Head Office: Università degli Studi di Padova

Department: Land, Environment, Agriculture and Forestry (TESAF)

Ph.D. COURSE IN: Land, Environment, Resources and Health (LERH)
SERIES: XXXVIII

Integrating Remote Sensing and Machine Learning for Monitoring and Managing Grasslands in North-East Italy

Thesis written with the financial contribution of the European Union - Next Generation EU - National Recovery and Resilience Plan (NRRP), M4C2 – Investment 1.4 "Potenziamento strutture di ricerca e creazione di "campioni nazionali di R&S" su alcune Key Enabling Technologies". "Centro Nazionale di Ricerca per le tecnologie dell'Agricoltura – Agritech"– CUP C93C22002790001.

Coordinator: Prof. Marco Borga
Supervisor: Prof. Francesco Marinello
Co-Supervisor: Prof. Stefano Macolino

Ph.D. student : Daniele Pinna

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Territorio e Sistemi Agro-Forestali (TESAF)

Ph.D. COURSE IN: Land, Environment, Resources and Health (LERH)
SERIES: XXXVIII

Integrazione di Telerilevamento e Machine Learning per il Monitoraggio e la Gestione di Prati e Pascoli nel Nord-Est Italia

Tesi redatta Tesi redatta con il contributo finanziario dell'Unione Europea - Next Generation EU - Piano Nazionale per la Ripresa e Resilienza (PNRR), M4C2 –Investimento 1.4: "Potenziamento strutture di ricerca e creazione di "campioni nazionali di R&S" su alcune Key Enabling Technologies". Programma "Centro Nazionale di Ricerca per le tecnologie dell'Agricoltura – Agritech" - CUP C93C22002790001.

Coordinatore: Ch. mo Prof. Marco Borga
Supervisore: Ch. mo Prof. Francesco Marinello
Co-Supervisore: Ch. mo Prof. Stefano Macolino

Dottorando: Daniele Pinna

Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA | Università degli Studi di Padova

Table of Contents

Abstract

Grasslands are key ecosystems supporting livestock production, biodiversity, and carbon storage, yet their productivity, botanical composition, and phenology are highly variable in space and time. This variability, driven by complex interactions among climate, soil, and management, makes their monitoring challenging through traditional approaches. This Ph.D. research aimed to assess the potential of remote sensing and machine learning for improving the monitoring and management of permanent meadows and pastures in North-East Italy. The main objective was to develop scalable and integrative methods to quantify productivity and phenological dynamics while accounting for environmental and management constraints. First, long-term satellite observations were combined with climatic, topographic, and soil data to identify the main drivers of grasslands and pastures dynamics across Alpine and pre-Alpine landscapes. Then, probabilistic machine learning models were applied to Sentinel-2 and PlanetScope imagery to estimate above-ground biomass (AGB) at high spatial and temporal resolution. The results demonstrated that the proposed models achieve high predictive accuracy and provide interpretable uncertainty estimates, supporting their operational use in precision management. Overall, this work demonstrates that integrating remote sensing and probabilistic modelling can effectively support data-driven, climate-smart grasslands management.

Riassunto

I prati e i pascoli sono ecosistemi fondamentali per la produzione zootecnica, la biodiversità e lo stoccaggio del carbonio, ma la loro produttività, composizione floristica e fenologia presentano un'elevata variabilità spaziale e temporale. Tale variabilità, dovuta a interazioni complesse tra clima, suolo e gestione, rende difficile il loro monitoraggio con i metodi tradizionali. Questa ricerca di dottorato ha valutato il potenziale del telerilevamento e del machine learning per migliorare il monitoraggio e la gestione dei pascoli nel Nord-Est Italia. L'obiettivo principale è stato sviluppare metodi scalabili e integrati per quantificare la produttività e le dinamiche fenologiche, considerando i vincoli ambientali e gestionali. In un primo momento, osservazioni satellitari di lungo periodo sono state integrate con dati climatici, topografici e pedologici per individuare i principali fattori che influenzano le dinamiche dei prati e pascoli alpini e prealpini. Successivamente, modelli probabilistici di apprendimento automatico sono stati applicati a immagini Sentinel-2 e PlanetScope per stimare la biomassa epigea (AGB) ad alta risoluzione spaziale e temporale. I risultati hanno mostrato un'elevata accuratezza predittiva e stime di incertezza interpretabili, confermando l'idoneità dei modelli per l'impiego pratico nella gestione di queste informazioni vegetative.

1. Introduction

1.1 Background

Grasslands and pastures are among the most extensive and ecologically significant terrestrial ecosystems globally. They span over 3 billion hectares, about 40% of Earth's land surface, and deliver a multitude of ecosystem functions critical to both natural systems and human societies (FAO, 2024). These ecosystems are particularly important in the context of the global carbon cycle. Grasslands act as both sources and sinks of greenhouse gases (GHGs), depending on land use and management intensity. Properly managed grasslands can function as effective carbon sinks, with soil organic carbon (SOC) storage playing a dominant role due to their perennial vegetation and limited soil disturbance (Bai and Cotrufo, 2022). Empirical evidence shows that European grasslands can sequester up to 0.77 g/C/m² annually (Soussana et al., 2010), while Swiss meadows have been reported to store up to 7 kg/C/m² in the top 50 cm of soil (Guillaume et al., 2022). In contrast to forests, which concentrate carbon in above-ground biomass, grasslands store most of their carbon below ground. This makes them potentially more resilient to climatic extremes such as droughts, wildfires, and pest outbreaks (Pelton et al., 2024). Furthermore, well-preserved grasslands offer valuable mitigation co-benefits by emitting comparatively lower levels of methane (CH_4) and nitrous oxide (N_2O) than intensive cropland or poorly managed grazing systems (Franzluebbers et al., 2020).

Beyond carbon cycling, grasslands are vital biodiversity reservoirs. Semi-natural grasslands, including permanent meadows and pastures, shaped historically by extensive grazing or low-input mowing, are among the most species-rich plant communities in temperate regions (Habel et al., 2013). Their conservation value extends to a wide array of fauna, including specialist insects and ground-nesting birds. These systems also contribute to other regulating and supporting services such as pollination, erosion control, and nutrient cycling (Andrieu et al., 2007; Prangel et al., 2024a). However, changes in grassland management, including overuse and abandonment, have led to widespread degradation of these habitats in recent decades (Cabernard et al., 2024). This degradation threatens not only biodiversity but also the stability of the ecological services that grasslands provide. Nonetheless, evidence suggests that restoration is possible and can be rapid under appropriate management. For instance, studies in Estonia have shown that the reintroduction of grazing or mowing in previously abandoned grasslands led to the recovery of multitrophic biodiversity within a few years (Liu et al., 2025; Prangel et al., 2024a).

In addition to their ecological importance, grasslands and pastures are crucial to sustainable food systems. They account for about 67% of global agricultural land (FAO, 2024) and provide up to half of the feed resources for ruminant livestock (De Vroey et al., 2023). Grazing-based systems, when appropriately managed, can deliver high productivity with relatively low input requirements and environmental footprints. Significantly, these systems also contribute to rural livelihoods, particularly in marginal areas such as mountainous or arid regions where cropping is not feasible. In North-East Italy, these ecosystems are crucial not only for forage production, but also for maintaining traditional landscapes and supporting high-nature-value farmland, especially in less productive and sloped areas (Hinojosa et al., 2019). In the Alps and pre-Alps, where cropland is limited due to high elevation and steep terrain, grasslands are widely distributed and provide a sustainable source of fodder while preventing erosion, preserving soil fertility, and contributing to the region's aesthetic and cultural character. Given their ecological and agricultural importance, the management of grasslands and pastures must be supported by accurate monitoring systems. To maintain ecosystem services, especially in mountainous regions of North-East Italy, where small productive areas, variable topography, and fragmentation are prevalent, continuous and spatially explicit data are required to support adaptive management decisions. This includes decisions on grazing rotations, hay cutting schedules, and conservation practices. Monitoring phenology and biomass is particularly critical for optimising pasture use, sustaining productivity, and informing policy measures such as the Common Agricultural Policy (CAP) and Agri-Environment Schemes (AES) application (Sartorello et al., 2020a).

Traditionally, grassland monitoring has relied on direct field-based methods, including destructive sampling (cut-and-dry approaches), visual inspections, and nondestructive tools such as rising plate meters (RPM) (Sanderson et al., 2001). Destructive methods provide highly accurate biomass estimates but are time-consuming, labour-intensive, and can be impractical for repeated sampling or large-scale applications, unless coupled with specific decision support systems (Hanrahan et al., 2017). Furthermore, RPMs and similar devices offer quicker assessments but are often affected by terrain unevenness, operator variability, and vegetation structure (Barnetson et al., 2020; Edvan et al., 2016). Proximal sensing tools, including optical sensors and spectroradiometers, have expanded the potential of field-based approaches, enabling non-destructive and high-resolution assessments of vegetation traits (Pullanagari et al., 2012). However, these tools still face limitations in accessibility, operational costs, and scalability, especially for use across fragmented landscapes or in rugged alpine terrain.

In recent years, technological advancements have significantly enhanced our capacity to monitor and manage grasslands. Remote sensing platforms, encompassing both satellite-based and unmanned aerial vehicle (UAV)-based systems, have become indispensable tools for vegetation monitoring.

UAVs have emerged as flexible, high-resolution tools for small- to medium-scale surveys. Recent developments highlight the versatility of UAVs for real-time livestock monitoring, behaviour tracking, and sustainable management. For example, UAV-based remote sensing can effectively assess grassland biomass, monitor livestock movements, and even assist in herding activities using technologies such as virtual fencing (Aquilani et al., 2022; Goliński et al., 2023) Moreover, UAV-driven methods facilitate automatic tracking of animal activity, significantly improving grazing management efficiency (Vayssade et al., 2019). Despite these advancements, UAV adoption remains constrained by operational costs, technical expertise requirements, and regulatory challenges, including airspace management and privacy concerns (Meiyan et al., 2022; Psomas et al., 2011).

Satellite sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat series, and Sentinel-2 have been instrumental in providing systematic, multispectral imagery with varying temporal and spatial resolutions. MODIS offers near-daily observations at moderate spatial resolutions, making it suitable for large-scale vegetation monitoring. The Landsat program, with its 30-meter spatial resolution and 16-day revisit cycle, has been pivotal in long-term ecological studies. Sentinel-2, part of the European Space Agency's Copernicus Programme, provides high-resolution (10–20 m) imagery with a 5-day revisit cycle, enabling detailed field-scale monitoring even in fragmented landscapes. The inclusion of red-edge spectral bands in Sentinel-2 enhances its capability to assess vegetation health and estimate aboveground biomass. Commercial platforms like PlanetScope further enhance monitoring capabilities by providing daily imagery at 3-meter resolution, facilitating near-real-time vegetation assessments.

Machine learning (ML) has become the backbone of modern pasture monitoring, enabling the extraction of quantitative and qualitative information from increasingly large and heterogeneous remote sensing (RS) datasets. Supervised models such as Random Forest (RF), Support Vector Machines (SVM), Linear Regression (LR), and Artificial Neural Networks (ANN) dominate the field, proving effective for both biomass and pasture quality estimation (Shahi et al., 2025). For instance, Chen et al. (2021) applied an ANN using Sentinel-2 imagery and climatic variables to estimate biomass in Australian dairy farms, achieving high predictive accuracy, while Bretas et al. (2023) integrated meteorological data with Sentinel-2 vegetation indices to model Brachiaria pasture biomass with RF ($R^2 = 0.85$). At smaller scales, Togeiro de Alckmin et al. (2021) tested RF, SVM, LR, and stacking ensemble models for pasture biomass estimation from handheld multispectral sensors, with the stacked model achieving an $R^2 = 0.79$ and RMSE = 405 kg DM ha^{-1}, outperforming traditional rising-plate meter measurements.

Beyond these classical algorithms, deep learning (DL) architectures, particularly convolutional neural networks (CNNs) and multi-layer perceptrons (MLPs), have expanded the analytical scope of RS data, benefiting from transfer learning with pre-trained models such as ImageNet to address complex pasture-related tasks (Coulibaly et al., 2022). Hybrid approaches combining multispectral (Sentinel-2, Landsat), radar (Sentinel-1), and UAV imagery are gaining traction, improving temporal continuity and robustness against cloud cover. However, model transferability remains limited, algorithms trained in one region or season often underperform elsewhere due to differing environmental conditions (Smith et al., 2023).

1.2 Research questions and objectives

Pastures and permanent meadows play a crucial role in supporting livestock systems, maintaining biodiversity, and providing essential ecosystem services. However, their productivity and phenology are subject to high spatial and temporal variability, often driven by complex interactions among climate, soil, and management factors. In this context, improved monitoring tools and predictive frameworks are necessary to support more informed and adaptive pasture management. This Ph.D. project addresses the following central research questions:

- How can the spatial and temporal variability of grassland productivity and phenology be effectively modelled and quantified across different environments with remote sensing tools?
- Which environmental and management factors most strongly influence pasture dynamics and productivity?
- Can remote sensing and machine learning be combined to produce accurate and transferable models for estimating above-ground biomass (AGB) in managed grasslands?

The *general objective* of this research is to develop scalable and integrative approaches for monitoring grassland productivity and phenology by combining remote-sensing data, field measurements, and advanced modelling techniques, with the overarching aim of improving the sustainability and efficiency of pasture management.

The *specific objectives* of the study are:

- To systematically assess and benchmark existing and emerging remote sensing technologies and machine learning models for grassland monitoring, identifying their operational strengths, limitations, and transferability across environmental contexts.
- To analyse the influence of environmental drivers on pasture dynamics, identifying the conditions that most strongly regulate productivity, phenology, and management outcomes to inform modelling and decision-support applications.

- To develop and validate novel modelling frameworks and operational tools integrating multi-source remote sensing data and machine learning approaches for high-resolution, scalable, and uncertainty-aware estimation of grassland biomass productivity.

By addressing these questions, the project seeks to advance the scientific understanding and practical monitoring of grassland ecosystems. Emphasis is placed on developing methods that are not only accurate and robust but also operationally applicable in real-world pasture management contexts. The research focuses on scalable tools that leverage multisource data, such as satellite imagery, meteorological inputs, and ground measurements, to capture the spatiotemporal dynamics of productivity. The ultimate goal is to contribute to the development of a decision-support framework that integrates predictive modelling with empirical observations to enhance the sustainability and efficiency of grasslands use.

1.3 Framework of integration of the research

Based on the specific objectives identified in Section 1.2, the PhD research was designed as a sequential framework progressing from conceptual review to model development and finally to operational application. The work integrates multiscale satellite observations, environmental modelling, and data-driven machine learning approaches to improve the monitoring and management of semi-natural grasslands.

The research pathway (Fig. 1.1) can be interpreted along two main axes: a spatial scale axis, ranging from local to regional applications, and a functional axis, evolving from monitoring to management. Each study builds upon the results of the previous one, creating a continuous methodological progression that links broad-scale understanding with field-level decision support.

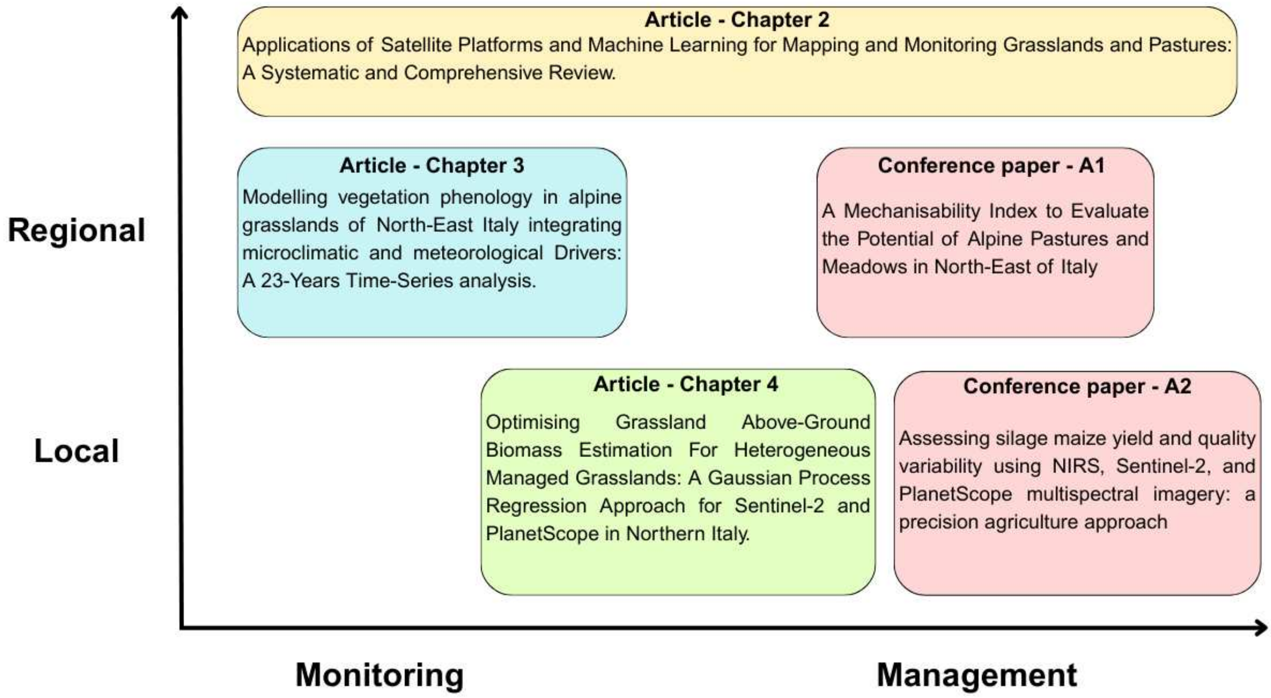


Figure 1.1 Cross-table representing the structure of the thesis. Papers in peer-reviewed journals are presented as Chapters of the thesis while conference papers are presented as annexes (A).

The first step of the PhD work addressed the initial objective: to assess existing and emerging remote-sensing technologies and machine-learning models applied to grassland monitoring. This systematic review established the conceptual and technological basis of the thesis by analysing more than two decades of studies on pasture and grassland monitoring. The work identified the most widely used satellite platforms (e.g., MODIS, Sentinel-2, PlanetScope, Landsat) and the main data-analysis techniques (e.g., Random Forest, Support Vector Machines, Gaussian Process Regression, Artificial Neural Networks). This review, published on the prestigious Smart Agricultural Technologies journal (IF:5.7), highlighted key research gaps, such as the need for multi-sensor data fusion, uncertainty quantification, and model transferability across environmental gradients, which guided the design of the subsequent experimental chapters. Thus, Chapter 2 represents the foundation of the thesis, encompassing technologies and techniques covering both regional and local scales, as well as monitoring and management applications.

The second step focused on the second objective: to analyse environmental and climatic scenarios and identify the conditions that most significantly influence pasture dynamics. Using long-term MODIS NDVI time series (2000–2023) and a suite of climatic and topographic variables, the study applied Generalised Additive Models (GAMs) to describe the non-linear relationships governing vegetation seasonality. This chapter represents the core phenological analysis of the thesis, bridging large-scale monitoring with local ecological processes. The modelling results revealed how multiple environmental drivers interact to shape grassland productivity and resilience, providing the analytical foundation for high-resolution biomass modelling in the following stage. This work was published as an article in the prestigious Computers and Electronics in Agriculture journal (IF: 8.9), thus significantly contributing to the field of phenological modelling of grassland vegetation in the region.

The third step corresponds to the third specific objective: integrating remote-sensing data and machine-learning models to estimate grassland above-ground biomass at high spatial and temporal resolution. This study developed and validated Gaussian Process Regression (GPR) models using Sentinel-2 and PlanetScope imagery. The approach enabled robust, uncertainty-aware estimation of above-ground biomass in heterogeneous managed grasslands at very high spatial and temporal resolution. This component transitions from broad-scale monitoring to local-scale, management-oriented applications, demonstrating how remote sensing observations and advanced machine learning can support operational decision-making in precision grassland management. This work was published as an article in the prestigious Precision Agriculture journal (IF: 6.6), being one of the first examples in the scientific literature of satellite-based above-ground biomass estimation at very high spatial resolution (up to 9m^2), and thus laying a path for the development of management applications in the European context, where very few public and commercial platforms are available.

Two conference papers extend the methodological framework into practical management and crop-specific contexts:

- Conference Paper A1: A Mechanisability Index to Evaluate the Potential of Alpine Pastures and Meadows in North-East of Italy. Presented at the 2023 IEEE International Workshop on Metrology for Agriculture and Forestry, Pisa, Italy, 6-8 November 2023.

This study develops an index to assess the mechanical accessibility of alpine pastures, linking biophysical factors (slope, elevation, and fragmentation) to management feasibility. It operationalises the findings of Chapter 3 by translating ecological and topographic information into practical indicators for land-use planning.

- Conference Paper A2: Assessing Silage Maize Yield and Quality Variability Using NIRS, Sentinel-2, and PlanetScope Multispectral Imagery: A Precision Agriculture Approach. Presented at the Remote Sensing for Agriculture, Ecosystems, and Hydrology XXVI Conference within the SPIE Sensors + Imaging symposium held in Edinburgh, UK, 16-19 September 2024.

This work extends the modelling framework to cultivated forage crops, demonstrating the transferability of remote-sensing and ML methods developed for pastures to agricultural systems. It strengthens the management dimension of the thesis by testing multispectral-proximal integration for yield and quality estimation.

The interconnection among the components of the PhD research is summarised in Figure 1.1. As previoulsy displayed in the figure, it presents the studies along two axes: the vertical axis (Local → Regional) represents the spatial scale of analysis, and the horizontal axis (Monitoring → Management) represents the functional evolution of the research. At the regional scale, Chapter 2 establishes the technological and methodological foundations, while Chapter 3 refines the understanding of environmental controls on grassland phenology. At the local scale, Chapter 4 and the two conference papers translate these insights into practical and management-oriented applications. Rather than constituting a linear analytical workflow, these components address grassland monitoring and management from complementary, interoperable perspectives, targeting different decision-making levels and user needs. Regional-scale analyses support strategic planning, resource allocation, and policy-oriented assessments, particularly in heterogeneous and environmentally constrained grassland systems, while local-scale modelling provides actionable, high-resolution information for farm-level management, grazing optimisation, and productivity assessment. This progression ensures coherence between the conceptual, analytical, and operational dimensions of the research. Collectively, the studies move from broad-scale observation to field-level application, enabling the joint use of regional contextual information and local productivity estimates to support data-driven, scalable tools for grassland monitoring and management under changing climatic and agronomic conditions. By linking large-scale environmental monitoring with local productivity modelling, the research contributes to translating remote-sensing and machine-learning advances into operational, evidence-based tools. These tools can support farmers, researchers, and policymakers in promoting adaptive, resource-efficient, and climate-resilient management of grasslands and forage systems.

Although the experimental applications presented in this thesis are rooted in Italian grasslands and pastures systems, the overall framework and methodologies developed are not context-specific. The integration of remote sensing, environmental modelling, and probabilistic machine-learning approaches is inherently scalable and transferable across regions, management regimes, and agro-ecological conditions. As such, the proposed framework provides a generalisable methodological blueprint that can be adapted to other grassland and forage systems worldwide, supporting comparable monitoring, modelling, and decision-support applications under diverse climatic, topographic, and management contexts

1.4 References

Andrieu, N., Josien, E., Duru, M., 2007. Relationships between diversity of grassland vegetation, field characteristics and land use management practices assessed at the farm level. Agric Ecosyst Environ 120, 359–369. https://doi.org/10.1016/J.AGEE.2006.10.022

Aquilani, C., Confessore, A., Bozzi, R., Sirtori, F., Pugliese, C., 2022. Review: Precision Livestock Farming technologies in pasture-based livestock systems. Animal. https://doi.org/10.1016/j.animal.2021.100429

Bai, Y., Cotrufo, M.F., 2022. Grassland soil carbon sequestration: Current understanding, challenges, and solutions. Science (1979) 377, 603–608. https://doi.org/10.1126/science.abo2380

Barnetson, J., Phinn, S., Scarth, P., 2020. Estimating Plant Pasture Biomass and Quality from UAV Imaging across Queensland's Rangelands. AgriEngineering 2, 523–543. https://doi.org/10.3390/agriengineering2040035

Bretas, I.L., Valente, D.S.M., de Oliveira, T.F., Montagner, D.B., Euclides, V.P.B., Chizzotti, F.H.M., 2023. Canopy height and biomass prediction in Mombaça guinea grass pastures using satellite imagery and machine learning. Precis Agric 24, 1638–1662. https://doi.org/10.1007/s11119-023-10013-z

Cabernard, L., Pfister, S., Hellweg, S., 2024. Biodiversity impacts of recent land-use change driven by increases in agri-food imports. Nat Sustain 7, 1512–1524. https://doi.org/10.1038/s41893-024-01433-4

Chen, Y., Guerschman, J., Shendryk, Y., Henry, D., Harrison, M.T., 2021. Estimating Pasture Biomass Using Sentinel-2 Imagery and Machine Learning. Remote Sens (Basel) 13. https://doi.org/10.3390/rs13040603

Coulibaly, S., Kamsu-Foguem, B., Kamissoko, D., Traore, D., 2022. Deep learning for precision agriculture: A bibliometric analysis. Intelligent Systems with Applications 16, 200102. https://doi.org/10.1016/J.ISWA.2022.200102

De Vroey, M., Radoux, J., Defourny, P., 2023. Classifying Sub-Parcel Grassland Management Practices by Optical and Microwave Remote Sensing. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15010181

Edvan, R., Bezerra, L., Marques, C., Carneiro, M.S., Oliveira, R., Ferreira, R., 2016. Methods for estimating forage mass in pastures in a tropical climate. Revista de Ciências Agrárias 39, 36–45. https://doi.org/10.19084/rca14100

Food and Agriculture Organization of the United Nations, 2024. World Food and Agriculture – Statistical Yearbook 2024. FAO. https://doi.org/10.4060/cd2971en

Franzluebbers, A.J., Wendroth, O., Creamer, N.G., Feng, G.G., 2020. Focusing the future of farming on agroecology. Agricultural & Environmental Letters 5, e20034. https://doi.org/https://doi.org/10.1002/ael2.20034

Goliński, P., Sobolewska, P., Stefańska, B., Golińska, B., 2023. Virtual Fencing Technology for Cattle Management in the Pasture Feeding System—A Review. Agriculture 13. https://doi.org/10.3390/agriculture13010091

Guillaume, T., Makowski, D., Libohova, Z., Elfouki, S., Fontana, M., Leifeld, J., Bragazza, L., Sinaj, S., 2022. Carbon storage in agricultural topsoils and subsoils is promoted by including temporary grasslands into the crop rotation. Geoderma 422, 115937. https://doi.org/10.1016/J.GEODERMA.2022.115937

Habel, J.C., Dengler, J., Janišová, M., Török, P., Wellstein, C., Wiezik, M., 2013. European grassland ecosystems: threatened hotspots of biodiversity. Biodivers Conserv 22, 2131–2138. https://doi.org/10.1007/s10531-013-0537-x

Hanrahan, L., Geoghegan, A., O'Donovan, M., Griffith, V., Ruelle, E., Wallace, M., Shalloo, L., 2017. PastureBase Ireland: A grassland decision support system and national database. Comput Electron Agric 136, 193–201. https://doi.org/10.1016/J.COMPAG.2017.01.029

Hinojosa, L., Tasser, E., Rüdisser, J., Leitinger, G., Schermer, M., Lambin, E.F., Tappeiner, U., 2019. Geographical heterogeneity in mountain grasslands dynamics in the Austrian-Italian Tyrol region. Applied Geography 106, 50–59. https://doi.org/10.1016/J.APGEOG.2019.03.006

Liu, S., Ward, S.E., Wilby, A., Manning, P., Gong, M., Davies, J., Killick, R., Quinton, J.N., Bardgett, R.D., 2025. Multiple targeted grassland restoration interventions enhance ecosystem service multifunctionality. Nat Commun 16, 3971. https://doi.org/10.1038/s41467-025-59157-8

Meiyan, S., Mengyuan, S., Qizhou, D., Xiaohong, Y., Baoguo, L., Yuntao, M., 2022. Estimating the maize above-ground biomass by constructing the tridimensional concept model based on UAV-based digital and multi-spectral images. Field Crops Res 282, 108491. https://doi.org/10.1016/J.FCR.2022.108491

Pelton, R.E.O., Kazanski, C.E., Keerthi, S., Racette, K.A., Gennet, S., Springer, N., Yacobson, E., Wironen, M., Ray, D., Johnson, K., Schmitt, J., 2024. Greenhouse gas emissions in US beef production can be reduced by up to 30% with the adoption of selected mitigation measures. Nat Food 5, 787–797. https://doi.org/10.1038/s43016-024-01031-9

Prangel, E., Reitalu, T., Neuenkamp, L., Kasari-Toussaint, L., Karise, R., Tiitsaar, A., Soon, V., Kupper, T., Meriste, M., Ingerpuu, N., Helm, A., 2024. Restoration of semi-natural grasslands boosts biodiversity and re-creates hotspots for ecosystem services. Agric Ecosyst Environ 374, 109139. https://doi.org/10.1016/J.AGEE.2024.109139

Psomas, A., Kneubühler, M., Huber, S., Itten, K., Zimmermann, N.E., 2011. Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of Grassland habitats. Int J Remote Sens 32, 9007–9031. https://doi.org/10.1080/01431161.2010.532172

Pullanagari, R.R., Yule, I.J., Tuohy, M.P., Hedley, M.J., Dynes, R.A., King, W.M., 2012. In-field hyperspectral proximal sensing for estimating quality parameters of mixed pasture. Precis Agric 13, 351–369. https://doi.org/10.1007/S11119-011-9251-4

Sanderson, M.A., Rotz, C.A., Fultz, S.W., Rayburn, E.B., 2001. Estimating Forage Mass with a Commercial Capacitance Meter, Rising Plate Meter, and Pasture Ruler. Agron J 93, 1281–1286. https://doi.org/https://doi.org/10.2134/agronj2001.1281

Sartorello, Y., Pastorino, A., Bogliani, G., Ghidotti, S., Viterbi, R., Cerrato, C., 2020. The impact of pastoral activities on animal biodiversity in Europe: A systematic review and meta-analysis. J Nat Conserv 56, 125863. https://doi.org/10.1016/J.JNC.2020.125863

Shahi, T.B., Balasubramaniam, T., Sabir, K., Nayak, R., 2025. Pasture monitoring using remote sensing and machine learning: A review of methods and applications. Remote Sens Appl 37, 101459. https://doi.org/10.1016/J.RSASE.2025.101459

Smith, H.D., Dubeux, J.C.B., Zare, A., Wilson, C.H., 2023. Assessing Transferability of Remote Sensing Pasture Estimates Using Multiple Machine Learning Algorithms and Evaluation Structures. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15112940

Soussana, J.F., Tallec, T., Blanfort, V., 2010. Mitigating the greenhouse gas balance of ruminant production systems through carbon sequestration in grasslands. Animal 4, 334–350. https://doi.org/10.1017/S1751731109990784

Togeiro de Alckmin, G., Kooistra, L., Rawnsley, R., Lucieer, A., 2021. Comparing methods to estimate perennial ryegrass biomass: canopy height and spectral vegetation indices. Precis Agric 22, 205–225. https://doi.org/10.1007/s11119-020-09737-z

Vayssade, J.A., Arquet, R., Bonneau, M., 2019. Automatic activity tracking of goats using drone camera. Comput Electron Agric 162, 767–772. https://doi.org/10.1016/J.COMPAG.2019.05.021

2. Applications of Satellite Platforms and Machine Learning for Mapping and Monitoring Grasslands and Pastures: A Systematic and Comprehensive Review

2.1 Abstract

Grasslands and pastures are critical ecosystems globally, essential for their agricultural and environmental roles. Extensive research from 2000 to 2022, comprising 504 articles, has explored the integration of remote sensing and statistical modelling for mapping and monitoring grassland and pasture ecosystems. These articles were sourced from the SCOPUS database and analysed using text mining and natural language processing techniques to investigate the evolution of publication trends over the past twenty-two years. The number of publications per year on this topic has grown consistently in the considered period, from 3 in 2000 to 93 in 2022, doubling their weight compared to the total number of Scopus publications. The quantitative analysis of satellite platform utilisation revealed the increasing importance of Sentinel-2, even though MODIS remained the most utilised satellite platform throughout the study period. The increasing availability of big data has helped spread the utilisation of machine learning algorithms, mostly in the last ten years. Among these, random forest appeared to be the most widespread for grassland and pasture studies. Researchers' primary interest in this field centres on the technologies and their applications. This is evidenced by cluster analysis, which reveals a dominant focus on terms related to the 'Instruments' (25.8%) and 'Parameters' (24.9%) categories. This analysis aims to outline the progression of research, offering insights that could be useful in forecasting future trends and facilitating stakeholder engagement in this sector.

2.2 Introduction

Grasslands are among the most critical and widespread ecosystems in the world. In 2020, they covered more than 3 million ha worldwide (*World Food and Agriculture – Statistical Yearbook 2023*, 2023a) and stored approximately 30% of terrestrial biomass (Bar-On et al., 2018). The role that these ecosystems play regarding climate-changing emissions is prominent (Sándor et al., 2023); pastures and grasslands can both stock and emit carbon dioxide (CO_2) (Derner and Schuman, 2007) and other greenhouse gases, such as methane (CH_4) and nitrous oxide (N_2O) (Franzluebbers et al., 2020). Several authors have investigated the role of grasslands as carbon sinks and highlighted their importance. Dass et al. (Dass et al., 2018)d that in California, grasslands might be more resilient and reliable than forests as carbon sinks. Evidence from European grasslands shows that the soil C sequestration rate can reach 0.77 g C m-2 (Soussana et al., 2010), while a Swiss study by Guillaume et al. (Guillaume et al., 2022) concluded that the first 50 cm of grassland topsoil contains approximately 7 kg C m-2 and that including temporary grassland in crop rotation may increase carbon storage in agricultural soils. Compared with croplands, grasslands and pastures have the potential to act as C and N sinks and mitigate GHG emissions in livestock systems, as C and N sequestration can offset GHG emissions (Guillaume et al., 2022; Soussana et al., 2010)**.** Recent studies developed remote sensing tools based on Sentinel-2 and Sentinel-5p to track and quantify GHG emissions (Ehret et al., 2022). Furthermore, grasslands are the cradle of vast plant and animal biodiversity (Bergman et al., 2008; Pokluda et al., 2012). Despite their role in mitigating greenhouse gas emissions, these ecosystems face significant threats from climate change. This is particularly evident in issues like erosion (Straffelini et al., 2024) and reduced grass productivity (Dibari et al., 2021a). From an agricultural point of view, using grasslands as pastures represents a vital resource. It is the least expensive way to produce fodder, providing nearly half the feed requirements for global livestock production (De Vroey et al., 2023) and covering 67% of agricultural land worldwide (*World Food and Agriculture – Statistical Yearbook 2023*, 2023a). The amount and quality of forage produced in pastures significantly impact animal welfare and final production (Cortner et al., 2019; Poffenbarger et al., 2017). Grasslands covering marginal areas in agricultural environments as riverbanks, also have potential as biomass supplier for anaerobic digestion plants (D. Boscaro et al., 2015). Changes in the botanical composition, production, and quality of pastures can affect final production, which can have important socioeconomic consequences for rural areas. The different characteristics of pastoral activities can change their effect on the grassland environment, both in terms of productivity and biodiversity. Sartorello et al. (2020b) identified four main categories of pastoral activities: (*i*) traditional management, defined as extensive grazing; (*ii*) overgrazing, representing land overexploitation; (*iii*) Agri-Environmental Schemes (AESs), reducing overgrazing towards traditional extensive management to mimic and/or enhance biodiversity conservation; and (*iv*) land abandonment following previous pastoral activities.

During the last decades, mountainous regions have experienced profound socioeconomic transformations (Price et al., 2017). These changes encompass land abandonment, a decline in agricultural practices, an ageing farming population, and increased urbanization (Sgroi and Modica, 2022). Consequently, pastoral systems have deteriorated, with unpalatable grasses taking over, shrubs encroaching, and reforestation occurring (Brilli et al., 2023; Dibari et al., 2021b). This degradation of pastoral landscapes highlights the need for sustainable management and restoration efforts to preserve these vital ecosystems (Malatesta et al., 2019; Price et al., 2017). To ensure accurate information about the state of grasslands, characterisation, monitoring, and management are necessary. Additionally, understanding the potential of nearby environments is crucial. Traditional ground-based methods are laborious and time-consuming (Zhu et al., 2023), while remote sensing is efficient and fast and thus has the potential to provide large amounts of useful real-time data to improve pasture and grassland management at both large (regional) and small (farm) scales. The activity of researchers in this field has substantially increased in recent years, together with the tools and possibilities offered by remote sensing technologies (Wang et al., 2022). Some of the most investigated topics include monitoring grassland degradation (Xu et al., 2020); estimating vegetation-related parameters such as aboveground biomass (Andersson et al., 2017; Barnetson et al., 2020; Psomas et al., 2011; Zeng et al., 2019; Zhou et al., 2023), leaf area indices (Jie Wang et al., 2019) or fractional vegetation cover (Andreatta et al., 2022a); and characterising grassland habitats (Bangira et al., 2023). Some of the most utilised satellite platforms include moderate-resolution imaging spectroradiometer (MODIS) (which is especially suitable for large-scale studies (Cai et al., 2021; Liu et al., 2017; Zhao et al., 2014)), Landsat satellites from NASA (Anderson et al., 1993; Kazar and Warner, 2013), and the Sentinel constellation from the ESA (Andreatta et al., 2022b; Cisneros et al., 2020; J. Wang et al., 2019). Remote sensing multispectral and hyperspectral images and radar data can also be used to predict grassland taxonomic and functional plant diversity, as well as phenological traits (Fauvel et al., 2020; Imran et al., 2021; Perrone et al., 2024). Fauvel et al. (Fauvel et al., 2020) used a combination of optical data from Sentinel-2 and synthetic aperture radar from Sentinel-1 to compute several diversity indices, while Imran et al. (Imran et al., 2021) focused on the application of fine-scale hyperspectral imaging to estimate grasslands biodiversity. Numerous studies have recently focused on integrating machine learning and deep learning techniques to extract valuable information about grassland vegetation from remote sensing data. Techniques such as random forest (Gao et al., 2020; Q. Wang et al., 2022), regression models (Dandikas et al., 2018), neural networks (Vawda et al., 2024), and clustering methods (Cui et al., 2024) have significantly enhanced the analysis and interpretation of complex datasets. These advancements have led to more accurate and comprehensive assessments of grassland ecosystems. In recent years, several authors have reviewed the status of grassland observation methods and applications based on remote sensing data and machine learning applications, the advancements in technology and methodology for retrieving various grassland biophysical data and management traits and recognised the primary unmet needs and forthcoming trends for development (Bazzo et al., 2023; Maake et al., 2023; Reinermann et al., 2020). The reviews included major advancements and trends from scientific researchers on the topic. However, a systematic quantitative review is missing. This review, similarly to other recent studies (Abad et al., 2021; Cogato et al., 2019; Sullivan et al., 2024; Trentin et al., 2024), would analyse trends and geographical and thematic clusters emerging from large quantities of scientific papers. In this research, the scientific literature was analysed using a quantitative method based on text-mining techniques. In fact, due to the vastness of the specific topic, the framework of the literature review should follow the guidelines presented in the PRISMA statement (Page et al., 2021). The present analysis aims to provide a comprehensive state-of-the-art review of the literature concerning the use of remote sensing technologies for characterising and monitoring grasslands and pastures. The specific objectives of this review are to (*i*) describe the trend of the publications throughout the years, (*ii*) highlight the significant gaps in the research relating to the use of remote sensing in this field, both in terms of topics and geographical areas, with a focus on satellite platforms and machine learning algorithms, and (*iii*) categorise the interrelationships among the grassland and pasture monitoring and management activities and remote sensing-based tools, with a focus on satellite platforms, reported in the literature using cluster analysis.

2.3 Materials and Methods

A systematic quantitative review was performed by extracting documents from the Scopus database (www.scopus.com). This approach relies on selecting explicit and reproducible survey methods (Pickering and Byrne, 2014), which allow a holistic view of the literature. It enables the presentation of a clear and complete picture of the state of the art. Additionally, it highlights key topics raised by the scientific community and facilitates cluster analysis.

The analytic methodology approach makes it possible to map the gaps, not only from a theoretical point of view but also from a methodological and geographic point of view.

The analysis was conducted through a text-mining approach. This involved examining the words present in the "title – abstract – keywords". The tools utilised included NumPy, Pandas, Matplotlib and Seaborn libraries in Python 3.0.11, Microsoft Excel and Gephi's (Gephi® Consortium, Compiegne, France) graphical representation tools, which are open-source network analysis software. Text mining involves extracting meaningful numerical indices from texts by scrutinising unstructured textual data. Analysing these indices statistically unlocks the understanding of the text, yielding substantial and quality insights (Bangira et al., 2023; Cogato et al., 2023, 2019).

2.3.1 Article Selection

The analysis focused on the remote sensing technologies and techniques utilised to map and characterise grasslands and pastures. The article selection was based on the combination of the terms "grassland" and "pasture" with the general terms "remote sensing", "UAV", "drones", and "satellite" and with the specific satellite platform names "Landsat", "Sentinel", "RapidEye", "Planet", "Worldview", "Gaofen", "Modis", "Alos", "Spot", "Avhrr", "Hyperion", "Ikonos", "Radarsat", "Envisat", "Formosat", "Pleiades" and "Quickbird". Only documents published in journals and written in English were utilised, and the considered timespan ranged from 2000 to 2022. The query string utilised is reported in Table 1.

A total of 504 articles were selected and downloaded using the .csv extension. Data relating to the following key metadata fields were selected: 'Title', 'Authors', 'Year', 'Source title', 'Affiliation', 'Cited by', and 'Author Keywords'.

Table 2.1 Script for the extraction of research papers.

Query String	((TITLE (pasture*) OR TITLE (grassland*)) AND (TITLE (satellite*) OR TITLE (uav*) OR TITLE (drone*) OR TITLE (sentinel) OR TITLE (landsat) OR TITLE (rapideye) OR TITLE (planet) OR TITLE (gaofen) OR TITLE (worldview) OR TITLE (modis) OR TITLE (spot) OR TITLE (avhrr) OR TITLE (hyperion) OR TITLE (ikonos) OR TITLE (alos) OR TITLE (radarsat) OR TITLE (envisat) OR TITLE (formosat) OR TITLE (pleiades) OR TITLE (quickbird) OR TITLE (remote AND sens*))) AND (PUBYEAR > 1999 AND PUBYEAR < 2023) AND (LIMIT-TO (SRCTYPE,"j")) AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"re")) AND (LIMIT-TO (LANGUAGE,"English"))

Due to the specific dataset and time period utilised, the study's limitations in terms of generalizability are clear. Even though these constraints are explicitly addressed, Scopus is a broad database, and several reviews are based on sole Scopus database (Abad et al., 2021; Preite et al., 2023; Trentin et al., 2024). According to Sampson et al. (Sampson et al., 2006), "Searching additional databases with overlapping coverage but fewer precision-enhancing features may reintroduce irrelevant material that has already been eliminated from the retrieval in the database with the fullest feature set". By concentrating on a well-defined dataset, we aim to ensure the robustness of our findings while acknowledging the inherent limitations of this approach.

2.3.2 Trend Analysis

The dataset downloaded from Scopus was imported using Python's panda's library, which is renowned for its powerful data manipulation capabilities and was instrumental in processing and analysing these data. Preliminary analysis involved descriptive statistical techniques to summarise publication trends over time and the geographical distribution of research contributions. Matplotlib and seaborn, Python's primary libraries for statistical graphics, facilitated the visualisation of these trends.

2.3.3 Quantitative analysis of satellite platforms and machine learning algorithms

This segment of the analysis focused on extracting specific information about satellite platforms and machine learning algorithms from the corpus of articles. The objective was to ascertain the prevalence and usage trends of various satellites and algorithms within the field. The search for the occurrence of satellite platforms and machine learning algorithms was exclusively conducted within the 'Abstract' section of each article in the dataset. This decision was based on the rationale that the abstracts provide a succinct and focused summary of the research, including key methodologies and technologies employed. The process involved the following detailed steps:

1. *Defining Keywords*: The first step was to compile comprehensive lists of keywords representing both satellite platforms and machine learning algorithms, as reported in Table 2.
2. *One-Time Count per Abstract*: To quantify the occurrence of these technologies, a distinct counting rule was applied: each term was counted only once per abstract, regardless of the number of times it appeared within that abstract.
3. *Frequency Tallying*: The Python script scanned each abstract for the presence of the predefined keywords, employing case-insensitive matching to ensure comprehensive detection. When a keyword was identified, it was counted for that abstract, and subsequent mentions within the same abstract were discarded.
4. *Data Aggregation and Interpretation*: The final tally provided insights into the prevalence of specific satellite platforms and machine learning algorithms within the field based on their utilisation in different studies.

Table 2.2 Lists of keywords employed for the quantitative analysis of satellite platforms and machine learning algorithms.

Topic	Keywords
Satellite Platforms	‘Sentinel’, ‘Landsat’, ‘Modis’, ‘Spot’, ‘AVHRR’, ‘Worldview’, ‘RapidEye’, ‘Planet’, ‘Radarsat’, ‘Quickbird’, ‘Alos’, ‘Gaofen’, ‘Hyperion’, ‘Envisat’, ‘Formosat’, ‘Hyspiri’, ’Ikonos’, ‘Prisma’, ‘Venus’
Machine Learning Algorithms	‘Linear Regression’, ‘Logistic Regression’, ‘Partial Least Square Regression’, ‘Decision Tree’, ‘Random Forest’, ‘Support Vector Machine’, ‘Naïve Bayes’, ‘K-Nearest Neighbors’, ‘Artificial Neural Network’, ‘Gradient Boosting’, ‘Generalized Additive Model’, ‘Convolutional Neural Network’, ‘Recurrent Neural Network’, ‘Principal Component Analysis’

2.3.4 Natural Language Processing, Clustering and Network Analysis

The textual content of titles, author keywords and abstracts underwent extensive natural language processing, implemented through several Python libraries:

1. *Text Cleaning and Normalization*: Utilising regular expressions (via Python's remodule), extraneous elements such as URLs, punctuation, and numerals were removed, and text normalization was achieved by converting all characters to lowercase.
2. *Tokenization and Stop Words Removal*: The nltk library (Natural Language Toolkit), a comprehensive suite for NLP in Python, facilitated the tokenization of text, splitting it into individual words and removing common stopwords.
3. *Stemming*: Nltk's PorterStemmer was applied to reduce word tokenisation of their root form, aiding in consolidating various morphological variants of words.
4. *Bigram Collocation* Analysis: Identification of frequently co-occurring word pairs (bigrams) was executed using nltk's BigramCollocationFinder.
5. *Word Frequency Counting*: The collection library's counter class was used to count the occurrence of each term within the corpus.

From the frequency distribution obtained by the analyses, we selected the 150 most common words. The CSV file was imported into Microsoft Excel for manual validation and clustering. The list of words was verified for synonyms and irrelevant words. The word list was ultimately limited to the 70 most important and recurring words. The definition of the clusters was based on an initial topic modelling analysis, which was then adapted by the authors in order to fit the relevant terms highlighted by the analysis.

The five identified clusters were 'Vegetation', 'Instruments', 'Environment', 'Parameters' and 'Management'; the words composing the five clusters are shown in Table 3.

Table 2.3 Cluster composition.

Cluster	**Lemmas**
Instruments	Model, Remote Sensing, Image, Satellite, Modis, Sentinel, R, Map, Landsat, Regression, Algorithm, Random Forest, UAV, Correlation, Hyperspectral, Multispectral
Parameters	Vegetation Index, Accuracy, Index, NDVI, AGB, Productivity, Spatial, Spectral, Time Series, Normalized Difference, LAI, GPP, Resolution, Temporal, Variability, Reflectance, Optical, Bands, Seasonal, EVI, RMSE

Vegetation	Grassland, Pasture, Biomass, Species, Phenology, Plant, Grass, Agricultural, Canopy, Forage
Environment	Soil, Ecosystem, Degradation, Water, Carbon, Ground, Drought, Land, Climate, Ecological, Precipitation, Fire, Steppe, Temperature, Alpine
Management	Estimation, Monitoring, Classification, Grazing, Mapping, Conservation, Livestock, N

The co-occurrence matrix was then constructed using these 70 terms. The matrix was a two-dimensional array in which both the rows and columns represented the selected words. For each document in the dataset, the script iterated through pairs of these top words, recording the instances where both words in a pair appeared together in the same document. The value in each cell of the matrix corresponded to the count of co-occurrences for the word pair it represented. This was achieved using Python's data manipulation capabilities, specifically with pandas and collections. The final co-occurrence matrix provided a quantified representation of the relationships between key terms in the corpus. This matrix served as the foundation for constructing an edge list, in which each edge represented a pair of co-occurring terms, and the weight of the edge corresponded to their co-occurrence frequency.

The edge list was formatted for compatibility with Gephi, facilitating the visualisation of the term network and the exploration of thematic clusters within the research field. The concluding stage involved utilising Gephi (Gephi® Consortium, Compiegne, France), which is open-source software for visualising the relationships between terms. In Gephi's graphical representation, terms are depicted as interconnected nodes (possibly with assigned weights), and the vectors, which can be directed or undirected, represent the links between these terms. Figure 1 depicts the conceptual flow of the analytical process.

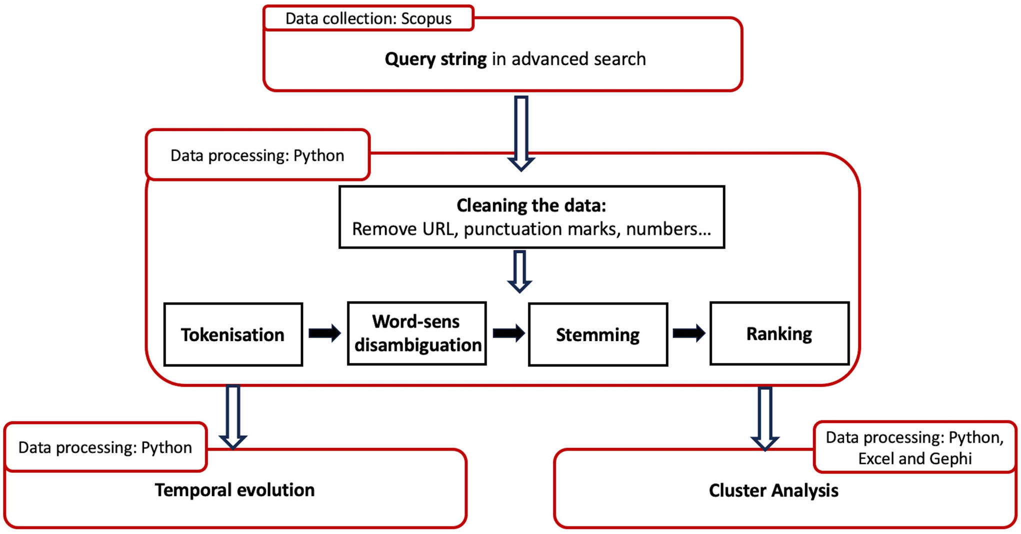


Figure 2.1 The conceptual flux of the analysis: model and software used.

2.4 Results

2.4.1. Analysis of Trends

The first consideration concerns the number of articles published per year in the considered period (2000-2022) concerning remote sensing in the fields of Earth and planetary sciences and agricultural and biological sciences.

Figure 2.2 compares them with the total number of publications published in Scopus during the same period. The number of total publications in Scopus has grown constantly, from 1,182,106 in 2000 to 3,971,310 in 2022. Even the number of publications concerning remote sensing in Earth and planetary sciences and agricultural and biological sciences has grown, going from 2,446 in 2000 to 13,962 in 2022, but following a less constant trend. The upwards trend followed an up-and-down trajectory until 2017, when it grew steeply and constantly, with the only exception being 2021, when it reached its maximum in 2022.

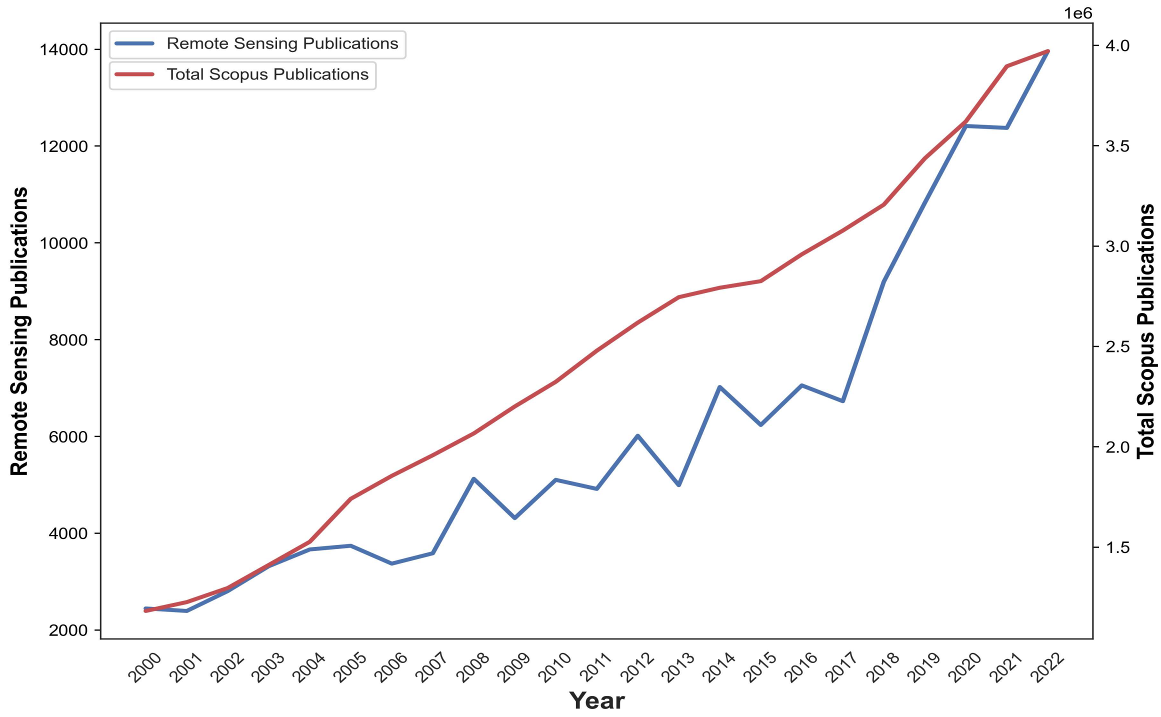


Figure 2.2 The evolution of total publications in Scopus (red line) and publications related to remote sensing in Earth and planetary and agricultural sciences (blue line) during 2000-2022.

In the second phase of the analysis, the focus moved from remote sensing applications to grassland and pasture monitoring and management topics and its ratio to the total number of publications in Scopus.

As shown in Figure 2.3, the number of publications concerning our research topic has increased in the last 20 years, increasing from 3 publications in 2000 to 93 in 2022. Additionally, the percentage of publications on the use of remote sensing for the management and monitoring of grasslands out of the total number of publications in Scopus increased significantly in 2022 and was more than two times greater than that in 2000. The increase in the number of publications is related to the general increase in the number of scientific publications and the growing interest in utilising remote sensing technologies, especially when related to environmentally relevant topics such as land management and monitoring of grasslands and pastures.

Figure 2.3 shows the number of publications on our research topic (blue histogram) and the ratio of the total number of publications in Scopus (red line plot). The trend over the years has increased, reaching its maximum for both the number of publications and the ratio in 2022.

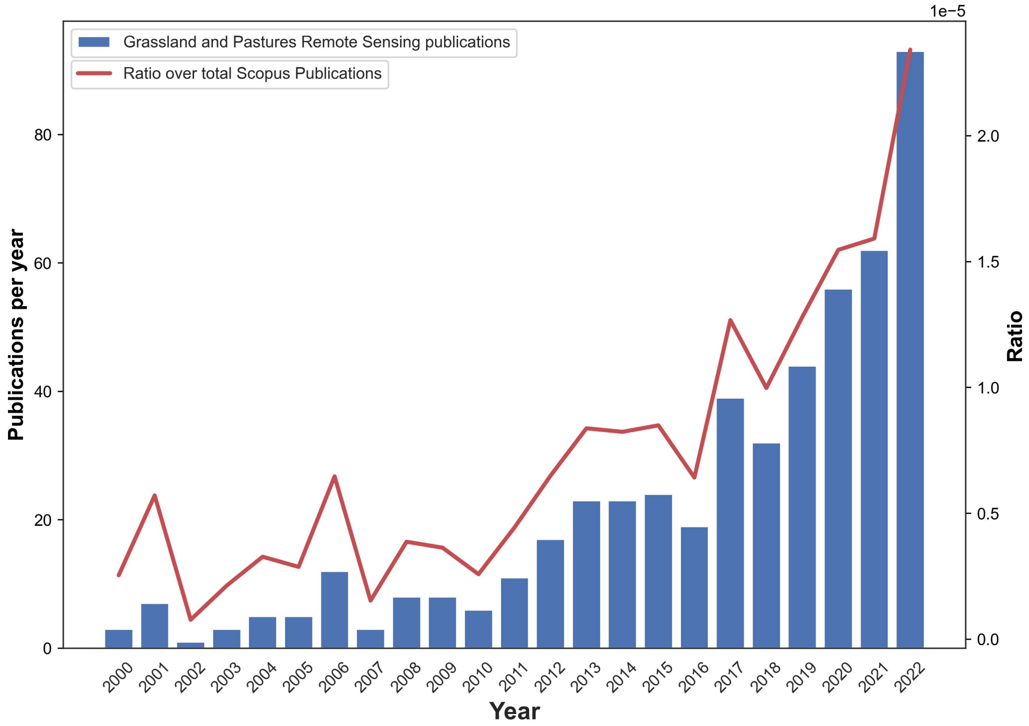


Figure 2.3 The evolution of publications related to remote sensing applied to grassland and pasture studies in the Earth and planetary and agricultural sciences topics (blue column) and their ratio by the number of total publications in Scopus during 2000-2022.

Another quantitative approach that has been used concerns the distribution of scientific publications across major contributing countries. The most important contributing country is China, with 142 publications corresponding to 28.2% of the total, followed by the United States with 77 (15.3%) and Germany with 51 (10.1%) publications. Table 2.4 lists the top 10 contributing countries in 2000-2022 and shows the evolution of the number of publications in each country.

Table 2.4 Lists of the top 10 contributing countries, their number of publications, percentage of the total and evolution in the period 2000-2022.

Country	**Total number of publications**	**Share of the total number of publications (%)**
China	142	28.2%
United States of America	77	15.3%
Germany	51	10.1%
Canada	31	6.2%
Australia	24	4.8%
Italy	24	4.8%
France	19	3.7%
Brazil	17	3.4%
South Africa	11	2.1%
New Zealand	10	2%
Others (38 countries)	98	19.4%

To compare the temporal trends of individual countries, Figure 2.4, highlighting the individual trend of the top 5 publishing countries, clearly shows how China has become the main protagonist actor over the last 20 years, passing from 0 publications in 2000-2004 to the absolute leader in recent years, quadrupling the number of publications of the second most important publishing country in 2022, with 36 publications versus 9 publications from the United States of America.

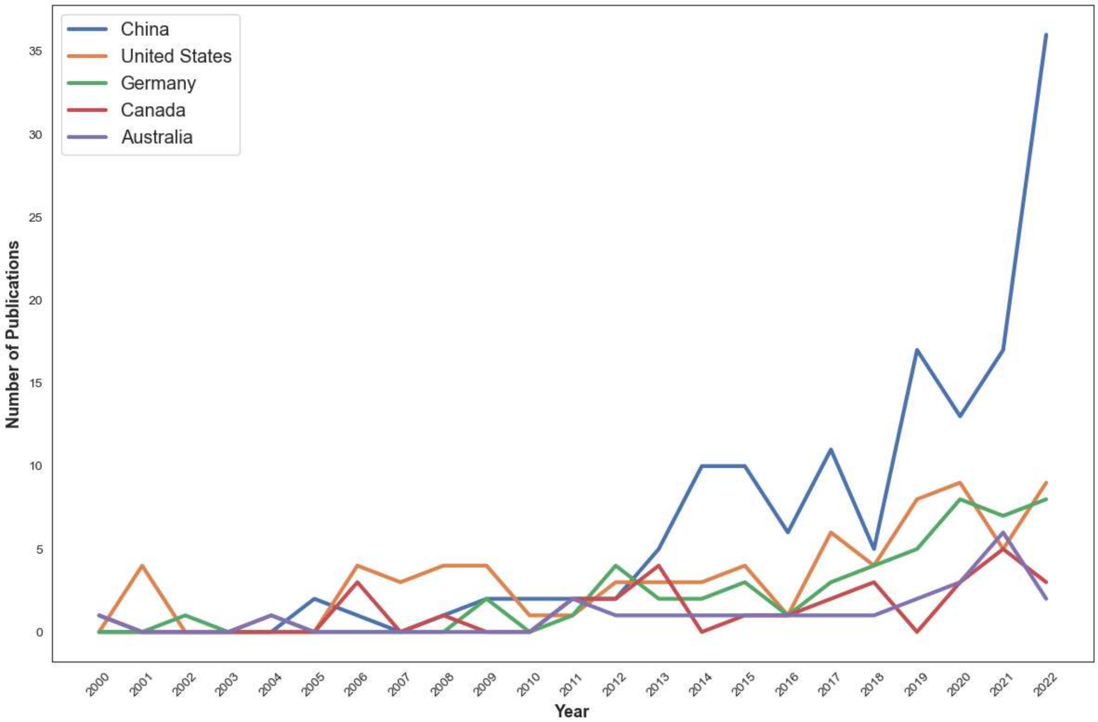


Figure 2.1 The evolution in terms of the number of publications during 2000-2022 of the top 5 contributing countries in regard to the research topic.

The journals that published the most papers on the topic were also analysed, and Table 5 reports the 10 most important journals by number of publications. The journal ‘Remote Sensing’ has a prominent role, as it published more than 20% of the articles, followed by the International Journal of Remote Sensing with 7.3% and Remote Sensing of Environment with 6.3%. Even if the publications are distributed across 159 different journals, the first five account for more than 44% of publications, showing that the topic is well identified and has a precise editorial placement.

Table 2.5. Lists of the top 10 publishing journals on the topic of remote sensing applied to grassland and pasture studies, their number of publications, percentage of total publications, impact factor, SJR and Cite Score (all metrics are available for 2022).

Journal	Total number of publications	Percentage	Impact Factor (2022)	SJR (2022)	Cite Score (2022)
Remote Sensing	103	20.4%	5.0	1.136	6.6
International Journal of Remote Sensing	37	7.3%	3.4	0.732	7.0
Remote Sensing of Environment	32	6.3%	13.5	4.057	24.8
Ecological Indicators	17	3.4%	6.9	1.396	10.3
International Journal of Applied Earth Observation and Geoinformation	17	3.4%	7.5	1.628	10.2
Canadian Journal of Remote Sensing	9	1.8%	2.6	0.619	3.9
GIScience and Remote Sensing	8	1.6%	6.7	1.340	9.4
ISPRS Journal of Photogrammetry and Remote Sensing	8	1.6%	12.7	3.308	19.2
Journal of Applied Remote Sensing	8	1.6%	1.7	0.388	3.4
Rangeland Ecology and Management	8	1.6%	2.3	0.776	4.6
Others (149 journals)	257	51%	-	-	-

2.4.2 Satellite Platforms and Machine Learning Quantitative Analysis

Satellite platforms and machine learning algorithms have a prominent role in the application of remote sensing technologies to characterise and monitor grasslands and pastures (Bangira et al., 2023; Ogungbuyi et al., 2023; Q. Wang et al., 2022).

A quantitative analysis was carried out to analyse and quantify the impact of individual satellite platforms and machine learning algorithms. UAVs are also important remote sensing data sources and were widely utilised in this review's research. The decision to focus on the analysis of satellite platforms was made after a preliminary quantitative analysis of the considered scientific articles. The number of articles including the words "UAV", "drone" or "unmanned aerial vehicle" in their abstracts was 55, 11% of the total, while the number of articles considering at least one of the considered satellite platforms was 445, 88% of the total.

Figure 2.5 shows the results of the quantitative analysis of the number of studies in which each satellite platform was considered. The 10 most recurring platforms are plotted. MODIS is the most utilised, appearing in 160 papers, followed by Sentinel-2 with 99 appearances and Landsat (all the Landsat satellites were analysed together), with 81. Among the other platforms not specifically cited in the plot, it is important to highlight the presence of Quickbird, which has been working since 2001, and the Gaofen constellation, a Chinese-based program first launched in 2013 that includes multispectral, radar and electro–optical sensors, which will become popular in the coming years.

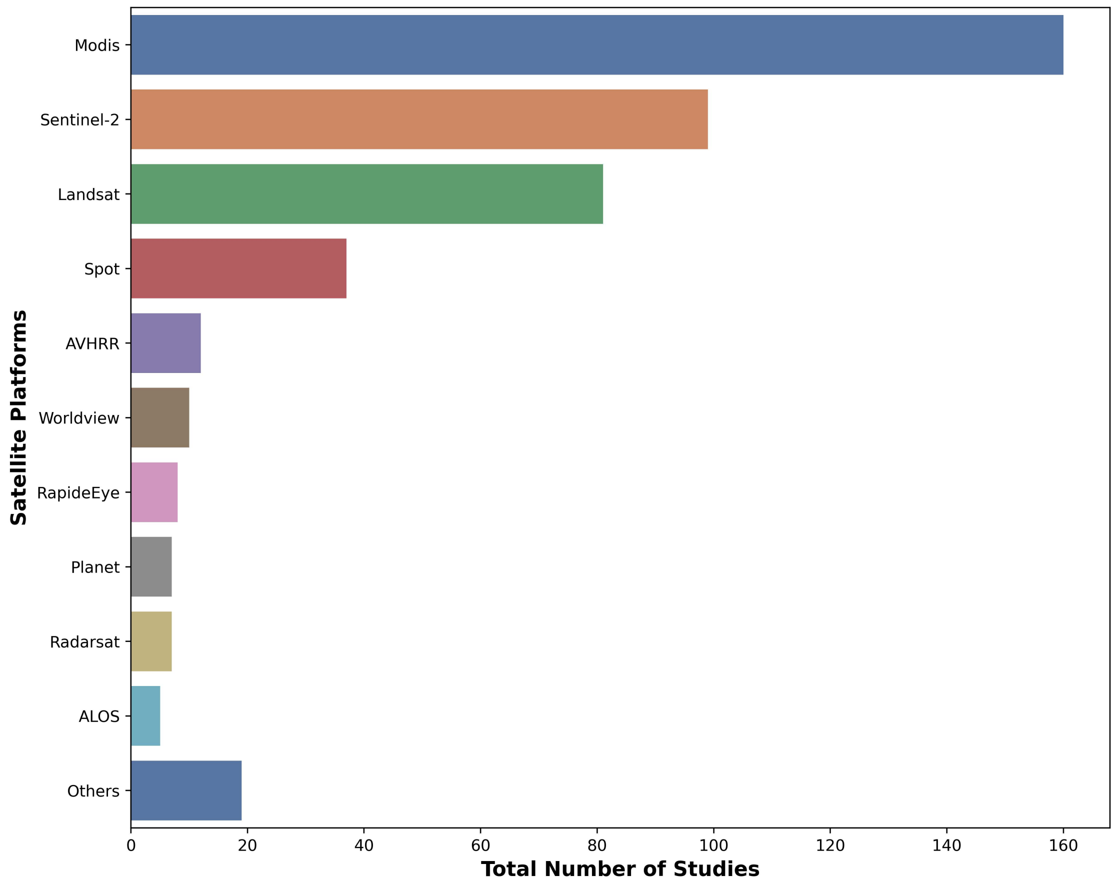


Figure 2.2 The top 10 utilised satellite platforms according to the number of studies in the field of remote sensing applied to grassland and pasture studies in 2000-2022.

Considering the evolution of the employment of different satellites in the characterisation and monitoring of grasslands and pastures over time, as shown in Figure 2.6, MODIS had led for many years, from 2008 to 2019, but since its launch on the 23rd of June 2015, Sentinel-2 has rapidly gained popularity among scientists and researchers, and since 2020, it has been the most utilised satellite platform in this field. In 2022, Sentinel-2 was the most considered satellite, utilised in 32 studies, almost double that of Modis, which stopped at 17.

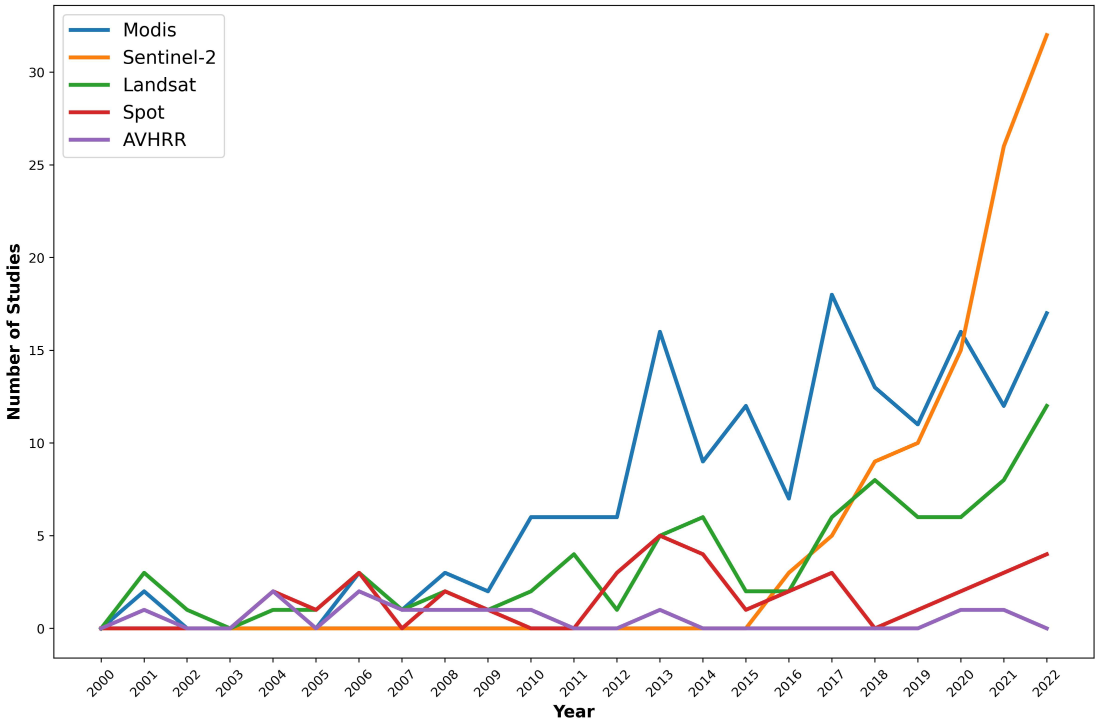


Figure 2.3. The evolution in terms of the number of studies during 2000-2022 of the top 5 satellite platforms in the research area.

Using satellite platforms by researchers in grassland and pasture studies reveals notable differences based on the authors' geographical origins. Figure 2.7 illustrates the proportional usage of various satellite platforms by different countries. Each pie chart on the map represents the distribution of satellite platform usage for a particular country, with the chart size corresponding to the total number of publications involving these platforms. Countries such as China, the USA, Canada, and Brazil predominantly rely on NASA's Modis and Landsat platforms. Modis is the most utilised platform in these countries, appearing in 59%, 48%, 59%, and 36% of the studies, respectively. Landsat also plays a significant role, being used in 18%, 29%, 20%, and 28% of the studies in these countries.
Conversely, Australia, South Africa, and European countries such as Germany, Italy, and France display a more balanced distribution across several platforms, with a relevant presence of Sentinel-2 and Spot. This distribution highlights the diverse approaches and capabilities of each country in utilising satellite technology for their respective needs.

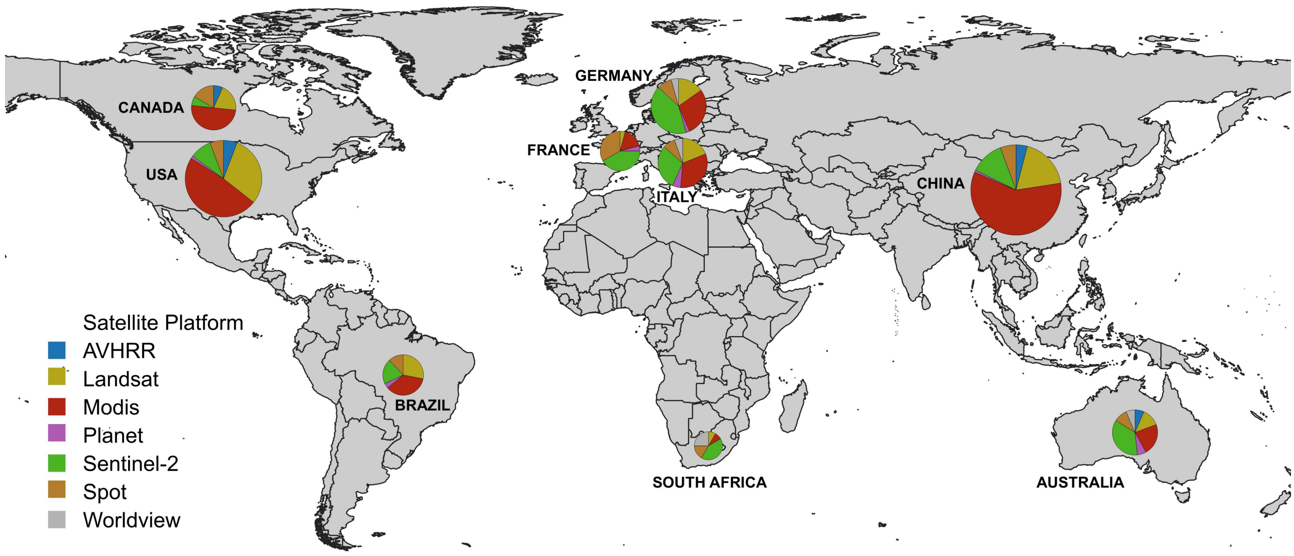


Figure 2.4. The proportion of the utilisation in grassland and pasture studies of satellite platforms in the most important contributing countries, the size of the pie chart is proportional to the total number of publications involving satellites platforms.

The use of advanced statistical modelling techniques as machine learning algorithms has gained much importance throughout the scientific world in recent years (Morais et al., 2023, 2021). The application of remote sensing technologies to the monitoring and characterisation of pastures and grasslands often requires the manipulation of large amounts of data, and machine learning algorithms are increasingly the most utilised tools by scientists and professionals in the sector (Smith et al., 2023; Xu et al., 2020). Figure 2.8 shows the results of the quantitative analysis of the number of studies in which each machine learning algorithm was utilised, and the 10 most recurring algorithms are plotted. Random Forest was the most common, with 66 papers, followed by linear regression and support vector machine, with 28 papers each.

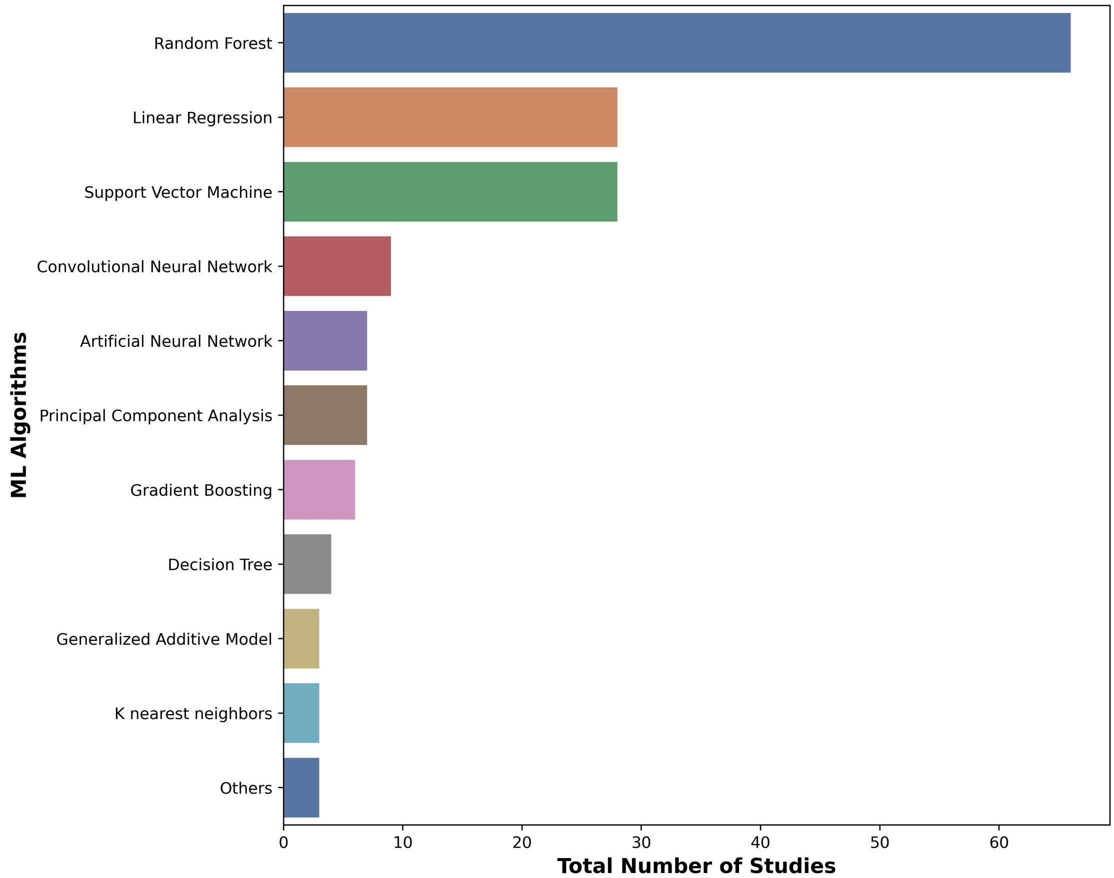


Figure 2.5 The top 10 utilised machine learning algorithms according to the number of studies in the field of remote sensing applied to grassland and pasture studies in 2000-2022.

Considering the evolution of machine learning algorithm utilisation over the past 20 years, it is important to note that, aside from linear regression, the first studies applying these techniques to monitor and characterise grasslands and pastures emerged only from 2012 onwards. Random Forest has gained significant popularity in the scientific community, with its usage increasing rapidly in recent years. Notably, these studies, excluding those involving linear regression, reached their peak in 2022, with 22 publications focusing on grassland and pasture monitoring and characterisation.

2.4.3. Cluster Analysis

The last part of the review focused on analysing the most recurring words in titles, abstracts and keywords of the published journal articles concerning the applications of remote sensing to characterise and monitor grasslands and pastures.

The titles, abstracts and keywords were processed using natural language processing techniques, and a list of words and their number of occurrences was extracted. As some bigrams, made of two words, can have great relevance considered together (for example, "Remote Sensing" or "Climate Change"), our research considered bigrams occurring more than 30 times as single words.

The first 70 terms and bigrams among the pre-processed terms were grouped into five conceptual clusters. The weight of a cluster was determined by the sum of the weights of the terms that belong to it over the total number of term occurrences.

The cluster with the highest weight was that with the theme "Instruments", at 25.83%, which included all words and concepts regarding all the tools involved in the applications of remote sensing technologies to the monitoring and characterisation of grasslands and pastures, including both data collection tools (e.g., 'Remote Sensing', 'Satellite', and 'UAV') and statistical data processing and modelling instruments (i.e., 'Model', 'Regression', and 'Algorithm'). The most common terms in the cluster were, in descending order, model (20%), remote sensing (13.8%), image (11.3%), satellite (7.9%), MODIS (7.7%), Sentinel (7.4%), R (5.2%), map (4.8%) and Landsat (3.5%). Even the "Parameters" cluster had a similar weight, 24.96%, as shown in Table *6*.

The cluster "Parameters" included the terms and concepts linked to the applications of remote sensing for grasslands and pastures that can calibrate and evaluate the quality of the technological applications, such as the vegetation index (11.9%), accuracy (9.9%), index (9.9%), NDVI (8.2%), AGB (7%), productivity (6.2%) and time series (4.6%).

As reported in Table 2.6, the "Vegetation" cluster was only slightly less important, with a weight of 24.15%. We collected words and bigrams concerning grassland and pasture vegetation and habitats and their components, such as grassland (51.9%), pasture (13.5%), biomass (8.4%), species (5.6%), plant (3.9%), grass (3.9%) and phenology (5%). The "Environment" (13.11%) and "Management" (11.94%) clusters had lower but still significant weights. The "Environment" cluster contains those terms and concepts related to the environmental elements and processes that characterise and affect grasslands and pastures; the most important terms in the cluster are soil (11.4%), ecosystem (11.3%), degradation (8.8%), water (8.7%), carbon (7.2%), ground (6.6%), drought (5.9%), land (5.9%), climate (5.7%) and ecological (5.6%). In the "Management" cluster, the terms that represent the most important actions involved in the applications of remote sensing to grassland and pasture monitoring and characterisation are estimation (34.9%), monitoring (17%), classification (13.9%), grazing (10.3%), mapping (8.5%) and conservation (5.5%).

Table 2.6 Clusters and their composing terms reported by highest occurrence frequency.

Cluster	Terms and Bigrams

Instruments (25.8%)	Model (20%), Remote Sensing (13.8%), Image (11.3%), Satellite (7.9%), Modis (7.7%), Sentinel (7.4%), R (5.2%), Map (4.8%), Landsat (3.5%), Regression (3.3%), Algorithm (2.9%), Random Forest (2.7%), UAV (2.6%), Correlation (2.3%), Hyperspectral (2.1%), Multispectral (2%)
Parameters (24.9%)	Vegetation Index (11.9%), Accuracy (9.9%), Index (8.6%), NDVI (8.2%), AGB (7%), Productivity (6.2%), Spatial (6%), Spectral (4.8%), Time Series (4.6%), Normalized Difference (3.7%), LAI (3.4%), GPP (3.1%), Resolution (3%), Temporal (3%), Variability (2.8%), Reflectance (2.5%), Optical (2.3%), Bands (2.3%), Seasonal (2.2%), EVI (2.2%), RMSE (2.1%)
Vegetation (24.1%)	Grassland (51.9%), Pasture (13.5%), Biomass (8.4%), Species (5.6%), Plant (3.9%), Grass (3.9%), Phenology (5%), Agricultural (2.8%), Canopy (2.4%), Forage (2.3%)
Environment (13.1%)	Soil (11.4%), Ecosystem (11.3%), Degradation (8.8%), Water (8.7%), Carbon (7.2%), Ground (6.6%), Drought (5.9%), Land (5.9%), Climate (5.7%), Ecological (5.6%), Precipitation (5.2%), Fire (4.8%), Steppe (4.42%), Temperature (4.4%), Alpine (4.1%)
Management (11.9%)	Estimation (34.9%), Monitoring (17%), Classification (13.9%), Grazing (10.3%), Mapping (8.5%), Conservation (5.5%), Livestock (4.9%), N (4.8%)

Thus, for each of the 70 most commonly used words in the title, keywords, and abstract sections, pairings were made with the other 69 words, resulting in 1,447 potential pairings. These pairings were analysed based on their frequency in the literature over 20 years and depicted visually, creating a complex network of relationships (Figure 2.8). Additionally, the occurrences of these combinations grouped by cluster are also presented in tabular form (Table 2.7).

Table 2.7 Co-occurrences of terms within and between clusters.

	Environment	**Instruments**	**Management**	**Parameters**	**Vegetation**
Environment	1986	1382	1003	2146	1669
Instruments	-	1186	1020	2031	1530
Management	-	-	664	1530	1208
Parameters	-	-	-	3932	2312
Vegetation	-	-	-	-	1412

Given the centrality of the term “grassland” in the research topic and its presence among the keyword used in the Scopus query, the “Vegetation” cluster was expected to have the highest number of co-occurrences. In contrast to expectations, the “Parameters” cluster showed the greatest number of co-occurrences, both between terms inside the cluster and those belonging to different clusters. The “Environment” and “Vegetation” clusters were second and third, respectively, in terms of co-occurrences both within and between the clusters. However, the reasons for their similar weights were different: while the environment cluster contained a large number of terms with very balanced weights. The vegetation cluster was characterised by the presence of the term “grassland”, also present among the keywords in the Scopus query, that ended as the most important and common term in the whole research topic, as clearly visible in Figure 2.9.

Interestingly, although the 'Instruments' cluster had the highest total occurrences as shown in Table 5, it ranked fourth in terms of co-occurrences, both within the cluster and with other clusters. Finally, the "Management" cluster was identified as having the least interactions with other terms, also showing the lowest number of connections among words within the same cluster.

As shown in Figure 2.9, the term showing the strongest connections was grassland. The connection between ‘grassland’ and ‘spatial’ was the strongest, but ‘model’, ‘satellite’, ‘monitoring’, ‘index’, ‘accuracy’ and ‘biomass’ also showed strong relationships with the central term of the topic. It is not surprising to see how these terms and their combinations are central to the research topic, as they represent some of the main instruments, parameters and management operations that are central in characterising and monitoring grasslands and pasture environments. Overall, despite the 'Instruments' cluster having more occurrences in the initial analysis, the 'Parameters' cluster showed greater centrality in terms of word relationships and bigrams.

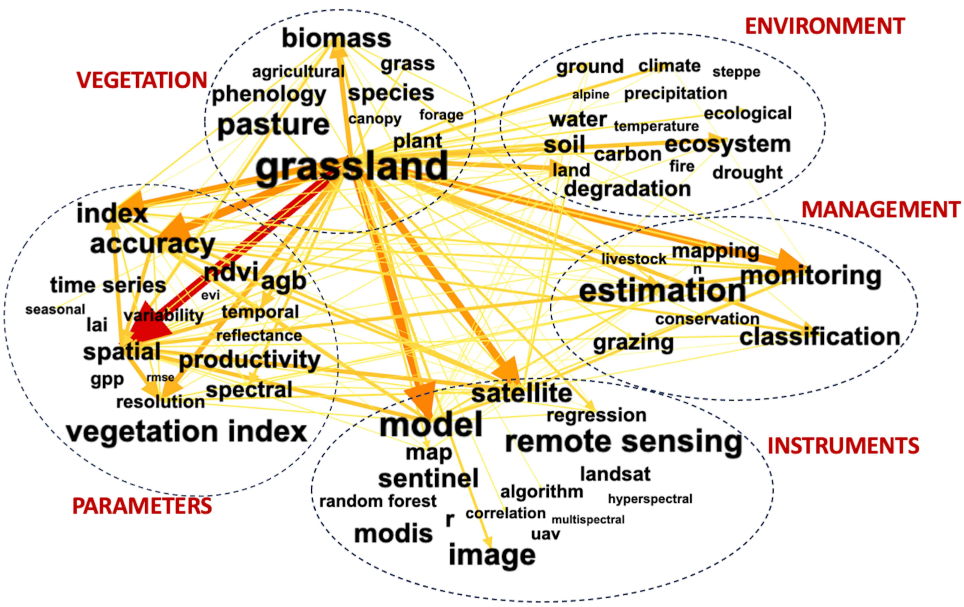


Figure 2.6 Graphical representation of the 70 most recurring words in the research topic. The font size is adjusted to be proportional to the number of occurrences, and the thickness and darkness of the connecting arrows represent the number of co-occurrences between linked terms.

The analysis of single pairs of terms, without considering the cluster to which they belong, allowed us to show which topics were the most related.

Table 2.8 was constructed by taking the first 30 pairs of terms by relationships. The term "grassland" and its relationships were excluded because they would hegemonise the table, making the analysis less interesting and significant. Some of the most recurring terms were "spatial", "Model", "index", "accuracy" and "resolution", suggesting that the focus of studies related to the characterisation and monitoring of grasslands and pastures using remote sensing is the development of tools and the evaluation and tuning of their technical parameters.

Table 2.8 Couples of terms with the highest co-occurrences (excluding 'grassland').

Source	**Target**	**Weight**	**Source**	**Target**	**Weight**
Spatial	Resolution	86	Monitoring	Index	54
Satellite	Spatial	82	Monitoring	Accuracy	54
Model	Spatial	81	Spatial	Estimation	53
Spatial	Index	80	Satellite	Resolution	53
Accuracy	Spatial	78	Satellite	Land	52
Spatial	Temporal	78	Biomass	Estimation	50
Monitoring	Spatial	72	Index	Resolution	49
Model	Accuracy	71	Pasture	Satellite	49
Model	Biomass	67	Model	Monitoring	48
Model	Index	65	Accuracy	Index	48
Accuracy	Satellite	65	Spatial	Soil	47
Biomass	Spatial	63	Index	Temporal	47
Monitoring	Satellite	63	Spatial	Land	47

Accuracy	Classification	61	Biomass	Index	47
Model	Estimation	54	Monitoring	Resolution	46

2.5 Discussion

The research utilised text-mining techniques to analyse the titles, abstracts, and keywords of each article. Key terms that were both critical and frequent were pinpointed and examined. Noteworthy connections were identified among specific terms as well as between aggregated clusters. Through a temporal analysis, the evolution of publications was tracked, highlighting topics that have become significant over time and relationships that have grown stronger.

2.5.1 Analysis of publication trends

Themes related to the application of remote sensing to Earth and agricultural sciences and, more specifically, to the characterisation and monitoring of grasslands and pastures, are relatively recent topics. During the last 22 years (2000-2022), interest in this topic has risen sharply, and the number of publications on remote sensing in the Earth and agricultural sciences grew slightly less than 11 times and more than 3 times the weight of total articles published in Scopus during the same period, respectively.

Analysing the publications concerning the applications of remote sensing to the characterisation and monitoring of grasslands and pastures, the growth was even sharper, with the number of publications growing by 31 times and the ratio of total Scopus publications growing by more than 9 times. Studies within this field have explored all facets related to the topic from various perspectives, as illustrated by the cluster analysis, with a focus on instruments and parameters.

The analysis of the geographic distribution of the literature carried out through the examination of the authors' affiliations (Table 3) revealed that the most active nations were China (28.3%), the USA (15.3%) and Germany (10.1%). China is unquestionably the largest contributor, with Figure 4 indicating a rapidly increasing number of publications. The analysis of the editorial distribution of publications highlighted that the selected topic has a precise editorial definition and allocation, as 44% of published articles are in three journals, with "Remote Sensing" collecting 20.4% of the considered articles on its pages.

2.5.2 Quantitative Analysis of Satellite Platforms and Machine Learning Algorithms

Satellite platforms are the most utilised source of remote sensing data (Ali et al., 2016; Clementini et al., 2020; Woodward et al., 2019), especially when related to the characterisation and monitoring of grasslands and pastures. Of the 503 papers collected for our analysis, 88% contained the word "satellite" in their title, abstract or keywords.

The quantitative analyses of the utilisation of the individual satellite platforms allowed us to identify some important facts, as shown in Figure 2.5. Due to the high versatility, daily revisit time and availability of multiple pre-processed products, the Moderate Resolution Imaging Spectroradiometer (MODIS), which is mounted on the Terra (launched in 1999) and Aqua (launched in 2002) satellites from NASA, is the most utilised tool in grassland and pasture studies.

However, when analysing the temporal trend in the use of satellite platforms, it is worth noting that Sentinel-2, which was launched in 2015 by the ESA, has gained massive popularity in the scientific community and has become the most utilised platform in recent years.

The availability of 13 spectral bands ranging from 443 to 2190 nm at high spatial resolution, ranging from 10 to 60 m, with a revisit time of 5 days, makes it a versatile and very suitable tool for Earth observation and monitoring studies (Fauvel et al., 2020; Hartmann et al., 2023; Kolecka et al., 2018; Jie Wang et al., 2019). Despite the high spatial resolution, the data from these satellites suffer from long time intervals and cloud cover interference (Whitcraft et al., 2015). Several studies have fused data from different satellite sources to enhance temporal continuity and fill in data gaps (Forsmoo et al., 2018). In smaller grassland fields, the need for high-resolution images is critical, and commercially available platforms such as Planet Scope and Worldview have gained significant importance in biomass estimation studies (Clementini et al., 2020; Schellberg et al., 2008). They offer daily revisit times with great spatial resolution (1.8 m for Worldview and 3 m for Planet Scope), even though the number of multispectral bands is still limited compared to that of Landsat and Sentinel platforms. While all satellite platforms provide valuable data for grassland and pasture monitoring, it is crucial to recognise the limitations and trade-offs associated with varying resolution levels. Mid- to low-resolution products, such as those from MODIS, are well-suited for large-scale applications due to their extensive coverage and frequent revisit times, which are particularly useful in global and regional monitoring efforts (Dieguez and Pereira, 2020; Reeves et al., 2021; Shrestha et al., 2020). However, these products may need more detailed precision for more localised studies, where fine-scale vegetation characteristics and management practices must be accurately assessed (Klingler et al., 2020; Xu et al., 2018). Conversely, high-resolution data from platforms like Sentinel-2, PlanetScope, and Worldview are essential for small-scale, detailed analyses, such as distinguishing pasture species or monitoring phenological changes at the paddock level (Marston et al., 2017; Raab et al., 2020; Shoko et al., 2019). Nonetheless, these high-resolution datasets come with increased resource demands, including higher data acquisition costs and processing time, which can limit their feasibility for widespread use (Eastwood et al., 2020). Therefore, selecting the appropriate resolution depends on the specific scale and objectives of the study, balancing the need for detailed information against the practical considerations of resource availability and processing capabilities.

With the rise in the availability of large volumes of raw and processed data, there is a growing need for statistical methods capable of handling this data and extracting significant and useful information. Machine learning algorithms have been increasingly utilised in studies related to the application of remote sensing technologies to the characterisation and monitoring of grasslands and pastures (De Rosa et al., 2021; Ford et al., 2023; Meng et al., 2020), as shown in Figure 2.8. The choice of data analysis methods is influenced by several factors, such as the size and heterogeneity of the study area, the number of predictor variables, and the study's objectives. Interestingly, random forest (Breiman, 2001) and linear regression, the most utilised methods, have very different characteristics. Linear regression has a better capacity for generalising the relationships among variables, thus producing robust models that can perform well on different datasets. On the other hand, the random forest algorithm is known to adapt more to the training dataset; it can detect and model complex relationships among multiple variables. These features make it very suitable for modelling biological variables such as biomass production, but at the same time, it has an intrinsic tendency to overfit the data (Hawkins, 2004), producing models that do not generalise well and thus are not useful for practical purposes (Hastie et al., 2009). Although some authors, such as Smith et al. (Smith et al., 2023), have recently attempted to assess the performance and transferability of several machine learning algorithms (particularly PCR, PLSR, LASSO, RF, SVM and GBM), the variability and complexity of grassland and pasture environments and features need further investigation and research.

2.5.3 Cluster and Network Analysis

The described cluster analysis shows that, during the last 20 years, the focus has been on the instruments and the evaluation and tuning of the parameters that are involved in the development of tools able to estimate, evaluate and monitor environmental variables, mainly vegetation characteristics such as biomass content and quality. The high importance of the "Instruments" and "Parameters" clusters, visible in Table 2.5, perfectly highlights the abovementioned concepts, and the focus of scientists involved in the applications of remote sensing to grassland and pasture studies has been on the technologies and their development. The analysis of the most significant relationships, apart from the centrality of "grassland", confirmed the centrality and importance of the "Parameters" cluster, but the increasing importance of the "Environment" cluster, characterised by the presence of a high quantity of words, was quite balanced in terms of occurrences.

2.6 Conclusions

Over the past two decades, there has been an increasing focus on utilising remote sensing technologies to map and monitor grasslands and pastures. This surge in interest has led to a deeper exploration of various and distinct topics within the broad domain of remote sensing applications in Earth and agricultural sciences. This research seeks to map these evolving trends by identifying and emphasizing the most significant terms or relationships based on their frequency in scientific publications.

The most important contributions are in three macroareas, China, North America, and Europe, while developing countries are less represented. Although the publications are spread across 159 different journals, the top five journals contribute to more than 44% of the publications. This indicates that the topic is well defined and possesses a clear editorial niche.

The quantitative analysis of satellite platforms showed that Moderate Resolution Imaging Spectroradiometer (MODIS) data are the most widespread among grassland and pasture studies, mainly because of its availability and daily revisit time, even though Sentinel-2, characterised by a higher spatial resolution and relatively low revisit time (5 days), has become the most common choice in recent years.

Additionally, the results of the review suggest that future efforts might focus on the implementation, testing, and tuning of instruments and parameters related to mapping and monitoring tasks, especially with a focus on biomass estimation and land degradation monitoring. The major concern that researchers are facing regards the development of models and tools that reach a high level of accuracy. The increasing availability of large amounts of data from satellite sources will demand tools that can analyse and extract valuable information. Thus, machine learning algorithms will be increasingly involved. The quantitative analysis of machine learning algorithms highlighted that the most commonly used algorithm for grassland and pasture studies is random forest, followed by linear regression and support vector machine methods.

Developing accurate and precise tools will aid farmers and policymakers in better managing grassland and pasture environments (French et al., 2015). This improvement will enhance these areas' agricultural and environmental roles. It is particularly crucial in regions where grasslands are a fundamental resource (Brilli et al., 2023). Despite the growing availability of free or cost-effective satellite data from both public entities (such as NASA, ESA, CNMS, etc.) and private companies (including AgroInsider, SPACETM, DataFarming, GeoGraze, pasture.io, etc.), leveraging this data for effective grassland monitoring involves significant associated costs. These costs include data preprocessing, which requires computing infrastructure for downloading, storing, and processing raw satellite data into analysis-ready formats. The entry of high-tech corporations with services and platforms such as Microsoft Planetary Computer, Google Earth Engine, Amazon Web Services, and Oracle Cloud Infrastructure, which offer cloud computing solutions to support digital agriculture and earth monitoring platforms, may facilitate the development of user-friendly tools and services as GeoServer platforms, but involving direct costs for users and developers (Eastwood et al., 2020). For example, while Microsoft Planetary Computer offers a free platform, users might still face costs depending on the amount of data processed and stored. Processing large datasets can necessitate an Azure Cloud Services subscription, with storage fees ranging from $0.003 to $0.15 per GB per month and data processing costs between $0.35 and $1.40 per GB per hour, depending on the service tier. Similarly, Google Earth Engine provides a free tier for researchers, but commercial users may encounter significant expenses. For instance, Google Cloud Platform charges between $0.006 and $0.023 per GB per month for data storage, while virtual machine costs start at $0.038 per vCPU per hour, offering scalable processing capacity. Amazon Web Services (AWS) also presents a flexible pricing structure, with storage costs ranging from $0.005 to $0.024 per GB per month. AWS provides a variety of virtual machine options, including free-tier services like AWS Lambda, but more robust solutions may come with additional costs. In comparison, Oracle Cloud Infrastructure offers competitive rates, with storage fees as low as $0.002 per GB per month, scaling up to $0.30 per GB for premium options. Their virtual machines are priced at $0.032 per vCPU per hour, making them a cost-effective choice for intensive data processing tasks. Alongside these global platforms, mostly USA-based, the European Union has developed a dedicated ecosystem of Earth Observation cloud infrastructures to reduce barriers to Copernicus data access. The Copernicus Data Space Ecosystem (CDSE), launched in 2023 under ESA coordination, serves as the official European gateway to the full Sentinel archive, offering free and open data access with monthly replenished processing credits and commercial packages for higher-demand users. Its underlying infrastructure is provided by CREODIAS (CloudFerro), which features pay-per-use virtual machines, free in-cloud EO data access, and €250 in starter credits for new users. WEkEO, one of five EU-funded DIAS platforms jointly operated by EUMETSAT, ECMWF, EEA, and Mercator Ocean International, provides federated access to Copernicus data alongside a free Essential tier and a no-cost Innovation Lab programme for eligible researchers. Paid tiers on WEkEO extend to GPU-enabled Virtual Machines and HPC resources on a pay-as-you-use basis.

Additionally, collecting ground data to calibrate and validate satellite models involves fieldwork expenses and laboratory analysis costs. Moreover, developing and refining machine learning algorithms demands computational resources and expert personnel, while operational monitoring necessitates ongoing staff support and robust data management systems.

Finally, the implemented semiautomatic methodology integrates natural language processing techniques to quantify the occurrence of concepts and their relationships within a specific research topic, allowing the identification and quantitative analysis of the main trends within a specific research topic and laying the path for similar work in the future.

Future efforts might focus on implementing, testing, and tuning instruments and parameters related to mapping and monitoring tasks, especially with a focus on biomass estimation and land degradation monitoring. The major concern that researchers are facing concerns the development of models and tools that reach a high level of accuracy. The increasing availability of large amounts of data from satellite sources will demand tools that can analyze and extract valuable information. Thus, machine learning algorithms will be increasingly involved.

Integrating data from various sources, advancing machine learning algorithms, assessing climate change impacts, conducting region-specific studies, developing user-friendly tools for stakeholders, and examining economic and environmental trade-offs are crucial areas for future research. These efforts will help improve the management of grassland and pasture environments, enhancing their agricultural and environmental roles.

2.7 References

Abad, J., Hermoso De Mendoza, I., Marín, D., Orcaray, L., Santesteban, L.G., 2021. Cover crops in viticulture. A systematic review (1): Implications on soil characteristics and biodiversity in vineyard. Oeno One 55, 295–312. https://doi.org/10.20870/OENO-ONE.2021.55.1.3599

Ali, I., Cawkwell, F., Dwyer, E., Barrett, B., Green, S., 2016. Satellite remote sensing of grasslands: From observation to management. Journal of Plant Ecology 9, 649–671. https://doi.org/10.1093/jpe/rtw005

Anderson, G.L., Hanson, J.D., Haas, R.H., 1993. Evaluating landsat thematic mapper derived vegetation indices for estimating above-ground biomass on semiarid rangelands. Remote Sens Environ 45, 165–175. https://doi.org/10.1016/0034-4257(93)90040-5

Andersson, K., Trotter, M., Robson, A., Schneider, D., Frizell, L., Saint, A., Lamb, D., Blore, C., 2017. Estimating pasture biomass with active optical sensors. Advances in Animal Biosciences 8, 754–757. https://doi.org/10.1017/S2040470017000838

Andreatta, D., Gianelle, D., Scotton, M., Dalponte, M., 2022a. Estimating grassland vegetation cover with remote sensing: A comparison between Landsat-8, Sentinel-2 and PlanetScope imagery. Ecol Indic 141. https://doi.org/10.1016/j.ecolind.2022.109102

Andreatta, D., Gianelle, D., Scotton, M., Vescovo, L., Dalponte, M., 2022b. Detection of grassland mowing frequency using time series of vegetation indices from Sentinel-2 imagery. GIsci Remote Sens 59, 481–500. https://doi.org/10.1080/15481603.2022.2036055

Bangira, T., Mutanga, O., Sibanda, M., Dube, T., Mabhaudhi, T., 2023. Remote Sensing Grassland Productivity Attributes: A Systematic Review. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15082043

Barnetson, J., Phinn, S., Scarth, P., 2020. Estimating Plant Pasture Biomass and Quality from UAV Imaging across Queensland's Rangelands. AgriEngineering 2, 523–543. https://doi.org/10.3390/agriengineering2040035

Bar-On, Y.M., Phillips, R., Milo, R., 2018. The biomass distribution on Earth. Proc Natl Acad Sci U S A 115, 6506–6511. https://doi.org/10.1073/pnas.1711842115

Bazzo, C.O.G., Kamali, B., Hütt, C., Bareth, G., Gaiser, T., 2023. A Review of Estimation Methods for Aboveground Biomass in Grasslands Using UAV. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15030639

Bergman, K.O., Ask, L., Askling, J., Ignell, H., Wahlman, H., Milberg, P., 2008. Importance of boreal grasslands in Sweden for butterfly diversity and effects of local and landscape habitat factors. Biodivers Conserv 17, 139–153. https://doi.org/10.1007/s10531-007-9235-x

Boscaro, D., Pezzuolo, A., Grigolato, S., Cavalli, R., Marinello, F., Sartori, L., 2015. Preliminary analysis on mowing and harvesting grass along riverbanks for the supply of anaerobic digestion plants in

North-Eastern Italy. Journal of Agricultural Engineering 46, 100–104. https://doi.org/10.4081/jae.2015.465

Breiman, L., 2001. Random Forests. Mach Learn 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brilli, L., Martin, R., Argenti, G., Bassignana, M., Bindi, M., Bonet, R., Choler, P., Cremonese, E., Della Vedova, M., Dibari, C., Filippa, G., Galvagno, M., Leolini, L., Moriondo, M., Piccot, A., Stendardi, L., Targetti, S., Bellocchi, G., 2023. Uncertainties in the adaptation of alpine pastures to climate change based on remote sensing products and modelling. J Environ Manage 336. https://doi.org/10.1016/j.jenvman.2023.117575

Cai, Z., Junttila, S., Holst, J., Jin, H., Ardö, J., Ibrom, A., Peichl, M., Mölder, M., Jönsson, P., Rinne, J., Karamihalaki, M., Eklundh, L., 2021. Modelling daily gross primary productivity with sentinel-2 data in the nordic region–comparison with data from modis. Remote Sens (Basel) 13, 1–18. https://doi.org/10.3390/rs13030469

Cisneros, A., Fiorio, P., Menezes, P., Pasqualotto, N., van Wittenberghe, S., Bayma, G., Nogueira, S.F., 2020. Mapping productivity and essential biophysical parameters of cultivated tropical grasslands from sentinel-2 imagery. Agronomy 10. https://doi.org/10.3390/agronomy10050711

Clementini, C., Pomente, A., Latini, D., Kanamaru, H., Vuolo, M.R., Heureux, A., Fujisawa, M., Schiavon, G., Frate, F. Del, 2020. Long-term grass biomass estimation of pastures from satellite data. Remote Sens (Basel) 12. https://doi.org/10.3390/rs12132160

Cogato, A., Cei, L., Marinello, F., Pezzuolo, A., 2023. The Role of Buildings in Rural Areas: Trends, Challenges, and Innovations for Sustainable Development. Agronomy 13. https://doi.org/10.3390/agronomy13081961

Cogato, A., Meggio, F., Migliorati, M.D.A., Marinello, F., 2019. Extreme weather events in agriculture: A systematic review. Sustainability (Switzerland) 11, 1–18. https://doi.org/10.3390/su11092547

Cortner, O., Garrett, R.D., Valentim, J.F., Ferreira, J., Niles, M.T., Reis, J., Gil, J., 2019. Perceptions of integrated crop-livestock systems for sustainable intensification in the Brazilian Amazon. Land use policy 82, 841–853. https://doi.org/10.1016/J.LANDUSEPOL.2019.01.006

Cui, K., Li, R., Polk, S.L., Lin, Y., Zhang, H., Murphy, J.M., Plemmons, R.J., Chan, R.H., 2024. Superpixel-Based and Spatially Regularized Diffusion Learning for Unsupervised Hyperspectral Image Clustering. IEEE Transactions on Geoscience and Remote Sensing 62, 1–18. https://doi.org/10.1109/TGRS.2024.3385202

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2018. Predicting methane yield by linear regression models: A validation study for grassland biomass. Bioresour Technol 265, 372–379. https://doi.org/10.1016/j.biortech.2018.06.030

Dass, P., Houlton, B.Z., Wang, Y., Warlind, D., 2018. Grasslands may be more reliable carbon sinks than forests in California. Environmental Research Letters 13. https://doi.org/10.1088/1748-9326/aacb39

De Rosa, D., Basso, B., Fasiolo, M., Friedl, J., Fulkerson, B., Grace, P.R., Rowlings, D.W., 2021. Predicting pasture biomass using a statistical model and machine learning algorithm implemented with remotely sensed imagery. Comput Electron Agric 180. https://doi.org/10.1016/j.compag.2020.105880

De Vroey, M., Radoux, J., Defourny, P., 2023. Classifying Sub-Parcel Grassland Management Practices by Optical and Microwave Remote Sensing. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15010181

Derner, J.D., Schuman, G.E., 2007. Carbon sequestration and rangelands: A synthesis of land management and precipitation effects. J Soil Water Conserv 62, 77.

Dibari, C., Costafreda-Aumedes, S., Argenti, G., Bindi, M., Carotenuto, F., Moriondo, M., Padovan, G., Pardini, A., Staglianò, N., Vagnoli, C., Brilli, L., n.d. Expected Changes to Alpine Pastures in Extent and Composition under Future Climate Conditions. https://doi.org/10.3390/agronomy10070926

Dibari, C., Pulina, A., Argenti, G., Aglietti, C., Bindi, M., Moriondo, M., Mula, L., Pasqui, M., Seddaiu, G., Roggero, P.P., 2021a. Climate change impacts on the alpine, continental and mediterranean grassland systems of italy: A review. Italian Journal of Agronomy 16. https://doi.org/10.4081/ija.2021.1843

Dibari, C., Pulina, A., Argenti, G., Aglietti, C., Bindi, M., Moriondo, M., Mula, L., Pasqui, M., Seddaiu, G., Roggero, P.P., 2021b. Climate change impacts on the alpine, continental and mediterranean grassland systems of italy: A review. Italian Journal of Agronomy 16. https://doi.org/10.4081/ija.2021.1843

Dieguez, F.J., Pereira, M., 2020. Uruguayan native grasslands net aerial primary production model and its application on safe stocking rate concept. Ecol Modell 430, 109060. https://doi.org/10.1016/J.ECOLMODEL.2020.109060

Eastwood, C., Dela Rue, B., Kerslake, J., 2020. Developing an approach to assess farmer perceptions of the value of pasture assessment technologies. Grass and Forage Science 75, 474–485. https://doi.org/https://doi.org/10.1111/gfs.12504

Egan, P.A., Price, M.F., 2017. Mountain Ecosystem Services and Climate Change A Global Overview of Potential Threats and Strategies for Adaptation. UNESCO, Paris.

Ehret, T., De Truchis, A., Mazzolini, M., Morel, J.M., D'Aspremont, A., Lauvaux, T., Duren, R., Cusworth, D., Facciolo, G., 2022. Global Tracking and Quantification of Oil and Gas Methane Emissions from Recurrent Sentinel-2 Imagery. Environ Sci Technol 56, 10517–10529. https://doi.org/10.1021/acs.est.1c08575

Fauvel, M., Lopes, M., Dubo, T., Rivers-Moore, J., Frison, P.L., Gross, N., Ouin, A., 2020. Prediction of plant diversity in grasslands using Sentinel-1 and -2 satellite image time series. Remote Sens Environ 237, 111536. https://doi.org/10.1016/J.RSE.2019.111536

Ford, J., Sadgrove, E., Paul, D., 2023. Developing an extreme learning machine based approach to weed segmentation in pastures. Smart Agricultural Technology 5. https://doi.org/10.1016/j.atech.2023.100288

Forsmoo, J., Anderson, K., Christopher, |, Macleod, J.A., Wilkinson, M.E., Brazier, | Richard, 2018. Drone-based structure-from-motion photogrammetry captures grassland sward height variability. J Appl Ecol 55, 2587–2599. https://doi.org/10.1111/1365-2664.13148

Franzluebbers, A.J., Wendroth, O., Creamer, N.G., Feng, G.G., 2020. Focusing the future of farming on agroecology. Agricultural & Environmental Letters 5, e20034. https://doi.org/https://doi.org/10.1002/ael2.20034

French, P., O'Brien, B., Shalloo, L., 2015. Development and adoption of new technologies to increase the efficiency and sustainability of pasture-based systems. Anim Prod Sci 55, 931–935. https://doi.org/https://doi.org/10.1071/AN14896

Gao, X., Dong, S., Li, S., Xu, Y., Liu, S., Zhao, H., Yeomans, J., Li, Y., Shen, H., Wu, S., Wu, S., Zhi, Y., 2020. Using the random forest model and validated MODIS with the field spectrometer measurement promote the accuracy of estimating aboveground biomass and coverage of alpine grasslands on the Qinghai-Tibetan Plateau. Ecol Indic 112. https://doi.org/10.1016/j.ecolind.2020.106114

Guillaume, T., Makowski, D., Libohova, Z., Elfouki, S., Fontana, M., Leifeld, J., Bragazza, L., Sinaj, S., 2022. Carbon storage in agricultural topsoils and subsoils is promoted by including temporary grasslands into the crop rotation. Geoderma 422, 115937. https://doi.org/10.1016/J.GEODERMA.2022.115937

Hartmann, A., Sudmanns, M., Augustin, H., Baraldi, A., Tiede, D., 2023. Estimating the temporal heterogeneity of mowing events on grassland for haymilk-production using Sentinel-2 and greenness-index. Smart Agricultural Technology 4. https://doi.org/10.1016/j.atech.2022.100157

Hastie, T., Tibshirani, R., Friedman, J., 2009. Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction. https://doi.org/https://doi.org/10.1007/978-0-387-84858-7

Hawkins, D.M., 2004. The Problem of Overfitting. J Chem Inf Comput Sci. https://doi.org/10.1021/ci0342472

Imran, H.A., Gianelle, D., Scotton, M., Rocchini, D., Dalponte, M., Macolino, S., Sakowska, K., Pornaro, C., Vescovo, L., 2021. Potential and limitations of grasslands α-diversity prediction using fine-scale hyperspectral imagery. Remote Sens (Basel) 13. https://doi.org/10.3390/rs13142649

Kazar, S.A., Warner, T.A., 2013. Assessment of carbon storage and biomass on minelands reclaimed to grassland environments using Landsat spectral indices. J Appl Remote Sens 7, 073583. https://doi.org/10.1117/1.jrs.7.073583

Klingler, A., Schaumberger, A., Vuolo, F., Kalmár, L.B., Pötsch, E.M., 2020. Comparison of Direct and Indirect Determination of Leaf Area Index in Permanent Grassland. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science 88, 369–378. https://doi.org/10.1007/s41064-020-00119-8

Kolecka, N., Ginzler, C., Pazur, R., Price, B., Verburg, P.H., 2018. Regional scale mapping of grassland mowing frequency with Sentinel-2 time series. Remote Sens (Basel) 10. https://doi.org/10.3390/rs10081221

Liu, S., Cheng, F., Dong, S., Zhao, H., Hou, X., Wu, X., 2017. Spatiotemporal dynamics of grassland aboveground biomass on the Qinghai-Tibet Plateau based on validated MODIS NDVI. Sci Rep 7. https://doi.org/10.1038/s41598-017-04038-4

Maake, R., Mutanga, O., Chirima, G., Sibanda, M., 2023. Quantifying Aboveground Grass Biomass Using Space-Borne Sensors: A Meta-Analysis and Systematic Review. Geomatics 3, 478–500. https://doi.org/10.3390/geomatics3040026

Malatesta, L., Tardella, F.M., Tavoloni, M., Postiglione, N., Piermarteri, K., Catorci, A., 2019. Land use change in the high mountain belts of the central Apennines led to marked changes of the grassland mosaic. Appl Veg Sci 22, 243–255. https://doi.org/https://doi.org/10.1111/avsc.12416

Marston, C.G., Aplin, P., Wilkinson, D.M., Field, R., O'Regan, H.J., 2017. Scrubbing Up: Multi-Scale Investigation of Woody Encroachment in a Southern African Savannah. Remote Sens (Basel) 9. https://doi.org/10.3390/rs9050419

Meng, B., Liang, T., Yi, S., Yin, J., Cui, X., Ge, J., Hou, M., Lv, Y., Sun, Y., 2020. Modeling Alpine Grassland above Ground Biomass Based on Remote Sensing Data and Machine Learning Algorithm: A Case Study in East of the Tibetan Plateau, China. IEEE J Sel Top Appl Earth Obs Remote Sens 13, 2986–2995. https://doi.org/10.1109/JSTARS.2020.2999348

Morais, T.G., Jongen, M., Tufik, C., Rodrigues, N.R., Gama, I., Fangueiro, D., Serrano, J., Vieira, S., Domingos, T., Teixeira, R.F.M., 2023. Characterization of portuguese sown rainfed grasslands using remote sensing and machine learning. Precis Agric 24, 161–186. https://doi.org/10.1007/s11119-022-09937-9

Morais, T.G., Teixeira, R.F.M., Figueiredo, M., Domingos, T., 2021. The use of machine learning methods to estimate aboveground biomass of grasslands: A review. Ecol Indic 130, 108081. https://doi.org/10.1016/J.ECOLIND.2021.108081

Ogungbuyi, M.G., Mohammed, C., Ara, I., Fischer, A.M., Harrison, M.T., 2023. Advancing Skyborne Technologies and High-Resolution Satellites for Pasture Monitoring and Improved Management: A Review. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15194866

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 10. https://doi.org/10.1186/s13643-021-01626-4

Perrone, M., Conti, L., Galland, T., Komárek, J., Lagner, O., Torresani, M., Rossi, C., Carmona, C.P., de Bello, F., Rocchini, D., Moudrý, V., Šímová, P., Bagella, S., Malavasi, M., 2024. "Flower power": How flowering affects spectral diversity metrics and their relationship with plant diversity. Ecol Inform 81. https://doi.org/10.1016/j.ecoinf.2024.102589

Pickering, C., Byrne, J., 2014. The benefits of publishing systematic quantitative literature reviews for PhD candidates and other early-career researchers. Higher Education Research & Development 33, 534–548. https://doi.org/10.1080/07294360.2013.841651

Poffenbarger, H., Artz, G., Dahlke, G., Edwards, W., Hanna, M., Russell, J., Sellers, H., Liebman, M., 2017. An economic analysis of integrated crop-livestock systems in Iowa, U.S.A. Agric Syst 157, 51–69. https://doi.org/10.1016/J.AGSY.2017.07.001

Pokluda, P., Hauck, D., Cizek, L., 2012. Importance of marginal habitats for grassland diversity: Fallows and overgrown tall-grass steppe as key habitats of endangered ground-beetle Carabus hungaricus. Insect Conserv Divers 5, 27–36. https://doi.org/10.1111/j.1752-4598.2011.00146.x

Preite, L., Solari, F., Vignali, G., 2023. Technologies to Optimize the Water Consumption in Agriculture: A Systematic Review. Sustainability 15. https://doi.org/10.3390/su15075975

Price, B., Kaim, D., Szwagrzyk, M., Ostapowicz, K., Kolecka, N., Schmatz, D.R., Wypych, A., Kozak, J., 2017. Legacies, socio-economic and biophysical processes and drivers: the case of future forest cover expansion in the Polish Carpathians and Swiss Alps. Reg Environ Change 17, 2279–2291. https://doi.org/10.1007/s10113-016-1079-z

Psomas, A., Kneubühler, M., Huber, S., Itten, K., Zimmermann, N.E., 2011. Hyperspectral remote sensing for estimating aboveground biomass and for exploring species richness patterns of Grassland habitats. Int J Remote Sens 32, 9007–9031. https://doi.org/10.1080/01431161.2010.532172

Raab, C., Riesch, F., Tonn, B., Barrett, B., Meißner, M., Balkenhol, N., Isselstein, J., 2020. Target-oriented habitat and wildlife management: estimating forage quantity and quality of semi-natural grasslands with Sentinel-1 and Sentinel-2 data. Remote Sens Ecol Conserv 6, 381–398. https://doi.org/https://doi.org/10.1002/rse2.149

Reeves, M.C., Hanberry, B.B., Wilmer, H., Kaplan, N.E., Lauenroth, W.K., 2021. An Assessment of Production Trends on the Great Plains from 1984 to 2017. Rangel Ecol Manag 78, 165–179. https://doi.org/10.1016/J.RAMA.2020.01.011

Reinermann, S., Asam, S., Kuenzer, C., 2020. Remote Sensing of Grassland Production and Management—A Review. Remote Sens (Basel) 12. https://doi.org/https://doi.org/10.3390/rs12121949

Sampson, M., McGowan, J., Cogo, E., Horsley, T., 2006. Managing database overlap in systematic reviews using Batch Citation Matcher: case studies using Scopus. J Med Libr Assoc 94, 461–463.

Sándor, R., Ehrhardt, F., Grace, P., Recous, S., Smith, P., Snow, V., Soussana, J.F., Basso, B., Bhatia, A., Brilli, L., Doltra, J., Dorich, C.D., Doro, L., Fitton, N., Grant, B., Harrison, M.T., Skiba, U., Kirschbaum, M.U.F., Klumpp, K., Laville, P., Léonard, J., Martin, R., Massad, R.S., Moore, A.D., Myrgiotis, V., Pattey, E., Rolinski, S., Sharp, J., Smith, W., Wu, L., Zhang, Q., Bellocchi, G., 2023. Residual correlation and ensemble modelling to improve crop and grassland models. Environmental Modelling & Software 161, 105625. https://doi.org/10.1016/J.ENVSOFT.2023.105625

Sartorello, Y., Pastorino, A., Bogliani, G., Ghidotti, S., Viterbi, R., Cerrato, C., 2020. The impact of pastoral activities on animal biodiversity in Europe: A systematic review and meta-analysis. J Nat Conserv 56, 125863. https://doi.org/https://doi.org/10.1016/j.jnc.2020.125863

Schellberg, J., Hill, M.J., Gerhards, R., Rothmund, M., Braun, M., 2008. Precision agriculture on grassland: Applications, perspectives and constraints. European Journal of Agronomy 29, 59–71. https://doi.org/10.1016/J.EJA.2008.05.005

Sgroi, F., Modica, F., 2022. Long-term changes in business models in inland and mountainous areas for the promotion of sustainable food systems. J Agric Food Res 10, 100451. https://doi.org/10.1016/J.JAFR.2022.100451

Shoko, C., Mutanga, O., Dube, T., 2019. Remotely sensed C3 and C4 grass species aboveground biomass variability in response to seasonal climate and topography. Afr J Ecol 57, 477–489. https://doi.org/https://doi.org/10.1111/aje.12622

Shrestha, S., Rahimzadeh-Bajgiran, P., De Urioste-Stone, S., 2020. Probing recent environmental changes and resident perceptions in Upper Himalaya, Nepal. Remote Sens Appl 18, 100315. https://doi.org/10.1016/J.RSASE.2020.100315

Smith, H.D., Dubeux, J.C.B., Zare, A., Wilson, C.H., 2023. Assessing Transferability of Remote Sensing Pasture Estimates Using Multiple Machine Learning Algorithms and Evaluation Structures. Remote Sens (Basel) 15. https://doi.org/10.3390/rs15112940

Soussana, J.F., Tallec, T., Blanfort, V., 2010. Mitigating the greenhouse gas balance of ruminant production systems through carbon sequestration in grasslands. Animal 4, 334–350. https://doi.org/10.1017/S1751731109990784

Straffelini, E., Luo, J., Tarolli, P., 2024. Climate change is threatening mountain grasslands and their cultural ecosystem services. Catena (Amst) 237. https://doi.org/10.1016/j.catena.2023.107802

Sullivan, C.S., Gemtou, M., Anastasiou, E., Fountas, S., 2024. Building trust: A systematic review of the drivers and barriers of agricultural data sharing. Smart Agricultural Technology 8. https://doi.org/10.1016/j.atech.2024.100477

Trentin, A.B., Cardoso, J.M.K., Ghisi, N.D.C., Rachid, C.T.C.D.C., Leite, D.C.D.A., 2024. Rooting for growth: Meta-analyzing the role of Endophytic fungi in plant growth. Sci Hortic 333. https://doi.org/10.1016/j.scienta.2024.113276

Vawda, M.I., Lottering, R., Mutanga, O., Peerbhay, K., Sibanda, M., 2024. Comparing the Utility of Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) on Sentinel-2 MSI to Estimate Dry Season Aboveground Grass Biomass. Sustainability (Switzerland) 16. https://doi.org/10.3390/su16031051

Wang, Jie, Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R.B., Chang, Q., 2019. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. ISPRS Journal of Photogrammetry and Remote Sensing 154, 189–201. https://doi.org/10.1016/j.isprsjprs.2019.06.007

Wang, J., Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R.B., Chang, Q., 2019. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. ISPRS Journal of Photogrammetry and Remote Sensing 154, 189–201. https://doi.org/10.1016/j.isprsjprs.2019.06.007

Wang, Q., Zhao, L., Wang, M., Wu, J., Zhou, W., Zhang, Q., Deng, M., 2022. A Random Forest Model for Drought: Monitoring and Validation for Grassland Drought Based on Multi-Source Remote Sensing Data. Remote Sens (Basel) 14. https://doi.org/10.3390/rs14194981

Wang, Z., Ma, Y., Zhang, Y., Shang, J., 2022. Review of Remote Sensing Applications in Grassland Monitoring. Remote Sens (Basel) 14. https://doi.org/10.3390/rs14122903

Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I., Justice, C.O., 2015. Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. Remote Sens Environ 156, 438–447. https://doi.org/10.1016/J.RSE.2014.10.009

Woodward, S.J.R., Neal, M.B., Cross, P.S., 2019. Preliminary investigation into the feasibility of combining satellite and gps data to identify pasture growth and grazing. Journal of New Zealand Grasslands 81, 47–54. https://doi.org/10.33584/jnzg.2019.81.404

World Food and Agriculture – Statistical Yearbook 2023, 2023. , World Food and Agriculture – Statistical Yearbook 2023. FAO. https://doi.org/10.4060/cc8166en

Xu, D., Koper, N., Guo, X., 2018. Quantifying the influences of grazing, climate and their interactions on grasslands using Landsat TM images. Grassl Sci 64, 118–127. https://doi.org/https://doi.org/10.1111/grs.12192

Xu, K., Su, Y., Liu, J., Hu, T., Jin, S., Ma, Q., Zhai, Q., Wang, R., Zhang, J., Li, Y., Liu, H., Guo, Q., 2020. Estimation of degraded grassland aboveground biomass using machine learning methods from terrestrial laser scanning data. Ecol Indic 108. https://doi.org/10.1016/j.ecolind.2019.105747

Zeng, N., Ren, X., He, H., Zhang, L., Zhao, D., Ge, R., Li, P., Niu, Z., 2019. Estimating grassland aboveground biomass on the Tibetan Plateau using a random forest algorithm. Ecol Indic 102, 479–487. https://doi.org/10.1016/j.ecolind.2019.02.023

Zhao, F., Xu, B., Yang, X., Jin, Y., Li, J., Xia, L., Chen, S., Ma, H., 2014. Remote sensing estimates of grassland aboveground biomass based on MODIS Net Primary Productivity (NPP): A case study in the Xilingol grassland of northern China. Remote Sens (Basel) 6, 5368–5386. https://doi.org/10.3390/rs6065368

Zhou, Y., Liu, T., Batelaan, O., Duan, L., Wang, Y., Li, X., Li, M., 2023. Spatiotemporal fusion of multi-source remote sensing data for estimating aboveground biomass of grassland. Ecol Indic 146. https://doi.org/10.1016/j.ecolind.2023.109892

Zhu, X., Bi, Y., Du, J., Gao, X., Zhang, T., Pi, W., Zhang, Y., Wang, Y., Zhang, H., 2023. Research on deep learning method recognition and a classification model of grassland grass species based on unmanned aerial vehicle hyperspectral remote sensing. Grassl Sci 69, 3–11. https://doi.org/10.1111/grs.12379

3. Modelling vegetation phenology in alpine grasslands of North-East Italy integrating microclimatic and meteorological Drivers: A 23-Years Time-Series analysis

3.1 Abstract

Grasslands in mountainous regions often exhibit unique variations in the amplitude and timing of their growing seasons in response to microclimatic and meteorological variability. This study investigates the environmental drivers of vegetation phenology in the Alpine grasslands of North-East Italy using a 23-year MODIS NDVI time series (2001–2023), harmonic regression, and Generalized Additive Models (GAMs). Results highlight a strong negative correlation between altitude and NDVI metrics, with higher elevations exhibiting lower vegetation cover, greater seasonal variability, and shorter growing seasons. Slope similarly constrains vegetation growth, while aspect influences microclimatic conditions but shows weaker overall effects. Temperature and solar radiation positively influence NDVI, enhancing vegetation vigor and extending the growing season, while precipitation plays a more limited role, with limited variability across the region. Wind exposure shortens the growing season, reducing NDVI values, while snowfall significantly delays and shortens vegetation growth, increasing seasonal variability. Soil properties, particularly water-holding capacity, pH, and texture, strongly influence vegetation dynamics, with silt-rich soils favoring longer and more stable growing periods. The developed GAMs demonstrated high predictive accuracy (R^2 = 0.68–0.90), effectively estimating phenological transitions, including the start and end of the vegetative season, with an accuracy of approximately 11 days. These findings enhance our understanding of the complex interactions shaping alpine pasture ecosystems and provide valuable tools for sustainable management, conservation, and climate adaptation strategies.

3.2. Introduction

Grasslands and pastures are vital ecosystems globally and locally, holding environmental, economic, and social significance and storing about 30% of Earth's terrestrial biomass. Global warming and shifts in precipitation patterns notably impact forage production, quality, and the botanical composition of rainfed pastures and meadows, rendering these systems highly susceptible to climate change (Brilli et al., 2023). Rising temperatures are causing grass vegetation communities to move to higher elevations, a trend more pronounced at higher altitudes. This shift threatens the survival of endemic species at the highest peaks of the Alps, as high-altitude species have fewer opportunities to migrate compared to those at lower elevations. Increased temperatures, reduced precipitation, and shorter snow cover periods have notably increased vegetation cover and plant growth in the Alpine belt. These changes result in lower nitrogen levels and reduced biomass crude protein content while promoting the spread of less palatable shrub species (Dibari et al., 2021a). Topographic features, including elevation, aspect, and position, play a crucial role in determining vegetation distribution and characteristics by altering the local environment, with a specific effect on water regime, solar radiation and thus macro and microclimatic conditions (Sundqvist et al., 2013). The effects of elevational gradients on plant communities have been extensively investigated, as ecologists have long recognised that plant communities exhibit distinctive responses to elevation (Whittaker, 1960). Increasing elevation is associated with a decrease in temperature, a reduction in land area, since higher altitudes typically have less total area and shifts in other abiotic factors (Sundqvist et al., 2013). These gradients influence various plant traits, such as decreasing specific leaf area, decreasing leaf nutrient concentrations, and increasing nutrient amounts per leaf area. In the Italian Alps, elevation showed varied effects on species composition and, thus, on vegetation phenology. Mesophilic grassland species are more common at lower elevations, while higher elevations host dry grassland species adapted to calcareous substrates with limited soil depth (Orlandi et al., 2016). The effects of slope and aspect on vegetation have also been widely investigated in various environments. Differences in regional slopes and aspects determine the timing and duration of snow cover in North American grasslands (Tennant et al., 2017), influencing alpine plants' growth timing and ecological processes. Research on the Tibetan Plateau has shown a trend toward shorter growth periods for vegetation at higher elevations, with green-up and dormancy dates varying according to the aspect and slope in meadow and steppe regions (An et al., 2018a). In the New Zealand Alps, Hua et al. (Hua et al., 2022) modelled the seasonal dynamics of grassland vegetation using NDVI, finding that the growing season started earlier and lasted longer at lower altitudes. In the climatic zone of the Italian Alps, most of permanent pastures and grasslands would naturally be covered by forests, but centuries of human activities have reshaped the landscape. Pastoralism has played a crucial role in maintaining these ecosystems, particularly in marginal, high-altitude areas where crop cultivation is challenging. However, declining pastoral activities have led to the abandonment of many grasslands, making them increasingly vulnerable to shrub encroachment and natural reforestation (Hinojosa et al., 2019). Although the natural treeline in this region is typically between 1900 and 2200 m a.s.l., traditional land use has allowed grassland ecosystems to persist at lower elevations. Protecting these landscapes means actively preventing their transition back into forested areas, ensuring the conservation of their ecological and cultural significance. Recognising their value, the European Union and regional authorities have designated many of these pastures as protected grazing areas. The Common Agricultural Policy (CAP) supports their maintenance through various initiatives, including investing in sustainable agricultural practices, biodiversity conservation, and measures to counteract land abandonment. However, effective management requires reliable monitoring tools to track vegetation dynamics and detect ecological shifts. The alpine grasslands of North-East Italy present a unique opportunity for research due to their ecological importance and the increasing challenges they face. Monitoring vegetation phenology is essential for understanding seasonal growth dynamics and guiding land-use decisions. Remote sensing-based phenological models can provide a scalable and efficient approach to tracking changes in plant development, offering valuable insights for optimising grazing schedules, timing haymaking operations, and implementing adaptive grazing management to sustain pasture productivity. Additionally, integrating phenological data into conservation and agricultural policies can enhance long-term resilience, ensuring the sustainability of these ecosystems in the face of environmental and socio-economic changes.

Multispectral satellite data, such as NDVI and other vegetation indices from AVHRR, MODIS, Landsat, and SPOT, has been widely used to monitor pasture growth and biomass at regional and global scales, supporting precision livestock farming and management (Ogungbuyi et al., 2024a), offering comparable performances to those obtained with proximal sensors (Thomson et al., 2023). Advanced statistical modelling has been deployed, integrating environmental data from remote and proximal sensing to estimate various vegetation features in agricultural and natural environments (Y. Liu et al., 2024). Ma et al. (2022) successfully developed Generalized Linear Models estimating phenological indicators of grassland vegetation in semiarid grassland of central Asia, integrating environmental data from ERA-5 climate reanalysis, reaching an accuracy of 3-5 days predicting start and end of the vegetative season. This study aims to investigate the main environmental drivers and model the vegetation phenology of alpine grasslands in North-East Italy using a 23-year MODIS NDVI time series and morphological and meteorological data. Specifically, the objectives of the study are:

1. Quantify the influence of environmental drivers, including soil features, meteorological variables and microclimatic differences shaped by topographical features (altitude, slope, and aspect), on the seasonal growth dynamics of alpine grassland vegetation.
2. Develop and validate statistical models that integrate environmental drivers to accurately predict the phenology of alpine grassland vegetation for monitoring and management purposes.

Despite the growing research on alpine grassland phenology, few studies have analysed long-term trends over multiple decades while integrating various environmental drivers. This study is novel in its use of a 23-year MODIS NDVI time series (2001–2023), which provides a uniquely extended dataset for assessing vegetation dynamics in alpine regions. Additionally, by combining harmonic regression to capture seasonal cycles and Generalized Additive Models (GAMs) to quantify environmental influences, this research offers a more refined and predictive understanding of phenological patterns than previous approaches. This study is based on the hypothesis that topographical, meteorological and edaphic factors are key drivers of alpine grassland phenology. Specifically, increasing altitude and slope are expected to shorten the growing season and reduce vegetation vigor due to their impact on microclimatic conditions, soil depth, fertility, and water availability. Steeper slopes often lead to shallower soils with lower water-holding capacity, further constraining plant growth. Higher temperatures and increased solar radiation are anticipated to enhance vegetation growth and extend the growing season by promoting photosynthesis. In contrast, greater snowfall and stronger winds may have a limiting effect, reducing vegetation productivity and shortening the growing season by delaying snowmelt, increasing evapotranspiration stress, and causing mechanical damage to plants.

3.3. Materials and methods

3.3.1 Study area

The present study is set in the Province of Trento (NUTS 2, ITH20), a region of north-eastern Italy, which stands out for its high climatic and morphological variability due to its mountainous terrain. The province's climate, classified as temperate, exhibits a wide range of microclimates, from Mediterranean influences in lower valleys to more continental and Alpine climates at higher altitudes. To ensure the highest level of accuracy, we only considered fields with an area greater than 6 hectares, thereby minimising the influence of neighbouring areas on MODIS surface reflectance values (250x250m pixels). Complying with this size condition, 890 fields were selected from the original 10852 fields, all classified as permanent pastures or grasslands (Figure 3.1), covering a total area of 54154 ha, from 828 to 2491 meters above sea level. The field land cover labelling was manually validated and made available by the province land authorities in a public database. At this climatic zone, most of permanent pastures and grassland areas would be occupied by forests, but centuries of human activities have reshaped the landscape. Pastoralism has long been a vital resource for the populations of Alpine regions, particularly in marginal areas such as high-altitude fields, where cultivating crops as cereals, grapevines, or apples is challenging. Figure 3.2 provides examples of permanent pastures and grasslands that characterise the region's landscape.

These pastures are now designated as protected grazing areas, reflecting their ecological and cultural importance. Both the European Union and the Province of Trento recognise their value. Through the Common Agricultural Policy (CAP), significant efforts have been made to support the maintenance of these areas, with direct and indirect investments aimed at promoting sustainable agricultural practices, preserving biodiversity, and preventing land abandonment.

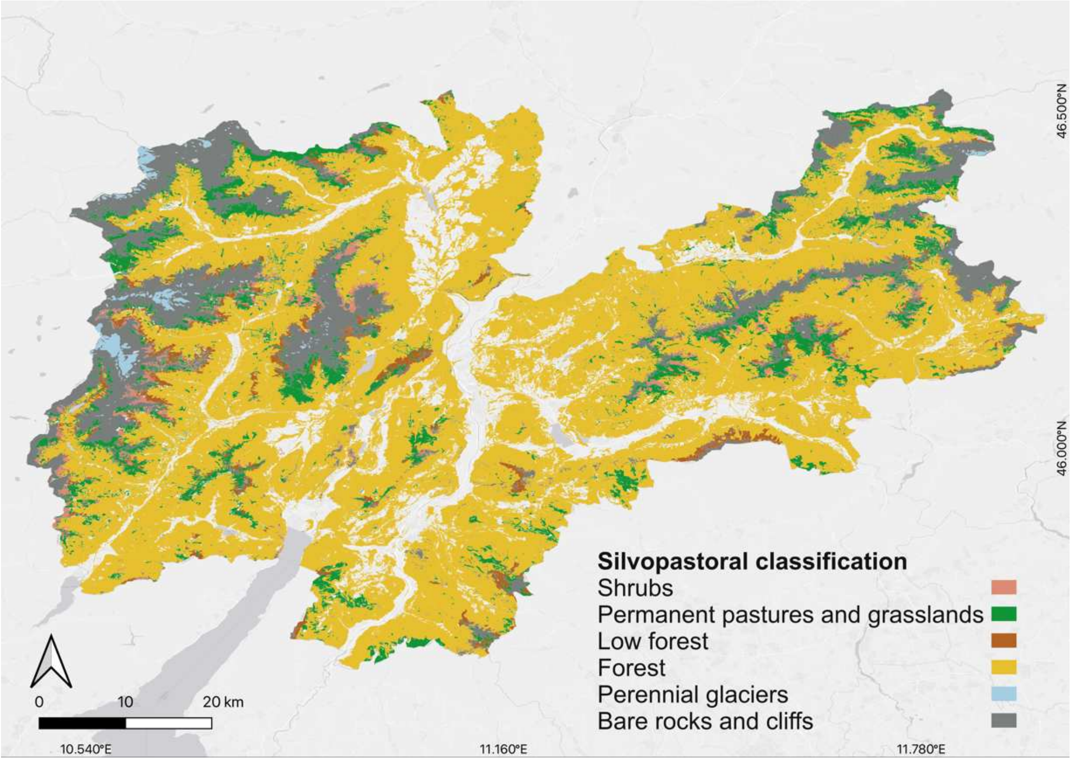


Figure 3.1 Silvopastoral classification of the Province of Trento, Italy, colored in green permanent pastures and grasslands, constituting the study area. Source: Portale Geocartografico Trentino.

Figure 3.2 Examples of alpine permanent grasslands and pastures of North-east Italy. Pasture on the left located at Malga Fosse, Siror (TN). Pasture on the right located at Malga Juribello, Passo Rolle (TN).

3.3.2 Data collection

The study aimed to analyse the temporal and spatial evolution of a vegetation-related spectral index as the Normalized Difference Vegetation Index (NDVI) over 23 years (2001-2023), and its relations with morphological and climatic variables.

Thus, data was collected from multiple sources, utilising the Google Earth Engine API as the primary tool to retrieve data. A monthly NDVI time series was produced by accessing the "MODIS/061/MOD13Q1" dataset, averaging monthly data on each of the selected 890 fields. Following the same approach, meteorological data was downloaded from the "ECMWF/ERA5_LAND/MONTHLY_AGGR" dataset. Topographical data such as mean altitude, slope, and aspect were retrieved from the "NASA/NASADEM_HGT/001" dataset.

Finally, data regarding various soil features, spatialised at 250m resolution, was collected from the OpenLandMap datasets produced by EnvirometriX Ltd. Since the root systems of species inhabiting the mountain grasslands of northern Italy primarily develop within the uppermost soil layers (Gonzalez-Ollauri et al., 2021), all soil data were extracted from band 0, representing the top 10 cm of soil. Table 1 lists the data collected, the spatial and temporal resolution and their sources.

The methodologies selected for this study were chosen to directly address the research objectives of quantifying the influence of environmental drivers on alpine grassland phenology and developing predictive models. Specifically, harmonic regression was used to extract key phenological metrics from NDVI time series, allowing for a detailed characterisation of seasonal vegetation dynamics. To establish predictive relationships between NDVI metrics and environmental variables, Generalized Additive Models (GAMs) were employed. GAMs were selected due to their ability to model both linear and nonlinear dependencies while maintaining interpretability.

Table 3.1 Collected data with measurement unit, spatial and temporal resolution and data source.

Variable	Measurement Unit	Spatial Resolution	Temporal Resolution	Source
NDVI	-	250m	Monthly	MODIS
Altitude	m.a.s.l.	30m	-	NASADEM
Slope	%	30m	-	NASADEM
Aspect	degrees	30m	-	NASADEM
Mean temperature	K	11km	Monthly	ERA5-ECMWF
Solar net radiation	J/m^2	11km	Monthly	ERA5-ECMWF
Total precipitation	m	11km	Monthly	ERA5-ECMWF
U component of wind	m/s	11km	Monthly	ERA5-ECMWF
V component of wind	m/s	11km	Monthly	ERA5-ECMWF

Snow cover	%	11km	Yearly	ERA5-ECMWF
Snowfall sum	m	11km	Monthly	ERA5-ECMWF
Soil organic carbon	g/kg	250m	-	OpenLandMap
Soil pH	-	250m	-	OpenLandMap
Soil water content at 33kPa (field capacity)	%	250m	-	OpenLandMap
Soil clay fraction	%	250m	-	OpenLandMap
Soil silt fraction	%	250m	-	OpenLandMap
Soil sand fraction	%	250m	-	OpenLandMap

U and V wind components were utilised to extract mean wind speed and prevalent wind direction. Wind speed and direction were calculated using the following formulas provided by ECMWF:

$$V\left(\frac{m}{s}\right) = \sqrt{u\ comp.^2 + v\ comp.^2}$$

$$\phi = mod\left(180 + \frac{180}{\pi} atan2(u\ comp., v\ comp.), 360\right)$$

Where ϕ is the wind direction in degrees in the range $0 < \phi < 360$.

As both aspect and prevalent wind direction are expressed in degrees but can only be interpreted as cardinal directions, all fields were classified using the following ranges:

- North: mean aspect or wind direction lower than 45 or greater than 315 degrees,
- East: mean aspect or wind direction between 45 and 135 degrees,
- South: mean aspect or wind direction between 135 and 225 degrees,
- West: mean aspect or wind direction between 225 and 315 degrees.

3.3.3 Data processing, analysis, visualisation

Python (version 3.11.8) libraries such as Pandas (v. 1.5.3), NumPy (v.1.23.5), SciPy (1.15.1), statsmodels (v.0.13.5) mpmath (v 1.2.1), pygam (v.0.9.0) and sklearn (v. 0.2) were used for data manipulation and statistical analysis, while Matplotlib (v. 3.7.0) and Seaborn (v. 0.12.2) were employed for plotting and visualisation. The geographical visualisation of data was performed using the open-source QGIS software (v. 3.28). In addition, the ggplot2 (v. 3.5.1), caret (v. 6.0-94), dplyr

(v. 1.1.4), tidyr (v. 1.3.1), Metrics (v. 0.1.4), ggpubr (v. 0.6.0), gridExtra (v. 2.3) and mgcv **(**v. 1.8**),** package in R (v. 4.2.2) was used to implement generalised additive models (GAMs) to build predictive models of NDVI metrics.

3.3.4 Harmonic Regression

This study employed harmonic regression to analyse the temporal evolution of Normalized Difference Vegetation Index (NDVI) data across 890 fields over 23 years (2001-2023).

Harmonic regression, distinguished by its ability to fit sinusoidal models to time-series data, is particularly adept at capturing the intrinsic seasonal patterns of vegetative growth and decline. The application of this approach, explicated by equation 1, facilitated a comprehensive examination of the periodic behaviour of NDVI signals, with specific emphasis on retrieving the amplitude and phase shift parameters for each field. These parameters are critical for understanding the phenological cycles of the observed vegetation, offering insights into the timing of vegetative growth and senescence cycles.

$$NDVI_t = \beta_0 + \beta_1 \cos(2\pi t) + \beta_2 \sin(2\pi t) + \in \qquad (1)$$

- t is the time in years,
- β0 is the baseline (mean NDVI level),
- β1 and β2 are coefficients that determine the amplitude and phase shift, estimated through OLS,
- ϵ is the error term.

From the models, critical features are extracted:

- Amplitude, representing the difference between peak and mean NDVI, calculated as $\sqrt{(\beta_2^1 + \beta_2^2)}$,
- Phase Shift (ϕ) represents the shift in time from the reference, expressed in radians and calculated as $arctan\,2\,(\frac{\beta_2}{\beta_1})$.
- Baseline: The coefficient β0, representing the mean NDVI over the whole 23-year period.

Based on the NDVI features investigated in this phase of the study were:

- Mean NDVI, the average NDVI value of each field in the 23-year period,
- NDVI peak, the highest detected NDVI value of each field in the 23 year period,

- Seasonal variability, computed as the normalised amplitude of the individual field harmonic regression divided by its baseline (mean), it represents the mean seasonal variability of the field,
- Day of Year of NDVI peak, calculated as: $DOY_{peak} = \frac{\phi}{2\pi} \times 365$, the phase shift of the model Φ, representing the difference between peak and mean NDVI, is calculated as $\phi = arctan\,2\,(\frac{\beta_2}{\beta_1})$

The start and end of vegetative season metrics are obtained as the days when NDVI becomes greater and smaller than the 0.4 threshold. The harmonic model described above predicts NDVI over a fine time grid (N=1000). Then, the SOS and EOS indices are identified using Python Numpy package and converted into fractional years $SOS/EOS_t = \frac{SOS/EOS_{index}}{N}$ and Day of the Year.

- Day of Year of start of season, calculated as $DOY_{SOS} = SOS_t \times 365$,
- Day of Year of end of season, calculated as $DOY_{EOS} = EOS_t \times 365$.

After calculations, the dataset restores to align with a format where each row corresponds to a unique field identifier (ID). It incorporates morphological characteristics, including elevation, slope, aspect category, mean temperature, mean solar net radiation, mean wind speed, mean yearly precipitation and snowfall. Additionally, this restructured dataset integrates Mean NDVI, NDVI peak in the whole period, Amplitude (reflecting NDVI seasonal variability), DOY of NDVI peak, DOY of start of season and DOY of end of season.

3.3.5 Kendall's Tau correlation analysis

To assess monotonic relationships between continuous environmental variables and NDVI metrics, Kendall's Tau correlation coefficient (τ) was employed. Kendall's Tau was selected because it is robust to outliers, appropriate for detecting monotonic associations, and accounts for tied ranks, common in ecological datasets. However, as Kendall's Tau does not inherently capture non-monotonic or complex non-linear relationships, additional analyses were conducted deploying the Generalized Additive Models as detailed in the following sections. The method compares all pairs of observations (X_i,Y_i) and (X_j,Y_j), where $i < j$, to determine whether the pairs are concordant or discordant:

- **Concordant Pairs**: If the ranks of both variables are in the same order ($X_i<X_j$ and $Y_i<Y_j$, or $X_i>X_j$ and $Y_i>Y_j$);
- **Discordant Pairs**: If the ranks are in opposite order ($X_i<X_j$ and $Y_i>Y_j$, or $X_i>X_j$ and $Y_i<Y_j$).
- Pairs tied in either X or Y are excluded from the concordant and discordant counts.

The coefficient (**τ**) is computed as:

$$\tau = \frac{C - D}{\frac{n(n-1)}{2}}$$

Where C is the number of concordant pairs, D is the number of discordant pairs, and n is the total number of observations. The resulting coefficient ranges from −1 (perfect negative monotonic correlation) to +1 (perfect positive monotonic correlation), with 0 indicating no monotonic association. Statistical significance was determined using p-values derived from permutation tests, with $\alpha=0.05$ as the threshold for significance. The analysis was conducted using the *scipy.stats.kendalltau* function in Python, and the results were summarised as correlation coefficients and corresponding p-values.

3.3.6 Kruskall-Wallis test

To evaluate the statistical significance of differences among categorical variables, such as aspect and prevalent wind direction, the **Kruskal-Wallis test** was performed. This non-parametric test compares the medians of multiple independent groups and is suitable for data that do not meet the assumptions of normality. The test statistic H is calculated as:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{g} n_i R_i^2 - 3(n+1)$$

where:

- g is the number of groups;
- n_i is the number of observations in the i^{th} group;
- R_i is the sum of ranks for the i^{th} group;
- n is the total number of observations across all groups.

The test assesses the null hypothesis (H_0) that the distributions of all groups are identical. If the resulting p-value is less than the significance threshold ($\alpha=0.05$), H_0 is rejected, indicating significant differences among the groups. The Kruskal-Wallis analysis was implemented using the *scipy.stats.Kruskal* function and results were reported as H statistics and corresponding p-values.

To quantify the effect size of the Kruskal-Wallis test, epsilon-squared (ε^2) was calculated for each variable. This measure indicates the proportion of variance in the dependent variable that can be attributed to the grouping factor (e.g., aspect or wind direction). The formula used for ε^2 is:

$$\varepsilon^2 = \frac{H - k + 1}{n - k}$$

where:

- H is the Kruskal-Wallis H-statistic,
- k is the number of groups,
- n is the total number of observations.

Epsilon-squared values range from 0 to 1, where higher values indicate a more substantial effect of the grouping variable on the response variable. The computation was performed for each NDVI metric to assess the magnitude of differences among aspect or wind direction categories.

3.3.7 Generalized Additive Models

To model the relationships between NDVI metrics (e.g., mean, peak, normalised amplitude, SOS, day of year, and EOS) and a combination of topographical, meteorological and soil-related variables, Generalized Additive Models (GAMs) were employed. GAMs were chosen for their ability to capture both linear and non-linear relationships while maintaining interpretability. A GAM can be expressed as:

$$g\big(E(y_i)\big) = \beta_0 + \sum_{j=1}^{p} f_j(x_{ij})$$

where:

- $\boldsymbol{g}$ is the link function relating the expected response E(yi) to the predictors;
- β_0 is the intercept;
- $\boldsymbol{f}_j(\boldsymbol{x}_{ij})$ are smooth functions of the continuous predictors $\boldsymbol{x}_{ij}$, estimated using penalised regression splines;
- $\boldsymbol{p}$ is the number of predictors.

Separate Generalized Additive Models (GAMs) were trained for each NDVI metric using the *mgcv* package in R. To ensure robust model evaluation, the dataset was split into training (80%) and test (20%) sets using stratified sampling based on altitude. The altitude variable was divided into three quantile-based bins (Low, Medium, High). The *CreateDataPartition()* function from the *caret* package maintained a representative distribution of elevation classes in both sets.
Model performance was assessed using the Root Mean Squared Error (RMSE) and the coefficient of determination (R^2), calculated on the test set to ensure a fair evaluation. The GAMs were built using cubic regression splines (bs='cs') as smoothing functions for all continuous predictors, allowing for flexible yet smooth fits. The default basis dimension (k=10) was used to provide sufficient flexibility

without overfitting, while model estimation was performed using Restricted Maximum Likelihood (REML) to optimise smoothness selection.

3.4 Results

3.4.1 Influence of altitude and slope on NDVI dynamics

The field distribution across both elevation and slope gradients was analysed to assess the robustness of the study and identify potential biases. Figure 3.3 illustrates that while there is some variation in sample sizes across gradients, such as a reduction in the number of fields at higher elevations and fewer fields on steep slopes, the overall distribution remains balanced enough to support meaningful conclusions.

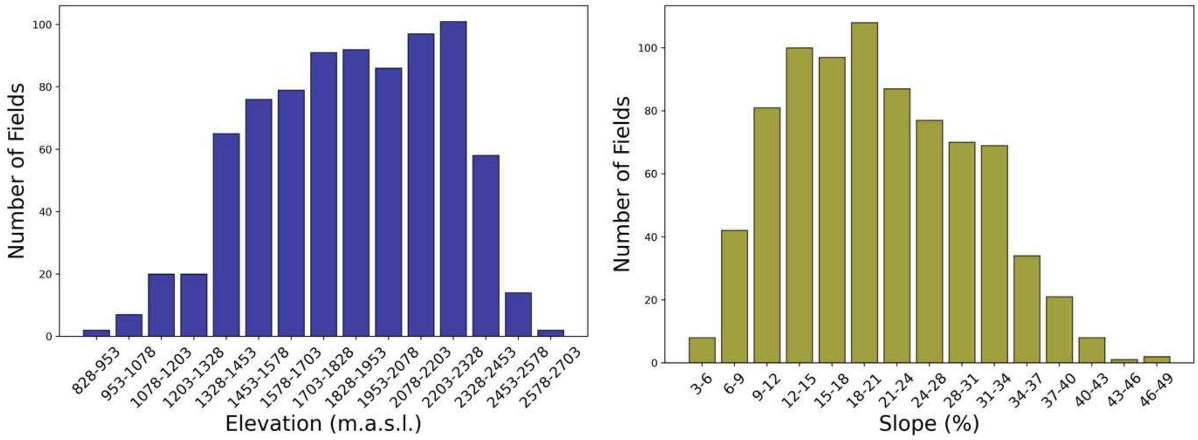


Figure 3.4 Barplots showing the distribution of studied fields along elevation and slope classes.

The scatterplots in Figure 4 illustrate the pronounced effects of altitude and slope on the dynamics of the Normalized Difference Vegetation Index (NDVI) across alpine grasslands, highlighting significant trends in mean NDVI, NDVI peak, and seasonal variability as influenced by these topographical factors. The mean NDVI value, representing the average over 23 years, exhibits a negative correlation with altitude, as indicated by Kendall's coefficient of -0.76. This demonstrates a substantial decline in NDVI with increasing altitude. A negative trend is observed for the mean seasonal peak NDVI, with a Kendall coefficient of -0.65, also revealing a non-linear relationship. NDVI peak values remain relatively stable up to approximately 2000 m a.s.l., followed by a sharp decline at higher altitudes. Conversely, altitude correlates positively with seasonal variability (τ=0.67).

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITÀ DEGLI STUDI DI PADOVA

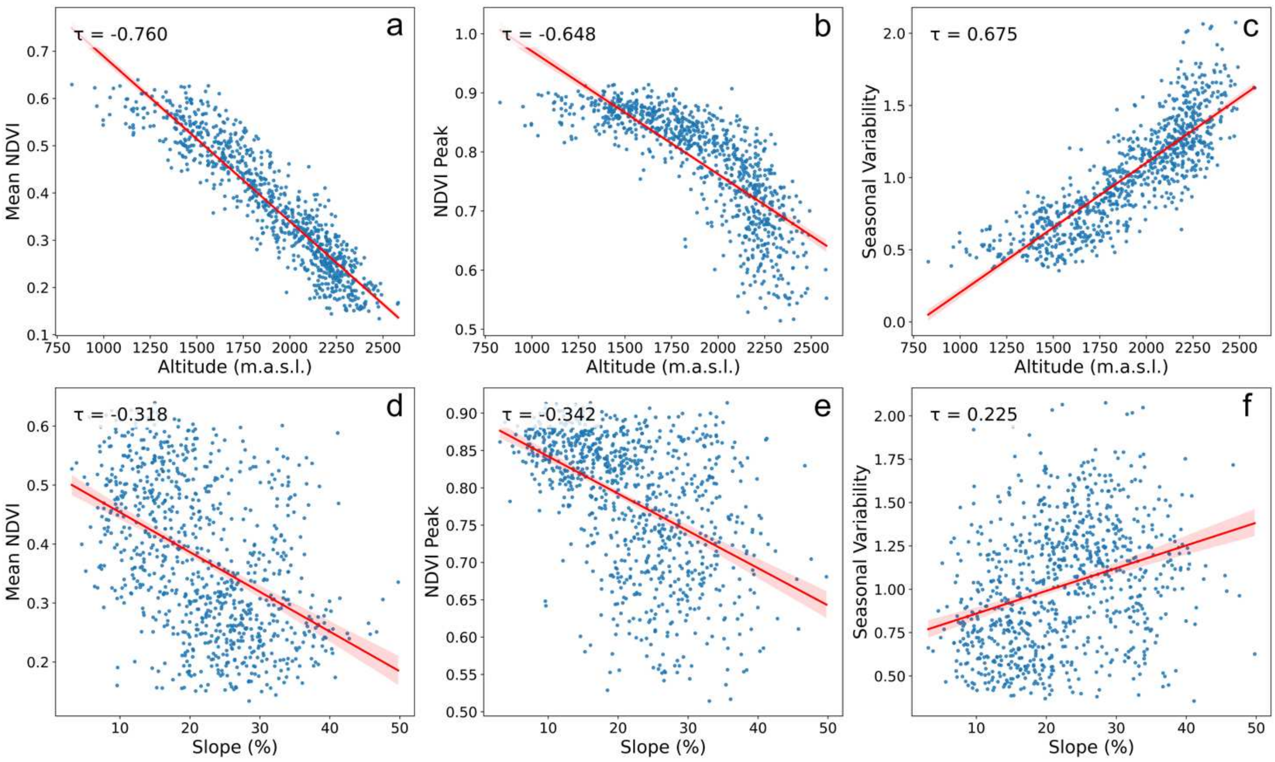


Figure 3.4 Scatterplots of correlations between Altitude (plots a, b, c) and Slope (plots d, e, f) with Mean NDVI, NDVI peak and Seasonal variability (normalised amplitude), Kendall's Tau coefficients are reported in each plot (τ).

The effects of increasing altitude on the seasonal timing of vegetation growth are visible in Figure 3.5. Altitude shows a positive correlation with the day of the start of the season (τ=0.75) and the day of the NDVI peak (τ=0.47), even if moderate. At the same time, it's negatively correlated with the day of the end of the season (τ=-0.72), indicating that high altitudes have a shortening and delaying effect on the vegetative season.

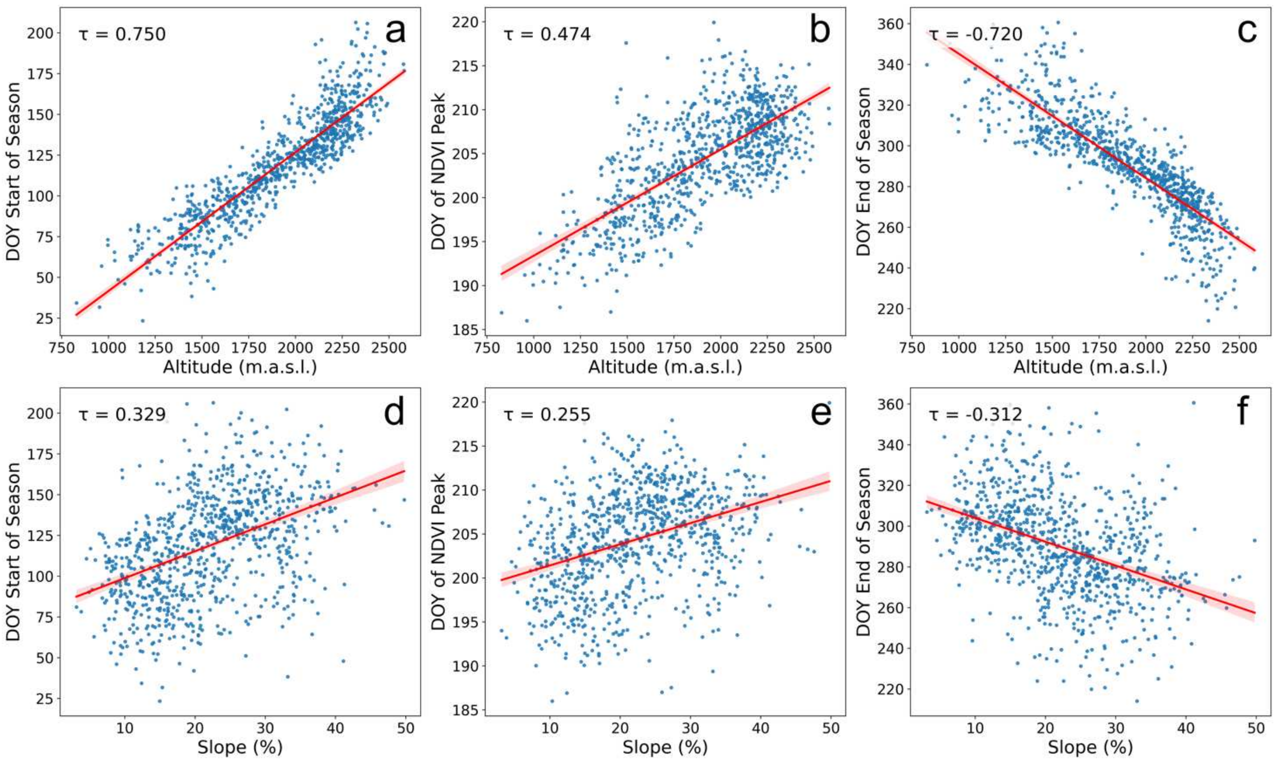


Figure 3.5 Scatterplots of correlations between Altitude (plots a, b, c) and Slope (plots d, e, f) with days of start, peak and end of the vegetative season, Kendall's Tau coefficients are reported in each plot (τ).

In parallel, the influence of slope on NDVI variables mirrors that of altitude to some extent, although with generally weaker correlations. An increment in slope is negatively correlated with mean NDVI and peak NDVI, while its correlation with seasonal variability is positive (Table 3.2). Also, the effects of increasing slope at the start, peak, and end of the season are similar to those at altitude, suggesting that steeper slopes favour later and shorter vegetative seasons.

Table 3.2 Kendall's Tau coefficients τ and p-values of correlations between Altitude and Slope with NDVI variables.

		Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)
Altitude (m.a.s.l.)	Kendall's coeff. τ	-0.761	-0.648	0.675	0.750	0.474	-0.720
	p value	$1.61e^{-252}$	$9.43e^{-184}$	$3.42e^{-199}$	$4.15e^{-245}$	$2.58e^{-99}$	$1.66e^{-225}$

Slope (%)	Kendall's coeff. τ	-0.318	-0.341	0.224	0.330	0.255	-0.312
	p value	$1.69e^{-36}$	$2.25e^{-24}$	$1.42e^{-34}$	$9.60e^{-33}$	$1.28e^{-13}$	$3.58e^{-35}$

3.4.2 Influence of aspect and geographical position on NDVI dynamics

The pastures and grasslands of our study all lie in a mountainous area, where aspect has a crucial influence on the local microclimates. The fields were classified into aspect categories: East, West, and South, as no fields faced North.

The influence of aspect on the NDVI dynamics was investigated, and a non-parametric Kruskal-Wallis test was performed to determine whether the differences between aspect categories were significant. As shown by the results of the Kruskal-Wallis test in Table 3.3, the differences among different aspect categories are significant regarding general mean NDVI, the NDVI peak, seasonal variability, and the start and end of the vegetative season. In contrast, the day of NDVI peak did not show significant differences. However, the effect sizes (ε^2) for these differences were low, indicating that while aspect shapes NDVI dynamics, its overall influence in our study area is relatively weak.

Table 3.3 H statistics, p-values and ε^2 of Kruskal-Wallis test between aspect categories and NDVI variables.

	Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)
H statistic	7.283	8.215	10.946	6.684	4.585	8.501
p-value	0.026	0.016	0.004	0.035	0.101	0.014
ε^2	0.006	0.007	0.010	0.005	0.003	0.007

The boxplots in Figure 3.6 reveal that west-facing fields show the highest mean NDVI values, followed by south and east. The result was the opposite when considering seasonal variability, with west-facing fields showing less seasonal variability than those in the south and east categories. When considering the seasonal NDVI peak, the south-facing fields showed the highest values, even if the number of outliers at the low end of the plot is evident. Interestingly, west-facing fields with the highest mean values also have the lowest seasonal peaks. West-facing fields also showed a longer growing season, with an earlier start and a later end of vegetative growth, followed by south and east-facing fields. At the same time, the differences in terms of seasonal peak timing were not significant.

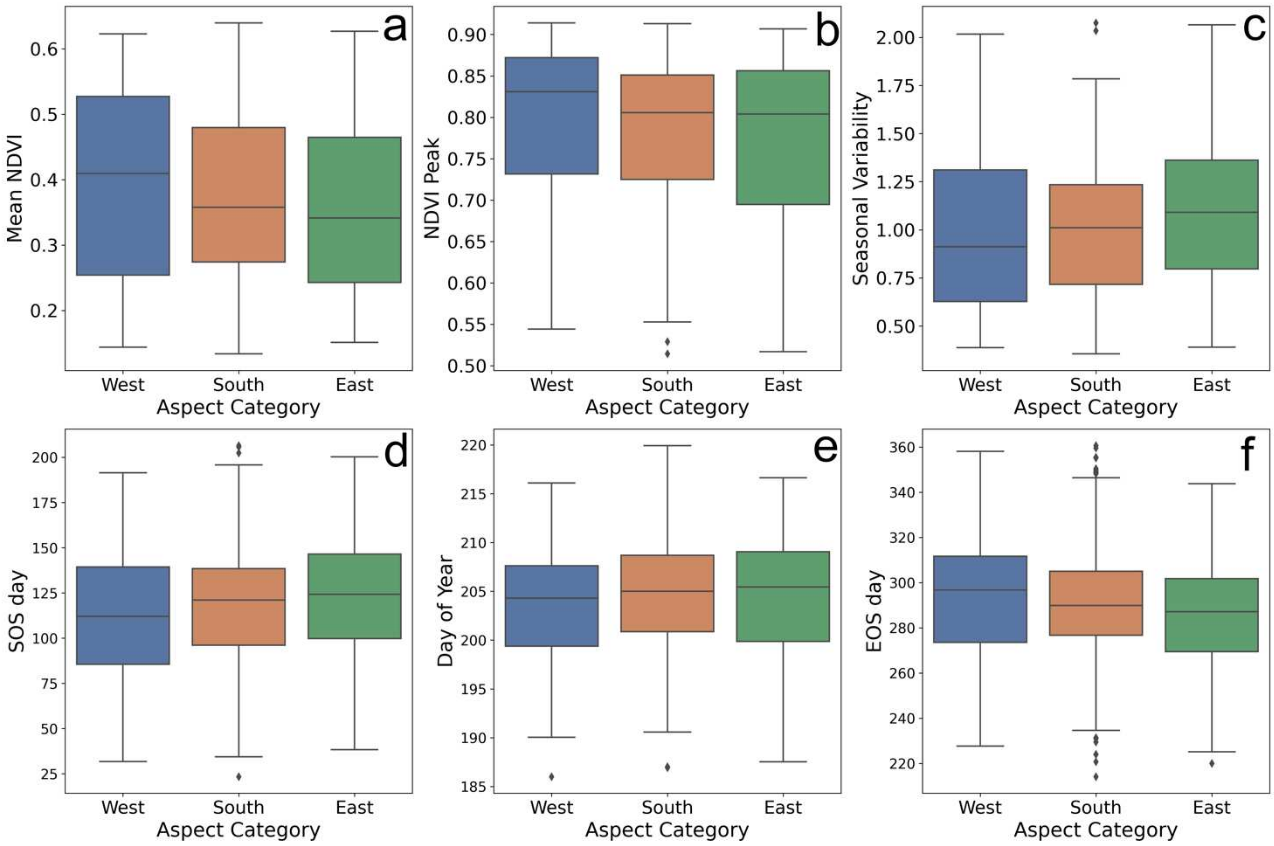


Figure 3.6 Boxplots of aspect category effects on NDVI variables: Mean NDVI (a), NDVI peak (b), Seasonal variability (c), Start of season SOS (d), Day of NDVI peak (e), End of season EOS (f).

The maps in Figures 3.7, 3.8 and 3.9 illustrate the geographical positions of the studied fields across the region, emphasising the variations in the timing of the vegetative season. Generally, grasslands in the southeastern region tend to experience an earlier season onset (Figure 3.7). In contrast, fields near high mountains, such as those in the Adamello-Brenta Park to the west and the Paneveggio Park to the east, show a later season start. The geographical distribution is similar also when referring to the peak (Figure 3.8) and the end of the season (Figure 3.9), even though in the latter case, the grasslands in the West of the region show an advanced end of the season compared to those in the East.

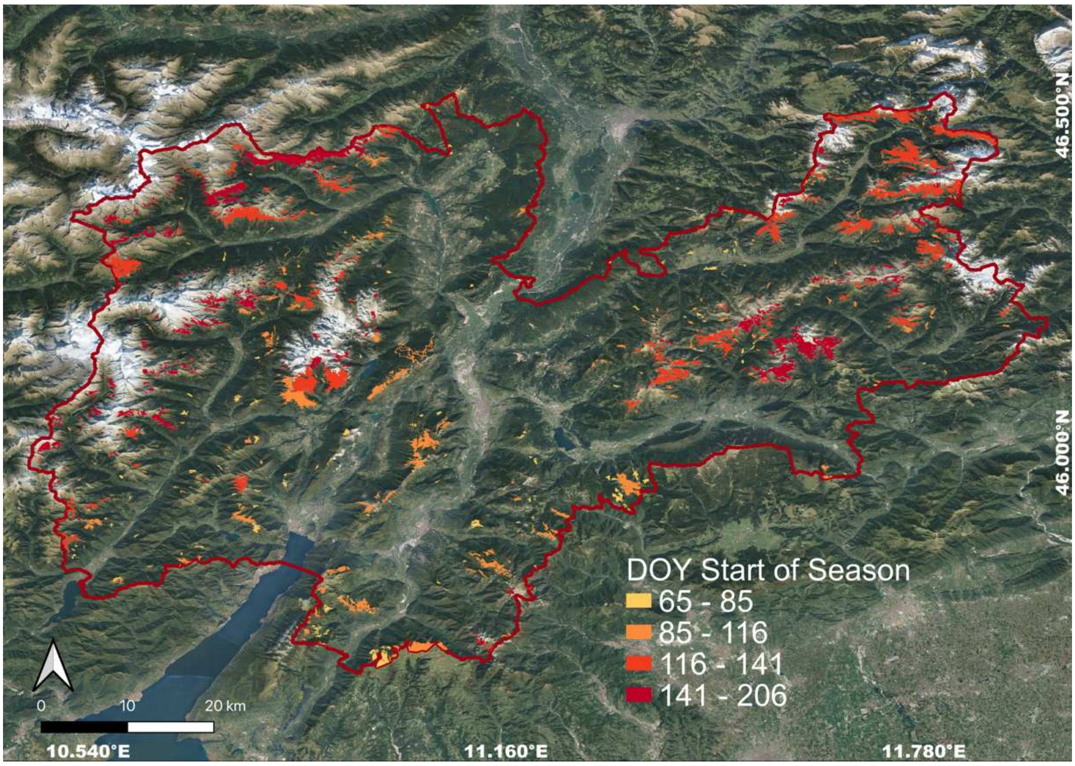


Figure 3.7 Map showing the start of the vegetative season (DOY) of pastures and grasslands in the Province of Trento, Italy.

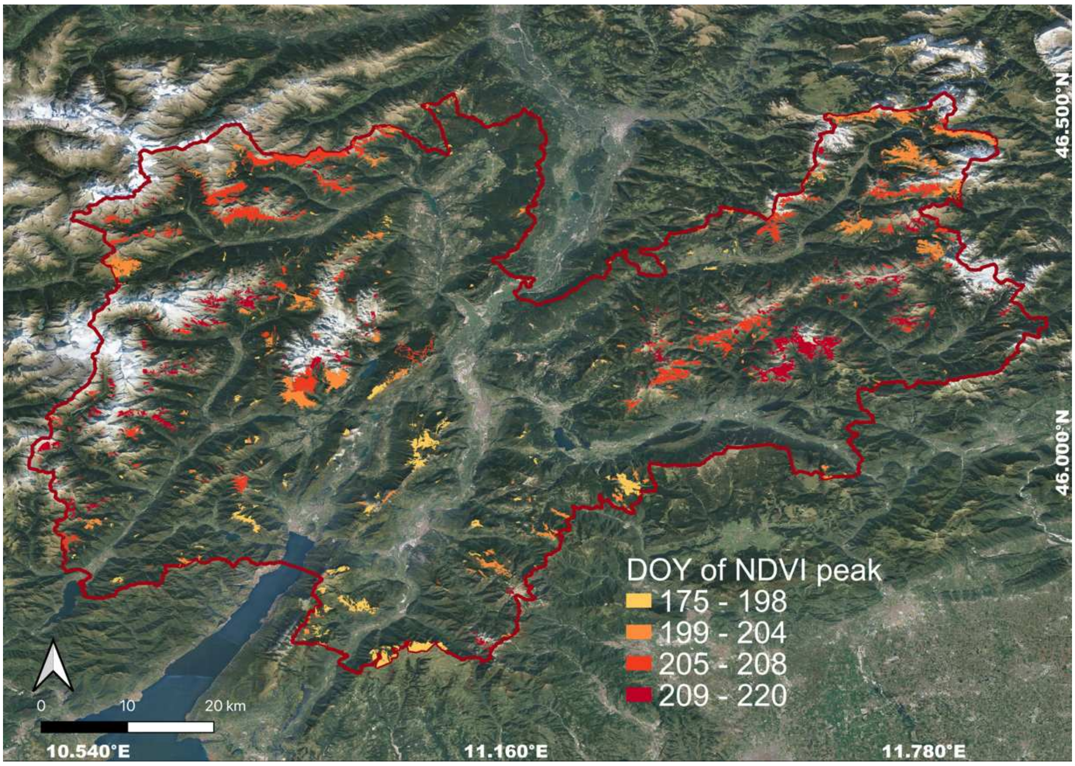


Figure 3.8 Map showing the peak of the vegetative season (DOY) of pastures and grasslands in the Province of Trento, Italy.

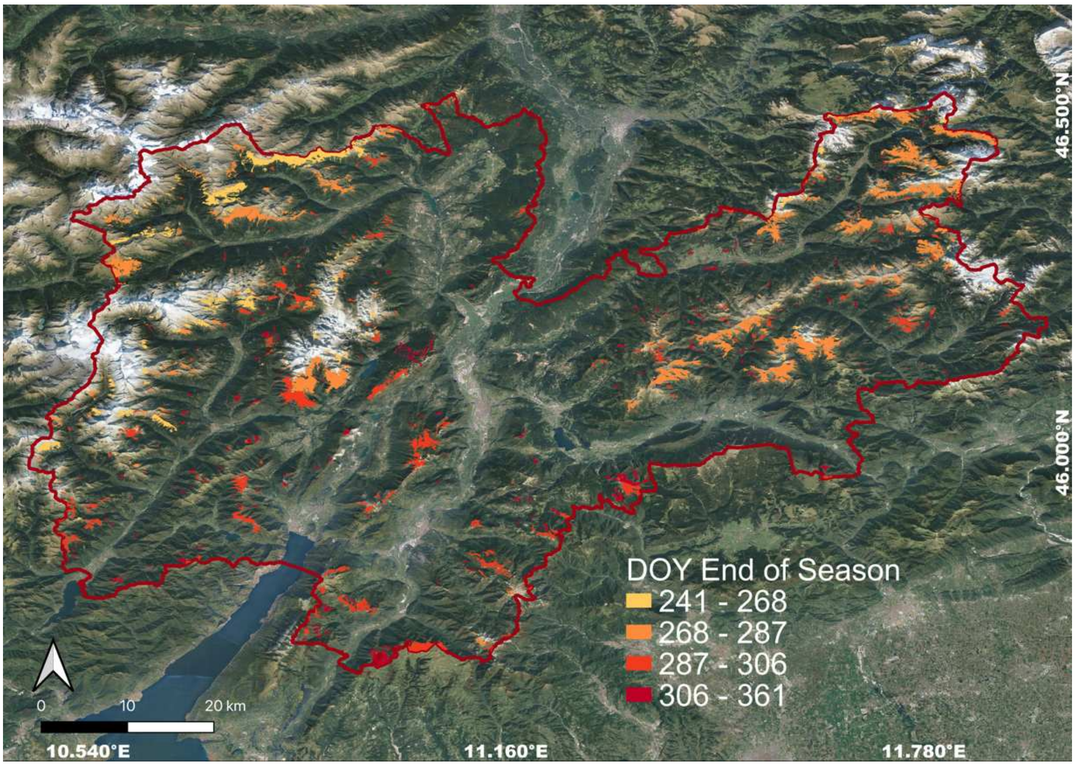


Figure 3.9 Map showing the end of the vegetative season (DOY) of pastures and grasslands in the Province of Trento, Italy.

3.4.3. Influence of meteorological variables on NDVI dynamics

The correlations between NDVI metrics and meteorological variables were analysed at the field level. The same NDVI metrics tested in Chapter 3.4.1 correlated with mean solar net radiation, mean 2m air temperature, mean wind speed, total precipitation and snowfall sum of the 23 years for each field. Figures 3.10 and 3.11 show the scatterplots between all NDVI metrics and the meteorological variables considered in this study. The influence of meteorological variables on NDVI was also analysed by testing the correlations between the mean NDVI of each year's summer growing season (April to October) with mean temperature, solar net radiation, precipitation metrics (total precipitation and number of rainy days), and snowfall metrics (total snowfall and percentage of snow cover) of the same year. The following paragraphs detail the analysis results, focusing on the meteorological variables.

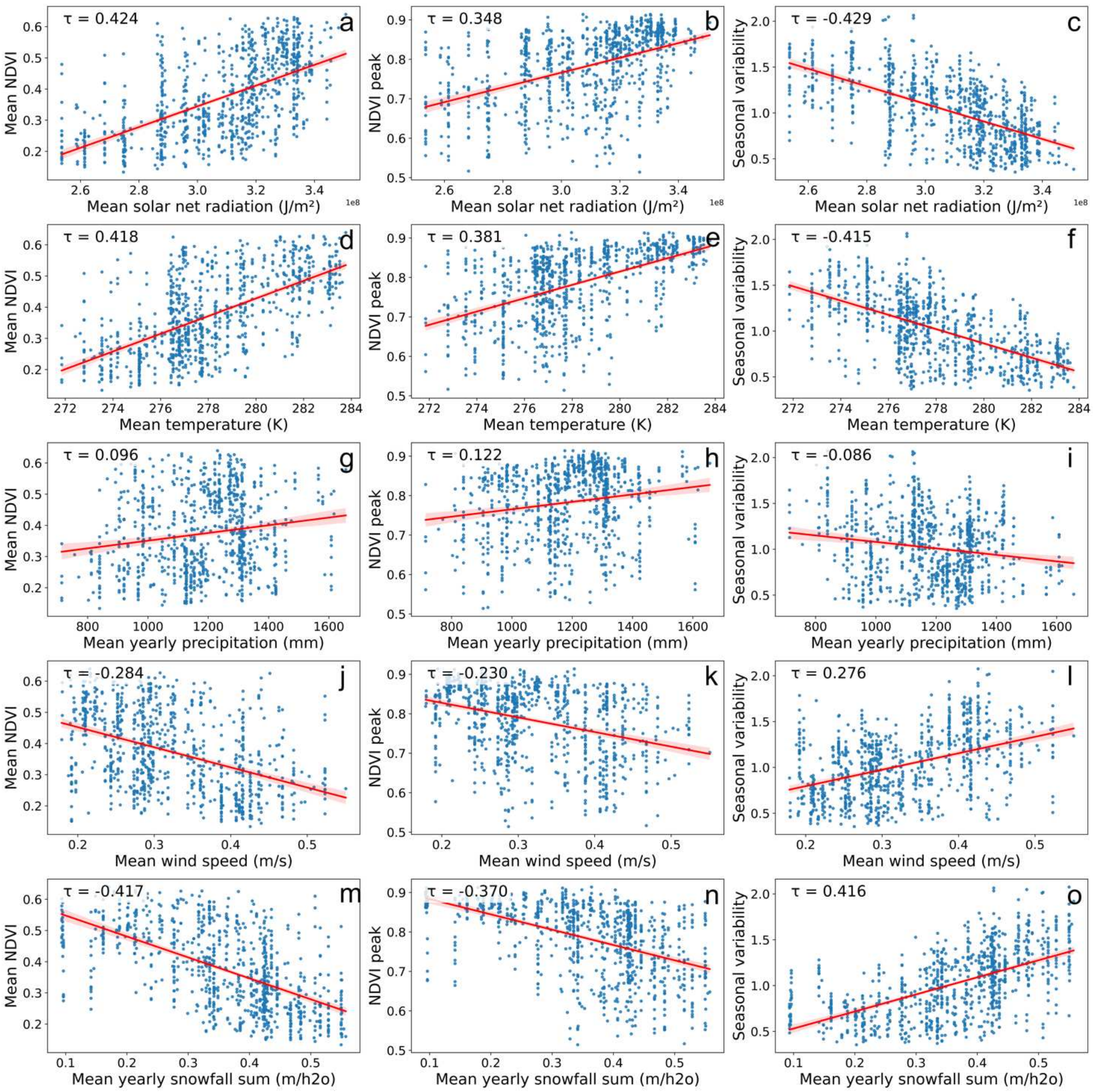


Figure 3.10. Scatterplots of correlations between mean solar net radiation (J/ m^2, plots a, b, c), mean temperature (K, plots d,e,f), mean yearly precipitation (mm, plots g, h, i), mean wind speed (m/s, plots j, k, l) and mean yearly snowfall sum (m of water equivalent, plots m, n, o) with mean NDVI, NDVI peak and Seasonal variability (normalised amplitude), Kendall's Tau coefficients are reported in each plot (τ).

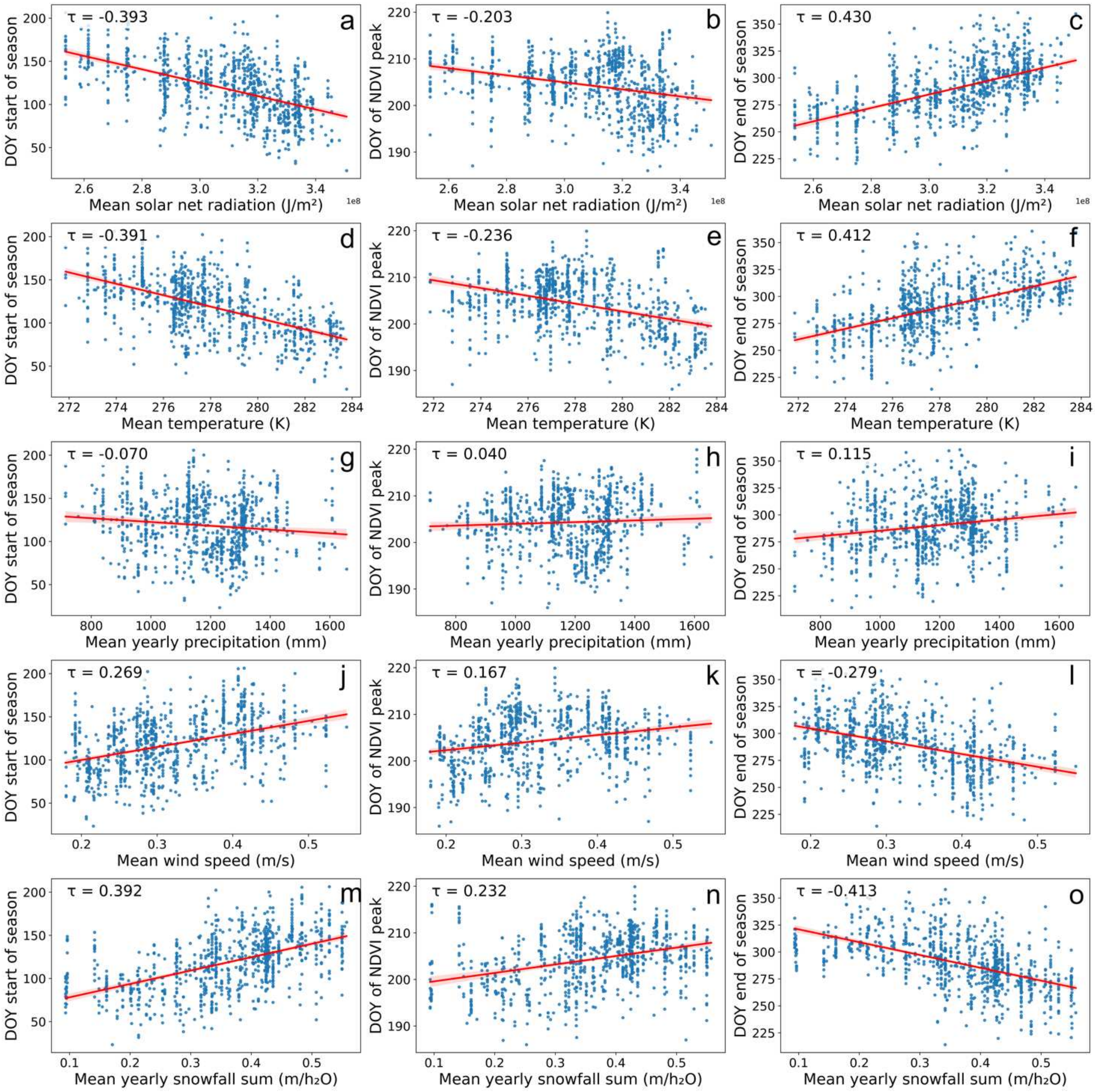


Figure 3.11. Scatterplots of correlations between mean solar net radiation (J/ m^2, plots a, b, c), mean temperature (K, plots d,e,f), mean yearly precipitation (mm, plots g, h, i), mean wind speed (m/s, plots j, k, l) and mean yearly snowfall sum (m of water equivalent, plots m, n, o) with Days of start, peak and end of the vegetative season, Kendall's Tau coefficients are reported in each plot (τ).

3.4.3.1 Effects of temperature and solar net radiation on NDVI

The relationship between NDVI dynamics, mean temperature, and solar net radiation was investigated at the field level through correlation analysis. Table 3.4 presents the Kendall correlation coefficients (τ) and associated p-values for the mean NDVI, NDVI peak, seasonal variability of NDVI, DOY of start of season, DOY of NDVI peak and DOY of end of season against mean temperature and solar net radiation over the 2001-2023 period. Both temperature and solar net radiation showed strong positive correlations with NDVI mean and peak, highlighting the role of temperature and solar radiation in enhancing overall vegetation health and productivity.

A significant negative correlation was found between seasonal variability and temperature and solar net radiation, suggesting that higher sun radiation values are associated with lower seasonal variability in vegetation greenness, implying more stable vegetation growth cycles in warmer conditions. The increasing temperature and amount of solar radiation show a lengthening effect on vegetation growth, advancing the start of the season and the day of NDVI peak (negative correlations) while delaying the end of the season (positive correlations).

Table 3.4 Kendall coefficients τ and p-values of correlations between mean Temperature (K) and Solar net radiation (J/m^2) with NDVI variables.

		Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)
Mean Temperature (K)	Kendall's coeff. τ	0.418	0.381	-0.415	-0.391	-0.236	0.412
	p value	$7.59e^{-77}$	$3.77e^{-64}$	$1.22e^{-75}$	$2.34e^{-67}$	$1.09e^{-25}$	$2.12e^{-74}$
Mean solar net radiation (J/m^2)	Kendall's coeff. τ	0.424	0.348	-0.429	-0.393	-0.203	0.430
	p value	$3.38e^{-79}$	$6.22e^{-54}$	$5.75e^{-81}$	$8.94e^{-68}$	$7.73e^{-20}$	$6.25e^{-81}$

3.4.3.2 Effects of precipitations on NDVI

The impact of mean yearly total precipitation on NDVI dynamics at the field level was investigated, and the analysis results are reported in Table 3.5.

Although the results highlight statistically significant relationships between mean yearly precipitation and mean NDVI, NDVI peak, seasonal variability, and start and end of the vegetative season, the low Kendall's coefficients reveal the absence of clear positive or negative relations between precipitation and vegetation growth magnitude and timing.

Table 3.5 Kendall coefficients τ and p-values of correlations between mean yearly precipitation (mm) with NDVI variables.

		Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)
Mean yearly precipitation (mm)	Kendall's coeff. τ	0.096	0.122	-0.086	-0.070	0.040	0.115
	p value	$2.01e^{-5}$	$6.26e^{-8}$	$1.36e^{-4}$	$1.99e^{-3}$	0.076	$3.18e^{-7}$

3.4.3.3 Effects of wind speed and direction on NDVI

The impact of wind speed on NDVI dynamics at the field level was investigated, and the results are reported in Table 6. Higher wind speeds showed a moderate negative correlation with both NDVI mean (τ = -0.28) and NDVI peak (τ = -0.23), which may suggest reduced vegetation productivity under windier conditions. At the same time, wind speed positively correlated with seasonal variability (τ = 0.28), indicating increased fluctuations in vegetation greenness. Additionally, wind speed influenced the timing of the growing season by slightly advancing the start (τ = 0.27) and the end of the season (τ = -0.28).

Table 3.6 Kendall coefficients τ and p-values of correlations between mean yearly precipitation (mm) with NDVI variables.

		Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day start of season (SOS)	Day of NDVI Peak	Day end of season (SOS)
Mean wind speed (m/s)	Kendall's coeff. τ	-0.284	-0.230	0.276	0.269	0.167	-0.279
	p value	$1.68e^{-36}$	$2.25e^{-24}$	$1.42e^{-34}$	$9.60e^{-33}$	$1.28e^{-13}$	$3.58e^{-35}$

The pastures and grasslands of our study lie in a mountainous region, where the prevalent wind direction plays an important role in shaping local microclimates and vegetation dynamics. The fields were classified into prevalent wind direction categories: East, West, and South, as there were no fields

with prevalent Northern winds. A non-parametric Kruskal-Wallis test was performed to determine whether differences among wind direction categories were significant in evaluating the effect of wind direction on NDVI dynamics. The results of the analysis are reported in Table 3.7.
The differences among wind direction categories were significant for all NDVI variables. However, the computed epsilon-squared values indicate that while wind direction has a measurable effect, its overall influence on NDVI dynamics in our study area is relatively low. The boxplots in Figure 3.12 reveal that fields exposed to prevalent south-blowing winds exhibited the highest mean NDVI. In contrast, seasonal variability was higher in fields characterised by west-blowing winds.
The timing of the growing season was also influenced by wind direction: fields characterised by west-blowing winds showed a shorter growing season with a later start and an earlier end of vegetative growth.

Table 3.7 H statistics and p-values of Kruskal-Wallis test between prevalent wind direction and NDVI variables.

	Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)
H statistic	46.247	17.307	51.072	36.526	34.822	58.10
p-value	$9.06e^{-11}$	0.0002	$8.13e^{-12}$	$1.17e^{-8}$	$2.74e^{-8}$	$2.42e^{-13}$
ε^2	0.0498	0.0172	0.0553	0.0389	0.0370	0.0632

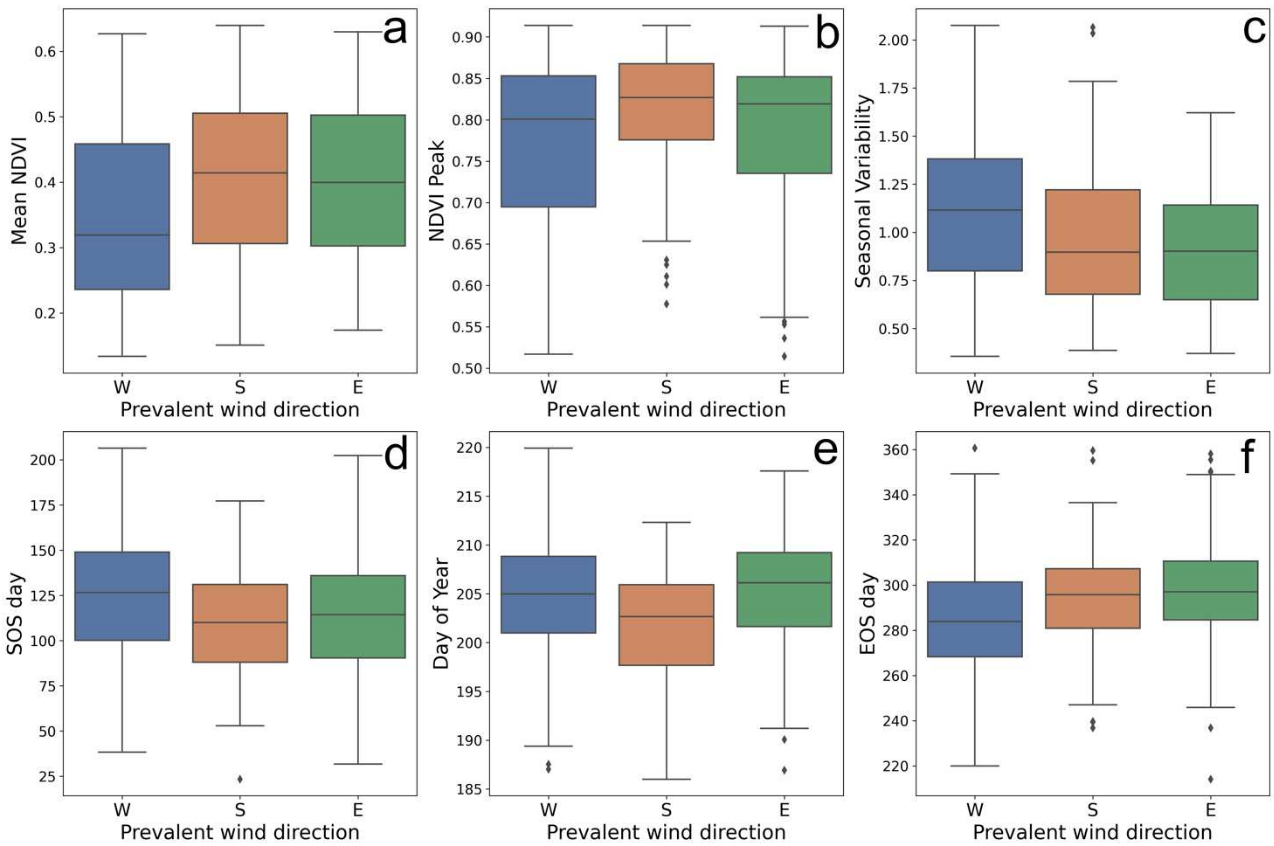


Figure 3.12. Boxplots of prevalent wind direction effects on NDVI variables: Mean NDVI (a), NDVI peak (b), Seasonal variability (c), Start of season SOS (d), Day of NDVI peak (e), End of season EOS (f).

3.4.3.4 Effects of snowfall on NDVI

The pastures and grasslands in our study area are situated in an alpine mountainous environment, with mean altitudes ranging from 800 to 2500 meters above sea level. Consequently, they are susceptible to snow cover during the winter season. However, the extent and amount of snowfall can vary significantly depending on the season, both in quantity and timing.

The influence of snowfall on NDVI dynamics at the field level was analysed and the findings are detailed in Table 3.8, indicating that snowfall amount significantly impacts NDVI dynamics. Higher snowfall amounts negatively affect vegetation growth, substantially reducing the mean and peak NDVI values. Additionally, increased snowfall contributes to more significant seasonal variability.

Increasing snowfall seems to significantly influence the timing of the vegetative season, with a shortening effect. Indeed, increasing snowfall is positively correlated with the day of the season's start (τ=0.39) and the day of the NDVI peak (τ=0.23), while negatively correlated with the end of the season (τ=-0.41), revealing a significant delay and shortening effect on vegetation development.

Table 3.8 Kendall coefficients τ and p-values of correlations between mean yearly snowfall sum (m of water equivalent) with NDVI variables.

		Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (SOS)
Mean yearly snowfall sum (m of water equivalent)	Kendall's coeff. τ	-0.417	-0.370	0.416	0.392	0.232	-0.413
	p value	$1.49e^{-76}$	$1.26e^{-60}$	$3.60e^{-76}$	$1.92e^{-67}$	$7.53e^{-25}$	$1.16e^{-74}$

3.4.4 Influence of edaphic factors on NDVI dynamics

Soil characteristics can significantly impact the growing conditions of mountain grassland vegetation. The impact of some edaphic factors was considered and analysed in the study. Kendall's coefficients and p-values of correlations between soil organic carbon, soil pH, soil water content at field capacity and soil clay, silt and sand fractions with NDVI variables are reported in Table 3.9.

In our study area, soil organic carbon content does not seem to influence vegetation growth and phenology, as correlation analysis reveals no significant relations with all the NDVI metrics.

In contrast, soil pH shows an effect on vegetation growth, both in terms of magnitude and timing. Most soils in our study area can be classified as moderately to slightly acidic, and neutral at the upper end, with pH values ranging from 5.17 to 6.75. Generally, higher pH seems to favour vegetation growth, positively correlating with mean and peak NDVI, also lowering seasonal variability, as shown in Figure 3.13 (d,e,f). More acidic soils also show an effect on the length and timing of the growing season, as lower pH values seem to postpone the start and peak of the season, while advancing the end, as visible in Figure 3.14 (d,e,f).

Table 3.9 Kendall's Tau coefficients τ and p-values of correlations between Soil organic carbon, soil pH, soil water content at 33kPa, soil clay fraction, soil silt fraction and soil sand fraction with NDVI variables.

		Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)

Soil organic carbon (g/kg),	Kendall's coeff. τ	-0.029	0.004	0.033	0.023	0.022	-0.015
	p value	0.386	0.489	0.253	0.595	0.898	0.765
Soil pH	Kendall's coeff. τ	0.301	0.298	-0.291	-0.318	-0.384	0.250
	p value	$2.84e^{-37}$	$3.50e^{-37}$	$4.39e^{-35}$	$2.21e^{-41}$	$2.21e^{-62}$	$4.08e^{-26}$
Soil water content at field capacity (%)	Kendall's coeff. τ	0.449	0.432	-0.403	-0.440	-0.291	0.452
	p value	$7.16e^{-86}$	$2.64e^{-80}$	$1.34e^{-69}$	$9.78e^{-83}$	$8.61e^{-38}$	$7.02e^{-87}$

Among all the edaphic factors considered in our study, soil water content at field capacity shows the most significant influence on grassland vegetation dynamics, positively influencing both magnitude and length of the growing season. Figure 3.13 (g,h,i) shows that fields with higher water storage capacity also show higher mean (τ=0.45) and peak (τ=0.43) NDVI values, also significantly reducing the seasonal variability (τ=-0.40). Water storage capacity also affects the growing season's length by advancing the start and peak and delaying the end of grassland vegetation growth, as visible in Figure 3.14 (g,h,i).

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITÀ DEGLI STUDI DI PADOVA

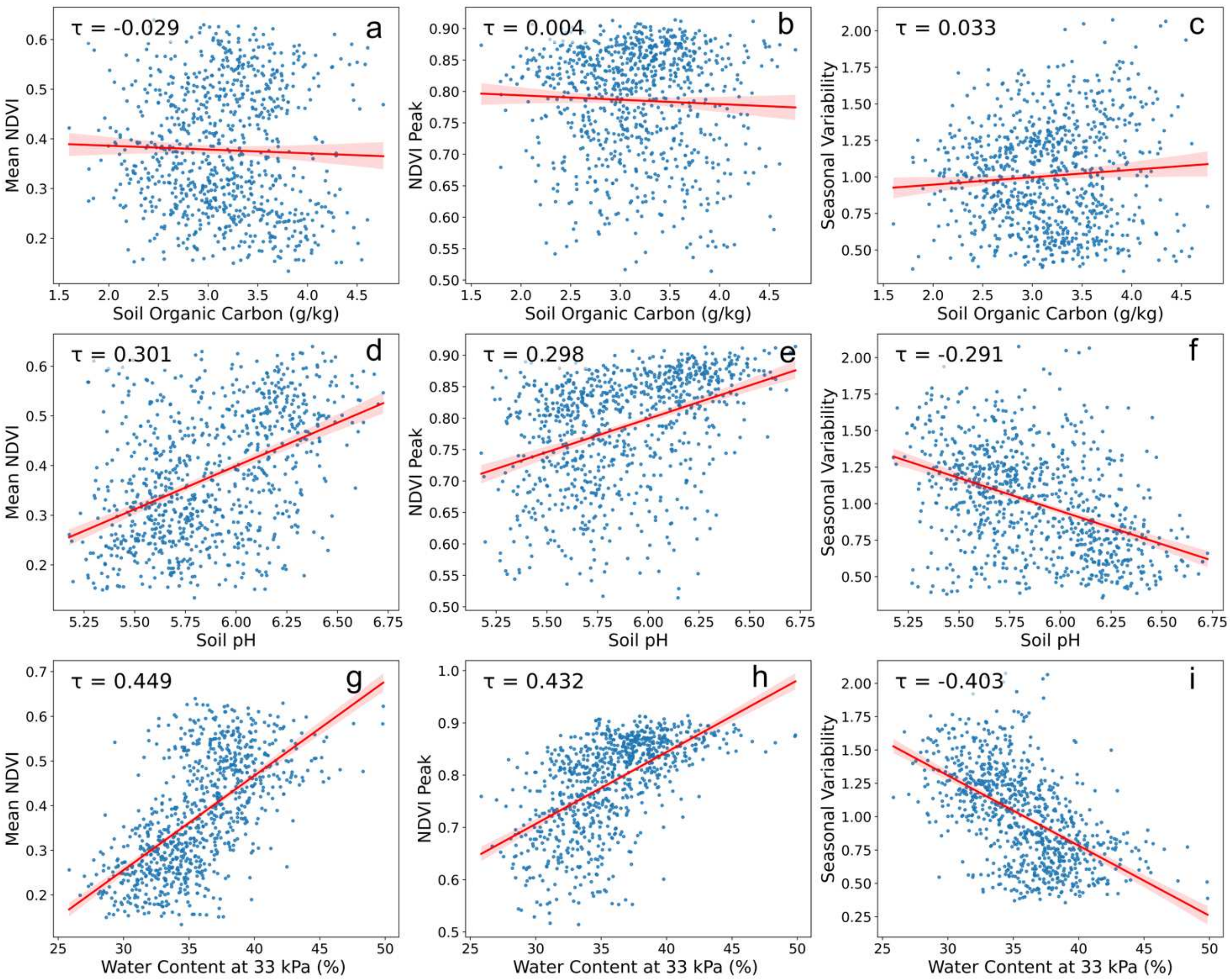


Figure 3.13. Scatterplots of correlations between soil organic carbon (kg/m^3, plots a,b,c), soil pH (plots d,e,f) and soil water content at 33kPa (%, plots g,h,i), with mean NDVI, NDVI peak and Seasonal variability (normalised amplitude), Kendall's Tau coefficients are reported in each plot (τ).

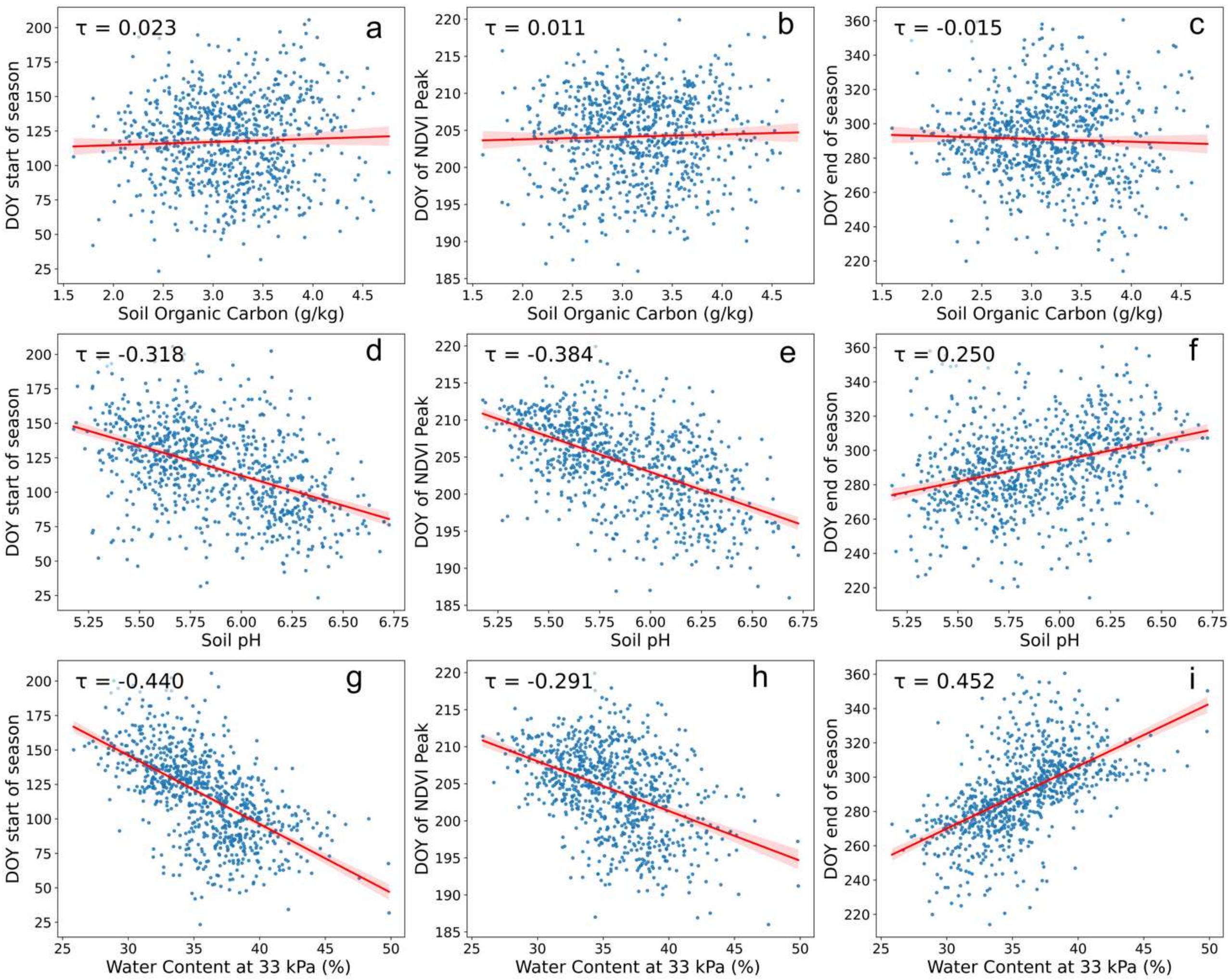


Figure 3.14 Scatterplots of correlations between soil organic carbon (kg/m^3, plots a,b,c), soil pH (plots d,e,f) and soil water content at 33kPa (%, plots g,h,i), with day of start, peak and end of the vegetative season, Kendall's Tau coefficients are reported in each plot (τ).

Soil texture and composition are one of the most critical drivers of plant growth. This study uses clay, silt and sand fractions of each field to classify soils according to the USDA soil texture triangle (Ditzler et al., 2017). The fields in our study area can be divided into three main soil classes: silt loam (471 fields), sandy loam (269 fields) and loam (150). The effects of soil texture groups on NDVI metrics were also tested using a Kruskal-Wallis test. The results are reported in Table 3.10, and the boxplots in Figure 3.15 show differences among soil groups.

Soil texture showed significant effects on all NDVI variables, even though modest ε^2 suggest that while these effects are statistically significant, soil texture alone does not fully explain the variability in vegetation dynamics. Generally, silt loam soils offer a better growing habitat for alpine grassland vegetation, with higher mean and peak NDVI values and lower seasonal variability observed. Fields with silt loam soils also favour longer growing seasons, but advancing the start and peak of vegetative

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Università degli Studi di Padova

season and delaying the end. Sandy loam soils seem less hospitable for mountain grassland vegetation in our study area, showing the lower mean NDVI and the higher seasonal variability, also shortening the length of the growing season. Finally, Loam soils, the least represented group, show very high variability for all the tested variables, showing the lower NDVI peaks and a very high seasonal variability.

Table 3.10. H statistics and p-values of Kruskal-Wallis test between USDA soil type and NDVI variables.

	Mean NDVI (23y)	NDVI peak (23y)	Seasonal variability (normalized amplitude)	Day of start of season (SOS)	Day of NDVI Peak	Day of end of season (EOS)
H statistic	75.02	94.57	45.08	80.16	84.48	65.53
p-value	$5.12e^{-17}$	$2.91e^{-21}$	$1.63e^{-10}$	$3.93e^{-18}$	$4.52e^{-19}$	$5.89e^{-15}$
ε^2	0.1028	0.1304	0.0607	0.1101	0.1162	0.0895

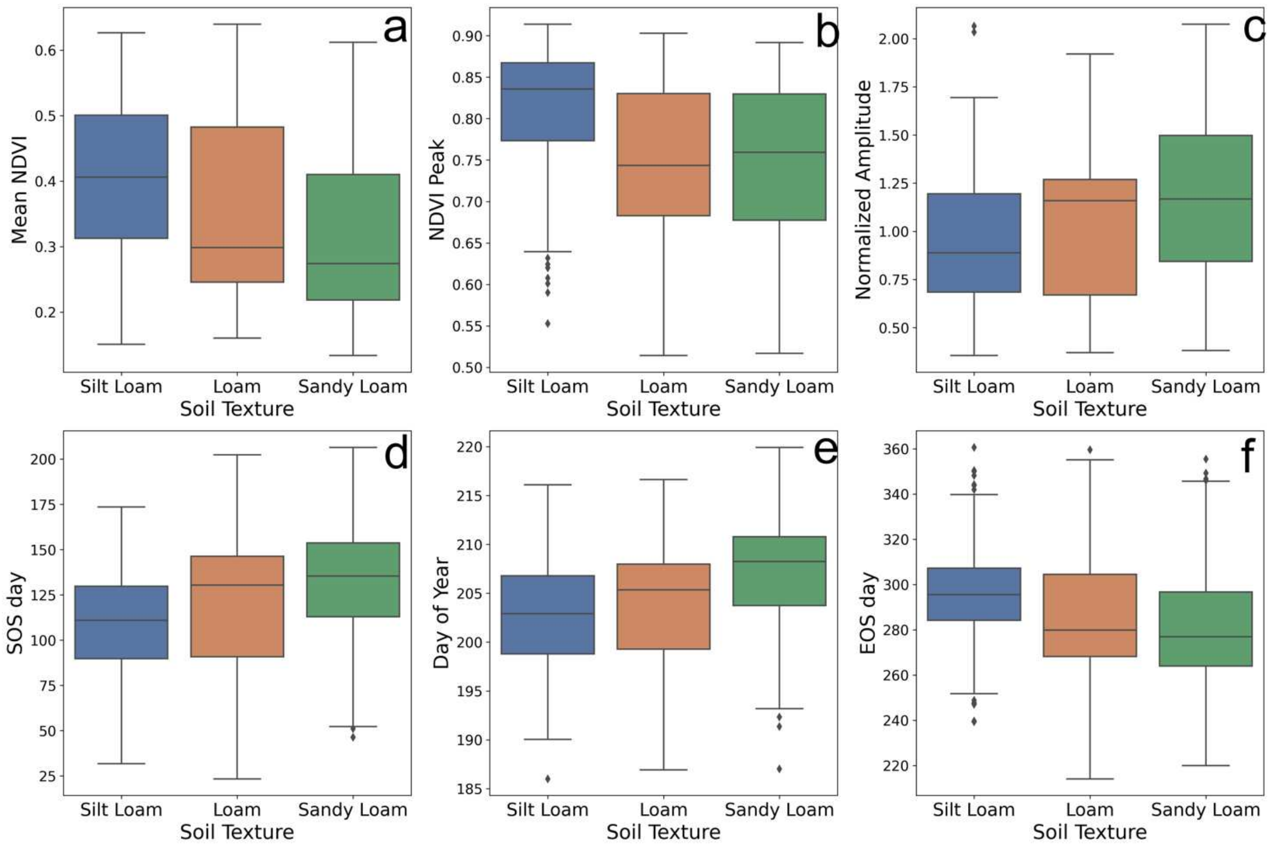


Figure 3.15 Boxplots of USDA soil classes effects on NDVI variables: Mean NDVI (a), NDVI peak (b), Seasonal variability (c), Start of season SOS (d), Day of NDVI peak (e), End of season EOS (f).

3.4.5 Environmental drivers cross-correlation analysis

Understanding the relationships among environmental variables is crucial before proceeding with phenological modelling, as strong correlations may indicate potential multicollinearity issues that could influence the reliability of subsequent analyses and conclusions. Based on Kendall's tau, the cross-correlation analysis provides insight into the interdependencies among key soil, climatic, and topographical factors, helping to identify redundant predictors and assess how different environmental drivers interact. The coefficients in the correlation matrix in Figure 3.16, highlight a relationship between soil texture and soil pH, with clay and silt fractions exhibiting a positive correlation with pH. This association is expected, as finer soil particles typically buffer soil acidity due to their higher cation exchange capacity and ability to retain moisture. Soil water holding capacity, a critical indicator of soil moisture availability, is influenced by both climatic and topographical factors, showing negative correlations with elevation and slope, suggesting that higher-altitude and steeper areas tend to have lower water retention capacity. Conversely, positive correlations with precipitation and mean temperature suggest that climate substantially modulates soil water dynamics.

Altitude emerges as a key variable affecting multiple environmental parameters, displaying a moderate inverse correlation with soil pH, clay content, and soil water retention, indicating that higher-elevation soils are generally more acidic, coarser in texture, and less capable of retaining moisture. Similarly, temperature and radiation are associated with soil water availability, as increasing temperature promotes evaporation and seasonal snowmelt processes, while radiation can enhance evapotranspiration, influencing soil moisture dynamics. The relationship between snowfall and temperature follows an expected inverse trend, where warmer conditions correspond to lower snowfall accumulation, reinforcing the well-established link between climate warming and reduced snow cover duration. Wind speed, another influential climatic driver, exhibits negative correlations with soil properties such as clay fraction and water retention. This suggests that areas with more substantial wind exposure are likely to have coarser soil textures, potentially due to the erosion of finer particles, leading to lower water-holding capacity. The influence of wind on soil conditions has essential implications for vegetation distribution, as it may contribute to the desiccation of exposed areas, further constraining plant growth in high-altitude environments.

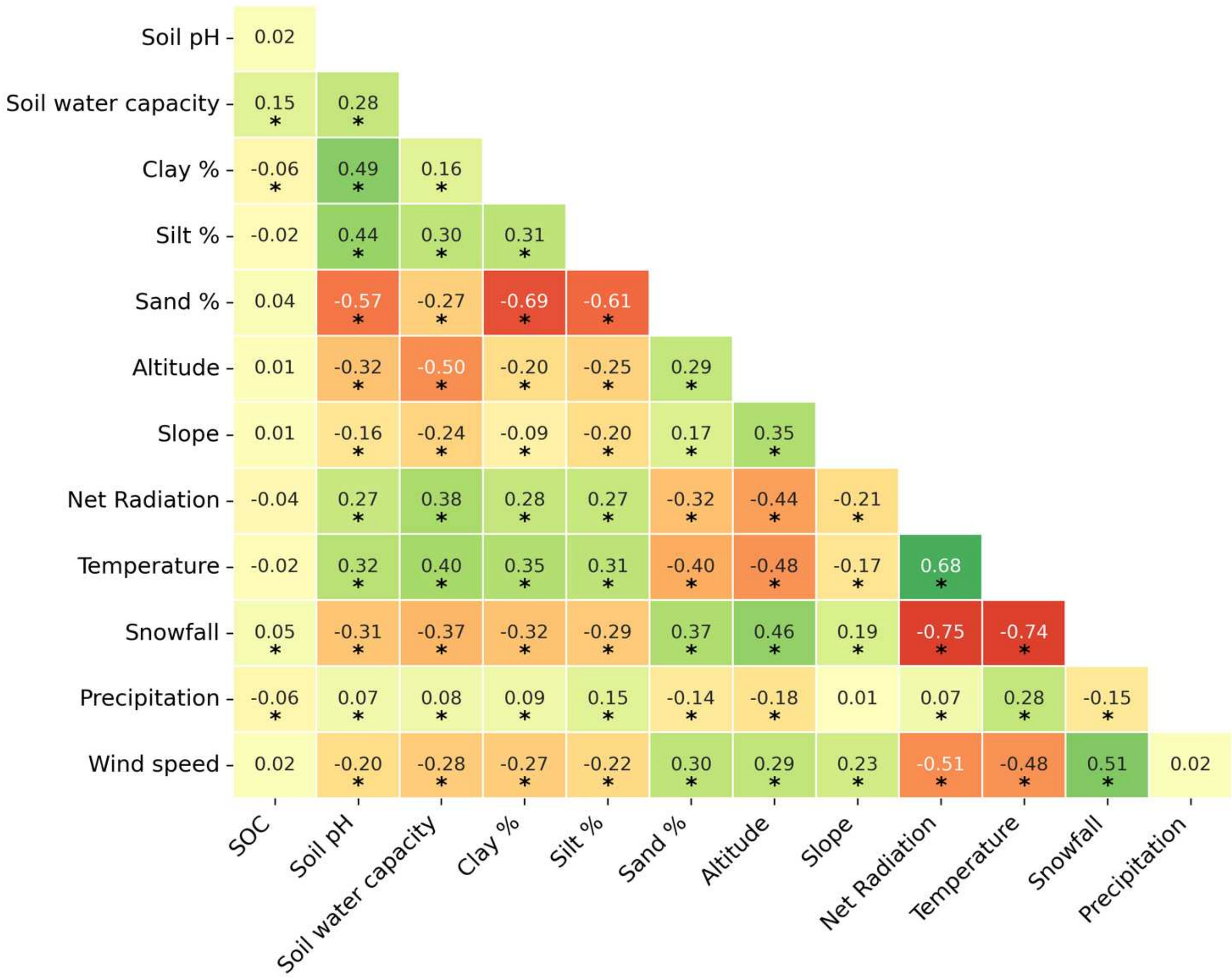


Figure 3.16 Kendall's tau correlation matrix of environmental drivers showing pairwise correlations among soil, climatic, and topographical variables. Asterisk indicates statistically significatn correlations (p-value <0.05).

3.4.6 Generalized additive models to predict vegetation's phenology

Generalized Additive Models (GAMs) were developed for six NDVI metrics: Mean NDVI, NDVI Peak, Normalized Amplitude, SOS day, Day of Year, and EOS day, using altitude, slope, solar net radiation, temperature, precipitation, wind speed, snowfall, soil organic carbon, soil pH, soil water content at field capacity and soil texture as predictors. In this modelling phase of the study, clay, silt and sand fraction were considered as separate continous variables, to be able to catch all the variability contained in the dataset. Table 3.11 reports the root mean square error (RMSE) and coefficient of determination (R^2) metrics, evaluatig the predicting performances of each model. Furthermore, estimated degrees of freedom (edf), F-statistic (F), p-value and level of significance are reported for each predictor. Finally, Figure 3.17 shows the scatterplots of observed vs predicted values of all models.

Table 3.11 Generalized Additive Models predicting NDVI metrics (Variable). RMSE = root mean square error; R^2 = coefficient of determination; edf = estimated degrees of freedom; F = F-statistic. Levels of significance: *** = p<0.001; ** = p<0.01; *=p<0.05.

Variable	RMSE	R^2	Predictor	edf	F	p-value	Sig.
Mean NDVI	0.040	0.900	Altitude	5.3	208.5	$<2.0e^{-16}$	***
			Slope	1.1	0.655	$8.39e^{-3}$	**
			Solar net radiation	1.8	0.984	$1.96e^{-4}$	***
			Mean temperature	1.1	0.262	0.065	*
			Mean precipitation	2.0	4.262	$<2.0e^{-16}$	***
			Mean wind speed	$3.2e^{-4}$	$7.6e^{-6}$	0.677	
			Mean snowfall	3.0	1.788	$8.81e^{-5}$	***
			Soil organic carbon	2.8	1.371	$2.18e^{-3}$	**
			Soil pH	2.5	1.326	$1.75e^{-3}$	**
			Soil water content	4.3	1.748	$1.94e^{-3}$	**
			Soil clay fraction	0.0	0.000	0.304	
			Soil silt fraction	4.0	3.303	$7.36e^{-7}$	***
			Soil sand fraction	3.6	1.755	$4.36e^{-4}$	***
NDVI peak	0.050	0.686	Altitude	4.7	66.24	$<2.0e^{-16}$	***
			Slope	1.0	0.416	0.031	**
			Solar net radiation	1.9	2.212	$9.8e^{-6}$	***
			Mean temperature	$8.2e^{-4}$	$8.1e^{-5}$	0.304	***
			Mean precipitation	1.6	2.298	$2.7e^{-6}$	***
			Mean wind speed	$2.7e^{-3}$	$2.6e^{-4}$	0.347	
			Mean snowfall	$2.0e^{-4}$	$9.8e^{-7}$	0.921	
			Soil organic carbon	2.1	0.627	0.040	*
			Soil pH	2.7	1.833	$1.93e^{-4}$	***
			Soil water content	$1.3e^{-4}$	$2.2e^{-6}$	0.754	
			Soil clay fraction	2.8	0.852	$2.1e^{-3}$	**
			Soil silt fraction	2.5	1.011	$5.9e^{-4}$	***
			Soil sand fraction	4.0	1.647	$4.8e^{-5}$	***
Seasonal Variability	0.136	0.846	Altitude	5.2	120.3	$<2.0e^{-16}$	***
			Slope	2.2	8.026	$<2.0e^{-16}$	***

			Solar net radiation	1.5	0.919	$1.6e^{-3}$	**
			Mean temperature	1.4	0.452	0.023	*
			Mean precipitation	2.0	2.455	$4.31e^{-7}$	***
			Mean wind speed	$2.1e^{-4}$	$1.2e^{-5}$	0.475	
			Mean snowfall	3.5	2.518	$3.7e^{-6}$	***
			Soil organic carbon	2.6	1.488	$1.2e^{-3}$	**
			Soil pH	$2.3e^{-4}$	$1.0e^{-5}$	0.533	
			Soil water content	4.1	2.363	$8.2e^{-5}$	***
			Soil clay fraction	4.2	1.492	$5.0e^{-4}$	***
			Soil silt fraction	2.8	0.696	0.017	*
			Soil sand fraction	3.9	1.899	$1.5e^{-5}$	***
DOY start of season	12.34	0.853	Altitude	4.5	167.8	$<2.0e^{-16}$	***
			Slope	1.3	1.266	$4.1e^{-4}$	***
			Solar net radiation	3.4	1.807	$2.2e^{-4}$	***
			Mean temperature	$8.6e^{-4}$	$1.8e^{-5}$	0.802	
			Mean precipitation	2.2	7.949	$<2.0e^{-16}$	***
			Mean wind speed	2.4	0.894	0.011	***
			Mean snowfall	2.6	1.939	$4.7e^{-5}$	***
			Soil organic carbon	2.6	1.054	$7.4e^{-3}$	**
			Soil pH	2.8	2.453	$1.2e^{-5}$	***
			Soil water content	0.8	0.190	0.108	
			Soil clay fraction	$5.0e^{-3}$	$3.5e^{-4}$	0.305	
			Soil silt fraction	4.0	3.469	$<2.0e^{-16}$	***
			Soil sand fraction	3.7	1.863	$1.8e^{-4}$	***
DOY of NDVI peak	3.150	0.715	Altitude	3.1	60.63	$<2.0e^{-16}$	***
			Slope	2.5	1.069	$7.3e^{-3}$	**
			Solar net radiation	3.8	3.739	$<2.0e^{-16}$	***
			Mean temperature	3.0	4.935	$<2.0e^{-16}$	***
			Mean precipitation	2.7	13.38	$<2.0e^{-16}$	***
			Mean wind speed	3.5	4.059	$<2.0e^{-16}$	***
			Mean snowfall	1.7	3.347	$<2.0e^{-16}$	***
			Soil organic carbon	2.3	0.708	0.031	*
			Soil pH	1.1	0.606	0.010	*

			Soil water content	0.9	0.380	0.034	*
			Soil clay fraction	1.6	4.031	<2.0e^{-16}	***
			Soil silt fraction	1.8e^{-4}	6.2e^{-6}	0.065	
			Soil sand fraction	2.3e^{-3}	1.2e^{-4}	0.049	*
DOY end of season	11.79	0.788	Altitude	4.7	90.5	<2.0e^{-16}	***
			Slope	1.4	1.285	5.3e^{-4}	**
			Solar net radiation	1.7	1.478	2.1e^{-5}	***
			Mean temperature	1.3	0.443	0.017	*
			Mean precipitation	1.5	1.498	3.2e^{-5}	***
			Mean wind speed	0.2	0.028	0.243	
			Mean snowfall	4.4e^{-3}	4.3e^{-4}	0.299	
			Soil organic carbon	2.33	0.795	0.022	*
			Soil pH	2.93	3.981	<2.0e^{-16}	***
			Soil water content	3.87	2.161	1.7e^{-4}	***
			Soil clay fraction	2.2e^{-3}	2.3e^{-4}	0.077	
			Soil silt fraction	3.99	4.435	<2.0e^{-16}	***
			Soil sand fraction	2.61	1.433	5.3e^{-4}	***

The models showed good performances, capturing a substantial proportion of the variance across all response variables, with R^2 values ranging from 0.686 to 0.900, as shown in Table 3.11 and Figure 3.17, indicating strong predictive ability. The models estimating vegetation growth exhibited strong predictive power, with Mean NDVI achieving the highest performance (R^2=0.900, RMSE=0.046), confirming that the selected environmental predictors effectively captured seasonal vegetation dynamics. Seasonal Variability was also well-explained (R^2=0.846, RMSE=0.136), reinforcing the robustness of the approach in capturing variations in vegetation. Conversely, NDVI Peak demonstrated the lowest predictive accuracy (R^2=0.686, RMSE=0.050). For models estimating phenological development, Start of Season and End of Season models explained a high proportion of variance (R^2=0.853 and R^2=0.788, respectively), with also very interesting RMSE of about 11 days. The model predicting the day of NDVI peak exhibited a lower RMSE (3 days), despite having a more moderate R^2 value (0.715). This difference is likely due to the variability in the range of observed values: while SOS and EOS varied considerably across the studied fields (110-day difference for SOS and 120-day difference for EOS), DOY Peak exhibited a much narrower range of only 34 days. This reduced variability likely contributed to the model's lower RMSE, despite explaining less overall variance than the SOS and EOS models. Across all models, altitude emerged as the most influential predictor, consistently showing high estimated degrees of freedom (edf) and intense significance levels ($p < 2.0e^{-16}$). Other significant predictors included slope, which, while significant in most models, had a relatively weaker effect than altitude. Solar net radiation was a strong determinant of vegetation indices, with significant contributions across all models ($p < 0.001$), confirming its importance in regulating vegetation productivity. Mean temperature and precipitation contributed to variations in NDVI metrics, particularly for seasonal dynamics such as DOY Peak and EOS, indicating the relevance of climatic factors in vegetation growth and senescence, but their contribution was less significant. Soil properties, including organic carbon, pH, water-holding capacity, and texture, significantly influenced vegetation dynamics. Higher soil organic carbon and water content supported plant productivity by improving nutrient availability and moisture retention, while soil texture affected water availability and root development. Silt-rich soils favored an earlier start of the growing season, while sandy soils, with lower water retention, were associated with delayed emergence. Soil properties also played a role in stabilizing seasonal vegetation trends, with higher water content helping buffer biomass fluctuations. Additionally, soil factors influenced the timing of NDVI peak and the end of the growing season, with better moisture retention extending vegetation activity.

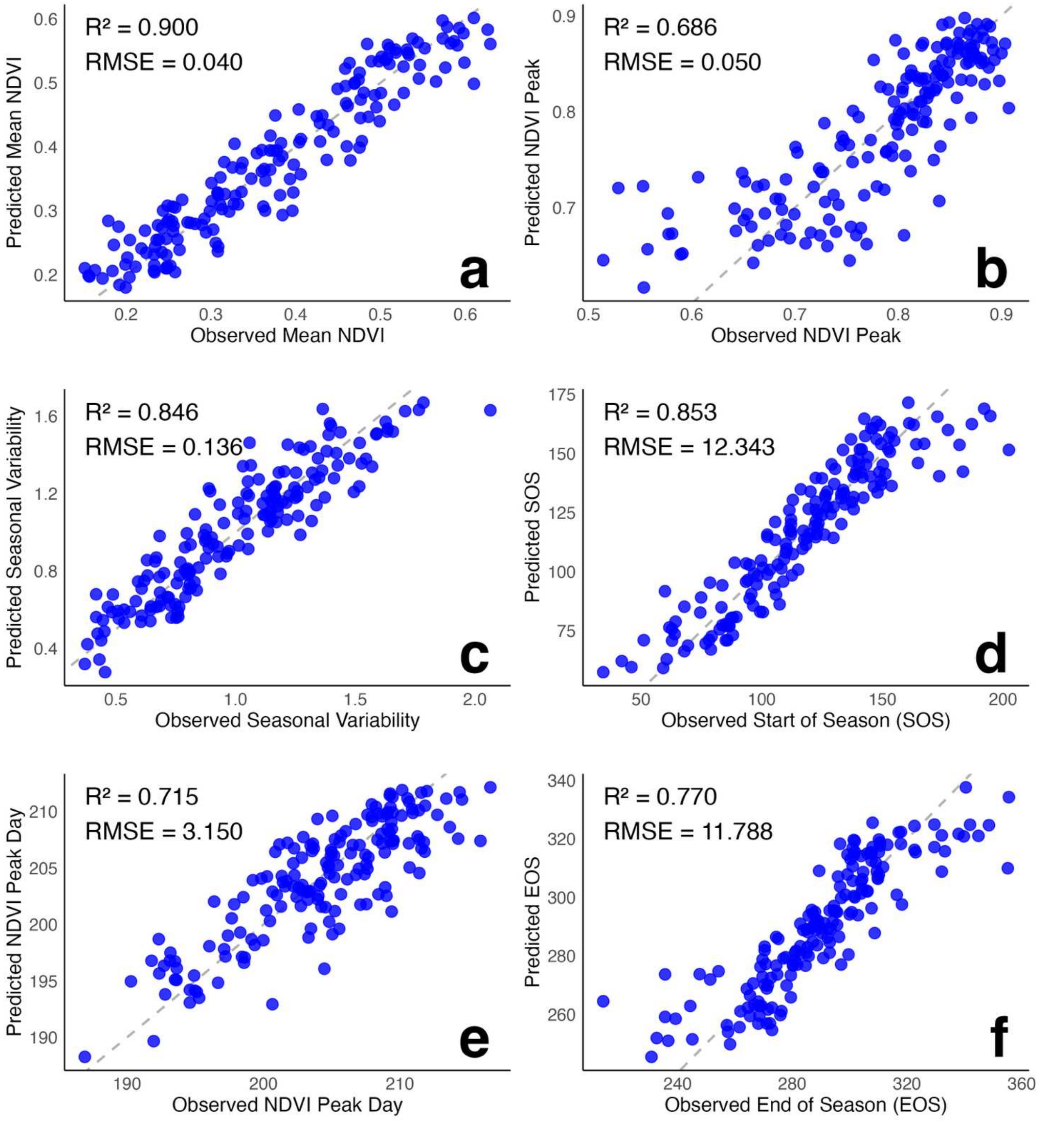


Figure 3.17. Scatterplots of observed vs predicted NDVI variables of GAM models, R2 and RMSE reported in each plot: a = Mean NDVI, b = NDVI peak, c = Seasonal Variability, d = Start of Season, e = NDVI Peak Day f = End of Season.

3.4.6.1 Models evaluation and diagnostics

The adequacy of the GAM models was assessed through residual diagnostics, including observed vs. fitted values, quantile-quantile (Q-Q) plots, and residual distribution analysis. These evaluations ensure that the models meet the assumptions required for reliable inference, as shown in Figures 3.18 and 3.19. The observed vs. fitted values plots show a strong alignment between predicted and actual values across most models. This indicates that the GAMs effectively capture the underlying relationships between vegetation indices and their predictors. However, slight deviations at extreme values suggest unexplained variability, potentially due to unaccounted factors such as land management practices or sub-seasonal climate fluctuations.

Q-Q plots were used to assess the normality of residuals. The results indicate that residuals generally follow a normal distribution, as evidenced by the close adherence of points to the diagonal line. Minor deviations at the tails are present, especially relatively to the NDVI peak and the SOS model, suggesting that the model may not fully capture some extreme values. Still, these effects are not substantial enough to undermine model validity.

The distribution of residuals was further evaluated through histogram analysis. Residuals exhibit a symmetric, bell-shaped distribution for all models, reinforcing the assumption of normally distributed errors. No severe skewness or heteroscedasticity was detected, suggesting that the models provide stable and unbiased estimates across the range of observed values.

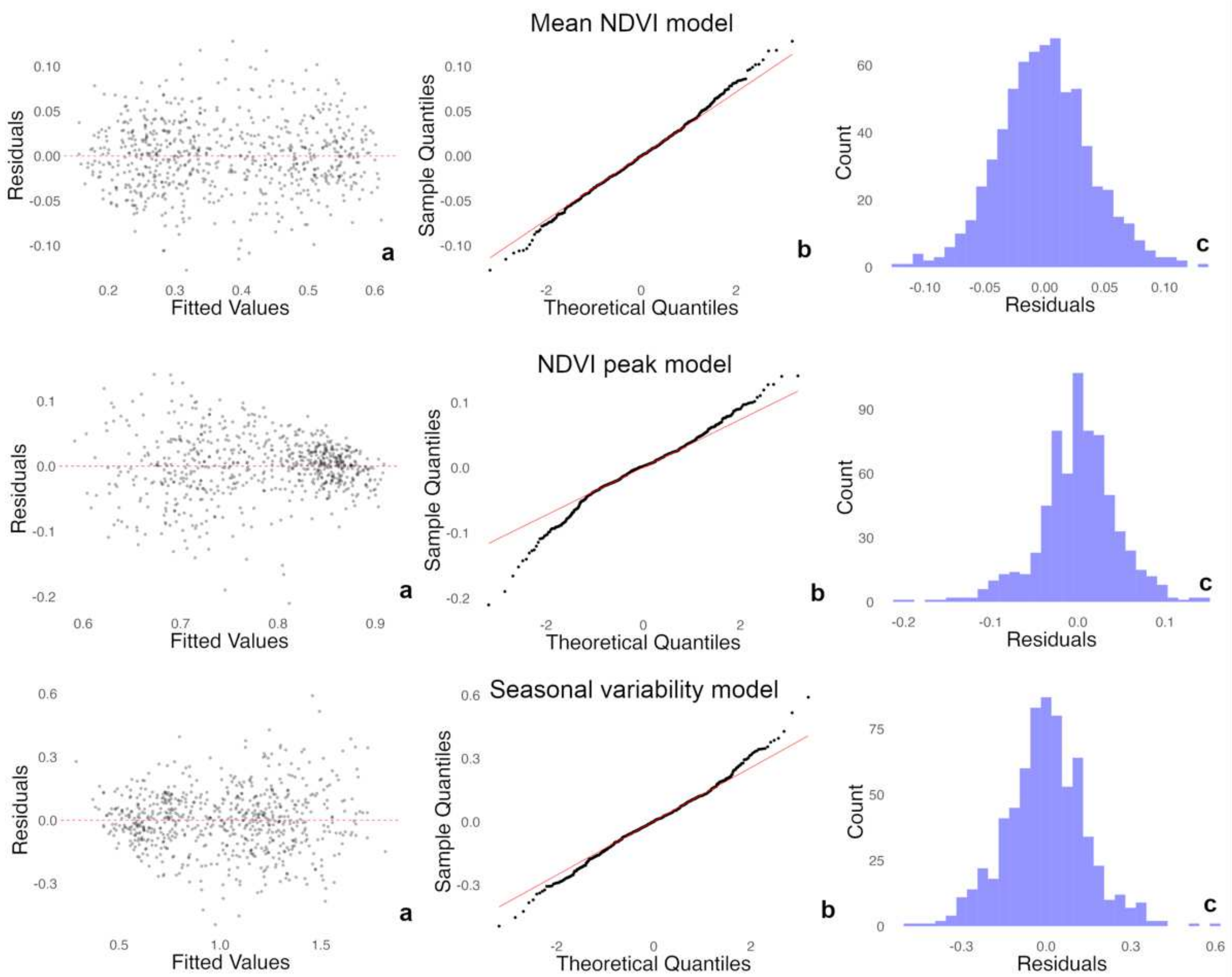


Figure 3.18. From left to right: Residual vs fitted plots (**a**), Quantile-Quantile plots (**b**) and Histogram of residuals distribution (**c**) for Mean NDVI, NDVI peak and seasonal variability GAM models.

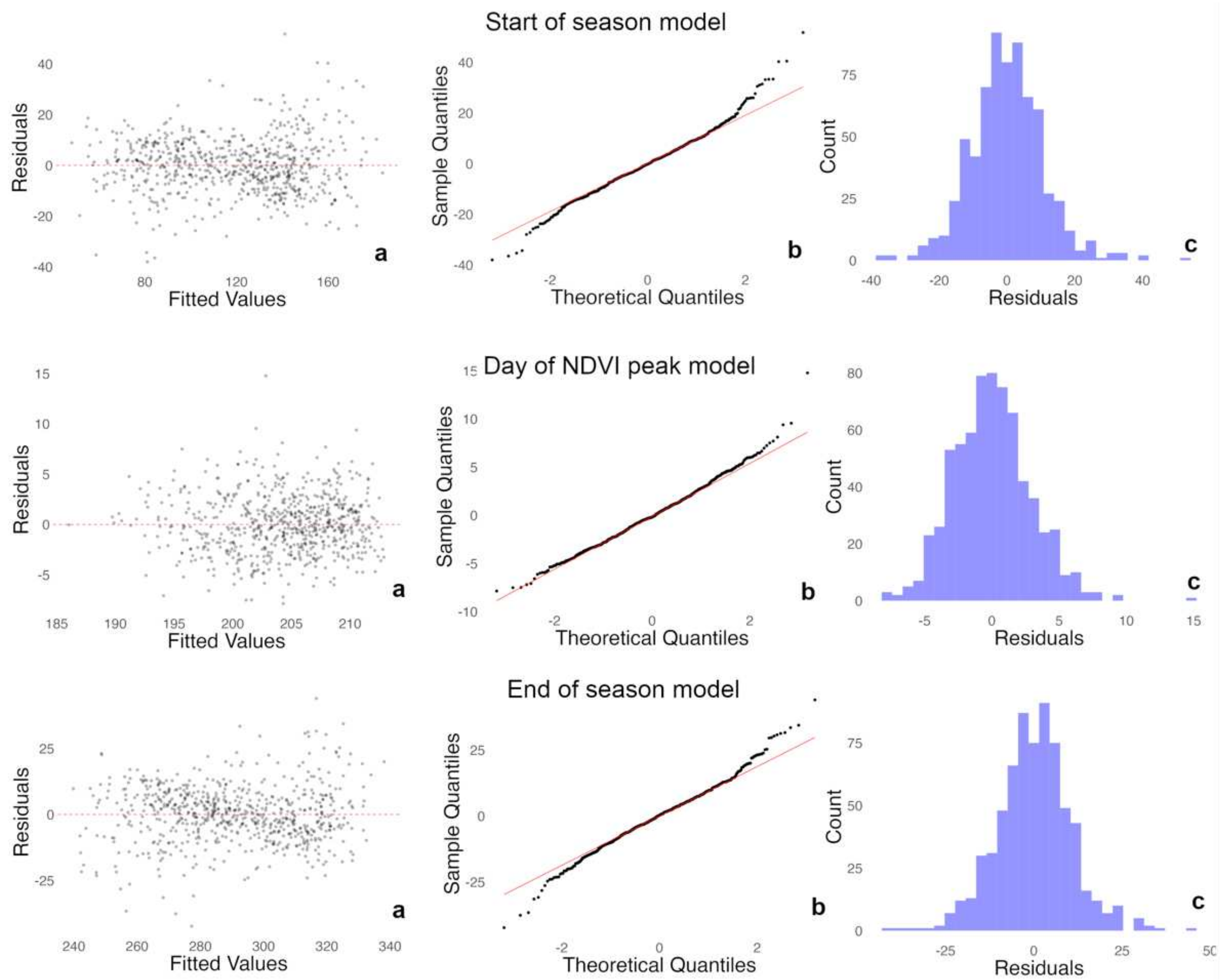


Figure 3.19. From left to right: Residual vs fitted plots (**a**), Quantile-Quantile plots (**b**) and Histogram of residuals distribution (**c**) for Start of season, Day of NDVI peak and End of season GAM models.

3.5 Discussion

The study highlights the significant influence of topographical, meteorological, and edaphic factors on the seasonal growth dynamics of alpine grasslands in North-East Italy. The findings indicate that these environmental drivers shape the magnitude of vegetation productivity and the timing of phenological events. The negative correlation between altitude and both mean and peak NDVI (τ = -0.76 and τ = -0.65) highlights the decrease in vegetation vigor at higher elevations (Figure 3.4a), which also affects NDVI peak (Figure 3.4b) and increases seasonal variability (Figure 3.4c). The complex effects of rising altitude on these ecosystems can be partially explained by lower temperatures, which also cause thicker and more persistent snowfall cover, or higher wind speed, but also include essential contributions by edaphic factors as nutrients availability, soil organic carbon, pH, soil water storage capacity and texture.

The correlation matrix in Figure 3.16 reveals that, in our study area, fields at higher altitude tend to have more acidic soils, with higher sand percentage and significantly lower water holding capacity. In contrast, the soil organic carbon contents seem to be less affected by the elevation gradient and do not influence grassland vegetation growth and phenology. However, such conclusions need further supporting evidence and possibly ground collected data. Similarly to what has been observed by Hua et al., (2022) on the mountain grasslands of New Zealand and by Primi et al., (2016) on Central Italy's Apennines mountain grasslands, increasing altitudes has a shortening effect on the vegetative season of alpine grasslands of our study area, as demonstrated by the correlations shown in Figure 3.5 (a,b,c). Furthermore, high altitude plants are often strongly limited in nutrients, which leads to greater root-to-shoot biomass allocation, which is caused by the necessity to maximise nutrient uptake (Hagedorn et al., 2019).

The latter effects on biomass allocation are also associated with increasing slope gradients, lower vegetation productivity, higher susceptibility to drought events (Måren et al., 2015) and shorter vegetative seasons (An et al., 2018b). Accordingly, in our study area, fields with steeper slopes also negatively correlated with mean and peak NDVI (τ = -0.32 and τ = -0.34), while showing greater seasonal variability (τ = 0.23) and shorter growing seasons (Figure 3.5 d,e,f). Previous findings by Wang et al., (2024) confirmed that steeper slopes support lower vegetation productivity, suggesting increased soil erosion, shallower soils and reduced water retention capacity. Although steeper fields tend to have lower water holding capacity (Figure 3.16), the relationships seem less relevant than those for the elevational gradient in our study area.

The effect of aspect on vegetation communities and growth has been widely investigated and shown effects on plant communities, vegetation structure and growth dynamics (Hua et al., 2022). However, our study was not able to strongly identify consistent effects of aspect on the vegetation of the alpine grasslands in our study area (Figure 3.6), likely due to the limitedness of the dataset. Future research may further explore the effects of aspects on Italian alpine grasslands by incorporating more fields with diverse orientations, expanding the study area, or utilising data with higher spatial resolution to capture more detailed variations.

Meteorological conditions have direct and indirect effects on vegetation growth dynamics. Higher temperatures and solar radiation promote plant growth in our study area, as reflected in the positive correlations between temperature and mean NDVI (τ = 0.42) and NDVI peak (τ = 0.38). However, while temperature enhances biomass production in lower and mid-altitude sites, its role can be less pronounced in higher elevations where soil water limitations and nutrient constraints dominate (Hagedorn et al., 2019) and can negatively affect vegetation in drought-prone areas (Wu et al., 2014). Temperature and radiation also play a significant role in regulating phenological cycles. Warmer conditions advanced the start of the season (τ = -0.39) and peak NDVI timing (τ = -0.23) while delaying the end of the season (τ = 0.41), as shown in Figure 11 (d,e,f). Accordingly, previous studies suggested warmer temperatures could accelerate vegetation greening (Hong et al., 2022), especially in low-temperature areas (Bai et al., 2024), identifying temperature as the primary climatic factor influencing herbaceous budburst (Chen et al., 2015).

Our findings indicate that precipitation has a limited direct influence on vegetation growth variability in the alpine grasslands of our study area. No clear correlation was found between mean NDVI and mean yearly precipitation at the field level (Table 3.6). Despite increasing precipitation variability in European alpine areas, our study area, with precipitation ranging from 760 mm in 2015 to 1734 mm in 2014, has remained resilient to drought events, even during extreme years such as 2003 and 2022. These results align with previous research indicating that precipitation alone is not always the key driver of grassland productivity. Lane et al., (1998) found soil texture to be a stronger determinant of vegetation structure than precipitation in the central U.S. grasslands. Similarly, Chelli et al., (2016) observed that Sub-Mediterranean grasslands maintained stable productivity even under extreme drought conditions. Forte et al., (2023) highlighted that alpine grasslands suffer more from early-season droughts than late-season ones, while Cancellieri et al., (2024) found that edaphic factors, rather than precipitation gradients, drive plant richness and productivity in the mountain grasslands of central Italy. Our study supports the latter hypothesis, as soil water-holding capacity is central in moderating climatic and topographical effects on vegetation growth. Soil water content at field capacity was positively correlated with both mean NDVI (τ = 0.45) and peak NDVI (τ = 0.43), as visible in Figure 3.16 (g,h,i). Fields with higher water retention showed lower seasonal variability and longer growing seasons (Figure 3.17 g,h,i). Wind exposure also interacts with these factors. Stronger winds were correlated with lower NDVI and shorter growing seasons (Figures 3.10 and 3.11, j,k,l), likely due to increased evapotranspiration, soil erosion, and mechanical stress on vegetation. In our region, prevalent wind direction did not affect grassland vegetation (Figure 3.12).

Snowfall is a significant factor influencing vegetation dynamics in mountainous areas; previous studies found that the length of the snow-free period is the primary driver of European mountain grassland biomass production variability (Choler, 2015). In our study, high snowfall amounts affected vegetative growth negatively, lowering the mean (τ=-0.42) and peak (τ=-0.37) NDVI values and enhancing seasonal variability (τ=0.42).

Snow acts as an insulator, regulating the temperature between the soil and air, thereby influencing the heat accumulation necessary for vegetation growth. Chen et al. (2015) identified snowfall as a secondary climatic factor influencing herbaceous budburst, while Wang et al. (2024), found a delaying effect of snow on grassland phenology, as also confirmed by our findings. Other studies found snowmelt timing to be a critical driver, closely aligned with SOS, as snow-free conditions determine the onset of favourable growing conditions, particularly at higher elevations where the growing season is shorter (Xie et al., 2020). The influence of snow cover duration and accumulation on vegetation dynamics is highly complex and varies across different regions. Contrary to our findings, Tomaszewska et al., (2020) observed a positive effect of prolonged snow cover on grassland productivity in the mountain ecosystems of Central Asia. This contrast may stem from the region's greater susceptibility to drought and arid conditions, where an extended snow season helps sustain soil moisture and vegetation growth.

As climate change leads to warmer temperatures, altered precipitation regimes, and reduced snow cover, alpine grasslands may experience shifts in productivity and phenology. Some studies predict that shorter winters and earlier snowmelt may lead to higher biomass production in the Alps (Jonas et al., 2008). However, the extent of this increase will depend on local soil properties, water availability, and nutrient cycling, as demonstrated by the abovementioned studies in Central Asia. The variability of these responses highlights the complexity of climate-vegetation interactions in mountain ecosystems, where biotic and abiotic factors play a crucial role.

Furthermore, the response of plant and soil microbial communities to climate warming is likely to introduce additional feedback that affects ecosystem functioning. As vegetation shifts to higher elevations, it will modify soil properties, influencing carbon and nutrient cycling. Changing temperature and precipitation patterns will also affect soil microbial communities, which are essential for organic matter decomposition and nutrient availability. In many mountain regions, warming is expected to accelerate soil organic matter decomposition, potentially leading to increased carbon emissions. However, the long-term effects on soil carbon storage remain uncertain, as the balance between carbon losses from decomposition and inputs from plant productivity is highly variable across different mountain systems (Hagedorn et al., 2019).

Thus, given the complexity of these interactions, there is a pressing need for models that can enhance our understanding, provide partial explanations, and accurately estimate vegetation productivity and phenology in these fragile and marginal seminatural ecosystems. Such models could serve as valuable tools for policymakers, institutions, and farmers, enabling them to make informed decisions for sustainable management and adaptation strategies in the face of ongoing climate change.

The high predictive accuracy of the GAM models developed in this study highlights their value for ecosystem monitoring, conservation, and climate change adaptation. With R^2 values ranging from 0.686 to 0.900 and low RMSE values across models (e.g. ~11 days for SOS and EOS), the models effectively captured vegetation growth patterns and phenological dynamics, as shown in Figure 3.17. The ability to predict vegetation growth magnitude and phenology with high accuracy has significant practical applications. In alpine pasture management, knowing the timing of peak productivity (R^2 = 0.715, RMSE = 3.150 days for DOY Peak) and seasonal transitions (R^2 = 0.853 for SOS, R^2 = 0.770 for EOS) can inform grazing strategies, preventing overgrazing and optimising forage use. These models also provide valuable tools for conservation planning, enabling land managers to anticipate vegetation shifts in response to climatic fluctuations. As climate change alters precipitation patterns and temperature regimes, predictive models like these will be essential for tracking phenological changes and identifying potential ecosystem vulnerabilities.

3.6 Conclusions

This study comprehensively assesses alpine grassland vegetation dynamics in North-East Italy, integrating 23 years of MODIS NDVI time series with topographical, meteorological, and soil-related variables. The long-term perspective enables a robust evaluation of both short-term fluctuations and long-term vegetation trends, offering critical insights into climate-driven shifts in these semi-natural ecosystems. The findings confirm the dominant role of topographical variables, particularly altitude and slope, in shaping vegetation productivity and phenology. Higher elevations exhibit lower NDVI values, delayed season onset, and shortened growing periods, while steeper slopes further constrain growth by reducing water availability and soil depth. Among climatic drivers, temperature and solar radiation were positively correlated with vegetation vigor and season length. At the same time, precipitation had limited direct influence, reinforcing the importance of soil water retention capacity over total rainfall amounts. Snowfall significantly influenced phenological cycles, delaying the growing season, while wind exposure reduced biomass accumulation through increased evapotranspiration and mechanical stress. The predictive models developed in this study, based on Generalized Additive Models (GAMs), demonstrated high accuracy, with R^2 values between 0.686 and 0.900 (Figure 3.17), effectively estimating vegetation growth patterns and seasonal transitions.

3.6.1 Limitations and perspectives

While the study provides strong insights, several limitations must be acknowledged. The reliance on MODIS NDVI data at 250m resolution may not fully capture fine-scale vegetation heterogeneity, potentially masking local variations caused by microclimatic or management differences at the small scale. Similarly, monthly NDVI time series limits the ability to detect rapid phenological changes, particularly during key transitions in the growing season or disturbing management events.

Although the models performed well statistically, the study lacks extensive field validation, which could further refine estimates of vegetation responses. Also, NDVI saturation remains a concern in dense vegetated areas, even though it may concern the NDVI peak analysis in our study mainly.

The study treats alpine grasslands as a homogeneous unit, potentially overlooking species-specific responses to environmental factors. Incorporating species composition data could enhance predictions and provide a more detailed understanding of vegetation dynamics. Additionally, the analysis does not consider the potential effects of rising atmospheric CO_2 concentrations, which could influence long-term productivity trends, particularly at high elevations. The role of lag effects between climatic drivers and vegetation responses remains an open question, requiring further exploration through time-lagged models.

Future research should improve spatial and temporal resolution by integrating higher-resolution satellite data and refining models to capture species-specific and long-term physiological responses. Expanding ground-based validation efforts and incorporating additional climatic and edaphic variables will enhance the accuracy and applicability of predictive models, strengthening the capacity to monitor, manage, and adapt alpine grassland ecosystems in a changing climate. Additionally, developing a self-powered, permanent IoT-based soil monitoring system, such as the one proposed by Ramson et al., (2021), could provide continuous, high-resolution data on soil moisture, temperature, and nutrient dynamics, offering deeper insights into the influence of environmental drivers on grassland vegetation phenology and further improving predictive modelling efforts.

3.6 References

An, S., Zhang, X., Chen, X., Yan, D., Henebry, G.M., 2018a. An exploration of Terrain effects on land surface phenology across the Qinghai-Tibet Plateau using Landsat ETM+ and OLI data. Remote Sens (Basel) 10. https://doi.org/10.3390/rs10071069

An, S., Zhang, X., Chen, X., Yan, D., Henebry, G.M., 2018b. An exploration of Terrain effects on land surface phenology across the Qinghai-Tibet Plateau using Landsat ETM+ and OLI data. Remote Sens (Basel) 10. https://doi.org/10.3390/rs10071069

Bai, Y., Liu, M., Zhou, J., Guo, Q., Wu, G., Li, S., 2024. Diverse responses of surface biogeophysical parameters to accelerated development and senescence of vegetation on the Mongolian Plateau. Science of The Total Environment 943, 173727. https://doi.org/10.1016/J.SCITOTENV.2024.173727

Brilli, L., Martin, R., Argenti, G., Bassignana, M., Bindi, M., Bonet, R., Choler, P., Cremonese, E., Della Vedova, M., Dibari, C., Filippa, G., Galvagno, M., Leolini, L., Moriondo, M., Piccot, A., Stendardi, L., Targetti, S., Bellocchi, G., 2023. Uncertainties in the adaptation of alpine pastures to climate change based on remote sensing products and modelling. J Environ Manage 336. https://doi.org/10.1016/j.jenvman.2023.117575

Cancellieri, L., Sperandii, M.G., Rosati, L., Bellisario, B., Franceschini, C., Aleffi, M., Bartolucci, F., Becker, T., Belonovskaya, E., Berastegi, A., Dengler, J., Filibeck, G., 2024. Drivers of vascular plant, bryophyte and lichen richness in grasslands along a precipitation gradient (central Apennines, Italy). Journal of Vegetation Science 35. https://doi.org/10.1111/jvs.13305

Chelli, S., Canullo, R., Campetella, G., Schmitt, A.O., Bartha, S., Cervellini, M., Wellstein, C., 2016. The response of sub-Mediterranean grasslands to rainfall variation is influenced by early season precipitation. Appl Veg Sci 19, 611–619. https://doi.org/https://doi.org/10.1111/avsc.12247

Chen, X., An, S., Inouye, D.W., Schwartz, M.D., 2015. Temperature and snowfall trigger alpine vegetation green-up on the world's roof. Glob Chang Biol 21, 3635–3646. https://doi.org/https://doi.org/10.1111/gcb.12954

Choler, P., 2015. Growth response of temperate mountain grasslands to inter-annual variations in snow cover duration. Biogeosciences 12, 3885–3897. https://doi.org/10.5194/bg-12-3885-2015

Dibari, C., Pulina, A., Argenti, G., Aglietti, C., Bindi, M., Moriondo, M., Mula, L., Pasqui, M., Seddaiu, G., Roggero, P.P., 2021. Climate change impacts on the alpine, continental and mediterranean grassland systems of italy: A review. Italian Journal of Agronomy 16. https://doi.org/10.4081/ija.2021.1843

Ditzler, C., Scheffe, K., Monger, H.C., 2017. Soil Survey Manual, in: USDA Handbook 18. Government Printing Office, Washington, D.C.

Forte, T.G.W., Carbognani, M., Chiari, G., Petraglia, A., 2023. Drought Timing Modulates Soil Moisture Thresholds for CO<inf>2</inf> Fluxes and Vegetation Responses in an Experimental Alpine Grassland. Ecosystems 26, 1275–1289. https://doi.org/10.1007/s10021-023-00831-7

Gonzalez-Ollauri, A., Hudek, C., Mickovski, S.B., Viglietti, D., Ceretto, N., Freppaz, M., 2021. Describing the vertical root distribution of alpine plants with simple climate, soil, and plant attributes. Catena (Amst) 203, 105305. https://doi.org/10.1016/J.CATENA.2021.105305

Hagedorn, F., Gavazov, K., Alexander, J.M., 2019. Above- and belowground linkages shape responses of mountain vegetation to climate change. Science (1979) 365, 1119–1123. https://doi.org/10.1126/science.aax4737

Hinojosa, L., Tasser, E., Rüdisser, J., Leitinger, G., Schermer, M., Lambin, E.F., Tappeiner, U., 2019. Geographical heterogeneity in mountain grasslands dynamics in the Austrian-Italian Tyrol region. Applied Geography 106, 50–59. https://doi.org/10.1016/J.APGEOG.2019.03.006

Hong, S., Zhang, Yichen, Yao, Y., Meng, F., Zhao, Q., Zhang, Yao, 2022. Contrasting temperature effects on the velocity of early- versus late-stage vegetation green-up in the Northern Hemisphere. Glob Chang Biol 28, 6961–6972. https://doi.org/https://doi.org/10.1111/gcb.16414

Hua, X., Ohlemüller, R., Sirguey, P., 2022. Differential effects of topography on the timing of the growing season in mountainous grassland ecosystems. Environmental Advances 8, 100234. https://doi.org/10.1016/J.ENVADV.2022.100234

Jonas, T., Rixen, C., Sturm, M., Stoeckli, V., 2008. How alpine plant growth is linked to snow cover and climate variability. J Geophys Res Biogeosci 113. https://doi.org/10.1029/2007JG000680

Lane, D.R., Coffin, D.P., Lauenroth, W.K., 1998. Effects of soil texture and precipitation on above-ground net primary productivity and vegetation structure across the Central Grassland region of the United States. Journal of Vegetation Science 9, 239–250. https://doi.org/https://doi.org/10.2307/3237123

Liu, Y., Fan, Y., Yue, J., Jin, X., Ma, Y., Chen, R., Bian, M., Yang, G., Feng, H., 2024. A model suitable for estimating above-ground biomass of potatoes at different regional levels. Comput Electron Agric 222, 109081. https://doi.org/10.1016/J.COMPAG.2024.109081

Ma, X.Q., Leng, P., Liao, Q.Y., Geng, Y.J., Zhang, X., Shang, G.F., Song, X., Song, Q., Li, Z.L., 2022. Prediction of vegetation phenology with atmospheric reanalysis over semiarid grasslands in Inner Mongolia. Science of The Total Environment 812, 152462. https://doi.org/10.1016/J.SCITOTENV.2021.152462

Måren, I.E., Karki, S., Prajapati, C., Yadav, R.K., Shrestha, B.B., 2015. Facing north or south: Does slope aspect impact forest stand characteristics and soil properties in a semiarid trans-Himalayan valley? J Arid Environ 121, 112–123. https://doi.org/10.1016/J.JARIDENV.2015.06.004

Ogungbuyi, M.G., Guerschman, J., Fischer, A.M., Crabbe, R.A., Ara, I., Mohammed, C., Scarth, P., Tickle, P., Whitehead, J., Harrison, M.T., 2024. Improvement of pasture biomass modelling using high-resolution satellite imagery and machine learning. J Environ Manage 356, 120564. https://doi.org/10.1016/J.JENVMAN.2024.120564

Orlandi, S., Probo, M., Sitzia, T., Trentanovi, G., Garbarino, M., Lombardi, G., Lonati, M., 2016. Environmental and land use determinants of grassland patch diversity in the western and eastern Alps under agro-pastoral abandonment. Biodivers Conserv 25, 275–293. https://doi.org/10.1007/s10531-016-1046-5

Primi, R., Filibeck, G., Amici, A., Bückle, C., Cancellieri, L., Di Filippo, A., Gentile, C., Guglielmino, A., Latini, R., Mancini, L.D., Ronchi, B., Piovesan, G., 2016. From Landsat to leafhoppers: A multidisciplinary approach for sustainable stocking assessment and ecological monitoring in mountain grasslands. Agric Ecosyst Environ 234, 118–133. https://doi.org/10.1016/j.agee.2016.04.028

Ramson, S.R.J., Leon-Salas, W.D., Brecheisen, Z., Foster, E.J., Johnston, C.T., Schulze, D.G., Filley, T., Rahimi, R., Soto, M.J.C.V., Bolivar, J.A.L., Bolivar, J.A.L., Malaga, M.P., 2021. A Self-Powered, Real-Time, LoRaWAN IoT-Based Soil Health Monitoring System. IEEE Internet Things J 8, 9278–9293. https://doi.org/10.1109/JIOT.2021.3056586

Sundqvist, M.K., Sanders, N.J., Wardle, D.A., 2013. Community and ecosystem responses to elevational gradients: Processes, mechanisms, and insights for global change. Annu Rev Ecol Evol Syst. https://doi.org/10.1146/annurev-ecolsys-110512-135750

Tennant, C.J., Harpold, A.A., Lohse, K.A., Godsey, S.E., Crosby, B.T., Larsen, L.G., Brooks, P.D., Van Kirk, R.W., Glenn, N.F., 2017. Regional sensitivities of seasonal snowpack to elevation, aspect, and vegetation cover in western North America. Water Resour Res 53, 6908–6926. https://doi.org/https://doi.org/10.1002/2016WR019374

Thomson, A., Jacobs, J., Morse-McNabb, E., 2023. Comparing the predictive ability of Sentinel-2 multispectral imagery and a proximal hyperspectral sensor for the estimation of pasture nutritive characteristics in an intensive rotational grazing system. Comput Electron Agric 214, 108275. https://doi.org/10.1016/J.COMPAG.2023.108275

Tomaszewska, M.A., Nguyen, L.H., Henebry, G.M., 2020. Land surface phenology in the highland pastures of montane Central Asia: Interactions with snow cover seasonality and terrain characteristics. Remote Sens Environ 240, 111675. https://doi.org/10.1016/J.RSE.2020.111675

Wang, W., Straffelini, E., Tarolli, P., 2024. 44% of steep slope cropland in Europe vulnerable to drought. Geography and Sustainability 5, 89–95. https://doi.org/10.1016/J.GEOSUS.2023.12.001

Wang, X., Li, Z., Xiao, J., Zhu, G., Tan, J., Zhang, Y., Ge, Y., Che, T., 2024. Snow cover duration delays spring green-up in the northern hemisphere the most for grasslands. Agric For Meteorol 355. https://doi.org/10.1016/j.agrformet.2024.110130

Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecol Monogr 30, 279–338. https://doi.org/https://doi.org/10.2307/1943563

Wu, D., Zhao, X., Zhao, W., Tang, B., Xu, W., 2014. Response of vegetation to temperature, precipitation and solar radiation time-scales: A case study over mainland Australia, in: International Geoscience and Remote Sensing Symposium (IGARSS). pp. 855–858. https://doi.org/10.1109/IGARSS.2014.6946559

Xie, J., Jonas, T., Rixen, C., de Jong, R., Garonna, I., Notarnicola, C., Asam, S., Schaepman, M.E., Kneubühler, M., 2020. Land surface phenology and greenness in Alpine grasslands driven by seasonal snow and meteorological factors. Science of the Total Environment 725. https://doi.org/10.1016/j.scitotenv.2020.138380

4. Optimising Grassland Above-Ground Biomass Estimation For Managed Grasslands: A Gaussian Process Regression Approach for Sentinel-2 and PlanetScope in Northern Italy

4.1 Abstract

Accurate and regular estimation of above-ground biomass (AGB) in grassland ecosystems is critical for sustainable grazing management, feed planning, and carbon monitoring. Recent advances in satellite remote sensing, particularly the increasing availability of high-resolution multispectral imagery with frequent revisit times and cloud-based processing tools, have enabled the scalable and timely estimation of biomass across diverse landscapes. This research investigates the use of Gaussian Process Regression (GPR) models, combined with multispectral satellite imageries from Sentinel-2 and PlanetScope, to predict Above-Ground Biomass (AGB) across three representative managed grassland systems in Northern Italy. Extensive manual AGB measurements ($n = 954$) were collected across three sites over 18 months, encompassing diverse topography and management regimes representatives of the region, including permanent meadows, lowland pastures, and alpine grasslands. At the pixel level, the model achieved an R^2 of 0.520 for Sentinel-2 and 0.514 for PlanetScope, with a mean absolute error (MAE) of ~ 400 kg/ha for both platforms. This is a common finding in heterogeneous grassland environments due to inherent spatial variability. However, averaging predictions at the field level significantly improved the accuracy, yielding $R^2 = 0.972$ (Sentinel-2) and 0.968 (PlanetScope) and MAE ≈ 60–120 kg DM ha^{-1} (≤ 10% relative error) for within-region applications, while pixel-level estimates remained affected by high local variability (CV ~30-32%). When aggregated at the field scale, our accuracy is comparable to that of current commercial pasture monitoring platforms, supporting the operational suitability of this approach for farm-scale decision support systems (DSS) and providing a solid basis for integration into measurement, reporting, and verification (MRV) protocols for carbon credits. As the model was calibrated over a limited temporal window, multi-year verification under varying climatic conditions will be required to confirm its temporal stability and operational reliability on the long-term.

4.2 Introduction

Semi-natural grasslands, including permanent pastures and meadows, are ecologically rich and multifunctional ecosystems that contribute to biodiversity conservation, soil stability, carbon sequestration, and landscape resilience (Prangel et al., 2024b). Managed grasslands, which cover about 70% of the world's agricultural area (FAO, 2023), offer substantial potential for climate mitigation. Notably, improved grassland management could enhance their capacity to act as carbon sinks, with sequestration potentials estimated at up to 699 megatons of CO_2eq annually (Bai and Cotrufo, 2022; Pelton et al., 2024). In addition to their ecological value, these ecosystems play a critical role in agricultural systems, providing forage for livestock and serving as the foundation for extensive grazing practices. Ruminant livestock production impacts Europe's economy, society, and environment (EU Commission, 2020), one of the leading global livestock production regions (FAOSTAT, 2023). According to the grazing livestock dataset by Malek et al., (2024), the North-East Italy region (NUTS 1=ITH) hosts more than 1 million grazing cattle livestock units (LSU)**.** Within this region, the province of Vicenza (NUTS 3=ITH32), which includes the study area of this research, accounts for around 66,500 cattle LSU, highlighting the importance of grass-based systems in the local landscape. As consumer awareness around sustainability grows, market preferences are shifting toward pasture-raised cattle, positioning permanent grasslands as key actors in agricultural production and ecological stewardship (Schulze et al., 2021).

Effective and timely above-ground biomass (AGB) monitoring tools and methods are required to optimise the management of these systems, reduce the risks of overgrazing and land degradation, and allow for innovative approaches such as adaptive grazing management. Traditional and modern field-based methods include cut-and-dry destructive sampling, visual assessment coupled with decision-support systems (Hanrahan et al., 2017), and rising/falling plate meters (McSweeney et al., 2022). Additionally, proximal sensing tools such as optical sensors and spectroradiometers have strong potential for forage characterisation (R.R. Pullanagari et al., 2012). However, these methods present essential limitations. They are often time-consuming, labour-intensive, and lack scalability, making them impractical for monitoring multiple fields and large geographic areas, and for deployment in practical tools by end-users. Despite these limitations, recent research highlights the complementary role of proximal sensing, providing valuable high-resolution ground truth data to train and validate satellite-based models. Unmanned Aerial Vehicle (UAV) based remote sensing has demonstrated its validity for estimating grassland biomass in combination with other sensing technologies. A wide range of sensing technologies, including multispectral (De Rosa et al., 2021), hyperspectral (Li et al., 2020), LiDAR (Yang et al., 2025), and structure-from-motion photogrammetric surveys (Sinde-González et al., 2021), have been successfully deployed. Nonetheless, the adoption of these technologies remains limited due to high cost, along with substantial expertise for data acquisition and processing, factors that hinder their widespread use among end-users. Moreover, the time required for data collection and analysis often makes these methods unsuitable for rapid, day-to-day management decisions.

The growing availability of public and commercial satellite platforms offering radar, multispectral, and hyperspectral data at high spatial resolution with frequent revisit times, combined with advanced cloud computing technologies that enable near real-time data processing and analysis, positions satellite-based remote sensing as a highly effective and scalable solution for the cost-efficient monitoring of grasslands (Pinna et al., 2024). A growing body of research has focused specifically on meadow management, demonstrating the feasibility of detecting mowing events using satellite data. In particular, the synergistic use of multispectral (e.g., Sentinel-2, Landsat 8) and radar-based features (e.g., Sentinel-1) has shown high accuracy in identifying both the timing and frequency of cuts (Lobert et al., 2021). However, most studies continue to treat permanent grasslands uniformly, without distinguishing whether the vegetation is grazed or cut, and instead focus more broadly on assessing vegetation characteristics such as above-ground biomass or canopy height. Since the early 2000s, multispectral satellite platforms such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Advanced Very-High Resolution Radiometer (AVHRR) have been successfully deployed for grassland biomass estimation (Hill et al., 2004). However, the coarse spatial resolution of these sensors (500m - 1 km) strongly limits their application. This is particularly relevant in areas where the agricultural landscape is highly fragmented and field sizes are tiny, such as Europe, and even more so in the Alpine region and Italy. The launch of the Sentinel-2 platform by ESA in 2015, with a spatial resolution of 10–20 meters and a 5-day revisit cycle, alongside the growing availability of both public and commercial satellite constellations such as PlanetScope, which provides daily imagery at 3-meter resolution, has significantly expanded the data available for developing grassland biomass monitoring tools and methods at the field and paddock level (Pinna et al., 2024).

Although some authors, such as Punalekar et al. (2018), have successfully deployed the PROSAIL radiative transfer model (RTM) to estimate grassland biomass, most previous studies have relied on vegetation indices, often coupled with advanced statistical modelling and machine learning algorithms. For instance, Qin et al. (2021) evaluated the Sentinel-2 and Landsat 8 vegetation indices to estimate fresh above-ground biomass (AGB) in Inner Mongolian grasslands at the field level. They applied a linear mixed-effects model on the Sentinel-2-derived Normalised Difference Phenology Index (NDPI) and achieved an $R^2 = 0.62$. In Florida, Smith et al. (2023) tested the efficacy of multiple Sentinel-2 spectral bands and vegetation indices, combined with support vector regression (SVR), in estimating grassland biomass and nitrogen content at the pixel level, yielding a maximum accuracy of R^2 = 0.73. Given the ability of machine learning algorithms to process high-dimensional inputs and capture complex non-linear relationships, an increasing number of studies now utilise the complete set of spectral bands to enhance model accuracy and capture subtle spectral information. Dusseux et al. (2022) compared several algorithms to estimate canopy height on French grasslands, achieving an R^2 of 0.56 with a Ridge Regression model. Chen et al. (2021) applied a multilayer perceptron neural network, combining Sentinel-2 and climate data, to monitor above-ground biomass (AGB) at the field level in five Tasmanian dairy farms, achieving an R^2 of 0.60 and an RMSE of 356 kg DM/ha. More recently, Ogungbuyi et al. (2024) developed the Spectral Pasture Estimation using Combined Techniques of Random-forest Algorithm for Features Optimisation and Retrieval (SPECTRA-FOR) framework using random forest models and data from Sentinel-2 and Planet SuperDove, achieving an R^2 of up to 0.92 and an RMSE of 346 kg DM/ha at the farm level across ten different Australian livestock farms.

Among machine learning approaches, Gaussian Process Regression (GPR) has emerged as a compelling method for modelling complex and nonlinear patterns, thanks to its probabilistic framework and ability to provide uncertainty estimates. Its usefulness has been demonstrated across diverse research fields, including the mapping of mangrove functional traits from Sentinel-2 imagery using hybrid active learning strategies (Jia et al., 2024), price prediction in agricultural markets (Jin and Xu, 2025),and energy consumption forecasting for electric vehicles (Hussain et al., 2025).

This study investigates the potential of Sentinel-2 and PlanetScope satellite imagery, combined with Gaussian Process Regression (GPR), for estimating above-ground biomass (AGB) across diverse and heterogeneous grassland systems in Northern Italy. Specifically, this pursues four key objectives:

(i) Evaluate the performance of GPR models in estimating Above-Ground Biomass (AGB) at both the pixel and field levels.

(ii) Compare the model performance based on inputs from Sentinel-2 and PlanetScope, to evaluate their respective advantages and limitations.

(iii) Examine the practical suitability and readiness of this modelling approach for farm-scale implementation under real-world conditions.

(iv) Identify areas for improvement and explore how this approach could be refined to enhance its accuracy, scalability, and operational utility.

This study advances grassland AGB mapping by applying Gaussian Process Regression (GPR) to heterogeneous European systems (meadows, lowland pastures, alpine pastures), a context where GPR has been little explored. Leveraging Sentinel-2 and PlanetScope with an extensive ground dataset (n = 954), we evaluate performance at pixel and field scales, showing that field-level aggregation attains operational accuracy for feed budgeting and adaptive grazing. We also benchmark against prior studies and commercial platforms, highlighting practical readiness in fragmented European landscapes.

4.3 Materials and Methods

4.3.1 Study Area

The sites selected for this study are in northeastern Italy and reflect three representative examples of the region's typical management practices adopted by graze-fed cattle farms. Two of the sites belong to the same farm based in Schio (45.68°N, 11.36°E), situated in a hilly transition zone where the Po Valley plains meet the foothills of the Alps, between 170 and 230 m.a.s.l. (Fig. 1).

The mean annual temperature in the Schio area is 13.8°C, and the average annual rainfall is 1303 mm (1994-2023 ARPAV). The first site is a 3.29 ha permanent meadow, primarily flat and mowed twice a year for haymaking purposes. It is occasionally fertilised with organic manure or slurry.

The botanical composition is dominated by *Lolium multiflorum*, *Lolium arundinaceum* (=*Festuca arundinacea*), *Lolium perenne*, and several other species typical of lowland meadows. The second site is a 2.97 ha permanent pasture with an average slope of 10°, used for grazing from March to November. It is subdivided into eight paddocks. Each is grazed in rotation by 12 beef cattle for three days before being moved to the next section. The pasture hosts a more diverse botanical composition, with many species belonging to the *Poaceae* family, including Cy*nosurus cristatus*, *Lolium perenne, Poa trivialis*, *Anthoxanthum odoratum, Agrostis stolonifera,* and *Festuca pratensis*. Species belonging to *Fabaceae* were *Trifolium repens*, *Trifolium pratense* and *Lotus corniculatus*. Other species were present, such as *Carex hirta*, *Taraxacum officinale*, *Agrimonia eupatoria*, *Daucus carota* and *Ranunculus bulbosus*. Overall, a high number of species were detected.

The third site is a 9.8 ha mountain pasture in Caltrano (45.80°n, 11.46°e), located on the nearby Asiago Plateau, at an elevation of approximately 1200 m.a.s.l., characterised by a highly irregular surface and an average slope of 13% (Fig. 2). During the summer of 2023 (June-September), it served as the farm’s alpine grazing area and was grazed by 40 beef cattle; however, it remained ungrazed in 2024. The mean annual temperature in this area is 6.29°c, and the annual rainfall is 1375 mm (ARPAV, 1994-2023 series data). The botanical composition of the flatter areas was characterised by *Phleum pratense*, *Festuca rubra*, *Agrostis stolonifera,* and *Deschampsia caespitosa,* all belonging to the *Poaceae* family, as well as *Trifolium pratense* and *Trifolium repens,* both belonging to the *Fabaceae* family. In these areas, given the high presence of nutrients, nitrophilous species had also been recorded, as well as *Trollius europaeus* and *Ranunculus acris.* In the steeper areas, the vegetation was characterised by the dominance of *Brachypodium rupestre*, *F. rubra*, *Achillea millefolium* and *Rhinanthus alectoroforus*.

The Schio sites were overlaid with a grid of points spaced 20 meters apart, resulting in 75 sampling points on the meadow and 78 sampling points on the pasture. Due to the larger size of the mountain pasture, 114 sampling points were randomly placed on the third site to extend the variability of the environment. The chosen sampling density across all sites aimed to balance spatial representativeness with the logistical feasibility of field data collection.

Accurate sampling location coordinates were estimated in GPX format and imported into a Garmin Montana 750i navigator, which supports both GPS and GALILEO constellations. Thus, the device often reached sub-meter accuracy in the Schio area and 2-3 meters accuracy in the Caltrano area.

4.3.2 Ground Data Collection

Starting from May 2023, four surveys have been conducted at the lowland farm in Schio (VI) and three at the mountain pasture in Caltrano (VI) to assess the above-ground biomass (AGB) content. Both the Schio meadow and pasture areas were surveyed on 23rd of May 2023, the 9th of October 2023, and the 4th of November 2024. Additionally, the meadow was surveyed on the 6th of May 2024, while the pasture was on the 20th of June 2024. The mountain pasture in Caltrano was surveyed on the 22nd of July 2023, the 28th of October 2023 and the 27th of July 2024.

At each point and date, the AGB of a 1x1m square sample area was cut and bagged, then oven-dried at 105°C for 48h to determine dry matter content (g/m^2) according to the widespread cut and dry method, for a total of 954 AGB points.

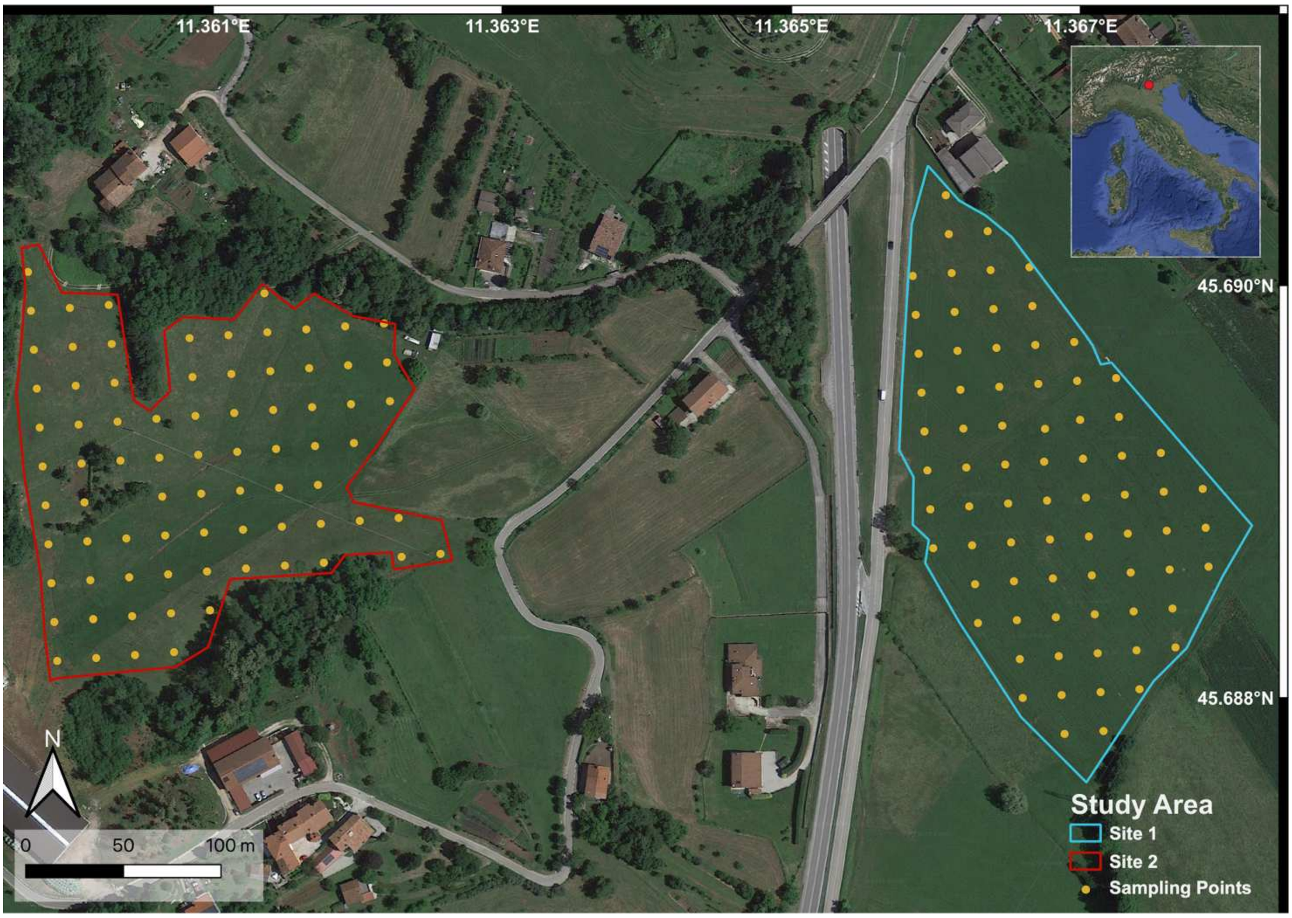


Figure 4.1 Site 1 (permanent meadow - blue line) and 2 (pasture - red line) in Schio (VI), 45.68N, 11.36E, between 170 and 230 m.a.s.l.

Figure 4.2 Site 3, mountain pasture in Caltrano (VI), 45.80N, 11.46E, ~ 1200 m.a.s.l.

4.3.3 Satellite Data Collection and Processing

For each survey date, multispectral images were collected from the Sentinel-2 and PlanetScope SuperDove sources. If a cloud-free image of the study area was not available on the exact date, the image captured at the closest to the survey date was used. A cloud-free image was available within a 10-day interval was considered for both satellite sources on all survey dates.

Sentinel-2 Level 1C images (Top-of-atmosphere reflectance) were downloaded from the Copernicus Data Space browser (https://browser.dataspace.copernicus.eu/) and subsequently processed with the Multi-Temporal Cloud Screening and Atmospheric Correction Software (MAJA), which is an advanced atmospheric correction software developed by CESBIO and CNES (Hagolle et al., 2017), to produce a Level 2A Surface Reflectance product. In this process, Band 1 (Coastal aerosol), Band 9 (Water Vapour) and Band 10 (Cirrus clouds) are used for atmospheric correction only and thus are eliminated from the final product, as shown in Table 4.1.

PlanetScope data were accessed and downloaded through the Planet Explorer portal (https://www.planet.com/explorer/) under an educational and research license granted by Planet to the authors. Level 3B surface reflectance products were clipped on the study area and downloaded through the portal. As with Sentinel-2, the PlanetScope Band 1 (Coastal aerosol) has been excluded from the analysis. Table 4.1 reports the selected spectral bands from both sources, highlighting their spectral characteristics and spatial resolution.

To sample spectral reflectance values from the collected series of Sentinel-2 and PlanetScope raster images on the sampling points, a Python (v. 3.10.8) script was developed using geopandas (v. 0.13.2), Rasterio (v. 1.3.8), Pandas (v. 1.5.3) and Glob2 (v. 0.7) packages.

Table 4.1. Selected bands from Sentinel-2 and PlanetScope platforms with spatial and spectral characteristics.

Satellite platform	Band	Spatial resolution (m)	Central wavelength (nm)	Bandwith (nm)	Data volume per Hectar (kB/ha)
Sentinel-2	B2 - Blue	10	490	65	0.838
	B3 - Green		560	35	
	B4 - Red		665	30	
	B5 - Red edge	20	705	15	
	B6 - Red edge II		740	15	
	B7 - Red edge III		783	20	
	B8 - NIR	10	842	115	
	B8A - NIR II	20	865	20	
	B11 - SWIR I		1610	90	
	B12 - SWIR II		2190	180	
PlanetScope Super Dove	B2 - Blue	3	490	50	8.89
	B3 – Green I		531	36	
	B4 - Green		565	36	
	B5 - Yellow		610	20	
	B6 - Red		665	31	
	B7 – Red edge		705	15	
	B8 - NIR		865	40	

4.3.4 Data Preparation and Preprocessing

The dataset containing AGB data, Sentinel-2 and PlanetScope imagery spectral bands was processed in R (v. 4.4.2), and any missing or invalid values were filtered out. To ensure a balanced representation across the AGB range, the dataset was stratified into quantile-based classes, and an 80/20 train-test split was performed using *createDataPartition()* from the *caret* package (v. 7.0.1). The characteristics of the training and test sets, including sample size, mean AGB values, and standard deviations, are summarised in Table 4.2. The predictor variables, corresponding to the spectral bands, were standardised by centring and scaling using *preProcess()* function from the *caret* package. The same scaling parameters were applied to the training and test datasets to maintain consistency.

Table 4.2 Summary statistics of the training and test sets after data preparation and preprocessing.

	Training	**Test**
N. of samples	763	191
Mean AGB (kg/ha)	1277.2	1261.2
Standard Deviation	822.6	785.1

4.3.5 Model Comparison and Selection

To evaluate the most suitable modelling approach for grassland biomass estimation, we compared four commonly adopted machine learning algorithms: Random Forest (RF), Support Vector Regression (SVR), multilayer perceptron neural networks (MLP), Lasso Regression, and Gaussian Process Regression (GPR). All models were trained on the same dataset following the previously described preprocessing steps and hyperparameter tuning by cross-validation. Based on this benchmarking process, GPR was selected as the final modelling approach because it provided the best trade-off among predictive accuracy, robustness, and the ability to quantify predictive uncertainty. Although the improvement in accuracy compared to RF was modest, the probabilistic framework of GPR offers a unique advantage by providing explicit uncertainty estimates, which are crucial for assessing model reliability, analysing the spatial structure and distribution of the model's errors and for future applications in environmental and agricultural management. While GPR is recognised to have scalability limitations for very large datasets due to its computational complexity, in this study (n = 954 observations), training and prediction were efficient and tractable. The computational performance metrics reported in Table 3 further demonstrate the suitability of GPR for this application, confirming that the model is suitable for farm and regional scale applications.

4.3.6 Gaussian Process Regressor Calibration, Training and Evaluation

Separate Gaussian Process Regression (GPR) models were trained for the Sentinel-2 and PlanetScope datasets to account for their distinct spectral properties. Prior to modelling, the dataset was cleaned by removing missing values and extreme outliers through an IQR-like trimming procedure based on the 5th–95th percentiles of biomass values. To ensure balanced representation across the full biomass range, data were stratified into quantile-based classes, and an 80/20 train–test split was applied. During model development, a range of kernel functions and their combinations were tested, including Matern (3/2 and 5/2), Exponential, Powered Exponential, Rational Quadratic, Triangular, Cubic, and Gaussian. Among these, the additive combination of Matern 5/2 and Exponential kernels consistently provided the best trade-off between accuracy and generalisation. Hyperparameters were optimised using cross-validation, and the nugget parameter was tuned across a grid of values (10^{-8}–10^{-4}), with 10^{-6} selected as optimal. Final model fitting was carried out with the gpkm() function from the GauPro package (v.0.2.14), and predictions on unseen test data were obtained using the predict() function, which outputs both mean predictions and posterior standard errors. Confidence intervals at the 95% level were computed as in equation 1, where γ is the predicted value and SE the posterior standard error.:

$$CI_{95} = \gamma \pm 1.96 \times SE \tag{1}$$

Model performance was evaluated using both a single 80/20 split and repeated stratified resampling (20 replicates), ensuring robustness to data partitioning. Additionally, bootstrap resampling with 10,000 iterations was employed to compute robust estimates of R^2, RMSE, and MAE, with 95% confidence intervals derived from the empirical distributions. In addition to evaluating mean performance metrics, we assessed the calibration of predictive uncertainty. A conformal prediction framework (Shafer and Vovk, 2008) has been adopted, which provides distribution-free predictive intervals with valid coverage guarantees. Specifically, we computed absolute standardised residuals from out-of-fold predictions obtained via repeated stratified splits, following the principles of cross-conformal prediction (Vovk, 2015). The out-of-fold residuals used for conformal calibration were obtained from the same 20 repeated stratified 80/20 splits employed in the robustness analysis, ensuring a common resampling design for both performance and uncertainty evaluation. The empirical 95th quantile of these residuals $q_{0.95}$ was then used to rescale the posterior standard errors, yielding calibrated intervals, as shown in equation (2), where $PI_{0.95}(x)$is the calibrated prediction interval, $\hat{y}(x)$ is the model prediction for input x, and $SE(x)$ is the posterior standard error.

$$PI_{0.95}(x) = \hat{y}(x) \pm q_{0.95} \times SE(x) \tag{2}$$

From these calibrated intervals, we derived four metrics to characterise predictive uncertainty at both pixel and field scales:

1. Coverage: the proportion of observed values falling inside the 95% prediction interval (ideal ≈ 0.95).
2. Interval width: the absolute span of the 95% interval (kg DM ha^{-1}), measuring the typical range of uncertainty.
3. Relative width: the interval width expressed as a fraction of the predicted mean, making uncertainty comparable across biomass levels.
4. Coefficient of variation (CV%): the posterior standard error divided by the predicted mean, expressed as a percentage, summarising relative precision.

To assess the relative contribution of each spectral band to above-ground biomass (AGB) prediction, we employed a SHAP framework (Lundberg and Lee, 2017) using the *fastshap* package (v.0.1.0). SHAP values quantify the marginal contribution of each predictor to model output, based on coalitional game theory, and provide a consistent measure of feature importance. For each model, a prediction wrapper was defined to generate numeric outputs from the trained GPR. Feature attribution was then computed on the independent test set to avoid training leakage, with a random subset of 200 samples from the training set used as the background dataset to stabilise conditional expectations. SHAP values were estimated through a Monte Carlo approximation with 400 simulations per feature, balancing stability and computational efficiency. Global feature importance was derived as the mean absolute SHAP value across all test observations for each band, providing a measure of its average contribution to predictions. To facilitate comparison, these values were normalised to percentages of the total SHAP magnitude, allowing the relative influence of spectral regions to be ranked.

A comprehensive residual analysis was conducted to evaluate model performance further. Residual distribution and structure were examined using residuals vs predicted, and QQ plots with *ggplot2* (v 3.5.1) and *qqplotr* (v 0.0.3), while normality and heteroscedasticity were tested via the Shapiro–Wilk and Breusch–Pagan tests from *lmtest* (v 0.9.4).

Spatial clustering of errors was assessed using Moran's I on the independent test set. Standardised residuals were defined as equation 3, where r_i is the standardised residual, y_i is the observed value at survey point i, $\hat{y}_i$ is the model prediction at point i and SE_i is the model's standard error for the given prediction.

$$r_i = (y_i - \hat{y}_i)/SE_i \quad (3)$$

Because multiple measurements were collected at the same survey point on different dates, residuals were aggregated to a single value per survey point using a precision-weighted mean, where weights w_i are calculated as in equation 4, with SE_i being the model's standard error for the given prediction.

$$w_i = 1/SE_i^2 \quad (4)$$

A k-nearest-neighbours spatial weights matrix (k=8) was constructed from the survey-point coordinates, and statistical significance was evaluated by permutation (9,999 permutations), implemented with the R package *spdep*. The overall workflow of the study, from site selection to model evaluation, is summarised in Fig. 2.

4.3.6 Deployment of GPR models to generate maps

The trained GPR models were operationalised on both Sentinel-2 and PlanetScope multispectral stacks to produce spatial maps of biomass and model uncertainty. For each scene, the acquisition date was parsed from the filename and the corresponding field polygon (ESRI Shapefile) was loaded. Raster stacks were read, clipped, and masked to the field boundary; band names were harmonised to match the models' expected inputs. Per-pixel reflectance with missing values were excluded, and the same processing saved from training was applied to ensure consistency. The fitted GPR models were then queried with predict so that, for each valid pixel, both the predicted AGB (kg DM ha^{-1}) and the associated posterior standard error were obtained. Outputs were written as two-band GeoTIFFs (Band 1: DM_mean, Band 2: DM_SE) using LZW compression, preserving the original CRS and field geometry. This identical deployment workflow across the two sensors guarantees comparability of biomass estimates and their uncertainty at the pixel scale.

4.3.7 Computational Performance and Scalability

To quantitatively assess the computational requirements of the Gaussian Process Regression (GPR) models, we benchmarked both the training and inference phases for Sentinel-2 and PlanetScope datasets on a standard workstation (Intel Core i5 2.3 GHz, 8 GB RAM, macOS 13.7). The analysis quantified training time, model memory footprint, peak RAM usage, and inference throughput, providing estimates of the maximum area that can be processed under typical hardware conditions (Table 3). Model training was timed using the *bench* R package, and peak memory usage was recorded with the *peakRAM* library, which tracks the maximum resident memory during model fitting. Inference throughput was computed as the number of pixels predicted per second and converted to hectares per hour based on each sensor's spatial resolution (10 m for Sentinel-2, 3 m for PlanetScope). From these values, we derived the estimated time required to process 100 ha (typical farm scale) and 10,000 ha (regional scale).

Training required approximately 2 minutes for both models (Table 4.3) with a modest peak memory usage of ~60 MB, confirming that datasets of this scale (≈ 1,000 samples) are well within the computational capabilities of standard desktop hardware. Inference efficiency was high, reaching 6,300–8,200-pixel predictions per second, corresponding to 20,000–26,600 ha per hour.

Consequently, a 100-ha area can be processed in under 20 seconds, and a 10,000-ha region in approximately 25–30 minutes, without parallelisation or GPU acceleration.

Table 4.3 Computational benchmarking of Gaussian Process Regression (GPR) models trained on Sentinel-2 and PlanetScope datasets.

Metric	**Sentinel-2**	**PlanetScope**
Training Time (s)	125.5	86.7
Model size (MB)	11.557	11.517
Peak RAM used (MB)	58.8	58.1
Processed pixels per second	6291	8210
Processed ha per hour	20382	26600
Time to process 100ha (s)	18	14
Time to process 10000 ha (min)	29	23

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
Università degli Studi di Padova

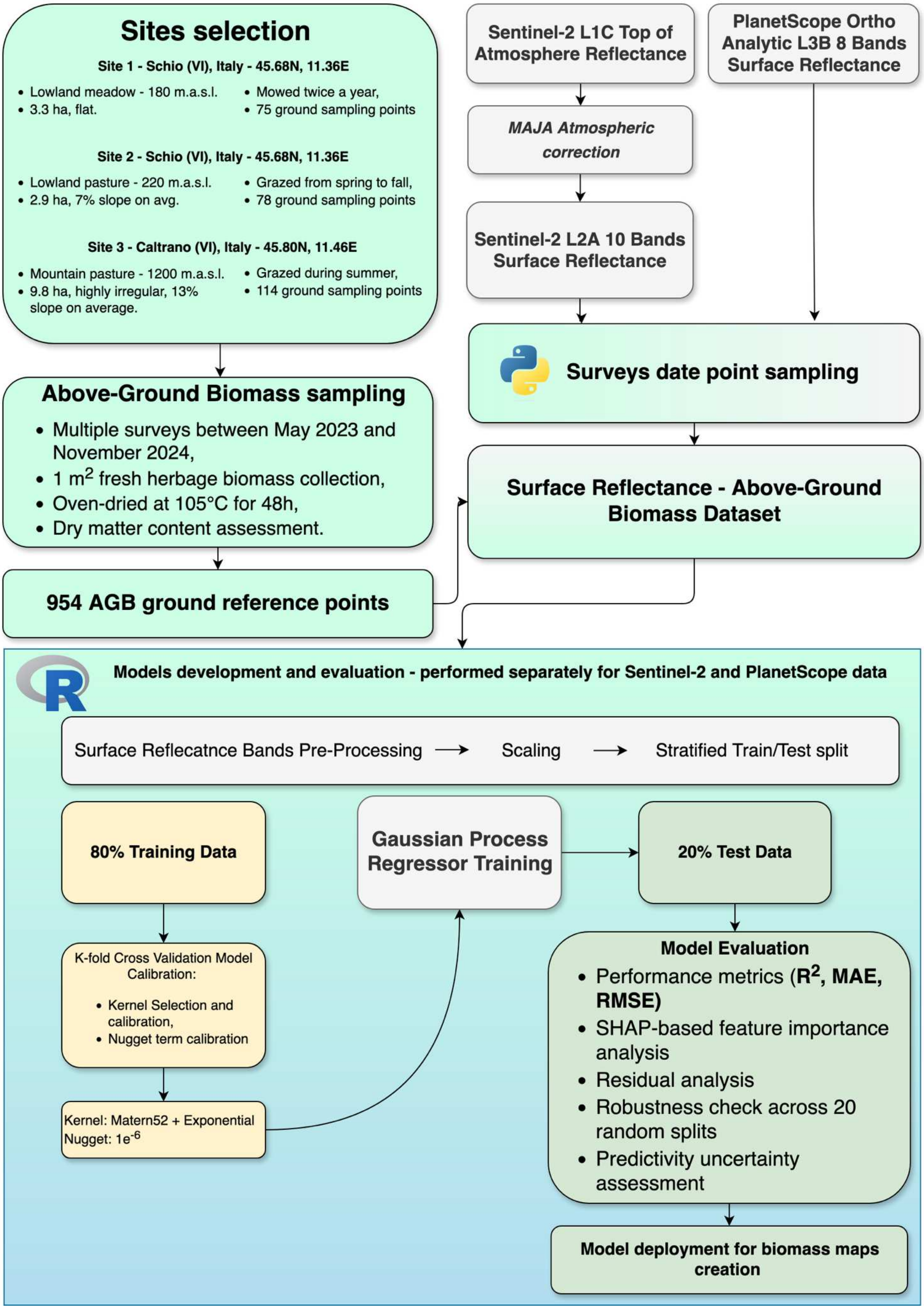


Figure 4.3 Workflow summarising the site selection, ground and satellite data collection, data preprocessing, Gaussian Process Regression (GPR) model development and validation.

4.4 Results

4.4.1. AGB Dataset Structure and Sampling Framework

The final dataset included above-ground biomass (AGB) measurements ranging from 12.3 to 4486.3 kg/ha, with a mean of 1279 kg/ha and a distribution skewed toward lower values (Fig. 4.4). The central 50% of the data was distributed between 753.6 and 1541.4 kg/ha, with a median of 1113.5 kg/ha, indicating a predominance of moderate biomass values across the surveyed area.

This variability reflects the intended representativeness of the dataset, which includes a permanent meadow sampled before and after the first and the second yearly mowing events, a permanent pasture monitored throughout its grazing cycle, and a mountain pasture surveyed during the summer grazing season. The high heterogeneity observed was both expected and desirable, as the study's objective was to develop a predictive model capable of generalising across a broad range of management conditions and vegetation states typical of semi-natural permanent grassland ecosystems.

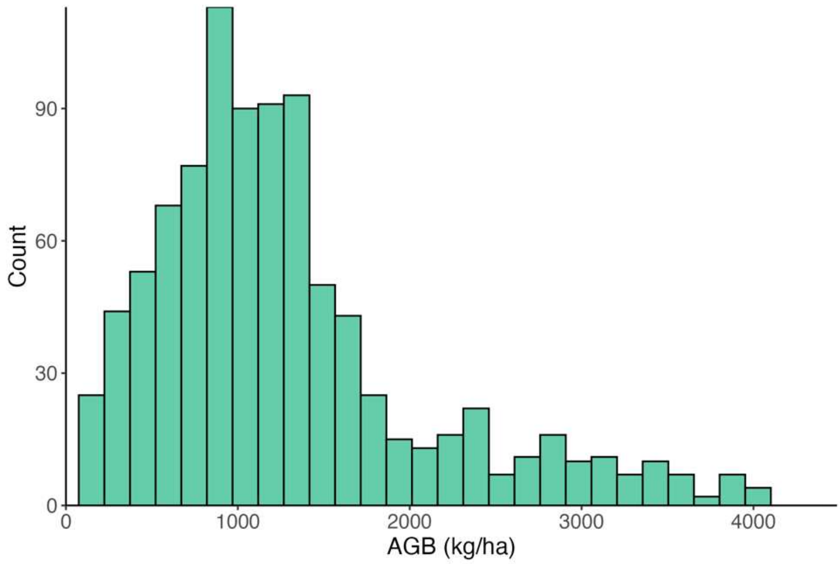


Figure 4.4 Histograms showing AGB data distribution

The dataset was split into training and unseen test sets using a stratified random sampling approach based on the quintiles of above-ground biomass (AGB). Specifically, the AGB values were divided into five strata using 20th, 40th, 60th, and 80th percentiles as breakpoints, ensuring that each AGB class was proportionally represented in both training and test subsets. This method effectively preserved the original distribution in both subsets, as illustrated in Figure 4.5a (Training set) and Figure 4.5b (Test set), ensuring that the full range of biomass conditions was equally represented equally during both the training and evaluation phases.

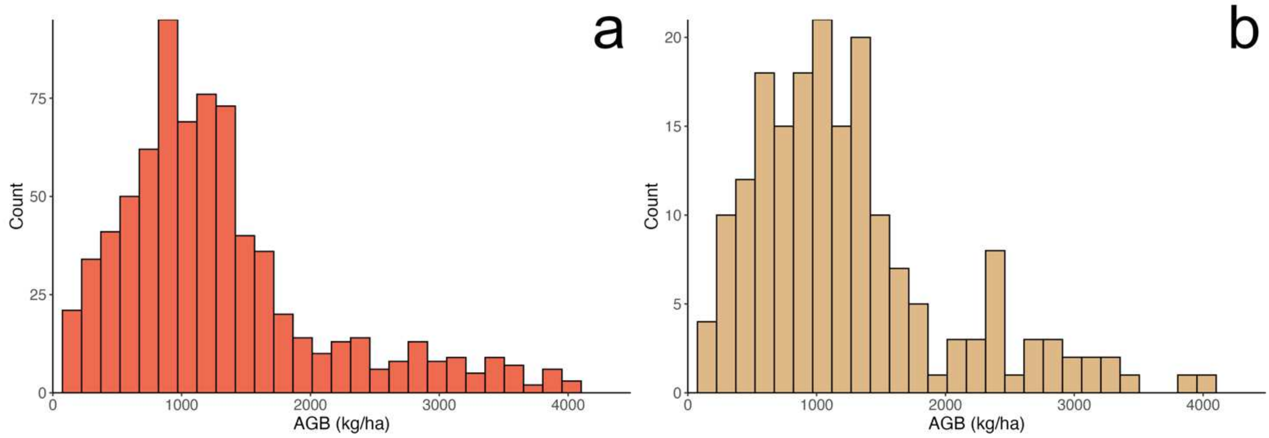


Figure 4.5 Histograms showing AGB data distribution in the Train set (a) and Test set (b).

4.4.2 Evaluation of AGB Prediction Models

Model performance was assessed using three commonly adopted regression metrics: the coefficient of determination (R^2), the mean absolute error (MAE), and the root mean square error (RMSE). Table 4.4 summarises the values of these metrics for both the Sentinel-2 and PlanetScope-based models, along with their associated 95% confidence intervals.

The Sentinel-2 model achieved an R^2 of 0.520, while MAE and RMSE were 396.9 kg/ha and 503.2 kg/ha, respectively, confirming moderate predictive ability and a tendency to underestimate higher AGB values (Fig. 6a). Although PlanetScope has a limited number bands, its model yielded a comparable R^2 of 0.528, along with comparable error metrics, including an RMSE of 498.1 kg/ha and an MAE of 395.1 kg/ha. Figure 6 displays scatterplots of predicted versus observed AGB values for both sensors. While both models successfully capture general trends in the data, predictions tend to cluster around mid-range biomass values, with underestimation at higher AGB values. This pattern reflects the skewed distribution of AGB in the dataset, where most observations fall within a moderate range. Despite the variability, the results demonstrate that both models offer valid estimates of AGB across a heterogeneous landscape and under diverse grassland management regimes.

Table 4.4 Coefficients of determination (R^2), mean absolute error (MAE) and root mean square error (RMSE), with the relative confidence intervals at 95%, for both Sentinel-2 and PlanetScope models at the pixel-level.

	Sentinel-2			**PlanetScope**		
		95%CI			95%CI	
		High	Low		High	Low
R2	**0.520**	0.630	0.349	**0.514**	0.637	0.313

MAE (kg/ha)	**396.9**	442.6	352.3	**400.1**	445.8	355.2
RMSE (kg/ha)	**503.2**	555.2	449.9	**506.1**	561.4	452.9

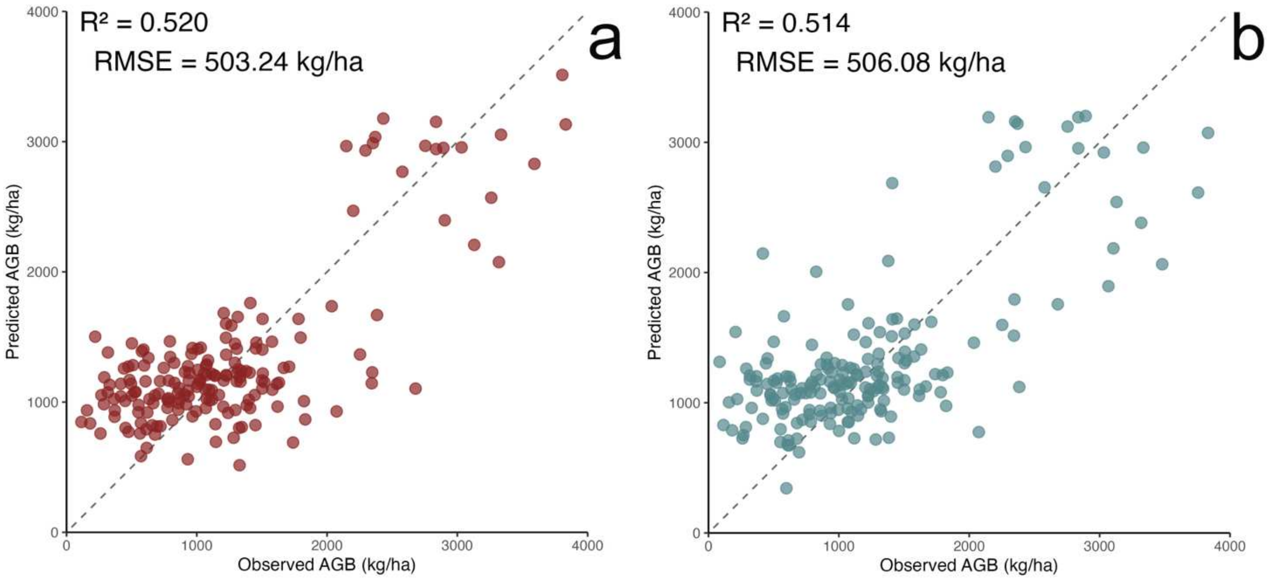


Figure 4.6 Scatterplot of Observed vs Predicted AGB values on the test dataset for Sentinel-2 (a) and PlanetScope (b) models.

4.4.3 AGB estimation at the field level

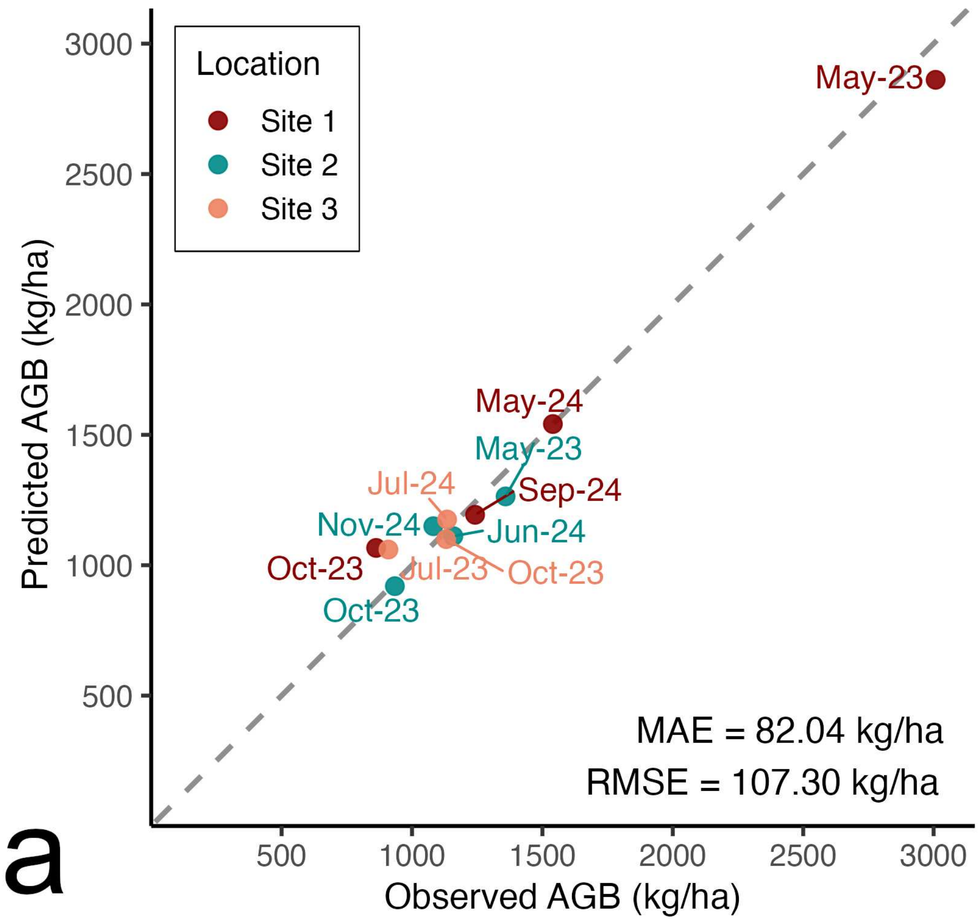
While point-based estimates provide valuable insights into model performance and allow sub-paddock field management and monitoring, field-level aggregation offers a more practical perspective for meadow and pasture management applications. In real-world settings, decisions are typically made at the paddock or field scale rather than at the level of individual pixels.

To evaluate the models under this more operational lens, AGB predictions were aggregated by field and survey date and compared with corresponding field-level mean observations. It is essential to note that the observed values in this analysis were derived from complete field measurements. In contrast, the predicted values were obtained exclusively from the unseen test subset, representing 20% of the observations. Nonetheless, the stratified sampling strategy adopted during data partitioning ensured adequate representation of all sites and survey dates within the unseen test set, enabling a robust and meaningful assessment of model performance at the field scale.

The Sentinel-2 and PlanetScope-based models demonstrated excellent performance in quantifying above-ground biomass at the field level across all sites and survey dates. The two models exhibited comparable predictive accuracy, consistent with the results observed at the pixel level. However, at the aggregated field scale, the Sentinel-2 model slightly outperformed the PlanetScope model, achieving an R^2 of 0.972 compared to 0.968, as shown in Figure 4.7. Table 4.5 highlights that the Sentinel-2 model yielded a mean absolute error (MAE) ranging from 61.2 to 117.5 kg/ha across the three sites, with percentage errors between 5.12% and 10.2%.

Table 4.5 Mean observed AGB (kg/ha), mean predicted AGB (kg/ha), mean absolute error (kg/ha) and mean percentage error (%) at the field level for all three sites and both Sentinel-2 (S2) and PlanetScope (PS) models.

Site	Mean observed AGB (kg/ha)	Mean predicted AGB (kg/ha)		Mean absolute error (kg/ha)		Mean percentage error (%)	
		S2	PS	S2	PS	S2	PS
Site 1	1658.8	1668.6	1675.2	104.2	117.5	10.2	12.7
Site 2	1132.6	1111.2	1228.9	61.20	73.72	5.12	7.75
Site 3	1058.4	1111.9	1089.1	80.72	93.84	8.45	9.21
Mean	1283.01	1296.2	1327.8	82.04	94.85	7.92	9.88



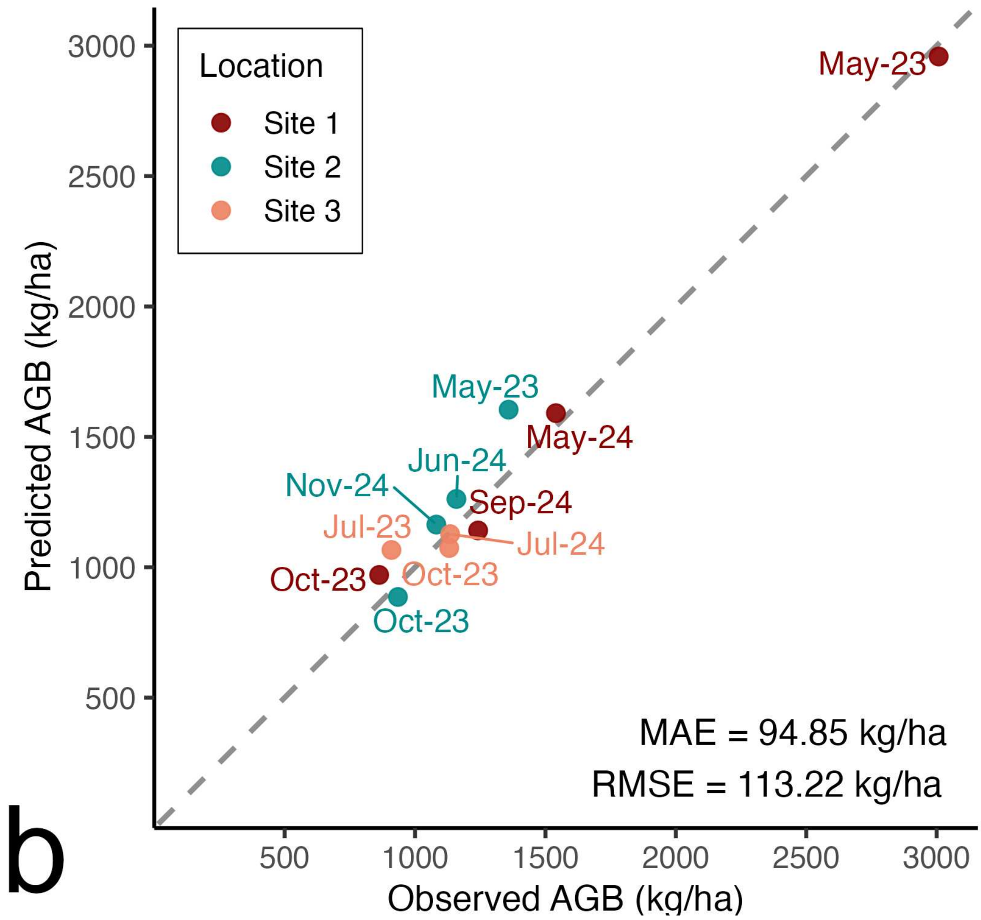


Figure 4.7 Scatterplot of predicted vs observed AGB (t/ha) values for Sentinel-2 (a) and PlanetScope (b) models at the field level.

4.4.3 Feature importance

The SHAP analysis of normalised absolute feature importance (expressed as each spectral region's percentage contribution to the total SHAP values) revealed consistent yet sensor-specific patterns in the spectral drivers of above-ground biomass prediction (Figure 4.8). For Sentinel-2 (Fig. 4.8a), the red-edge (B5–B7, 705nm to 783nm) and shortwave infrared (B11–B12, 1610nm to 2190nm) regions contributed most to model predictions, reflecting their sensitivity to canopy structure, chlorophyll content, and vegetation water status. The near-infrared (B8, 842nm) also provided substantial information, whereas the visible region (B2–B4, 490nm to 665nm) contributed minimally, indicating that these wavelengths are of limited relevance for biomass estimation in the study area. For PlanetScope (Fig. 4.8b), a similar pattern emerged, with the near-infrared (PLB8-865nm) region dominating the contribution spectrum, complemented by moderate influence from the visible (PLB2-PLB6, 490nm to 665nm) and red-edge (PLB7-705nm) regions.

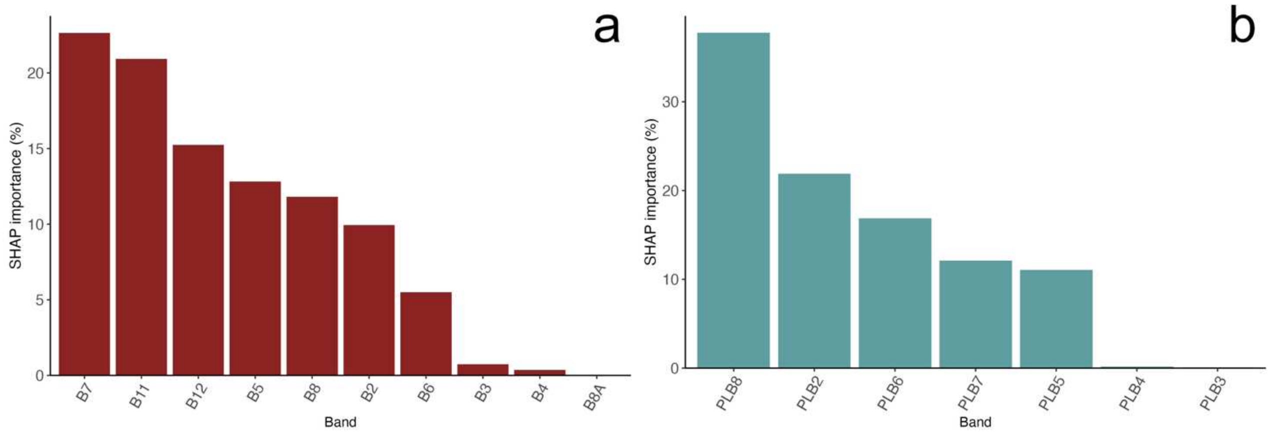


Figure 4.8 Normalised SHAP-based feature importance for Sentinel-2 (a) and PlanetScope (b) models.

4.4.4 Residuals, Robustness and Uncertainty Analysis

Residual analysis was first conducted to evaluate the adequacy of the Gaussian Process Regression (GPR) models. The histogram of residuals for the Sentinel-2 model (Fig. 4.9a) shows a slight left skew, with most residuals concentrated around zero and a longer positive tail. This skewness is also visible in the Q–Q plot (Fig. 4.10a), where deviations are most pronounced at the distribution tails. The Shapiro–Wilk test confirmed a significant deviation from normality ($W = 0.9683$, $p < 0.001$). By contrast, the PlanetScope residuals (Fig. 4.9b, Fig. 4.10b) appeared more symmetrically distributed and closely aligned with the theoretical quantiles, with the Shapiro–Wilk test confirming normality ($W = 0.9947$, $p = 0.745$).

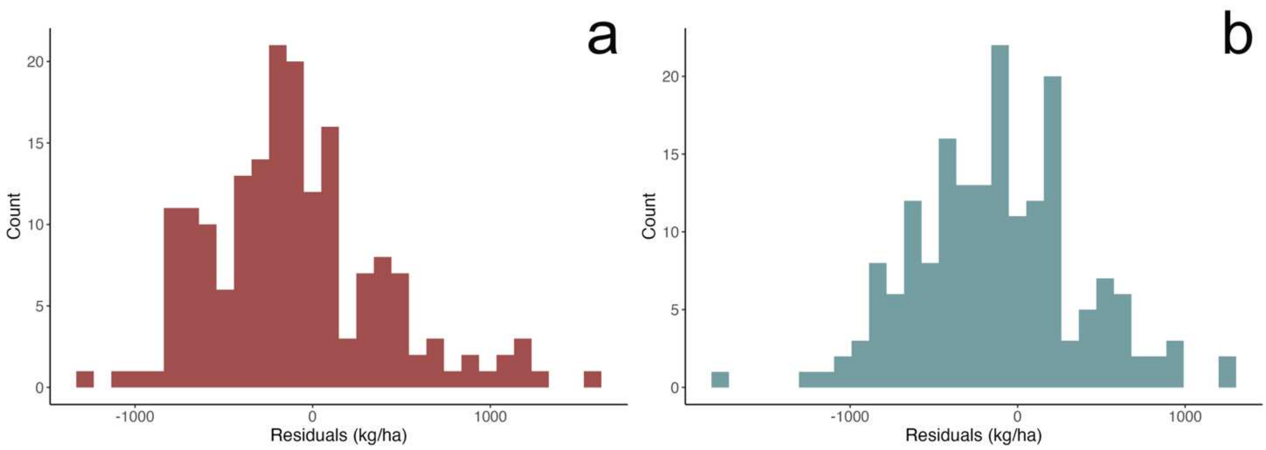


Figure 4.9 Histogram of residuals distribution for Sentinel-2 (a) and PlanetScope (b) models.

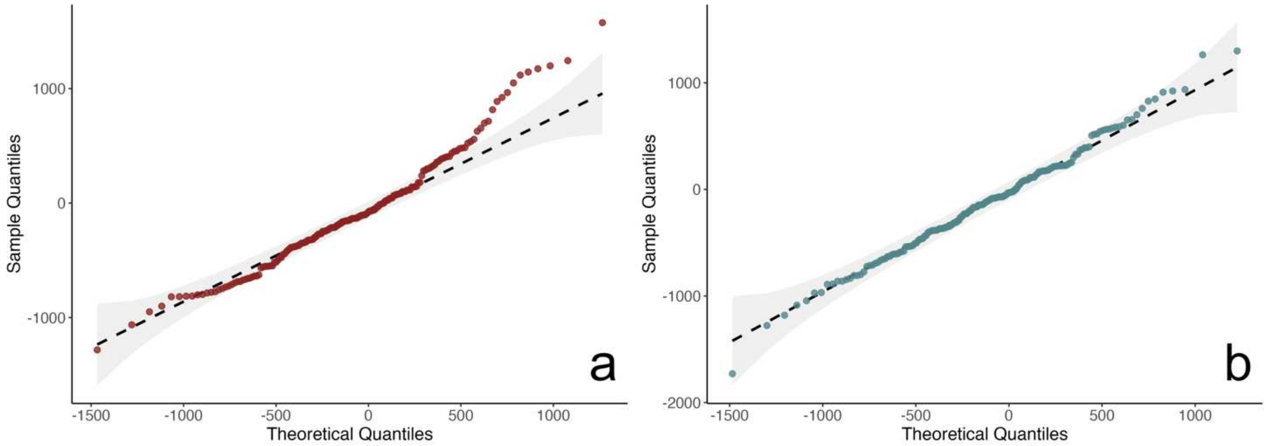


Figure 4.10 Q-Q plots of residuals for Sentinel-2 (a) and PlanetScope (b) models.

GPR does not require residuals to follow a normal distribution, as it assumes Gaussian shape of the joint distribution of predictions rather than individual residuals. Moderate departures from normality, therefore, do not compromise validity. Both models satisfied the assumption of homoscedasticity: the Breusch–Pagan test yielded non-significant results (Sentinel-2: BP = 0.3231, $p = 0.569$; PlanetScope: BP = 1.733, $p = 0.431$), and residuals vs. predicted plots (Fig. 4.11) showed no systematic variance patterns.

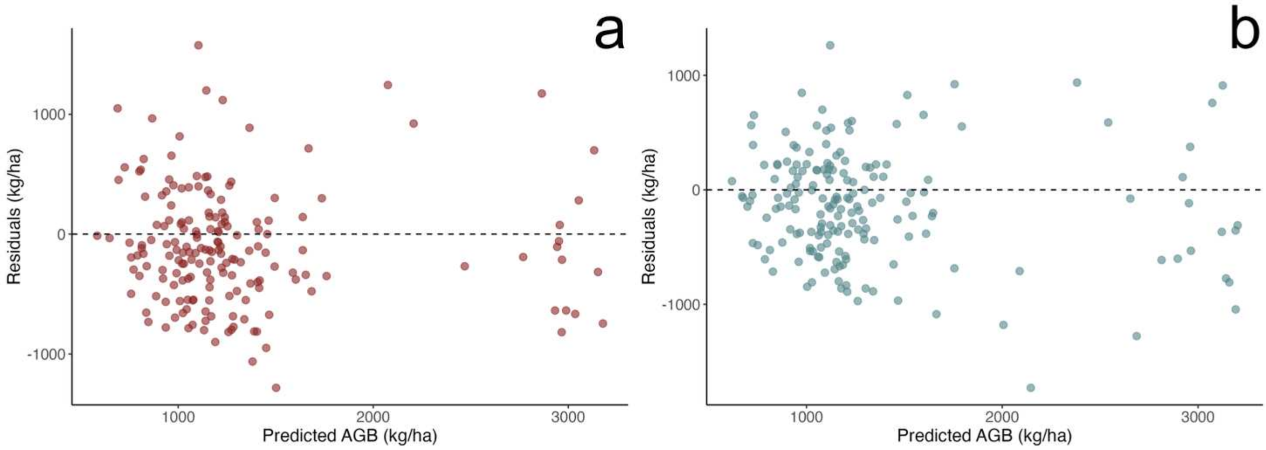


Figure 4.11 Scatterplot of residuals vs predicted AGB values for Sentinel-2 (a) and PlanetScope (b) models.

To evaluate robustness to data partitioning, we repeated the stratified 80/20 train–test split 20 times, re-fitting the models for each fold. Table 4.5 reports the mean and standard deviation of R^2, MAE, and RMSE across the 20 refits. Overall performance was stable across folds, confirming that the results reported in Table 4.3 are not sensitive to a particular partition. PlanetScope exhibited slightly lower variability than Sentinel-2, indicating marginally greater stability in the choice of split, while achieving comparable mean accuracy.

Table 4.6 Coefficients of determination (R2), mean absolute error (MAE) and root mean square error (RMSE) of compared GPR models across 20 repeated stratified 80/20 train–test splits for Sentinel-2 and PlanetScope models

	Sentinel-2			**PlanetScope**		
	R2	MAE (kg/ha)	RMSE (kg/ha)	R2	MAE (kg/ha)	RMSE (kg/ha)
Mean	0.502	406.2	502.3	0.513	401.2	504.8
Standard deviation	0.057	22.3	38.2	0.084	19.2	28.2

In terms of predictive uncertainty, we applied conformal prediction, a distribution-free framework that recalibrates predictive intervals to achieve valid empirical coverage (Vovk, 2015). Using out-of-fold residuals obtained from the same 20 repeated stratified 80/20 splits employed in the robustness analysis, we computed the 95th percentile of the standardised residuals to derive the conformal scaling factor. We rescaled the posterior standard errors to obtain calibrated intervals, as in Eq. (2). After calibration, both models produced conservative and well-calibrated uncertainty estimates (Table 4.7). At the pixel level, the median PI95 width remained relatively wide, 1472.4 kg DM ha^{-1} for Sentinel-2 and 1366.2 kg DM ha^{-1} for PlanetScope, with median relative PI95 widths of 1.32 and 1.22, and median CV% of 30.1% and 28.5%, respectively, reflecting substantial uncertainty at fine pixel-scale. Aggregation at the field level markedly reduced uncertainty: median PI95 widths fell to 342.6 (Sentinel-2) and 324.8 kg DM ha^{-1} (PlanetScope), with median relative widths of 0.29 and 0.27, and CV% of 7.3% and 6.9%. Coverage at the field scale reached 1.00 for both sensors, confirming reliable calibration at operational decision scales.

Table 4.7. Predictive uncertainty summary for GPR models using Sentinel-2 and PlanetScope at pixel and field level. Metrics report calibrated 95% prediction interval (PI95) coverage, median PI95 width (kg DM ha^{-1}), median relative PI95 width (PI95 width ÷ predicted mean), and median coefficient of variation (CV%) (SE ÷ mean × 100).

	Sentinel-2		**PlanetScope**	
	Pixel-level	**Field-level**	**Pixel-level**	**Field-level**

Calibrated prediction interval (PI95) coverage (%)	96.6 %	100 %	95.3 %	100 %
Median prediction interval (PI95) width (kg/ha)	1472.4	342.6	1366.2	324.8
Median relative prediction interval (PI95) width	1.32	0.32	1.22	0.30
Median coefficient of variation (CV%)	30.1 %	7.3 %	28.5%	6.9 %

4.4.5 Spatial visualisation of predictions

To complement the numerical model evaluation, we generated spatial maps of predicted biomass, along with the associated GPR posterior standard errors. Figure 4.12 shows the biomass and standard error maps produced at all sites with reflectance data collected on 15/08/2023 for Sentinel-2 and 13/08/2023 for PlanetScope at all sites.

Across all sites, Sentinel-2 and PlanetScope yielded broadly consistent biomass patterns, with PlanetScope maps capturing finer within-field heterogeneity due to their higher spatial resolution.

While at site 1, the standard error maps show inconsistent spatial structure between the two models and therefore do not provide information on the error's spatial distribution. In contrast, uncertainty maps for sites 2 and 3 reveal a clear spatial structure. Higher error values occur along field boundaries and in areas characterised by the presence of trees and their shadows. The level of uncertainty remains high at the pixel level for both platforms, despite PlanetScope's model generally showing less uncertainty across predictions.

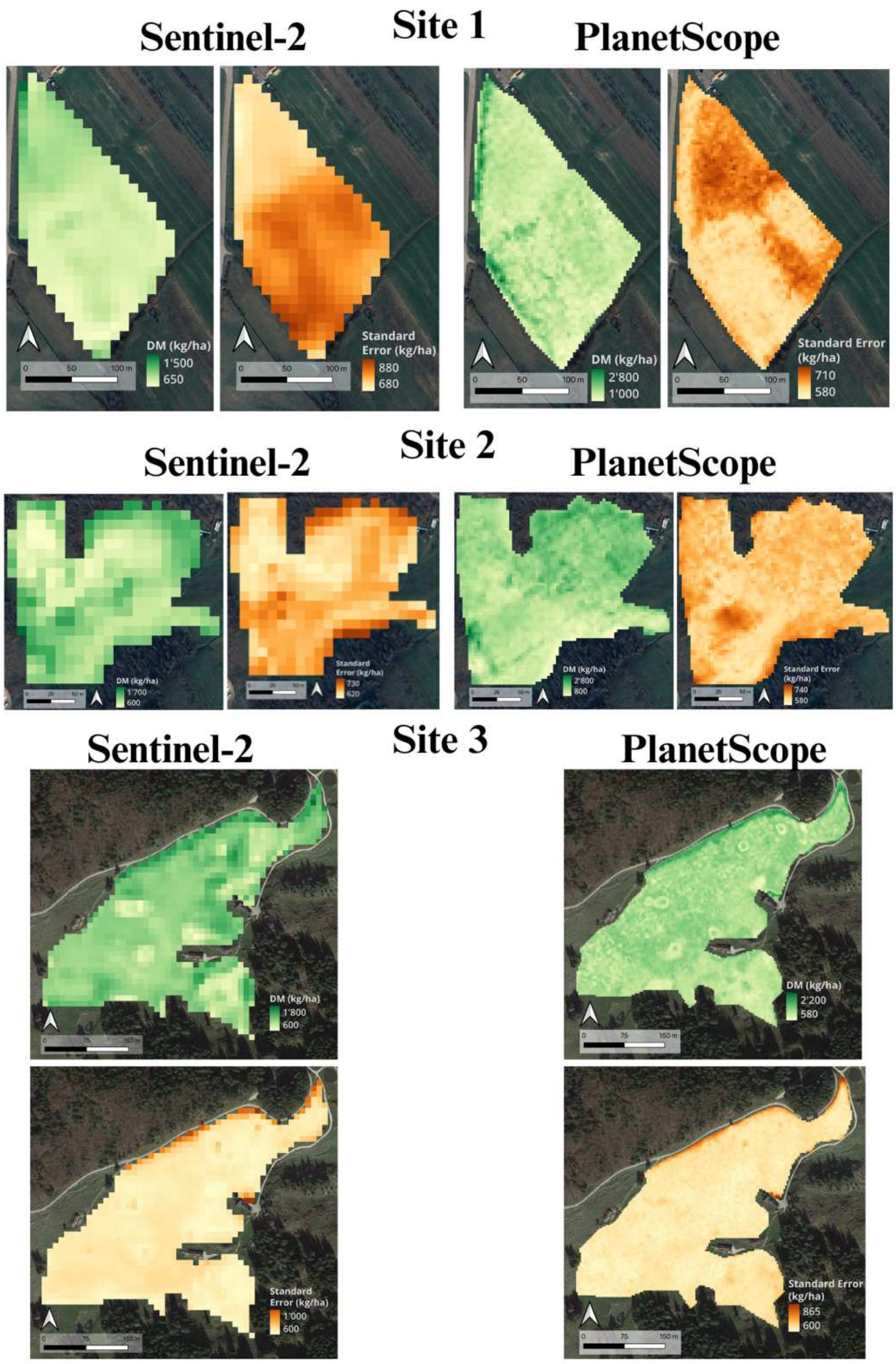


Figure 4.12 Spatial maps of predicted above-ground biomass and Gaussian Process Regression posterior standard error for Sentinel-2 and PlanetScope models across the three study sites. Each pair of panels shows predicted AGB (kg DM ha^{-1}) and corresponding uncertainty (SE; kg DM ha^{-1}) for representative dates. Sentinel-2 estimates are produced with imagery collected on the 15/08/2023 while PlanetScope imagery used for the estimates was collected on the 13/08/2022.

4.5 Discussion

4.5.1 Model Comparison at Pixel and Field Levels

The first objective of this study was to evaluate the performance of GPR models in estimating AGB at both pixel and field scales, while the second was to compare satellite platform's data performance. As shown in Table 4.3 and Figure 4.6, the Sentinel-2 and PlanetScope models achieved similar predictive accuracy at the pixel level, with R^2 values of 0.520 and 0.514 and RMSE slightly above 500 kg/ha. These results align with those reported by Smith et al. (2023), who conducted the only prior study focused on biomass estimation at a comparable very high spatial resolution. Although their models achieved slightly higher R^2 values (up to 0.72), the reported RMSEs were around 700 kg/ha, and the study was conducted within a homogeneous field environment, unlike the present work, which encompasses multiple fields with diverse environmental and management conditions.

Several factors influence model accuracy at this scale, with multiple sources of potential error affecting prediction reliability, including geometric distortions in satellite imagery, positional inaccuracies in ground reference data, coarse and limited spectral bads and spectral saturation. Both Sentinel-2 and PlanetScope imagery undergo geometric correction to ensure spatial accuracy, especially in areas with complex terrain. Sentinel-2 products are orthorectified using Digital Elevation Models at 30m resolution to correct distortions caused by slope and elevation. Similarly, PlanetScope Ortho Scene products apply orthorectification using DEMs with resolution varying from 30 to 90 m and ground control points to align imagery to accurate ground positions. These processing steps are efficient in flat or uniformly sloped areas; however, their accuracy may decrease in rugged and topographically complex environments, such as Site 3 in our study, where extreme elevation changes and surface heterogeneity can introduce residual geometric distortions (Peng et al., 2024).

Additionally, the GNSS receiver used for geolocating sampling points achieved sub-meter accuracy at Sites 1 and 2. In contrast, at the more rugged and mountainous Site 3, positional accuracy declined to 2–3 meters, introducing an additional source of potential error in model validation and pixel-level comparison. However, including such challenges enhances our approach's robustness, demonstrating its capacity to handle more significant variability and improving its potential for broader generalisation across diverse grassland environments. Beyond geometric correction, the spatial maps in Figure 4.12 reveal clear adjacency and edge effects at the pixel scale. Especially in Sites 2 and 3, higher standard errors cluster along field borders and near roads, buildings, and tree lines, where mixed pixels blend canopy reflectance with non-grass targets. Consequently, even with accurate co-registration, border pixels and pixels adjacent to structures exhibit larger uncertainty because their spectra integrate multiple surface types. The model's accuracy significantly improved when aggregating the model estimations at the field level and averaging the AGB predicted values by field and survey date. As seen at the pixel level, the differences in terms of performance between the satellite platforms were negligible. As summarised in Table 4.4 and shown in Figure 4.7, the RMSE was slightly above 100 kg/ha for both Sentinel-2 and PlanetScope models, with mean percentage error consistently below 10% on all sites. Despite these marginal differences, the overall performance of both models can be considered highly robust and practically equivalent for operational applications. These results are consistent and outperform the ones collected in Australia by Ogungbuyi et al., (2024) at the farm level, using both Sentinel-2 and PlanetScope data in a Random Forest-based model, reinforcing the potential of these two satellite platforms' applications for pasture biomass estimation and monitoring.

The SHAP-based variable importance analysis of the Sentinel-2 model (Fig. 4.7a) indicated that the red-edge bands (B5-B7) and the shortwave infrared bands (B11-B12) were the most influential in predicting above-ground biomass, followed by the near-infrared band B8. These spectral regions are closely linked to canopy structure, chlorophyll absorption, and water status, which are fundamental determinants of biomass. This pattern is broadly consistent with the findings of Dusseux et al. (2022), who also identified red-edge, near infrared and short-wave infrared bands as key contributors. For PlanetScope (Fig. 4.7b), the near-infrared band (PLB8) accounted for the largest share of predictive importance, with additional contributions from the blue (PLB2), red (PLB6), and red-edge (PLB7) regions. This outcome reinforces the critical role of the red-edge domain and suggests that PlanetScope's broad visible bands provide complementary information. Similarly, Ogungbuyi et al. (2024) highlighted the dominant effect of the NIR band in PlanetScope-based biomass estimation. More generally, previous research has confirmed the central role of red-edge and NIR regions in capturing biomass variability; for example, Sibanda et al. (2015) demonstrated that Sentinel-2's red-edge and NIR bands outperform broader-band sensors, with strong correlations to biomass across different fertiliser treatments. However, the strong multicollinearity among adjacent spectral bands, an inherent characteristic of spectral data, limits the interpretability of feature-importance methods such as SHAP when evaluated at the level of single bands. The correlation analysis revealed that Sentinel-2 bands within the red-edge and NIR regions exhibit Pearson correlation coefficients exceeding 0.9, while SWIR bands (B11–B12) are also highly correlated ($r \approx 0.8$). PlanetScope bands displayed a similar pattern, with correlations above 0.8 across most of the visible-to-NIR range, reflecting their broad spectral response functions. These results confirm that successive bands capture overlapping spectral information, making it inappropriate to interpret their SHAP importance in isolation. For this reason, our interpretation focuses on spectral regions, visible, red-edge, NIR, and SWIR, rather than on individual wavelengths.

An additional point of comparison between Sentinel-2 and PlanetScope concerns data volume and operational costs. As reported in Table 2, the average data volume per hectare was 0.838 kB for Sentinel-2 and 8.89 kB for PlanetScope, reflecting PlanetScope's higher spatial resolution and denser pixel structure. From a practical standpoint, this translates into greater storage and processing requirements when scaling up analyses across large regions. In terms of costs, Sentinel-2 imagery is freely available through the Copernicus programme, while PlanetScope is a commercial service. The actual subscription fees depend on spatial and temporal coverage and on specific contractual agreements, making direct comparisons challenging. Nevertheless, Sozzi et al. (2021) estimated that the cost of PlanetScope data acquisition is approximately 0.011 € per hectare. In addition to cost and data volume, practical deployment also depends on computational scalability, which was explicitly evaluated for both models (Table 3). Despite GPR's cubic scaling with training sample size, the model remains computationally tractable and operationally scalable for field- and regional-scale biomass monitoring applications. While providing useful reference, accurately quantifying the economic and operational cost of producing a commercial, end-user-oriented biomass monitoring platform lies beyond the scope of this study.

4.5.2 Operational Suitability for Farm-scale Biomass Monitoring

The third objective of this study was to evaluate the readiness of the GPR-based modelling approach for practical, farm-scale applications. Defining acceptable accuracy for above-ground biomass (AGB) grassland estimates depends mainly on their intended application, particularly in informing grazing decisions, feed budgeting, and agri-environmental compliance. In practice, errors must be low enough to avoid misleading farmers when planning stocking rates or harvesting schedules.

To date, most of the practical research, development, and validation of satellite-based pasture biomass estimation tools has been conducted in Australia, where precision grazing technologies are more widely adopted and tested under actual farm conditions. According to Correa-Luna et al. (2024), pasture biomass estimates with a mean absolute error (MAE) in the range of 293–342 kg DM/ha, or approximately 11–13% of observed values, are considered sufficient for operational use in pasture-based systems. These values are consistent with the performance of current commercial satellite-based tools used in Australian grazing systems. For instance, the Pastures from Space platform explains 70–75% of the variance in Feed on Offer (FOO) and reports a mean error of ±300 kg DM/ha (Smith et al., 2011). However, due to its reliance on NDVI data extracted from MODIS, its spatial resolution of 6.25 ha/pixel makes it unsuitable for medium and small-scale farms that characterise the European and Italian agricultural landscape. PastureSmarts (https://app.pasturesmarts.com), a commercial platform based on unspecified satellite data, reports a relative error of approximately

13.5% across a biomass range of 410–5777 kg DM/ha. However, it does not disclose the spatial resolution used. Pasture.io (https://pasture.io), the only commercial platform currently available in some European countries, though not yet in Italy, claims independently tested mean absolute errors (MAE) as low as ± 56 kg DM/ha for field conditions with biomass levels up to 3500 kg DM/ha. This platform works at the 0.012 ha resolution and relies on PlanetScope and Sentinel-2 constellations data, together with weather, soil, and farm management information (Correa-Luna et al., 2024), and therefore is the most suitable benchmark for comparing our proposed approach. While the inclusion of these additional data sources understandably enhances performance, our approach demonstrates that operationally useful accuracy can be achieved using only multispectral satellite imagery, without requiring external environmental or management information. Although our pixel-level (~0.010 ha for Sentinel-2, ~0.0009 ha for PlanetScope) predictions resulted in MAEs of around 440 kg DM/ha for both Sentinel-2 and PlanetScope models (Fig. 4.6, Table 4.4), exceeding the error margins reported by commercial platforms, performance improved substantially when estimates were aggregated at the field scale. In our sites, ranging from 3 to 10 ha, our models consistently achieved MAEs close to or below 100 kg DM/ha and relative errors well under 10% (Figure 4.7, Table 4.5), thus meeting or exceeding the acceptable standards for practical use. In addition to improved accuracy, field-level aggregation substantially reduces predictive uncertainty: conformal-calibrated 95% intervals shrink from wide, pixel-scale bands to narrow, operationally meaningful ranges at the field scale, with median relative interval widths and CV% dropping to 0.30–0.32 and 6–7%, respectively (Table 7). These results indicate that meaningful uncertainty reduction occurs only when predictions are aggregated at the field scale (approximately 3–10 ha). This range therefore represents the minimum operational unit for obtaining reliable, decision-ready estimates from multispectral satellite data with the proposed model. Continued refinement of model structure, expansion of training datasets, and targeted recalibration are expected to progressively improve prediction accuracy at finer spatial resolutions.

4.5.3 Limitations and Future Perspectives

Satellite-based multispectral imagery faces well-documented limitations in estimating above-ground biomass (AGB) in grasslands, particularly at high biomass levels. A key issue is spectral saturation, where reflectance signals become insensitive to further increases in vegetation density. This typically occurs when the leaf area index (LAI) exceeds 3 in grassland systems, beyond which further increases in biomass produce minimal changes in spectral response (Peng et al., 2023). As a result, remote sensing models tend to underestimate high AGB values since the sensor signal no longer captures actual biomass variation (Sa et al., 2024). Due to this saturation issue, the abovementioned Australian

Pasture from Space platform becomes less reliable than the 2000 kg/ha threshold (Donald et al., 2010). Another limitation of satellite-based biomass estimation approaches is mixed pixels and coarse spectral and spatial resolution. If a sensor's pixels cover a mix of grass, soil, and other features, the spectral value is an average that may not represent biomass well. Past studies noted that high-resolution data (e.g. 30 m Landsat) suffered from mixed-pixel effects and saturation problems in pasture biomass mapping (Chen et al., 2021). While Sentinel-2's (10 m) and PlanetScope's (3 m) higher spatial resolution helps mitigate this issue to some extent, even 10 m pixels can be affected in areas with complex terrain, high spatial variability, and surface rock presence, conditions typical of alpine mountain pastures, such as Site 3 of our study area. In addition, the distribution of our field training data was skewed, with relatively few samples representing very high and very low biomass levels (Fig. 4). The combination of soil and rock disturbance, spectral saturation and lack of training data at the extremes of biomass distributions have a strong effect on the reliability of our models at different biomass levels. In our study, for both Sentinel-2 and PlanetScope, RMSE values were lowest (270–320 kg ha^{-1}) within the 1000–1500 kg ha^{-1} interval and remained below 450 kg ha^{-1} up to about 2000 kg ha^{-1}. In contrast, errors increased sharply exceeding 700 kg ha^{-1} for both very low biomass pixels below 500 kg ha^{-1} and for dense canopies above 2000 kg ha^{-1}.

Furthermore, mixed pixels and canopy shadows can dampen reflectance in the NIR and red-edge regions, further contributing to underestimation in very dense plots. Finally, Gaussian Process Regression, being a smooth non-parametric interpolator, naturally regresses extreme values toward the mean when these are underrepresented in the training set, which likely reinforced the systematic underprediction at the upper tail of the biomass distribution. Such limitations are common to multispectral remote sensing approaches, where both sensor constraints and sampling imbalances restrict the capacity to fully capture variability at the highest biomass levels. Moreover, although 954 AGB samples were collected over 18 months, they corresponded to only 267 distinct spatial positions. This design increased the temporal depth of observations and ensured that seasonal and management variability was well captured, including permanent meadows sampled before and after mowing, rotational fields across multiple cuts, and alpine pastures during summer grazing. However, the limited number of independent sites may restrict the spatial generalizability of the models, as fewer distinct landscapes were represented. All sites were located within northern Italy, providing robust insights into the region's managed grasslands, but generalisation to other climatic, management, or vegetation regimes has not yet been tested. The model is therefore best suited to northern Italy and potentially applicable to neighbouring Alpine and pre-Alpine regions with comparable conditions, while broader validation and refinement will be necessary to adapt the framework to other ecosystems. Given the small number of independent sites, leave-one-site-out or leave-one-season-out validation would have removed too much relevant information and introduced strong site-specific biases. Because multiple temporal observations were collected at the same spatial locations, the stratified 80/20 split adopted in this study allowed samples from identical coordinates to appear in both the training and test sets. As a result, the reported performance reflects within-site interpolation, that is, the model's ability to predict new observations under familiar spatial and management contexts, rather than out-of-site generalisation to entirely unseen areas. This design choice was made to preserve distributional balance across biomass levels, seasons, and management regimes, ensuring a robust assessment of model behaviour within the study region. We therefore consider the adopted approach sufficiently reliable for within-region applications, while acknowledging that it does not fully address spatial autocorrelation or evaluate true out-of-site predictive capability. Future work will explicitly test spatial generalisation using independent-site validation as additional data become available. Beyond model accuracy, the uncertainty analysis highlighted that pixel-level predictions were associated with wide confidence intervals, reflecting the combined effects of spectral saturation, geolocation mismatches, and canopy heterogeneity. However, aggregation at the field scale markedly reduced interval width and relative uncertainty, producing values of operational relevance for grazing and feed management. Further reductions in predictive uncertainty will require enlarging the training dataset to better capture rare but ecologically meaningful conditions, particularly at very high biomass levels. Such data expansion could benefit from participatory sampling approaches, where farmers and stakeholders contribute to field measurements, thus strengthening model calibration across management practices, vegetation types, and environmental gradients. Future research should also incorporate spatially explicit validation schemes, such as leave-one-location-out or spatial blocking, to strengthen model generalizability. Although GPR proved to be the most accurate and robust model in this study, it does not scale efficiently to vast datasets due to the cubic computational complexity of kernel inversion. This limitation was not restrictive in our case, given the moderate dataset size (n = 954), but may become critical in regional or continental-scale applications. For such contexts, scalable kernel approximations such as sparse GPR, inducing-point methods, or distributed training frameworks could retain the benefits of probabilistic predictions while ensuring computational tractability.

In addition to optical data, incorporating ancillary data such as topography, weather and management information greatly enhances biomass estimation models. Topography, rain, temperature, soil moisture, and management events drive grassland growth and phenology (Pinna et al., 2025). By explicitly adding these inputs, models can account for productivity differences that spectral signals alone can't explain, reaching higher estimation accuracies as done by previously mentioned commercially available decision support systems (Correa-Luna et al., 2024; Hanrahan et al., 2017).

Collecting high spatial resolution and weather data while capturing within-field variability, can be challenging. Still, the development of self-powered, permanent IoT-based air and soil monitoring systems, such as the one proposed by Ramson et al. (2021) could provide continuous, high-resolution data on temperature, precipitation, soil moisture, and nutrient dynamics, offering valuable support for precision management and improved model calibration in complex agroecosystems. Furthermore, advancements in deep learning are enabling continuous biomass monitoring by effectively addressing challenges such as cloud cover through the integration of Sentinel-1 SAR and Sentinel-2 optical data (Lasko, 2022). Additionally, integrating empirical reflectance-based models with physically based radiative transfer models (RTM) offers a promising avenue for enhancing the estimation of qualitative features of forage, such as moisture content and nutritional features, which are crucial for livestock management. The spectral sensitivity of vegetation to qualitative traits, such as protein, carotenoid, and anthocyanin content, is well established, and hybrid approaches that combine empirical modelling with radiative transfer theory have shown strong potential for accurately retrieving these biochemical properties (Dehghan-Shoar et al., 2024). The growing availability of hyperspectral satellite missions, such as the Italian Space Agency's PRISMA and the German Aerospace Center's EnMAP, provides high spectral resolution data, unlocking new possibilities for enhancing satellite-based monitoring of vegetation characteristics. Finally, future work should quantify the digital footprint of these models and any operational platforms using a DF-LCA approach based on Data Processing Units to aggregate performance, energy, and value indicators per unit of data across the whole data life cycle, to ensure virtual-environment sustainability (Huang et al., 2025).

4.6 Conclusions

This study provides robust evidence supporting the use of Gaussian Process Regression (GPR) models in combination with high-resolution multispectral satellite imagery from Sentinel-2 and PlanetScope constellations for estimating above-ground biomass (AGB) in diverse managed grassland ecosystems typical of European and Italian livestock farms.

Although pixel-level predictions exhibited moderate accuracy affected by geolocation uncertainty, terrain complexity, and spectral saturation in dense canopies, the models demonstrated strong generalisation capability across sites and dates. Remarkably, when predictions were aggregated at the field scale, model performance improved significantly, reaching mean absolute errors well below 100 kg DM/ha and relative errors under 10%. These values fall within the thresholds required for operational use in decision-support contexts, such as feed budgeting and adaptive grazing management. They appear to be comparable to the performance reported for commercially available decision-support systems, although direct quantitative benchmarks are often not publicly available.

However, these accuracies represent within-site interpolation performance, reflecting the model's ability to predict new observations within familiar spatial contexts rather than out-of-site generalisation. Therefore, while the reported results are robust for within-region applications, further validation across independent locations will be required to assess true spatial transferability. Furthermore, the dataset covers only two growing seasons, and the model has not yet been validated over multiple years or under contrasting climatic conditions. Consequently, multi-year verification will be necessary to evaluate its temporal stability and ensure reliable long-term performance for operational applications.

The comparable performance of Sentinel-2 and PlanetScope platforms further highlights that both medium- and high-resolution multispectral sensors can support accurate AGB estimation, with the choice of platform depending on trade-offs between spatial detail, revisit frequency, and cost. Nonetheless, prediction uncertainty at the pixel scale remains non-trivial, even after calibration, owing to geolocation mismatch, canopy heterogeneity, and spectral saturation; aggregating to the field level markedly tightens intervals, but broader, more diverse training data (potentially via participatory/crowd-sourced measurements) and ancillary drivers (weather, soil, management) will be needed to improve model's usability further. Future research should focus on integrating physically based radiative transfer models (RTM) with empirical approaches and high-resolution hyperspectral satellite data, potentially in combination with IoT-based in situ monitoring systems. This integration could enhance the generalisability and accuracy of biomass predictions, particularly under extreme environmental conditions or in heterogeneous landscapes. Moreover, the high spectral sensitivity of hyperspectral sensors, coupled with RTM-informed modelling, holds strong potential for retrieving qualitative traits of forage vegetation, including nutritional parameters, offering valuable insights for livestock nutrition and pasture quality assessment.

Finally, due to its ability to provide consistent, transparent, and spatially explicit biomass estimates, the modelling framework proposed in this study also has the potential to contribute to developing reliable Measurement, Reporting, and Verification (MRV) protocols within carbon farming and climate mitigation initiatives, as suggested by the FAO's protocol (2020).

4.7 Supplementary material

Table S1. Coefficients of determination (R2), mean absolute error (MAE) and root mean square error (RMSE) of compared model approaches tested for both Sentinel-2 and PlanetScope models: Random Forest (RF), Support Vector Regression (SVR), Multilayer Perception Neural Network (MLP), Lasso Regression and Gaussian Process Regression (GPR).

	Sentinel-2					PlanetScope				
	RF	*SVR*	*MLP*	*Lasso*	*GPR*	*RF*	*SVR*	*MLP*	*Lasso*	*GPR*
R2	0.412	0.398	0.351	0.406	0.530	0.481	0.415	0.280	0.330	0.528
MAE (kg/ha)	454.4	442.3	476.2	461.6	392.7	433.6	448.5	493.7	484.2	395.1
RMSE (kg/ha)	593.1	600.8	623.4	596.5	497.3	557.7	591.2	571.4	558	498.1

Table S2. Model performance by observed biomass classes on the test set. Each class represents a range of observed above-ground biomass (AGB, kg DM ha⁻¹). For each interval, the number of samples and the root mean square error (RMSE, kg ha⁻¹) are reported for Sentinel-2 and PlanetScope Gaussian Process Regression (GPR) models.

	Observed AGB (kg/ha) class				
	<500	500-1000	1000-1500	1500-2000	<2000
N. of points in test set	25	57	63	19	26
RMSE (kg/ha)					
Sentinel-2	733.3	403.9	268.4	525.8	789.7
PlanetScope	785.1	425.1	315.2	444.8	718.9

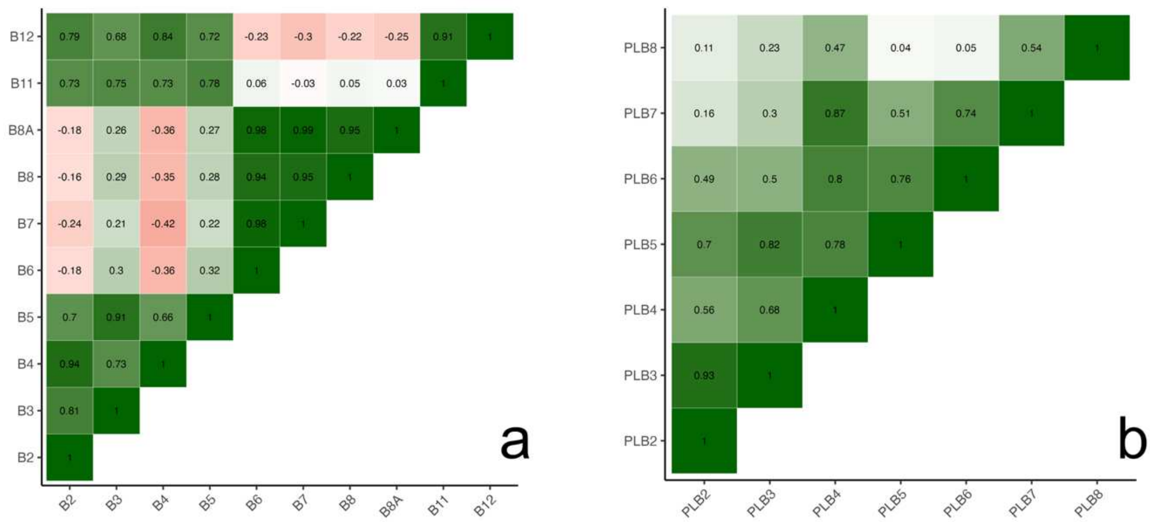


Figure S1. Pearson correlation matrices for spectral bands used in model training: (a) **Sentinel-2** (bands B2–B12) and (b) **PlanetScope** (bands PLB2–PLB8). Each cell reports the pairwise correlation coefficient (*r*) between spectral bands, based on the training dataset.

4.8 References

Bai, Y., & Cotrufo, M. F. (2022). Grassland soil carbon sequestration: Current understanding, challenges, and solutions. *Science*, *377*(6606), 603–608. https://doi.org/10.1126/science.abo2380

Chen, Y., Guerschman, J., Shendryk, Y., Henry, D., & Harrison, M. T. (2021). Estimating Pasture Biomass Using Sentinel-2 Imagery and Machine Learning. *Remote Sensing*, *13*(4). https://doi.org/10.3390/rs13040603

Correa-Luna, M., Gargiulo, J., Beale, P., Deane, D., Leonard, J., Hack, J., et al. (2024). Accounting for minimum data required to train a machine learning model to accurately monitor Australian dairy pastures using remote sensing. *Scientific Reports*, *14*(1), 16927. https://doi.org/10.1038/s41598-024-68094-3

De Rosa, D., Basso, B., Fasiolo, M., Friedl, J., Fulkerson, B., Grace, P. R., & Rowlings, D. W. (2021). Predicting pasture biomass using a statistical model and machine learning algorithm implemented with remotely sensed imagery. *Computers and Electronics in Agriculture*, *180*. https://doi.org/10.1016/j.compag.2020.105880

Dehghan-Shoar, M. H., Kereszturi, G., Pullanagari, R. R., Orsi, A. A., Yule, I. J., & Hanly, J. (2024). A physically informed multi-scale deep neural network for estimating foliar nitrogen concentration in vegetation. *International Journal of Applied Earth Observation and Geoinformation*, *130*. https://doi.org/10.1016/j.jag.2024.103917

Development, E. C. D.-G. for A. and R., Peyraud, J.-L., & MacLeod, M. (2020). *Future of EU livestock – How to contribute to a sustainable agricultural sector ? – Final report*. Publications Office. https://doi.org/doi/10.2762/3440

Donald, G. E., Gherardi, S. G., Edirisinghe, A., Gittins, S. P., Henry, D. A., & Mata, G. (2010). Using MODIS imagery, climate and soil data to estimate pasture growth rates on farms in the south-west of Western Australia. *Animal Production Science*, *50*(6), 611–615. https://doi.org/10.1071/AN09159

Dusseux, P., Guyet, T., Pattier, P., Barbier, V., & Nicolas, H. (2022). Monitoring of grassland productivity using Sentinel-2 remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*, *111*. https://doi.org/10.1016/j.jag.2022.102843

Food and Agriculture Organization. (2020). *A protocol for measurement, monitoring, reporting and verification of soil organic carbon in agricultural landscapes. A protocol for measurement, monitoring, reporting and verification of soil organic carbon in agricultural landscapes*. Rome: FAO. https://doi.org/10.4060/cb0509en

Hagolle, O., Huc, M., Desjardins, C., Auer, S., & Richter, R. (2017, December). MAJA Algorithm Theoretical Basis Document. Zenodo. https://doi.org/10.5281/zenodo.1209633

Hanrahan, L., Geoghegan, A., O'Donovan, M., Griffith, V., Ruelle, E., Wallace, M., & Shalloo, L. (2017). PastureBase Ireland: A grassland decision support system and national database. *Computers and Electronics in Agriculture*, *136*, 193–201. https://doi.org/10.1016/J.COMPAG.2017.01.029

Hill, M. J., Donald, G. E., Hyder, M. W., & Smith, R. C. G. (2004). Estimation of pasture growth rate in the south west of Western Australia from AVHRR NDVI and climate data. *Remote Sensing of Environment*, *93*(4), 528–545. https://doi.org/10.1016/J.RSE.2004.08.006

Huang, Q., Wang, X., Gao, Q., Carraro, A., Pezzuolo, A., & Marinello, F. (2025). How to assess the digitization and digital effort: A framework for Digitization Footprint (Part 1). *Computers and Electronics in Agriculture*, *230*, 109875. https://doi.org/10.1016/J.COMPAG.2024.109875

Hussain, I., Ching, K. B., Uttraphan, C., Tay, K. G., & Noor, A. (2025). Evaluating machine learning algorithms for energy consumption prediction in electric vehicles: A comparative study. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-94946-7

Jia, M., Guo, X., Zhang, L., Wang, M., Wang, W., Lu, C., et al. (2024). Mapping mangrove functional traits from Sentinel-2 imagery based on hybrid models coupled with active learning strategies. *International Journal of Applied Earth Observation and Geoinformation*, *130*. https://doi.org/10.1016/j.jag.2024.103905

Jin, B., & Xu, X. (2025). Machine Learning Green Gram Price Predictions. *Materials Circular Economy*, *7*(1), 18. https://doi.org/10.1007/s42824-025-00175-8

Lasko, K. (2022). Gap Filling Cloudy Sentinel-2 NDVI and NDWI Pixels with Multi-Frequency Denoised C-Band and L-Band Synthetic Aperture Radar (SAR), Texture, and Shallow Learning Techniques. *Remote Sensing*, *14*(17). https://doi.org/10.3390/rs14174221

Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., et al. (2020). Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 161–172. https://doi.org/10.1016/J.ISPRSJPRS.2020.02.013

Lobert, F., Holtgrave, A.-K., Schwieder, M., Pause, M., Vogt, J., Gocht, A., & Erasmi, S. (2021). Mowing event detection in permanent grasslands: Systematic evaluation of input features from Sentinel-1, Sentinel-2, and Landsat 8 time series. *Remote Sensing of Environment*, *267*. https://doi.org/10.1016/j.rse.2021.112751

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 2017-December, pp. 4766–4775). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85044542379&partnerID=40&md5=78f2bc16fd361e274004b0f78b3ef44a

Malek, Ž., Romanchuk, Z., Yashchun, O., & See, L. (2024). A harmonized data set of ruminant livestock presence and grazing data for the European Union and neighbouring countries. *Scientific Data*, *11*(1), 1136. https://doi.org/10.1038/s41597-024-03983-w

McSweeney, D., Delaby, L., O'Brien, B., Ferard, A., Byrne, N., McDonagh, J., et al. (2022). Dynamic algorithmic conversion of compressed sward height to dry matter yield by a rising plate meter. *Computers and Electronics in Agriculture*, *196*, 106919. https://doi.org/10.1016/J.COMPAG.2022.106919

Ogungbuyi, M. G., Guerschman, J., Fischer, A. M., Crabbe, R. A., Ara, I., Mohammed, C., et al. (2024). Improvement of pasture biomass modelling using high-resolution satellite imagery and machine learning. *Journal of Environmental Management*, *356*, 120564. https://doi.org/10.1016/J.JENVMAN.2024.120564

Pelton, R. E. O., Kazanski, C. E., Keerthi, S., Racette, K. A., Gennet, S., Springer, N., et al. (2024). Greenhouse gas emissions in US beef production can be reduced by up to 30% with the adoption of selected mitigation measures. *Nature Food*, *5*(9), 787–797. https://doi.org/10.1038/s43016-024-01031-9

Peng, J., Zeiner, N., Parsons, D., Féret, J.-B., Söderström, M., & Morel, J. (2023). Forage Biomass Estimation Using Sentinel-2 Imagery at High Latitudes. *Remote Sensing*, *15*(9). https://doi.org/10.3390/rs15092350

Peng, S., Wang, Z., Lu, X., & Liu, X. (2024). Hybrid inversion of radiative transfer models based on topographically corrected Landsat surface reflectance improves leaf area index and aboveground biomass retrievals of grassland on the hilly Loess Plateau. *International Journal of Digital Earth*, *17*(1). https://doi.org/10.1080/17538947.2024.2316840

Pinna, D., Pezzuolo, A., Cogato, A., Pornaro, C., Macolino, S., & Marinello, F. (2024). Applications of satellite platforms and machine learning for mapping and monitoring grasslands and pastures: A systematic and comprehensive review. *Smart Agricultural Technology*, *9*, 100571. https://doi.org/10.1016/J.ATECH.2024.100571

Pinna, D., Sozzi, M., Pornaro, C., Macolino, S., Pezzuolo, A., & Marinello, F. (2025). Modelling vegetation phenology in alpine grasslands of North-East Italy integrating microclimatic and meteorological Drivers: A 23-Years Time-Series analysis. *Computers and Electronics in Agriculture*, *232*, 110154. https://doi.org/10.1016/J.COMPAG.2025.110154

Prangel, E., Reitalu, T., Neuenkamp, L., Kasari-Toussaint, L., Karise, R., Tiitsaar, A., et al. (2024). Restoration of semi-natural grasslands boosts biodiversity and re-creates hotspots for ecosystem services. *Agriculture, Ecosystems & Environment*, *374*, 109139. https://doi.org/10.1016/J.AGEE.2024.109139

Pullanagari, R. R., Yule, I. J., Tuohy, M. P., Hedley, M. J., Dynes, R. A., & King, W. M. (2012). In-field hyperspectral proximal sensing for estimating quality parameters of mixed pasture. *Precision Agriculture*, *13*(3), 351–369. https://doi.org/10.1007/s11119-011-9251-4

Punalekar, S. M., Verhoef, A., Quaife, T. L., Humphries, D., Bermingham, L., & Reynolds, C. K. (2018). Application of Sentinel-2A data for pasture biomass monitoring using a physically based radiative transfer model. *Remote Sensing of Environment*, *218*, 207–220. https://doi.org/10.1016/J.RSE.2018.09.028

Qin, Q., Xu, D., Hou, L., Shen, B., & Xin, X. (2021). Comparing vegetation indices from Sentinel-2 and Landsat 8 under different vegetation gradients based on a controlled grazing experiment. *Ecological Indicators*, *133*. https://doi.org/10.1016/j.ecolind.2021.108363

Ramson, S. R. J., Leon-Salas, W. D., Brecheisen, Z., Foster, E. J., Johnston, C. T., Schulze, D. G., et al. (2021). A Self-Powered, Real-Time, LoRaWAN IoT-Based Soil Health Monitoring System. *IEEE Internet of Things Journal*, *8*(11), 9278–9293. https://doi.org/10.1109/JIOT.2021.3056586

Sa, R., Nie, Y., Chumachenko, S., & Fan, W. (2024). Biomass Estimation and Saturation Value Determination Based on Multi-Source Remote Sensing Data. *Remote Sensing*, *16*(12). https://doi.org/10.3390/rs16122250

Schulze, M., Spiller, A., & Risius, A. (2021). Do consumers prefer pasture-raised dual-purpose cattle when considering meat products? A hypothetical discrete choice experiment for the case of minced beef. *Meat Science*, *177*, 108494. https://doi.org/10.1016/J.MEATSCI.2021.108494

Shafer, G., & Vovk, V. (2008). A Tutorial on Conformal Prediction. *J. Mach. Learn. Res.*, *9*, 371–421.

Sibanda, M., Mutanga, O., & Rouget, M. (2015). Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. *ISPRS Journal of Photogrammetry and Remote Sensing*, *110*, 55–65. https://doi.org/10.1016/J.ISPRSJPRS.2015.10.005

Sinde-González, I., Gil-Docampo, M., Arza-García, M., Grefa-Sánchez, J., Yánez-Simba, D., Pérez-Guerrero, P., & Abril-Porras, V. (2021). Biomass estimation of pasture plots with multitemporal UAV-based photogrammetric surveys. *International Journal of Applied Earth Observation and Geoinformation*, *101*, 102355. https://doi.org/10.1016/J.JAG.2021.102355

Smith, H. D., Dubeux, J. C. B., Zare, A., & Wilson, C. H. (2023). Assessing Transferability of Remote Sensing Pasture Estimates Using Multiple Machine Learning Algorithms and Evaluation Structures. *Remote Sensing*, *15*(11). https://doi.org/10.3390/rs15112940

Smith, R. C. G., Matthew, A., Stewart, G., Steve, G., Duncan, W., Stefan, M., et al. (2011). Near real-time Feed On Offer (FOO) from MODIS for early season grazing management of Mediterranean annual pastures. *International Journal of Remote Sensing*, *32*(16), 4445–4460. https://doi.org/10.1080/01431161.2010.487078

Sozzi, M., Kayad, A., Gobbo, S., Cogato, A., Sartori, L., & Marinello, F. (2021). Economic Comparison of Satellite, Plane and UAV-Acquired NDVI Images for Site-Specific Nitrogen Application: Observations from Italy. *Agronomy*, *11*(11). https://doi.org/10.3390/agronomy11112098

Vovk, V. (2015). Cross-conformal predictors. *Annals of Mathematics and Artificial Intelligence*, *74*(1), 9–28. https://doi.org/10.1007/s10472-013-9368-4

World Food and Agriculture – Statistical Yearbook 2023. (2023a). *World Food and Agriculture – Statistical Yearbook 2023*. FAO. https://doi.org/10.4060/cc8166en

World Food and Agriculture – Statistical Yearbook 2023. (2023b). *World Food and Agriculture – Statistical Yearbook 2023*. FAO. https://doi.org/10.4060/cc8166en

Yang, Y., Zhang, M., Li, J., Wang, X., Yan, Y., Xin, X., & Xu, D. (2025). The performance of a canopy relative height model (CRHM) in natural grassland aboveground biomass estimation using unmanned aerial vehicle data. *Computers and Electronics in Agriculture*, *233*, 110137. https://doi.org/10.1016/J.COMPAG.2025.11013

5. Comprehensive discussion

5.1 How the research addressed its objectives and questions

The present research aimed to explore the potential of remote sensing and machine learning in improving the monitoring and management of grasslands and pastures. The work originated from the recognition that these ecosystems are crucial to ecological stability and agricultural sustainability. Yet, their dynamics are complex, variable, and challenging to quantify using traditional field-based methods alone. The overarching goal was therefore to develop integrative approaches capable of translating multisource data, ranging from satellite imagery to environmental variables, into accurate, transferable, and operational models of grassland productivity and phenology.

The first step of this work, reported in Chapter 2, involved a comprehensive analysis of existing sensing technologies and modelling approaches. Over the past two decades, the growing availability of multispectral and hyperspectral imagery, combined with open-access archives and cloud computing platforms, has revolutionised vegetation monitoring. The review of Chapter 2 confirmed that sensors such as MODIS, Landsat, Sentinel-2 and others, provide complementary strengths: MODIS offers valuable temporal depth, while Sentinel-2 and Landsat deliver the spatial detail required for local applications. Machine learning algorithms, particularly Regression-based approaches, Random Forest, Support Vector Machines, and Artificial Neural Networks among many others, have emerged as standard modelling approaches for deriving vegetation traits such as biomass, leaf area index, and chlorophyll content. However, despite these advances, some limitations persist. The integration of multiple sensors remains underexploited, uncertainty quantification remains rare, and the transferability of models across sites, seasons, and management systems is often limited. These observations framed the subsequent stages of the research and provided a clear rationale for testing new, uncertainty-aware and scalable approaches.

The following phase, reported in Chapter 3, focused on understanding how climatic, topographic, and soil factors shape the seasonal development of mountain grasslands in North-East Italy. By analysing a 23-year time series of MODIS-derived NDVI, in combination with detailed topographical and environmental datasets, the study reconstructed the temporal dynamics of vegetation productivity across alpine and pre-alpine landscapes and identified the main drivers governing their phenological patterns. The results revealed that grassland phenology is highly sensitive to elevation and slope, which modulate temperature and radiation regimes, thereby strongly influencing the onset, duration, and intensity of the growing season. Altitude and steepness were associated with shorter, less vigorous cycles, while higher solar radiation and temperature tended to prolong growth. Snowfall, on the other hand, delayed green-up and condensed the period of active vegetation. Beyond their influence on phenological dynamics, the topographical and structural characteristics of mountain grasslands also play a decisive role in determining their management feasibility. Steep slopes, variable field morphology, and limited accessibility not only shape vegetation patterns but also constrain the type and intensity of agricultural practices that can be realistically implemented. To address this aspect, a complementary analysis was conducted (Annex 1) to evaluate the mechanisability potential of Alpine and pre-Alpine pastures in North-East Italy. By integrating morphological variables such as slope and altitude with accessibility indicators including distance to roads, field size, and shape complexity, the study developed a spatial framework to quantify the feasibility of mechanised operations in these landscapes. The results highlighted that steep terrain and poor accessibility are the predominant factors limiting mechanisation, creating a mosaic of productive and marginal areas within the same region.

Driven by the ambition to translate remote sensing and modelling advances into practical farm-scale management tools for the grasslands and pastures of North-East Italy, the research then focused on the quantitative estimation of above-ground biomass at very high spatial and temporal resolution. In this phase, as reported in Chapter 4, probabilistic machine learning models were applied to satellite imagery from Sentinel-2 and PlanetScope platforms to derive biomass estimates that support operational management. Gaussian Process Regression proved particularly suitable because of its flexibility in modelling non-linear relationships and its intrinsic capacity to represent predictive uncertainty. The analysis demonstrated that biomass can be estimated with high accuracy when data from multispectral sensors are properly processed and when predictions are aggregated at the field scale. Indeed, aggregation was shown to significantly reduce random variability and yield uncertainty intervals narrow enough for practical decision-making. The use of conformal prediction techniques further strengthened the reliability of these estimates, offering statistically valid and interpretable confidence ranges. The comparison between Sentinel-2 and PlanetScope imagery revealed that both sensors perform well for biomass estimation. Yet, they cater to slightly different needs: while Sentinel-2 provides robust coverage and cost-effective monitoring, PlanetScope's daily revisit rate and finer resolution make it more suitable for those applications in which high-frequency output is required.

In parallel, part of the methodological framework was applied to a more intensively managed context, with the aim of assessing its robustness and adaptability beyond semi-natural grasslands. This complementary work focused on Zea mays, the most widespread forage crop in the region, to test how the modelling approach developed for grassland biomass estimation would perform under conditions of high input intensity and well-defined phenological cycles. By using Sentinel-2 and PlanetScope imagery, the study assessed the capacity of these platforms to capture within-field variability in yield, moisture, and nutritional quality. The results indicated that spectral responses were strongly dependent on the crop's growth stage: early vegetative and pre-flowering phases provided the best performance for yield prediction, while pre-harvest conditions were more informative for moisture and compositional attributes, such as fibre and starch content. These findings confirmed that the spectral modelling principles established for grasslands are adaptable to intensively managed cropping systems.

5.2 Limitations and future directions

While the present research achieved its main objectives and advanced the integration of remote sensing and machine learning for grassland monitoring, several limitations emerged that provide valuable guidance for future work. These limitations can be grouped into three main categories: data availability and scale, methodological constraints, model transferability and operational integration. The accuracy and temporal continuity of remote-sensing-based monitoring remain inherently limited by sensor characteristics and data availability. Despite the complementary strengths of MODIS, Sentinel-2, and PlanetScope, trade-offs persist between spatial detail, temporal frequency, and data consistency. Cloud cover, sun-sensor geometry effects, and snow masking particularly affect optical observations in alpine and subalpine environments, leading to seasonal gaps that can influence trend analyses and phenological reconstructions. Future studies should explore data fusion approaches combining multispectral, synthetic aperture radar (e.g., Sentinel-1), and meteorological information to ensure more complete temporal coverage and improved resilience to atmospheric and illumination constraints.

Although the models developed in this thesis demonstrated strong predictive performance, several methodological challenges remain. The Gaussian Process Regression framework, while powerful, can become computationally demanding with large datasets or multi-temporal applications. Dimensionality reduction and sparse approximations may be necessary to ensure scalability for continuous regional monitoring. In addition, the field data used for calibration, though extensive, were necessarily limited in spatial coverage and sampling frequency. Expanding ground-truth campaigns to include broader gradients of vegetation composition, soil type, and management intensity would help refine model robustness and reduce site-specific bias. Furthermore, future work could integrate physiological or radiative transfer models to better link spectral responses with plant functional traits, improving interpretability and generalisation across species and regions.

The models developed here were designed for high interpretability and adaptability, but their transfer to different ecological or management contexts remains challenging. Differences in spectral behaviour, soil background, and canopy structure can limit cross-site applicability. To overcome this, future efforts should focus on harmonising input data through standardised pre-processing pipelines and domain adaptation techniques, enabling cross-regional applications. At the operational level, integrating the probabilistic modelling framework into decision-support systems (DSS) would represent a crucial step toward practical use. This would allow uncertainty-aware biomass estimates and phenological indicators to directly guide management practices such as grazing rotation, cutting schedule, or fertiliser application.

Beyond methodological aspects, future research should also address the socio-ecological dimension of grassland monitoring, bridging technological innovation with the needs of farmers, cooperatives, and policymakers. The integration of satellite-derived indicators into agri-environmental schemes or carbon accounting frameworks (e.g. MRV systems) would strengthen the contribution of remote-sensing-based models to sustainability assessment and climate mitigation policies.

6. Conclusions

Grasslands and pastures are essential for sustaining livestock production, maintaining biodiversity, and providing key ecosystem services. Yet, their strong spatial and temporal variability makes them difficult to monitor with traditional approaches. This research addressed this challenge by developing and testing integrative frameworks that combine remote sensing and machine learning to monitor productivity and phenology in the grasslands of North-East Italy. The results confirmed that the combined use of satellite observations and probabilistic modelling represents a robust and scalable strategy for assessing vegetation dynamics and estimating above-ground biomass. Multi-scale analyses revealed how climatic, topographic, and soil factors jointly control the growth and resilience of mountain grasslands, while mechanisability assessments clarified how terrain structure limits the feasibility of management in Alpine landscapes. At finer scales, Gaussian Process Regression applied to Sentinel-2 and PlanetScope imagery demonstrated high predictive accuracy and reliable uncertainty estimation, providing operationally useful information for precision grassland management. Overall, the outcomes of this research contribute to advancing both scientific understanding and practical management of grasslands under changing environmental conditions. By linking ecological processes, management feasibility, and predictive modelling, the study supports the transition toward data-driven, climate-smart pasture systems. Nevertheless, some limitations remain. The models developed here require further validation across different regions, vegetation types, and management regimes to ensure their generalisation. Expanding ground observations and integrating complementary data sources, such as radar and proximal sensing, would improve robustness and continuity. The next step should be to translate these advances into accessible, farmer-friendly decision-support tools capable of providing timely and interpretable information for adaptive management. In conclusion, this thesis demonstrates that the integration of remote sensing and machine learning can transform grassland monitoring from a simple observation-based assessment into an operational component of sustainable agricultural planning. Strengthening the connection between scientific innovation, field practice, and policy design will be key to promoting resilient and efficient forage systems in the years to come.

Annex 1 - A Mechanisability Index to Evaluate the Potential of Alpine Pastures and Meadows in North-East of Italy

A1.1 Abstract

Management of pastures and meadows in the Alpine area of Northern Italy often implies mechanizable practices. In order to optimise and achieve economic sustainability of such practices, agricultural machinery implementation is needed. Thus, the present work introduces a methodology aimed to assess the potential for agricultural mechanisation of pastures and meadows based on rational and quantitative criteria. In particular, a mechanisability index is here proposed, based on GIS analysis of landscape parameters as mean and maximum slope, size, mean altitude, distance from the nearest road and shape regularity of the fields. For the scope, 11.093 fields in the Province of Trento were identified and considered to analyse their mechanisability potential. The results give evidence of a medium to low mechanisability potential of pastures and meadows in the area, the most limiting factors being elevated slopes, altitude and distance from roads. The methodology here reported might be applied to other mountainous areas, helping the decision-making processes of institutions and farmers.

A1.2 Introduction

Sustainable and innovative practices have become increasingly important as livestock systems strive to become smarter (Tassinari et al., 2021) and more efficient. In fact, in the last years, livestock production is changing fast towards smart systems (Cillis et al., 2018), driven by the rapid pace of technological advancements as Internet of Things (IoT) and data management (Bovo et al., 2020; Pinotti et al., 2021). Livestock production needs to pay more attention to limit the use of natural resources expressed as the footprint per product, for example water footprint, land (arable or total land) footprint (Pinotti et al., 2021). However, livestock farming impacts can range widely in terms of resource depletion, degree of intensification, socio-economic context, and many other aspects that influence the sustainability of the livestock production sector (Poffenbarger et al., 2017). To sustain the production of domestic animals, pasture and meadow management is essential for ensuring sufficient fodder quantity and quality. In addition, several nations have used the incorporation of cattle into cropping systems to advance agricultural sustainability on an ecological level (Garrett et al., 2017; Moraine et al., 2017; Poffenbarger et al., 2017). Livestock farming systems based on pastures and meadows are still important in many regions all around the world, among which the countries of the European Mediterranean Basin where such livestock systems play an important role in the management and conservation of large High Nature Value farmland, mostly located in less productive areas (Cortner et al., 2019). Among these areas, in the alpine region of Northern Italy, the livestock industry plays an economic and cultural pivotal role. Cattle sheep, and goats are reared in the high-altitude areas, utilising the grasslands that cover the valleys and slopes. However, the harsh winters, characterised by heavy snowfall and freezing temperatures, render it nearly impossible for livestock to graze on fresh vegetation. This is where haymaking emerges as a vital practice, ensuring that animals have access to adequate nutrition throughout the year. At present, extensively utilised hay meadows are widespread in the alpine environment, where farmers employ heavy machinery to store hay in modern haylofts. The hay is typically baled into large rounded or square shapes weighing up to 1,400 kg. The decline of traditional hay-making practices has been expedited by the introduction of silage during the 20th century. Silage, whether stored in permanent silos or round bales, offers farmers greater flexibility in adapting to environmental uncertainties (T. Nowacki, 1978). By way of example, production of grass silage in Austria increased significantly over time. In 1988, only 77 grass silage bales were produced, while a decade later, that number rose to 4.5 million bales. By 2006, it reached 5.6 million silage bales. In mountainous regions, silage is utilised on 95% of pastures, particularly on the so-called "best soils" that cover 70% of the area (Špulerová et al., 2019). Even though certain regions with extreme natural conditions or remote areas have managed to preserve traditional hay-making structures, silage, and modern haymaking practices necessitate modern machinery. Thus, it is important to consider the accessibility of meadows and their suitability with respect to implementation of mechanised operations. Although an expansion of mechanisation has been obtained, mostly thanks to the improvement of machine performances, an assessment of the real potential of alpine meadow mechanisation is still not well defined for the Italian North-Eastern Alpine region. Traditional mechanisation indices are typically calculated ex-post. By way of example, Nowacki (W. Nowacki, 1978) proposed estimation of a mechanisation index on the basis of the time related proportion between machine, human and animal work in field operations. Other approaches consider. Other works introduce implementation of life cycle assessment in order to quantify the overall sustainability of a mechanization process (Boscaro et al., 2018; Davide Boscaro et al., 2015). Other authors recommend the ratio between the mechanised operation and the total cultivated area as a mean to be implemented in the description of the amount of mechanisation in a cultivated area. A mechanisation index was created by Zangeneh et al. (2010) based on an artificial neural network model to measure how much a farm's amount of machine work deviates from values at the regional level. Based on the slope local length of auto-correlation, Maheshwari and Tripathi (2019) evaluated a novel method for calculating a mechanisation index that accounts for the strength of machines and animals as well as the time it takes for humans, machines, and animals to carry out cultivation operations. Analysis from Geographic Information Systems (GIS) is essential to the management of agriculture. GIS enhances land suitability evaluation through dedicated spatial analysis [16]. Previous research revealed that geospatial data might be used to assess the prospective usage and mechanical accessibility of agricultural land. For instance, Jasinski (2005) suggested a method based on five environmental and landscape criteria (slope, soil type, precipitation, distance from roadways, and land-cover type before conversion) to evaluate the possibility of Brazilian land-use conversion to mechanised agricultural farming. Cogato et al. (2020). developed a mechanisability index for the Italian viticultural regions, the following parameters were rated from 0 (not mechanisable) to 100 (highly mechanisable): mean slope, block shape, length-width ratio, headland size, training system, and row spacing; the mechanisability of each vineyard was then calculated by multiplying the ranking of each parameter. The approaches discussed address the physical landscape perspective of land suitability for agricultural mechanisability. To the best of our knowledge, there is not a method for assessing mechanisability based on precise, quantifiable factors related to machinery accessibility, especially related to mountainous marginal areas as most of pastures and meadows in the alpine areas of northeast Italy are.

A1.3 Materials and Methods

The present study focuses on an area located in the Province of Trento (NUTS 2, ITH20), chosen as a target in order to test a mechanisability analysis approach. For the scope, 11093 fields classified as pastures or grasslands were considered and analysed in the study (Fig. A1.1). Fields land cover labelling was manually validated and made available by the province land authorities in a public database (https://siat.provincia.tn.it). Selected fields underwent a multistage GIS analysis, in order to include the following parameters: mean slope, size of the field (m^2), mean altitude (masl), distance from the nearest road (m), and convexity value to evaluate the regularity of the field shapes.

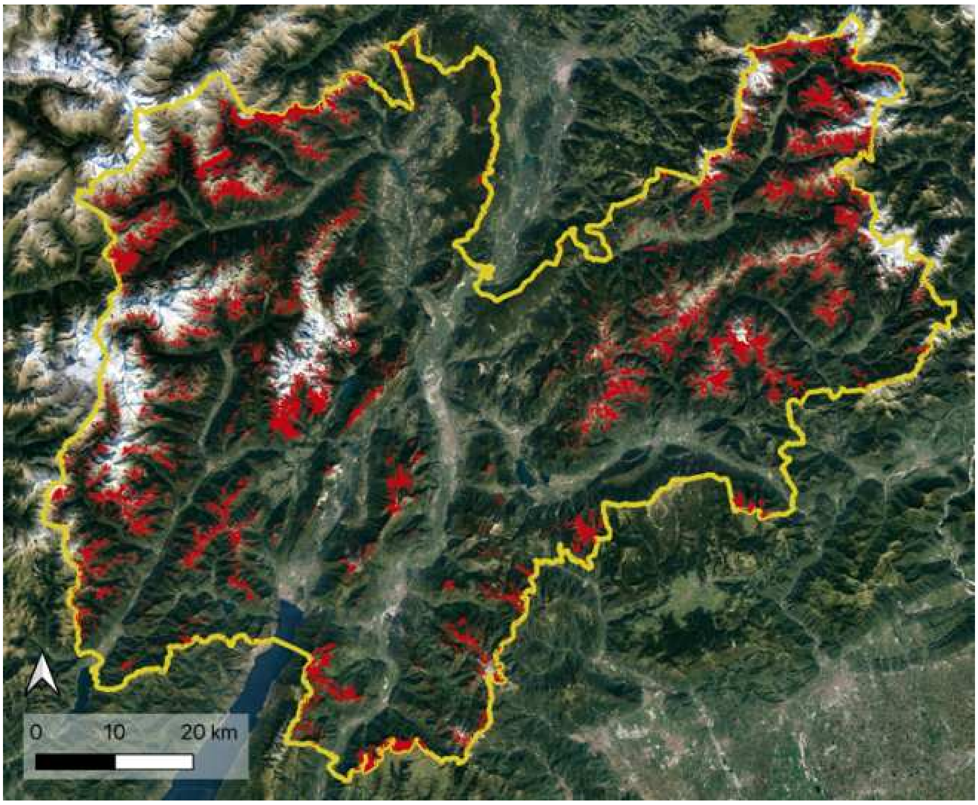


Fig. A1.1 Aerial image of the study area, the yellow line represents the border of the Province of Trento (TN); the red areas represent pastures and grasslands selected for the analysis reported in the present manuscript.

The analysis was performed in the QGIS environment (QGIS 3.28 Firenze version), an open-source geoinformation software that enables management and analysis of geographic data. Once the database was imported into Excel, each parameter was graded from zero (mechanisation no or hardly possible) to 100 (highest mechanisation potential), depending on how much it contributed to mechanisability. The ranking of the individual factors (fi) was then multiplied to determine the mechanisability of each field (M.I.) (1).

$$M.I. = \prod_{i=1}^{n} f_i \qquad (1)$$

To avoid redundancy of implemented parameters, a correlation analysis was carried out among the coefficients selected to calculate the mechanisability index. A simple regression analysis was additionally carried out in order to evaluate the weight of each factor on the index outcomes. JMP Pro 17 statistical program was used to conduct the statistical analysis.

Slope

The slope of the field has a significant impact on how agricultural tractors should be routed and manoeuvred (Hameed et al., 2013). A Digital Terrain Model (DTM) with a spatial resolution of 10 meters was used to calculate the average and the maximum field slope (Statuto et al., 2017). The DTM is available online at http://tinitaly.pi.ingv.it (accessed on June 2023) and was built by the Istituto Nazionale di Geofisica e Vulcanologia (INGV). For pasture mechanisation, the slope is frequently the limiting element and one that cannot be changed. Self-levelling and tilt-correction tools make it possible to work on steep terrain (Fortune and Tawanda, 2013; Hunter, 1993), although stability risks can appear when the slope is greater than 10. Six classes were selected based on the combination of mean and maximum slope values, as shown in Table A1.1.

Table A1.1 Categorisation by mechanical accessibility of the slopes in the pastures of the province of Trento.

Slope (%)	Class Ranking (%)
Mean < 5 and Max < 10	100
Mean <10 and Max <20	75
Mean <15 and Max <30	60
Mean <20 and Max <40	40
Mean <30 and Max <40	20
Max > 40	0

Size

The size of the field has a great impact on the mechanisability of meadows. Even though also very small fields might be in principle managed through mowing, conditioning and silage operations, it is difficult to keep acceptable efficiency when fields dimensions are small and transfer times exceed operation time. Thus, if the area do not exceeds a given threshold (which might change as a function of the distance from the farm centre) it might be not economically convenient for the farm owner to adopt mechanised technologies. The area (m^2) of the selected fields was calculated using the field calculator tool on QGIS. Size values were eventually ranked following the classification reported in Table A1.2.

Table A1.2 Categorisation by mechanical accessibility of the size (m^2) of pastures in the province of Trento.

Size (m^2)	Class Ranking
>10000	100
5000÷10000	90

1000÷5000	70
<1000	50

Altitude

The altitude of the field has a significant influence on the amount of biomass produced and on the accessibility by agricultural machines. A Digital Terrain Model (DTM) with a spatial resolution of 10 meters was used to calculate the average field altitude. The DTM is available online at http://tinitaly.pi.ingv.it (accessed on June 2023) and was built by the Istituto Nazionale di Geofisica e Vulcanologia (INGV). Even though the relation between the mean altitude of the field and its accessibility might not be linear, an assumption was made considering that the higher is the altitude of a field, the lower will be its mechanisability. Thus, four altitude classes were chosen, as shown in Table A1.3.

Table A1.3 Categorisation by mechanical accessibility of the altitude (masl) of pastures in the province of Trento.

Altitude (masl)	Class Ranking (%)
<1000	100
1000÷1500	90
1500÷2000	80
>2000	75

Distance from the nearest road

The distance of a field from the nearest road is necessarily a significant limiting factor for its mechanisation. Indeed, even though modern machinery might be able to cross small offroad trails, a road network is necessary to cover relatively long distances and rapidly reach the meadows. Estimation of the distance of each selected area from the nearest road, was allowed by a public database of the Province of Trento road network (https://download.geofabrik.de/europe/italy.html, accessed on June 2023). Then, the distance value (m) for each field from the closest road was quantfied, taking advantage of a specific tool available in the QGIS package, namely "Distance to the nearest hub (lines to hub)". The distances were then divided into four classes as shown in Table A1.4.

Table A1.4 Categorisation by distance from the nearest road of pastures in the province of Trento.

Distance from nearest road (m)	Class Ranking (%)

<50	100
50÷100	80
100÷150	60
>150	50

Shape regularity

The shape of the fields holds great importance in the suitability of a field to be mechanised. Fields with regular shapes are easier to mow or to harvest, as the necessary number of turns and manoeuvres are significantly lower. In order to assess the regularity of the shape of the fields included in the analysis, the presence of convex portions of the borders was considered (see for instance Fig. A1.2). Convexity is defined as the ratio of perimeters of the convex hull and the original contour, where the convex hull is the minimal convex covering of an object. The measure takes a maximum value of 1 for a circle, objects that have complicated and irregular boundaries have lower convexity (Peura and Iivarinen, 1997). Three classes were then selected, based on the convexity value, as shown in Table A1.5.

Fig A1.2 An example of an area characterized by an highly convex shape.

Table A1.5 Categorisation by shape regularity (convexity) of pastures in the province of Trento.

Shape regularity (convexity)	Class Ranking (%)
>0.867	100
0.744÷0.867	80
<0.744	60

A1.4 Results and Discussion

Preliminary Analysis of the Parameters

A preliminary analysis of the contributed parameters was conducted to verify the significance of the parameters in the determination of the calculated index and the absence of strong correlations among the selected variables. A correlation matrix containing the five selected parameters was created, and the Row-wise method estimates the correlations. The resulting coefficients of determination (R^2) are sufficiently low (between 0.0003 and 0.174) to exclude significant correlation as shown in Tab. A1.6 and in the scatterplot in Fig. A1.3.

Table A1.6 Coefficients of determination (R^2) calculated among the contributing factors.

Correlations	Size	Road dist.	Slope	Altitude	Shape
Size	1.000	0.0003	0.001	0.009	0.112
Road dist.	0.0003	1.000	0.025	0.174	0.007
Slope	0.0001	0.025	1.000	0.171	0.006
Altitude	0.009	0.174	0.171	1.000	0.066
Shape	0.112	0.007	0.006	0.066	1.000

Level of Mechanisability of Pastures in the Province of Trento

The index distribution obtained over the fields selected in the study area is reported in Fig. 4. As clearly visible, the mechanisability of the fields in this is generally very low, 52% of the fields obtained a Mechanisability Index score below 0.1 and there is a clear trend with very few fields obtaining good scores; only 0.1% of the fields obtained a score above 0.5. These results are not surprising as the area and the fields selected for the present study are in the alpine region and mainly along the slopes of the mountains, where mechanisation is always hard to carry out.

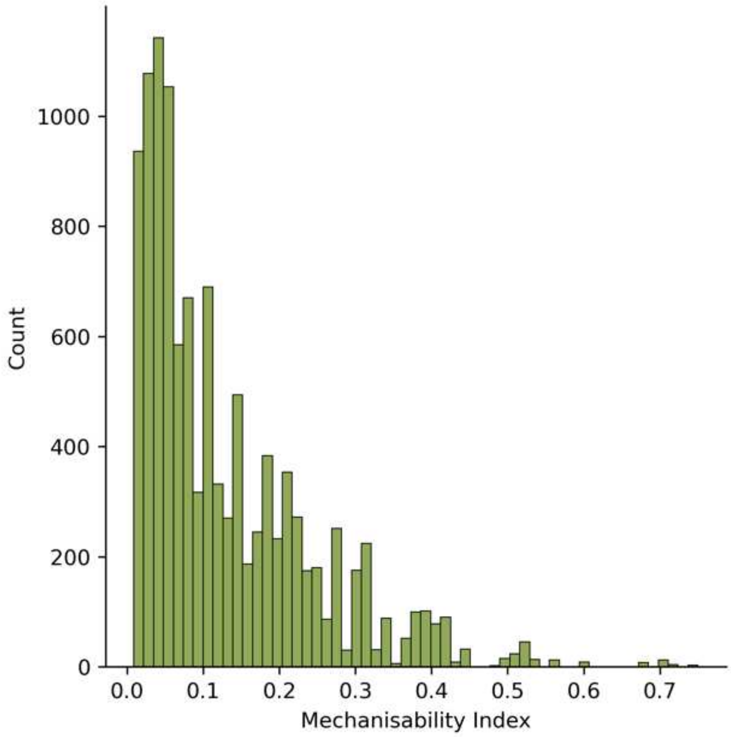


Fig. A1.3 Histogram representing the distribution of the Mechanisability Index values calculated on the 11093 fields in the study area.

Main factors influencing the mechanisability of Pastures in the Province of Trento

The results reported in Fig. A1.3 give a clear evidence of the actual condition: pastures of the Province of Trento are in general exhibiting a low mechanisability potential. To evaluate how the five different parameters influenced the calculation of the mechanisability index, the parameters were screened to evaluate their importance in determining the mechanisability index by means of a Response Screening platform on JMP Pro. Through the Response Screening tool, we automatised the process of testing each response of the Mechanisability Index against all the five factors performing simple linear regression. As shown in Table A1.7, all factors significantly influenced the index outcomes, with negligible p-values (<10-9), but the difference among the weight of their effects was clear.

Indeed, as expected, given the mountainous nature of the study area, the slope of the fields was very high as shown in Fig. A1.4 and this it was the most important factor influencing the low mechanisability of pastures, followed by altitude and distance from the nearest road.

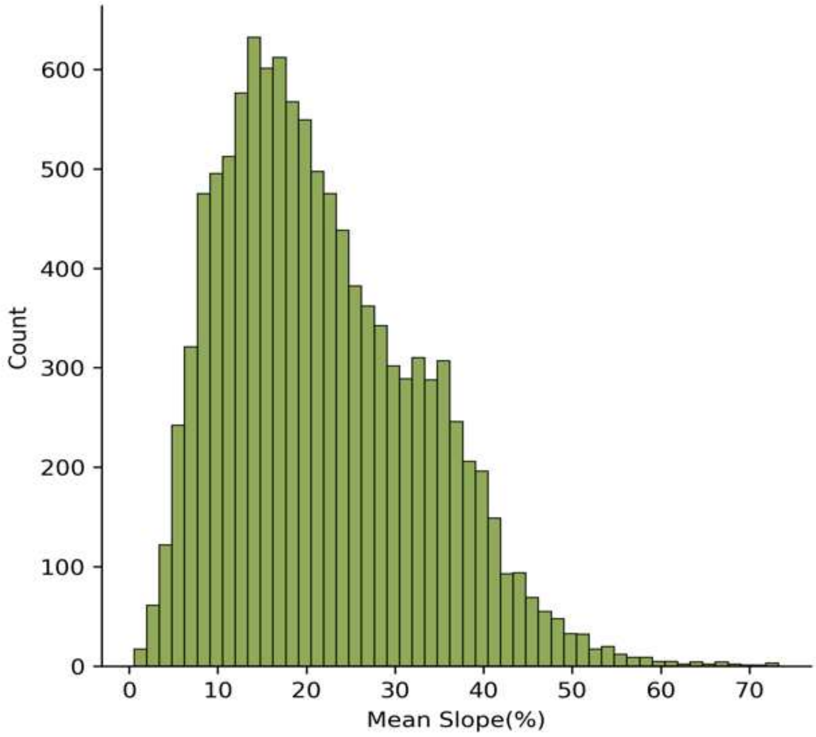


Fig A1.4 Histogram representing the distribution of the mean slope values calculated on the 11093 fields in the study area.

Despite being statistically significant, the effect of the shape (convexity) and size of the field was way less important in defining the index values.

Table A1.7 Response of each factor to the Mechanisability Index

Response to Mechanisability Index	
Variable	R^2
Slope	0.484
Altitude	0.323
Road Distance	0.161
Shape	0.064
Size	0.005

A1.5 Conclusions

The study created a GIS-based methodology to automatically classify alpine pastures and meadows according to their suitability for mechanisation. The developed index was based on five factors: mean slope, mean altitude, field size, shape regularity, and distance from the nearest road. Among these factors, the slope is the most limiting factor for the accessibility of pastures and meadows, followed by altitude and distance from the nearest road. Generally, the mechanisability of pastures and meadows in the Province of Trento is very low due to the mountainous characteristics of the area, but knowing the potential of each field to be mechanised might influence managerial decisions. The potential of this methodology to be applied to other alpine areas or to be modified to evaluate other characteristics of fields is very promising. It might be used for institutions or policymakers to evaluate the effectiveness of agricultural policies, but also by companies that manage large area, in order to optimise the organization of the farm fleet and the priorities in the acquisition of new machinery.

A1.6 References

Boscaro, D., Pezzuolo, A., Grigolato, S., Cavalli, R., Marinello, F., Sartori, L., 2015. Preliminary analysis on mowing and harvesting grass along riverbanks for the supply of anaerobic digestion plants in North-Eastern Italy. Journal of Agricultural Engineering 46, 100–104. https://doi.org/10.4081/jae.2015.465

Boscaro, D., Pezzuolo, A., Sartori, L., Marinello, F., Mattioli, A., Bolzonella, D., Grigolato, S., 2018. Evaluation of the energy and greenhouse gases impacts of grass harvested on riverbanks for feeding anaerobic digestion plants. J Clean Prod 172, 4099–4109. https://doi.org/10.1016/j.jclepro.2017.02.060

Bovo, M., Benni, S., Barbaresi, A., Santolini, E., Agrusti, M., Torreggiani, D., Tassinari, P., 2020. A Smart Monitoring System for a Future Smarter Dairy Farming, in: 2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor). pp. 165–169. https://doi.org/10.1109/MetroAgriFor50201.2020.9277547

Cillis, D., Pezzuolo, A., Marinello, F., Sartori, L., 2018. Field-scale electrical resistivity profiling mapping for delineating soil condition in a nitrate vulnerable zone. Applied Soil Ecology 123, 780–786. https://doi.org/10.1016/J.APSOIL.2017.06.025

Cogato, A., Pezzuolo, A., Sørensen, C.G., De Bei, R., Sozzi, M., Marinello, F., 2020. A GIS-based multicriteria index to evaluate the mechanisability potential of Italian vineyard area. Land (Basel) 9, 1–17. https://doi.org/10.3390/land9110469

Cortner, O., Garrett, R.D., Valentim, J.F., Ferreira, J., Niles, M.T., Reis, J., Gil, J., 2019. Perceptions of integrated crop-livestock systems for sustainable intensification in the Brazilian Amazon. Land use policy 82, 841–853. https://doi.org/10.1016/J.LANDUSEPOL.2019.01.006

Fortune, T.C.F., Tawanda, D., 2013. An Assessment of Agricultural Mechanization Index and Evaluation of Agricultural Productivity of some Fast Track Resettlement Farms in Bindura District of Mashonaland Central Province: Zimbabwe. International Journal Of Social Sciences \& Interdisciplinary Research 2.

Garrett, R.D., Niles, M.T., Gil, J.D.B., Gaudin, A., Chaplin-Kramer, R., Assmann, A., Assmann, T.S., Brewer, K., de Faccio Carvalho, P.C., Cortner, O., Dynes, R., Garbach, K., Kebreab, E., Mueller, N., Peterson, C., Reis, J.C., Snow, V., Valentim, J., 2017. Social and ecological analysis of commercial integrated crop livestock systems: Current knowledge and remaining uncertainty. Agric Syst 155, 136–146. https://doi.org/10.1016/J.AGSY.2017.05.003

Hameed, I.A., Bochtis, D., Sørensen, C.A., 2013. An optimized field coverage planning approach for navigation of agricultural robots in fields involving obstacle areas. International Journal of Advanced Robotic Systems 10. https://doi.org/10.5772/56248

Hunter, A.G.M., 1993. A review of research into machine stability on slopes. Safety science 16, 325–339.

Jasinski, E., Morton, D., DeFries, R., Shimabukuro, Y., Anderson, L., Hansen, M., 2005. Physical landscape correlates of the expansion of mechanized agriculture in Mato Grosso, Brazil. Earth Interact 9, 1–18. https://doi.org/10.1175/EI143.1

Maheshwari, T.K., Tripathi, A., 2019. Quantification of Agricultural Mechanization Status for Etawah District of Uttar Pradesh, India. Int J Curr Microbiol Appl Sci 8, 659–666. https://doi.org/10.20546/ijcmas.2019.805.077

Moraine, M., Duru, M., Therond, O., 2017. A social-ecological framework for analyzing and designing integrated crop-livestock systems from farm to territory levels. Renewable Agriculture and Food Systems 32, 43–56. https://doi.org/10.1017/S1742170515000526

Nowacki, T., 1978. Methodology Used by ECE Countries in Forecasting Mechanization Developments. Geneva, Switzerland.

Nowacki, W., 1978. Methodology Used by ECE Countries in Forecasting Mechanization Developments.

Peura, M., Iivarinen, J., 1997. Efficiency of simple shape descriptors. Proceedings of the Third International Workshop on Visual Form 443–451.

Pinotti, L., Luciano, A., Ottoboni, M., Manoni, M., Ferrari, L., Marchis, D., Tretola, M., 2021. Recycling food leftovers in feed as opportunity to increase the sustainability of livestock production. J Clean Prod 294, 126290. https://doi.org/10.1016/J.JCLEPRO.2021.126290

Poffenbarger, H., Artz, G., Dahlke, G., Edwards, W., Hanna, M., Russell, J., Sellers, H., Liebman, M., 2017. An economic analysis of integrated crop-livestock systems in Iowa, U.S.A. Agric Syst 157, 51–69. https://doi.org/10.1016/J.AGSY.2017.07.001

Špulerová, J., Kruse, A., Branduini, P., Centeri, C., Eiter, S., Ferrario, V., Gaillard, B., Gusmeroli, F., Jurgens, S., Kladnik, D., Renes, H., Roth, M., Sala, G., Sickel, H., Sigura, M., Štefunková, D., Stensgaard, K., Strasser, P., Ivascu, C.M., Öllerer, K., 2019. Past, present and future of hay-making structures in Europe. Sustainability (Switzerland) 11. https://doi.org/10.3390/su11205581

Statuto, D., Cillis, G., Picuno, P., 2017. Using historical maps within a GIS to analyze two centuries of rural landscape changes in southern Italy. Land 6. https://doi.org/10.3390/land6030065

Tassinari, P., Bovo, M., Benni, S., Franzoni, S., Poggi, M., Mammi, L.M.E., Mattoccia, S., Di Stefano, L., Bonora, F., Barbaresi, A., Santolini, E., Torreggiani, D., 2021. A computer vision approach based on deep learning for the detection of dairy cows in free stall barn. Comput Electron Agric 182, 106030. https://doi.org/10.1016/J.COMPAG.2021.106030

Zangeneh, M., Omid, M., Akram, A., 2010. Assessment of agricultural mechanization status of potato production by means of artificial Neural Network model. Aust J Crop Sci 4, 372–377.

Annex 2 - Assessing silage maize yield and quality variability using NIRS, Sentinel-2, and PlanetScope multispectral imagery: a precision agriculture approach

A2.1 Abstract

The current advances in remote sensing and the availability of high-resolution and high-frequency multispectral images from satellite platforms can enhance the implementation of Precision Agriculture (PA). This study used vegetation indices (VIs) derived from Sentinel-2 and PlanetScope images and linear regression models to assess silage maize yield (Zea Mays), moisture, and qualitative spatial variability regarding Neutral Detergent Fiber (NDF), Acid Detergent Fiber (ADF), and starch content at the field scale. Maize was harvested by a combined harvester equipped with a yield monitor and a NIRS sensor on a 13-ha study field in North-East Italy, which collected more than 10,000 georeferenced yield, moisture, and quality observation points. 11 Sentinel-2 and 28 PlanetScope images were collected, covering the whole cultural cycle, from April to August 2020). Green Normalized Difference Vegetation Index (GNDVI) and Normalized Difference Vegetation Index (NDVI) were calculated for each image. Multiple regression models were tested for yield, moisture, NDF, ADF, and starch content. The capabilities of models to assess crop variability in terms of R2 were evaluated throughout the growing season. PlanetScope GNDVI was the best predictor of yield variability very early in the season, 63 days after sowing (R2=0.33). PlanetScope GNDVI provided the highest R2 value of 0.48 for crop moisture content at harvest day while ADF, NDF and starch content variability was best predicted by PlanetScope NDVI with R2 between 0.13 and 0.23. Sentinel-2 and PlanetScope VI's showed similar predictive capabilities, with an early season peak for yield and quality parameters between 63 and 79 days after sowing. Based on the results, there is a potential for assessing crop moisture variability in the proximity of harvest, which is crucial for the ensilage process, while the early season yield and quality assessment can provide major information for precise fertilization management.

A2.2 Introduction

Understanding within-field yield variability is vital for farmers, as it helps refine management strategies, improve profitability, assess land rental values, and facilitate insurance evaluations (Kross et al., 2015). The cornerstone of precision agriculture (PA) is yield monitoring, which is foundational for defining management zones (MZ) based on past yield data. Creating MZs necessitates a thorough understanding of soil properties and a record of prior management practices, while yield monitoring assesses the overall effectiveness of these practices by the season's end. Satellite remote sensing (RS) provides information on various crop and vegetation indicators, such as crop health (Kogan et al., 2012), biomass estimation (Dusseux et al., 2015), soil moisture levels (Baldwin et al., 2017) and crop phenological stages (An et al., 2018b), which are crucial for agricultural and environmental management. Despite the widespread application of these technologies on regional and national levels, their use at the field level is less common due to high costs, the focus of research funding on overall production, and challenges with obtaining accurate ground-truth data (Lobell et al., 2015). Traditional methods for crop yield monitoring, are often impractical for large fields due to their time-consuming, labour-intensive, and destructive nature. Modern yield monitoring systems, which can be mounted on harvesters, are widespread but provide yield maps only at harvest. The rapid advancements in RS technologies have sparked interest in using satellite and aerial data to assess within-field variability and predict crop yields throughout the growing season (Kayad Ahmed G. AND Al-Gaadi, 2016; Mokhtari et al., 2018; Sartore et al., 2023). Developing empirical relationships between vegetation indices (VIs) and crop yields offers a straightforward and practical method to assess within-field variability (W. Liu et al., 2024; Sibley et al., 2014). However, these empirical models often face limitations when applied across different fields or growing seasons due to spatial and temporal constraints (Báez-González et al., 2002; Lobell, 2013). Research on predicting yields within commercial maize fields has advanced significantly due to the widespread use of yield monitors that track mass flow, moisture content, and geo-coordinates in real time (Deines et al., 2021; Kross et al., 2020). These yield maps help identify areas of varying productivity within a field and, when analysed over several years, can define yield stability zones (Kharel et al., 2019). Research on predicting corn (Zea mays L.) grain yield and its variability across space and time has been extensive (Maestrini and Basso, 2018; Salazar et al., 2008; Xu et al., 2019). For example, Lobell et al. (Lobell et al., 2015) developed a yield prediction model using Landsat 5 and Landsat 7 satellite imagery combined with weather data and crop simulations, achieving an overall R^2 of 0.35 for corn yields across different U.S. fields. In another study, Shanahan et al. (Shanahan et al., 2001) found that GNDVI, derived from aircraft imagery during the mid-grain filling stage, showed strong correlations (0.7 to 0.92) with corn yield in plot experiments. Another study conducted in Brazil and the U.S. used Sentinel-2 imagery to predict corn yield at the field level, revealing that NDRE, GNDVI, and NDVI effectively predicted yield variability, with an R^2 of 0.32 for a universal corn yield estimation model (Schwalbert et al., 2018). Although the research on the estimation of corn yield is extensive, there is not much knowledge on the capabilities of satellite-based vegetation indices estimating moisture and qualitative parameters of crops. In this study, we test the capabilities of NDVI and GNDVI indices calculated from Sentinel-2 and PlanetScope platforms to assess within-field corn variability in terms of yield, moisture, neutral detergent fibre (NDF), acid detergent fibre (ADF) and starch content throughout the whole crop cycle, from sowing to harvest.

A2.3 Materials and Methods

Study Area

This study was conducted on a 14-ha field in Cona (VE), North Italy (Figure A2.1). The study field lies within the Mediterranean climatic zone; the mean temperature during the considered period was 18.5°C, and the total rainfall was 169mm. The corn planting date was 29th of March 2020 and harvest was done between the 11th and the 12th of August 2020.

Figure A2.1. Study Field in Cona (VE), North Italy.

Yield data collection

A John Deere 9900i combine harvester (Deere & Company, Moline, Illinois, USA) was outfitted with a John Deere 2630 yield monitor and a HarvestLab 3000 NIR sensor. This setup could record approximately 10,000 data points across the entire study field as shown in Figure A2.2. With a harvesting width of 4.5 meters, the yield monitor recorded data every second, corresponding to an average harvester speed of 1.5 meters per second in the harvesting direction, as shown in Figure A2.2(a, c). The Near Infrared Spectroscopy sensor mounted on the combined harvester also collected moisture (%), neutral detergent fiber (NDF), acid detergent fiber (ADF) and starch content for each yield point.

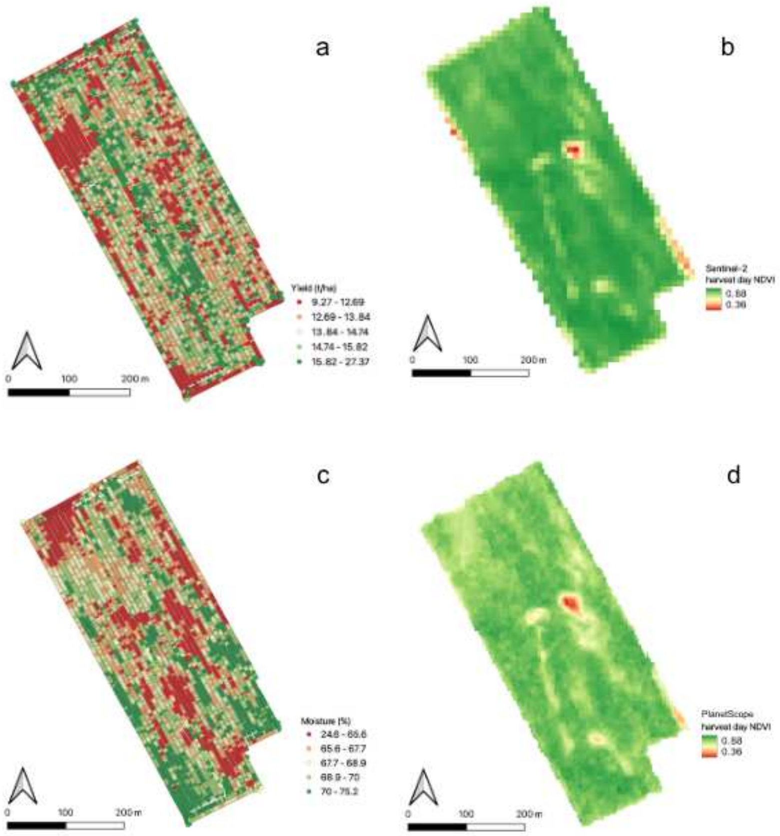


Figure A2.2. Yield (a) and Moisture (c) points collected by the John Deere 9000i combined harvester, NDVI collected by Sentinel-2 platform in the proximity of harvest day (08/08/2020) (c) and PlanetScope platform at harvest day (08/08/2020) (d)

Satellite Imagery collection and statistical analysis

Multispectral data was collected from two satellite sources, Sentinel 2 from ESA and the PlanetScope constellation from Planet. Sentinel 2 is a constellation of two satellites launched in 2017 that provide multispectral data with 13 bands between 443 and 2190nm, a five-day revisit time, and a spatial resolution between 10 and 60m. PlanetScope is a constellation that includes multiple satellites and offers multispectral images with four spectral bands ranging from 455 to 860nm, with a spatial resolution of 3m and a daily revisit time. All the images were filtered, and only those with cloud cover under 10% were included in the study. A total of 11 Sentinel 2 and 28 PlanetScope images were collected, encompassing the whole cultural cycle, from seeding date (1st of April, 2020), to harvest date (12th of August, 2020). The list of collected images, with dates and vegetative stage of the crop are listed in Table A2.1. Normalized Difference Vegetation Index (NDVI) and Green Normalized Vegetation Index (GNDVI) were calculated for each image and sampled for the yield points using QGIS (v. 3.28), Figure A2.2(b) shows an example of Sentinel-2 NDVI values over the field at the proximity of harvest date (8th of August), while Figure A2.2(c) shows PlanetScope NDVI values collected at the same date. The collected dataset was analysed and visualised using the Numpy, Pandas, Sci-kit learn, Matplotlib and Seaborn Python libraries. The regression models were trained and tested in the Python environment using the Scikit-learn package.

Table A2.1. List of images collected from satellite platforms during the growing season.

Sentinel-2		PlanetScope			
Date	Age (Stage)	Date	Age (Stage)	Date	Age (Stage)
08/04/2020	7 (VE)	04/04/2020	3 (VE)	05/06/2020	65 (R1)
23/04/2020	22 (V3)	06/04/2020	5 (VE)	21/06/2020	81 (R3)
08/05/2020	37 (V5)	09/04/2020	8 (VE)	03/07/2020	93 (R4)
02/06/2020	62 (V8)	10/04/2020	9 (VE)	04/07/2020	94 (R4)
22/06/2020	82 (R3)	11/04/2020	10 (V1)	09/07/2020	99 (R4)
27/06/2020	87 (R3)	16/04/2020	15 (V2)	11/07/2020	101 (R5)
07/07/2020	97 (R4)	22/04/2020	21 (V3)	15/07/2020	105 (R5)
12/07/2020	102 (R5)	24/04/2020	23 (V3)	20/07/2020	110 (R5)
27/07/2020	117 (R6)	30/04/2020	29 (V4)	30/07/2020	120 (R6)
01/08/2020	122 (R6)	07/05/2020	36 (V5)	06/08/2020	127 (R6)
08/08/2020	129 (R6)	12/05/2020	41 (V5)	07/08/2020	128 (R6)
		17/05/2020	46 (V6)	08/08/2020	129 (R6)
		21/05/2020	50 (V7)	10/08/2020	131 (R6)
		03/06/2020	63 (V8)	11/08/2020	132 (R6)

A2.4 Results and Discussion

The calculated vegetation indices were used to train and test linear regression models to determine which of the collected images was the best yield, moisture, and qualitative parameters estimator. The Sentinel-2 vegetation indices showed limited capabilities in predicting yield (Figure A2.3a), even though, interestingly, it reached the maximum R2 value on the 27th of June for both NDVI and GNDVI (R2=0.164 and 0.176 respectively), 87 days after sowing with maize crop at the R3 (Milk) phenological stage. PlanetScope-derived vegetation indices, especially GNDVI, showed better predicting capabilities, reaching a maximum R2 of 0.33, maintaining the early-season peak, in this case on the 5th of June, as visible in Figure A2.3b.

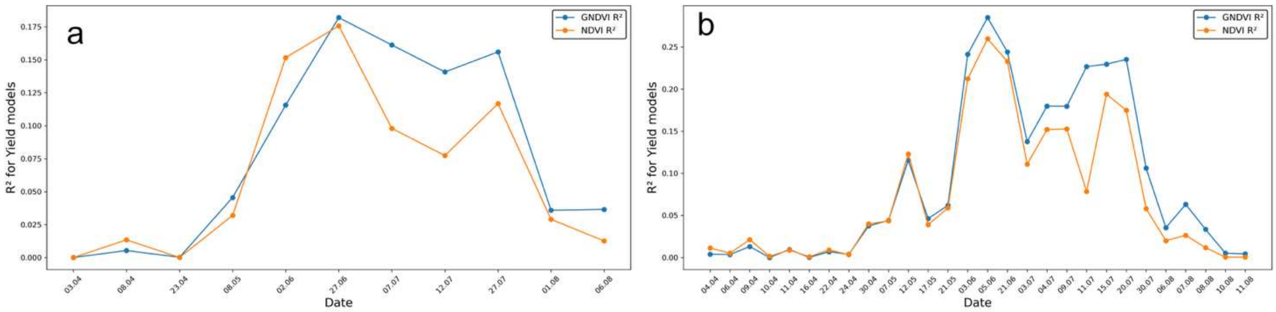


Figure A2.3 R^2 values of regression models for maize yield using GNDVI and NDVI from Sentinel-2 (a) and PlanetScope (b) at different dates throughout the crop cycle.

GNDVI demonstrated the strongest correlation between vegetation index (VI) and corn grain yield, consistent with the findings of Kayad et al. (Kayad et al., 2019). Furthermore, both Schwalbert et al. (Schwalbert et al., 2018) and Peralta et al. (Peralta et al., 2016) highlighted that GNDVI, along with NDVI, effectively captured the potential for predicting corn yield in their models. The present research found that both VIs from Sentinel-2 and PlanetScope imagery are particularly effective in capturing the spatial variability of corn grain yield within a field during a specific period, during the third month from the sowing date. In this period, the maize crop goes through the first three stages of its reproductive development, silking (R1), blister (R2) and milk (R3) stages.

Similar results have been observed in Central European and North American environments, highlighting this timeframe as ideal for corn yield prediction across different regions (Maestrini and Basso, 2018; P. Bognár Cs. Ferencz and Ferencz, 2011; Shanahan et al., 2001). The potential to assess crop yield variability in the early stages of development, opens the possibility to differentiate fertilization and pest management within the field, optimizing the production process and maximizing the efficiency of the productive inputs (Méndez-Vázquez et al., 2019).

The crop's moisture content is an important parameter to consider, especially when it is destined for the silage process (Bal et al., 1997). Therefore, assessing moisture content variability within the field in the proximity of the harvest date can provide useful information for potential differential harvest zones. In our study, NDVI and GNDVI showed comparable performances when predicting the crop's moisture content near the harvest date (Figure A2.4).

The Sentinel-2 platform's NDVI model performed slightly better than GNDVI, reaching an R2 equal to 0.37 (Figure A2.4a), while Planet Scope's GNDVI model obtained the highest accuracy, showing an R2 of 0.48 at the harvest date (Figure A2.4b).

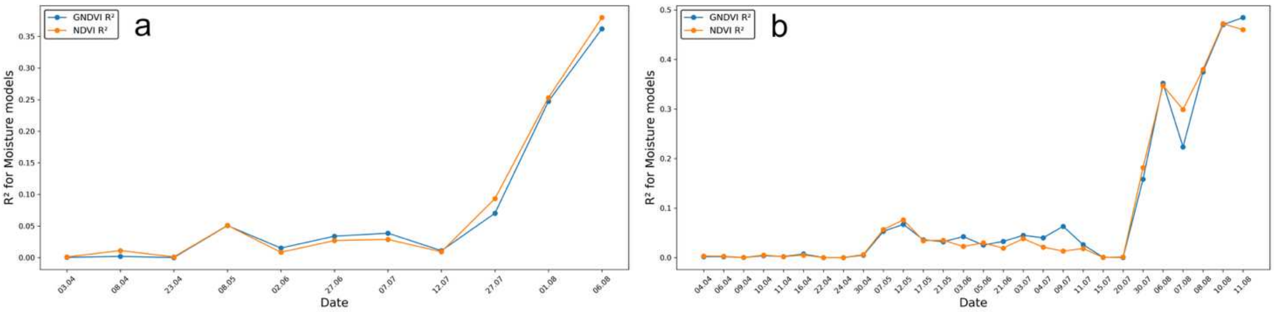


Figure A2.4. R² values of regression models for maize moisture content using GNDVI and NDVI from Sentinel-2 (a) and PlanetScope (b) at different dates throughout the crop cycle.

The potential of Near Infrared Spectrometry Sensors (NIRS) to assess crucial qualitative parameters of forage maize as Neutral Detergent Fiber (NDF), Acid Detergent Fiber (ADF) and Starch content has been investigated and assessed in the past by multiple authors (Berauer et al., 2020; Decruyenaere et al., 2009; Starks et al., 2004). In our study we aimed to verify the ability of regression models based on Sentinel-2 and PlanetScope vegetation indices to assess maize qualitative parameters throughout the crop cycle. The chemical composition of the forage affects the silage process 30 and the palatability and nutritional characteristics of the final product (Broderick, 2003; Der Bedrosian et al., 2012). Even though the relationships between cultural practices and qualitative parameters of forage are not as straightforward as for yield and moisture content, there is a potential for the delineation of differential management zones also based on qualitative parameters. In our study, when using the Sentinel-2 platform, NDVI models showed higher accuracies than GNDVI, both for NDF, ADF and Starch content estimation (Figures A2.5a, A2.6a and A2.7a). In the case of the PlanetScope platform, NDVI and GNDVI accuracies throughout the crop cycle overlapped for almost the whole season (Figures A2.5b, A2.6b and A2.7b).
Figure A2.5 shows that Neutral Detergent Fiber variability can be estimated with the highest accuracy in the proximity of harvest date, even though regression models reached modest R2 of 0.09 and 0.13 for Sentinel-2 and PlanetScope models, respectively.

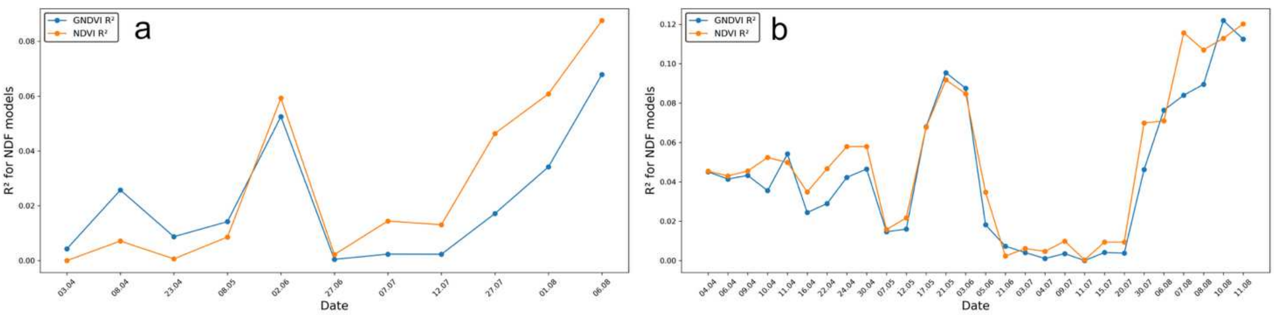


Figure A2.5. R² values of regression models for maize NDF content using GNDVI and NDVI from Sentinel-2 (a) and PlanetScope (b) at different dates throughout the crop cycle.

The models estimating Acid Detergent Fiber (ADF) showed slightly better accuracy, also reaching their maximum just before harvest date. Sentinel-2 NDVI model had an R2 of 0.12 while PlanetScope GNDVI reached 0.15, as shown in Figure A2.6. It is noteworthy to highlight that in both cases, even though with modest R2 values, the accuracy of the models has a peak in the period between May and June, 2 to 3 months after sowing, when the crop is approaching the end of its vegetative stages.

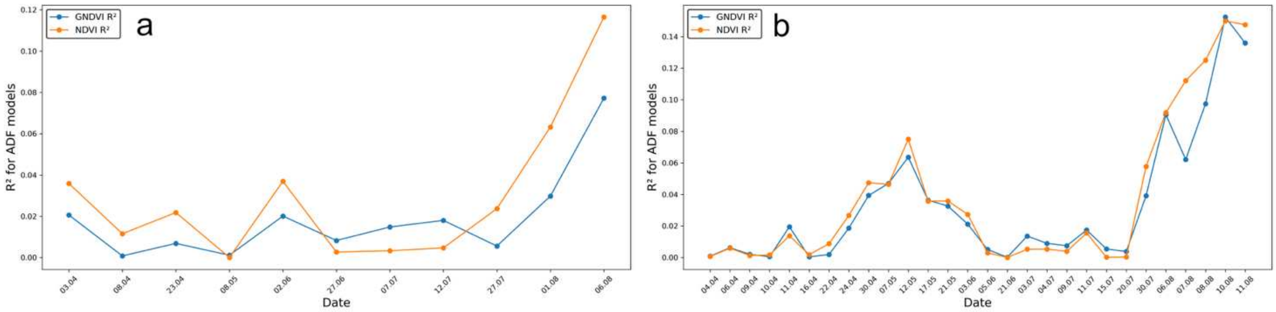


Figure A2.6. R² values of regression models for maize ADF content using GNDVI and NDVI from Sentinel-2 (a) and PlanetScope (b) at different dates throughout the crop cycle.

Similarly to what has been noticed for NDF and ADF, the models estimating Starch content show their better performances at harvest date. Sentinel-2 NDVI reach a maximum R2 of 0.13 while PlanetScope NDVI touch an R2 of 0.23, as visible in Figure A2.7. Both for Sentinel and PlanetScope models, their ability to assess within field variability is present in the first two months of the crop cycle, during its vegetative stages.

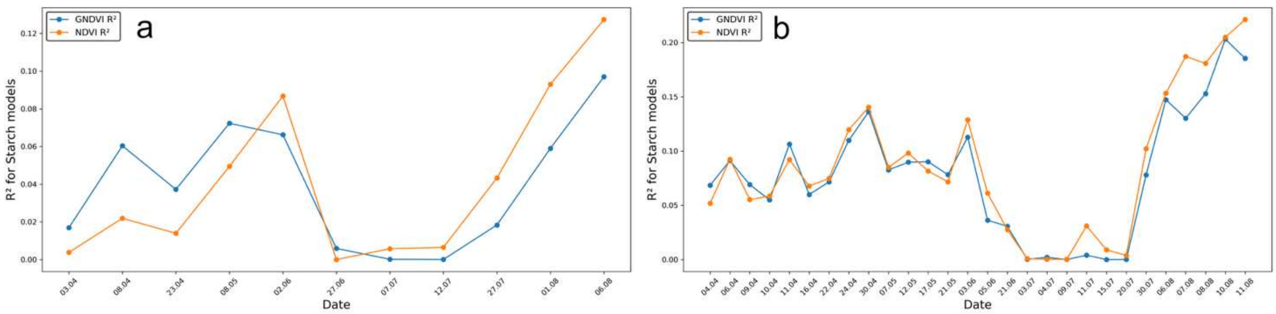


Figure A2.7. R² values of regression models for maize Starch content using GNDVI and NDVI from Sentinel-2 (a) and PlanetScope (b) at different dates throughout the crop cycle.

Although the relationships among cultural practices as fertilization, irrigation and pest management are not straightforward, the ability to assess and predict variability of forage qualitative variability early in the season opens to the possibility of testing and developing precision agriculture strategies and management practices.

A2.5 Conclusions

This study demonstrates the potential of using high-resolution multispectral imagery from Sentinel-2 and PlanetScope to assess within-field variability in maize yield, moisture content, and forage quality parameters (NDF, ADF, and starch content). The research highlighted that vegetation indices, particularly GNDVI and NDVI, derived from these satellite platforms, can effectively predict crop yield and moisture content, with PlanetScope GNDVI showing superior performance, especially early in the season and near harvest time. Although the predictive capabilities for qualitative parameters like NDF, ADF, and starch were modest, the findings suggest that early-season monitoring could inform differential management strategies, optimizing precision agriculture practices. Overall, integrating remote sensing data into Precision Agriculture could enhance decision-making, improving crop management and productivity at the field scale.

A2.6 References

An, S., Zhang, X., Chen, X., Yan, D., Henebry, G.M., 2018. An exploration of Terrain effects on land surface phenology across the Qinghai-Tibet Plateau using Landsat ETM+ and OLI data. Remote Sens (Basel) 10. https://doi.org/10.3390/rs10071069

Báez-González, A.D., Chen, P., Tiscareño-López, M., Srinivasan, R., 2002. Using Satellite and Field Data with Crop Growth Modeling to Monitor and Estimate Corn Yield in Mexico. Crop Sci 42, 1943–1949. https://doi.org/https://doi.org/10.2135/cropsci2002.1943

Bal, M.A., Coors, J.G., Shaver, R.D., 1997. Impact of the Maturity of Corn for Use as Silage in the Diets of Dairy Cows on Intake, Digestion, and Milk Production. J Dairy Sci 80, 2497–2503. https://doi.org/10.3168/jds.S0022-0302(97)76202-7

Baldwin, D., Manfreda, S., Keller, K., Smithwick, E.A.H., 2017. Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States. J Hydrol (Amst) 546, 393–404. https://doi.org/10.1016/j.jhydrol.2017.01.020

Berauer, B.J., Wilfahrt, P.A., Reu, B., Schuchardt, M.A., Garcia-Franco, N., Zistl-Schlingmann, M., Dannenmann, M., Kiese, R., Kühnel, A., Jentsch, A., 2020. Predicting forage quality of species-rich pasture grasslands using vis-NIRS to reveal effects of management intensity and climate change. Agric Ecosyst Environ 296. https://doi.org/10.1016/j.agee.2020.106929

Broderick, G.A., 2003. Effects of varying dietary protein and energy levels on the production of lactating dairy cows. J Dairy Sci 86, 1370–1381. https://doi.org/10.3168/jds.S0022-0302(03)73721-7

Decruyenaere, V., Lecomte, Ph., Demarquilly, C., Aufrere, J., Dardenne, P., Stilmant, D., Buldgen, A., 2009. Evaluation of green forage intake and digestibility in ruminants using near infrared reflectance spectroscopy (NIRS): Developing a global calibration. Anim Feed Sci Technol 148, 138–156. https://doi.org/10.1016/j.anifeedsci.2008.03.007

Deines, J.M., Patel, R., Liang, S.Z., Dado, W., Lobell, D.B., 2021. A million kernels of truth: Insights into scalable satellite maize yield mapping and yield gap analysis from an extensive ground dataset in the US Corn Belt. Remote Sens Environ 253, 112174. https://doi.org/10.1016/J.RSE.2020.112174

Der Bedrosian, M.C., Nestor, K.E., Kung, L., 2012. The effects of hybrid, maturity, and length of storage on the composition and nutritive value of corn silage. J Dairy Sci 95, 5115–5126. https://doi.org/10.3168/jds.2011-4833

Dusseux, P., Hubert-Moy, L., Corpetti, T., Vertès, F., 2015. Evaluation of SPOT imagery for the estimation of grassland biomass. International Journal of Applied Earth Observation and Geoinformation 38, 72–77. https://doi.org/10.1016/J.JAG.2014.12.003

Kayad, A., Sozzi, M., Gatto, S., Marinello, F., Pirotti, F., 2019. Monitoring Within-Field Variability of Corn Yield using Sentinel-2 and Machine Learning Techniques. Remote Sens (Basel) 11. https://doi.org/10.3390/rs11232873

Kayad Ahmed G. AND Al-Gaadi, K.A.A.N.D.T.E.A.N.D.M.R.A.N.D.Z.A.M.A.N.D.K.C., 2016. Assessing the Spatial Variability of Alfalfa Yield Using Satellite Imagery and Ground-Based Data. PLoS One 11, 1–15. https://doi.org/10.1371/journal.pone.0157166

Kharel, T.P., Maresma, A., Czymmek, K.J., Oware, E.K., Ketterings, Q.M., 2019. Combining Spatial and Temporal Corn Silage Yield Variability for Management Zone Development. Agron J 111, 2703–2711. https://doi.org/https://doi.org/10.2134/agronj2019.02.0079

Kogan, F., Salazar, L., Roytman, L., 2012. Forecasting crop production using satellite-based vegetation health indices in Kansas, USA. Int J Remote Sens 33, 2798–2814. https://doi.org/10.1080/01431161.2011.621464

Kross, A., McNairn, H., Lapen, D., Sunohara, M., Champagne, C., 2015. Assessment of RapidEye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops. International Journal of Applied Earth Observation and Geoinformation 34, 235–248. https://doi.org/10.1016/J.JAG.2014.08.002

Kross, A., Znoj, E., Callegari, D., Kaur, G., Sunohara, M., Lapen, D.R., McNairn, H., 2020. Using Artificial Neural Networks and Remotely Sensed Data to Evaluate the Relative Importance of Variables for Prediction of Within-Field Corn and Soybean Yields. Remote Sens (Basel) 12. https://doi.org/10.3390/rs12142230

Liu, W., Mõttus, M., Gastellu-Etchegorry, J.P., Fang, H., Atherton, J., 2024. Seasonal and vertical variation in canopy structure and leaf spectral properties determine the canopy reflectance of a rice field. Agric For Meteorol 355, 110132. https://doi.org/10.1016/J.AGRFORMET.2024.110132

Lobell, D.B., 2013. The use of satellite data for crop yield gap analysis. Field Crops Res 143, 56–64. https://doi.org/10.1016/J.FCR.2012.08.008

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens Environ 164, 324–333. https://doi.org/10.1016/J.RSE.2015.04.021

Maestrini, B., Basso, B., 2018. Predicting spatial patterns of within-field crop yield variability. Field Crops Res 219, 106–112. https://doi.org/10.1016/J.FCR.2018.01.028

Méndez-Vázquez, L.J., Lira-Noriega, A., Lasa-Covarrubias, R., Cerdeira-Estrada, S., 2019. Delineation of site-specific management zones for pest control purposes: Exploring precision agriculture and species distribution modeling approaches. Comput Electron Agric 167. https://doi.org/10.1016/j.compag.2019.105101

Mokhtari, A., Noory, H., Vazifedoust, M., Bahrami, M., 2018. Estimating net irrigation requirement of winter wheat using model- and satellite-based single and basal crop coefficients. Agric Water Manag 208, 95–106. https://doi.org/10.1016/j.agwat.2018.06.013

P. Bognár Cs. Ferencz, Sz.P.G.M.G.T.D.H.J.L.B.S.P.S., Ferencz, O.E., 2011. Yield forecasting for wheat and corn in Hungary by satellite remote sensing. Int J Remote Sens 32, 4759–4767. https://doi.org/10.1080/01431161.2010.493566

Peralta, N.R., Assefa, Y., Du, J., Barden, C.J., Ciampitti, I.A., 2016. Mid-Season High-Resolution Satellite Imagery for Forecasting Site-Specific Corn Yield. Remote Sens (Basel) 8. https://doi.org/10.3390/rs8100848

Salazar, L., Kogan, F., Roytman, L., 2008. Using vegetation health indices and partial least squares method for estimation of corn yield. Int J Remote Sens 29, 175–189. https://doi.org/10.1080/01431160701271974

Sartore, L., Boryan, C.G., Chen, L., 2023. Utilizing Land Cover, Satellite and Agricultural Survey Data to Produce Early Season Crop Acreage Estimates, in: International Geoscience and Remote Sensing Symposium (IGARSS). pp. 3434–3437. https://doi.org/10.1109/IGARSS52108.2023.10282795

Schwalbert, R.A., Amado, T.J.C., Nieto, L., Varela, S., Corassa, G.M., Horbe, T.A.N., Rice, C.W., Peralta, N.R., Ciampitti, I.A., 2018. Forecasting maize yield at field scale based on high-resolution satellite imagery. Biosyst Eng 171, 179–192. https://doi.org/10.1016/J.BIOSYSTEMSENG.2018.04.020

Shanahan, J.F., Schepers, J.S., Francis, D.D., Varvel, G.E., Wilhelm, W.W., Tringe, J.M., Schlemmer, M.R., Major, D.J., 2001. Use of Remote-Sensing Imagery to Estimate Corn Grain Yield. Agron J 93, 583–589. https://doi.org/https://doi.org/10.2134/agronj2001.933583x

Sibley, A.M., Grassini, P., Thomas, N.E., Cassman, K.G., Lobell, D.B., 2014. Testing Remote Sensing Approaches for Assessing Yield Variability among Maize Fields. Agron J 106, 24–32. https://doi.org/https://doi.org/10.2134/agronj2013.0314

Starks, P.J., Coleman, S.W., Phillips, W.A., 2004. Determination of forage chemical composition using remote sensing. Journal of Range Management 57, 635–640. https://doi.org/10.2111/1551-5028(2004)057[0635:dofccu]2.0.co;2

Xu, J., Meng, J., Quackenbush, L.J., 2019. Use of remote sensing to predict the optimal harvest date of corn. Field Crops Res 236, 1–13. https://doi.org/10.1016/j.fcr.2019.03.003