned settings till these sections till 3.3.6 completed the level 1 of the methods. In this section following the classification, accuracy assessment, a final qualitative and quantitative evaluation was conducted to solidify the finding related to the best classifiers and identifying the forest cover change. Upon analysing accuracy both quantitatively by measuring it (10-fold classification averaged results OA, Kappa, Precision, Recall, F1-score) as well as by visualization, the most appropriate classifier was identified.

3.3.8 Seasonal Scale Validation of Best Classifier Performance (Level 2)

To further validate the consistency of the best selected classifier, level 2 analysis also comprised applying the same classification workflow across seasonal time periods (Summer Window: May–July). This level 2 analysis ensured temporal as well as methodological consistency of forest cover change identification, strengthening the study result as an aggregate. Each step conducted in level 1 repeated here executed previously in the same manner with this shorter timescale. This is done to ensure classifiers stability in a shorter time, compare annual and seasonal forest cover trends etc.

3.3.9 Forest Cover Change Analysis and Transition Matrix

Area estimation was done using the Random Forest classifier which was found out as the most suitable classifier for the study using Excel. This is done for the quantification of the forest cover trend. This process comprises cover estimation, land cover change analysis, time series analysis and change visualization utilizing transition matrices in GEE and ArcGIS pro. From the four classified images, the forest area change analysis was estimated to ensure both absolute and relative forest cover change. The area estimation followed pixel-based calculation in transition matrix, where each classified pixel was converted into square km using Excel and GEE. The base equation for area calculation the study followed:

$$\text{Forest Area (sq km)} = \text{Pixel Count} \times \left(\frac{\text{Spatial Resolution}^2}{1{,}000{,}000} \right) \quad (6)$$

Where each pixel represents $30m \times 30m$ ($900\ m^2$) spatial resolution.

For tracking land cover changes over time, the transition matrix is a widely used approach. Runfola & Pontius, (2013) elaborated the concept of using transition matrices to illustrate changes in land cover in their work on the flow matrix. They clarify how transition matrices might be used to quantify changes in land cover over time, which is aligned with concepts taken into consideration for this study. Transition matrices provide these representations of how land cover changes over time.

During this study, reclassification of forest and non-forest classes using numerical encoding as forest =1, non-forest= 2. Then the classified images were overlaid to derive pixel-wise changes. The total transitional values were aggregated to compute the total area (sq km) undergoing each type of change using GEE and then exported as GeoTIFF to ArcGIS pr where change trends were further analysed. Based on the derived raster data of transitional forest cover change, pixel counts were calculated. Then using the area calculation equation no-6 four transitional area changes were calculated. The results were structured into a transition matrix table, detailing the area (sq km) for each land cover transition. Both results of pixel counts, and four transitional forest cover changes derived from GEE and Excel were compared and matched.

3.3.10 NDVI Analysis & Climate Trends (2010-2014)

NDVI trends (2010–2024):
The Normalized Difference Vegetation Index (NDVI) is a widely used vegetation index derived from remotely sensed data. It quantifies an intense vegetation index that measures the density, health, and greenness of vegetation (Rhew et al., 2011). Its values typically range from -1 to +1. Values closer to 1 indicate dense, healthy vegetation, while lower values (closer to -1) suggest barren land, water bodies, or non-vegetated surface.

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{Red})}{(\mathrm{NIR} + \mathrm{Red})} \tag{7}$$

Where,

NIR = Near Infrared which reflects healthy vegetation.

Red is absorbed by chlorophyll in vegetation
(Boiarskii & Hasegawa, 2019).

In this study, the NDVI band was normalised and converted to real values using a scale factor of 0.0001. The MODIS NDVI (MOD13A3) dataset was used to filter the data from 2010 to 2024. The yearly mean NDVI raster data was assessed using GEE. The mean NDVI values have been computed for all the years for tracking behaviour of vegetation over a continuous time period. And exporting yearly NDVI values as CSV for further analysis was conducted by using Python and ArcGIS pro for trend analysis.

Climate trends analysis:
Month-wise and annual trends in Precipitation, Temperature-2m, and LST have been computed using Graphical visualization via Python Matplotlib. To analyse climate trends, python-based Google collab scripts were used to filter climate datasets from 2010-2024. Then monthly and yearly data aggregated for individual trend analysis and further analysis. For graphical representation of climate trends precipitation, temperature 2m and LST were plotted using python's Matplotlib.

3.3.11 Pearson Correlation Analysis: NDVI and Climatic Variables

To analyse the relationship between vegetation dynamics and climatic parameters, the Pearson correlation coefficient (r) was used to calculate the sensitivity of NDVI to precipitation, air temperature (2m), and land surface temperature (LST) over time. The Pearson correlation coefficient is a widely accepted statistical indicator employed to quantify the linear relationship between two continuous variables (Benesty et al., 2009).

$$r = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}} \quad (8)$$

Where,

X_i and Y_i are individual data points.

$\overline{X}$ and $\overline{Y}$ are the means of respective variables.

r values range from –1 to +1. Here, –1 indicates strong negative correlation, +1 indicates strong positive correlation with 0 indicating no correlation (Zhou et al., 2016).

For testing NDVI's sensitivity towards climatic factors, Pearson correlation (r) and relevant p-values have been computed using python-based Google Colab. To ensure accurate correlation analysis, a structured data pipeline was developed using python. Some Python packages are utilized for importing data. For instance, Google colab file uploader allows for manual file import, Matplotlib is used for visualization, and Pandas is utilized to handle tabular data. The python script processed an excel file containing NDVI and climate data by ensuring to remove missing values, handling date-time inconsistencies. Month wise and yearly NDVI-climate correlation was computed as a final product. At the end, the results were visualised using time-series plots with dual y-axes.

3.3.12 Final Correlation Analysis: Forest Cover Change and Climatic Variables

The final step was macro-scale forest cover trend analysis using climatic variables like total precipitation, 2m air temperature, and land surface temperature (LST). This enabled us to detect the long-term association between climate variability and forest area trends from the period 2010-2024. Data compilation is done based on the previously extracted values, stored properly in excel with clearly structured browns for each year. Column names were standardised, and non-numeric values were removed. The process was conducted using python and excel. Forest area, non-forest area and precipitation were plotted as bar graphs (left y axis) and Temperature-2m and LST were visualized as line graphs (right y axis). The methodological framework employed in this study includes multiple tools and software for image preprocessing, classification, accuracy assessment and climate-NDVI correlation analysis.

The various tools and software utilized for image processing, classification, accuracy assessment and analysis are listed in Table no 4.

Tool	Software
Google Earth Engine (GEE)	Image preprocessing, NDVI analysis
Python, (matplotlib, scikit learn, pandas etc libraries)	Statistical analysis, accuracy assessment; graphical representation, climate data extraction
ArcGIS pro (version 3.4.0)	Mapping and visualization, change detection
Excel (version 2502)	Data compilation, chart representation

Table 4: Tools and software-packages used in this study

Chapter 4. Results and Discussion

This chapter presents a complete review of forest cover changes and their relationship with climatic variables in the Castellón province. It has been systematically structured into some main sections. We considered the first section in two levels for better tracking of the steps or understanding. The first section level 1 addresses comparing performance of different machine learning classifiers (RF, SVM, and CART) based on various accuracy metrics, confusion matrices, and 10-fold-cross-validation techniques for choosing the best-performing classifier. At the end of the first section in level 2 we repeated the processes for the seasonal scale analysis. It helps us to assess how the model is performing in a short time scale period (summer season) and comparing with the level 1 result of the annual longer period.

The second section focuses on addressing forest cover change, including statistics of changed cover, trends with a temporal context, and transition matrices for quantification of spatiotemporal variations. The third section addresses the general trends of NDVI and other climatic variables. Then assesses their relationship between NDVI with climatic parameters such as precipitation, air temperature (2m), and LST (land surface temperature) with spatial correlation. Next for extending this correlation for testing climatic fluctuations' sensitivity of NDVI regarding seasonal and interannual variations Pearson correlation coefficient results are presented.

Finally, the association between forest cover dynamics and climate variables is assessed on a broader scale, integrating classification results with long-term climatic variables trend. The systematic approach ensures a clear understanding of the interdependencies between forest cover changes and climatic variability in Castellón province.

4.1 Comparison of Classifiers Based on Different Years (Level 1)

The whole classification is conducted in 2 segments; Level 1, which is a comparison of classifiers based on different years (2010, 2015, 2020, 2024), and Level 2 is conducted using the best classifiers based on summer season scale (May-July, 2024).

4.1.1 Performance Metrics Comparison

Initially, the three supervised classifiers (RF, SVM, and CART) were assessed over four different years (2010, 2015, 2020, 2024) in Castellon province. The overall classification accuracy was assessed using a hold-out validation approach with 70%-30% (train-test) split. The results of this assessment are summarized in table 5.

The CART model consistently showed 100% accuracy across the years, which looks ideal, but raised concerns about potential overfitting. Similarly, RF displayed high accuracy, particularly from 2015 onward, reaching 100%. In contrast SVM consistently underperformed compared to RF and CART, with lower accuracy scores ranging from 0.84 to

0.87 and a lower Kappa coefficient, indicating weakness in handling the dataset's spectral variability.

The classification performance metrics derived using the hold-out (70%-30%) (training-testing set) split is summarized in Table 5.

Year	Classifier	Overall Accuracy	Kappa	Precision	Recall	F1-score
2010	RF	.99	.98	0.99	1	0.99
2010	SVM	.84	.67	0.78	0.75	0.77
2010	CART	1	1	1	1	1
2015	RF	1	1	1	1	1
2015	SVM	0.86	0.72	.86	.89	.87
2015	CART	1	1	1	1	1
2020	RF	1	1	1	1	1
2020	SVM	.84	.69	.85	.82	.84
2020	CART	1	1	1	1	1
2024	RF	1	1	1	1	1
2024	SVM	.87	.74	.86	.89	.87
2024	CART	1	1	1	1	1

Table 5: Classification performance metrics based on hold-out (70% -30% split) validation

According to some prior studies overfitting can be an issue in accuracy assessment. Among them for example, one study stated that the continuous 100% accuracy is often signed as overfitting (Julien & Sobrino, 2009). Despite this general overfitting behaviour, the most notable observation is the potential overfitting in CART and RF, which showing high accuracy, may not necessarily deliver the better real world classification performance. However, due to inconsistencies and unreliability in results, this hold-out method was later discarded. Later, to solve this, a more robust 10-fold-cross-validation approach was used for selecting the best performing classifier. In the next step, we conducted a confusion metrics comparison as the overall accuracy, Kappa and F1 score gave a numerical summary but the confusion matrix shows us a detailed breakdown of misclassifications (false positive and false negative).

4.1.2 Confusion Matrices Analysis

Confirming overfitting pattern CART's perfect accuracy and RF's near-perfect score are clearly reflected in the confusion matrices. Identifying model weakness SVM's higher misclassification rates become more evident. All these confusion metrics are attached in the Appendix as Figure A1, A2, A3, A4. These figures showing the confusion matrices for the year 2010, 2015, 2020, 2024 is giving us a deeper insight into the classification performance of RF, SVM and CART.

CART consistently produces zero misclassification across all the years of observation. Confirming the overfitting concerns raised earlier. RF produces a small number of false negatives and false positives, but the overall performance remains extremely high. And SVM struggles the most, showing higher false positives and false negatives compared to RF and CART. This matched the lower accuracy and kappa values seen in the previous performance Table 5. SVM struggles more in certain years, possible due to variations in spectral reflectance across different Landsat sensors. All together these insights of raw performance metrics lead us to the decision that the method cannot be used in further steps. So, this total hold-out approach is discarded for the study. The next step is about cross-validation for selecting an appropriate model and confirming the best generalisation performance.

4.1.3 10-Fold-Cross-Validation

The 10-fold-cross-validation results provide a more robust and realistic assessment of classifier performance by ensuring the models are tested on multiple training-validation splits. This helps mitigate the risk of overfitting and improves generalisation mainly when limited samples are available.

RF consistently performs the best, with accuracy ranging from 0.92 to 0.96 across the years. It has huge Kappa values (0.84-0.96), indicating strong alignment beyond random chance. Precision, recall, F1 scores also showed high (>0.90), expressing its balances false positives and false negatives well. Overall, RF's high accuracy and Kappa values, combined with balanced recall and precision, confirm it as the best model for generalization. It is the ideal option for classifying forests in Castellón Province since it may handle complex spectral interactions without overfitting.

By combining some techniques like-bootstrap aggregating, reduced variance, averaging these kinds of techniques of RF, it can create a robust model that can capture complex patterns in data while performing well on testing data. By successfully balancing the trade-off between generalization and model complexity, Random Forests are less likely to overfit than individual decision trees (Sage, 2018).

SVM consistently underperformed relative to the RF and CART. SVM has the lowest accuracy (0.80-0.84) and the weakest Kappa scores (0.60-0.72) across all years. In this

case precision, recall and F1 score are also the lowest, confirming that SVM struggles more with distinguishing between frosted and non-forested areas. SVM's lower accuracy suggests it struggles with spectral complexity. SVM's adaptability in handling complicated, mixed spectrum data is limited by its reliance on linear decision boundaries or hyperplanes. For this reason, SVM's linear method may fail to effectively separate and classify the data points (Burges, 1998). The classification performance metrics obtained through the 10-fold-cross-validation approach, are presented in Table no 6.

Year	Classifier	Average Overall Accuracy	Average Kappa	Average Precision	Average Recall	Average F1-score
2010	RF	0.92	0.85	0.91	0.95	0.93
2010	SVM	0.80	0.60	0.81	0.78	0.79
2010	CART	0.89	0.77	0.90	0.89	0.89
2015	RF	0.95	0.90	0.96	0.94	0.95
2015	SVM	0.85	0.72	0.86	0.88	0.86
2015	CART	0.92	0.82	0.93	0.91	0.92
2020	RF	0.96	0.92	0.99	0.94	0.96
2020	SVM	0.81	0.62	0.82	0.82	0.81
2020	CART	0.94	.89	0.95	0.95	0.95
2024	RF	0.92	0.84	0.94	0.90	0.92
2024	SVM	0.84	0.68	0.84	0.86	0.85
2024	CART	0.88	0.75	0.90	0.85	0.87

Table 6: Performance metrics of classifiers based on 10-fold-cross-validation

And about the last one CART, compared to the initial result RF showed perfect accuracy, CART's cross-validation scores dropped to 0.88-0.94, proving earlier overfitting concerns. In 2024, its accuracy is the lowest (0.88), it is showing less stability over time. Kappa values (0.77-0.94) confirm CART's reduced ability to generalize access to different training-validation splits. Cross validation demonstrated that CART's performance had declined, proving that its previously flawless performances may have been inflated. But it remains a strong choice after RF.

Decision trees, including CART, sometimes can be unstable due to the slight changes in data that could produce an entire new tree. This tendency raises the possibility that CART may struggle with variability over time or in different datasets. Also, (Sage, 2018) stated that

CART can be used for prediction tasks but require techniques like cross validation for improving generalization (Cheng et al., 2018). So overall, this analysis provides a strong justification for selecting RF as the best classifier. Despite its relatively minor instability, CART remains a reasonable alternative. And SVM is not suitable as a classifier in our study.

4.1.4 Fold-wise Confusion Matrices Analysis

This section builds upon the cross-validation results, providing a fold-wise confusion matrix analysis for each classifier over the study year. It helps us to examine misclassification trends and identify strengths and weaknesses of each model. The confusion metrics for the 10-fold cross-validation of the RF and CART classifiers for the year 2010, illustrating classification performance across multiple folds are shown in Figure 4 and Figure 5 respectively.

For 2010

RF

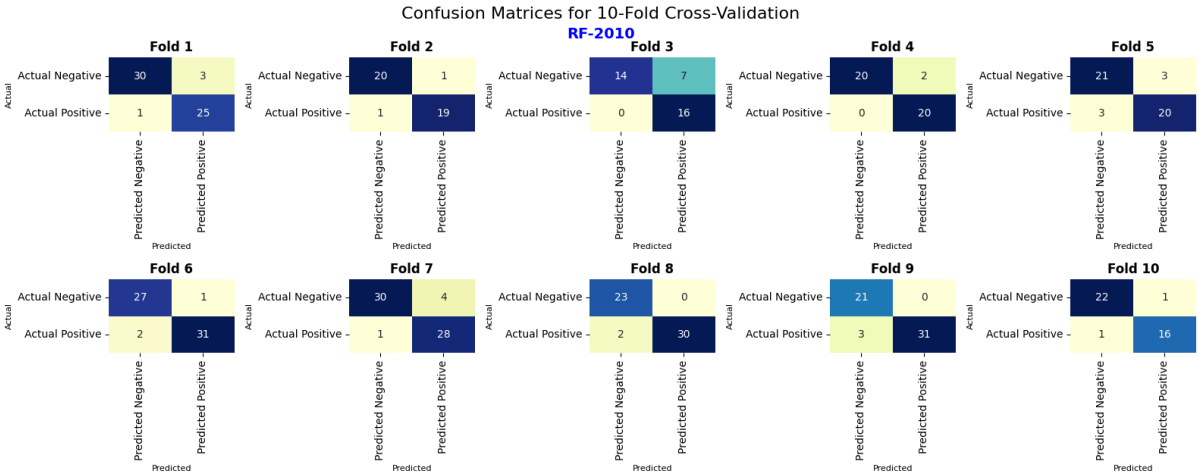


Figure 4: Confusion metrics for 10-fold-cross-validation-RF 2010

CART

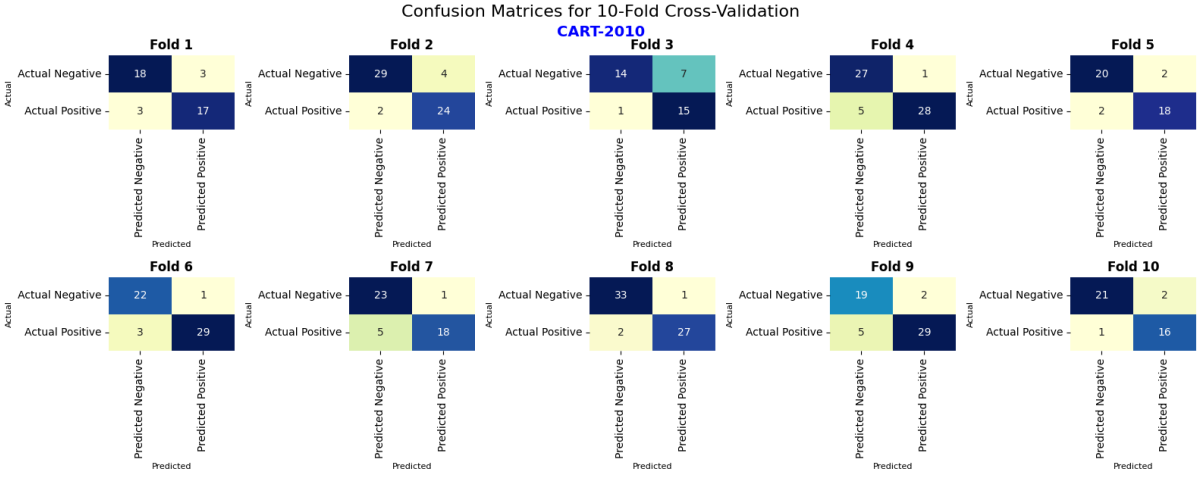


Figure 5: Confusion metrics for 10-fold-cross-validation-CART 2010

RF performs well in most folds, with minimal false positives and false negatives. RF shows consistently good accuracy across all folds. SVM shows higher misclassification rates compared to RF and CART. There are noticeable false positives (non-forest misclassified as forest) and false negatives, proving weak classification. Performance fluctuated significantly between folds, showing sensitivity to variations in the datasets. CART also performed well, but misclassification errors were slightly more visible compared to RF. In 2015, RF showed near perfect classification in most folds. Performance slightly improved compared to 2010, proving RF adapts well over time. SVM is still falling behind with significant misclassification rates. CART performance drops slightly compared to RF, as it showed instability across different folds. In 2020, RF showed the best performance, but a small increase in false positives was noticeable. Some folds show some minor misclassification, maybe because of spectral variability in the satellite images compared to previous years. SVM continues to struggle. Many folds show significant false positives, that means it is unable to handle spectral variability efficiently. At this point it is evident that SVM is not suitable for this classification task. And, for CART false positive increases in multiple folds. False positives are indeed an important metric in evaluating classification performance in binary classification tasks like this research (Klusowski, 2020).

In 2024, CART's weakness became more prominent, and RF's results remained superior. RF showed minimal classification errors. Some slight increase in false negatives appears but overall proves a consistent stability and adaptability. And, for CART, its instability becomes more noticeable in 2024. False positives increase significantly where CART fails to distinguish between forest and non-forest areas. The dropped performance further proves that CART's previous near perfect results were due to overfitting rather than true predictive power. Also, CART finds it challenging to generalize, proving it a less reliable choice than RF. The confusion metrics for the 10-fold cross-validation of the RF and CART classifiers for the year 2024, illustrating classification performance across multiple folds are shown in Figure 6 and Figure 7, respectively.

2024

RF

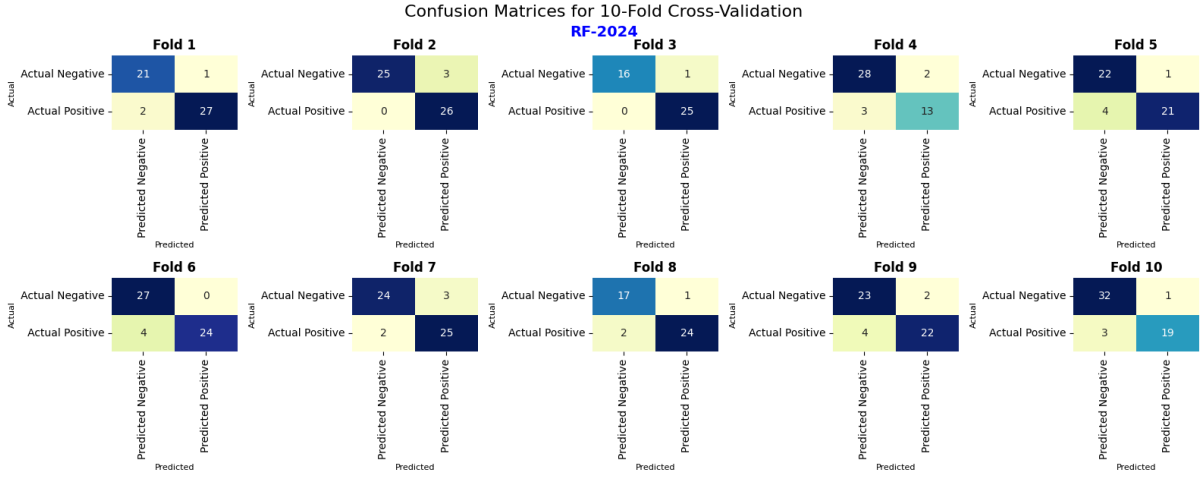


Figure 6: Confusion metrics for 10-fold-cross-validation-RF 2024

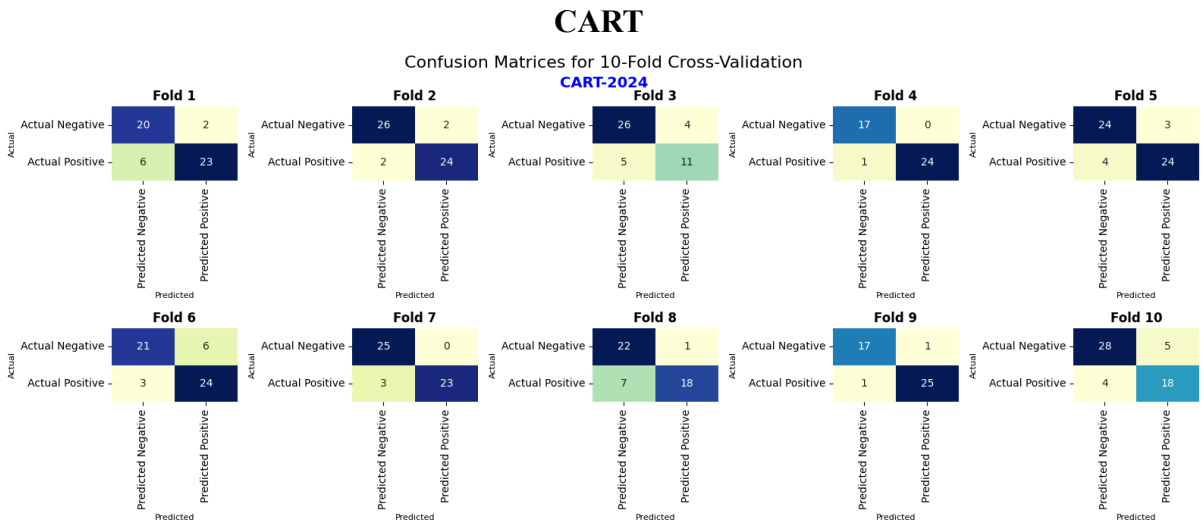


Figure 7: Confusion metrics for 10-fold-cross-validation-CART 2024

So, overall RF was consistently good, CART, initially appearing competitive but became insatiable over time. SVM continues to prove highest misclassification rates. Considering the facts, RF is the most dependable and applicable model, while highlighting CART's weakness in long-term trends and SVM's overall limitations.

4.1.5 Area Analysis

The section represents estimated forest and non -forest areas for each classifier. The areas are calculated based on classification to examine spatio-temporal change of forest cover and assess the reliability of different classifiers in estimating land cover changes. Across all the classifiers, a clear declining trend in the forest area is evident. The rate varies from across classifiers, showing differences in how each algorithm interprets forest cover.

The steady trend implies that RF reliably represents increasing deforestation. In the earlier years RF predicted moderate forest cover, but in later years less. That might be for reflecting real life forest cover changes. The trend aligns with the urbanisation and land use shifts in Castellon province. Study revealed that between 1985 and 2020, Castellón's built-up areas grew by 110.37%, creating 56.59 km² of newly developed urban growth (Sobrino et al., 2024). SVM predicts a consistently higher first area than RF and CART in later years, especially in 2020 (3365.59 sq km) and 2024 (2407.02 sq km). Also, it estimates a slight increase in forest from 2015 to 2020, which is unrealistic. Again, the forest area in 2024, SVM is significantly higher than the RF and CART, suggesting the potential overestimation. The reason for SVM's this trend is: SVM might be less sensitive to subtle spectral differences between forests and non-forests because of its reliance on decision boundaries.

A study reveals that the categorization made by SVM for complex vegetation may miss subtle spectral differences (Sobrino et al., 2024). Also, another study stated that when SVM and RF were compared for vegetation classification, the results showed that SVM was able to catch certain minor differences, as it performed better for classes that are not uniform in a

spectral sense and had similar spectral characteristics with their surroundings (Sabat-Tomala et al., 2020). So, lack of spectral adaptability in SVM likely contributes to these inconsistencies over time.

CART initially estimates higher forest cover than RF (2015: 3721,95 sq km) but in later years (2020 and 2024), it predicts the lowest forest cover among all classifiers. But in 2024, It estimates only 1760.43 sq km of forest lower than both RF and SVM and the amount of non-forest highest among all classifiers. This suggests that CART is extremely sensitive to variations in spectrum, CART's decision-tree nature makes it prone to hard splits as CART may over-segment regions and incorrectly label partially degraded forests as non-forest (Bishop & Nasrabadi, 2006). Also, some reasons can be like spectral variations due to soil exposure, seasonal changes or urban expansion may have caused CART to overestimate deforestation.

The estimated forest and non-forest areas for different classifiers over the years 2010, 2015, 2020, 2024, derived from classification results are summarized in Table 7.

Year	Classifier	Forest Area (sq km)	Non-Forest Area (sq km)
2010	RF	3509.27	3127.28
2010	SVM	3279.27	3357.28
2010	CART	3234.17	3402.38
2015	RF	3315.69	3320.86
2015	SVM	2659.71	3976.84
2015	CART	3721.95	2914.60
2020	RF	2198.09	4438.46
2020	SVM	3365.59	3270.96
2020	CART	1794.63	4841.92
2024	RF	1783.78	4852.77
2024	SVM	2407.02	4229.53
2024	CART	1760.43	4876.12

Table 7: Forest and non-forest area estimation based on classification (2010-2024)

So, based on the overall discussion, RF remains the most balanced and reliable classifier, capturing gradual deforestation without extreme underestimating (CART) or excessive overestimation (SVM). RF prominent with most reliable estimates, showing gradual and consistent changes over time.

4.1.6 Visual comparison of Classified Maps

This section is about examining classification output. It is for reviewing the classified images for each classifier and year, so that we can search for spatial distribution, classification consistency, land-use transitions, boundary artifacts, misclassification trends, spatial inconsistencies, etc. while analysing land use dynamics in Castellón Province. Here to be noted that, all the classified maps are compared to each year's images respectively. Here, the classified images from three classifiers are presented in figure 9, 10, 11 comparing them with the reference of Landsat image Figure 8.

Figure 8: Landsat reference image of Castellón Province for visual comparison

4.1.6.1 RF classified maps

The classification results obtained using RF classification for the years 2010, 2015, 2020, 2024 are depicted in Figure 9, illustrating the spatial distribution of forest and non-forest areas across Castellón Province**.**

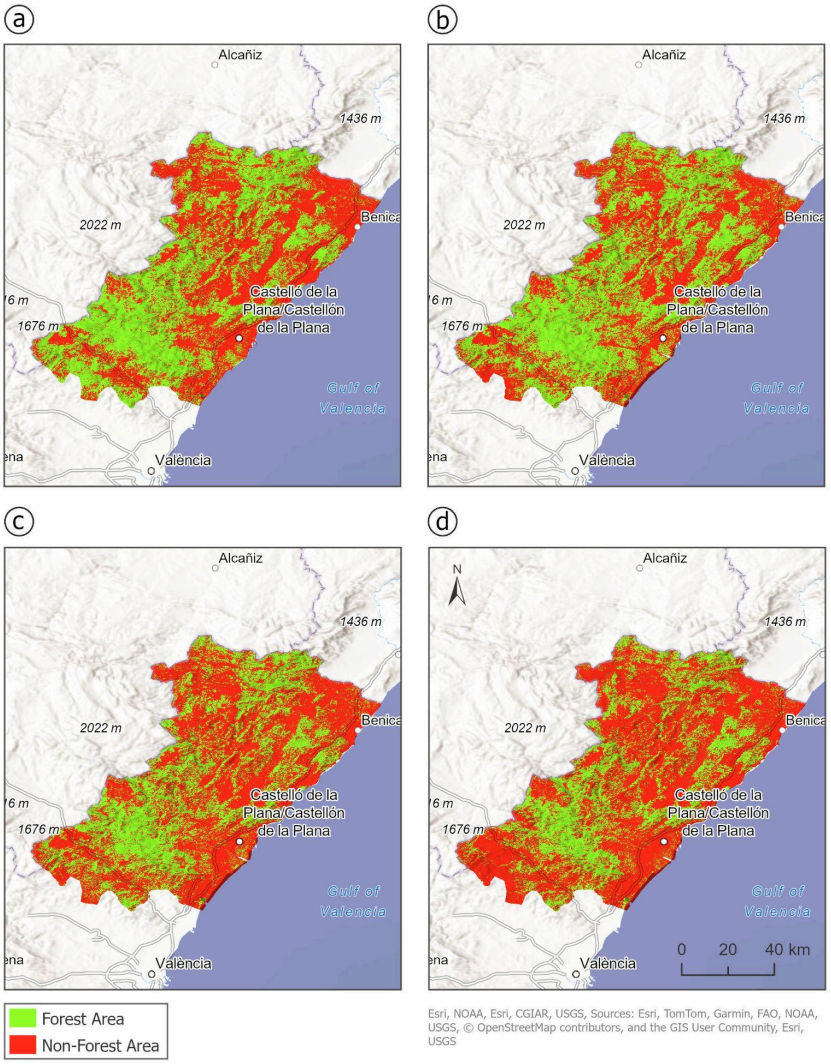


Figure 9: RF classified maps for the years a) 2010, b) 2015, c) 2020, and d) 2024 in Castellón Province

Spatial aspects:

RF depicts the balanced classification of forest cover, closely aligning with the reference image. Key forested regions, particularly in western and northwestern parts of the province, are accurately mapped. Natural Park of the Sierra de Espadán (Southwest of Castellon de la plana) retains its dense forest coverage across all the years. Penyagolosa Natural Park (north central part of Province) is consistently classified as a forest, reflecting the satellite image's depiction of reality. Alto Mijares (Northwest) and Alcalatén (west central) also show stable forest cover with RF, with insignificant decrease over time. In the eastern and southwest coastal regions (Castellón de la Plana, Benicàssim, and Oropesa del Mar), urban and semi urban areas are appropriately categorised as non-forest. Similar to satellite image, the southern part of the province's agricultural zones (Vila-real, Nules, and La Vall d'Uixó) remains as non-forest, with only scattered green patches visible. Forest cover is maintained in

protected areas like Sierra de Espadán and Penyagolosa Natural Park, though deforestation is advancing on their peripheries.

Trend over the years:
Gradual deforestation in the northwest and western mountain regions is well captured by RF, especially between 2020 and 2024. Urban expansion around the eastern coastal and central parts of the province is reflected in decreasing forest area.

4.1.6.2 SVM Classified Maps

The classification results obtained using SVM classification for the years 2010, 2015, 2020, 2024 are depicted in Figure 10, illustrating the spatial distribution of forest and non-forest areas across Castellón Province**.**

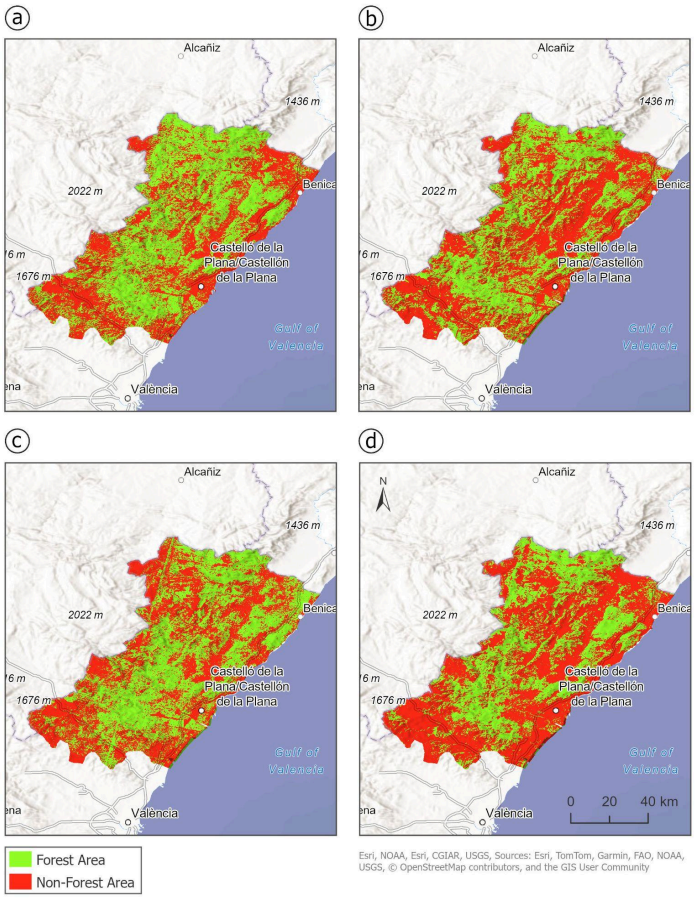


Figure 10: SVM classified maps for the years a) 2010, b) 2015, c) 2020, and d) 2024 in Castellón Province

Spatial aspects:
SVM frequently over classified forest cover, especially in Castellon Province's Northwest, east and coastal regions. The mediterranean shrubland on the northeastern shore, close to Peñíscola, is mistakenly categorised as deep forest. It is inaccurate to count urban greenery as

continuous woodland in Vila-real (southern part) and Castellón de la Plana (eastern part). In the southeast region of La Vall d'Uixó, low-density vegetation and orchards appeared as forest, despite being agricultural areas. Because of NDVI's sensitivity to mixed pixels, coastal areas close to Oropesa del Mar and Torreblanca (the eastern coastal zone) show more forest than in Landsat image.

Trend over the years:
SVM predicts higher forest cover in 2010 and 2015 compared to RF and reference images. Deforestation is visible in 2015 and 2014 but, compared to RF, forest reduction is less noticeable. It struggles to differentiate between urban greenery, mixed vegetation and true forest, leading to misclassification of forest area.

4.1.6.3 CART Classified Maps

The classification results obtained using CART classification for the years 2010, 2015, 2020, 2024 are depicted in Figure 11, illustrating the spatial distribution of forest and non-forest areas across Castellón Province.

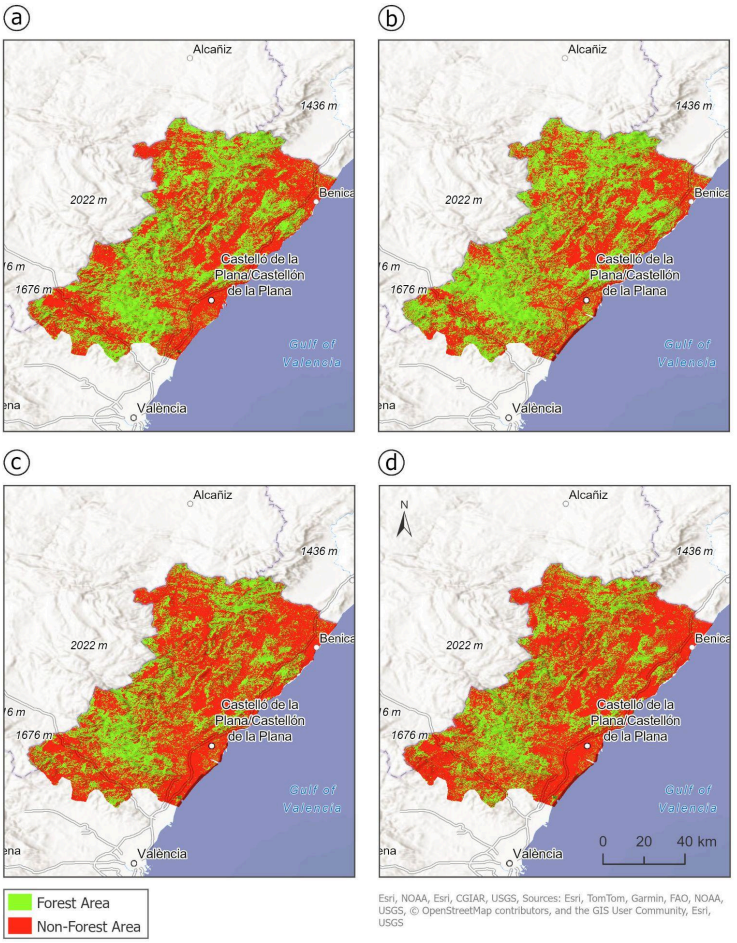


Figure 11: CART classified maps for the years a) 2010, b) 2015, c) 2020, and d) 2024 in Castellón province

Spatial aspects:
CART is classifying less forest area compared to RF and SVM. Spare tree cover and degraded forests in Penyagolosa Natural Park (north central) and Sierra de Espadán (southwest) are misclassified as non-forest. Some inland mountainous areas (Alt Maestrat, northwest) showing significant forest loss which is not as noticeable in the satellite view. Compared to RF and SVM, the urban and industrial expansion near Castellón de la Plana (east) and Villarreal (southeast) is well detected. Most of the agricultural land in the southern portion of the province (Nules, La Vall d'Uixó) is categorized as non-forest, with little misclassification.

Trend over the years:
The CART classification trends show a fluctuating trend in forest cover across the study period. The forests are stable in 2010 and 2015 with proportionally more forests in mountainous areas and natural parks. But from 2020 and 2024, the classification indicates significant deforestation. The overall CART classification indicates a mixed pattern of forest decline and regeneration, with a visible expansion of non-forest areas in recent years. CART's results in comparison to RF, the forest reduction appears more severe in 2020 and 2024. Aggressively, sometimes addressing fragmented forest as non-forest.

4.1.6.4 Summary of Visual Inspection of Classified Maps

After overall inspection, classified maps by different classifiers were compared in different areas. The example of this inspection was in appendix Figure A.22. RF successfully captures the finest details of the landscape, closely resembling the actual forest and non-forest distribution seen in the satellite imagery. It maintains strong boundary differentiations, small vegetation patches and mixed land cover types are well preserved. SVM tends to overestimate forest areas- classifying non-forest regions incorrectly as forests. Also, some regions appeared fragmented, indicating poor boundary definition between forest and non-forest areas. SVM struggles to detect finer details. On the other hand, CART fails to capture finer variations in land cover, leading to misclassifications. The output appears overly generalized, missing small first patches and misclassifying non-forest areas. This results in lower classification accuracy compared to RF.

Overall, RF remains the best-performing classifier, showing high spatial accuracy and realistic forest decline. SVM overestimates forest cover, particularly in mixed land use zones, CART underestimates the forest area, making it more suitable for deforestation risk assessment. Urban expansion and land-use transitions are well captured in RF but misclassified in SVM and CART.

So, RF is showing the most reliable classifier for forest and non-forest cover differentiation by accurately representing the Castellon Province's diverse land, from mountain forest to urban zones and agricultural lands.

4.1.7 Summary of Best Classifier

To determine the most suitable classifier of forest cover classification, a comparative evaluation of RF, SVM, and CART was conducted before.

Here, RF consistently performs the best in classifying metric categories making it the most reliable classifier. Several studies stated that RF performed well for forest classification. (Aziz et al., 2024) showed in their study that SVM fails at dealing with mixed landscapes which results in moderate area consistency, but low reliability while reasonably accurate, CART over predicts deforestation, lowering long-term reliability. Another study by (Praticò et al., 2021) demonstrated in their study that RF was the best classifier in their work with overall accuracy of 88%, while SVM with 83%, and CART with 80% accuracy. So, RF outperforms in all metric categories in the most reliable classifier in previous studies that completely aligns with us. Following Table 8 presents a comparative summary of their performance based on some key observations.

Metric	RF	SVM	CART
Average accuracy	0.94 (High)	0.83 (Low)	0.91 (Medium-High)
Kappa	0.88 (High)	0.66 (Low)	0.81 (Medium)
Area Consistency across years	High	Medium	Low
Area Estimate Reliability	High	Low	Medium
Visual inspection Quality	High	Medium	Low

Table 8: Comparative performance evaluation of RF, SVM and CART

4.2 Additional Analyses for the Best classifier (Level 2)

4.2.1 Seasonal Analysis of RF Model (May-July)

This section assessed how well our best performed RF model worked on a seasonal scale of summer months (May-July). This study assessed how seasonal changes influence the accuracy of classification and evaluated model performance on estimation of forest and non-forest areas. The methodology remains the same, utilizing 10-fold-cross-validation, confusion matrix analysis, area estimations and mapping assessments as in previous sections.

4.2.2 Performance Metrics and Confusion Metrics Analysis

A seasonal assessment was conducted for the period of May-July, 2024 to evaluate the consistency of classifier performance under seasonal variation.

RF remains high classification accuracy (90%) and strong Kappa (0.80), indicating good agreement with ground truth. Precision and recall values remain balanced, suggesting RF effectively distinguishes between forest and non-forest areas in summer. SVM shows lower kappa (0.70), reflecting moderate misclassification issues. CART performance remains decent but slightly weaker than RF. So, despite seasonal variations RF consistently performs as the reliable model for summer season classification. Compared with the annual classification performance Table no 6, this seasonal classification maintains high performance but is slightly lower than the full year model. Overall, RF remains stable in a short temporal window.
Table 9 summarizes the key findings of performance matrices from the 10-fold-cross validation results for RF.

(May-July) of the Year	Classifier	Average Overall Accuracy	Average Kappa	Average Precision	Average Recall	Average F1-score
2024	RF	0.90	0.80	0.92	0.89	0.90
2024	SVM	0.83	0.70	0.83	0.85	0.84
2024	CART	0.87	0.75	0.90	0.85	0.87

Table 9: Seasonal performance evaluation of RF, SVM, CART (May-July, 2024)

The details of the confusion matrix are attached in the Appendix. It is showing the same as RF is showing the best results.

4.2.3 Area Estimations Analysis in Summer Season

It was evident that Forest area reduced over time. It showed lowest values recorded in 2020 and 2024 (from 3416 sq km to 1829 sq km). The slight recovery in 2024 (from 1829 sq km to 2108 sq km) suggests restoration initiatives or classification misclassification/refinements. But, like the annual one, summer seasonal classification supports the overall decrease in forest areas.

The seasonal forest and non-forest area estimation for the summer period (May to July) from 2010 to 2024 are summarized in Table 10.

(May-July) of the Year	Forest Area (sq km)	Non-Forest Area (km)	Total Area (sq km)
2010	3216.41	3330.82	6547.23
2015	3416.21	3220.34	6636.54
2020	1829.80	4806.74	6636.54
2024	2108.50	4528.05	6636.54

Table 10: Seasonal forest and non-forest area estimations for summer of 2010-2024

4.2.4 Visual Representation of Summer Season Classification

The seasonal forest cover classification for May to July 2024, generated using the RF classifier, is shown in Figure 12. The maps provide spatial representation of forest areas compared to non-forest areas, indicating seasonal changes in vegetative coverage within Castellón Province.

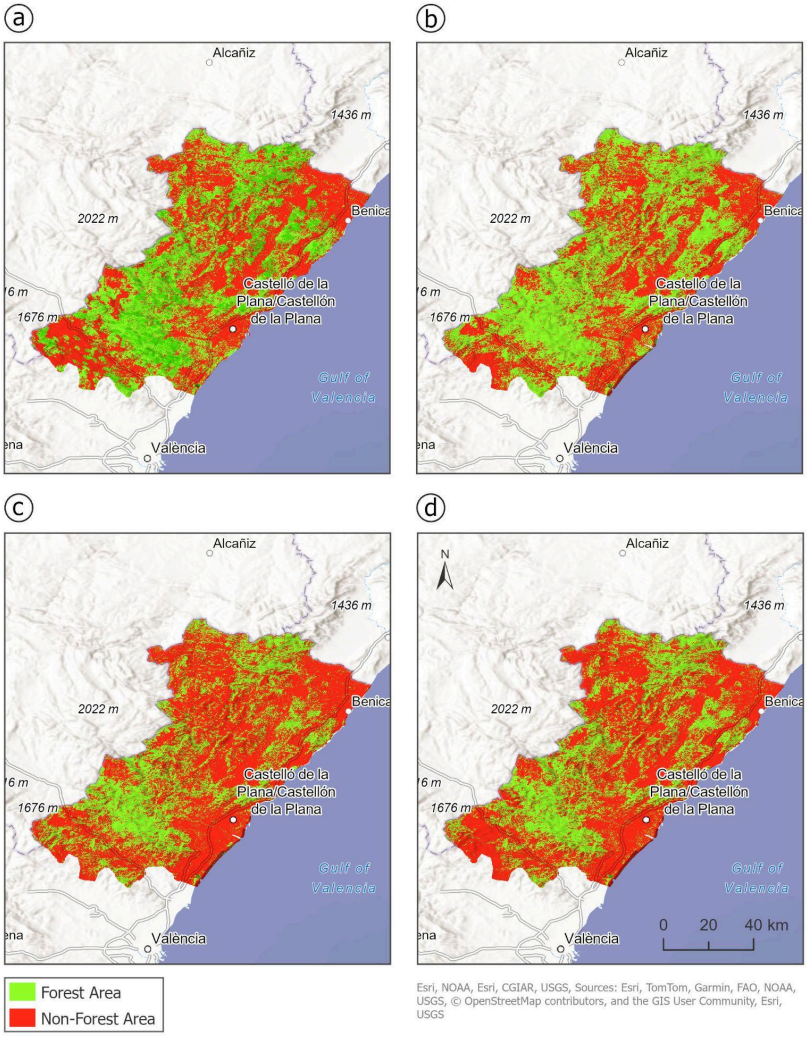


Figure 12: RF classified maps for seasonal forest cover (May–July, 2024)

Careful inspection of the classified summer maps from May to July confirms that RF accurately distinguishes between forests and non-forested areas. The eastern coastal regions like: Castellón de la Plana, Benicàssim and Oropesa del Mar are correctly identified as non-forest areas. Mountainous forests in western and northwestern areas, such as the Sierra de Espadán and Penyagolosa Natural Park maintain accurate classification. Many sparse forests and many semi-urban vegetation patches in areas such as Vila-real and La Vall d'Uixó were classified precisely, thereby avoiding overgeneralizations. RF reliably performs spatial classification, when applied to this seasonal, small-scale analysis, thus showing its effectiveness. But comparison with the full-year classification maps highlights shifts in vegetation appearance and potential seasonal misclassification. The next section is describing some potential issues related to these observations.

4.2.5 Discussion on Summer Seasonal Influence on Classification Performance

Increased temperatures, less precipitation, and decreased soil moisture during the summer months (May–July) may cause visible change in forest reflectance. Because vegetation stress can change spectral characteristics, classifying forests can be a little more difficult. RF still captures the general forest structure but might misclassify temporarily dry forests as non-forest due to spectral variations over seasonal time.

A study supports that in the seasonal tropics, due to phenology vegetation depicts large reflectance, which can complicate land cover change monitoring (J. Liu et al., 2015). Another study noted that seasonality can lead to possible misclassification of change because of forest phenology (Silveira et al., 2018). And for the justification of the unusual decreased amount of non-forest area in 2024 can be that in the summer, mediterranean drylands and coastal scrublands may reflect different signatures in summer, causing minor classification shifts.

One study stated that like many parts of Mediterranean Spain, Castellón Province has exceptionally dry summers. Increased fire risk and vegetation stress could arise from this seasonal aridity. So due to the fire and all these reasons the overall deforestation may have occurred in our study period (De Castro et al., 2005). Another thing is that the different types of forests, tree species in Castellón Province, particularly Mediterranean forests, pine-dominated areas, and mixed woodland, may respond differently to seasonal variations and eventually affect classification accuracy. For instance, Aleppo pine found in Sierra de Espadán, Penyagolosa, and coastal hills near Benicàssim, experience moisture stress in summer, sometimes leading to as misclassification as forests due to increased soil exposure. Then Holm oak dominates in higher elevation like: Alto Mijares and Penyagolosa, maintaining a stable canopy year-round, reducing misclassification in that region. Different types of pine species in mountainous zones like Penyagolosa and Els Ports, keeping dense canopies with minimal spectral variations, ensuring classification accuracy. The riparian forest types are located near the Mijares and Palancia rivers; they are subjected to summer water stress, producing spectral shifts that can lead to classification errors.

Overall, oak and mountain pine forests are the most stable forest types in RF classification, while Aleppo Pine, riparian forests, and shrublands are more susceptible to seasonal misclassification, especially during summer. Despite these changes, RF continues maintaining high classification consistency with overall reduction in forest areas in the summer season that are supported by many studies.

4.3 Correlation Between Small-Scale Summer Season (May-July) and Full-Year Classification Results

This section connects seasonal-scale and annual-scale forest cover trends.

Key similarities:

- RF remains the most reliable classifier in both full-year (2024) and summer window (May-July 2024).
- Both forested and non-forested regions preserve spatial stability, ensuring the stability of RF across time scales.
- Clear decline in forest cover. Major deforestation zones (e.g., Penyagolosa Natural Park, southern Espadán region, and scattered lowland forests) were clear in both scale results.
- In both datasets, urbanized areas close to Castellón de la Plana, Vila-real, and Benicàssim show comparable growth tendencies.

Key differences:

- Accuracy slightly dropped to .90 than full years dataset due to seasonal vegetation stress, changing in spectral reflectance etc.
- Spectral alterations in drought-sensitive tree species (Aleppo Pine, riparian vegetation) can lead to some summertime misclassifications.
- The full-year RF classification estimated 1783.78 sq km of forest in 2024, whereas the summer window estimates 2108.50 sq km.
  This is technically opposite to the expected facts. Some reasons can be:
  - ❖ The full-year RF classification combines multi-season composite images which blend information from the whole year. That may cause forest as non-forests in the full year datasets.
    Research found that the frequencies of Landsat data with 0-5% cloud cover are very high in dry seasons compared to wet seasons. So, here in the full scale window where it is using a combination of all dry and wet season images, so it can be lower than the dry season's specific high frequencies images (Ayanlade & Nwaezeigwe, 2018)
  - ❖ The full-year classification relies on an annual composite of Landsat images; therefore, many regions may be considerably obscured by cloud cover or large seasonal atmospheric phenomena (e.g., winter moisture, haze, or fog in higher elevations like Penyagolosa).

- In summer, the inland dry zones of Sierra de Espadán (southwest) and some Aleppo Pine-dominated areas close to Peñíscola (northeast) are mistakenly classified as non-forest. This does not appear in full-year classification, addressing that summer specific vegetation loss affects reflectance values. But the evergreen forest remains stable.
- Mediterranean shrublands display higher misclassification rates in summer. In the full-year dataset, these zones are more accurately classified.

4.4 Forest Cover Change Analysis

This section evaluates the temporal forest cover changes in Castellón Province over a 14-year period (2010, 2015, 2020, and 2024) using RF classifiers, which was previously identified as the most reliable model. This section provides a comprehensive understanding of forest dynamics in the study regions.

4.4.1 Forest Area Statistics

This section highlights the temporal dynamics of forest cover changes over the study period. Table 11 provides a comprehensive summary of forest and non-forest areas in Castellón Province as derived from RF-based classification for the years 2010, 2015, 2020, 2024.

Year	Forest Area (sq km)	Non-Forest Area (km)	Total Area (sq km)
2010	3509.27	3127.28	6636.55
2015	3315.69	3320.86	6636.55
2020	2198.09	4438.46	6636.55
2024	1783.78	4852.77	6636.55

Table 11: Forest area statistics for Castellon Province (2010-2024)

Forest cover has declined significantly over the years, with a total loss of 1725.49 sq km between 2010 and 2024. The substantial growth of non-forest areas demonstrates the conversion of forest lands into other land uses. There was the most significant reduction in forest area, with 1117.60 sq km of forest loss between 2015 and 2020.

In the earlier parts we mentioned anthropogenic activities or other factors for forest reduction. In this section, we want to discuss some forest fire incidents that happened in our study period which can be a factor for forest loss. A significant forest fire happened in March 2023 near Villanueva de Viver, Castellón Province destroying 3000 hectares of land (Ottolini et al., 2024). A significant forest fire broke out in Gatova, Valencia (which includes Castellón

province) on June 28, 2017 (European Commission, 2017).

In addition, the Copernicus Emergency Management Service produces yearly reports on wildfire activity in Europe, including more detailed analyses on the significant fires of Spain. Such reports would give much-needed context and data regarding frequency, scale, and impact of the wildfires in areas like Castellón Province. So, all these massive fire events might have affected the forest area and possibly are a working factor behind the forest loss in Castellón Province (Jones et al., 2024).

4.4.2 Changes in Forest Cover

Following Table 12 presents the quantified forest cover changes over different time intervals, highlighting significant reduction in forest area.

Time Period	Forest Change (sq km)	Percentage Change
2010-2015	-193.58	-5.52%
2015-2020	-1117.60	-33.70%
2020-2024	-414.31	-18.85%

Table 12: Changes in forest cover across different time periods (2010-2024)

The period between 2015 and 2020 experienced the most rapid forest decline in Castellón Province, followed by a slight reduction in forest loss rates in the last four years. In 2010-2015 a relatively slow rate of forest decline (5.52%), likely because of slow changes in land use in agricultural and peri-urban areas. In 2015-2020 highest rate of deforestation (33.70%) suggesting a significant decrease in forest cover due to anthropogenic activities, urban expansion, and agricultural expansion Forest loss has slowed (-18.85%), suggesting either conservation initiatives, reforestation programs, or a decrease in land conversion.

For this justification of the improvement in forest loss, some events are found such as Red Eléctrica reported on a reforestation initiative. In Altura, Castellón, they accomplished reforesting of woodland. "The restoration of the landscape and wildlife habitat, and to the protection of biodiversity" was the stated goal of this €154,000 initiative. This suggests that the area is making some attempts to restore its forests (Redeia, 2021).

Also, a case study in 2022 stated that the second half of the 20th century has been characterised by the rural abandonment in several regions of Castellón province. It involved an intensive migration movement from inland areas towards coasts due to the general collapse of traditional agriculture and livestock activities, which caused a massive increase in forest cover in abandoned rural areas (Delgado-Artés et al., 2022b). The situation may continue in the later periods.

4.4.3 Time Series Analysis

The time-series analysis section visually represents table 11 and trends in forest and non-forest area changes over time. Figure 13 illustrates the temporal changes in forest and non-forest cover across Castellón Province from 2010 to 2024.

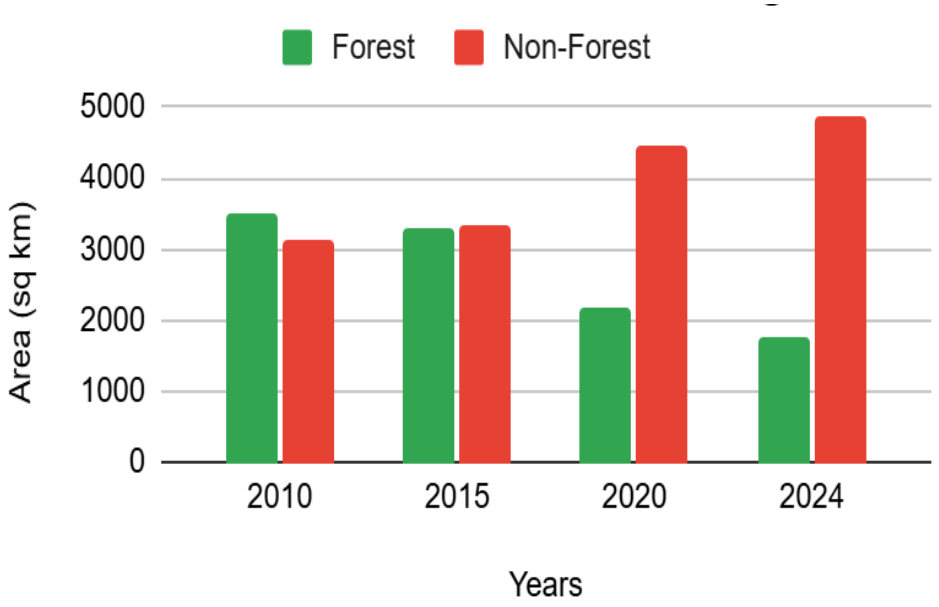


Figure 13: Time-series analysis of forest and non-forest cover change (2010-2024)

A sharp decline in forest cover is seen in 2020 from 2010 (2198 sq km from 3509.27 sq km), as revealed by the earlier numerical results. The stabilization of non-forest area expansion in 2024 suggests a possible slowdown in deforestation. It matches the observed pattern of land-use transitions in eastern and southeastern parts of the province with more developed urbanization.

4.4.4 Transition Matrices Analysis

Using transition metrics, this section provides a comprehensive analysis of forest cover change in Castellón Province for the three periods (2010-2015, 2015-2020, 2020-2024). This analysis examined the pixel-based area changes to quantify forest stability, deforestation, afforestation, and non-forest persistence. To facilitate these three transition matric maps (Figure 14,15,16) illustrate the spatial patterns of forest cover changes, while Table 13 quantifies these transitions based on pixel count and their equivalent areas (sq km). The following Table 13 presents pixel count based area calculation of forest cover change across three distinct periods: 2010-2015, 2015-2020 and 2020-2024.

The pixel count column represents the actual number of pixels classified under each transition category. Since each Landsat pixel represents 30m x 30m on the ground, pixel counts were converted into actual land area, sq km. It also allows a more detailed

interpretation of classification accuracy and spatial consistency. This helps validate the reliability of the classification results over time.

The following table 13 shows the extent of stable forest, deforestation, reforestation etc. in a comprehensive way.

Transition Type	Pixel Count (2010-2015)	Area (sq km) (2010-2015)	Pixel Count (2015-2020)	Area (sq km) (2015-2020)	Pixel Count (2020-2024)	Area (sq km) (2020-2024)
Forest → Forest (stable forest)	4,253,147	3827.83	3,723,997	3351.60	2,907,338	2616.60
Forest → Non-Forest (deforestation)	471,229	424.11	1,341,506	1207.35	934,383	840.94
Non-Forest → Forest (reforestation)	812,356	731.12	117,724	105.95	381,154	343.04
Non-Forest → non-Forest (stable non-forest)	4,140,005	3,726.00	4,493,510	4044.16	5,453,862	4908.47

Table 13: Transition matrix of forest cover change (2010-2024)

The following sections provide an organized discussion of three transition maps and their interpretations.

4.4.4.1 Forest Cover Change Dynamics (2010-2015)

Between 2010 and 2015, the province experienced a net loss of forest cover, though afforestation rates seem relatively higher during this period. This section shows the transitions of forest and non-forest areas in Castellón province during 2010-2015.

471,229 pixels (424.11 sq km) of forest were lost, mainly in Castellón de la Plana, Almassora and Burriana. The urban agricultural expansion played key roles here. Small patches of forest degradation were also visible near Villafranca del Cid and Atzeneta del Maestrat in the interior regions. 731.12 sq km of non-forest transitioned into forest, primarily along the mountainous regions of Penyagolosa Natural Park; Serra d'Espadà Natural Park. It indicates reforestation efforts or natural growth of forest. Despite deforestation, 3827.83 sq km of forest remained unchanged, with high range in the interior mountainous areas and parts of Alto Palancia.

Overall, this period witnessed moderate deforestation countered by notable afforestation efforts. The forest cover change patterns for 2010-2015 are visualized in Figure 14, depicting the spatial extent of deforestation and afforestation across Castellón province.

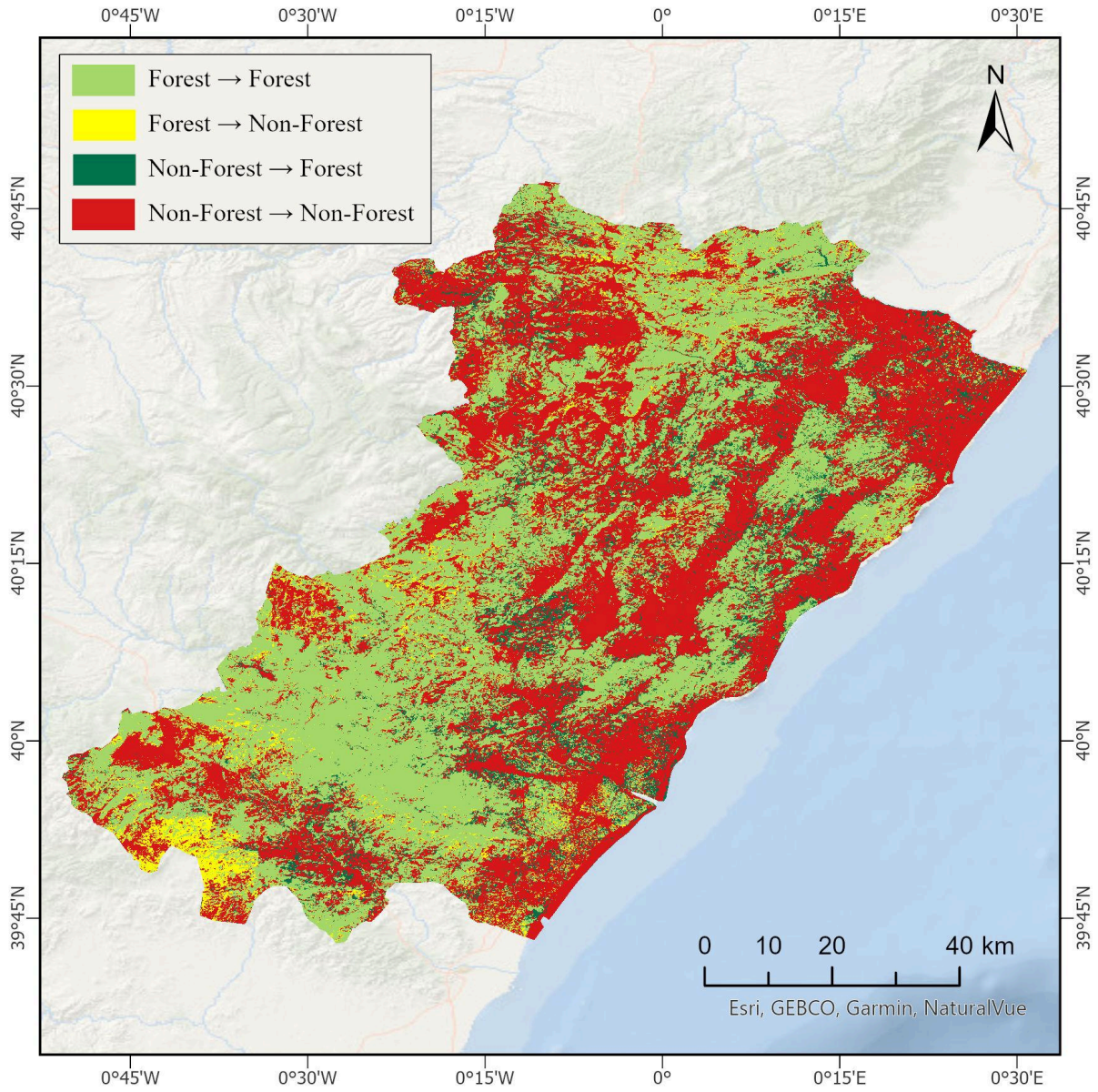


Figure 14: Forest cover changes dynamics in Castellón (2010-2015)

4.4.4.2 Forest Cover Change Dynamics (2015-2020)

Between 2015 and 2020, this period represents the most significant phase of deforestation in the study. As the result showed the highest deforestation rate and lowest afforestation efforts among all the periods.

In this period, 1207.35 sq km of forest cover was lost. It marks a significant increase from the previous period. The most intense deforestation occurred in coastal regions (Castellón de la Plana, Benicàssim, Oropesa del mar) and agricultural expansions in Plana Alta regions. Only

105.95 sq km of afforestation occurred. That indicated a sharp decline compared to 2010-2015. Regrowth was mainly shown in Penyagolosa and Serra d'Espadà Natural Park, but reforestation was much lower. For the stable forest areas: the forest retention rate dropped, covering only 3351.60 sq km. This is significantly lower than the previous period. More areas converted to urban settlements and agricultural fields. Also, in this period reduced core forest zones into Alsto Mijares and El Ports were noticed. The following Figure 15 shows the forest cover changes patterns in Castellón during 2015-2020.

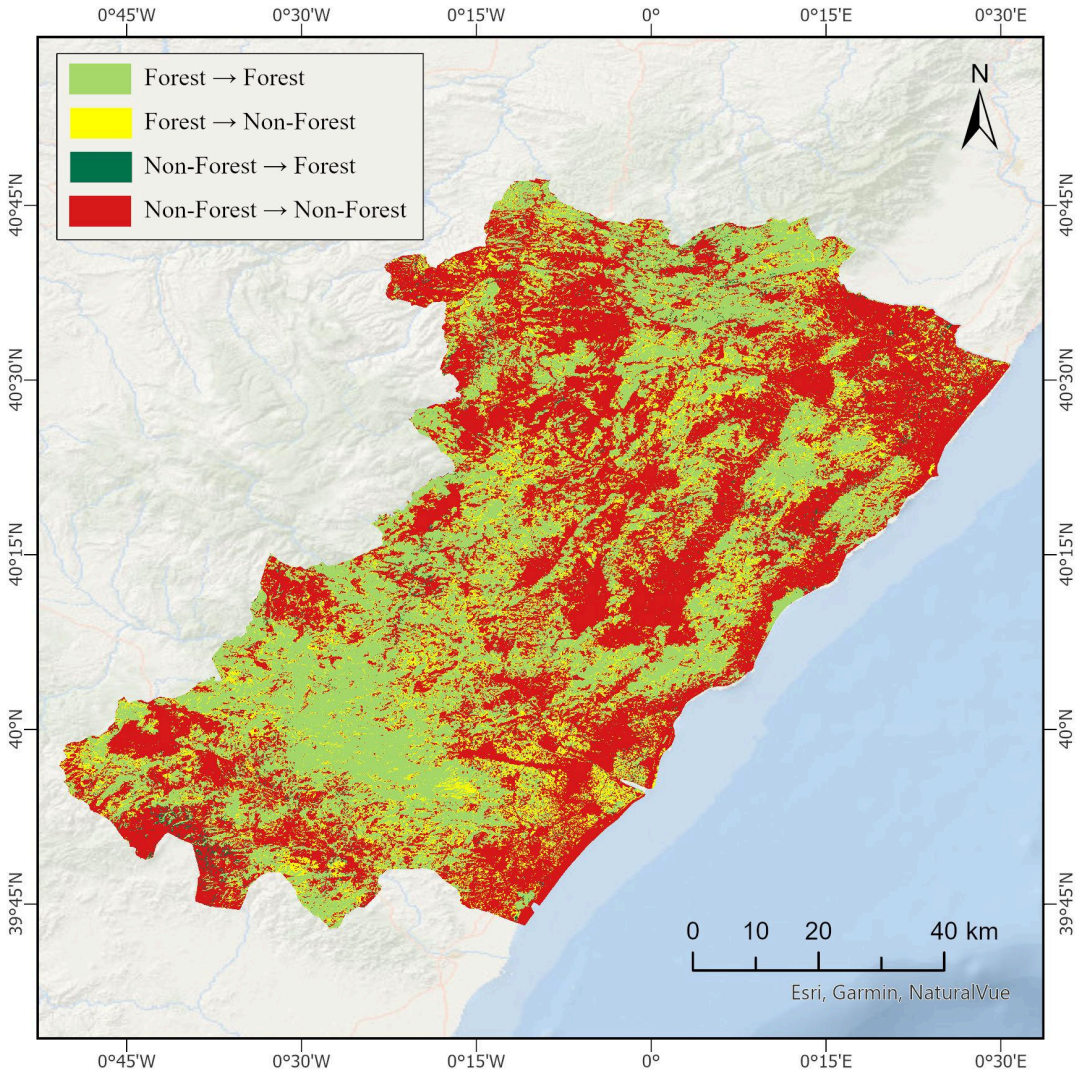


Figure 15: Forest cover change dynamics in Castellón (2015-2020)

Overall, this period is marked by widespread deforestation and the lowest afforestation rates, likely due to the increased human activity and climate variability.

4.4.4.3 Forest Cover Change Dynamics (2020-2024)

During this period, afforestation improved slightly. Along with this, deforestation was less compared to 2015-2020. The forest cover change patterns for 2020-2024 are visualized in

Figure 16, depicting the spatial extent of deforestation and afforestation across Castellón province.

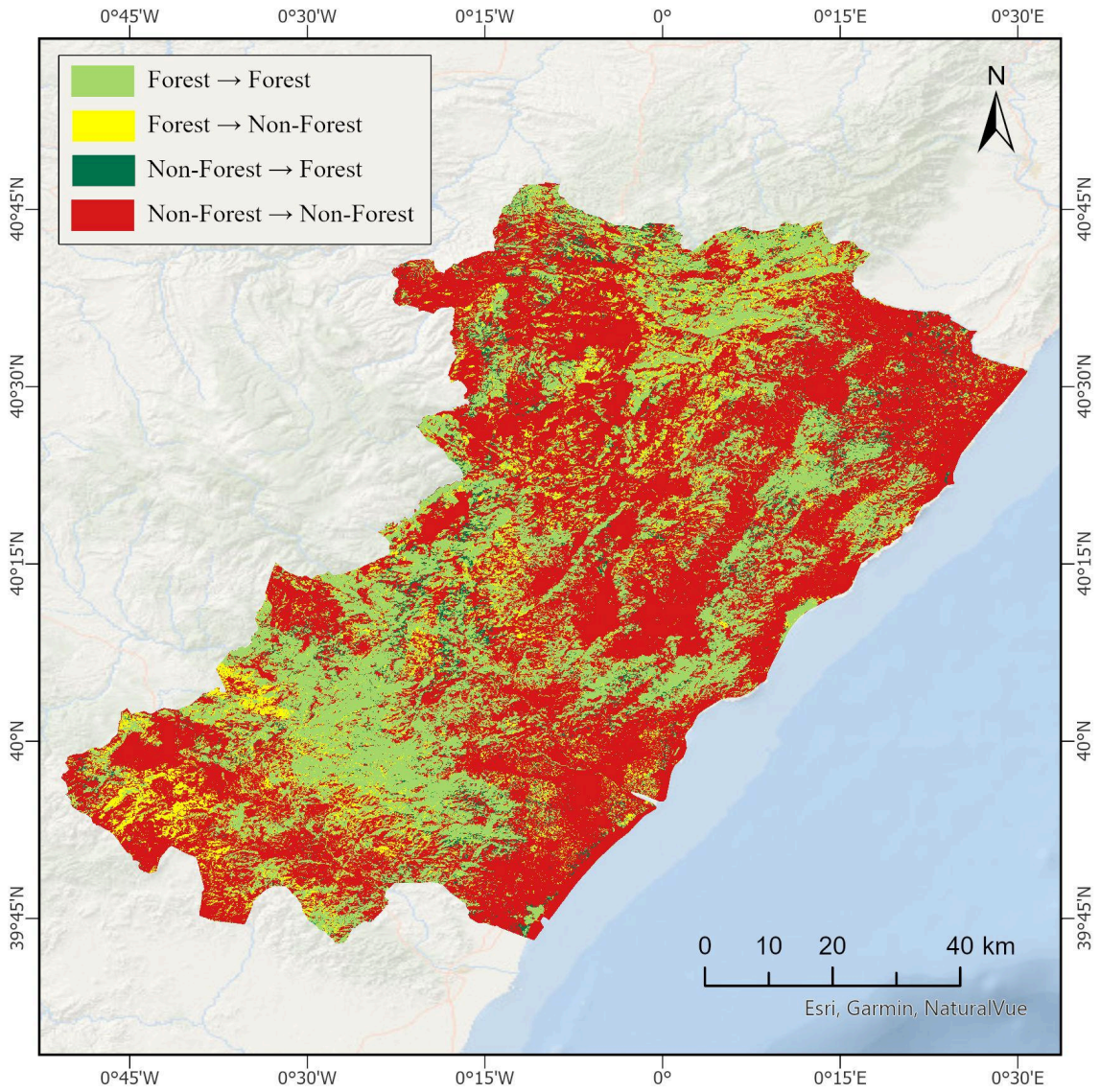


Figure 16: Forest cover change dynamics in Castellón (2020-2024)

In this period, 840.94 sq km of forest was lost, marking a slight reduction from the previous period. But deforestation still occurred in lowlands regions of Vall d'alba, Onda and parts of coastal areas. Afforestation of 343.04 sq km of new forest emerged, mainly in Serra Calderona and parts of Els ports, indicating improved conservation measures. 2616.60 sq km of forests remained intact, but overall stability continued to decline. The province still faced continued fragmentation, particularly in transition zones between agriculture and forested landscapes.

This period highlights a slight recovery in afforestation efforts, although deforestation remains an ongoing concern for the province.

Key observations:

- Among all the three periods, forest stability decreased with an area of 424.10 sq km (in 2010-2015) to 840.94 sq km (in 2020-2024). This indicated progressive deforestation of forest areas over time.
- The highest forest loss was recorded in 2015-2020, mainly due to urban expansion, agricultural land conversion and forest fire (specially in 2016 and after periods) events.
- The 2010-20105 period had the highest reforestation rate (731. 12 sq km). But the rate declined significantly to 105.95 sq km in 2015-2020. But again in 2020-2024 it recovered to (343.04 sq km). It suggests that forest management initiatives were reinforced.
- Major deforestation hotspots were in Castellón de la Plana, Almassora and Burianna and may be mostly driven by human activities.
- Forest expansion was mainly concentrated in Penyagolosa and Serra d'Espadà Natural Park.

Thus, the pixel-count based area calculation reinforces the trends seen in land cover transitions, confirming that non-forest areas are expanding more rapidly than forested areas are recovering. The scale of pixel counts based area calculation confirms that lost forest is not usually regained and, therefore, requires sustainable land-use planning.

4.5 Relationship between NDVI and Climatic Factors

At first the general trend of NDVI and climatic variables are examined separately to have a better understanding before conducting the correlation analysis.

4.5.1 Spatial Variation of NDVI/NDVI Trend

The NDVI trend map shown in Figure 17 visually depicts interannual variations in vegetation health across Castellón Province. It represents the distribution of mean NDVI across all years (2010-2024), underlining different vegetation cover levels. Darker green shades correspond to higher NDVI values, representing denser, healthier vegetation, while lighter shades represent sparse vegetation or degraded areas.

The highest values (above 0.5) are consistently observed in the northwestern and the southern part. Large parts of this area in the northwest are covered by the Penyagolosa Natural Park; other mountain ranges, like Serra d'Espadà Natural Park in the south, show almost the same vegetation density. These natural reserves exhibit persistent high NDVI values, indicating stable healthy vegetation cover. The centre and the inland areas, including Alcora etc., exhibit NDVI values ranging from 0.40 to 0.51. These areas have a mix of agricultural land, shrublands and scattered forest patches. In contrast, lower NDVI (<0.4) is dominant in the urban and industrial zones of the coastal cities of Castellón de la Plana, Benicàssim, and Oropesa. The consistent low NDVI values in this area corresponds to urban expansion,

industrial activity and agricultural practices etc. that limits vegetation density. The spatial distribution of NDVI variations across Castellón Province for the period of 2010-2024 is depicted in Figure 17, highlighting interannual changes in vegetation cover.

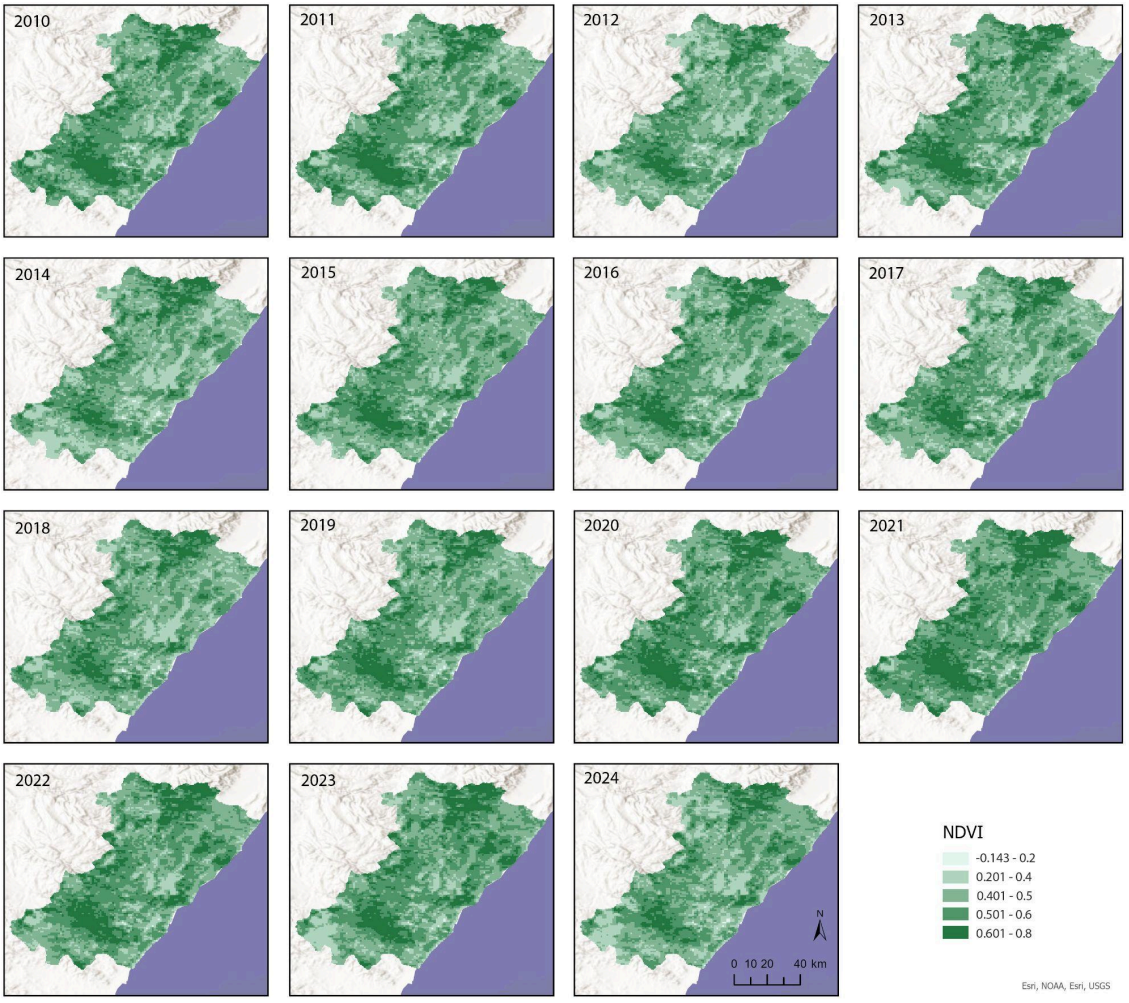


Figure 17: Spatial Distribution of NDVI trend across Castellón Province (2010-2024)

The mean NDVI values extracted from MODIS datasets over a 14-year period gave the necessary perspective into long-term vegetation changes. The data attached in Appendix Table A1. In the temporal pattern the NDVI started at 0.511 in 2010 and went through periodic fluctuations in the 14-year period. 2014 was the year with the minimum NDVI of 0.475, which reflects a probable dry period or increased land-use changes. In 2017 and 2018, NDVI values were 0.486 and 0.488, respectively, during a time of reduced precipitation coupled with increased temperature. This was followed by a notable recovery in the year 2020 and 2021, when NDVI reached its peak at 0.523 and 0.533, respectively. This period coincided with relatively higher amounts of rainfall (shown in Figure no 18) that allowed vegetation to regenerate. However, the NDVI declined again in 2024 to 0.486, indicating further climatic pressures or changes in land use affecting vegetation stability. The yearly NDVI maps provide a comprehensive insight into spatial and temporal vegetation dynamics, highlighting areas of stability, degradation, and regeneration.

4.5.2 Precipitation Trend

The annual precipitation variations across Castellón Province are illustrated in Figure 18, depicting fluctuations in precipitation patterns over the study period (2010-2024). The green line represents the total precipitation, and the red line indicates the linear regression trend, indicating potential long-term precipitation shifts.

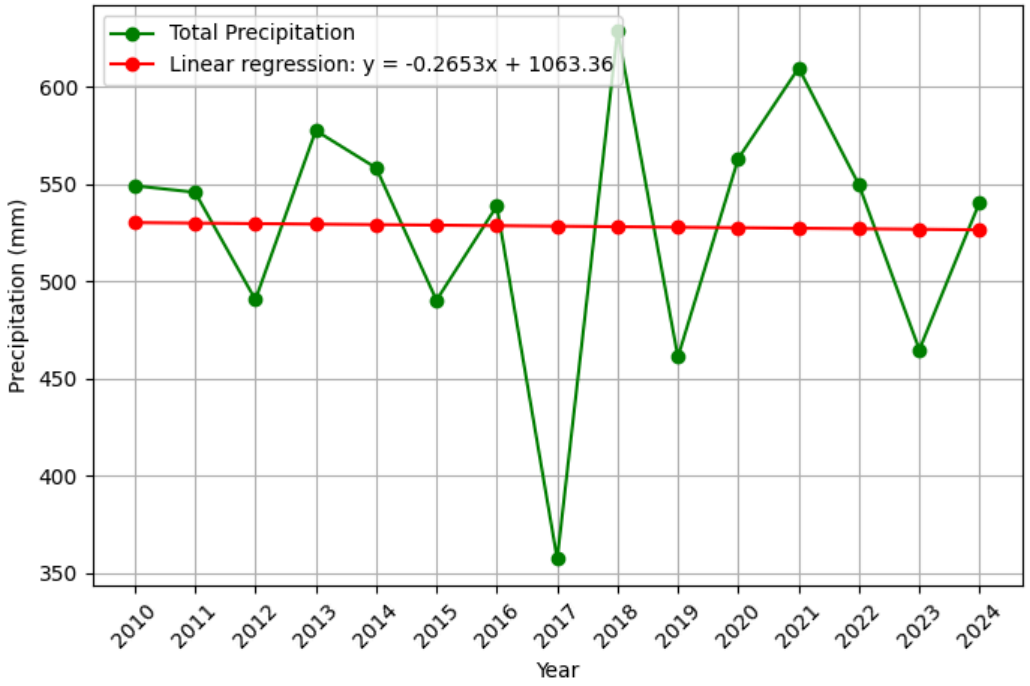


Figure 18: Annual Precipitation trend in Castellón Province (2010-2024)

The trend in precipitation in Castellón Province over the 2010–2024 period is highly variable inter-annually, with significant fluctuations in total precipitation. Regression analysis (red line) shows a negative trend (0.2653 mm per year), which means that the precipitation has been reduced over time. The year 2017 was highly dry, receiving minimum recorded rainfall of 357 mm, corresponding to a remarkable decline in NDVI. On the other hand, 2018 is representative of the maximum recorded rainfall of 628 mm; it resulted in partial recovery in respect to NDVI. Such fluctuation shows that vegetation cover is highly vulnerable to rainfall variability.

4.5.3 Temperature-2m Trend

The Temperature variation at 2 metres above the surface across Castellón Province is depicted in Figure 19. It illustrates an increasing trend in near-surface air Temperature or tropospheric temperature from 2010-2024.The green line represents the annual Temperature 2m values. And the red line represents the linear regression trend, indicating a gradual rise in temperature over the years.

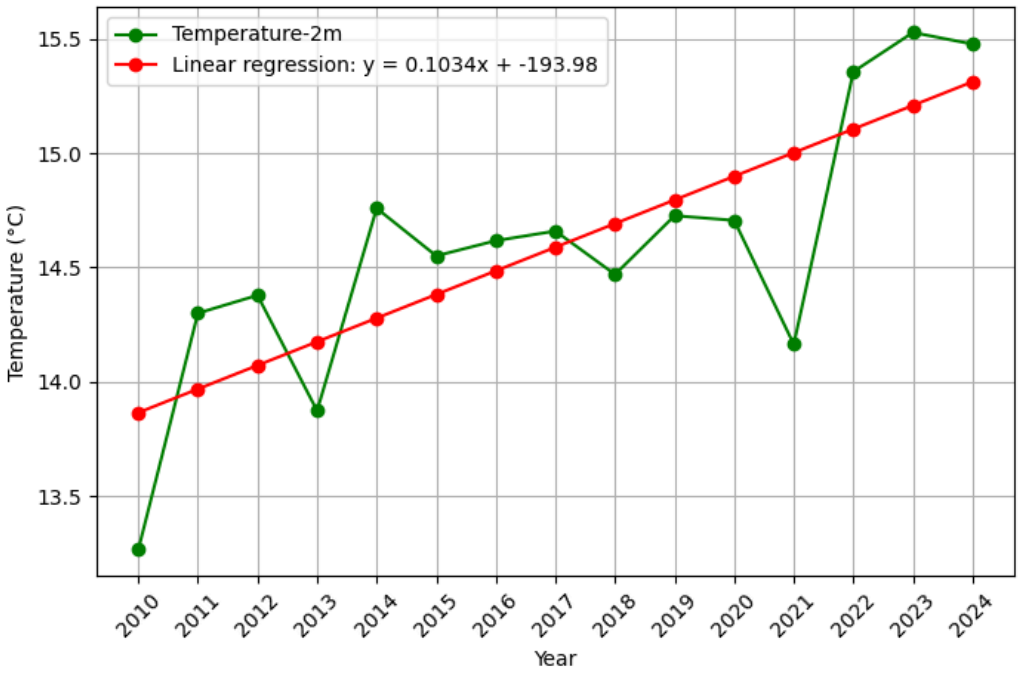


Figure 19: Temperature-2m trend in Castellón Province (2010-2024)

The 2m temperature trend indicates a clear warming pattern, with an increase of about 0.1034°C per year shown in regression analysis. The temperature has risen from 13.26°C in 2010 to 15.47°C in 2024, showing the long-term rise in temperature. (Miller et al., 2023) stated that the climate of Spain has been experiencing warming trends year by years. It is observed that from 2016, the temperatures have increased significantly, with 2022 and 2023 having the highest recorded mean temperature, close to 15.5°C. Especially in the semiarid parts of eastern Castellón, where the increase in temperature, together with the decline in rainfall, may lead to lower NDVI and higher levels of land degradation.

The semi-arid region in Castellon has a hot semi-arid climate with mild winters and hot, dry summers is well documented in Wikipedia (Wikipedia, 2025). Similar trends of precipitation and temperature were found in Serbia, showing almost similar trends for both climatic variables (Baumgertel et al., 2024). Seasonal changes also manifest in evident warming-up during the summer months, influencing vegetation stress. This coincides with the rise in wildfire risk, especially within regions of woodland heathlands, such as Serra d'Espadà and Penyagolosa, which have seen recorded cases of fire throughout history.

4.5.4 LST Trend

The following Figure 20 presents the temporal trend of Land Surface Temperature (LST) in Castellón Province between 2020 and 2024. The green lime represents yearly fluctuations, interannual variability. The red line indicates the linear regression line highlights an overall increasing trend in LST.

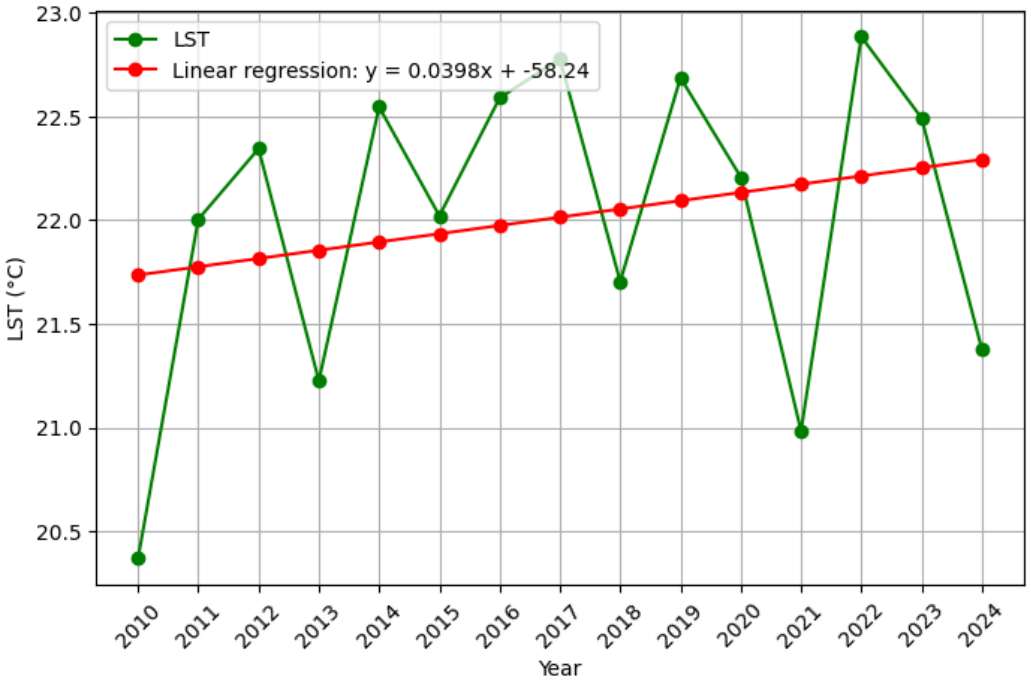


Figure 20: Land Surface Temperature (LST) trend in Castellón Province (2010-2024)

First, the LST trend aligns closely to air temperature, showing a progressive increase. And according to a regression model showing an increase in value by approximately 0.0398°C every year within this period and represents the strong effect of an intensified land surface heating. The years with extreme heat events, especially in 2017 (22.77°C), in 2022 (22.88°C), and in 2023(22.49°C), which eventually point toward vegetation stress. The 2022 LST value (21.37°C) is notably high, indicating that the land is more prone to degradation and the vegetation less resilient. This holistic NDVI trend suggests a declining vegetation cover, which may be strongly correlated with increased temperature and decreased precipitation. Also, heat accumulation due to higher LST in the province can also have an effect in reduced vegetation.

Here, the section provides a very important basis for the following section on the correlation analysis of NDVI and climatic factors, where a qualitative assessment can be made of these relationships.

4.6 Correlation Analysis of Vegetation Coverage and Climate Factors

This study further develops an analysis of spatial correlations between NDVI (vegetation cover) and three key climatic factors using Pearson's correlation: Precipitation, Temperature (2m), and Land Surface Temperature (LST) for 2024. We have computed yearly Pearson correlations and the correlation maps generated from 12-months average for the year 2010, 2015, 2020, and 2024 (the derived results are attached in Appendix Table A2 to Table A6); however, we have presented here only the result for the year 2024 to discuss in detail in this section. Based on the results Pearson correlation is mapped throughout the province of

Castellón and then analysed by relevant coefficient (r) values and relevant statistical significance (p) values.

4.6.1 NDVI vs Precipitation

The relationship between vegetation and precipitation is a key factor in understanding ecological stability. Figure 21 illustrates the spatial correlation between NDVI and Precipitation across Castellón Province.

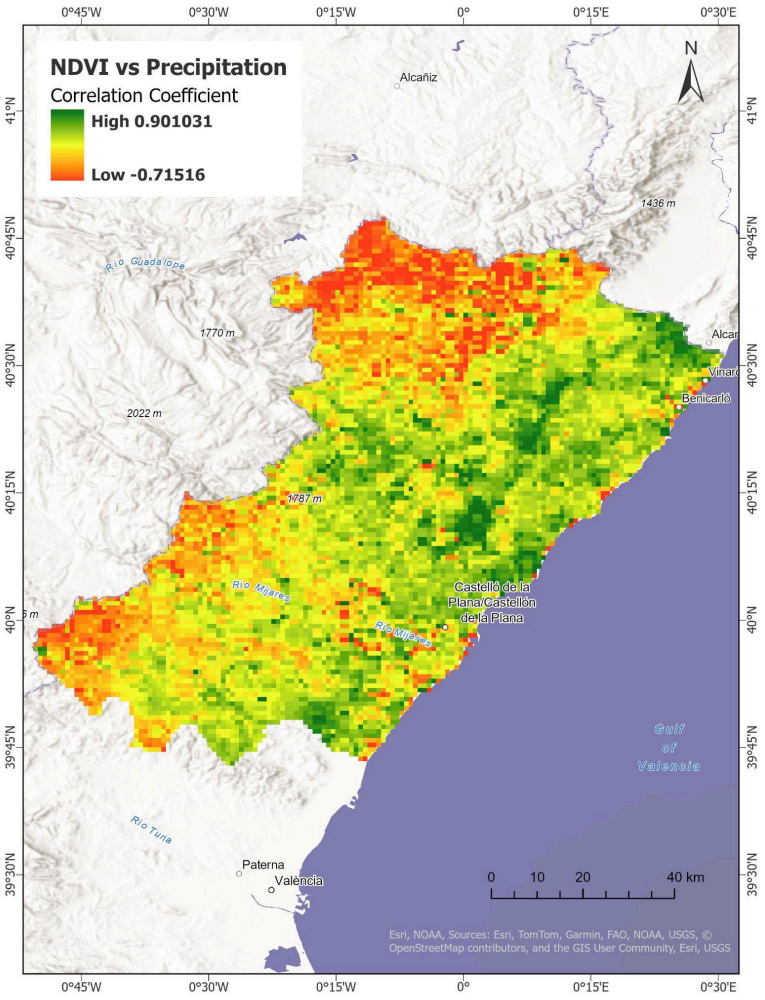


Figure 21: NDVI-Precipitation correlation map for Castellón Province (2024)

The NDVI-Precipitation correlation map shows a significant strong positive relationship in most of the regions that fall within values from -0.715 to 0.901. Average correlation was measured as 0.3 (Moderate Positive) with standard deviation of 0.18. And statistical significance (showed p values <0.05) was significant in many areas.

Regions of high Positive Correlation (r > 0.7) includes the northwest highlands and the central forested areas including Sierra de Espadán Natural Park, Upper Palancia, and Penyagolosa Natural Parks. All these areas are highly dependent upon rainfall, meaning

increased vegetation with increased rainfall. Moderate Positive Correlation ($0.3 < r \leq 0.7$) includes most of the province and shows this correlation, meaning that, although precipitation is a contributing factor to vegetation growth, other factors such as soil conditions and land use may also affect NDVI.

Negative Correlation in Coastal and Urban Areas ($r < -0.3$) includes areas like: Castellón de la Plana, Benicarló, and Vinaròs in the east show negative or weak correlations, probably due to irrigated agriculture and urban infrastructure reducing dependence on precipitation. The analysis showed strong p-value significance where 87.16% of the pixels are positively correlated; among them, 39.34% showed statistical significance, highlighting that precipitation has a major control on vegetation activity. A study revealed the direct influence of precipitation on crop productivity which is a key indicator of vegetation activity (Pitarch, 2024). Additionally, one study gave scientific evidence to the general positive correlation between NDVI and precipitation, showing that increased rainfall is usually accompanied by increased vegetation productivity measured by NDVI (Wang et al., 2003).

4.6.2 NDVI vs Temperature-2m

Temperature variations play a crucial role in influencing vegetation health and growth patterns. Figure 22 presents the spatial Pearson correlation between NDVI and 2m air temperature. The NDVI-Temperature (2m) correlation map reveals a relationship with values ranging from -0.967 to 0.849 among them mostly showing a negative relationship.

For this relationship average correlation was -0.639 (Moderate Negative) with standard deviation of 0.226 and this showed strong statistical significance (p-value < 0.05) which was found in most regions. Strong Negative Correlation ($r < -0.7$) was found in lowland agricultural and urban areas including areas like-Vila-real, Almassora, and Segorbe. This indicates that higher temperatures negatively impact vegetation, likely due to heat stress and increased evapotranspiration. Moderate Negative Correlation ($-0.7 \leq r < -0.3$) was seen in large portions of the province, highlighting how higher temperatures decrease NDVI overall, while the resilience shows for some other areas. Positive Correlations ($r > 0.3$) was evident in the regions like- the plateau part of Peñagolosa Natural Park, Sierra de Espadán.

70.1% of the province had a negative correlation, from which 20.22% was statistically significant. It therefore supports findings from a study where they found that temperature significantly contributes to the productivity of vegetation and negatively impairs areas that are hot enough with some studies (Cowles et al., 2018).

Another study proved the positive relationship between NDVI vs Precipitation and Strong negative correlation between NDVI vs temperature by using Pearson Correlation Coefficient which completely aligns with our study (Measho et al., 2019). The following Figure 22 presents the spatial Pearson correlation between NDVI and 2m air temperature.

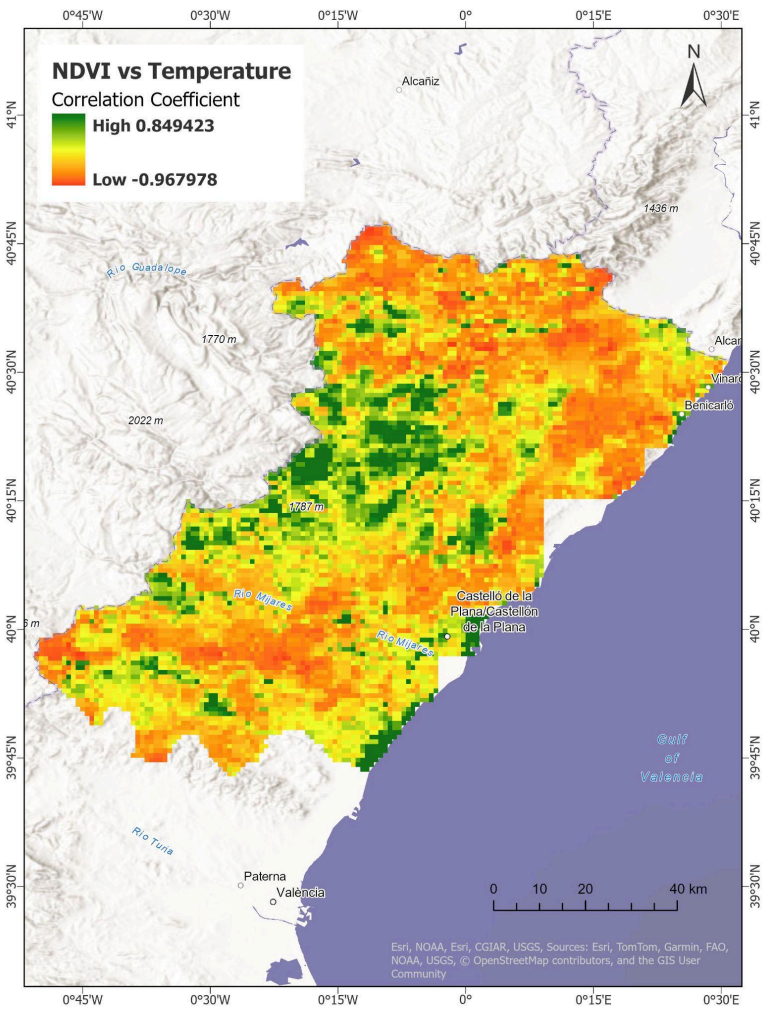


Figure 22: NDVI-Temperature-2m correlation map for Castellón Province (2024)

4.6.3 NDVI vs LST

Land Surface Temperature (LST) significantly influences vegetation productivity by affecting evapotranspiration rates and soil moisture availability. Figure 23 illustrates the spatial correlation between NDVI and LST across Castellón Province in 2024. In the NDVI-LST correlation map this study noticed a very strong negative relationship. The value of correlation ranges from -0.995 to 0.763. In the province, the average correlation was -0.789 (Strong Negative). The results showed strong statistical significance with p-value < 0.05 which was significant for most of the study area.

Very Strong Negative ($r < -0.9$) was found in large urbanized or low-elevation areas, in Castellón de la Plana, in Vila-real, in Benicàssim, or in coastal ones, which highlights that the higher land surface temperatures reduce vegetation fitness, stress the plants, and lower NDVI. Moderate to Strong Negative ($-0.7 \leq r < -0.3$) was in most regions of the study area, therefore reinforcing that high LST contributes negatively to the vegetation. The following Figure 23 illustrates the spatial correlation between NDVI and LST across Castellón Province in 2024.

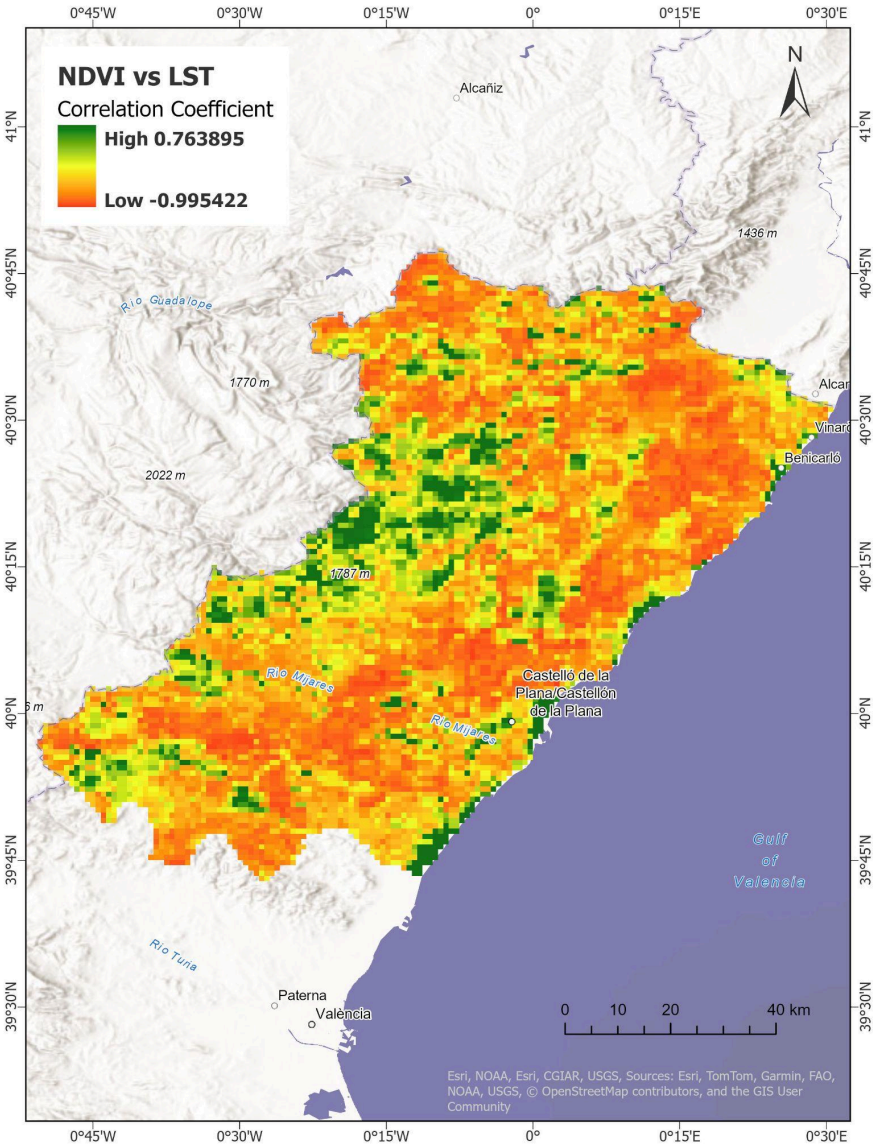


Figure 23: NDVI-LST correlation map for Castellón Province (2024)

The positively correlated areas include the highlands of Penyagolosa and some areas in Sierra de Espadán fall under the category >0.3 indicating weak to moderate correlation and implying that temperature is not a limiting factor on vegetation growth in these areas. During the study most of the negative correlations showed a statistical significance which is an implication of strong land surface temperature control on vegetation activity which align with some prior studies (Oloyede et al., 2021; S. A. Shah, 2022; Song et al., 2018)**.**

So overall, during spatial analysis on a yearly scale-precipitation is the strongest positive driver of NDVI (87.16% positive correlation, 39.34% statistically significant) with high dependence in forested and inland areas. Temperature (2m) has mixed effects on NDVI, with negative correlations in hotter urban areas and positive correlations in cooler highlands. LST has the most substantial negative influence on vegetation, particularly in coastal and urban zones, demonstrating the effects of urban heat islands and increased evapotranspiration.

4.7 NDVI's Sensitivity to Climatic Variables Over Time

We move towards a temporal correlation analysis instead of spatial focus to have a more comprehensive outlook of how the relationships evolved over time. This section aims to analyse the temporal variations of NDVI's relationship with precipitation, temperature, and LST over the period 2010–2024, using monthly aggregated data. Almost all the values meet the $p < 0.05$ threshold of significance, ensuring reliability.

4.7.1 NDVI vs Precipitation

Understanding the relationship between NDVI and precipitation is crucial for estimating vegetation response to variations in rainfall. The results show fluctuating correlations with most of the values varying between 0.1 and 0.5, showing a moderate positive association. Figure 24 presents the mean monthly NDVI average over all province and monthly summed total precipitation correlation coefficients during the 2010-2024 period.

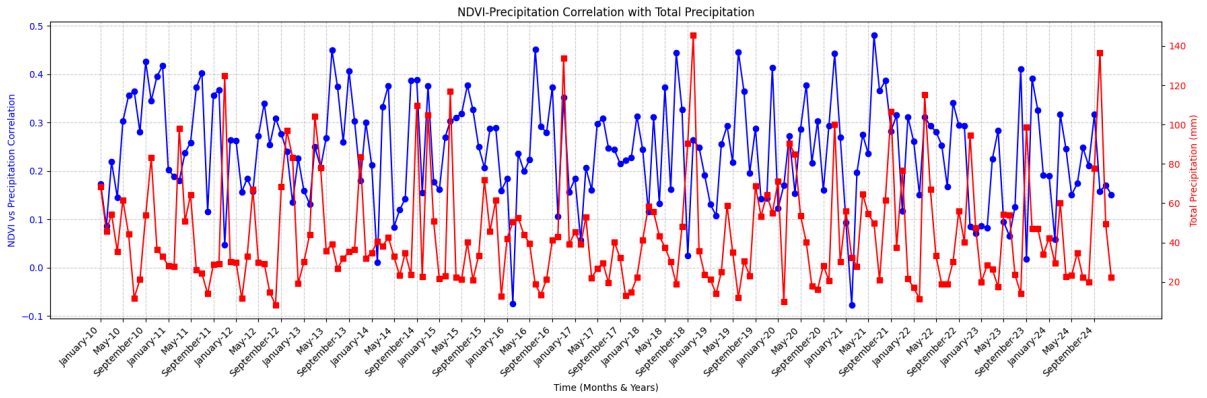


Figure 24: Monthly correlation between NDVI and Precipitation with Precipitation trends (2010-2024)

The Pearson correlation coefficients between NDVI and precipitation are fluctuating during the period under study of 2010-2024. Most values fall between 0.1 and 0.5, though some are negative for a few months. Notably, strong positive correlations (>0.4) appear in early growing seasons (spring months), while lower correlations occur in late summer and winter. The NDVI-Precipitation relationship is strongest during March–May, reflecting vegetation dependence on spring rainfall for growth. Certain years (2013, 2016, and 2018) exhibit high correlation (0.4–0.5), while others (2022, 2024) show weaker relationships (0.1–0.2), suggesting changing precipitation dependence. While NDVI showed moderate to strong precipitation correlation in earlier years, this trend has weakened in recent years (2020–2024), indicating a shift in vegetation reliance on precipitation.

The reason can be-If precipitation occurs in short bursts (extreme events) rather than steady rain, NDVI might not fully benefit from it. So, overall NDVI on a multi-annual scale show: a seasonal dependence on precipitation, particularly in the early growing season, but this

relationship has weakened over time. So, precipitation may not be the driving factor here in this scenario. There may be some other driving factors.

4.7.2 NDVI vs Temperature-2m

The relationship between Temperature-2m and NDVI is predominantly negative, with correlation values ranging from -0.3 to -0.5. This indicates that the increase in temperature predominantly results in the decline of vegetation health. This would suggest a general decrease in NDVI with an increased temperature. The variations in NDVI-Temperature correlation trends over the study period are shown in Figure 25. It represents the mean monthly NDVI average over the province and monthly mean Temperature of the province.

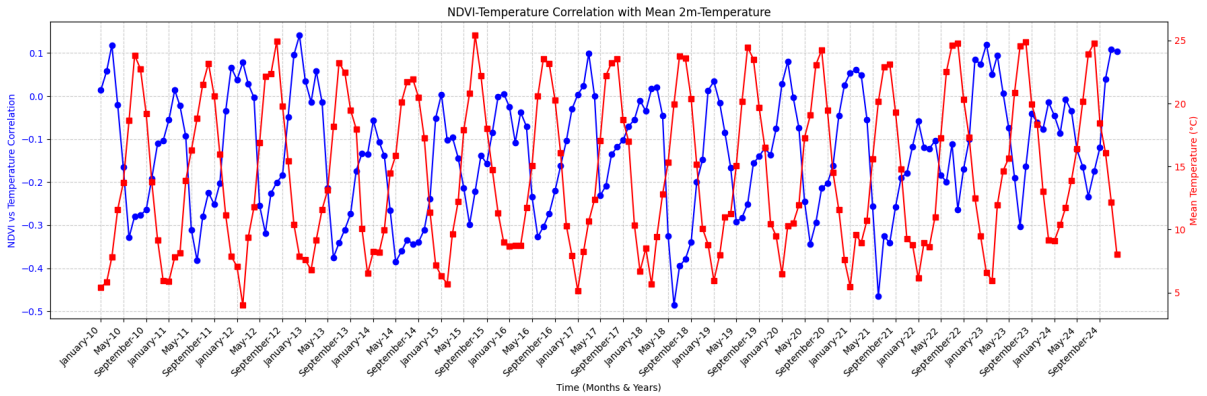


Figure 25: Monthly correlation between NDVI and Temperature-2m with Temperature trends (2010-2024)

The overall relationship between NDVI and temperature is inversely related; thus, higher the temperature, more the vegetation stress. Strongest negative correlations (- 0.5) occur during summer months (June-August, with higher temperatures contributing to vegetation water loss and stress. Weaker correlations (- 0.2 to -0.3) occur during the winter months, suggesting that colder temperatures are less limiting for vegetation.

Recent trends (2020-2024) indicate a little weaker negative correlation, suggesting vegetation may slowly be adapting to rising temperatures. High summer temperatures induce vegetation stress, hence the negative NDVI values due to increased evapotranspiration rates.

In winter, temperature is less limiting since the plant activity is already reduced, which explains the weaker and complex correlation. Some studies showed winter cannot affect plant activity (Yamori et al., 2014) and some suggested that winter can significantly impact plant activity (Larran et al., 2023). Recent weakening of the negative correlation (2020-2024) may indicate a vegetation adaptation to the rise in temperatures. Overall, strong negative correlation of NDVI and Temperature indicates temperature-driven vegetation stress, more precisely during the summer months. Slight recent weakening of the negative correlation indicates vegetation adaptation due to rising temperatures.

4.7.3 NDVI vs LST

The correlation between NDVI and LST exhibits a high negative correlation, and its values ranging between -0.4 and -0.8. This indicates that, in comparison to air temperature, vegetation health is significantly impacted negatively by LST. Figure 26 illustrates monthly trends in correlation of NDVI-LST throughout the whole observation period.It represents the mean monthly NDVI average over the whole province and monthly mean LST of the province.

.

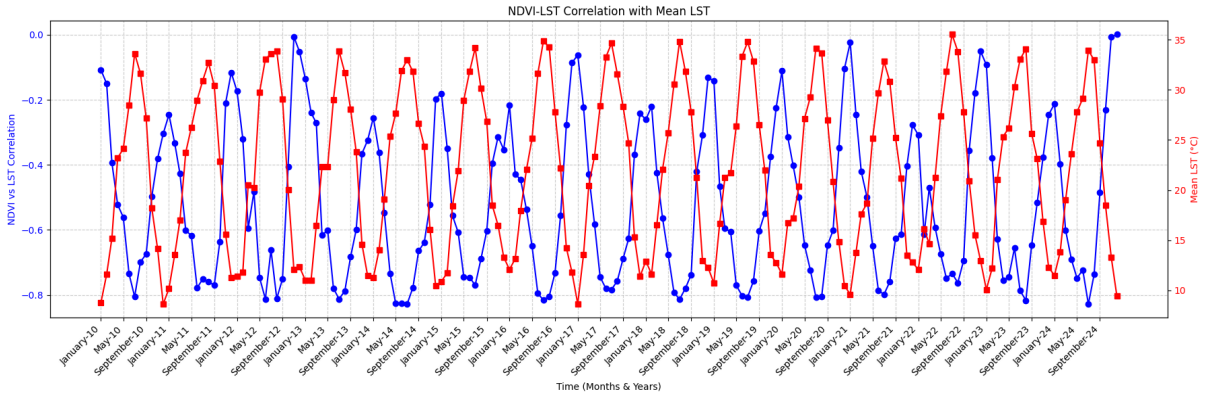


Figure 26: Monthly correlation between NDVI and Land Surface Temperature (LST) with LST trends (2010-2024)

Strongly negative correlations in the range between -0.6 and -0.8 occur during the peak summer month, thus confirming that vegetation activity is strongly inhibited by high surface temperature. Weaker correlations (-0.3 to -0.5) exist during transitional months where LST fluctuations moderate.

Annual fluctuations follow the seasonal course of surface temperature and, hence, further reinforce the influence of LST on NDVI. This result is supported by many research papers. One study analysed the relationship between NDVI and LST on a global scale. They note that annual oscillations in NDVI closely follow the seasonal patterns of surface temperature (Julien & Sobrino, 2009). So, it can be said that The LST impacts water availability in the soil, leading to evaporation and dryness of the soil; hence, reducing NDVI.

LST reflects conditions of heat waves, which cause severe vegetation stress, reduced photosynthesis, and thus lower NDVI. Recent trends do not show any marked weakening in NDVI-LST correlation, thus indicating that LST remains the dominant limiting factor for vegetation.

4.8 Association of Forest Cover Change and Climatic Variables

To better understand the interaction between forest cover change dynamics and climatic variables over time, a macro-level comparison analysis was conducted. This evaluation considers forest and non-forest class trends, along with temperature, precipitation, and land

surface temperature variations. Figure 27 illustrates the comparative trends for all of these variables during the study period (2010-2024).

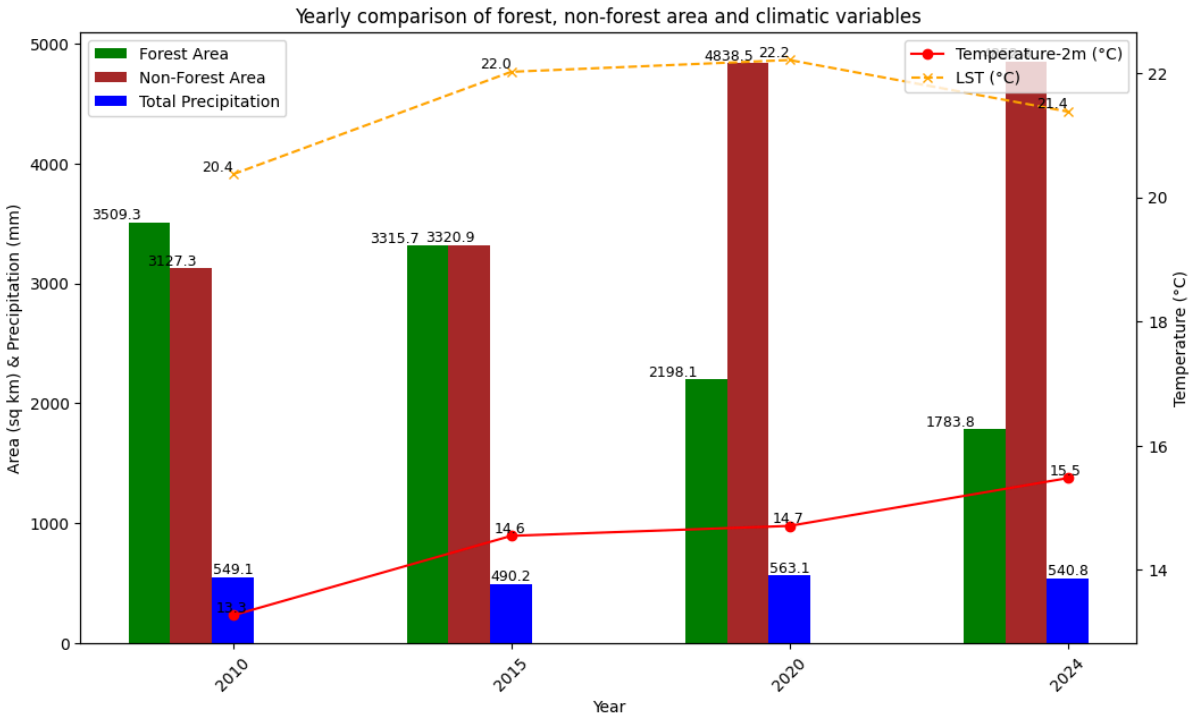


Figure 27: Comparative trends of forest and non-forest cover and climatic variables in Castellón Province (2010-2024)

Based on the results illustrated in the graph shows that Forest area shows a significant reduction over 14 years. Correspondingly non forest areas increased from 3127.28 sq km in 2010 to 4852 sq km in 2024.The mean 2m- temperature has steadily increased from 13°C to 15.5°C, a 2.21°C rise over 14 years. LST peaked in 2020 at 22.21°C, dropped slightly in 2024 to 21.83 °C. This could mean that LST values responded to forest cover loss, but at the same time, they could be influenced by other surface characteristics. Annual precipitation has shown fluctuation, basically decreasing from 549.12 mm in the year 2010 to 490.22 mm in 2015, rising to 563.09 mm in the year 2020, and slightly reducing again to 540.8 mm in 2024.

Overall, the temperature increase is aligned with the extreme reduction in forest cover. Loss of forest area leads to increased temperatures due to reduced evapotranspiration and increased heat absorption from the already exposed land surfaces. The reduction in forest area from 2010 to 2024 was 1725 km² which eventually caused a gradual rise in temperature. Therefore, it has a positive feedback loop wherein rising temperatures also promote forest stress and lead to more degradation.

LST changes reflect changes in the surface, and in Castellón, peak LST values for 2020 are coinciding with the largest loss in the area covered by forests. The minor decrease of LST in 2024, despite further losses in forest cover, would hint at other active factors in the temperature regulation, such as surface moisture availability. Unlike the temperature,

precipitation trends do not show any direct relationship with the loss of forest. Increased precipitation in 2015-2020 did not lead to the recovery of forests and showed that other factors such as land-use change, fire, or human interference were much stronger. However, long-term regional hydrologic cycles could be influenced by forest losses.

This spatiotemporal analysis of changes in forest cover and climate dynamics within Castellón Province clearly underlines the fact that loss has been a dominant trend in forests over the last 14 years, having a strong linkage with rising temperature and LST. Though precipitation fluctuation does not directly correspond with the change in forest condition, land-use changes, and warming-up trends are found to be major causes of deforestation.

Chapter 5. Conclusions and Recommendation

5.1 Conclusion

This research examined forest cover change in Castellón Province by using satellite-based remote sensing and machine learning classifiers (RF, SVM, and CART) over a 14-year period. Then it focused on to find out forest cover change and investigate the relationship between remote sensing index NDVI and climatic variables such as Precipitation, Temperature and LST. NDVI highlighted the importance of climate variability in forest cover dynamics, indicating the necessity of sustainable forest management.

Spatiotemporal trends in this study combined Google Earth Engine (GEE), Python-based statistical analysis, and GIS mapping to quantify land cover transitions and climate-vegetation interactions successfully. Seasonal analysis further validated forest cover fluctuations due to climate change, with increased spectral variability and classification uncertainty during summer season.

The key findings derived from the result based on research question no 1:

- 10-fold-cross-validation showed RF maintaining stable average accuracy (0.92-0.96), indicating its reliability in classification compared to SVM and CART.
- Visual analysis of classified maps confirmed that RF effectively captured forest cover trends, accurately differentiating forest and non-forest area.
- Between 2010 and 2024, Castellón Province lost approximately 1725.49 sq km of forest.
- The highest forest loss occurred between 2015-2020. In this period 1117.60 sq km of forest was lost, marking a 33% reduction in forest area.
- In the period of 2020-2024, deforestation slowed with 343.04 sq km of new forest emerging.
- Forest expansion was mostly concentrated in protected areas like Penyagolosa and Serra d'Espadà Natural Park.
- Castellón de la Plana, Almassora and Burianna and the coastal regions were identified as major deforestation hotspots.
- Historical wildfire events, such as the Villanueva de viver fire in 2023.
- The proportion of stable forest decreased from 3827.83 sq km (2010-2015) to 2616.60 sq km (2002-2024) showed progressive deforestation over time. It points towards continuous degradation.
- Reforestation efforts were higher at 731.12 sq km in the period of 200-20215, which decreased to 105.95 sq km in the period of 2015-2020. It indicates reduced reforestation activities.

The key findings derived from the result based on research question no 2:

- Overall NDVI trend indicated vegetation decline. NDVI started at 0.511 in 2020 and dropped to 0.486 in 2024. It suggests increased stress in vegetation.

- Despite regional NDVI declines, Penyagolosa and Serra d'Espadà Natural Park maintained high NDV values.
- Annual precipitation showed interannual fluctuations but a general decreasing trend (-0.2653mm per year) was observed.
- Increase of temperature was evident. The 2m-Temperature increased from 13.26°C in 2010 to 15.47 °C in 2024.
- Land Surface Temperature increased by 0.0398°C per year which peaks recorded in 2020 and 2022. Overall LST exhibited a significant upward trend.
- A combination of increased both types of temperature and decreased precipitation contributed to declining NDVI and resulted in increased deforestation.
- NDVI exhibited positive correlation with precipitation. Higher rainfall (example 2018) aligns with NDVI recovery, while dry years with low precipitation (example 2017) corresponds to vegetation stress. NDVI-precipitation correlation was strongest in mountainous areas. That led to increased vegetation productivity in inland and mountainous regions.
- NDVI showed a negative correction with both Temperatures-2m. Almost all the years showed, higher temperatures resulted in lower NDVI values. Summer months exhibited the strongest negative correlation. Strongest negative correlation is observed consistently in urban and agricultural zones.
- NDVI had the strongest negative correlation with LST. Increased LST contributed to reduced vegetation productivity, especially in coastal and lowland regions. Overall, urban and coastal areas demonstrated the strongest inverse relationships, highlighting the role of urban heat islands.
- Forest cover loss corresponded with rising temperatures. Increased both types of temperature were associated with declining forest area.
- The highest LST values in 2020 aligned with the most extensive deforestation period (2015-2020).
- Precipitation fluctuation did not directly influence forest loss.
- Urbanization, expansion of built-up areas, and other land-uses shifts (natural or human-induced) etc. were significant contributors to forest reduction.

Overall deforestation remains an on-going challenge in this province. While reforestation showed minor improvement in recent years, overall deforestation is constantly increasing. Rising Temperature and urban expansion collectively contributed to forest degradation. Efforts like- reforestation, conservation initiatives and improved land-use planning are crucial in this stage to mitigate the deforestation and climate change impacts.

Finally, this research provided a comprehensive machine-learning-based assessment of forest cover change and its climatic influences in Castellón Province while also addressing a crucial research gap. The outcome of this study will contribute to support more informed decision making for adaptive measurements, conservation planning, and sustainable land management in Mediterranean forest ecosystems. Future research needs to expand datasets with high-resolution imagery, incorporate more climate variables, and explore socio-economic drivers of deforestation to improve predictive modelling capacities.

5.2 Limitation of the Study and Future Works

Limitation:

The study has some drawbacks despite its strength:

- The study used random sampling for forest and non-forest categorization locations without considering spectral histograms, which may lead to misclassifications specially in transitional vegetation zones.
- By using just binary forest/non-forest categories and without distinguishing between additional forest types like shrubland, mixed forests, or degraded forests, the categorization scheme reduced biological specificity according to the forest definition. Also forests are heterogeneous, so different types of forest are also not taken into consideration.
- Because of sensor limitations, one classed image had striping that could not be fixed in time.
- Also, minor radiometric differences could exist between Landsat 5 and 8 as they have different sensor calibrations.
- Cloud filtering only employed images with less than 10% cloud cover as the images were seen clearer; no other cloud masking techniques were used, which would have affected the categorization accuracy during very humid or heavily cloudy conditions.
- The correlation analysis may have geographic inconsistencies because the MODIS-derived NDVI (1 km resolution) was correlated with CHIRPS and ERA5 climate variables with different resolutions (5 km and 9 km), but while exporting average response was taken).
- Accuracy evaluations were based exclusively on remote sensing validation methods because field validation of the forest classification data was not carried out. Only assessed with statistical validation techniques and visual inspection.
- The combined correlation raster data for all years could not be exported due to the error related to earth engine memory limit, so individual data were analysed.
- Variations in seasonal vegetation phenology may have influenced classification stability, even though RF did well in both yearly and summer-specific classifications. As they have some discrepancies in the result of forest area.
- The forest cover transition matrix was based on RF classified images, so the accuracy is dependent on the classification performance. Misclassification could lead to over or under-estimation of forest-loss/gain in a specific area.

Recommendations for Future works:

- Incorporating expanded classification to different forest types and other classes based on their spectral histogram
- Conducting in-situ validation using GPS-referenced forest data to improve classification accuracy with the real world
- Future study can integrate sentinel-2 imagery with Landsat for higher quality spatial and temporal assessments

- For improving classification accuracy, exploring deep learning techniques such as convolutional neural Network (CNNs)
- The findings can be integrated into global forest-climate models, improving predictions of future forest responses to climate variability.

References

Ajayi, O. G., & Ojima, A. (2022). Performance evaluation of selected cloud occlusion removal algorithms on remote sensing imagery. *Remote Sensing Applications: Society and Environment*, *25*, 100700. https://doi.org/10.1016/j.rsase.2022.100700

Ayanlade, A., & Nwaezeigwe, J. O. (2018). *Effects of cloud cover variation on satellite data utilisation: A case dtudy of landsat dataset*.

Aziz, G., Minallah, N., Saeed, A., Frnda, J., & Khan, W. (2024). Remote sensing based forest cover classification using machine learning. *Scientific Reports*, *14*(1), 69. https://doi.org/10.1038/s41598-023-50863-1

Baniya, B., Tang, Q., Huang, Z., Sun, S., & Techato, K. (2018). Spatial and Temporal Variation of NDVI in Response to Climate Change and the Implication for Carbon Dynamics in Nepal. *Forests*, *9*(6), Article 6. https://doi.org/10.3390/f9060329

Baumgertel, A., Lukić, S., Caković, M., Lazić, I., Tošić, M., Momirović, N., Pandey, S., Bezdan, A., Blagojević, B., & Djurdjević, V. (2024). Spatio-Temporal Analysis of Vegetation Response to Climate Change, Case Study: Republic of Serbia. *International Journal of Environmental Research*, *18*(2), 21. https://doi.org/10.1007/s41742-024-00571-z

Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson Correlation Coefficient. In I. Cohen, Y. Huang, J. Chen, & J. Benesty, *Noise Reduction in Speech Processing* (Vol. 2, pp. 1–4). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-00296-0_5

Bindajam, A. A., Mallick, J., AlQadhi, S., Singh, C. K., & Hang, H. T. (2020). Impacts of vegetation and topography on land surface temperature variability over the semi-arid mountain cities of Saudi Arabia. *Atmosphere*, *11*(7), 762. https://doi.org/10.3390/atmos11070762

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4). Springer.

Boiarskii, B., & Hasegawa, H. (2019). Comparison of NDVI and NDRE indices to detect differences in vegetation and chlorophyll content. *J. Mech. Contin. Math. Sci*, *4*, 20–29. https://doi.org/10.26782/jmcms.spl.4/2019.11.00003

Breil, M., Schneider, V. K., & Pinto, J. G. (2024). The effect of forest cover changes on the regional climate conditions in Europe during the period 1986–2015. *Biogeosciences*, *21*(3), 811–824. https://doi.org/10.5194/bg-21-811-2024

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

Breiman, L. (2017). *Classification and regression trees*. Routledge. https://www.taylorfrancis.com/books/mono/10.1201/9781315139470/classification-regression-trees-leo-breiman

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, *2*(2), 121–167. https://doi.org/10.1023/A:1009715923555

Bush, E., & Lemmen, D. S. (2019). *Canada's changing climate report*. https://cir.nii.ac.jp/crid/1360584346968280448

Cheng, Z., Nakatsugawa, M., Hu, C., Robertson, S. P., Hui, X., Moore, J. A., Bowers, M. R., Kiess, A. P., Page, B. R., Burns, L., Muse, M., Choflet, A., Sakaue, K., Sugiyama, S., Utsunomiya, K., Wong, J. W., McNutt, T. R., & Quon, H. (2018). Evaluation of classification and regression tree (CART) model in weight loss prediction following head and neck cancer radiation therapy. *Advances in Radiation Oncology*, *3*(3), 346–355. https://doi.org/10.1016/j.adro.2017.11.006

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cowles, J., Boldgiv, B., Liancourt, P., Petraitis, P. S., & Casper, B. B. (2018). Effects of increased temperature on plant communities depend on landscape location and precipitation. *Ecology and Evolution*, *8*(11), 5267–5278. https://doi.org/10.1002/ece3.3995

D'Annunzio, R., Lindquist, E., & MacDicken, K. G. (2014). Global forest land-use change from 1990 to 2010: An update to a global remote sensing survey of forests. *Food and Agriculture Organization of the United Nations*.

De Castro, M., Martín-Vide, J., & Alonso, S. (2005). The climate of Spain: Past, present and scenarios for the 21st century. *A Preliminary General Assessment of the Impacts in Spain Due to the Effects of Climate Change. Madrid (Spain), Ministerio de Medio Ambiente*, 1–62.

De Luca, G., M. N. Silva, J., Di Fazio, S., & Modica, G. (2022). Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for land cover mapping in a Mediterranean region. *European Journal of Remote Sensing*, *55*(1), 52–70. https://doi.org/10.1080/22797254.2021.2018667

Delgado-Artés, R., Garófano-Gómez, V., Oliver-Villanueva, J.-V., & Rojas-Briales, E. (2022a). Land use/cover change analysis in the Mediterranean region: A regional case study of forest evolution in Castelló (Spain) over 50 years. *Land Use Policy*, *114*, 105967. https://doi.org/10.1016/j.landusepol.2021.105967

Delgado-Artés, R., Garófano-Gómez, V., Oliver-Villanueva, J.-V., & Rojas-Briales, E. (2022b). Land use/cover change analysis in the Mediterranean region: A regional case study of forest evolution in Castelló (Spain) over 50 years. *Land Use Policy*, *114*, 105967. https://doi.org/10.1016/j.landusepol.2021.105967

Díaz-García, P., & Regos, A. (2024). Assessing Land-Cover Changes in the Natural Park 'Fragas do Eume' over the Last 25 Years: Insights from Remote Sensing and Machine Learning. *Land*, *13*(10), Article 10. https://doi.org/10.3390/land13101601

European Comission. (2017, July 5). *Forest Fires in Castellón Province (2017-07-05)*. European Commission, Joint Research Centre (JRC). http://data.europa.eu/89h/c15446e7-96b7-49f5-9b61-cb919a4417fd

FAO. (2022). *Chapter 2 Forests and trees provide vital goods and ecosystem services but are undervalued in economic systems*. https://doi.org/10.4060/cb9360en

Farmonaut. (2024, October 7). *Revolutionizing Sustainable Forest Management: Harnessing Satellite Imagery for Advanced Monitoring and Conservation* -. https://farmonaut.com/remote-sensing/revolutionizing-sustainable-forest-management-harnessing-satellite-imagery-for-advanced-monitoring-and-conservation/

Flood, N. (2013). Seasonal composite Landsat TM/ETM+ images using the medoid (a

multi-dimensional median). *Remote Sensing*, *5*(12), 6481–6500. https://doi.org/10.3390/rs5126481

Gao, B.-C. (2009). *Atmospheric Correction Algorithms for Remote Sensing of Land and Water Surfaces*. https://doi.org/10.1016/j.rse.2007.12.015

Gök, D. T., Scherler, D., & Wulf, H. (2024). Land surface temperature trends derived from Landsat imagery in the Swiss Alps. *The Cryosphere*, *18*(11), 5259–5276. https://doi.org/10.5194/tc-18-5259-2024

Holmgren, P., & Marklund, L. G. (2007). National forest monitoring systems: Purposes, options and status. In *Forestry and climate change* (pp. 163–173). https://doi.org/10.1079/9781845932947.0163

Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing*, *23*(4), 725–749. https://doi.org/10.1080/01431160110040323

Jiang, L., Guli·Jiapaer, Bao, A., Guo, H., & Ndayisaba, F. (2017). Vegetation dynamics and responses to climate change and human activities in Central Asia. *Science of The Total Environment*, *599–600*, 967–980. https://doi.org/10.1016/j.scitotenv.2017.05.012

Jiménez-Olivencia, Y., Ibáñez-Jiménez, Á., Porcel-Rodríguez, L., & Zimmerer, K. (2021). Land use change dynamics in Euro-mediterranean mountain regions: Driving forces and consequences for the landscape. *Land Use Policy*, *109*, 105721. https://doi.org/10.1016/j.landusepol.2021.105721

Jog, S., & Dixit, M. (2016). Supervised classification of satellite images. *2016 Conference on Advances in Signal Processing (CASP)*, 93–98. https://doi.org/10.1109/CASP.2016.7746144

Jones, M. W., Kelley, D. I., Burton, C. A., Di Giuseppe, F., Barbosa, M. L. F., Brambleby, E., Hartley, A. J., Lombardi, A., Mataveli, G., McNorton, J. R., Spuler, F. R., Wessel, J. B., Abatzoglou, J. T., Anderson, L. O., Andela, N., Archibald, S., Armenteras, D., Burke, E., Carmenta, R., … Xanthopoulos, G. (2024). State of Wildfires 2023–2024. *Earth System Science Data*, *16*(8), 3601–3685. https://doi.org/10.5194/essd-16-3601-2024

Julien, Y., & Sobrino, J. A. (2009). The Yearly Land Cover Dynamics (YLCD) method: An analysis of global vegetation from NDVI and LST parameters. *Remote Sensing of Environment*, *113*(2), 329–334. https://doi.org/10.1016/j.rse.2008.09.016

Khudhur, M. H., Aziz, N. A., & Alwan, I. A. (2024). Comparative study of supervised classification methods of land cover mapping using remote sensing data: A case study in Al-Hawija district/Iraq. *AIP Conference Proceedings*, *3105*(1). https://doi.org/10.1063/5.0213746

Klusowski, J. (2020). Sparse Learning with CART. *Advances in Neural Information Processing Systems*, *33*, 11612–11622.

Larran, A. S., Pajoro, A., & Qüesta, J. I. (2023). Is winter coming? Impact of the changing climate on plant responses to cold temperature. *Plant, Cell & Environment*, *46*(11), 3175–3193. https://doi.org/10.1111/pce.14669

Li, C., Wang, J., Wang, L., Hu, L., & Gong, P. (2014). Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat

Thematic Mapper Imagery. *Remote Sensing*, *6*(2), Article 2. https://doi.org/10.3390/rs6020964

Liu, J., Hagan, D. F. T., & Liu, Y. (2020). Global land surface temperature change (2003–2017) and its relationship with climate drivers: AIRS, MODIS, and ERA5-land based analysis. *Remote Sensing*, *13*(1), 44. https://doi.org/10.3390/rs13010044

Liu, J., Heiskanen, J., Aynekulu, E., & Pellikka, P. K. E. (2015). Seasonal variation of land cover classification accuracy of Landsat 8 images in Burkina Faso. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-7-W3*, 455–460. 36th International Symposium on Remote Sensing of Environment (Volume XL-7/W3) - 11–15 May 2015, Berlin, Germany. https://doi.org/10.5194/isprsarchives-XL-7-W3-455-2015

Liu, Q., Yang, Z., Han, F., Wang, Z., & Wang, C. (2016). NDVI-based vegetation dynamics and their response to recent climate change: A case study in the Tianshan Mountains, China. *Environmental Earth Sciences*, *75*(16), 1189. https://doi.org/10.1007/s12665-016-5987-5

Mallinis, G., Gitas, I. Z., Tasionas, G., & Maris, F. (2016). Multitemporal monitoring of land degradation risk Due to soil loss in a fire-prone Mediterranean landscape using multi-decadal Landsat imagery. *Water Resources Management*, *30*, 1255–1269.

Martínez-Vega, J., Mili, S., & Gallardo, M. (2021). Modelling Land Use and Land Cover Changes in the Mediterranean Agricultural Ecosystems. In *Modeling for Sustainable Management in Agriculture, Food and the Environment*. CRC Press.

Maxwell, A. E., Warner, T. A., & Guillén, L. A. (2021). Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—Part 1: Literature review. *Remote Sensing*, *13*(13), 2450. https://doi.org/10.3390/rs13132450

Measho, S., Chen, B., Trisurat, Y., Pellikka, P., Guo, L., Arunyawat, S., Tuankrua, V., Ogbazghi, W., & Yemane, T. (2019). Spatio-Temporal Analysis of Vegetation Dynamics as a Response to Climate Variability and Drought Patterns in the Semiarid Region, Eritrea. *Remote Sensing*, *11*(6), Article 6. https://doi.org/10.3390/rs11060724

Mihăilă, D., Bistricean, P.-I., Sfîcă, L., Horodnic, V.-D., Prisăcariu, A., & Amihăesei, V.-A. (2024). Summer Discrepancies between 2 m Air Temperature and Landsat LST in Suceava City, Northeastern Romania. *Remote Sensing*, *16*(16), 2967. https://doi.org/10.3390/rs16162967

Millard, K., & Richardson, M. (2015). On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote Sensing*, *7*(7), 8489–8515. https://doi.org/10.3390/rs70708489

Miller, K. A., Santillo, D., & Johnston, P. (2023). *Climate change and extreme weather events in Spain, Greenpeace*.

Oloyede, A. O., Olatunbosun, D. E., Asuquo, P. M., Udo, U. E., & Essien, I. O. (2021). Correlation Analysis of Vegetation and Land Surface Temperature in Uyo, Nigeria Using Satellite Remote Sensing and Python-Based Geographic Information System. *Science and Technology Publishing*, *5*(2632–1017).

Oo, T. K., Arunrat, N., Sereenonchai, S., Ussawarujikulchai, A., Chareonwong, U., & Nutmagul, W. (2022). Comparing four machine learning algorithms for land cover classification in gold mining: A case study of Kyaukpahto Gold Mine, Northern

Myanmar. *Sustainability*, *14*(17), 10754.
Ottolini, I., Salesa, D., del Romero Renau, L., & Fernández, N. S. (2024). *Kindling change: Shaping a New Fire Culture in Mediterranean socioenvironmental systems from the roots*. *18*.
Pandya, P. A., & Gontia, N. K. (2025). Correlation vs. Kappa statistic: A new perspective on comparing meteorological drought indices. *Theoretical and Applied Climatology*, *156*(1), 19.
Peris-Llopis, M., Mola-Yudego, B., Berninger, F., Garcia-Gonzalo, J., & González-Olabarria, J. R. (2024). Impact of species composition on fire-induced stand damage in Spanish forests. *Scientific Reports*, *14*(1), 8594.
Pitarch, J. (2024, November 14). *Torrential rains in the north of Castellón have had an impact on the citrus and vegetable harvest*.
Praticò, S., Solano, F., Di Fazio, S., & Modica, G. (2021). Machine learning classification of mediterranean forest habitats in google earth engine based on seasonal sentinel-2 time-series and input image composition optimisation. *Remote Sensing*, *13*(4), 586. https://doi.org/10.3390/rs13040586
Psistaki, K., Tsantopoulos, G., & Paschalidou, A. K. (2024). An Overview of the Role of Forests in Climate Change Mitigation. *Sustainability*, *16*(14), Article 14. https://doi.org/10.3390/su16146089
Qiu, S., Zhu, Z., Olofsson, P., Woodcock, C. E., & Jin, S. (2023). Evaluation of Landsat image compositing algorithms. *Remote Sensing of Environment*, *285*, 113375. https://doi.org/10.1016/j.rse.2022.113375
Raman, R., Manalil, S., Dénes, D. L., & Nedungadi, P. (2024). The role of forestry sciences in combating climate change and advancing sustainable development goals. *Frontiers in Forests and Global Change*, *7*. https://doi.org/10.3389/ffgc.2024.1409667
Rani, N. R. A., & Inayathulla, M. (2024). Machine Learning Algorithms for Classifying Land Use and Land Cover. In *ResearchGate*. https://doi.org/10.1007/978-981-99-9610-0_20
Reddy, C. S., Jha, C. S., & Dadhwal, V. K. (2013). Assessment and monitoring of long-term forest cover changes in Odisha, India using remote sensing and GIS. *Environmental Monitoring and Assessment*, *185*(5), 4399–4415. https://doi.org/10.1007/s10661-012-2877-5
Redeia. (2021). *Red Eléctrica finalises the reforestation of 32.4 hectares of woodland in Altura (Castellón)*. Red Eléctrica.
Rhew, I. C., Vander Stoep, A., Kearney, A., Smith, N. L., & Dunbar, M. D. (2011). Validation of the Normalized Difference Vegetation Index as a Measure of Neighborhood Greenness. *Annals of Epidemiology*, *21*(12), 946–952. https://doi.org/10.1016/j.annepidem.2011.09.001
Rudrapal, D., & Subhedar, M. (2015). Land cover classification using support vector machine. *Int. J. Eng. Res*, *4*(09), 584–588.
Runfola, D. S. M., & Pontius, R. G. (2013). Measuring the temporal instability of land change using the Flow matrix. *International Journal of Geographical Information Science*, *27*(9), 1696–1716. https://doi.org/10.1080/13658816.2013.792344
Ruslan, R. A., & Arbaiy, N. (2024). Handling Imbalanced Datasets in Machine Learning:

Challenges, Approaches, and Best Practices. *Journal of Applied Science, Technology and Computing*, *2*(1), 20–27.

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

Sage, A. (2018). *Random forest robustness, variable importance, and tree aggregation*.

Santoro, A., Piras, F., Fiore, B., Bazzurro, A., & Agnoletti, M. (2024). Forest-Cover Changes in European Natura 2000 Sites in the Period 2012–2018. *Forests*, *15*(2), Article 2. https://doi.org/10.3390/f15020232

Savastru, D. M., Zoran, M. A., & Savastru, R. S. (2019). Satellite remote sensing detection of forest vegetation land cover changes and their potential drivers. *Remote Sensing for Agriculture, Ecosystems, and Hydrology XXI*, *11149*, 540–550. https://doi.org/10.1117/12.2532882

Shah, I. A., Khan, H., Muhammad, Z., Ullah, R., Iqbal, S., Nafidi, H.-A., Bourhia, M., & Salamatullah, A. M. (2024). Evaluation of climate change impact on plants and hydrology. *Frontiers in Environmental Science*, *12*, 1328808. https://doi.org/10.3389/fenvs.2024.1328808

Shah, S. A. (2022). Statistical Analysis of Land Surface Temperature and Normalized Difference Vegetation Index Relation-ship Based on Remote Sensing. *Journal of Agriculture and Aquaculture*, *4*(3).

Silveira, E. M., Bueno, I. T., Acerbi-Junior, F. W., Mello, J. M., Scolforo, J. R. S., & Wulder, M. A. (2018). Using spatial features to reduce the impact of seasonality for detecting tropical forest changes from Landsat time series. *Remote Sensing*, *10*(6), 808. https://doi.org/10.3390/rs10060808

Sobrino, J. A., Gimeno, S., Crisafulli, V., & Sobrino-Gómez, Á. (2024). Analysing Land Cover Change in the Valencian Community through Landsat Imagery: From 1984 to 2022. *Land*, *13*(7), 1072. https://doi.org/10.3390/land13071072

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Song, Z., Li, R., Qiu, R., Liu, S., Tan, C., Li, Q., Ge, W., Han, X., Tang, X., & Shi, W. (2018). Global land surface temperature influenced by vegetation cover and PM2. 5 from 2001 to 2016. *Remote Sensing*, *10*(12), 2034.

Tang, W., Zhou, J., Ma, J., Wang, Z., Ding, L., Zhang, X., & Zhang, X. (2024). TRIMS LST: A daily 1 km all-weather land surface temperature dataset for China's landmass and surrounding areas (2000–2022). *Earth System Science Data*, *16*(1), 387–419. https://doi.org/10.5194/essd-16-387-2024

Tran, D. X., Pla, F., Latorre-Carmona, P., Myint, S. W., Caetano, M., & Kieu, H. V. (2017). Characterizing the relationship between land use land cover change and land surface temperature. *ISPRS Journal of Photogrammetry and Remote Sensing*, *124*, 119–132. https://doi.org/10.1016/j.isprsjprs.2017.01.001

Vapnik, V. (1998). *Statistical Learning Theory now plays a more active role: After the general analysis of learning processes, the research in the area of synthesis of optimal*

algorithms was started. These studies, however, do not belong to history yet. They are a subject of today's research activities. https://www.math.pku.edu.cn/teachers/ganr/course/pr/Ref/scholkopf99_svm_intro.pdf

Wang, J., Rich, P. M., & Price, K. P. (2003). Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA. *International Journal of Remote Sensing*, *24*(11), 2345–2364. https://doi.org/10.1080/01431160210154812

Wikipedia. (2025). *Castellón de la Plana—Wikipedia*. https://en.wikipedia.org/wiki/Castell%C3%B3n_de_la_Plana

Xiao, H., Liu, J., He, G., Zhang, X., Wang, H., Long, T., Zhang, Z., Wang, W., Yin, R., Guo, Y., Cheng, B., & Cao, Q. (2022). Data-Driven Forest Cover Change and Its Driving Factors Analysis in Africa. *Frontiers in Environmental Science*, *9*. https://doi.org/10.3389/fenvs.2021.780069

Yamori, W., Hikosaka, K., & Way, D. A. (2014). Temperature response of photosynthesis in C3, C4, and CAM plants: Temperature acclimation and temperature adaptation. *Photosynthesis Research*, *119*(1), 101–117. https://doi.org/10.1007/s11120-013-9874-6

Yan, Y., Mao, K., Shi, J., Piao, S., Shen, X., Dozier, J., Liu, Y., Ren, H., & Bao, Q. (2020). Driving forces of land surface temperature anomalous changes in North America in 2002–2018. *Scientific Reports*, *10*(1), 6931.

Zang, J., Qiu, F., Zhang, Y., Shang, R., & Liang, Y. (2025). A dataset of forest regrowth in globally key deforestation regions. *Scientific Data*, *12*(1), 154.

Zhang, X., Flato, G., Kirchmeier-Young, M., Vincent, L., Wan, H., Wang, X., Rong, R., Fyfe, J., Li, G., & Kharin, V. V. (2019). Changes in temperature and precipitation across Canada. *Canada's Changing Climate Report*, 112–193.

Zhong, L., Ma, Y., Salama, Mhd. S., & Su, Z. (2010). Assessment of vegetation dynamics and their response to variations in precipitation and temperature in the Tibetan Plateau. *Climatic Change*, *103*(3), 519–535. https://doi.org/10.1007/s10584-009-9787-8

Zhou, H., Deng, Z., Xia, Y., & Fu, M. (2016). A new sampling method in particle filter based on Pearson correlation coefficient. *Neurocomputing*, *216*, 208–215. https://doi.org/10.1016/j.neucom.2016.07.036

A Appendix

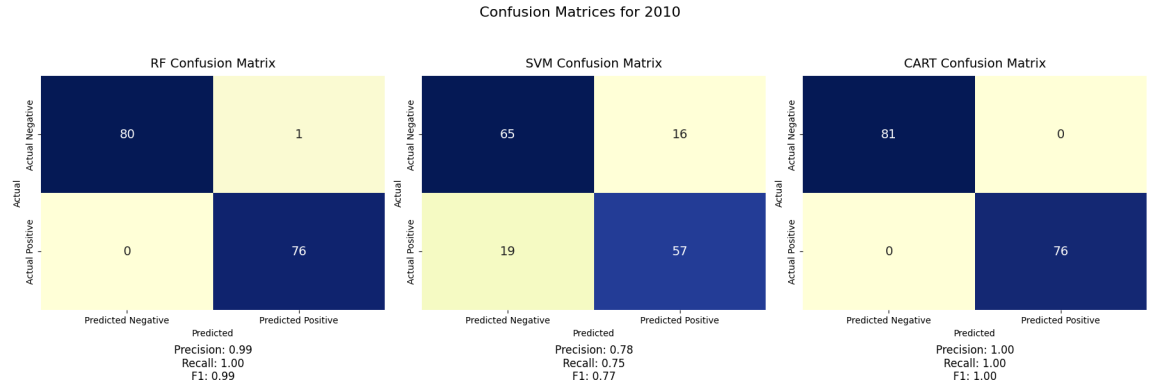


Figure A.1: Confusion metrics for RF, SVM, CART across 2010 (before cross-validation)

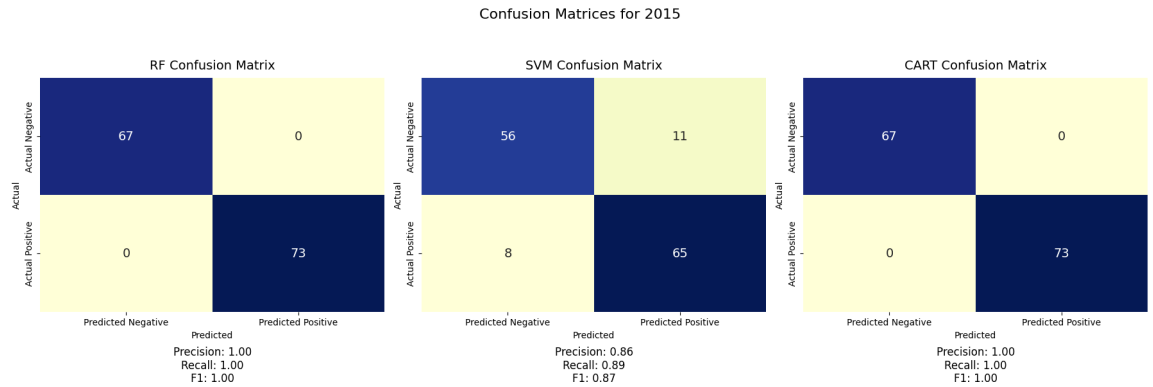


Figure A.2: Confusion metrics for RF, SVM, CART across 2015 (before cross-validation)

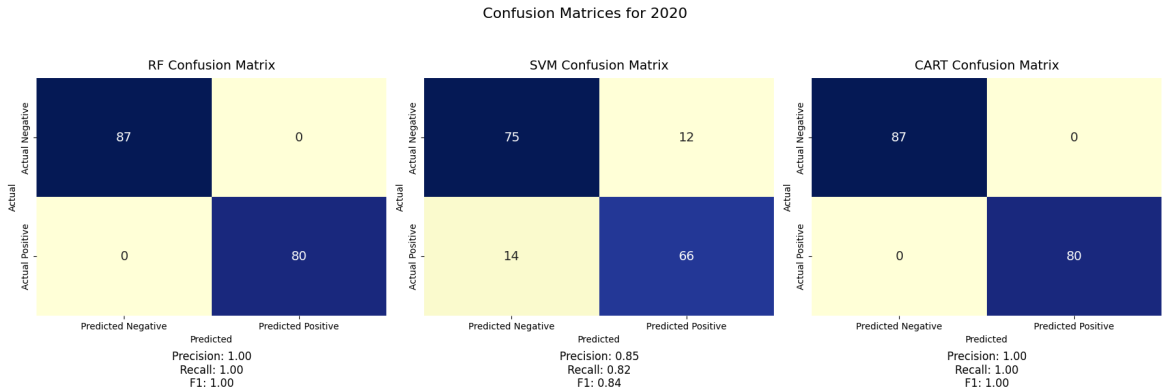


Figure A.3: Confusion metrics for RF, SVM, CART across 2020 (before cross-validation)

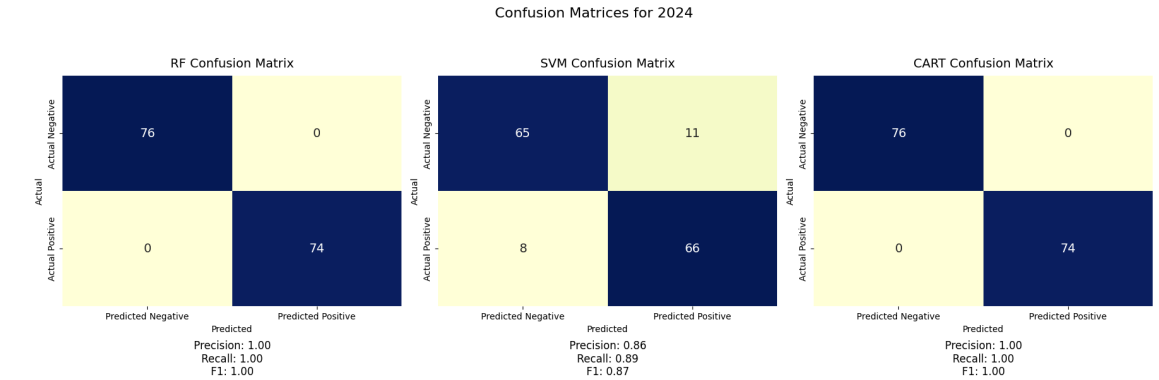


Figure A.4: Confusion metrics for RF, SVM, CART across 2024 (before cross-validation)

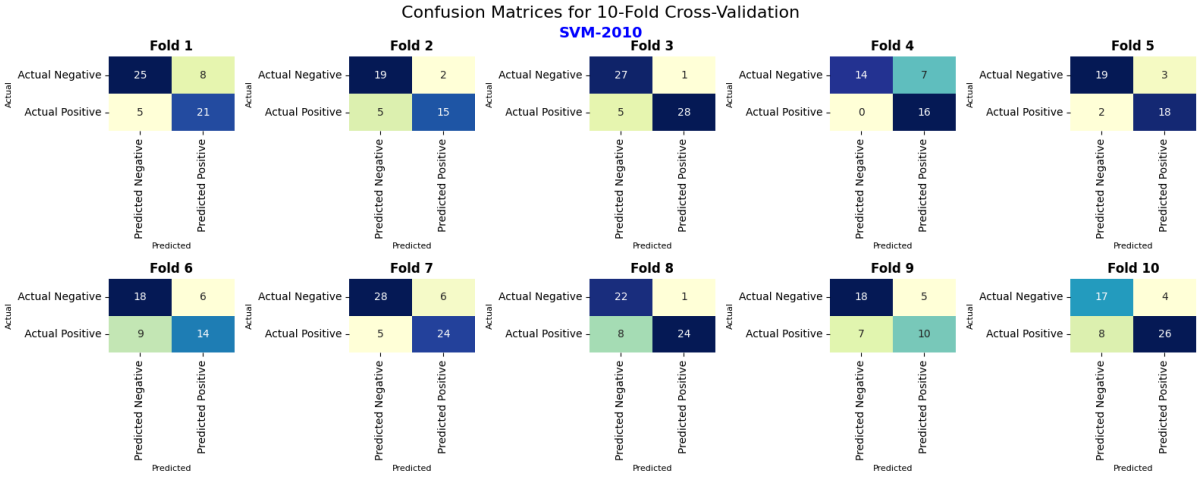


Figure A.5: Confusion metrics for 10-fold-cross-validation for SVM 2010

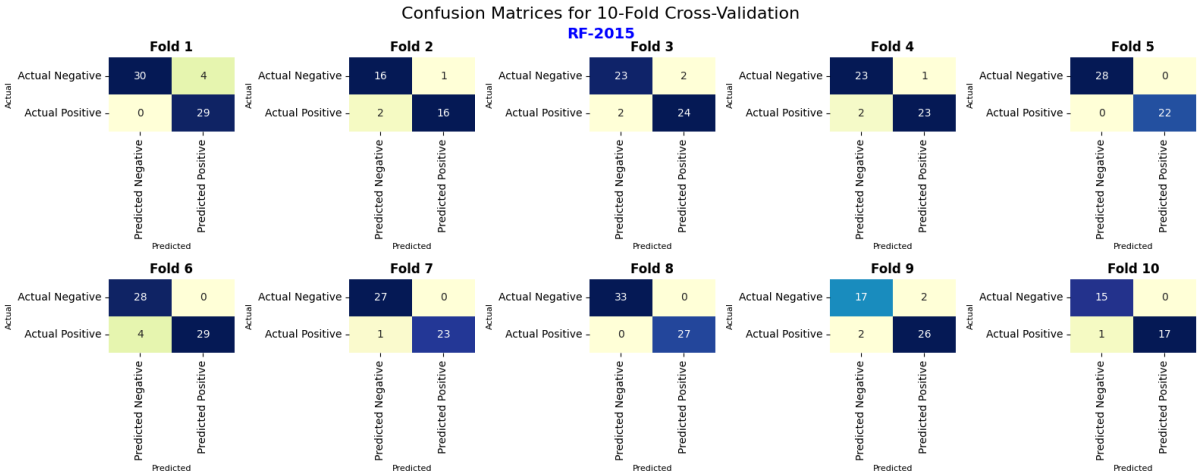


Figure A.6: Confusion metrics for 10-fold-cross-validation for RF 2015

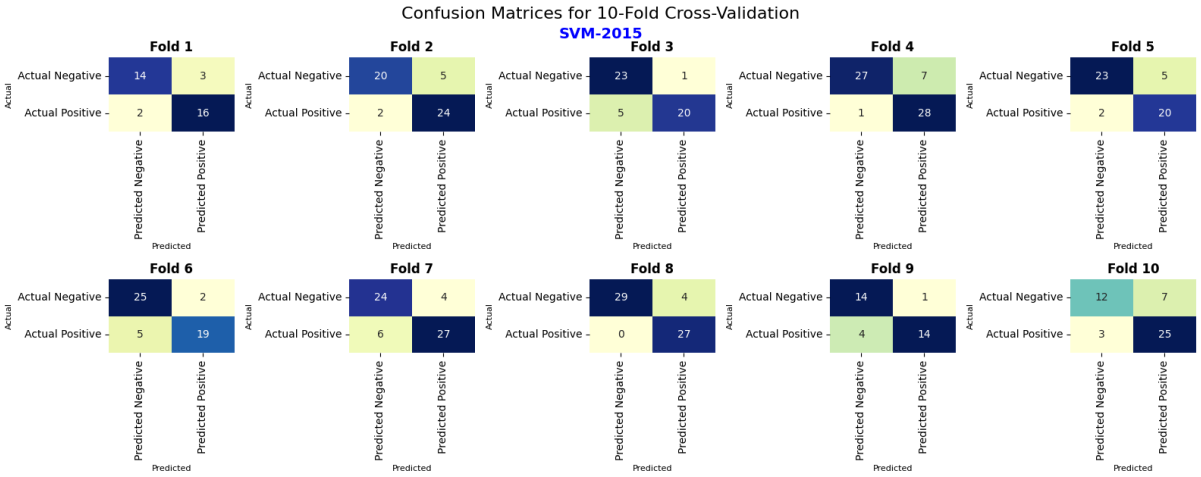


Figure A.7: Confusion metrics for 10-fold-cross-validation for SVM 2015

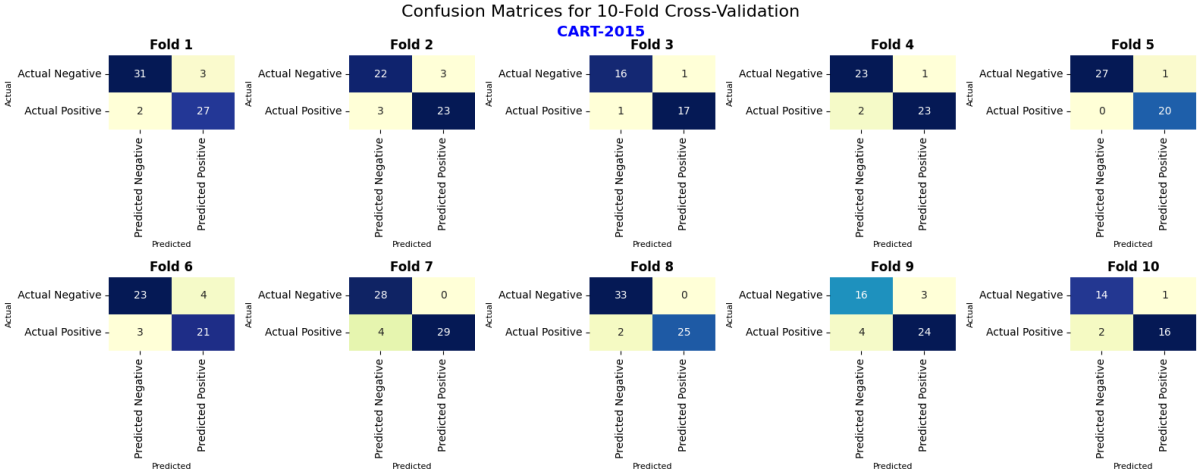


Figure A.8: Confusion metrics for 10-fold-cross-validation for CART 2015

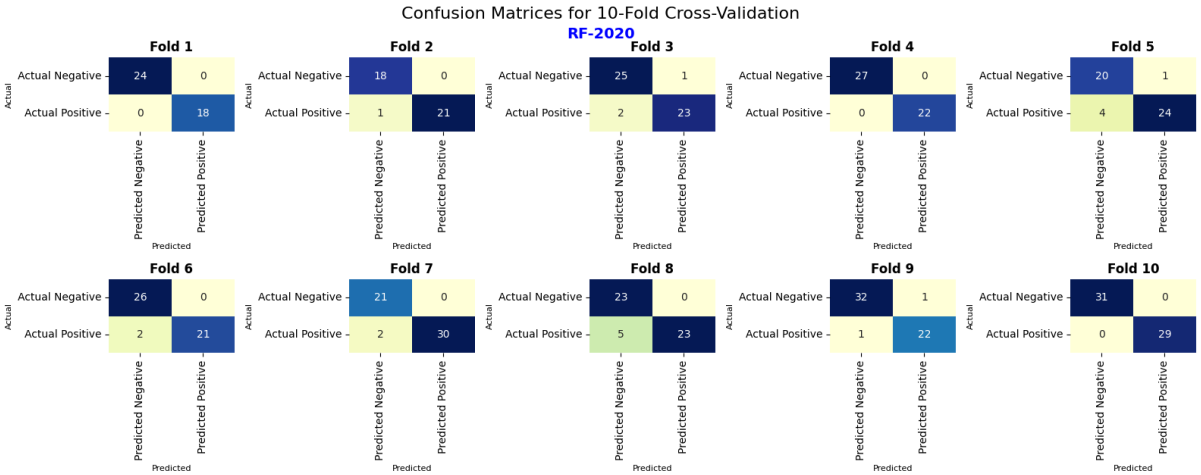


Figure A.9: Confusion metrics for 10-fold-cross-validation for RF 2020

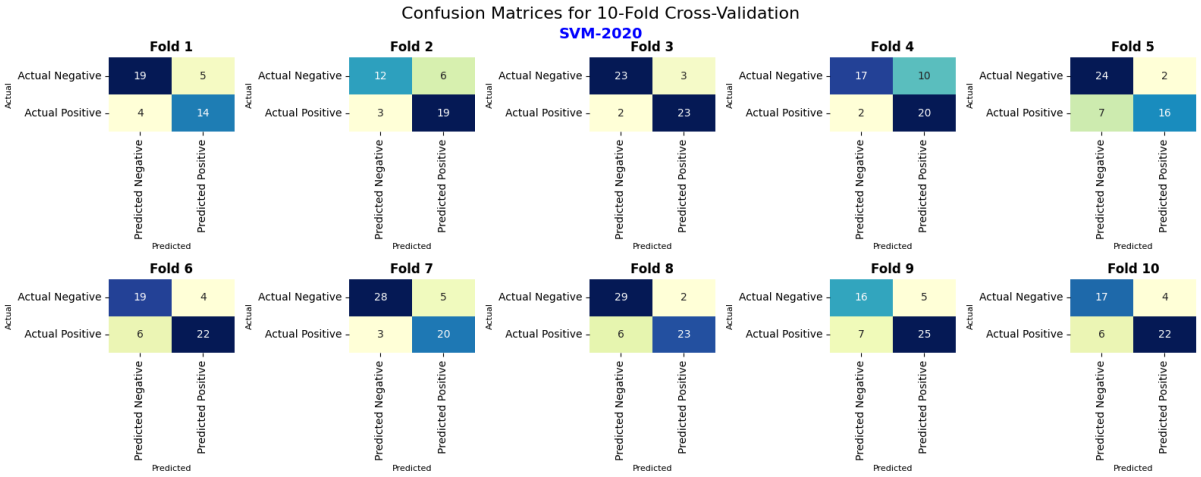


Figure A.10: Confusion metrics for 10-fold-cross-validation for SVM 2020

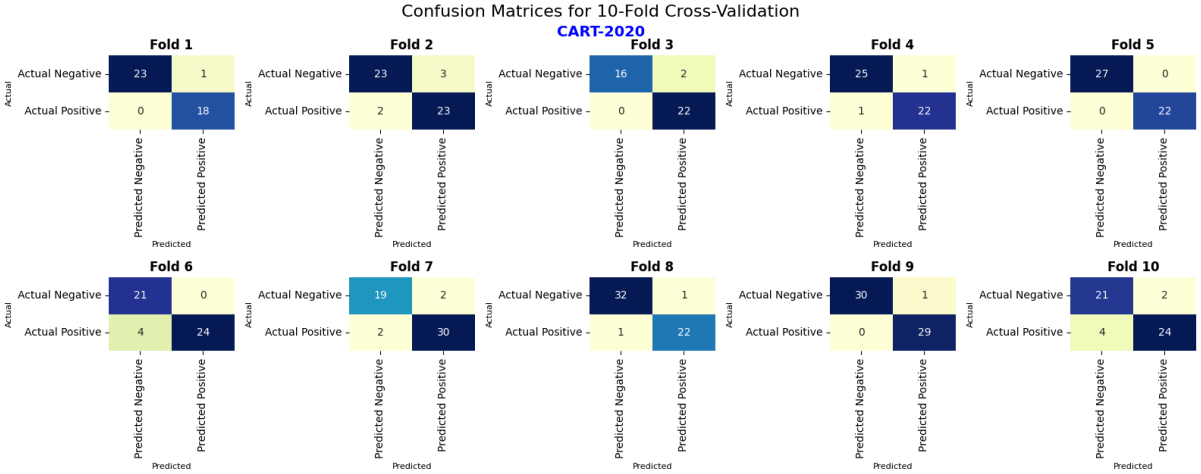


Figure A.11: Confusion metrics for 10-fold-cross-validation for CART 2020

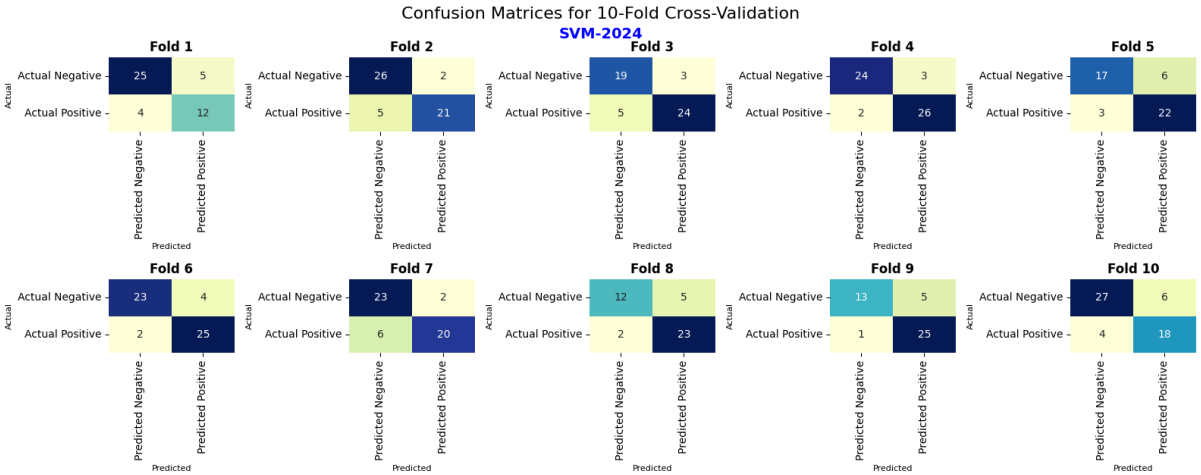


Figure A.12: Confusion metrics for 10-fold-cross-validation for SVM 2024

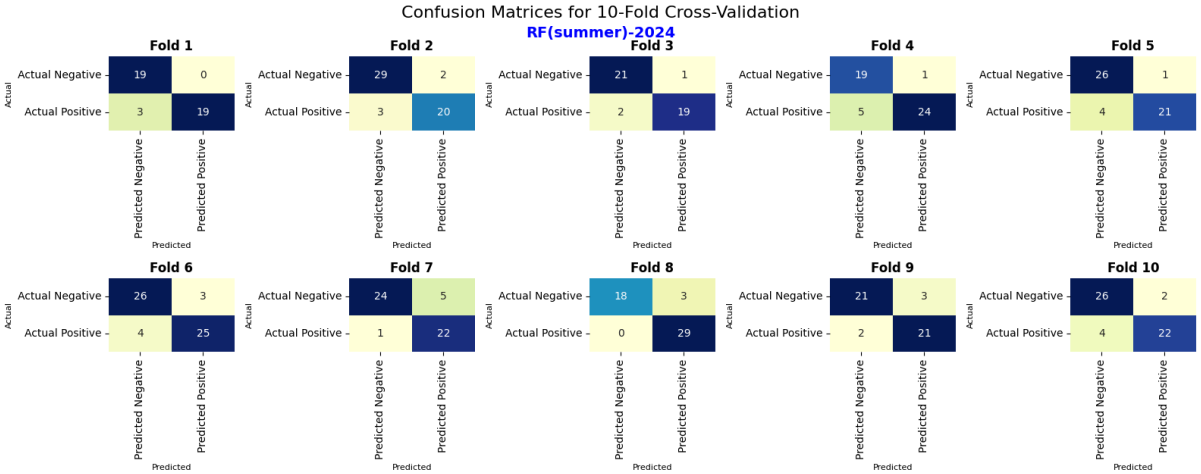


Figure A.13: Confusion metrics for RF in seasonal classification (May-July, 2024)

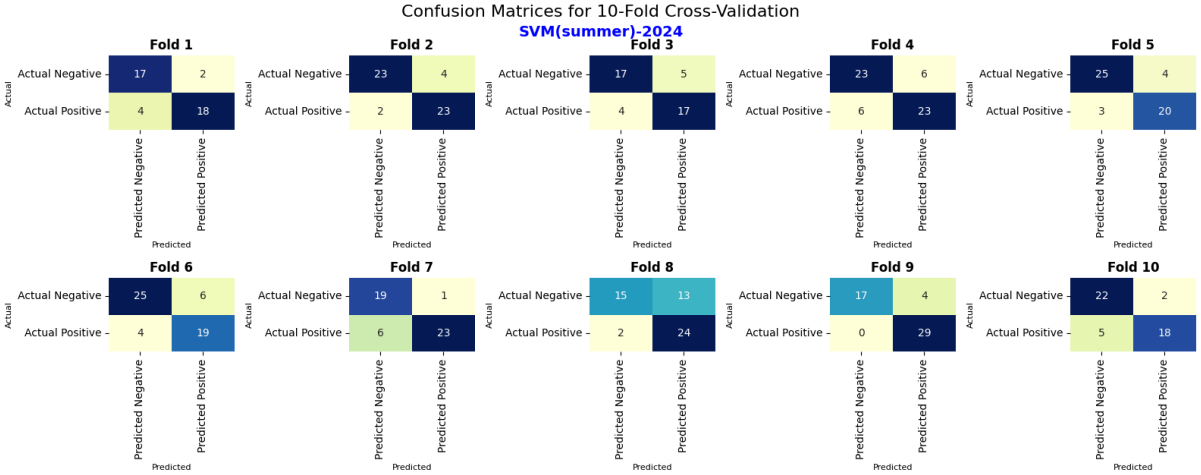


Figure A.14: Confusion metrics for SVM in seasonal classification (May-July, 2024)

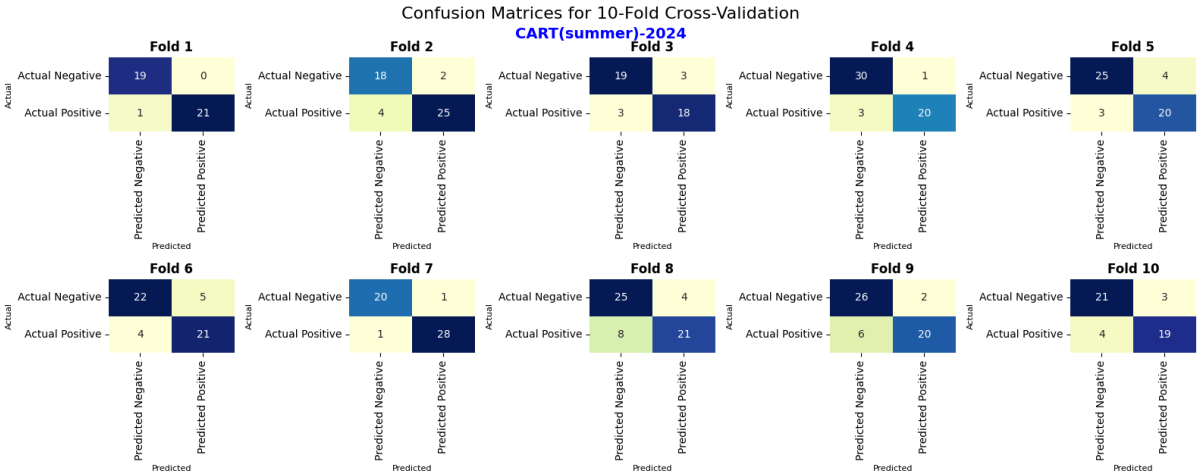


Figure A.15: Confusion metrics for CART in seasonal classification (May-July, 2024)

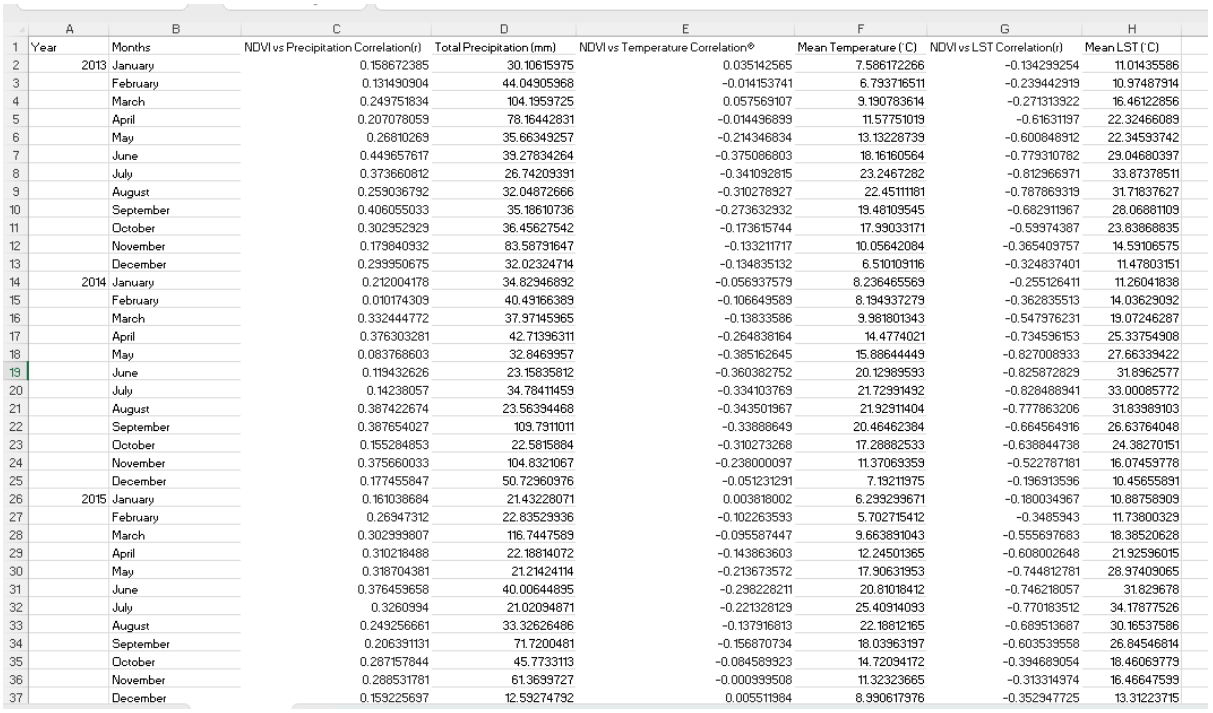
Yearly Aggregated Values:

	Year	Total Precipitation (mm)	Mean Temperature (°C)	Mean NDVI	Mean LST (°C)
0	2010.0	549.1157462685632	13.265189243215275	0.5105552351941226	20.369017855161246
1	2011.0	545.7550895761257	14.299839923113325	0.5058991676576821	22.00393669228265
2	2012.0	490.4913729980224	14.37816311806165	0.481182171380927	22.34403203786144
3	2013.0	577.5018224155154	13.87279807484498	0.5064351874013815	21.22562377351585
4	2014.0	558.2943746369348	14.758770073046662	0.47457479783382045	22.545418773231233
5	2015.0	490.22446334873433	14.550649296669917	0.4984103233138611	22.020509339212254
6	2016.0	538.7081713755756	14.61715486106226	0.4950288576303461	22.58871631379236
7	2017.0	357.1516283021764	14.659219561544774	0.48618695455994687	22.776947604387942
8	2018.0	628.5655585284021	14.46950348844888	0.48832375299440356	21.70342469689689
9	2019.0	461.0188910086192	14.726416106029603	0.5002858246062836	22.686077683905182
10	2020.0	563.0944267712083	14.706056906046683	0.5227312326120609	22.207059225198122
11	2021.0	609.7091970464778	14.163578378686225	0.5325576621039957	20.983190905524907
12	2022.0	549.5156393262369	15.356435397665166	0.5153129659202826	22.883396632486665
13	2023.0	464.7822032094366	15.528128723497105	0.5127855399905727	22.491776246833126
14	2024.0	540.7964659033446	15.477216761268664	0.4855079779353484	21.37709525177388

Figure A.16: Yearly aggregated values for total precipitation, mean temperature, mean LST and mean NDVI

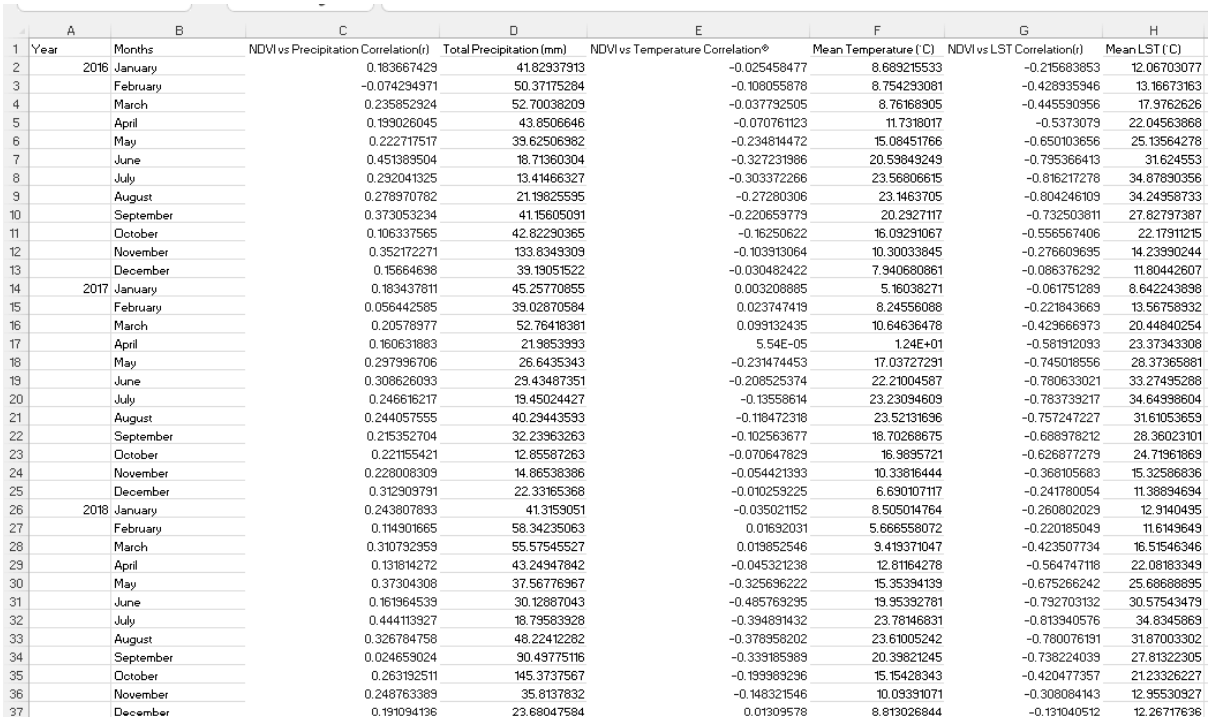
	A	B	C	D	E	F	G	H
1	Year	Months	NDVI vs Precipitation Correlation(r)	Total Precipitation (mm)	NDVI vs Temperature Correlation®	Mean Temperature (°C)	NDVI vs LST Correlation(r)	Mean LST (°C)
2	2010	January	0.173464381	68.40360033	0.01411201	5.426400257	-0.1075602	8.748962458
3		February	0.0859212	45.82109615	0.057776543	5.830974634	-0.150395209	11.57971762
4		March	0.218859329	54.42250848	0.118463962	7.82465858	-0.393252109	15.18778177
5		April	0.144870383	35.48502365	-0.019691215	11.58064458	-0.522493969	23.20639012
6		May	0.302300036	61.61879413	-0.164420982	13.70256124	-0.561107525	24.1242362
7		June	0.355728797	44.17251247	-0.328877773	18.65131504	-0.735422455	28.48647401
8		July	0.365033557	11.78529061	-0.279025327	23.82490003	-0.805564351	33.63400868
9		August	0.280052035	21.1655849	-0.27665449	22.75694481	-0.693129046	31.65537748
10		September	0.425521619	53.85953783	-0.264148156	19.18685468	-0.674101518	27.18906697
11		October	0.344935645	83.07761241	-0.191213538	13.78066722	-0.497821344	18.22114393
12		November	0.395491892	36.24437169	-0.110703383	9.183232934	-0.381332433	14.15186177
13		December	0.417428232	33.05981362	-0.103570644	5.946053374	-0.30448898	8.617880469
14	2011	January	0.201826409	28.01493104	-0.054857926	5.913719503	-0.245384481	10.17583196
15		February	0.188310255	27.73323671	0.013755779	7.800139206	-0.333610232	13.55529097
16		March	0.179653042	98.01215066	-0.022628066	8.143170197	-0.427128074	17.00445334
17		April	0.237165555	50.9075165	-0.093341581	13.88012988	-0.600523009	23.75549817
18		May	0.258794136	64.19516668	-0.310469194	16.27792487	-0.618543457	26.24987491
19		June	0.372269296	25.9653381	-0.382473014	18.82824014	-0.778361491	28.97818038
20		July	0.402493759	24.20414263	-0.280167351	21.50543513	-0.751355462	30.91962581
21		August	0.114974751	13.88197942	-0.224405079	23.18611519	-0.759735569	32.70541396
22		September	0.356438937	28.74532395	-0.251883817	20.58564058	-0.77042215	30.41453034
23		October	0.3672888	29.15101958	-0.202806129	15.99208115	-0.635653652	22.89056323
24		November	0.046426871	124.8700275	-0.035276385	11.16505827	-0.209253726	15.56921379
25		December	0.263624433	30.07425677	0.066347393	7.879228749	-0.115964838	11.29730892
26	2012	January	0.262227408	29.9997552	0.037584602	7.053658233	-0.173766347	11.42224391
27		February	0.15661607	11.46397729	0.077999052	4.000886534	-0.319856238	11.77335338
28		March	0.183831735	32.86159358	0.028540396	9.366837949	-0.594487799	20.53826246
29		April	0.157647267	66.95606817	-0.003556336	11.77065634	-0.482696506	20.2601392
30		May	0.272755343	29.70978949	-0.254092527	16.91781897	-0.746920417	29.77191023
31		June	0.339696938	29.24903257	-0.31817819	22.17314672	-0.81277149	33.06195123
32		July	0.254026063	14.65055875	-0.225754405	22.35374764	-0.661355263	33.59757506
33		August	0.30873174	8.094977959	-0.200625043	24.92377695	-0.811297034	33.87072549
34		September	0.276666793	68.41676846	-0.183815256	19.81086579	-0.751269412	29.10575313
35		October	0.239646633	96.83917268	-0.048693548	15.45462574	-0.405261473	20.03257676
36		November	0.134614244	83.20477563	0.095519019	10.41460194	-0.006185966	12.06451487
37		December	0.226080644	19.04490322	0.140965975	7.882283693	-0.0517908	12.34376845

Figure A.17: Per month aggregated values for NDVI with precipitation Pearson correlation value (r) and total precipitation, NDVI with precipitation Pearson correlation value (r) and mean temperature, NDVI with precipitation Pearson correlation value (r) and mean LST (2010-2012)



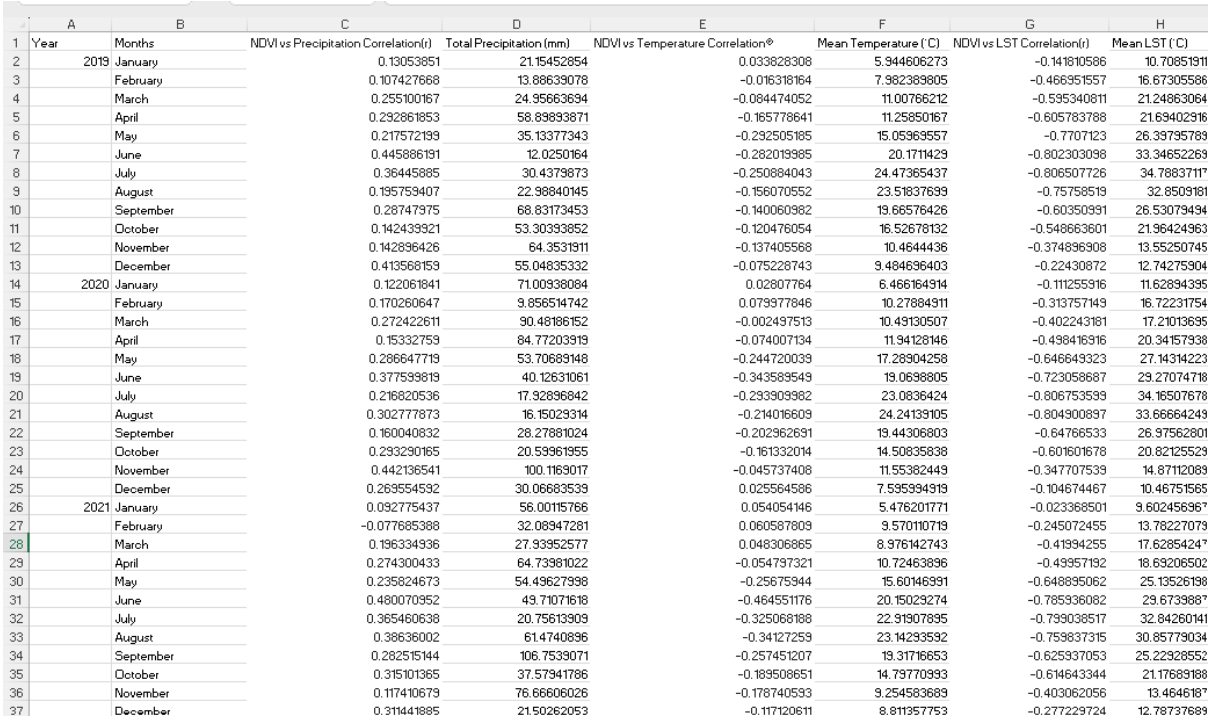
	A	B	C	D	E	F	G	H
1	Year	Months	NDVI vs Precipitation Correlation(r)	Total Precipitation (mm)	NDVI vs Temperature Correlation®	Mean Temperature (°C)	NDVI vs LST Correlation(r)	Mean LST (°C)
2	2013	January	0.158672385	30.10615975	0.035142565	7.586172266	-0.134239254	11.01435586
3		February	0.131490904	44.04905968	-0.014153741	6.793716511	-0.239442919	10.97487914
4		March	0.249751834	104.1959725	0.057563107	9.190783614	-0.271313922	16.46122856
5		April	0.207078059	78.16442831	-0.014496899	11.57751019	-0.61631197	22.32466089
6		May	0.26810269	35.66349257	-0.214346834	13.13228739	-0.600848912	22.34593742
7		June	0.449657617	39.27834264	-0.375086803	18.16160564	-0.779310782	29.04680397
8		July	0.373660812	26.74209391	-0.341092815	23.2467282	-0.812966971	33.87378511
9		August	0.259036792	32.04872666	-0.310278927	22.45111181	-0.787863319	31.71837627
10		September	0.406055033	35.18610736	-0.273632932	19.48109545	-0.682911967	28.06881109
11		October	0.302952929	36.45627542	-0.173615744	17.99033171	-0.59374387	23.83868835
12		November	0.173840932	83.58791647	-0.133211717	10.05642084	-0.365409757	14.59106575
13		December	0.299950675	32.02324714	-0.134835132	6.510109116	-0.324837401	11.47803151
14	2014	January	0.212004178	34.82946892	-0.056937579	8.236465569	-0.255126411	11.26041838
15		February	0.010174309	40.49166389	-0.106649589	8.194937279	-0.362835513	14.03629092
16		March	0.332444772	37.97145965	-0.13833586	9.981801343	-0.547976231	19.07246287
17		April	0.376303281	42.71396311	-0.264838164	14.4774021	-0.734596153	25.33754908
18		May	0.083768603	32.8469957	-0.385162645	15.88644449	-0.827008933	27.66339422
19		June	0.119432626	23.15835812	-0.360382752	20.12989593	-0.825872829	31.8962577
20		July	0.14238057	34.78411459	-0.334103769	21.72991492	-0.828488941	33.00085772
21		August	0.387422674	23.56394468	-0.343501967	21.92911404	-0.777863206	31.83989103
22		September	0.387654027	109.7911011	-0.33886649	20.46462384	-0.664564916	26.63764048
23		October	0.155284853	22.5815884	-0.310273268	17.28882533	-0.638844738	24.38270151
24		November	0.375660033	104.8321067	-0.238000097	11.37069359	-0.522787181	16.07459778
25		December	0.177455847	50.72960976	-0.051231291	7.19211975	-0.196913596	10.45655891
26	2015	January	0.161038684	21.43228071	0.003818002	6.239239671	-0.180034967	10.88758909
27		February	0.26947312	22.83529936	-0.102263593	5.702715412	-0.3485943	11.73800329
28		March	0.302999807	116.7447589	-0.095587447	9.663891043	-0.555697683	18.38520628
29		April	0.310218488	22.18814072	-0.143863603	12.24501365	-0.608002648	21.92596015
30		May	0.318704381	21.21424114	-0.213673572	17.90631953	-0.744812781	28.97409065
31		June	0.376459658	40.00644895	-0.298228211	20.81018412	-0.746218057	31.829678
32		July	0.3260994	21.02094871	-0.221328129	25.40914093	-0.770183512	34.17877526
33		August	0.249256661	33.32626486	-0.137916813	22.18812165	-0.689513687	30.16537586
34		September	0.206391131	71.7200481	-0.156870734	18.03963197	-0.603539558	26.84546814
35		October	0.287157844	45.7733113	-0.084589323	14.72094172	-0.394689054	18.46069779
36		November	0.288531781	61.3693727	-0.000999508	11.32323665	-0.313314974	16.46647599
37		December	0.159225697	12.59274792	0.005511984	8.990617976	-0.352947725	13.31223715

Figure A.18: Per month aggregated values for NDVI with precipitation Pearson correlation value (r) and total precipitation, NDVI with precipitation Pearson correlation value (r) and mean temperature, NDVI with precipitation Pearson correlation value (r) and mean LST (2013-2015)



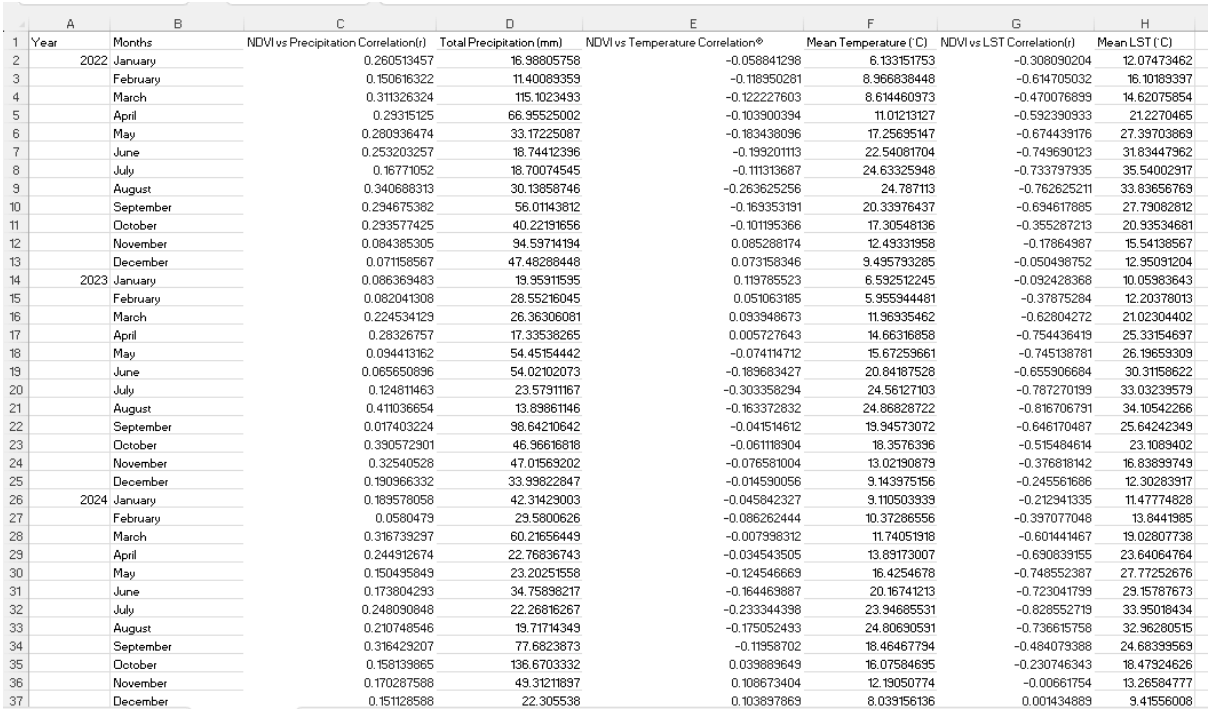
	A	B	C	D	E	F	G	H
1	Year	Months	NDVI vs Precipitation Correlation(r)	Total Precipitation (mm)	NDVI vs Temperature Correlation®	Mean Temperature (°C)	NDVI vs LST Correlation(r)	Mean LST (°C)
2	2016	January	0.183667429	41.82937913	-0.025458477	8.689215533	-0.215683853	12.06703077
3		February	-0.074294971	50.37175284	-0.108055878	8.754293081	-0.428935946	13.16673163
4		March	0.235852924	52.70038209	-0.037792505	8.76168905	-0.445590956	17.9762626
5		April	0.199026045	43.8506646	-0.070761123	11.7318017	-0.5373079	22.04563868
6		May	0.222717517	39.62506982	-0.234814472	15.08451766	-0.650103656	25.13564278
7		June	0.451389504	18.71360304	-0.327231986	20.59843249	-0.795366413	31.624553
8		July	0.292041325	13.41466327	-0.303372266	23.56806615	-0.816217278	34.87890356
9		August	0.278970782	21.19825595	-0.27280306	23.1463705	-0.804246109	34.24958733
10		September	0.373053234	41.15605091	-0.220659779	20.2927117	-0.732503811	27.82797387
11		October	0.106337565	42.82290365	-0.16250622	16.09291067	-0.556567406	22.17911215
12		November	0.352172271	133.8349309	-0.103913064	10.30033845	-0.276609695	14.23990244
13		December	0.15664698	39.19051522	-0.030482422	7.940680861	-0.086376292	11.80442607
14	2017	January	0.183437811	45.25770855	0.003208885	5.16038271	-0.061751289	8.642243898
15		February	0.056442585	39.02870584	0.023747419	8.24556088	-0.221843669	13.56758932
16		March	0.20578977	52.76418381	0.093132435	10.64636478	-0.429666973	20.44840254
17		April	0.160631883	21.9853993	5.54E-05	1.24E+01	-0.581912093	23.37343308
18		May	0.297396706	26.6435343	-0.231474453	17.03727291	-0.745018556	28.37365881
19		June	0.308626093	29.43487351	-0.208525374	22.21004587	-0.780633021	33.27495288
20		July	0.246616217	19.45024427	-0.13558614	23.23094609	-0.783739217	34.64998604
21		August	0.244057555	40.29443593	-0.118472318	23.52131696	-0.757247227	31.61053659
22		September	0.215352704	32.23963263	-0.102563677	18.70268675	-0.688978212	28.36023101
23		October	0.221155421	12.85587263	-0.070647829	16.9895721	-0.626877279	24.71961869
24		November	0.228008309	14.86538386	-0.054421393	10.33816444	-0.368105683	15.32586836
25		December	0.312909791	22.33165368	-0.010259225	6.690107117	-0.241780054	11.38894694
26	2018	January	0.243807893	41.3159051	-0.035021152	8.505014764	-0.260802029	12.9140495
27		February	0.114901665	58.34235063	0.01692031	5.666558072	-0.220185049	11.6149649
28		March	0.310792959	55.57545527	0.019852546	9.419371047	-0.423507734	16.51546346
29		April	0.131814272	43.24947842	-0.045321238	12.81164278	-0.564747118	22.08183349
30		May	0.37304308	37.56776967	-0.325696222	15.35394139	-0.675266242	25.68688895
31		June	0.161964539	30.12887043	-0.485763295	19.95392781	-0.792703132	30.57543479
32		July	0.444113927	18.79583928	-0.394891432	23.7814683	-0.813940576	34.8345869
33		August	0.326784758	48.22412282	-0.378958202	23.61005242	-0.780076191	31.87003302
34		September	0.024659024	90.49775116	-0.339185989	20.39821245	-0.738224039	27.81322305
35		October	0.263192511	145.3737567	-0.199989296	15.15428343	-0.420477357	21.23326227
36		November	0.248763389	35.8137832	-0.148321546	10.09391071	-0.308084143	12.95530927
37		December	0.191094136	23.68047584	0.01309578	8.813026844	-0.131040512	12.26717636

Figure A.19: Per month aggregated values for NDVI with precipitation Pearson correlation value (r) and total precipitation, NDVI with precipitation Pearson correlation value (r) and mean temperature, NDVI with precipitation Pearson correlation value (r) and mean LST (2016-2018)



	A	B	C	D	E	F	G	H
1	Year	Months	NDVI vs Precipitation Correlation(r)	Total Precipitation (mm)	NDVI vs Temperature Correlation®	Mean Temperature (˚C)	NDVI vs LST Correlation(r)	Mean LST (˚C)
2	2019	January	0.13053851	21.15452854	0.033828308	5.944606273	-0.141810586	10.70851911
3		February	0.107427668	13.88639078	-0.016318164	7.982389805	-0.466951557	16.67305586
4		March	0.255100167	24.95663694	-0.084474052	11.00766212	-0.595340811	21.24863064
5		April	0.292861853	58.89893871	-0.165778641	11.25850167	-0.605783788	21.69402916
6		May	0.217572199	35.13377343	-0.292505185	15.05969557	-0.7707123	26.39795789
7		June	0.445886191	12.0250164	-0.282019985	20.1711429	-0.802303098	33.34652269
8		July	0.36445885	30.4379873	-0.250884043	24.47365437	-0.806507726	34.78837117
9		August	0.195759407	22.98840145	-0.156070552	23.51837699	-0.75758519	32.8509181
10		September	0.28747975	68.83173453	-0.140060982	19.66576426	-0.60350991	26.53079434
11		October	0.142439921	53.30393852	-0.120476054	16.52678132	-0.548663601	21.96424963
12		November	0.142896426	64.3531911	-0.137405568	10.4644436	-0.374896908	13.55250745
13		December	0.413568159	55.04835332	-0.075228743	9.484696403	-0.22430872	12.74275904
14	2020	January	0.122061841	71.00938084	0.02807764	6.466164914	-0.111255916	11.62894395
15		February	0.170260647	9.856514742	0.079377846	10.27884911	-0.313757149	16.72231754
16		March	0.272422611	90.48186152	-0.002497513	10.49130507	-0.402243181	17.21013695
17		April	0.15332759	84.77203919	-0.074007134	11.94128146	-0.498416916	20.34157938
18		May	0.286647719	53.70689148	-0.244720039	17.28904258	-0.646649323	27.14314223
19		June	0.377599819	40.12631061	-0.343589549	19.0698805	-0.723058687	29.27074718
20		July	0.216820536	17.92896842	-0.293909362	23.0836424	-0.806753599	34.16507678
21		August	0.302777873	16.15029314	-0.214016609	24.24139105	-0.804900897	33.66664249
22		September	0.160040832	28.27881024	-0.202962691	19.44306803	-0.64766533	26.97562801
23		October	0.293290165	20.59961955	-0.161332014	14.50835838	-0.601601678	20.82125529
24		November	0.442136541	100.1169017	-0.045737408	11.55382449	-0.347707539	14.87112089
25		December	0.263554532	30.06683539	0.025564586	7.595394319	-0.104674467	10.46751565
26	2021	January	0.092775437	56.00115766	0.054054146	5.476201771	-0.023368501	9.602456967
27		February	-0.077685388	32.08947281	0.060587809	9.570110719	-0.245072455	13.78227079
28		March	0.196334936	27.93952577	0.048306865	8.976142743	-0.41994255	17.62854247
29		April	0.274300433	64.73981022	-0.054797321	10.72463896	-0.49357192	18.69206502
30		May	0.235824673	54.49627398	-0.25675944	15.60146991	-0.648895062	25.13526198
31		June	0.480070952	49.71071618	-0.464551176	20.15029274	-0.785936082	29.6739887
32		July	0.365460638	20.75613909	-0.325068188	22.91907895	-0.799038517	32.84260141
33		August	0.38636002	61.4740896	-0.34127259	23.14293592	-0.759837315	30.85779034
34		September	0.282515144	106.7539071	-0.257451207	19.31716653	-0.625937053	25.22928552
35		October	0.315101365	37.57941786	-0.189508651	14.79770993	-0.614643344	21.17689188
36		November	0.117410679	76.66606026	-0.178740593	9.254583689	-0.403062056	13.4646187
37		December	0.311441885	21.50262053	-0.117120611	8.811357753	-0.277229724	12.78737689

Figure A.20: Per month aggregated values for NDVI with precipitation Pearson correlation value (r) and total precipitation, NDVI with precipitation Pearson correlation value (r) and mean temperature, NDVI with precipitation Pearson correlation value (r) and Mean LST (2019-2021)



	A	B	C	D	E	F	G	H
1	Year	Months	NDVI vs Precipitation Correlation(r)	Total Precipitation (mm)	NDVI vs Temperature Correlation®	Mean Temperature (˚C)	NDVI vs LST Correlation(r)	Mean LST (˚C)
2	2022	January	0.260513457	16.98805758	-0.058841298	6.133151753	-0.308090204	12.07473462
3		February	0.150616322	11.40089359	-0.118950281	8.966838448	-0.614705032	16.10189397
4		March	0.311326324	115.1023493	-0.122227603	8.614460973	-0.470076899	14.62075854
5		April	0.29315125	66.95525002	-0.103900394	11.01213127	-0.592390933	21.2270465
6		May	0.280936474	33.17225087	-0.183438096	17.25695147	-0.674439176	27.39703869
7		June	0.253203257	18.74412396	-0.199201113	22.54081704	-0.749690123	31.83447962
8		July	0.16771052	18.70074545	-0.111313687	24.63325948	-0.733797935	35.54002917
9		August	0.340688313	30.13858746	-0.263625256	24.787113	-0.762625211	33.83656769
10		September	0.294675382	56.01143812	-0.169353191	20.33976437	-0.694617885	27.79082812
11		October	0.293577425	40.22191656	-0.101195366	17.30548136	-0.355287213	20.93534681
12		November	0.084385305	94.59714194	0.085288174	12.49331958	-0.17864987	15.54138567
13		December	0.071158567	47.48288448	0.073158346	9.495793285	-0.050498752	12.95091204
14	2023	January	0.086369483	19.95911595	0.119785523	6.592512245	-0.092428368	10.05383643
15		February	0.082041308	28.55216045	0.051063185	5.955944481	-0.37875284	12.20378013
16		March	0.224534129	26.36306081	0.093948673	11.96935462	-0.62804272	21.02304402
17		April	0.28326757	17.33538265	0.005727643	14.66316858	-0.754436419	25.33154697
18		May	0.094413162	54.45154442	-0.074114712	15.67259661	-0.745138781	26.19659309
19		June	0.065650896	54.02102073	-0.189683427	20.84187528	-0.655906684	30.31158622
20		July	0.124811463	23.57911167	-0.303358294	24.56127103	-0.787270199	33.03239579
21		August	0.411036654	13.89861146	-0.163372832	24.86828722	-0.816706791	34.10542266
22		September	0.017403224	98.64210642	-0.041514612	19.94573072	-0.646170487	25.64242349
23		October	0.390572901	46.96616818	-0.061118904	18.3576396	-0.515484614	23.1089402
24		November	0.32540528	47.01569202	-0.076581004	13.02190879	-0.376818142	16.83893749
25		December	0.190966332	33.99822847	-0.014590056	9.143975156	-0.245561686	12.30283917
26	2024	January	0.189578058	42.31423003	-0.045842327	9.110503939	-0.212941335	11.47774828
27		February	0.0580479	29.5800626	-0.086262444	10.37286556	-0.397077048	13.8441985
28		March	0.316739297	60.21656449	-0.007398312	11.74051918	-0.601441467	19.02807738
29		April	0.244912674	22.76836743	-0.034543505	13.89173007	-0.690839155	23.64064764
30		May	0.150495849	23.20251558	-0.124546669	16.4254678	-0.748552387	27.77252676
31		June	0.173804293	34.75898217	-0.164469887	20.16741213	-0.723041799	29.15787673
32		July	0.248090848	22.26816267	-0.233344398	23.94685531	-0.828552719	33.95018434
33		August	0.210748546	19.71714349	-0.175052493	24.80690591	-0.736615758	32.96280515
34		September	0.316429207	77.6823873	-0.11958702	18.46467794	-0.484079388	24.68399569
35		October	0.158139865	136.6703332	0.039889649	16.07584695	-0.230746343	18.47924626
36		November	0.170287588	49.31211897	0.108673404	12.19050774	-0.00661754	13.26584777
37		December	0.151128588	22.305538	0.103897869	8.039156136	0.001434889	9.41556008

Figure A.21: Per month aggregated values for NDVI with precipitation Pearson correlation value (r) and total precipitation, NDVI with precipitation Pearson correlation value (r) and mean temperature, NDVI with precipitation Pearson correlation value (r) and mean LST (2022-2024)

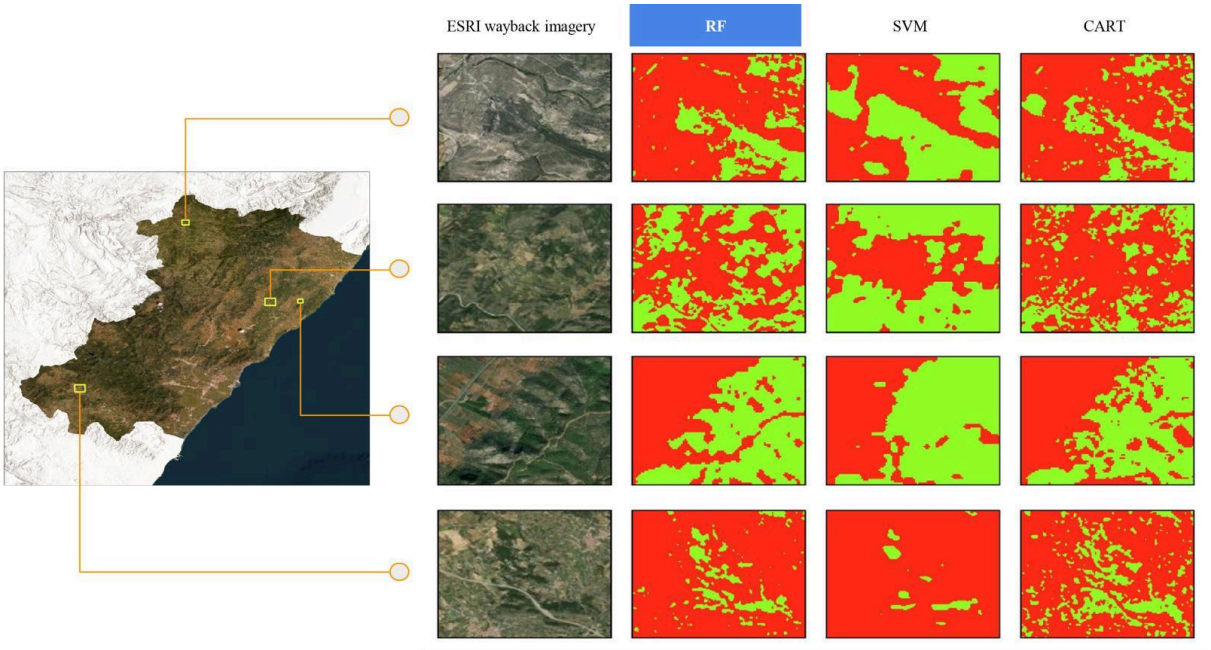


Figure A.22: Visual inspection of classifiers by comparing results of three models

Masters Program in **Geospatial Technologies**

Supported by:
Education and Culture
ERASMUS MUNDUS