POLITECNICO DI TORINO

Master Degree Agritech Engineering

Master Degree Thesis

Satellite-derived NDVI and climate projections for sustainable pasture biomass management: a case study in the Madrid region, Spain

Tutors:

prof. Carlos Gregorio Diaz Ambrona

prof. Elena Belcore

Candidate:

Bruno Toro

POLITECNICO DI TORINO

Master Degree Agritech Engineering

Master Degree Thesis

Satellite-derived NDVI and climate projections for sustainable pasture biomass management: a case study in the Madrid region, Spain

Tutors:

prof. Carlos Gregorio Diaz Ambrona

prof. Elena Belcore

Candidate:

Bruno Toro

Summary

Grasslands are among the world's most vital ice-free ecosystems, providing critical services such as carbon sequestration, biodiversity support, and pollination. However, these systems face increasing pressure from climate change, characterized by rising temperatures and altered precipitation patterns. Predictive models that integrate historical data with future climate scenarios are essential for anticipating shifts in biomass dynamics and developing adaptive management strategies to preserve the ecological and economic functions of grasslands.

This thesis, developed in collaboration with the CEIGRAM Research Center in Madrid (central Spain), evaluates the efficacy of satellite-derived vegetation indices, specifically the Normalized Difference Vegetation Index (NDVI), as reliable proxy for pasture biomass production under climate change. It addresses a key limitation in current agricultural insurance models, which often rely on long-term NDVI averages that may not accurately capture real-time forage availability in a changing climate. Through a combined analysis of satellite imagery and field measurements, this research performs a statistical comparison between NDVI trends and directly measured biomass growth.

The study is grounded in a mathematical model that integrates soil properties and climatic data to simulate pasture biomass. Field campaigns provide crucial data for model calibration, ensuring that spectral indices robustly reflect actual on-the-ground biomass. The research was carried out across three areas of the Community of Madrid, Spain, each chosen for their distinctive climatic and edaphic features; these locations encompass a gradient of environmental conditions, ranging from humid mountainous zones to semi-arid agricultural landscapes. The findings aim to determine if incorporating direct biomass estimation can enhance the precision of pasture monitoring systems; as shown in the chapter 4 a correlation between NDVI and pasture biomass confirms that satellite data can be effectively and scientifically employed as a reliable indicator for monitoring the status, productivity, and health of pastoral ecosystems in a continuous and large-scale manner. This research contributes to optimizing remote sensing tools for sustainable rangeland management, with potential implications for improving agricultural insurance frameworks, informing policy, and supporting climate adaptation strategies.

Contents

Summary **I**

Contents **II**

List of Figures **IV**

List of Tables **VI**

1 Introduction **1**
1.1 Agriculture and climate change 1
1.2 Importance of pasture 3
1.3 State of the art and thesis' goal 6

2 Materials and case study **9**
2.1 Case study 9
2.2 Data utilized 12
2.2.1 NDVI 12
2.2.2 Climate models 14

3 Methodology **20**
3.1 Data analysis and processing 20
3.1.1 Pre-analysis of the dataset 20
3.1.2 Regression 22
3.1.3 Correlation 26
3.2 Software 28
3.2.1 Functions and formulas 28
3.3 Workflow 31
3.3.1 Data preparation 32
3.3.2 Regression and correlation analysis 33

4 Results and discussion **40**
4.1 Regression results 42
4.2 Correlation results 48

5 Conclusions **54**

APPENDIX A **56**

Bibliography **64**

List of Figures

1.1 Contribution of crops and livestock activities to total non-CO_2 emissions from agriculture in 2018 (5.3 Gt CO_2eq) [1]. 2
1.2 Grassland soil carbon sequestration [2] 4

2.1 Geographic distribution of pasture types (tree-dominated, shrub-dominated, and herbaceous pastures) within the study area [3]. 10
2.2 Traditional field survey in Buitrago del Lozoya, May 2025. 11
2.3 Traditional field survey in Colmenar Viejo, May 2025. 11
2.4 Traditional field survey in Tielmes, May 2025. 12
2.5 Spectral resolution healthy and stressed vegetation 13

3.1 Box plot, graphical representation 22
3.2 Pearson correlation values 27
3.3 Code correlation function 30
3.4 Graphical workflow 31
3.5 Code illustration part 1 34
3.6 Code illustration part 2 35
3.7 Code illustration part 3 36
3.8 Code illustration part 4 37
3.9 Code illustration part 5 38
3.10 Code illustration part 6 39

4.1 Time series area 12280; top model Access, below model Cnrm 41
4.2 Regression analysis model MIRO area 12280: top automatic regression, below seasonal regression 44
4.3 Automatic regression area 11033, model CNRM 46
4.4 Seasonal regreggion area 9862, model MPI 46
4.5 Correlation's heatmap based on regression line: top-left area 12280, top-right area 9862, below area 11033 49

4.6 Heatmap correlation based on dataset area 12280: left automatic, right season 50
4.7 Heatmap correlation based on dataset area 11033: left automatic, right season 51
4.8 Heatmap correlation based on dataset area 9862: left automatic, right season . 51

List of Tables

2.1 MOD09Q1 layers 14
2.2 Description of scenarios IPCC 17
2.3 Model characteristics 19

4.1 Results area 12280 43
4.2 Slopes' values for NDVI and pasture across the study areas 48

Chapter 1

Introduction

1.1 Agriculture and climate change

For millennia, agriculture has been the backbone of human civilization, ensuring food security, shaping cultural landscapes, and enabling socioeconomic development. Yet, in the 21st century, this ancient sector stands at a critical crossroads, increasingly strained by the multifaceted consequences of climate change. Rising global temperatures, altered precipitation patterns, more frequent and intense extreme weather events and the growing atmospheric concentration of greenhouse gases (particularly carbon dioxide (CO_2)) are already reshaping agroecosystems worldwide, undermining both productivity and the long-term stability of global food systems [4, 5].

Taking as an example the wheat culture, thermal analyses indicate that wheat yields decline substantially as temperatures exceed optimal thresholds, with estimated reductions of 5-15$\%$ additional degree Celsius [6, 7]. Simultaneously, increased variability in precipitation patterns and a higher frequency of drought events are undermining rainfed agricultural systems, particularly in tropical and subtropical regions [8, 9].
Droughts have already led to significant global agricultural losses, with average yield reductions of 9–10% for major cereals [10]. Beyond heat and water stress, the rising incidence of extreme weather events—such as heatwaves and intense rainfall—has caused direct damage to crops, especially during critical stages like flowering and grain set [11, 12]. Indirect effects include the poleward expansion of pests and pathogens, facilitated by warmer conditions [13]. Furthermore, elevated atmospheric CO_2 concentrations have been linked to declines in the nutritional quality of staple crops, with reductions in protein, zinc, and iron content [14]. Coastal agricultural zones face additional threats from soil salinization due to sea-level rise and saltwater intrusion into coastal aquifers [15]. Collectively, these challenges threaten

global food security, with projections suggesting substantial declines in agricultural productivity by mid-century, especially in already vulnerable regions [16]. Adaptation measures, such as breeding crop varieties resistant to abiotic stress and implementing more sustainable farming practices, are increasingly recognized as essential for mitigating negative impacts [17].

The relationship between climate and agriculture is not unidirectional. While agriculture is deeply vulnerable to climatic variability, it is also a significant contributor to anthropogenic climate change. According to the Food and Agriculture Organization, agricultural activities, including land use change, enteric fermentation, and the application of synthetic fertilizers, account for approximately 23% of total global greenhouse gas emissions, [1], a more detelied description of the different contribution is illustraed in figure 1.1. This complex, bidirectional relationship necessitates a nuanced and interdisciplinary approach that not only examines the vulnerabilities of agricultural systems but also identifies pathways for both adaptation and mitigation.

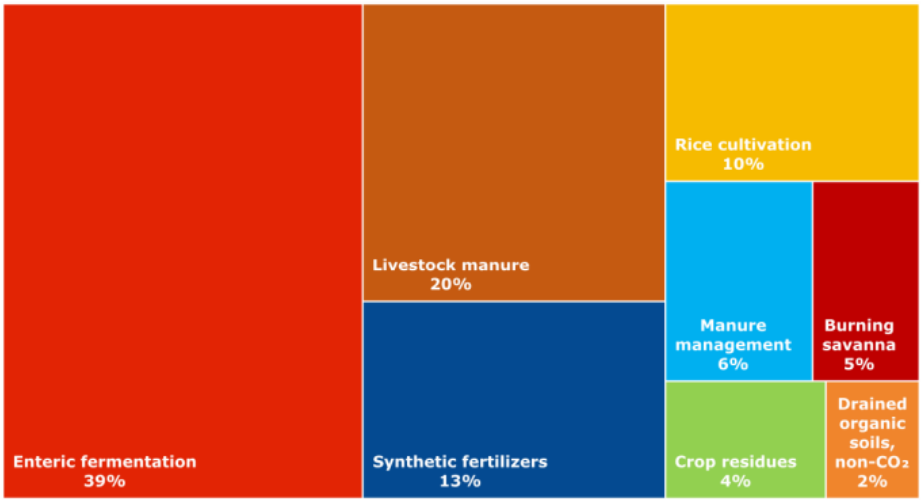

Figure 1.1: Contribution of crops and livestock activities to total non-CO_2 emissions from agriculture in 2018 (5.3 Gt CO_2eq) [1].

A growing body of scientific research has documented the adverse effects of climate change on crop yields, particularly in low-income and tropical regions where adaptive capacity is limited. Lobell et al. [18] estimate that, without significant adaptation measures, global warming could reduce the yields of key staple crops such as maize, wheat, and rice by 5–10% by mid-century, disproportionately affecting already food-insecure populations. Similarly, Schlenker and Roberts [19] highlight the nonlinear effects of temperature increases on crop productivity, with even short periods of extreme heat causing substantial yield losses. At the same time, certain forms of intensive agriculture are accelerating environmental degrada-

tion. The widespread use of nitrogen-based fertilizers contributes to the emission of nitrous oxide (N_2O), a greenhouse gas with a global warming potential nearly 300 times greater than that of CO_2 [20]. Livestock farming emits large quantities of methane (CH_4), while deforestation for agricultural expansion not only releases stored carbon but also diminishes the land's natural capacity for carbon sequestration [21].

However, agriculture is not solely a victim or a driver of climate change, it also holds transformative potential to become part of the solution. The concept of *climate-smart agriculture* (CSA), introduced by Lipper et al. [22], emphasizes the integration of adaptation, mitigation, and food security objectives. Practices such as conservation tillage, agroforestry, and integrated nutrient management can enhance soil health, improve water use efficiency, and reduce GHG (green house gases) emissions while increasing resilience to climatic stressors [23, 24]. Moreover, recent advancements in agricultural technology and data science are opening new frontiers. The development of drought-resistant crop varieties, the deployment of precision agriculture tools, and the use of AI-driven climate forecasting models are enabling more informed, proactive responses to climate variability [7]. These innovations, when embedded in inclusive policy frameworks and supported by robust rural extension systems, have the potential to steer agriculture toward greater sustainability and climate resilience.

1.2 Importance of pasture

Grassland ecosystems represent one of the planet's most vital biomes, delivering indispensable ecosystem services that sustain both environmental stability and human livelihoods. These semi-natural systems play a disproportionately large role in global carbon storage, with estimates indicating that they contain up to one-third of the world's terrestrial soil carbon [2, 25]. Beyond carbon sequestration, grasslands contribute to climate regulation, biodiversity preservation, hydrological functions and sustainable agricultural production, making their conservation a priority for global sustainability efforts.

Pastures represent a fundamental resource for sustainable livestock systems, providing not only the primary source of forage but also contributing to animal health, welfare, and natural feeding behaviors. They sustain biodiversity by hosting a wide range of plant and animal species, enhance soil fertility through nutrient cycling, and support carbon sequestration, thus playing a pivotal role in climate regulation. Pastures also help preserve traditional rural landscapes and cultural heritage, while ensuring the production of high-quality animal products with distinctive nutritional and organoleptic properties. By reducing the need for external feed inputs, they strengthen farm self-sufficiency and economic resilience, while simultane-

ously delivering essential ecosystem services such as erosion control, water regulation, and habitat provision.

The dynamics of carbon cycling in grasslands (also shown in figure 1.2) are governed by intricate interactions between biotic and abiotic factors. Plant diversity enhances soil organic carbon (SOC) storage through increased root biomass and microbial activity, while variables such as soil composition, climatic conditions and land-use practices further modulate sequestration efficiency [26, 2].

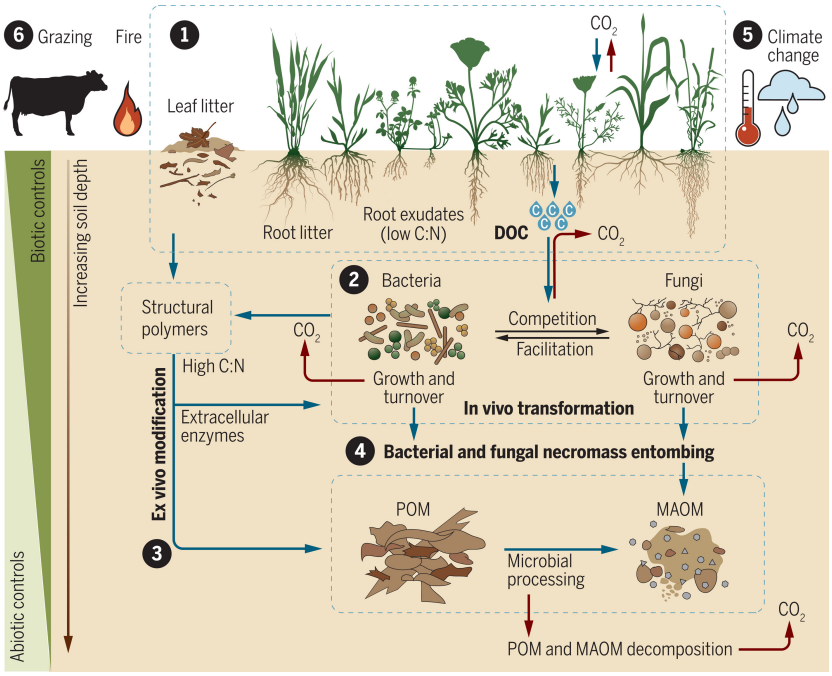

Figure 1.2: Grassland soil carbon sequestration [2]

In particular, studies in seminatural Nordic grasslands demonstrate that optimized grazing regimes can simultaneously sustain above-ground biodiversity and improve below-ground carbon retention, illustrating the potential for management strategies that reconcile ecological integrity with climate mitigation objectives [27].

Biodiversity serves as a pillar of the functionality of the grassland ecosystem. Long-term experimental research underscores that the richness of plant species exerts a more substantial influence on carbon storage than elevated CO_2 levels or nitrogen deposition. For instance, a 19-year grassland experiment revealed that biodiversity-driven carbon accumulation surpassed the effects of other global change factors, reinforcing the critical role of species diversity in ecosystem resilience [25]. Systematic reviews further identify carbon sequestration

and forage production as among the most extensively studied grassland ecosystem services, highlighting their dual agricultural and environmental significance.

Grazing management exerts a dualistic impact on grassland productivity and biodiversity. Moderate grazing intensities have been shown to promote belowground biomass and plant diversity by alleviating competitive dominance among species. Conversely, excessive or poorly managed grazing depletes vegetative cover, diminishes biodiversity, and disrupts carbon accumulation processes [28]. These findings emphasize the necessity of context-appropriate grazing systems to maintain ecosystem functionality.
Restoration and sustainable management of grasslands hold substantial potential for climate change mitigation. Biodiversity-focused restoration alone could sequester between 2.3 and 7.3 gigatonnes of CO_2 equivalent annually, with additional contributions from optimized grazing practices (148–699 megatonnes CO_2e/year) and legume integration (147 megatonnes CO_2e/year) [2]. Such nature-based solutions offer cost-effective strategies to enhance carbon storage while supporting rural economies, particularly in regions where grasslands underpin livestock production and food security. Despite their ecological and socioeconomic importance, grasslands face persistent threats from land-use conversion, afforestation, and policy neglect. In Europe and Nordic countries, grasslands remain indispensable for meat and dairy production, yet their conservation status is increasingly precarious [27]. Effective governance frameworks that integrate biodiversity and climate objectives are urgently needed to safeguard these ecosystems.

Functional diversity further improves grassland resilience by allowing complementary resource use and stabilizing productivity under environmental stressors. Variations in plant traits, such as rooting depth, leaf morphology, and flowering phenology, facilitate nutrient cycling, forage quality, and resistance to invasive species [29]. This diversity of traits buffers ecosystems against climate extremes, ensuring sustained service provision under changing conditions. Nevertheless, translating scientific insights into actionable policies remains a challenge. Socioeconomic barriers, divergent stakeholder interests, and institutional limitations often hinder the implementation of grassland conservation measures [30]. Market-based incentives and participatory governance models are increasingly advocated to embed ecosystem service valuations into land-use planning, thereby aligning ecological preservation with local livelihood needs.

In summary, grasslands occupy a unique niche in global sustainability efforts, offering unparalleled benefits for carbon management, biodiversity conservation, and agricultural resilience. Their extensive distribution, deep-rooted vegetation, and multifunctional capacity position them as indispensable natural climate solutions. However, realizing this potential

requires concerted efforts to preserve species diversity, adopt regenerative land-use practices, and institutionalize ecosystem service frameworks. Only through integrated management and policy coherence can grasslands continue to fulfill their critical role in planetary stewardship.

1.3 State of the art and thesis' goal

The measurement of pasture biomass through remote sensing has evolved from simple NDVI-based correlations to sophisticated multi-sensor approaches, yet fundamental challenges persist in operational applications. Traditional field measurements, still considered the gold standard, typically involve destructive sampling or visual estimation methods, which provide absolute biomass values but suffer from limited spatial representativeness. This limitation becomes particularly acute when attempting to validate satellite-derived estimates, as demonstrated by Dusseux et al. [31] in French grasslands, where the spatial mismatch between 20×20 cm quadrats and 10 m Sentinel-2 pixels introduced significant errors. Similar challenges were also discussed in Mediterranean dehesas by Hernández Díaz-Ambrona et al. [32], highlighting the role of vegetation indices for forage estimation in heterogeneous systems. Recent advances have demonstrated that machine learning models (e.g., Random Forest, XGBoost) trained on a combination of satellite-derived vegetation indices and meteorological data can significantly improve the accuracy of dry matter estimation in intensively managed grazing systems [33].

The search for robust vegetation indices has historically revolved around the Normalized Difference Vegetation Index (NDVI), which has proven to be a remarkably effective indicator of canopy greenness and pasture productivity across diverse ecosystems [34, 35]. NDVI is not the only index used but also indices such as the Enhanced Vegetation Index (EVI) and EVI2 [36] incorporate corrections for soil and atmospheric effects, showing improved performance in sparse or degraded grasslands. Similarly, the Soil-Adjusted Vegetation Index (SAVI) and its modifications [37] have been effective in arid systems with high bare soil fractions. The integration of long-term satellite data, particularly from the Sentinel-2 and Landsat missions, allows for the development of robust biomass estimation models that can account for seasonal and interannual variability, providing reliable information for sustainable pasture management over time [38].

Recent work has demonstrated the added value of integrating hyperspectral information. Narrowband indices such as the Normalized Difference Red Edge (NDRE) and Chlorophyll Indices [39] provide better sensitivity to nitrogen content and photosynthetic activity, com-

plementing NDVI in contexts where nutrient monitoring is essential. With Sentinel-2, red-edge bands enable indices like the Inverted Red-Edge Chlorophyll Index (IRECI), which correlate strongly with pasture biomass [40]. As shown by Chen et al. [41], the use of multi-index machine learning models leveraging Sentinel-2 improves biomass estimation accuracy compared to NDVI alone, particularly under conditions of high grazing pressure. The relationship between NDVI and biomass is fundamentally non-linear and can be effectively captured through optimized non-linear models, such as polynomial regressions, which have been shown to significantly outperform simple linear correlations for accurate biomass estimation [42].

The temporal resolution of satellite observations is another crucial factor. While MODIS data provided the first continental-scale time series of pasture dynamics [43], their coarse resolution limited their utility in heterogeneous grazing systems. Sentinel-2's 5-day revisit cycle represents a significant step forward, though even this may miss rapid biomass changes during rotational grazing events, as documented by Kallel et al. [44]. The fusion of Sentinel-2 with PlanetScope data further enhances temporal resolution to near-daily levels [45], although this comes at the cost of increased processing complexity. The value of high-temporal resolution data is profound for capturing dynamic processes. Time-series analysis of coarse-resolution data like AVHRR NDVI has been foundational for quantifying spatio-temporal patterns of agricultural productivity over large, fragmented landscapes, revealing trends and responses to environmental drivers that are invisible at shorter time scales [46].

Beyond optical data, the fusion of radar and optical observations represents a promising frontier. Sentinel-1's C-band SAR offers all-weather capability, with moderate correlations to biomass [47], while more recent approaches combining radar and optical indices in gradient boosting frameworks have shown stronger predictive power [48]. In the context of sustainable grazing, applications such as the use of vegetation indices for Voisin rational grazing management [49] demonstrate the operational relevance of these tools in adapting to climate change.

Despite technological advances, ground-based calibration remains indispensable. The dependence on field sampling continues to be a limiting factor [32], as species composition strongly modulates the relationship between spectral signals and actual forage availability. NDVI, while simple, cost-effective, and widely available, remains a cornerstone of biomass estimation due to its robustness and historical consistency across studies [34]. Its enduring relevance lies in its ability to provide a reliable baseline against which newer indices and machine learning models can be evaluated.

The integration of multi-scale observations, including UAV (Uncrewed Aerial System)-based remote sensing, now bridges the gap between field plots and satellite pixels [50].

LiDAR-equipped UAVs provide structural information that, when combined with multispectral indices, improves calibration and accounts for within-pasture heterogeneity. These hierarchical approaches underscore that while remote sensing offers unparalleled spatial and temporal coverage, its effectiveness ultimately depends on field-based validation and ecologically informed interpretation.

In this thesis, the objective is to evaluate whether satellite-derived vegetation indices, particularly the NDVI, can serve as reliable indicators for pasture biomass growth under climate change scenario. This assessment is conducted through a combined analysis of satellite data and field measurements, aiming to support more sustainable rangeland management practices. Current insurance models often rely on decadal NDVI averages to assess pasture productivity, neglecting real-time biomass dynamics, this approach may lead to inaccuracies in estimating forage availability, particularly under changing climatic conditions. By performing a statistical comparison between NDVI trends and measured biomass growth, this study seeks to determine whether incorporating direct biomass estimation could improve the precision of pasture monitoring systems.
The research is based on a mathematical model developed by the tutor who assisted in the development of this thesis project [3], which integrates soil characteristics and climate data to assess pasture data. Field campaigns provide essential calibration data, ensuring that spectral indices accurately reflect on-the-ground biomass conditions.

Chapter 2

Materials and case study

2.1 Case study

This case study focuses on the evaluation of the dynamics of pasture biomass within the Madrid Community, Spain - a region characterized by diverse climatic conditions and land use patterns along a distinct north–south gradient. The study aims to explore how grassland productivity and NDVI index are correlated.
The analysis is grounded in the SIMPAST model [3], which integrates both climatic and soil parameters to estimate potential biomass production. Historical climate records from 1975 to 2021, alongside projections up to 2100 under four Shared Socioeconomic Pathways (SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5), provide the foundation for simulations. Inputs such as daily temperature extremes, precipitation, and soil water capacity are used to assess net primary productivity and water use efficiency across the study areas, as shown in figure 2.1. To capture the environmental heterogeneity of the region three study zones with different climatic and edaphic characteristics were defined in the Community of Madrid, within each zone, 5x5 km grids were established.

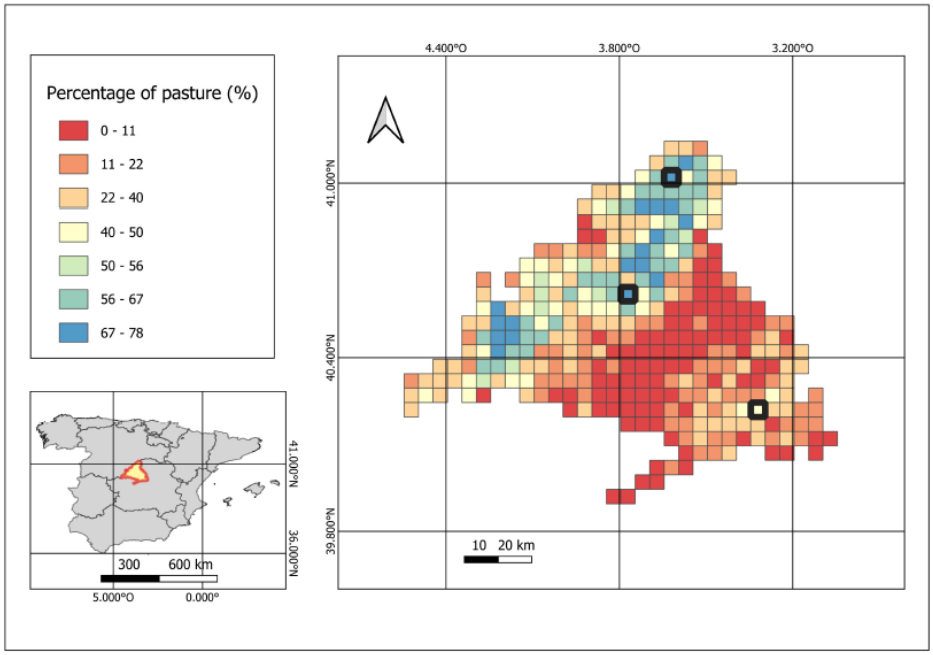

Figure 2.1: Geographic distribution of pasture types (tree-dominated, shrub-dominated, and herbaceous pastures) within the study area [3].

One of the limitation to not include other regions was the proximity to soil pit available information, as needed for the biomass model. The selection process utilized satellite imagery and field verification to ensure the representativeness of each site. The zones are:

- Buitrago del Lozoya, northern zone (12280): characterized by a mountain climate with lower temperatures and higher annual precipitation. In this region, water availability plays a key role in the dynamics of herbaceous biomass, influencing its seasonality and the recovery capacity of pastures after drought periods.

Figure 2.2: Traditional field survey in Buitrago del Lozoya, May 2025.

- Colmenar Viejo, central zone (11033): located in the transition between dehesa (wooded pastureland) and grassland systems, this area features moderately fertile soils and an intermediate climatic regime. Research in this zone will assess how interactions between climate, soil and management influence forage productivity and quality.

Figure 2.3: Traditional field survey in Colmenar Viejo, May 2025.

- Tielmes, southern zone (9862): Characterized by a drier climate and higher agricultural pressure, with shallower soils and more limited water availability. In this context, the results of this study could be used to analyze how aridity conditions affect herbaceous biomass and what management strategies could mitigate drought effects.

Figure 2.4: Traditional field survey in Tielmes, May 2025.

2.2 Data utilized

2.2.1 NDVI

The use of satellite imagery in agriculture has enabled farmers to monitor their fields efficiently, assess crop health, detect pests and diseases, and evaluate soil moisture levels—all

without the need for extensive ground-based data collection.
In this context, numerous spectral indices have been developed, including several specifically designed for vegetation health monitoring, such as NDVI.
The NDVI is a widely used metric to assess vegetation health and density [51]. It quantifies photosynthetic activity by analyzing the contrast between near-infrared (NIR) and red light reflectance: healthy vegetation strongly reflects NIR while absorbing red light. The NDVI is calculated as follows:

$$NDVI = \frac{NIR - Red}{NIR + Red}[51] \tag{2.1}$$

The result of this formula 2.1 ranges from -1 to +1, with negative values indicating cloud cover or water presence, values close to 0 which represent bare soil and values greater than 0 ranging from sparse vegetation (0.1 - 0.5) to dense green vegetation (0.6 and above). In fact, as an example, the figure below shows that stressed plants have a difference between near-infrared and red slower compared to healthy ones.

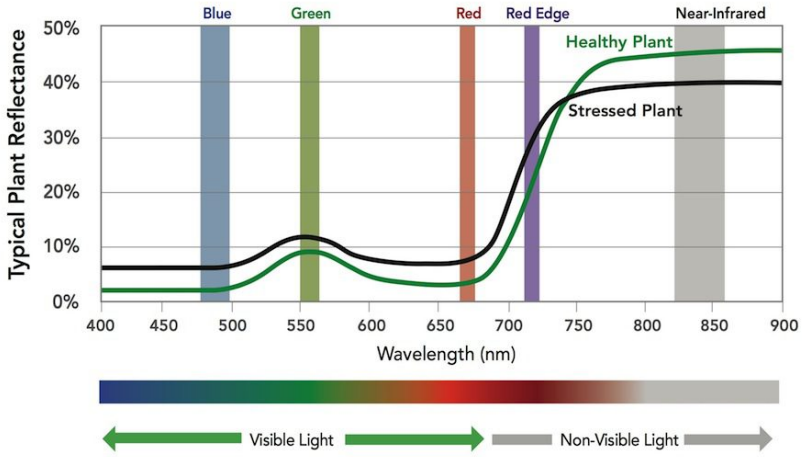

Figure 2.5: Spectral resolution healthy and stressed vegetation [52]

The NDVI data used in this study are retrieved from NASA's tool APPEARS (Application for Extracting and Exploring Analysis Ready Samples) [53], a tool designed to facilitate the efficient processing of geospatial data. Among the various datasets available, it has been used MODIS (moderated resolution imaging spectroradiometer); it is an instrument aboard the Terra and Aqua satellites which captures global data across 36 spectral bands, with complete coverage every one to two days. Thanks to its wide spectral range, MODIS supports research in diverse fields such as vegetation analysis, land use changes, oceanography, sea surface temperature monitoring, and cloud studies. Additionally, it plays a crucial role in tracking wildfires, natural disasters, and even oil spills. One of MODIS's key strengths is its

rapid data accessibility. Real-time data streams are available via direct broadcast stations that receive transmissions straight from the satellite. For near real-time access, NASA's Land, Atmosphere Near Real-time Capability for EOS (LANCE) system delivers select MODIS products within just three hours of observation.

Our object of interest can be recovered by the MOD09Q1 version 6.1 product. It provides surface spectral reflectance estimates for bands 1 and 2 of the MODIS on board the Terra satellite, with a spatial resolution of 250 meters. These data are corrected for atmospheric effects, such as gases, aerosols and Rayleigh scattering, and include two additional quality levels. For each pixel, a single value is selected from all available acquisitions over the 8-day composite period. The pixel selection is based on specific criteria, including the absence of clouds and the solar zenith angle.
As shown in the table the MOD09Q1 contains different layers but the NDVI can be calculated by the first two.

Science Data Sets (HDF Layers (4))	Units	Data Type	Fill Value	Valid Range	Scale Factor
sur_refl_b01:250m Surface Reflectance Band 1 (620-670 nm)	Reflectance	16-bit signed integer	-28672	-100 - 16000	0.0001
sur_refl_b02: 250m Surface Reflectance Band 2 (841-876 nm)	Reflectance	16-bit signed integer	-28672	-100 - 16000	0.0001
sur_refl_state_250m: 250m Reflectance State QA (*see Table 13*)	Bit Field	16-bit unsigned integer	65535	0 - 57343	NA
sur_refl_qc_250m: 250m Reflectance Band Quality (*see Table 9*)	Bit Field	16-bit unsigned integer	65535	0 - 32767	NA

Table 2.1: MOD09Q1 layers

2.2.2 Climate models

Climate models, are mathematical models capable of simulating the complex interactions among key climate drivers, including *atmospheric dynamics, ocean circulation, land surface processes, and cryospheric changes* (such as ice sheet and glacier evolution) [54]. Among the different models there are the general circulations models (GCM). They consist of separating the Earth's surface into a three-dimensional grid of cells connected to each other; in fact the results of processes modeled in each cell are passed to neighboring ones to model the exchange of matter and energy over time. The size of the grid cell defines the resolution of the model: the smaller the size of the grid cells, the higher the level of detail in the model.

Many critical physical processes - such as cloud formation, radiative transfer, precipitation mechanisms, and atmospheric and oceanic turbulence - occur at scales finer than the

model grid can directly represent. To account for these sub-grid-scale phenomena, GCMs incorporate parameterizations: mathematically formulated approximations derived from extensive theoretical and observational studies. These parameterizations essentially function as specialized "sub-models" within the broader climate model framework, enabling the representation of small-scale processes' impacts on the global climate system.

The parameterization approach reflects the necessary compromise between physical realism and computational feasibility, allowing climate models to simulate Earth's complex climate system while remaining computationally tractable. For example, one single model can contain four different sub-models which describe, respectively, ocean, land, atmosphere, and sea ice.

To assess the climate variability of the region and its impact on pasture productivity, daily data on temperature (maximum, minimum, and mean), precipitation, solar radiation, and potential evapotranspiration (ET0) have been collected. The data are sourced by AdapteCCa [55], it is a platform from which it is possible to know, visualize and download past and projected climate data. AdapteCCa include local *weather stations, IPCC (AR6) scenarios SSP*, (Shared Socioeconomic Pathways, SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5) and *climate models derived from CMIP6*. The last two will be briefly described below.

Climate scenarios are indispensable in climate change research, offering a framework to assess the long-term consequences of human activities on Earth's systems. By simulating potential futures, they enable researchers to evaluate how near-term decisions—whether policy-driven, technological, or behavioral—could alter environmental outcomes decades ahead. Among these tools, Shared Socioeconomic Pathways (SSPs), which were officially adopted by the IPCC starting with the Sixth Assessment Report (AR6, 2021-2023) [56] stand out as a comprehensive suite of scenarios that extend beyond traditional emission projections to incorporate the complex interplay of societal, economic, and political factors shaping climate vulnerability and resilience.

Developed as an advancement over earlier climate modeling frameworks, the SSPs provide structured narratives alongside quantitative projections, outlining how global society might evolve throughout the 21st century. Unlike purely emissions-focused approaches, they emphasize the underlying drivers of climate change, including governance, inequality, technological innovation, and resource management. This multidimensional perspective allows researchers to explore how divergent development trajectories could either facilitate or hinder climate mitigation and adaptation efforts.

The SSP framework comprises five core pathways, each representing a plausible future with

distinct socioeconomic conditions and emission levels. Below is a detailed breakdown of their key characteristics:

- SSP1 ("Sustainability"): This pathway envisions a world increasingly committed to environmental protection, equity, and global cooperation. Societies prioritize education, clean energy innovation, and sustainable consumption, leading to low greenhouse gas emissions and high adaptive capacity. Technological advances are widely shared, reducing material intensity and fostering resilience.

- SSP2 ("Middle of the Road"): Reflecting historical trends, this scenario assumes uneven progress without transformative policy shifts. While some regions achieve moderate development, challenges like inequality, resource scarcity, and climate vulnerability persist globally. Emissions and adaptation efforts remain intermediate, mirroring current trajectories.

- SSP3 ("Regional Rivalry"): A fragmented world marked by geopolitical tensions, weak institutions, and limited cooperation. High population growth in low-income countries coincides with underinvestment in education and technology. Environmental degradation accelerates, while political and economic barriers stifle both mitigation and adaptation.

- SSP4 ("Inequality"): This pathway highlights a future of stark disparities. Technological progress and economic growth benefit a privileged minority, leaving marginalized populations exposed to climate risks. Access to adaptation resources is highly unequal, creating pockets of resilience amid widespread vulnerability.

- SSP5 ("Fossil-Fueled Development"): Driven by rapid economic growth and energy-intensive lifestyles, this scenario prioritizes industrialization over sustainability. Despite high adaptive capacity due to technological and financial resources, delayed mitigation locks in high emissions, exacerbating long-term climate impacts.

These pathways are not deterministic forecasts but rather exploratory tools to assess how societal choices—from governance to technology adoption—interact with climate systems. Each SSP combines qualitative narratives (e.g., political cooperation, equity trends) with quantitative data (e.g., GDP, population, energy demand) to feed into integrated assessment models (IAMs). The Intergovernmental Panel on Climate Change (IPCC) routinely employs SSPs in its reports to analyze risks and inform policy recommendations. By bridging socioeconomic dynamics with biophysical impacts, the SSP framework helps decision-makers weigh trade-offs between development priorities and climate action.

The true value of the SSP framework lies in its flexibility and its ability to be combined with different levels of greenhouse gas forcing. When merged with Representative Concentration Pathways (RCPs) or similar radiative forcing targets, researchers can evaluate not just what kind of climate future might result from certain emissions levels, but also how societal characteristics affect both the likelihood of those outcomes and our ability to manage them.
The integrated SSP-RCP framework (alternatively designated as SSPX-Y scenarios) represents a methodological advancement in climate projection by coupling the foundational Shared Socioeconomic Pathways (SSPs) with the radiative forcing targets established by the Representative Concentration Pathways (RCPs) from the IPCC's Fifth Assessment Report (AR5).

This hybrid approach systematically applies the climate forcing constraints of RCP scenarios to the socioeconomic trajectories outlined in the SSP baseline scenarios. The resulting matrix produces five principal SSP-RCP scenario combinations:

	Near term, 2021–2040		Mid-term, 2041–2060		Long term, 2081–2100	
Scenario	Best estimate (°C)	*Very likely* range (°C)	Best estimate (°C)	*Very likely* range (°C)	Best estimate (°C)	*Very likely* range (°C)
SSP1-1.9	1.5	1.2 to 1.7	1.6	1.2 to 2.0	1.4	1.0 to 1.8
SSP1-2.6	1.5	1.2 to 1.8	1.7	1.3 to 2.2	1.8	1.3 to 2.4
SSP2-4.5	1.5	1.2 to 1.8	2.0	1.6 to 2.5	2.7	2.1 to 3.5
SSP3-7.0	1.5	1.2 to 1.8	2.1	1.7 to 2.6	3.6	2.8 to 4.6
SSP5-8.5	1.6	1.3 to 1.9	2.4	1.9 to 3.0	4.4	3.3 to 5.7

Table 2.2: Description of scenarios IPCC

The Coupled Model Intercomparison Project Phase 6 (CMIP6), instead, is a major international effort coordinated by the World Climate Research Program (WCRP) to advance global climate modeling [57]. It serves as a foundation for understanding Earth's climate system by bringing together the world's leading climate research institutions to develop, compare, and refine climate models. These models simulate interactions between the atmosphere, oceans, land surface, and ice to project future climate conditions under different scenarios. Within CMIP6, climate models like CNRM-ESM2-1 produce simulations across multiple SSPs to evaluate divergent climate futures under varying socioeconomic and emission conditions. Six different models have been used to have a better results and compare them, the models are:

1. **ACCESS CM2**
   Access is a family of related models, configured for specific applications, to meet operational and research needs from weather forecasting to climate projections created by CSIRO (Commonwealth Scientific and Industrial Research Organisation) [58].

2. **CNRM ESM2**

   The CNRM-ESM2-1 represents the second-generation Earth system model developed by the CNRM/CERFACS modeling consortium, was released in 2017 by Centre National de Recherches Meteorologique [59]. Building upon the physical-dynamical framework of its predecessor, the CNRM-CM6-1 coupled climate model, this advanced system incorporates comprehensive interactions between physical atmospheric-oceanic processes and biogeochemical components.

3. **EC EARTH3 VEG**

   The EC-Earth modeling framework originated as a coupled atmosphere-ocean general circulation model (AOGCM), derived from the seasonal forecasting system developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) [60]. The EC-Earth3 represents the third generation of this modeling system and serves as the foundational configuration for standard-resolution atmosphere-ocean physical modeling. However, EC-Earth3 is not a single compact model but rather a suite of eight distinct configurations, each tailored for specific scientific applications. These variants were employed by the EC-Earth consortium in its contributions to the sixth phase of the Climate Model Intercomparison Project (CMIP6). Among these, the "ec-earth3-veg" configuration stands out as one of the actively used versions, incorporating dynamic vegetation processes to better represent land-atmosphere interactions.

4. **MIROC6**

   The MIROC6 climate model (Model for Interdisciplinary Research on Climate, version 6) was developed by a Japanese consortium led by the University of Tokyo, in collaboration with the National Institute for Environmental Studies (NIES) and the Japan Agency for Marine-Earth Science and Technology (JAMSTEC) [61]. MIROC6 incorporates advanced components for representing the climate system, including high-resolution atmospheric and oceanic dynamics, biogeochemical processes, and land-atmosphere interactions. The model is designed to improve the simulation of complex climate phenomena, such as interannual variability (e.g., ENSO) and long-term projections under different anthropogenic forcing scenarios.

5. **MPI ESM1**

   The MPI-ESM1-2 (Max Planck Institute Earth System Model version 1-2) represents a state-of-the-art climate modeling system developed by Germany's Max Planck Institute for Meteorology (MPI-M) [62]. Notably, MPI-ESM1-2 exists in multiple configurations, with the LR (low resolution) and HR (high resolution) versions being particularly widely used in CMIP6 experiments.

6. **MRI ESM2**

   Developed by Japan's Meteorological Research Institute (MRI), the MRI-ESM2.0 builds upon the institution's long-standing expertise in climate modeling while introducing significant advancements in Earth system representation [63].

In the next table are shown the different characteristics of each model; they all contain the four different submodels already mentioned, where it is specified the name of the mathematical model used to characterized the corresponed climate driver and the resolution. It is possible to see for example the CNRM-ESM2 and EC-EARTH3-VEG use the same model for ocean or that in MIRCO6 it is used the same one for ocean and sea ice.

Model	Sub Models							
	Ocean		Atmopsphere		Land		Sea ice	
	Name	Resoultion (km)	Name	Resoultion(km)	Name	Resoultion (km)	Name	Resoultion (km)
ACCESS CM2	Access-om 2	100	MetUM-HadGEM3 -GA7.1	250	CABLE2.5	250	CICE5.1.2	100
CNRM ESM2	Nemo 3.6	100	Arpege 6.3	250	Surfex 8.0c	250	Gelato 6.1	100
EC EARTH3 VEG	Nemo 3.6	100	IFS cy36r4	100	HTESSEL	100	LIM3	100
MIROC6	Coco 4.9	100	CCSR AGCM	250	MATSIRO6.0	250	COCO4.9	100
MPI ESM1 HR	Mpiom 1.63	50	ECHAM6.3	100	JSBACH3.20	100	unnamed	50
MRI ESM2	Mri.com 4.4	100	MRI-AGCM3.5	100	HAL 1.0	100	MRI.COM4.4	100

Table 2.3: Model characteristics

Chapter 3

Methodology

This section outlines the methodology employed to establish the correlation between pastures and NDVI. First, will be presented the data in raw form. Subsequently, the statistical approach used for the study will be presented; the software tools, functions, and formulas applied in the analysis will also be discussed. Finally, the complete workflow will be described, providing a comprehensive overview of the entire analytical process.

3.1 Data analysis and processing

3.1.1 Pre-analysis of the dataset

Each study area was assigned a unique identifier code: *Buitrago del Lozoya* North (12280), *Colmenar Viejo* central (11033), and *Tielmes* South (9862). For each area, an Excel file containing NDVI data and six additional Excel files, one for each climate model, were provided. Since the data structure remains consistent across all areas and models, this section will describe the organization for a single area and a single model; the same approach will be applied in the Workflow section.

The NDVI dataset included the following fields: date, month, year, ID, mean NDVI, median NDVI, SGolay N5 P2, and SGolay N7 P2. The ID column corresponds to the study area, labeled as described above. The columns SGolay N5 and N7 represent corrected NDVI values obtained through the Savitzky-Golay (SGolay) filter. The Savitzky-Golay (S-Golay) filter is a digital signal processing method widely employed for data smoothing and numerical differentiation while preserving essential signal features, such as peak amplitudes and widths, which are often distorted by conventional moving-average filters [64]. The mathematical foundation of the S-Golay filter relies on constructing a polynomial of degree *p*

over a window of *2m+1* points centered at each data sample; the one used in our case uses a second polynomial degree and a window of 7 points, the central observation, the previous three and the following three. The observation period spans from February 18, 2000, to April 6, 2024, with data collected at 8-day intervals. For each interval, the best-quality observation was selected and subsequently refined using the SGolay correction.

Regarding the climate models, the six provided files, each corresponding to a different model, contained data for all three study areas (identified by the ID column) and the four IPCC scenarios. The datasets included records of day, month, year, soil characteristics, IPCC scenario, mean/min/max temperature, and precipitation. Unlike the NDVI data, which were aggregated at 8-day intervals, the climate model observations were recorded daily. The temporal coverage of the climate models extends from 1975 to 2100, with observed data from 1975 to 2021 and projected climate simulations, aligned with the scenarios outlined in Chapter 2, from 2021 to 2100.

Based on this explanation, we have to deal with time series; to do that we use box-plots. Box-plots, also known as box-and-whisker plots, are a fundamental graphical tool in descriptive statistical analysis, as they summarize the distribution of a dataset through robust indicators such as the median, quartiles, and potential outliers.
As shown in figure 3.1, the boxplot structure consists of a central box and two lines, called "whiskers," extending from it. The box represents the interquartile range (IQR), bounded by the first quartile (Q1) and the third quartile (Q3), with an internal line marking the median (Q2). The whiskers typically extend up to 1.5 times the IQR above and below the quartiles, though alternative conventions may define them differently. Data points falling beyond this range are considered potential outliers and are represented as individual points or asterisks.

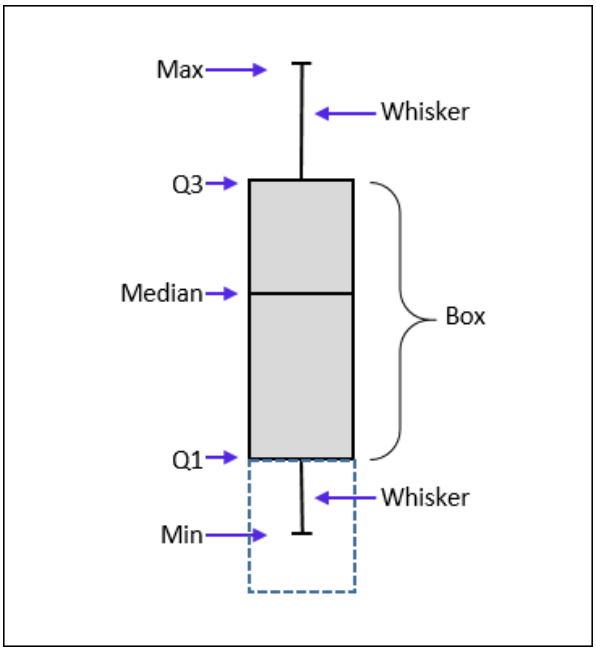

Figure 3.1: Box plot, graphical representation

The primary functions of the boxplot are to analyze data variability, identify skewness, and detect outliers. Unlike other graphical tools, such as histograms, the boxplot provides a robust distribution summary due to its low sensitivity to extreme values. This makes it particularly useful for comparing multiple datasets, allowing for quick assessment of differences in central tendency and dispersion.

When applied to time series, box-plots gain particular significance by enabling the examination of how a variable's distribution evolves over time. This is achieved by segmenting the data into defined temporal intervals—such as months, seasons, or years—and representing the distributional characteristics for each period. This approach proves especially valuable for identifying trends, cyclical patterns, or structural shifts, as it extends beyond mere central tendency analysis to assess dispersion, skewness, and the presence of outliers in each interval.

3.1.2 Regression

Regression analysis stands as one of the foundational methodologies in statistics and machine learning, serving to model and examine relationships between a dependent variable (or outcome) and one or more independent variables (predictors) [65]. At its core, this approach seeks to understand how variations in predictors influence the outcome, enabling both prediction and inference. Its formulation is here represented:

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{3.1}$$

Where:

- y : dependent variable;
- x : independent variable;
- β_0 : intercept;
- β_1 : angular coefficient;
- ε : residual.

While standard regression assumes a uniform relationship across all data points, real-world phenomena often exhibit structural changes or threshold effects - precisely where piecewise regression becomes indispensable.
Piecewise regression, also called segmented regression, extends conventional regression analysis by accommodating situations where the relationship between variables is non linear and changes at specific intervals of the independent variable. Rather than imposing a single parametric form across the entire data range, this method identifies breakpoints (or knots) where the nature of the association fundamentally shifts, fitting distinct regression models to each homogeneous segment; breakpoints can be either manually selected or automatically searched. For the presence of these breakpoints the formulation changes as follow:

$$y = \beta_0 + \beta_1 x + \beta_2 (x - c_1)^+ + \beta_3 (x - c_2)^+ + \dots \tag{3.2}$$

Where:

- y : dependent variable;
- x : independent variable;
- β_0 : intercept;
- $\beta_1 x$: first linear part with slope β_1;
- $\beta_2 (x - c_1)^+$: change in slope after the first breakpoint;
- $\beta_3 (x - c_2)^+$: change in slope after the second breakpoint.

One of the purpose of this thesis is to compare the different slopes (so β_0, β_1, etc.) among all the datasets, between NDVI and the different models which predict pasture and correlate them. To evaluate the goodness of the regression we use different parameters: R-square (multiple and adjusted), standard error and P value.

R-squared (R^2) is one of the most widely used metrics, quantifying the proportion of variance in the dependent variable accounted for by the independent variables [65]. However, distinctions exist between multiple R-squared and adjusted R-squared, each serving specific purposes in model assessment.
Multiple R-square ranges from 0 to 1, it provides an immediate assessment of how well the regression model captures the observed data patterns. The formula is here described:

$$R^2_{multiple} = 1 - \frac{SSE}{SST} \tag{3.3}$$

Where:

- SSE: Residual Sum of Squares. It represents the sum of the squares of the differences between the observed values and the estimated values, low SSE means that the models fit in a good way the data.
- SST: Total Sum of Squares. It is the sum of the squares of the differences between the observed values and their mean.

However, a critical limitation of R-squared is that it invariably increases when additional predictors are included in the model, regardless of their actual contribution. This property can lead to model overfitting, where the model appears to perform well on the training data but fails to generalize to new data.

Adjusted R-squared addresses this limitation by incorporating a penalty for the number of predictors relative to the sample size [65]. Unlike R-squared, which can only increase or stay the same with added variables, adjusted R-squared may decrease when irrelevant predictors are included. This adjustment makes it particularly valuable for comparing models with different numbers of predictors or when working with limited data. The penalty term ensures that only meaningful variables that genuinely improve the model's explanatory power contribute positively to the metric.

$$R^2_{adjusted} = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} \tag{3.4}$$

Where:

- n: sample size;
- k: number of estimated. regression coefficients excluding the intercept β_0

The distinction between these two measures has important practical implications. While R-squared provides an initial assessment of model fit, adjusted R-squared offers a more rigorous evaluation by accounting for model complexity.

The *standard error (SE)* is a fundamental measure in regression analysis that quantifies the precision of an estimated parameter, such as a regression coefficient or the model's predicted values [65]. It represents the estimated standard deviation of the sampling distribution of the statistic, reflecting how much the estimated value would vary across different samples drawn from the same population. A lower standard error indicates greater precision in the estimate, while a higher standard error suggests greater uncertainty.

The *p-value*, or probability value, is a fundamental statistical measure that quantifies the evidence against a null hypothesis [65]. It represents the probability of obtaining a result at least as extreme as the one actually observed in the sample data, assuming the null hypothesis is true.
In research practice, the p-value serves as a decision-making tool. Researchers establish a predetermined significance level, typically 0.05 (called α), and compare the obtained p-value with this threshold. When the p-value falls below α, we conclude there is sufficient statistical evidence to reject the null hypothesis in favor of the alternative. Conversely, a p-value above the threshold indicates the data do not provide enough evidence against the null hypothesis. Proper interpretation of p-values requires careful attention to several crucial aspects. First and foremost, it's essential to understand that the p-value does not represent the probability that the null hypothesis is true, nor the probability that the alternative hypothesis is false. Moreover, the p-value does not provide information about the practical importance or magnitude of the observed effect, but only about its statistical significance.
In regression analysis, for example, the p-value associated with each coefficient allows us to evaluate whether the corresponding predictor variable has a statistically significant effect on the response variable. However, it's important to remember that statistical significance doesn't automatically imply practical relevance. A small p-value might result from a very large sample detecting negligible differences, while a larger p-value in an underpowered study might mask potentially important effects.

Depending on the value of this parameter it is possible to associate a number of stars to the result which indicate the statistical significance of a test result, more in details:

- p value > 0,05: no symbol;
- p value <= 0,05: *;
- p value <= 0.01: **;
- p value <= 0,001: ***.

3.1.3 Correlation

Correlation between two variables, X and Y, exists when they exhibit systematic covariation—that is, when changes in X correspond to predictable changes in Y (and vice versa). If this covariation follows a straight-line pattern, the relationship is termed linear correlation; if it follows a curved pattern, it is referred to as nonlinear correlation. The absence of correlation occurs when Y fluctuates randomly with respect to X (or vice versa), showing no discernible pattern. A linear correlation is considered positive if increases in X are associated with increases in Y, whereas it is negative if increases in X correspond to decreases in Y.
To quantify the strength and direction of a linear relationship, correlation coefficients are employed. Various linear correlation coefficients exist, each suited to different measurement scales, differing in computational formulas but sharing identical interpretation principles. These coefficients range between -1 and +1, conveying two key pieces of information: the strength of the relationship (indicated by the absolute value) and its direction (indicated by the sign). For example in the figure below 3.2 coefficient closer to ±1 reflects a stronger linear association, whereas values near zero suggest a weaker relationship. For instance, a coefficient of -0.7 denotes a strong negative correlation, while 0.3 indicates a weak positive correlation. A perfect linear relationship ($r = \pm 1$) implies that a unit change in X corresponds to a constant proportional change in Y. However, a coefficient near zero does not necessarily imply statistical independence; rather, it suggests the absence of a linear association, as X and Y may still be related through a nonlinear function, which standard linear correlation measures cannot detect. In such cases, it is more precise to state that no linear correlation exists.

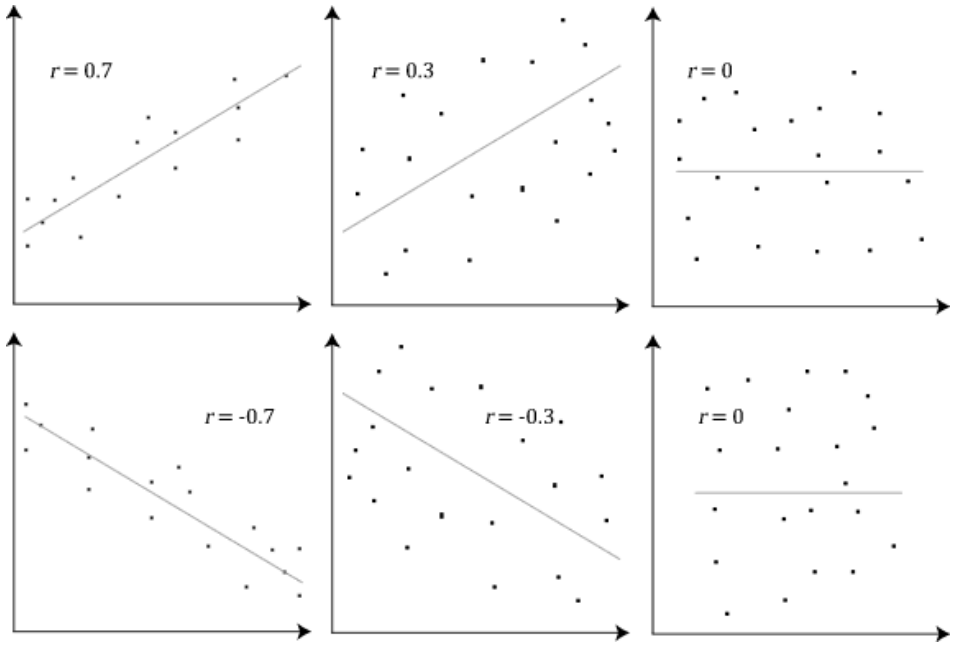

Figure 3.2: Pearson correlation values

Among these measures, Pearson's correlation coefficient is a parametric statistic assessing the degree of linear dependence between two continuous variables under the assumption of bivariate normality [66]. It yields values between -1 and +1 because its formula 3.5 derives from the standardization of the covariance with respect to the variability of the individual variables. Since Pearson's method relies on distributional assumptions, particularly normality and the fact that the variance of the error is constant whithin a model, its validity is contingent upon meeting these conditions. Consequently, it is most reliable when applied to data conforming to these requirements, as violations may lead to biased or misleading inferences.

$$r = \frac{COV(X,Y)}{\sigma_x \sigma_y} \tag{3.5}$$

Where:

- COV(X,Y): covariance of the variables X and Y;
- σ_x: standard deviation of X;
- σ_y: standard deviation of Y.

3.2 Software

The software that has been used for the statistical analysis is R Studio. R is a language and environment for statistical computing and graphics [67]. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.
R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible. In the next section will be now explained the functions used to elaborate the dataset.

3.2.1 Functions and formulas

aggregate: Splits the data into subsets, computes summary statistics for each, and returns the result in a convenient form. The usage is aggregate(x, dataset, FUN), where "x" is the *formula* (such as y $\sim$ x) where the y variables are numeric data to be split into groups according to the grouping x variables, "dataset" is the dataset from which the data are used and "FUN" is a function to compute the summary statistics which can be applied to all data subsets (in our case will be the mean function but it will be explained better in the next paragraph).

predict: generic function for predictions from the results of various model fitting functions. The usage is predict(object) where "object" stands for a model object for which prediction is desired.

lm: is used to fit linear models, including multivariate ones. It can be used to carry out regression, single stratum analysis of variance and analysis of covariance. The usage is lm(formula, data), both already explained.
Models for lm are specified symbolically. A typical model has the form response $\sim$ terms where response is the (numeric) response vector and terms is a series of terms which specifies a linear predictor for response. A terms specification of the form first + second indicates all the terms in first together with all the terms in second with duplicates removed. A specification of the form first:second indicates the set of terms obtained by taking the interactions of all terms in first with all terms in second. The specification first*second indicates the cross of first and second. This is the same as first + second + first: second. For this type of analysis it has been use the first one.

pmin and **pmax**: take one or more vectors as arguments, recycle them to common length and return a single vector giving the 'parallel' maxima (or minima) of the argument vectors.

cbind: take a sequence of vector, matrix or data-frame arguments and combine by columns or rows, respectively. These are generic functions with methods for other R classes.

cor: provides the correlation of two vectors. The usage is cor(x,y, use="complete.obs") where x and y are the vectors to be correlated and use complete.obs to avoid NA values.

find_best_segmented <- function(data, y_var, x_var, max_breakpoints = n): this function is designed to systematically identify breakpoints in relationships between variables. The process begins by dynamically creating a regression formula and fitting a baseline linear model, which serves as the foundation for subsequent segmented analyses. The function extracts the relevant variables' values, computes the total range of the independent variable and the total sum of squares—key quantities for evaluating model fit through R-squared. This metric is manually calculated using the traditional statistical definition, ensuring more robust and consistent results compared to automated implementations.

The algorithm employs an incremental strategy for breakpoint detection, sequentially testing models with an increasing number of breakpoints up to a user-specified maximum (max_break point). This parameter is particularly relevant in specific applications, such as seasonal data analysis where a known number of breakpoints might be expected (e.g., 4 for quarterly data). For each tested configuration, breakpoints are initially placed at quantiles of the independent variable's distribution while maintaining a minimum spacing of 15% of the total domain width—a critical constraint to prevent overfitting and ensure statistical significance of the identified segments. The core segmentation relies on the segmented package, which implements an iterative algorithm. Starting from initial breakpoint positions (psi_init), the method progressively adjusts their locations to maximize alignment between linear segments and observed data until convergence (when further improvements become negligible). The entire process is wrapped in a robust error-handling mechanism using tryCatch, ensuring analysis continuity even with problematic configurations by returning NULL instead of failing.

Model estimation occurs through the primary call *segmented(mod, seg.Z = as.formula(paste("∼", x_var)), psi = psi_init, control = seg.control(display = FALSE))*, where both the base model (mod) containing the variable relationship and the target independent variable (x_var) for breakpoint detection are specified. This bidirectional approach maintains clear overall model specification while focusing structural change analysis on the variable of

interest.
For each successfully estimated model, the function performs detailed result analysis—computing not just the global R-squared but also segment-specific slopes and their statistical significance. These metrics are systematically compared across alternative models, with variables storing the best configuration being progressively updated. The selection criteria intelligently balance goodness-of-fit (prioritizing models with R-squared improvements >0.02) and model complexity. The final output provides a comprehensive summary, including the optimal model, breakpoint locations, and detailed segment characteristics with transition analyses.

correlation_function <- function(df1, df2, col1, col2, group_col1="new_label", group_col2="label", groups): is the prototype of the function developed to analyze the correlation between pairs of variables from two different data frames, while maintaining alignment based on shared groups or categories.
The function takes the list of groups of interest (the groups parameter). For each group, it searches the first data frame (df1) for all rows belonging to that group (using the column specified by group_col1) and extracts the values from the column you want to correlate (col1). At the same time, it performs the same operation in the second data frame (df2), selecting rows from the same group (this time using group_col2) and the values from col2. Then, for each group, the function computes the correlation coefficient between these two sets of values. The use = "complete.obs" argument ensures that only complete cases are used, automatically excluding missing values from both datasets.

The final output is a set of correlation coefficients—one for each group—telling you how linearly related the two selected variables are within each category. This is particularly useful when you suspect that the relationship between two variables might vary across different groups, allowing you to systematically test for these differences.

```r
#function
correlation_function <- function(df1, df2, col1, col2, group_col1="new_label",
                                 group_col2="label", groups){
  sapply(groups, function(g) {
    cor(
      df1[df1[[group_col1]] %in% g, col1],
      df2[df2[[group_col2]] %in% g, col2],
      use = "complete.obs"
    )
  })
}
```

Figure 3.3: Code correlation function

3.3 Workflow

In the picture below (figure 3.4) a graphical workflow is presented to simplify the lecture and have a better illustration of this section.

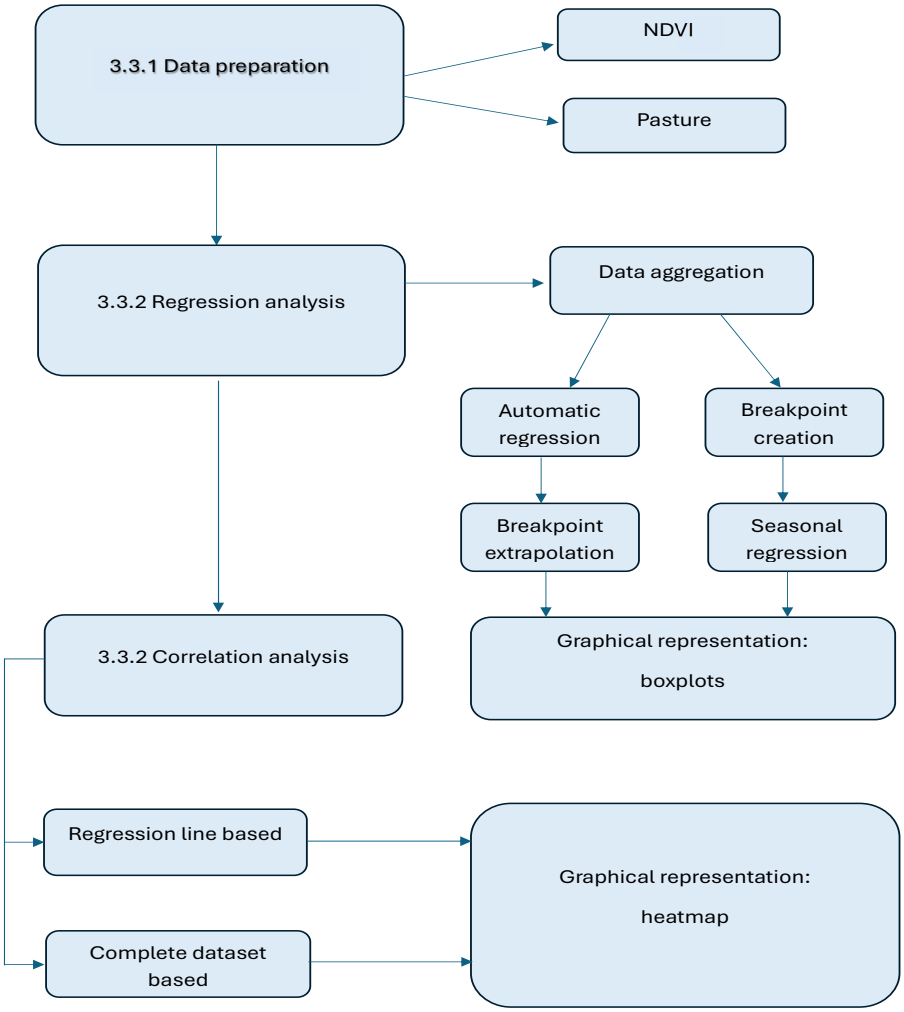

Figure 3.4: Graphical workflow

3.3.1 Data preparation

The initial phase of the analysis focused on NDVI data, organized into one single file and divided in three for each study area to streamline calculations. Each file contains a unique identifier code in the "ID" column. As said, the hydrological calendar (starting in September) was adopted to align with the biomass growth cycle, as vegetation typically begins to recover after the summer stagnation period, ensuring observations commence following a stationary phase.

To examine the correlation between the parameters of interest, a regression analysis was performed. To do that preliminary boxplots were generated (described in Section 3.1.1), with the x-axis representing the observations days of each year and the y-axis displaying the corresponding NDVI values. For this purpose, the new Excel files for each area were structured to include data from 5 September 2000 to 6 April 2024, excluding records from 2 February 2000 to August 2000, as they pertain to the preceding hydrological year. This approach yielded a 25-year observation period with recurring dates (e.g., 5 September or 13 September across different years). To facilitate grouping, a "new label" column was introduced, assigning values from 1 to 46 for each 8-day interval within the hydrological year. September start dates were labeled as 1, while late August dates were labeled as 46. Leap years were accommodated by maintaining the same label for the corresponding 8-day composite observation, for our interests it does not represent a problem cause the NDVI value is the best out of 8 days of observations.
Given that pasture data were provided in cumulative form, NDVI values were similarly transformed from discrete to cumulative. A new column, labeled "Cum Sgolay n7 from sept," was created, where the first cell matched the SGolay-filtered value for that row, and subsequent cells represented the cumulative sum of the preceding values. This was repeated annually, ensuring that each label "1" corresponded to the initial SGolay value of the hydrological year.

Regarding pasture data, they were incorportaed with the climate ones cause they are retrieved from the model [3] in six separate files, corresponding to different climate models, each containing data for the three study areas under four scenarios. To simplify processing, new Excel files were generated for each model and area, retaining only data from 1 September 2000 to 11 April 2024 under the SSP245 scenario (18 files total, 6 for each area). This alignment was necessary because NDVI data were limited to this period, and the SSP245 scenario most closely followed historical trends. Unlike NDVI data, which were collected every 8 days, climatic data were recorded daily. To compare the datasets, an 8-day moving average was computed for the column containing the data related to the biomass called

"pasture corregido" which is the one of our interest. Starting from 1 September, the average of each 8-day interval was calculated (e.g., days 1–8, 9–16, etc.), with care taken to handle year transitions. These averaged values were then associated with the corresponding NDVI reference dates. A dedicated "id day" column was created in Excel, numbering rows in 8-day blocks (1 to 1086), enabling the subsequent calculation of averages in R for all dates sharing the same "id day" value. Excel microsoft sofware was used for the data preparation.

3.3.2 Regression and correlation analysis

In the R environment, three distinct files were created for each area under investigation. Initially, the necessary packages—namely readxl, ggplot2, segmented, and corrplot—were imported to facilitate the reading of Excel files (.xls format), generate graphical representations, perform regression analyses, and compute correlations, respectively.
The NDVI data files and the corresponding climate model datasets were loaded and stored in variables labeled "area*" (where the asterisk denotes the specific area: 1, 2, or 3, corresponding to 12280, 11033, and 986) and the relevant climate model name, also indexed by the asterisk to indicate the associated area. For illustrative purposes, this section focuses on area 12280 (area1) and the climate model ACCESS (access1), as the computational procedure remains consistent across all cases.

Subsequently, the aggregate function was employed to compute the mean values of pasture data. The operation *PST_access1 <- aggregate(Access1$Pasture_Corregido ∼Access1$id day, Access1, mean)* signifies that the variable "PST_access1" stores the mean of the "Pasture_Corregido" variable, grouped by the id day identifier, as derived from the Access1 dataset. After converting "PST_access1" into a dataframe, a "label" column was appended, containing sequential values from 1 to 46, to align with the structure of the NDVI data. Additionally, the column names were modified, replacing "Access1$Pasture_corregido" with "Pasture_Cum" and "Access1$id day" with "id day" for clarity.

To ensure the graphical representation was interpretable, a scaling factor was calculated, given the disparity in magnitude between NDVI and pasture values (20 versus 2000). This factor, denoted as "scal_acc1," was derived as the ratio of the maximum values of the two datasets and subsequently applied to the NDVI values within the area1 dataframe.

```
############################ ndvi data ##################################
#upload NDVI data
area1<-read_xlsx("NDVI_area12280.xlsx", sheet= "Foglio2")

###################### MODEL ACCESS CM2 ####################################

#upload model ACCESS for pasture
Access1<-read_xlsx("correct_ACCESS_CM2_area12280.xlsx")

#aggregate date
PST_access1<-aggregate(Access1$Pasture_Corregido~ Access1$`id day`, Access1, mean)
df_acc1<-data.frame(PST_access1)
df_acc1$label <- rep_len(1:46, length.out = nrow(df_acc1))
names(df_acc1)[c(1,2)]<-c("id_day", "Pasture_Cum")
View(df_acc1)

#scale factor to overlay the graphs
scal_acc1<-max(df_acc1$Pasture_Cum,na.rm = TRUE)/max(area1$Cum_Sgolay_n7_from_sept , na.rm = TRUE)
area1$sgolay_n7_p2_norm<-scal_acc1*area1$Cum_Sgolay_n7_from_sept +3000
```

Figure 3.5: Code illustration part 1

Two distinct regression approaches were implemented. The first utilized a custom function (detailed in Section 3.2) to automatically identify optimal breakpoints for fitting the data, while the second employed fixed breakpoints corresponding to seasonal transitions (specifically at points 3, 14, 26, and 37). For the former, the function prototype *regr_pst_acc1 <- find_best_segmented(df_acc1, "Pasture_Cum", "label", max_breakpoints = 4)* was applied, storing the output—a list-type variable—under the name "regr_pst_acc1". A similar naming convention was adopted for other models; for instance, the MIROC6 model output would be labeled regr_pst_miro1, where "regr" denotes regression, "pst" refers to pasture, and the suffix indicates the climate model and area number.
In essence, this function performs a regression of "Pasture_Cum" against "label" within the df_acc1 dataframe, with a predefined maximum of four breakpoints to facilitate comparison not only with NDVI results but also with the seasonal regression, which likewise incorporates four breakpoints.

An analogous regression was conducted for NDVI data using the command *regr_ndvi1 <- find_best_segmented(area1, "sgolay_n7_p2_norm", "new_label", max_breakpoints = 4)*, a procedure performed only once, cause it would have been redound.

The regression outputs, stored as list-type variables, contain several fields, of which "model" and "breakpoint" are of primary interest. The former encapsulates the regression results, with values extracted and estimated via the predict function in a new column of the corresponded dataset called "pred", while the latter records the breakpoint locations. Both sets of values were subsequently appended to the pasture and NDVI dataframes for further analysis.

```r
# regression on NDVI data
regr_ndvi1 <- find_best_segmented(area1, "sgolay_n7_p2_norm", "new_label",
                                  max_breakpoints = 4)

# regression on pasture data
regr_pst_acc1 <- find_best_segmented(df_acc1, "Pasture_Cum", "label",
                                     max_breakpoints = 4)

# add prediction on dataframe
df_acc1$pred <- predict(regr_pst_acc1$model)
area1$pred <- predict(regr_ndvi1$model)

# add subset for breakpoints (bp)
bp_pst_acc1 <- df_acc1[round(df_acc1$label) %in% round(regr_pst_acc1$breakpoints), ]
bp_ndvi1 <- area1[round(area1$new_label) %in% round(regr_ndvi1$breakpoints), ]
```

Figure 3.6: Code illustration part 2

Regarding the seasonal regression, a vector containing four breakpoints was created. These values were determined based on the id day column in the NDVI dataframe (ranging from 1 to 46), corresponding to seasonal transitions. Specifically:

1. "3": 22/09, autumn equinox;
2. "14": 19/12, winter solstice;
3. "26": 22/03, spring equinox;
4. "37": 18/06, summer solstice.

To ensure precise placement of the breakpoints, the dataframes were manually partitioned into five segments using the pmin and pmax functions mentioned above. Following this segmentation, a piecewise regression was performed using the lm function: *seas_acc1 <- lm(Pasture_Cum ∼ seg1 + seg2 + seg3 + seg4 + seg5, data = df_acc1)*.
As with the automated regression, the predict function was applied to the seas_acc1 variable to generate the regression line, in this case the new column is called "pred2". This procedure was executed once for the NDVI data and repeated for each climate model under analysis.

```r
# season regression
season_bp<-c(3,14,26,37)

#PASTURE
#segments creation
df_acc1$seg1 <- pmin(df_acc1$label, season_bp[1])
df_acc1$seg2 <- pmin(pmax(df_acc1$label - season_bp[1], 0), season_bp[2] - season_bp[1])
df_acc1$seg3 <- pmin(pmax(df_acc1$label - season_bp[2], 0), season_bp[3] - season_bp[2])
df_acc1$seg4 <- pmin(pmax(df_acc1$label - season_bp[3], 0), season_bp[4] - season_bp[3])
df_acc1$seg5 <- pmax(df_acc1$label - season_bp[4], 0)

# Regression
seas_acc1 <- lm(Pasture_Cum ~ seg1 + seg2 + seg3 + seg4 + seg5, data = df_acc1)
df_acc1$pred2 <- predict(seas_acc1)

#breakpoints
bp_seas_acc1 <- df_acc1[df_acc1$label %in% season_bp, ]

#NDVI
#segments creation
area1$seg1 <- pmin(area1$new_label, season_bp[1])
area1$seg2 <- pmin(pmax(area1$new_label - season_bp[1], 0), season_bp[2] - season_bp[1])
area1$seg3 <- pmin(pmax(area1$new_label - season_bp[2], 0), season_bp[3] - season_bp[2])
area1$seg4 <- pmin(pmax(area1$new_label - season_bp[3], 0), season_bp[4] - season_bp[3])
area1$seg5 <- pmax(area1$new_label - season_bp[4], 0)

# Regression
seas_ndvi1 <- lm(sgolay_n7_p2_norm ~ seg1 + seg2 + seg3 + seg4 + seg5, data = area1)
area1$pred2 <- predict(seas_ndvi1)

#breakpoints
bp_seas_ndvi1 <- area1[area1$new_label %in% season_bp, ]
```

Figure 3.7: Code illustration part 3

Regarding the correlation analyses, instead, the previously mentioned *correlation_function <- function(df1, df2, col1, col2, group_col1="new_label", group_col2 ="label", groups)* function was employed. Four distinct types of correlations were conducted—two for each regression type (automatic and seasonal)—with one based on the regression line and the other on the entire dataset. Again, for brevity, only the correlations pertaining to one regression type will be explained, as the methodology remains consistent across both, differing only in the data source.

Considering the automatic regression, an additional objective was to compare the different climate models. However, since these models exhibited breakpoints at varying positions and better explined in the next chapter—identified by the automatic regression function—a solution was required to standardize the analysis. To address this, the breakpoints derived from the NDVI dataset were used instead. These were extracted via a list named period_auto1.
Regarding the correlation analysis for the regression line, a new dataset was first generated for each model (as well as for the NDVI data). This dataset retained only the columns containing the regression line values—specifically, the "pred" column for the automatic regression and "pred2" for the seasonal regression. The naming convention followed the structure df_pr_ndvi1 (where "df" denotes the dataframe, "pr" signifies the predicted values, and the

suffix indicates either NDVI or the climate model of interest).

To streamline processing and consolidate all Pearson correlation indices, a variable was created using the cbind function. This variable, named matrix_cor_auto1, stored the output of the correlation function. For example: *Access = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_ac1, col1 = "pred", col2 = "pred", groups = periods_auto1)* This approach computed the correlation between the pred variable from the first dataframe and the identically named pred variable in the second dataframe, segmented according to the NDVI breakpoints. Similarly, for the full-dataset correlation, another variable—matrix_cor_auto1.2—was generated. Here, the input dataframes (area1 and df_acc1) were passed alongside grouping variables (Cum_Sgolay_n7_from_sept and Pasture_Cum) to compute correlations across the complete dataset.

```
# subset data for correlation by regression
df_pr_ndvi1 <- area1[1:46, c("new_label", "pred", "pred2")]
df_pr_ac1 <- df_acc1[1:46, c("label", "pred", "pred2")]
df_pr_cnrm1 <- df_cnrm1[1:46, c("label", "pred", "pred2")]
df_pr_ec1 <- df_ec1[1:46, c("label", "pred", "pred2")]
df_pr_miro1<- df_miro1[1:46, c("label", "pred", "pred2")]
df_pr_mpi1<-df_mpi1[1:46, c("label", "pred", "pred2")]
df_pr_mri1<- df_mri1[1:46, c("label", "pred", "pred2")]

#define breakpoints
periods_auto1 <- list(
  "6/09 - 8/10" = 1:5,
  "9/10 - 23/04" = 6:30,
  "24/04 - 10/06" = 31:36,
  "11/06 - 12/07" = 37:40,
  "13/07 - 29/08" = 41:46
)
```

Figure 3.8: Code illustration part 4

```r
#build matrix with correlation indexes based on regression line
matrix_cor_auto1 <- cbind(
  Access = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_ac1,
    col1 = "pred", col2 = "pred", groups = periods_auto1
    ),
  CNRM = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_cnrm1,
    col1 = "pred", col2 = "pred", groups = periods_auto1
  ),
  EC = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_ec1,
    col1 = "pred", col2 = "pred", groups = periods_auto1
  ),
  MIRO = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_miro1,
    col1 = "pred", col2 = "pred", groups = periods_auto1
  ),
  MPI = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_mpi1,
    col1 = "pred", col2 = "pred", groups = periods_auto1
  ),
  MRI = correlation_function(df1 = df_pr_ndvi1, df2 = df_pr_mri1,
    col1 = "pred", col2 = "pred", groups = periods_auto1
  )
)
```

Figure 3.9: Code illustration part 5

The final step involved the creation of graphical representations, beginning with regression plots followed by correlation heatmaps, both of which will be instrumental in the subsequent results analysis chapter. The initial concept was to consolidate, for each model under consideration, the NDVI boxplots, their corresponding regression lines, and the biomass boxplots into a single comprehensive graph. This approach was adopted to provide a clear visualization of the observed trends.

To achieve this, the ggplot function from the ggplot2 package was employed. Within this framework, three additional functions were utilized: geom_boxplot for the boxplots, geom_line for the regression line, and geom_point to mark the breakpoints. In these visualizations, the NDVI data was assigned a green color scheme, while each model was distinguished by a unique color.
The structure of these three functions follows a consistent pattern, requiring the definition of the data source, the x-axis (which, across all data frames, corresponded to the "label" column), and the y-axis. Additional parameters were also specified to enhance the graphical output, including color schemes, line thickness, and the handling of outliers. Finally, a secondary y-axis was incorporated for the NDVI data, facilitated by the previously mentioned scaling factor.

```
ggplot() +
  # Pasture data
  geom_boxplot(data = df_acc1, aes(x = as.factor(label), y = Pasture_Cum),
               alpha = 0.5, coef = 0.3, col = "red") +
  geom_line(aes(x = df_acc1$label, y = df_acc1$pred),
            col = "blue", linewidth = 1) +
  geom_point(data = bp_pst_acc1, aes(x = label, y = pred),
             col = "green", size = 2) +
  # NDVI data
  geom_boxplot(data = area1, aes(x = as.factor(new_label), y = sgolay_n7_p2_norm),
               col = "green", alpha = 0.3) +
  geom_line(aes(x = area1$new_label, y = area1$pred),
            col = "blue", linewidth = 1) +
  geom_point(data = bp_ndvi1, aes(x = new_label, y = pred),
             col = "red", size = 2) +

  # Axis formatting
  scale_x_discrete(labels = area1$`days/mm`, name = "Observations") +
  scale_y_continuous(
    name = "Pasture",
    sec.axis = sec_axis(~ (. - 3000)/scal_acc1, name = "NDVI")) +

  # Title
  labs(title = sprintf("Area 12280 - Model Access ")) +
  theme_minimal() +
  theme(
    axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5),
    axis.title.y= element_text(color="red"),
    axis.title.y.right = element_text(color ="green"),
    plot.title = element_text(hjust = 0.5, face = "bold"),
    plot.subtitle = element_text(hjust = 0.5)
  )
```

Figure 3.10: Code illustration part 6

Chapter 4

Results and discussion

In this section, the findings of this study will be examined. First, it is presented the complete dataset to provide an initial overview of the temporal dynamics of NDVI and pasture productivity across the observed years. The figure 4.1 shows the timeseries related to the area 12280 of the models ACCESS and CNRM, this area, as explained in section 2.1, is the one with more biomass in fact, it has a mean of 3000 NPP (net primary productivity). As in all studies, all the models have been considered, but in this section just some example will be represented, and the complete results are shown in Appendix A.
Considering the graph in the top, the NDVI is depicted in green, and pasture biomass is depicted in blue, both variables exhibit a generally consistent trend, both within individual years and throughout the study period. However, notable differences in their temporal patterns emerge upon closer inspection. More in details the NDVI follows a more linear path compared to pasture, this can be explained by the fact that for the former it is always possible to appreciate a change every observation time, while regarding the latter, there are periods where there the biomass does not grow and so the cumulative data does not change. Moreover, the NDVI shows a more constant behavior than pasture biomass, this could be explicable by many reasons as a different usage of the pasture, climate conditions or water scarcity.
Additionally, it is possible to notice a difference between the two models (just from the 2020 cause are predictions and not observation data), precisely because different sub-models are used.

Regarding the area 11033 and area 9862 similar considerations can be done; instead, the main differences are nominal values of both NDVI and pasture. For the area 11033 are the NDVI is approximately 13 and pasture 1000 NPP, while for the area 9862 the NDVI is in mean around 10 and pasture 500 NPP.

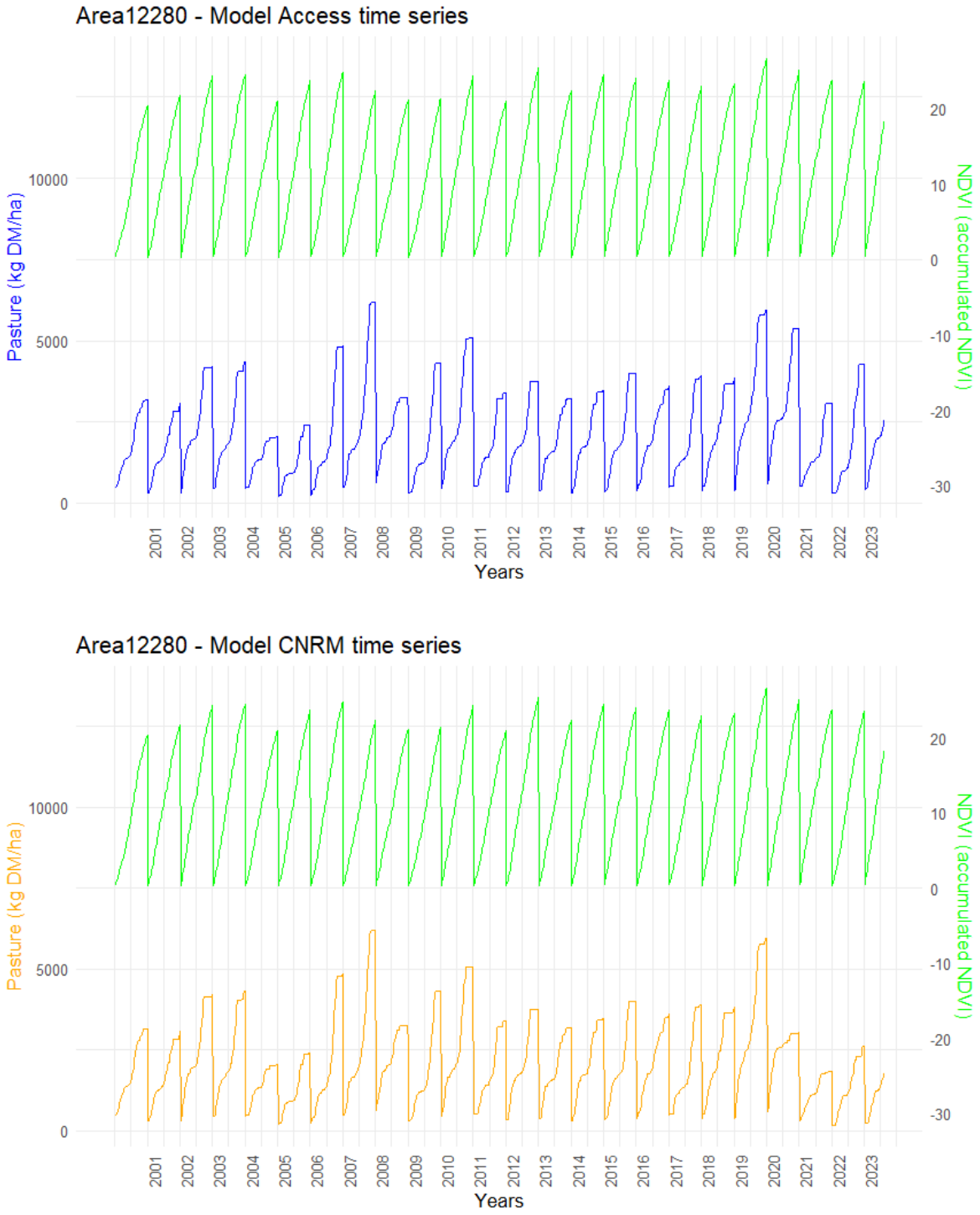

Figure 4.1: Time series area 12280; top model Access, below model Cnrm

4.1 Regression results

This section presents the results of the regression analysis. For each study area, the findings for the NDVI will be detailed first, followed by those for the pasture data.

Area 12280

NDVI

The automated regression analysis for the NDVI yielded a highly significant R^2 value of 0.97. Breakpoints were identified at the following temporal positions: 08/10, 23/04, 02/06, and 04/07, as indicated by the red points in the figure 4.2. The extended duration between the first and second breakpoint, covering more than half of the year, suggests a period of consistent NDVI growth from autumn through spring. This trend is characterized by a stable slope of 125.07, which is statistically significant (p-value < $2x10^{-16}$). This slope coefficient indicates an increase of 125.07 units on the ordinate axis for each step on the abscissa, which in this case corresponds to an 8-day interval. The high statistical significance of the slopes for all other regression segments, further detailed in Appendix A, confirms the robustness of the model. A similarly high R^2 value of 0.97 was obtained using seasonal regression, with breakpoints consistently identified at 22/09, 19/12, 22/03, and 18/06 across all model variations.

The notably high incremental values over an 8-day period are attributable to the application of a scaling factor used for graphical representation. Furthermore, negative slopes are not observed in this analysis (except in certain instances of constrained seasonal regression) as the data are cumulative. Consequently, the comparative analysis of trend intensity is best achieved by evaluating the relative magnitude of the slope coefficients.

Pasture

Regarding the pasture data, a notable observation is the slight variation in breakpoint positions (green points in the figure 4.2) across the different models, which are generally located around 30/09, 25/11, 22/03, and 10/06. The R^2 values demonstrate moderate explanatory power, ranging from 0.76 for the ACCESS model to 0.72 for the EARTH EC model.
In contrast to the NDVI, the pasture growth trend is not constant. The slopes between the second and third breakpoints, and particularly after the final breakpoint, are comparatively

low, with values of 31 and 14, respectively. The most substantial growth occurs in the fourth period, with an incremental value of 163.97, which is associated with a highly significant p-value (***).
Statistically, the first and final regression segments for all six models were found to be non-significant. The non-significance of the initial segment may be explained by the limited number of data points available for that period. For the final segment, the presence of several influential outliers, as visible in the figure, is likely responsible for the lack of statistical significance.

The seasonal regression analysis follows a congruent trend, yielding comparable R^2 values and slope coefficients to those obtained through automated regression, as comprehensively detailed in Appendix A.

	Automatic regression			Season regression		
	Value	St. error	P value	Value	St. error	P value
Slope 1	66.887	69.213	0.3341	46.96	90.56	0.604
Slope 2	92.12	8.572	2x10^-16 ***	124.06	24.2	3.5x10^-7 ***
Slope 3	27.718	6.73	4.1x10^-5 ***	31.56	7.56	4x10^-5 ***
Slope 4	171.448	7.644	2x10^-16 ***	173.59	12.35	<2x10^-16 ***
Slope 5	2.488	11.6621	0.8311	13.085	14.4	0.36

Table 4.1: Results area 12280

A comparative analysis reveals a strong congruence in both the magnitude and statistical significance of the slope coefficients for the fourth (Slope 4) and third (Slope 3) temporal segments across the models. As previously elucidated, the first and final slope estimates lack statistical significance, a finding consistent with the earlier explanations concerning data paucity in the initial segment and the influence of outliers in the terminal one.

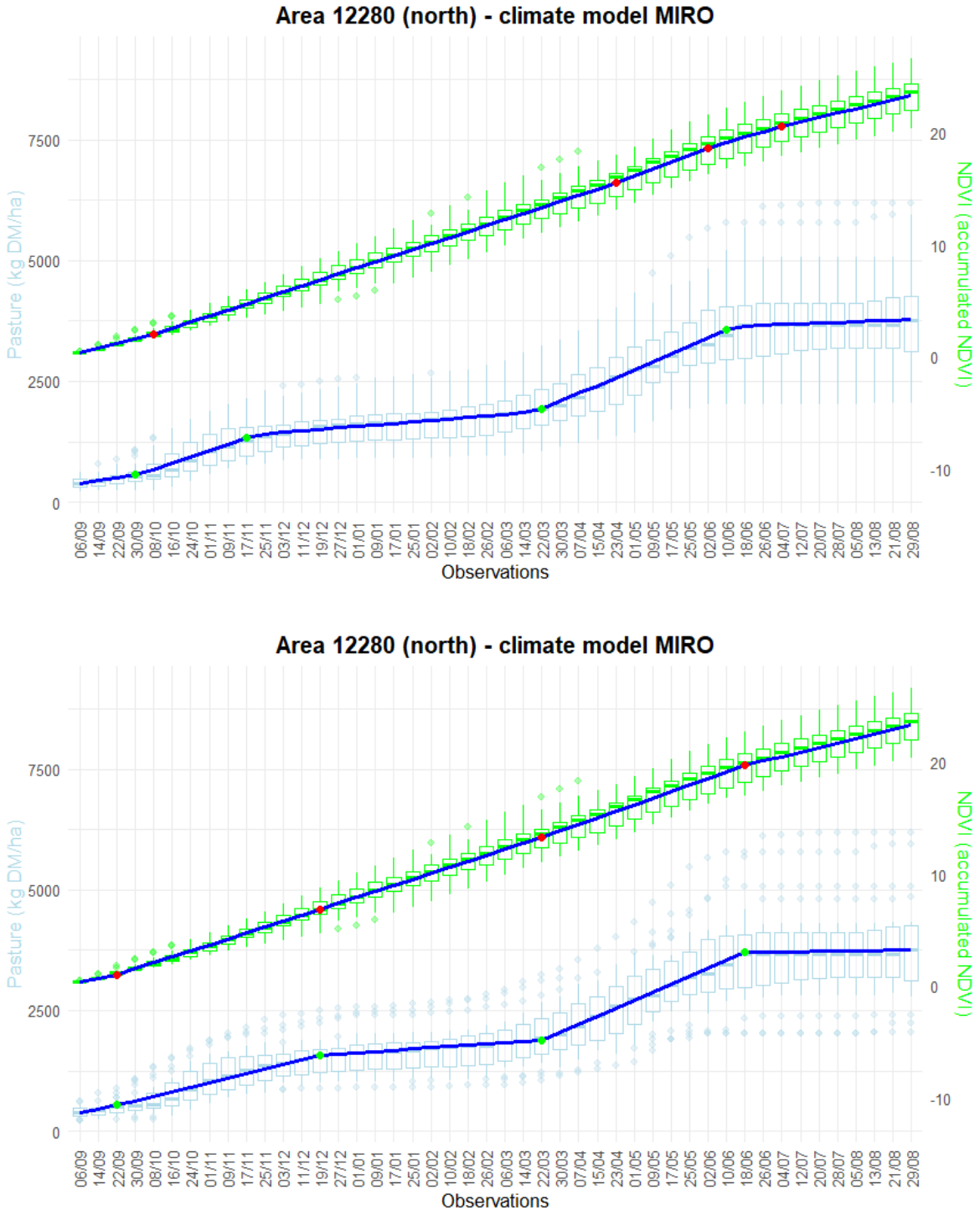

Figure 4.2: Regression analysis model MIRO area 12280: top automatic regression, below seasonal regression

AREA 11033

NDVI

Area 11033 exhibits the most fragmented automatic regression for NDVI, as shown in the figure 4.3; however, the lowest R-square value, equal to 0.95, remains highly satisfactory. In this case as well, the trend is linear, with breakpoints more evenly distributed at 16/10, 1/01, 30/03, and 2/06. As a result, the slopes are relatively similar, indicating a consistent growth pattern throughout the year. Notably, all slopes show a high level of statistical significance, as reported in Appendix A.

Pasture

The R-squared values in this case are lower, ranging from 0.58 for the MIRO model to 0.659 for the MRI model. It is useful to note that the positions of the breakpoints vary across the different models; in some models the first is located at 22/09 and in others at 27/12, (identified with green points in figure 4.3 regarding the model CNRM), while the EC model identifies only three breakpoints in the dataset. As explained in Chapter 2, the various models employ different mathematical formulations, which can lead to divergent results. As expected, the slope values, which represent the rate of biomass increase, are lower compared to those in area 12280, as this is an area with lower biomass.

The seasonal regression demonstrates, for all models and across other areas, the greatest growth occurs during autumn and a period of stasis during summer, with the specific values reported in Appendix A.

AREA 9862

NDVI

Both the automatic and seasonal regression analysis revealed an Rsquared value of 0.97. The model's slope parameters were found to be consistent and demonstrated a high degree of statistical significance (p-value < 0.001).

Pasture

Consistent with its classification as the area with the most intensive agricultural use, region

9862 was predicted to have reduced biomass. This prediction is strongly supported by the model's output, which shows significantly diminished slope values (approximately 25).

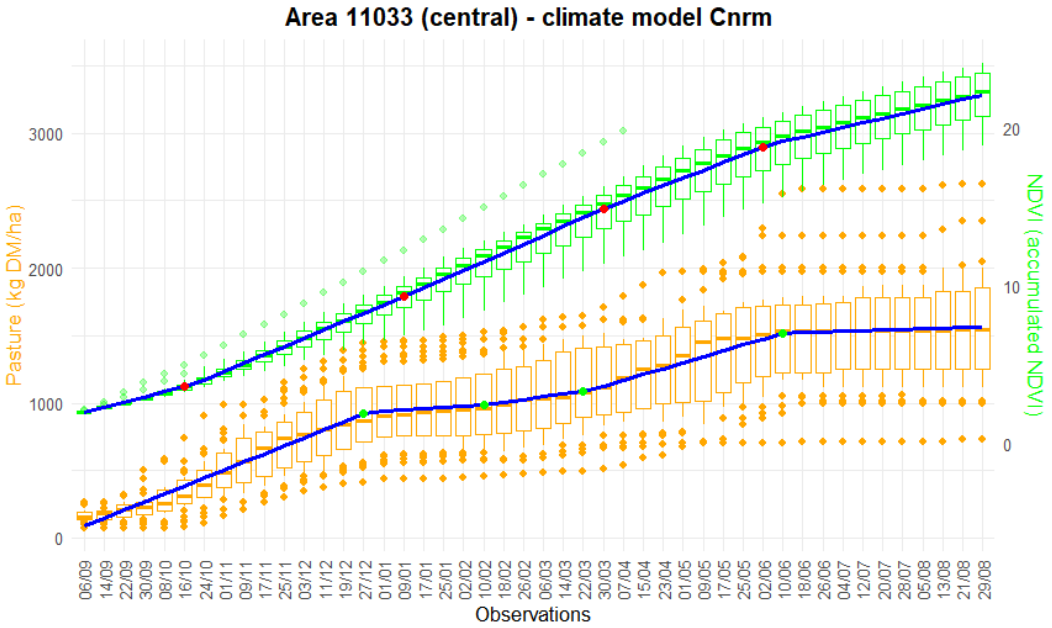

Figure 4.3: Automatic regression area 11033, model CNRM

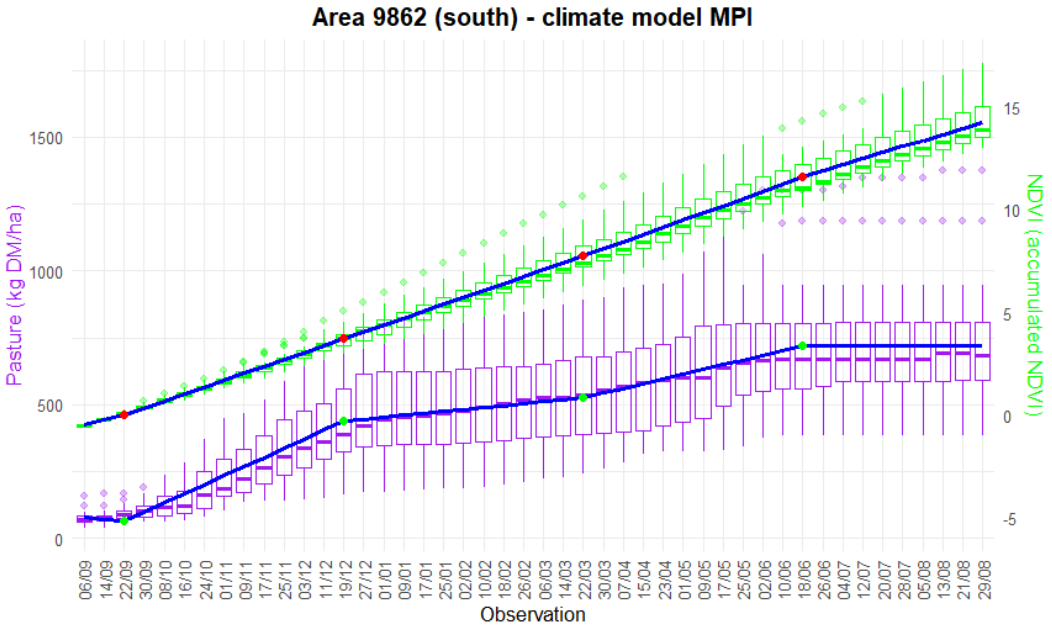

Figure 4.4: Seasonal regreggion area 9862, model MPI

It is possible to see how considering all the three study areas, although the climate models display slight variations, particularly in the initial part of the curves, the overall trend within each area remains consistent across all figures. The NDVI exhibits a linear increase throughout the year, with the highest values recorded in the northern area (12280) and the lowest in the southern area (9862), in line with the description provided in Chapter 2. Since the values are cumulative, the increase or decrease in NDVI is inherently variable, in contrast to the behavior observed for biomass.

With regard to biomass, the geographical gradient once again reflects the description outlined in Chapter 2, with higher values found in the northern mountainous area and lower values in the southern area, which is more strongly influenced by agricultural activities. Unlike the NDVI, the biomass trend is not linear; however, the general pattern is consistent across all areas, even though with different magnitudes. A common characteristic, as illustrated in Appendix A, is that, with the exception of a few models, the breakpoints occur in comparable positions within the same area, and the derived slope values are similarly aligned. Overall, the results highlight a phase of stationarity during the summer period (following the fourth breakpoint in the figures) across all areas. This can be explained by the fact that in summer pastures stop growing due to lack of water and too much temperature.

When comparing NDVI and pasture biomass, the analysis refers to seasonal regressions in order to consider equivalent periods. The table 4.2 reports NDVI and biomass values for each area, where "pasture" estimates represent the average of six climate models. In bold, the higher growth rate between NDVI and biomass is highlighted. It emerges that the steepest slopes, corresponding to the highest temporal growth, are found in spring for area 12280, with slope values across the different models averaging 168.48; instead, in areas 11033 and 9862, the highest slopes occur in autumn, with values of 65.454 and 34.5, respectively. These are the only seasons in which biomass growth exceeds NDVI growth, as also evident from the figures 4.2, 4.3 and 4.4. In winter period, all areas display lower slope values compared to spring and autumn, though these are not to be considered stationary as in the case of summer. The reduction in slope during winter can be attributed to limiting factors such as low temperatures, climatic events including snow or frost, and reduced solar radiation.

	AREA 12280		AREA 11033		AREA 9862	
	NDVI	PASTURE	NDVI	PASTURE	NDVI	PASTURE
06/09 – 22/09 End of summer	**99,64**	71,5455	**38,37**	8,3965	**22,95**	2,99667
22/09 – 19/12 Autumn	**125,07**	93,36017	61,55	**65,454**	26,66	**34,50667**
19/12 – 22/03 Winter	**142,84**	27,843	**65,75**	16,1035	**25,33**	7,3894
22/03 – 18/06 Spirng	109,81	**168,4848**	**57,14**	42,12717	**28,14**	16,72917
18/06 – 29/08 Summer	**89,23**	7,056	**34,45**	1,779833	**22,8**	1,0135

Table 4.2: Slopes' values for NDVI and pasture across the study areas

4.2 Correlation results

As previously stated, four types of correlation analyses between NDVI and pasture biomass were performed, two for each regression type. To do that, it has been selected the periods identified by the NDVI to have a comparison among the climate models. For both automatic and seasonal regressions, the results based on the regression line (labeled 'regression line' in the figure 4.5) produced statistically significant but functionally irrelevant results for the purposes of this investigation. Since these correlations are derived from regression lines fitted to cumulative data, they inherently produce Pearson correlation coefficients approaching 1. For illustrative purposes, the figure includes heatmaps for the three areas: those derived from automatic regressions for areas 12280 and 9862, and the seasonal regression for area 11033. It is evident that, regardless of the area, climate model, or time period considered, the correlation values are consistently close to 1, indicating a very strong relationship.

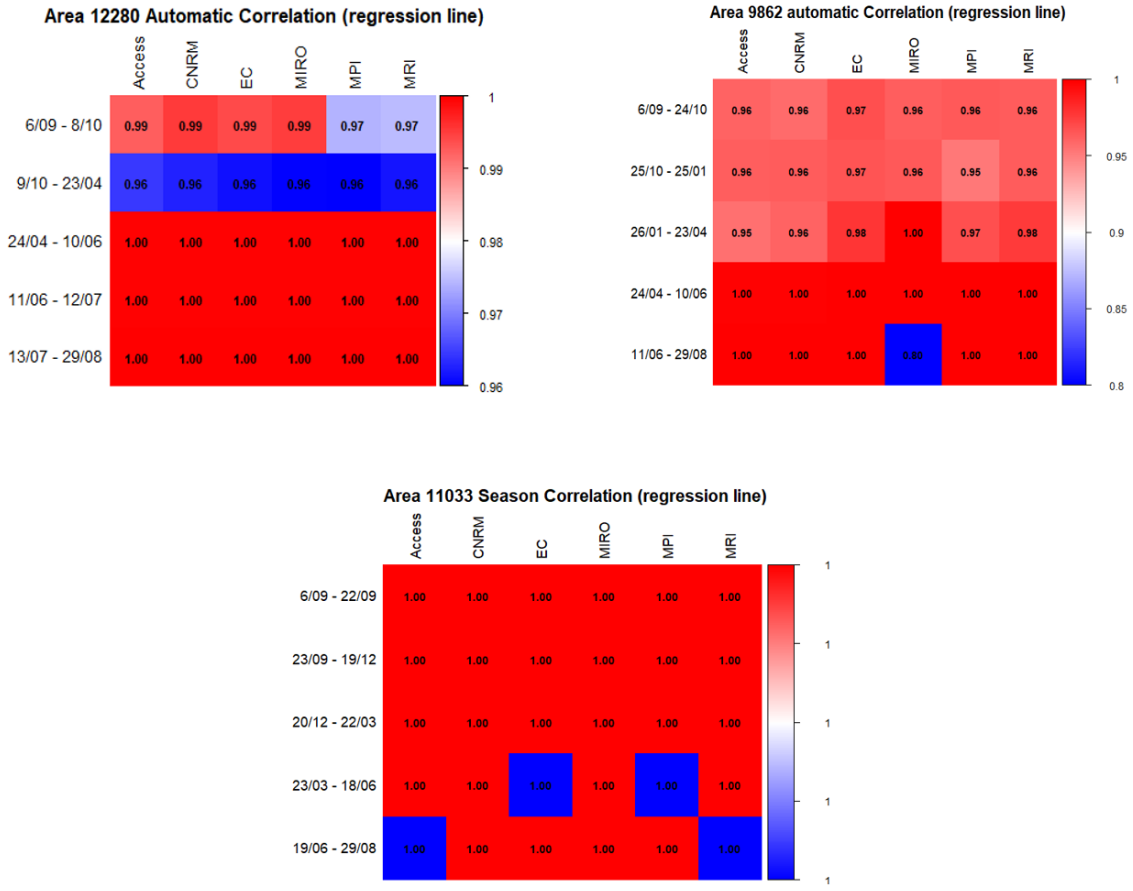

Figure 4.5: Correlation's heatmap based on regression line: top-left area 12280, top-right area 9862, below area 11033

In contrast, heat maps generated from complete datasets yield significant and interesting results. As said in chapter 3 it has been carried out 2 types of correlations, reflecting the methodology used in the regression analysis section, a detailed area-by-area examination will be performed to elucidate these findings.

AREA 12280

Regarding the automatic correlation it is possible to notice also by the figure 4.6 (the left one), the CNRM model consistently demonstrates the lowest correlation values across the entire annual cycle, establishing it as the least effective in capturing the relationship between the variables. Conversely, the MRI model emerges as the most robust, achieving a peak correlation coefficient of 0.80 for the period spanning 9/10 to 23/04.

These correlation findings are consistent with the results of the regression analysis. The

stronger correlation aligns with periods where the slopes of the NDVI and pasture biomass trends are more similar. The weaker correlation, though still indicating a discernible linear relationship (Pearson's *r* = 0.40), corresponds to the final period characterized by a stationary trend in pasture biomass.

The seasonal correlation analysis, designated as "season correlation (data)" in the figure 4.6, provides the most significant insight. This methodological approach allows for a direct comparison of data within identical temporal windows, thereby controlling for seasonal variability. This analysis yields higher average correlation indices than the previous method, while reaffirming the performance hierarchy among the models: CNRM remains the poorest performer and MRI the strongest. Notably, this seasonal framework reveals elevated correlation values, approaching 0.70, even during the spring period.

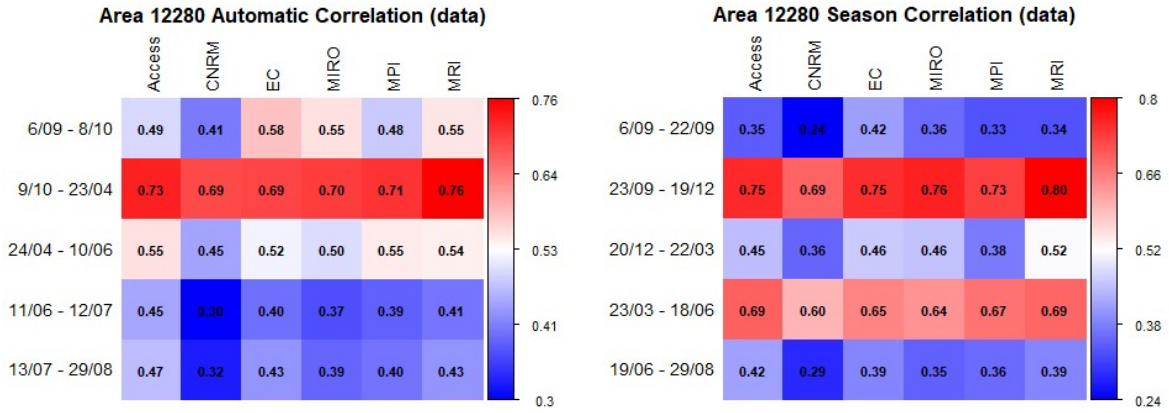

Figure 4.6: Heatmap correlation based on dataset area 12280: left automatic, right season

AREA 11033

The analysis reveals a pronounced temporal decline in correlation strength. For instance, the highest-performing model, MIROC, exhibits a substantial decrease in the Pearson correlation coefficient, from *r* = 0.66 in the initial period to *r* = 0.27 in the final period. This declining trend is consistently observed across all other models, confirming a systematic pattern of weakening relationship over time.

The ACCESS model consistently demonstrates the poorest performance, with a notably low correlation of *r* = 0.05 during the period from 10/01 to 30/01, indicating a negligible linear relationship between the variables.

This performance hierarchy between models remains consistent in the seasonal regression analysis, with MIRO again identified as the most robust and ACCESS as the least effective. A key finding from this seasonal evaluation is the stark contrast in model performance between periods. The autumn season is characterized by a strong correlation, with coefficients

reaching approximately *r* = 0.70. In contrast, for the remainder of the annual cycle, the Pearson correlation indices suggest only a weak relationship.

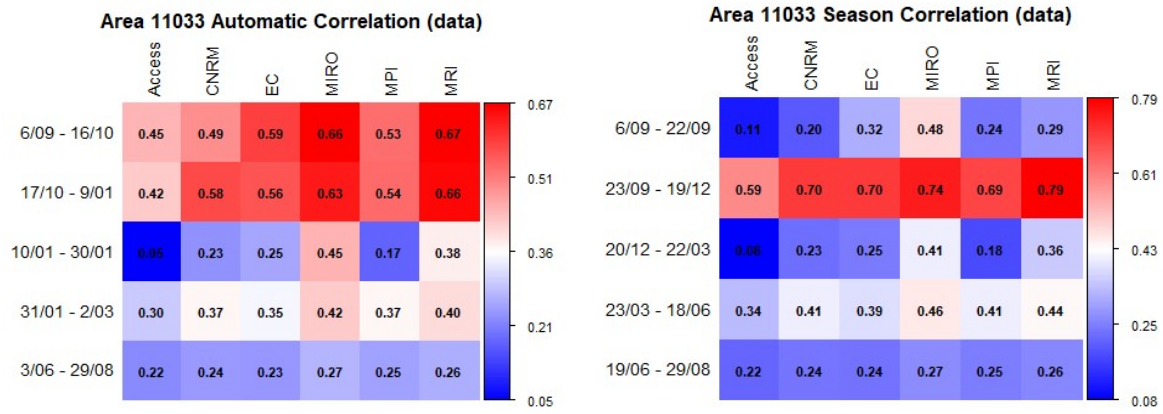

Figure 4.7: Heatmap correlation based on dataset area 11033: left automatic, right season

AREA 9862

Consistent with the regression analysis, the area 9862 exhibits a trend analogous to that observed in area 11033. Specifically, the automated regression results indicate that the initial two periods are characterized by higher correlation coefficients compared to the latter periods. Within this area, the EC model demonstrates the strongest performance, whereas the ACCESS model yields the weakest results.

This pattern is further corroborated by the seasonal regression analysis. As illustrated in the accompanying figure, the autumn season consistently presents the highest correlation indices. Conversely, the remainder of the annual cycle is marked by a notably weak correlation.

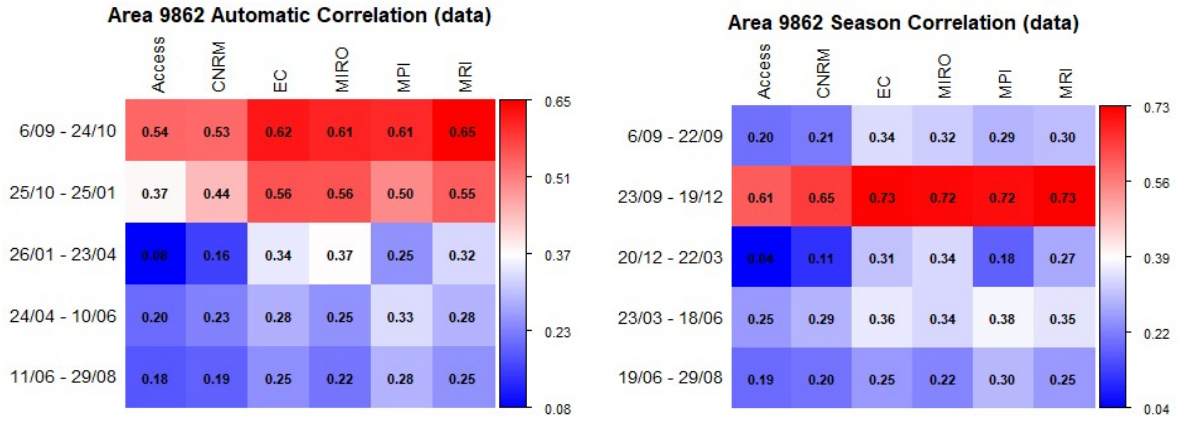

Figure 4.8: Heatmap correlation based on dataset area 9862: left automatic, right season

The analysis of correlations between NDVI and pasture biomass clearly demonstrates that the relationship between satellite-derived NDVI and the actual productivity of pastoral ecosystems is not constant but varies markedly according to season and vegetative phenology. Among the different temporal windows, the autumn period (23 September – 19 December) emerges unequivocally as the most favorable phase for the operational use of NDVI. During this season, correlation values frequently exceed 0.70, with peaks of up to 0.80, indicating a robust and predictable linear relationship between the spectral index and field-measured biomass.
This pattern is consistent with the ecological dynamics of Mediterranean pastures, where the first autumn rains terminate the summer dormancy and trigger a rapid resurgence of herbaceous vegetation. NDVI captures this flush of photosynthetically active green biomass with high fidelity, providing a reliable proxy for estimating pasture productivity.

In spring, the relationship remains statistically significant, though more variable between study areas, with correlations generally ranging between 0.4 and 0.6. Such variability suggests that local factors—including soil type, floristic composition, and management practices—exert a stronger influence, thereby reducing the uniformity of the spectral response. In contrast, both summer and late summer, as well as the winter period, reveal much weaker correlations, often falling below 0.3 and in some cases approaching zero. In summer, the accumulation of senescent, non-photosynthetic biomass leads to a clear decoupling between NDVI and total biomass, while in winter the limited vegetative activity prevents the index from adequately reflecting the true productivity of the pastures.

Inter-site comparisons further highlight the non-universal nature of the NDVI–biomass relationship. Area 12280 stands out for its greater stability, maintaining moderate to strong correlations not only in autumn but also in winter and spring, thus indicating a more consistent link between vegetation indices and productivity throughout the year. Conversely, Area 11033 shows an intermediate pattern, with strong correlations confined to autumn, weaker relationships in spring, and negligible predictability in summer. Area 9862 presents the most extreme case, where NDVI is a reliable proxy exclusively in the fall and loses predictive power throughout the year. The comparison of climate models reinforces these findings, revealing a clear performance hierarchy. The MRI model consistently delivers the highest correlations, confirming its robustness and reliability, whereas CNRM and ACCESS systematically produce weaker coefficients, in some instances approaching non-significance. This pattern, consistent with earlier regression analyses, underscores the importance of model selection as a critical factor in achieving accurate biomass estimation from satellite data.

Overall, these results provide scientific validation for the use of NDVI as an effective tool to monitor the status, productivity, and ecological health of pasture ecosystems on a continuous and large-scale basis. However, its efficacy as a proxy for total biomass is strongly modulated by seasonal phenology. While autumn constitutes the optimal window for its application, in spring the index must be interpreted with caution, and in summer and winter its utility is considerably limited. This evidence highlights a key implication for agro-pastoral management: NDVI-based maps are highly useful and easily accessible for decision-making, yet their interpretation requires critical calibration to seasonal dynamics and, when necessary, integration with complementary sources of information. In this sense, both the timing of satellite data acquisition and the choice of interpretive model emerge as decisive elements in transforming NDVI from a powerful observational index into a robust and operationally reliable metric in support of rangeland management.

Chapter 5

Conclusions

This thesis set out to investigate the efficacy of the Normalized Difference Vegetation Index (NDVI) as a reliable proxy for estimating pasture biomass under the evolving pressures of climate change. The research was motivated by a critical gap in current agricultural insurance models, which often rely on decadal NDVI averages that fail to capture the real-time dynamics of forage availability, thereby potentially undermining the resilience of pastoral systems in a warming world. Through a rigorous statistical analysis that integrated satellite-derived NDVI data from multiple climate models with field-measured biomass across three distinct areas in the Community of Madrid (Spain), this work provides a refined and scientifically robust answer to this central question.

The findings of this study unequivocally demonstrate that the relationship between NDVI and pasture biomass is not a static or universal constant but is profoundly mediated by seasonal phenology and local ecological context. The autumn period (23 September – 19 December) emerged as the unequivocally optimal temporal window for application, characterized by strong and reliable correlations frequently exceeding 0.70. This is directly attributable to the Mediterranean climatic regime of the study area, where autumn rains break summer dormancy, triggering a flush of photosynthetically active green biomass that the NDVI index captures with high fidelity. In this phase, satellite data transcends its role as a mere indicator of "greenness" to become a powerful and trustworthy tool for estimating actual biomass, offering land managers and policymakers a reliable basis for decision-making.

Conversely, the summer and late-summer periods revealed the fundamental limitation of NDVI, with correlations often diminishing to non-significant levels. This attenuation results from a decoupling of the spectral signal from total biomass; while senescent, non-photosynthetic material remains present on the ground, it is largely invisible to an index designed to detect chlorophyll activity. This critical insight confirms that NDVI is an excellent proxy for photosynthetically active green biomass but a poor one for total biomass during

periods of plant senescence. The spring season presented an intermediate scenario, with moderate correlations that exhibited significant variability between the studied areas, suggesting that local factors such as soil composition, specific floristic communities, and management practices gain greater influence, modulating the NDVI-biomass relationship.

The comparative analysis of the three study areas further enriched this understanding, highlighting that the stability of the NDVI-biomass link is itself site-specific. Area 12280, north, demonstrated remarkably consistent correlations across multiple seasons, indicating a more predictable and stable ecological dynamic. In contrast, Areas 11033 and 9862 (central and south respectively), exhibited a more extreme pattern, with strong predictability confined almost exclusively to autumn. This geographic variability underscores the necessity of moving beyond one-size-fits-all models and towards context-specific calibrations for effective large-scale monitoring. Furthermore, the consistent performance hierarchy among the climate models, with the MRI model systematically outperforming others like CNRM and ACCESS, underscores that the choice of the underlying climatic data is not a neutral detail but a critical factor determining the accuracy of the final biomass estimation.

In conclusion, this thesis validates the scientific use of satellite-derived NDVI for continuous, large-scale monitoring of pasture ecosystems, thereby contributing to the broader fields of climate-smart agriculture and sustainable rangeland management. The research confirms that when applied judiciously, with a critical understanding of its seasonal strengths and limitations, NDVI is an invaluable tool. The key to its operational success lies in the conscious integration of phenological awareness: autumn provides a window of high reliability, spring requires cautious interpretation supplemented by local knowledge, and summer demands the integration of complementary data sources or alternative methodologies. Therefore, this work provides a clear framework for enhancing the precision of agricultural insurance models and pastoral management strategies. It recommends a shift from static, long-term averages towards dynamic, phenologically-calibrated monitoring systems that can better reflect the true state of pasture productivity in an era of climate change, ultimately supporting the resilience of both agricultural livelihoods and the vital ecosystem services provided by grasslands.

Future research should focus on integrating multi-sensor approaches and developing phenologically calibrated monitoring. Combining optical indices like NDVI with radar data could be particularly useful for improving senescent biomass detection.

APPENDIX A

Tables of all results

Acronyms used:

- BP: breakpoint
- N°: number
- Automatic: regression based on breakpoints provied by automatic formula
- Season: regression based on season breakpoints
- St. err. : standard error

NDVI RESULTS

	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	V	ST. ERR.	P VALUE
AREA 12280	0,9723	0,921	1	5,44	1	99,642	17,89	3,25 x 10^-8 ***	0,9723	0,9722	1	3	1	75,28	30,53	0,0139 **
			2	29,55	2	125,07	1,67	0 ***			2	14	2	123,008	3,782	< 2 x 10^-16 ***
			3	35,39	3	142,84	13,81	6,05 x 10^-5 ***			3	26	3	124,06	2,969	< 2 x 10^-16 ***
			4	39,48	4	109,81	25,85	2,34 x 10^-5 ***			4	37	4	135,332	3,373	< 2 x 10^-16 ***
					5	89,23	10,92	8,86 x 10 ***					5	93,59	5,146	< 2 x 10^-16 ***
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQAURE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 11033	0,9615	0,9612	1	6,44	1	38,37	7,36	2,27 x 10^-7 ***	0,9614	0,9612	1	3	1	17,22	16,65	0,301
			2	16,76	2	61,55	3,39	2,11 x 10^-6 ***			2	14	2	55,95	2,06	< 2 x 10^-16 ***
			3	27,39	3	64,75	2,93	3,94 x 10^-89 ***			3	26	3	66,1	1,62	< 2 x 10^-16 ***
			4	35,31	4	57,14	4,82	1,76 x 10^-30 ***			4	37	4	55,335	1,84	< 2 x 10^-16 ***
					5	34,45	3,01	7,36 x 10^-29 ***					5	31,57	2,81	< 2 x 10^-16 ***
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 9862	0,9715	0,9712	1	7,14	1	22,95	2,26	3,76 x 10^-23 ***	0,9715	0,9713	1	3	1	18,783	6,464	0,00374 **
			2	18,54	2	26,66	1,14	6,83 x 10^-98 ***			2	14	2	25,875	0,8	< 2 x 10^-16 ***
			3	29,58	3	25,33	1,14	1,98 X10^-89 ***			3	26	3	25,835	0,628	< 2 x 10^-16 ***
			4	36,15	4	28,14	2,31	5,01 X10^-32 ***			4	37	4	26,783	0,7139	< 2 x 10^-16 ***
					5	22,8	1,34	3,54 X10^-57 ***					5	22,862	1,089	< 2 x 10^-16 ***

PASTURE RESULTS: ACCESS CM2

	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	V	N °	VALUE	ST. ERR.	P VALUE
AREA 12280	0.7665	0.7645	1	4,47	1	53,534	58,52	0,36	0,759	0,759	1	3	1	62,05	70,66	0,38
			2	10,48	2	128,87	31,282	4,08 x 10^-8 ***			2	14	2	93,288	8,752	<2 x 10^-16 ***
			3	26,28	3	34,765	7,09	1,11 x 10^-10 ***			3	26	3	26,564	6,871	0,000117 ***
			4	36,67	4	183,21	14,59	7,81 x 10^-34 ***			4	37	4	179,279	7,805	<2 x 10^-16 ***
					5	13,96	14,717	0,342					5	9,958	11,907	0,403
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 11033	0,5539	0,5502	1	3,74	1	17,098	56,4	0,761	0,5537	0,5517	1	3	1	6,286	43,024	0,88
			2	13,59	2	64,29	8,78	4,82x10^-13 ***			2	14	2	61,11	5,32	<2 x 10^-16 ***
			3	25,74	3	14,57	6,67	0,029 *			3	26	3	15,01	4,18	0,000349 ***
			4	36,21	4	47,029	7,69	1,37 x 10^-9 ***			4	37	4	45,661	4,751	<2 x 10^-16 ***
					5	3,78	8,97	0,67					5	0,925	7,24	0,898
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 9862	0,4879	0,4836	1	4,7	1	9,89	18,433	0,591	0,4874	0,485	1	3	1	-6,28	22,23	0,77
			2	14,5	2	32,73	4,53	1,02 x 10^-12 ***			2	14	2	31,797	2,754	<2 x 10^-16 ***
			3	26,3	3	5,563	3,44	0,1067			3	26	3	6,885	2,163	0,00149 ***
			4	35,7	4	19,62	5,38	2,83 x 10^-4 ***			4	37	4	17,569	2,45	1,56 x 10^-12***
					5	1,639	4,014	0,683					5	0,165	3,747	0,964

PASTURE RESULTS: CNRM ESM2

	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 12280	0,7249	0,7226	1	4,58	1	54,394	61,03	0,372	0,7239	0,7227	1	3	1	63,274	73,725	0,391
			2	10,61	2	128,17	32,622	9,07x10^-5 ***			2	14	2	91,412	9,131	<2x10^-16 ***
			3	25,61	3	31,933	8,155	9,58x10^-5 ***			3	26	3	28,803	7,169	6,28x10^-5 ***
			4	36,35	4	162,52	13,163	7,62x10^-33 ***			4	37	4	162,102	8,142	<2x10^-16 ***
					5	14,384	15,34	0,348					5	4,862	12,423	0,696
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 11033	0,6475	0,6447	1	15,2	1	59,67	3,965	1,05x10^-46 ***	0,6478	0,6462	1	3	1	2,36	35,7	0,9473
			2	21,5	2	9,086	15,839	0,566			2	14	2	65,721	4,422	<2x10^-16 ***
			3	26,4	3	21,16	20,95	0,312			3	26	3	16,147	3,471	3,7x10^-6 ***
			4	36,1	4	43,79	7,389	4,15x10^-9 ***			4	37	4	41,74	3,943	<2x10^-16 ***
					5	4,24	7,45	0,59					5	1,405	0,234	0,8154
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 9862	0,5289	0,525	1	4,81	1	10,03	17,07	0,556	0,5283	0,5261	1	3	1	-7,86	20,6	0,7
			2	14,6	2	33,97	4,202	1,67x10^-15 ***			2	14	2	32,998	2,55	<2 x 10^-16 ***
			3	26,27	3	5,683	3,192	0,075 *			3	26	3	7,2605	2,003	0,0003 ***
			4	35,68	4	18,384	4,989	2,4 x 10^-4 ***			4	37	4	16,355	2,275	1,22x10^-12 ***
					5	1,596	3,718	0,667					5	0,244	3,471	0,943

PASTURE RESLTS: EC-EARTH VEG

AREA 12280

AUTOMATIC								SEASON							
R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
0,72	0,7175	1	4,4	1	62,65	63,97	0,327	0,7189	0,7176	1	3	1	81,82	77,281	0,2899
		2	10,5	2	130,1	34,198	1,5x10^-4 ***			2	14	2	92,6	3,716	<2x10^-16 ***
		3	25,7	3	32,137	8,549	1,79x10^-4 ***			3	26	3	27,925	7,515	0,000213 ***
		4	36,5	4	168,07	13,8	4,64x10^-32 ***			4	37	4	168,909	8,535	<2x10^-16 ***
				5	17,496	16,08	0,27					5	10,315	13,022	0,428

AREA 11033

AUTOMATIC								SEASON							
R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
0,5835	0,5807	1	14,6	1	59,08	5	2,25x10^30 ***	0,584	0,582	1	3	1	8,047	40,688	0,843
		2	25,6	2	14,64	7,19	0,0419 *			2	14	2	63,864	5,039	<2x10^-16 ***
		3	36,1	3	43,5	5,97	3,03x10^-9 ***			3	26	3	16,265	3,956	4,24x10^-5 ***
		4	-	4	4,714	0,555	0,57			4	37	4	42,13	4,494	<2x10^-16 ***
				5	-	-	-					5	2,294	6,856	0,738

AREA 9862

AUTOMATIC								SEASON							
R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
0,5427	0,5389	1	4,85	1	12,5	17,37	0,47	0,5421	0,54	1	3	1	-4,97	20,96	0,812
		2	14,51	2	35,05	4,276	6,98x10^-16 ***			2	14	2	34,231	2,596	<2x10^-16 ***
		3	27,49	3	7,07	2,879	0,0142 **			3	26	3	7,734	2,038	0,000156 ***
		4	36,6	4	18,6	5,093	2,727x10^-4 ***			4	37	4	16,55	2,315	1,61x10^-12 ***
				5	1,821	4,368	0,676					5	1,947	3,532	0,581

PASTURE RESULTS: MIROC6

	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 12280	0,7286	0,7259	1	4,49	1	62,71	61,1	0,304	0,7272	0,7259	1	3	1	80,245	73,827	0,277
			2	10,44	2	132,99	32,66	5x10^-5 ***			2	14	2	93,797	9,144	<2X10^-16 ***
			3	25,69	3	31,06	8,165	1,5x10^-5 ***			3	26	3	26,646	7,179	0,000216 ***
			4	36,39	4	163,97	13,18	2,7x10^-33 ***			4	37	4	164,564	8,156	<2X10^-16 ***
					5	14,012	15,367	0,362					5	6,251	12,44	0,6154
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 11033	0,5865	0,583	1	4,36	1	33,122	34,186	0,332	0,5862	0,5844	1	3	1	17,271	0,419	0,675
			2	12,69	2	77,413	11,795	8,17x10^-11 ***			2	14	2	69,878	5,107	<2X10^-16 ***
			3	29,29	3	20,806	3,7969	5,29x10^-8 ***			3	26	3	16,365	4,009	4,79x10ì-5 ***
			4	36,14	4	46,723	14,757	1,58x10^-3 ***			4	37	4	38,332	8,418	<2X10^-16 ***
					5	4,49	8,59	0,601					5	3,847	0,554	0,58
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 9862	0,5325	0,5296	1	4,82	1	11,91	17,21	0,488	0,5332	0,5311	1	3	1	-5,75	20,736	0,78
			2	13,42	2	37,108	4,986	1,66X10^-13 ***			2	14	2	35,46	2,568	<2X10^-16 ***
			3	39,42	3	10,525	1,017	5,45X10^-24 ***			3	26	3	7,576	2,016	0,000181 ***
			4	44	4	0,785	12,431	0,949			4	37	4	14,26	2,29	6,81X10^-10 ***
					5	-0,81	55,59	0,988					5	1,173	3,494	0,737

PASTURE RESULTS: MPI ESM1

	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 12280	0,761	0,7591	1	3,96	1	49,96	90,56	0,604	0,7601	0,759	1	3	1	66,887	69,213	0,3341
			2	10,71	2	124,06	24,2	3,5x10^-7 ***			2	14	2	92,12	8,572	<2x10^-16 ***
			3	25,73	3	31,56	7,65	4x10^-5 ***			3	26	3	27,718	6,73	4,1x10^-5 ***
			4	36,29	4	173,59	12,35	2,64x10^-36 ***			4	37	4	171,448	7,644	<2x10^-16 ***
					5	13,085	14,4	0,36					5	2,488	11,662	0,8311
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 11033	0,6324	0,6293	1	4,31	1	25,13	30,71	0,413	0,632	0,63	1	3	1	6,4	37,05	0,862
			2	13,47	2	69,277	8,86	1,31x10^-14 ***			2	14	2	64,29	4,589	<2x10^-16 ***
			3	25,67	3	15,079	5,74	8,77x10^-3 ***			3	26	3	15,75	3,602	1,35x10^-5 ***
			4	35,71	4	46,394	7,64	1,8x10^-9 ***			4	37	4	43,98	4,092	<2x10^-16 ***
					5	4,244	6,688	0,525					5	0,078	6,24	0,99
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 9862	0,5819	0,5784	1	5,22	1	12,08	11,258	0,283	0,5819	0,5784	1	3	1	12,08	11,258	0,283
			2	13,85	2	37,12	5,493	2,28X10^-11 ***			2	14	2	37,124	5,493	2,28X10^-11 ***
			3	26,31	3	6,716	2,638	0,011 *			3	26	3	6,7169	2,6389	0,0011 ***
			4	35,75	4	19,35	4,653	3,344X10^-5 ***			4	37	4	19,355	4,653	3,44x10^-5 ***
					5	1,577	3,467	0,6493					5	1,577	3,467	0,649

PASTURE RESULTS: MRI ESM2

	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 12280	0,7689	0,7669	1	3,62	1	45,114	88,578	0,6106	0,7677	0,7666	1	3	1	74,997	67,735	0,2685
			2	11,3	2	125,85	19,329	1,14x10^-10 ***			2	14	2	96,944	8,389	<2X10^-16 ***
			3	24,73	3	30,79	8,305	2,2x10^-4 ***			3	26	3	29,402	6,586	8,88x10^-6 ***
			4	37,01	4	161,8	12,083	6,25x10^-38 ***			4	37	4	164,607	7,481	<2X10^-16 ***
					5	16,46	14,088	0,242					5	8,462	11,413	0,4586
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 11033	0,6597	0,6568	1	3,6	1	18,04	46,63	0,698	0,6595	0,6579	1	3	1	10,015	35,574	0,7784
			2	13,5	2	71,73	7,26	4,26x10^-22 ***			2	14	2	67,861	4,406	<2X10^-16 ***
			3	25,3	3	16,75	5,514	0,024 ***			3	26	3	17,084	3,459	9,1x10^-7 ***
			4	36,5	4	40,899	6,361	1,92x10^-10 ***			4	37	4	40,92	3,929	<2X10^-16 ***
					5	4,106	7,416	0,579					5	2,13	5,994	0,7224
	AUTOMATIC								SEASON							
	R SQUARE		BP		SLOPE				R SQUARE		BP		SLOPE			
	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE	MULTIPLE	ADJUSTED	N °	VALUE	N °	VALUE	ST. ERR.	P VALUE
AREA 9862	0,5965	0,5932	1	4,52	1	11,19	15,849	0,48	0,5962	0,5943	1	3	1	-5,2	19,12	0,7853
			2	14,48	2	36,31	3,9	7,1x10^-20 ***			2	14	2	35,43	2,368	<2X10^-16 ***
			3	25,43	3	6,563	3,379	0,0523 *			3	26	3	8,164	1,859	1,24X10^-5 ***
			4	36,6	4	16,7	3,3418	1,178x10^-6 ***			4	37	4	16,286	2,11	2,8X10^-14 ***
					5	1,424	3,985	0,72					5	0,975	3,221	0,7621

Bibliography

[1] FAO. Emissions due to agriculture: Global, regional and country trends 1990–2019, 2021. Food and Agriculture Organization of the United Nations.

[2] Yongfei Bai and M. Francesca Cotrufo. Grassland soil carbon sequestration: Current understanding, challenges, and solutions. *Science*, 377(6606):603–608, 2022. ISSN 0036-8075, 1095-9203. URL https://www.science.org/doi/10.1126/science.abo2380.

[3] Marcos Aragon M Pizarro, Carlos Diaz-Ambrona, A.M. Tarquis, M. Ana, Andrés F. Almeidana-ñuay, and Ernesto Sanz. Modelling biomass projections in grasslands of central spain under climate change scenarios. 2025.

[4] IPCC. Climate change and land: An ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. https://www.ipcc.ch/srccl/, 2019. Intergovernmental Panel on Climate Change.

[5] Tim Wheeler and Joachim von Braun. Climate change impacts on global food security. *Science*, 341(6145):508–513, 2013. ISSN 1095-9203. doi: 10.1126/science.1239402. URL https://doi.org/10.1126/science.1239402.

[6] S. Asseng et al. Rising temperatures reduce global wheat production. *Nature Climate Change*, 5(2):143–147, 2015. ISSN 1758-6798. doi: 10.1038/nclimate2470. URL https://www.nature.com/articles/nclimate2470.

[7] Chuang Zhao, Bing Liu, Shilong Piao, Xuhui Wang, David B. Lobell, Yao Huang, and Senthold Asseng. Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences*, 114 (35):9326–9331, 2017. doi: 10.1073/pnas.1701762114. URL https://www.pnas.org/doi/full/10.1073/pnas.1701762114.

[8] J. Rockström et al. Managing water in rainfed agriculture: The need for a paradigm shift. *Managing water in rainfed agriculture: The need for a paradigm shift*, 97(4): 543–550, 2010. ISSN 0378-3774.

[9] Aiguo Dai. Increasing drought under global warming in observations and models. *Nature Climate Change*, 3(1):52–58, 2013. ISSN 1758-6798. doi: 10.1038/nclimate1633. URL https://www.nature.com/articles/nclimate1633.

[10] Corey Lesk, Pedram Rowhani, and Navin Ramankutty. Influence of extreme weather disasters on global crop production. *Nature*, 529(7584):84–87, 2016. ISSN 1476-4687. doi: 10.1038/nature16467.

[11] Edmar I. Teixeira, Guenther Fischer, Harrij Van Velthuizen, Christof Walter, and Frank Ewert. Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*, 170:206–215, 2013. ISSN 0168-1923. doi: 10.1016/j.agrformet.2011.09.002. URL https://www.sciencedirect.com/science/article/pii/S0168192311002784.

[12] Kevin E. Trenberth, Aiguo Dai, Gerard Van Der Schrier, Philip D. Jones, Jonathan Barichivich, Keith R. Briffa, and Justin Sheffield. Global warming and changes in drought. *Nature Climate Change*, 4(1):17–22, 2014. ISSN 1758-678X, 1758-6798. URL https://www.nature.com/articles/nclimate2067.

[13] Daniel P. Bebber, Mark A. T. Ramotowski, and Sarah J. Gurr. Crop pests and pathogens move polewards in a warming world. *Nature Climate Change*, 3(11):985–988, 2013. ISSN 1758-6798. URL https://www.nature.com/articles/nclimate1990.

[14] S. S. Myers et al. Increasing CO2 threatens human nutrition. *Nature*, 510(7503): 139–142, 2014. ISSN 0028-0836, 1476-4687. URL https://www.nature.com/articles/nature13179.

[15] Pichu Rengasamy. World salinization with emphasis on Australia. *Journal of Experimental Botany*, 57(5):1017–1023, 2006. ISSN 1460-2431, 0022-0957. URL http://academic.oup.com/jxb/article/57/5/1017/641287/World-salinization-with-emphasis-on-Australia.

[16] IPCC. *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report*. Cambridge University Press, 2022.

[17] S. Mark Howden, Jean-François Soussana, Francesco N. Tubiello, Netra Chhetri, Michael Dunlop, and Holger Meinke. Adapting agriculture to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50): 19691–19696, 2007. ISSN 1091-6490. doi: 10.1073/pnas.0701890104.

[18] David B. Lobell, Wolfram Schlenker, and Justin Costa-Roberts. Climate trends and global crop production since 1980. *Science*, 333(6042):616–620, 2011. doi: 10.1126/science.1204531.

[19] Wolfram Schlenker and Michael J. Roberts. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37):15594–15598, 2009. ISSN 0027-8424. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2747166/.

[20] Pete Smith and et al. Agriculture, forestry and other land use (afolu). In *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the IPCC*. Cambridge University Press, 2014. URL https://www.ipcc.ch/report/ar5/wg3/.

[21] Timothy Searchinger and et al. Creating a sustainable food future. https://www.wri.org/research/creating-sustainable-food-future, 2019. World Resources Institute.

[22] Leslie Lipper, Philip Thornton, Bruce M. Campbell, Thomas Baedeker, Ademola Braimoh, Moses Bwalya, and Robert Hottle. Climate-smart agriculture for food security. *Nature Climate Change*, 4(12):1068–1072, 2014. ISSN 1758-678X, 1758-6798. URL https://www.nature.com/articles/nclimate2437.

[23] Jules Pretty, Tim Benton, Zareen Bharucha, Lynn Dicks, Cornelia Flora, Charles Godfray, Dave Goulson, Susan Hartley, Nicolas Lampkin, Carol Morris, Gary Pierzynski, P. V. Vara Prasad, John Reganold, Johan Rockstrom, Pete Smith, Peter Thorne, and Stephen Wratten. Global assessment of agricultural system redesign for sustainable intensification. *Nature Sustainability*, 1, 08 2018. doi: 10.1038/s41893-018-0114-0.

[24] Miguel A. Altieri, Clara I. Nicholls, Alejandro Henao, and Marcos A. Lana. Agroecology and the design of climate change-resilient farming systems. *Agronomy for Sustainable Development*, 35(3):869–890, 2015. URL https://hal.science/hal-01309778.

[25] Yulong Zhao, Zhongwei Liu, and Jianguo Wu. Grassland ecosystem services: A systematic review of research advances and future directions. *Landscape Ecology*, 35(5): 1053–1072, 2020. doi: 10.1007/s10980-020-00995-3.

[26] L. Liu, E.J. Sayer, M. Deng, and S. Piao. The grassland carbon cycle: Mechanisms, responses to global changes, and potential contribution to carbon neutrality. *Science of The Total Environment*, 849:157748, 2022. doi: 10.1016/j.scitotenv.2022.157748.

[27] Ann Norderhaug, Karina E. Clemmensen, Paul Kardol, Anna Gudrun Thorhallsdottir, and Iulie Aslaksen. Carbon sequestration potential and the multiple functions of Nordic grasslands. *Climatic Change*, 176(5):55, 2023. ISSN 1573-1480. URL https://doi.org/10.1007/s10584-023-03537-w.

[28] Fengfeng Cao, Weibin Li, Yuan Jiang, Xiaoling Gan, Chuanyan Zhao, and Jiancheng Ma. Effects of grazing on grassland biomass and biodiversity: A global synthesis. *Field Crops Research*, 306:109204, 02 2024. doi: 10.1016/j.fcr.2023.109204.

[29] Sølvi Wehn, Knut Anders Hovstad, and Line Johansen. The relationships between biodiversity and ecosystem services and the effects of grazing cessation in semi-natural grasslands. *Web Ecology*, 18(1):55–65, 2018. ISSN 1399-1183. URL https://we.copernicus.org/articles/18/55/2018/.

[30] Yongyao Li, Kangning Xiong, Shuzhen Song, and Wenfang Zhang. Realizing the value of grassland ecosystem services: Global practice and its inspiration for the karst desertification control area. *Frontiers in Sustainable Food Systems*, 7, 2024. ISSN 2571-581X. URL https://www.frontiersin.org/journals/sustainable-food-systems/articles/10.3389/fsufs.2023.1240431/full.

[31] Pauline Dusseux, Laurence Hubert-Moy, Thomas Corpetti, and Françoise Vertes. Evaluation of spot imagery for the estimation of grassland biomass. *International Journal of Applied Earth Observation and Geoinformation*, 38:72–77, 06 2015. doi: 10.1016/j.jag.2014.12.003.

[32] Carlos Gregorio Hernández Díaz-Ambrona, Ana Tarquis Alfonso, and Juan Rodríguez. SelecciÓn de Índices de vegetaciÓn para la estimaciÓn de la producciÓn herbÁcea en dehesas = selection of vegetation indices to estimate pasture production in dehesas. *Pastos*, 44:6–18, 10 2015.

[33] Guilherme Defalque, Ricardo Santos, Davi Bungenstab, Diego Echeverria, Alexandre Dias, and Cristiane Defalque. Machine learning models for dry matter and biomass

estimates on cattle grazing systems. *Computers and Electronics in Agriculture*, 216: 108520, 2024. ISSN 0168-1699. URL https://www.sciencedirect.com/science/article/pii/S0168169923009080.

[34] M. Pulido, S. Schnabel, J.F. Lavado-Contador, et al. The role of vegetation indices in pasture biomass estimation under mediterranean conditions. *Agricultural Systems*, 149: 103–113, 2016. doi: 10.1016/j.agsy.2016.07.017.

[35] M. Balzarini et al. Remote sensing for pasture biomass monitoring in heterogeneous landscapes. *Remote Sensing*, 12(13):2160, 2020. doi: 10.3390/rs12132160.

[36] Z Jiang, A Huete, K Didan, and T Miura. Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112(10):3833–3845, 2008. ISSN 00344257. URL https://linkinghub.elsevier.com/retrieve/pii/S0034425708001971.

[37] A.R. Huete. A soil-adjusted vegetation index. *Remote Sensing of Environment*, 25(3):295–309, 1988. URL https://www.sciencedirect.com/science/article/pii/003442578890106X.

[38] Chiara Clementini, Andrea Pomente, Daniele Latini, Hideki Kanamaru, Maria Raffaella Vuolo, Ana Heureux, Mariko Fujisawa, Giovanni Schiavon, and Fabio Del Frate. Long-Term Grass Biomass Estimation of Pastures from Satellite Data. *Remote Sensing*, 12(13):2160, 2020. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/12/13/2160.

[39] P.S. Thenkabail et al. Hyperspectral vegetation indices. In *Hyperspectral Remote Sensing of Vegetation*, pages 123–176. CRC Press, 2013. ISBN 9781439845370. doi: 10.13140/2.1.2985.3127.

[40] William James Frampton, Jadunandan Dash, Gary Watmough, and Edward James Milton. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82: 83–92, 2013. ISSN 09242716. URL https://linkinghub.elsevier.com/retrieve/pii/S092427161300107X.

[41] Yun Chen, Juan Guerschman, Yuri Shendryk, Dave Henry, and Matthew Tom Harrison. Estimating Pasture Biomass Using Sentinel-2 Imagery and Machine Learning. *Remote Sensing*, 13(4):603, 2021. ISSN 2072-4292. URL https://www.mdpi.com/2072-4292/13/4/603.

[42] Ghislain Tontibomma Bambara, André Kiema, Valérie M.C. Bougouma-Yameogo, Abroulaye Sanfo, and Wièmè Some. Development of a model for estimating the biomass of pastures using normalized difference vegetation index (NDVI): The case of the Niassa pastoral zone in Burkina Faso. *GSC Advanced Research and Reviews*, 20(2):157–166, 2024. ISSN 25824597. URL https://gsconlinepress.com/journals/gscarr/node/2788.

[43] Xiaoyang Zhang, Mark A. Friedl, Crystal B. Schaaf, Alan H. Strahler, John C.F. Hodges, Feng Gao, Bradley C. Reed, and Alfredo Huete. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84(3):471–475, 2003. ISSN 00344257. URL https://linkinghub.elsevier.com/retrieve/pii/S0034425702001359.

[44] Abdelaziz Kallel, Mauro Dalla Mura, Sana Fakhfakh, and Najmeddine Romdhane. Physics-based fusion of sentinel-2 and sentinel-3 for higher resolution vegetation monitoring. *IEEE Transactions on Geoscience and Remote Sensing*, PP:1–1, 01 2023. doi: 10.1109/TGRS.2023.3257219.

[45] Bernardo Cândido, Ushasree Mindala, Hamid Ebrahimy, Zhou Zhang, and Robert Kallenbach. Integrating Proximal and Remote Sensing with Machine Learning for Pasture Biomass Estimation. *Sensors*, 25(7):1987, 2025. ISSN 1424-8220. URL https://www.mdpi.com/1424-8220/25/7/1987.

[46] M. J. Hill and G. E. Donald. Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. *International Journal of Remote Sensing*, 24(14):3011–3030, 2010. doi: 10.1080/01431161003743181. URL https://www.tandfonline.com/doi/full/10.1080/01431161003743181.

[47] Mathilde De Vroey, Julien Radoux, and Pierre Defourny. remote sensing grassland mowing detection using sentinel-1 time series: Potential and limitations. *Remote Sensing*, 13:348, 01 2021. doi: 10.3390/rs13030348.

[48] Yizhe Li and Xinqing Xiao. Deep learning-based fusion of optical, radar, and lidar data for advancing land monitoring. *Sensors*, 25:4991, 08 2025. doi: 10.3390/s25164991.

[49] P. Manzano et al. Application of vegetation indices obtained from satellite images for the management of the voisin rational grazing. *Rangeland Ecology and Management*, 2024. In press.

[50] Wagner Santos and other. Use of Unmanned Aerial Vehicles for Monitoring Pastures and Forages in Agricultural Sciences: A Systematic Review. *Drones*, 8(10):585, 2024. ISSN 2504-446X. URL https://www.mdpi.com/2504-446X/8/10/585.

[51] Mich Ann Arbor. *Proceedings of the sixth International Symposium on Remote Sensing of Environment, Oct. 13-16, 1969*. Ann Arbor : Willow Run Laboratories, Environmental Research Institute of Michigan, 1969. URL http://archive.org/details/trent_0116401042951.

[52] Alyssa Pfaff. Capturing Deeper Ag Data Analytics with NDVI Imagery, May 2023. URL https://sentera.com/resources/articles/ndvi-imagery/.

[53] NASA Earth Science Data Systems. AppEEARS | NASA Earthdata, August 2024. URL https://www.earthdata.nasa.gov/data/tools/appeears. Publisher: Earth Science Data Systems, NASA.

[54] Bader D. C., Covey C., Gutowski W. J., Held I. M., Kunkel K. E., and Miller R. L. *Climate Models: An Assessment of Strengths and Limitations*. Climate Change Science Program (U.S.) eBooks, 2008.

[55] Agencia estatal de meteorologia. Visor de Escenarios de Cambio Climático. URL https://escenarios.adaptecca.es/#&model=CMIP6-spain.ACCESS-CM2_r1i1p1f1_ESD-RegBA&variable=tasmin&scenario=ssp245&temporalFilter=year&layers=AREAS&period=NEAR_FUTURE&anomaly=RAW_VALUE&ids=11,6.

[56] Intergovernmental Panel On Climate Change (Ipcc). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 1 edition, July 2023. ISBN 978-1-009-15789-6. doi: 10.1017/9781009157896. URL https://www.cambridge.org/core/product/identifier/9781009157896/type/book.

[57] Coupled Model Intercomparison Project. CMIP - Coupled Model Intercomparison Project, 2022. URL https://wcrp-cmip.org/.

[58] Commonwealth Scientific and Industrial Research Organisation (CSIRO). ACCESS-CM2. URL https://research.csiro.au/access/about/cm2/.

[59] Roland Séférian et al. Development and evaluation of CNRM Earth system model – CNRM-ESM1. *Geoscientific Model Development*, 9(4):1423–1453, April 2016. ISSN

1991-9603. doi: 10.5194/gmd-9-1423-2016. URL https://gmd.copernicus.org/articles/9/1423/2016/.

[60] A European community Earth System Model. EC-Earth – EC-Earth a Global Climate Model. URL https://ec-earth.org/.

[61] Hiroaki Tatebe et al. Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, 12(7):2727–2765, July 2019. ISSN 1991-959X. doi: 10.5194/gmd-12-2727-2019. URL https://gmd.copernicus.org/articles/12/2727/2019/. Publisher: Copernicus GmbH.

[62] Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2 - Mauritsen - 2019 - Journal of Advances in Modeling Earth Systems - Wiley Online Library. URL https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018MS001400.

[63] Seiji aand others Yukimoto. The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *Journal of the Meteorological Society of Japan. Ser. II*, 97(5):931–965, 2019. doi: 10.2151/jmsj.2019-051.

[64] Abraham. Savitzky and M. J. E. Golay. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36(8):1627–1639, July 1964. ISSN 0003-2700, 1520-6882. doi: 10.1021/ac60214a047. URL https://pubs.acs.org/doi/abs/10.1021/ac60214a047. Publisher: American Chemical Society (ACS).

[65] Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Viving. *Introduction to linear regression analysis*. Wiley series in probability statistics. Wiley, New York, 5rd ed. edition, 2012. ISBN 978-0-470-54281-1.

[66] Jules J. Berman. Chapter 4 - understanding your data. In Jules J. Berman, editor, *Data Simplification*, pages 135–187. Morgan Kaufmann, Boston, 2016. ISBN 978-0-12-803781-2. doi: https://doi.org/10.1016/B978-0-12-803781-2.00004-7. URL https://www.sciencedirect.com/science/article/pii/B9780128037812000047.

[67] R: What is R? URL https://www.r-project.org/about.html.

Acknowledgement

I would like to express my sincere gratitude to my thesis supervisor, Prof. Carlos Gregorio Diaz Ambrona, and to the entire team at the CEIGRAM research center. I am deeply thankful for the opportunity to work on such a challenging and interesting topic, for their invaluable support, and for graciously hosting me during this period.

A special thank you goes to my tutor, Prof. Elena Belcore, for her continued support, her prompt responses, and her consistently valuable suggestions, which were essential to the completion of this work.

Finally, I wish to extend my heartfelt thanks to Ernesto Sanz for his daily, always kindly offered support, without which this journey would have been far more difficult.