



UNIVERSITÀ
POLITECNICA
DELLE MARCHE

DIPARTIMENTO DI SCIENZE AGRARIE ALIMENTARI E AMBIENTALI
CORSO DI LAUREA IN: SCIENZE FORESTALI, DEI SUOLI E DEL PAESAGGIO
(LM-73)

**Verso un sistema esperto di classificazione
automatica della vegetazione e degli
habitat. Il caso studio della ZSC Gola di
Frasassi.**

**Towards an expert system for the automatic classification of vegetation
and habitats. The case study of the ZSC Gola di Frasassi.**

TIPO TESI: sperimentale

Studente:
Sofia Vignoli

Relatore:
Dott. Simone Pesaresi

Correlatori:
Dott. Giacomo Quattrini
Prof.ssa Simona Casavecchia

Anno Accademico 2023-2024



UNIVERSITÀ
POLITECNICA
DELLE MARCHE

DIPARTIMENTO DI SCIENZE AGRARIE ALIMENTARI E AMBIENTALI
CORSO DI LAUREA IN: SCIENZE FORESTALI, DEI SUOLI E DEL PAESAGGIO
(LM-73)

**Verso un sistema esperto di classificazione
automatica della vegetazione e degli
habitat. Il caso studio della ZSC Gola di
Frasassi.**

**Towards an expert system for the automatic classification of vegetation
and habitats. The case study of the ZSC Gola di Frasassi.**

TIPO TESI: sperimentale

Studente:
Sofia Vignoli

Relatore:
Dott. Simone Pesaresi

Correlatori:
Dott. Giacomo Quattrini
Prof.ssa Simona Casavecchia

Anno Accademico 2023-2024

UNIVERSITÀ POLITECNICA DELLE MARCHE
DIPARTIMENTO DI SCIENZE AGRARIE ALIMENTARI E AMBIENTALI
CORSO DI LAUREA IN: SCIENZE FORESTALI, DEI SUOLI E DEL PAESAGGIO (LM-73)
Via Brecce Bianche – 60131 Ancona (AN), Italy

"Tu non sai: ci sono betulle che di notte levano le loro radici, e tu non crederesti mai che di notte gli alberi camminano o diventano sogni..."

"Pensa che in un albero c'è un violino d'amore. Pensa che un albero canta e ride. Pensa che un albero sta in un crepaccio e poi diventa vita..."
(Alda Merini)

Ringraziamenti

Al termine di questo mio percorso universitario desidero ringraziare tutte quelle persone che hanno contribuito alla realizzazione di questo lavoro.

Un ringraziamento speciale, per la sua grande disponibilità e professionalità, al mio relatore, il Dott. Simone Pesaresi, che mi ha sempre seguito, sia durante la stesura della tesi triennale che in quella magistrale mettendo a disposizione il suo lavoro, le sue attività di ricerca e le sue esperienze unitamente ai suoi collaboratori, nonché miei correlatori, il Dottorando Giacomo Quattrini e la Prof.ssa Simona Casavecchia. Vorrei inoltre ringraziare il Prof. Adriano Mancini, correlatore nella tesi triennale, per il suo supporto tecnico anche in questo lavoro e Nicole Hofmann per il prezioso aiuto fornitomi nella fase conclusiva di questa tesi nell'individuazione e riconoscimento delle specie botaniche presenti nelle foto scattate da drone.

Oltre al mio relatore ed ai miei correlatori, desidero ringraziare la mia famiglia per l'amore ed il sostegno trasmesso e senza il quale non avrei di certo raggiunto questo importante traguardo.

In particolare, vorrei rivolgere un sincero ringraziamento ai miei genitori per tutti i sacrifici che avete fatto per permettermi di arrivare fino a qui. Grazie perchè oltre ad avermi trasmesso valori importanti, mi avete fatto conoscere posti meravigliosi durante i nostri viaggi di famiglia. Vi ringrazio anche perchè, a modo vostro, non mi avete mai fatto mancare il vostro sostegno nelle scelte che ho intrapreso finora e perchè mi spronate ogni giorno a dare il meglio di me in ogni situazione. Spero di avervi reso orgogliosi e, anche se non ve lo dico mai, sappiate che vi voglio bene.

A mia sorella Alessandra che, pur essendo così diversa da me, ha saputo trasformare questa diversità nel nostro punto di forza. L'aver condiviso come me l'esperienza universitaria ti ha permesso di comprendere più di chiunque altro i miei stati d'animo prima di ogni esame e di aver avuto la capacità di trovare, seppur non sempre, parole di conforto. Questa è anche l'occasione per dirti che, anche se a volte possa sembrare severa nei tuoi confronti, sappi che lo faccio per il tuo bene e desidero il meglio per te perchè te lo meriti.

Ai miei nonni materni, Livio e Renata, che a pochi giorni dall'inizio di questo mio percorso magistrale, in una notte di metà settembre, hanno perso quanto costruito in

quasi 60 anni di vita insieme ma l'amore per la propria famiglia ha permesso loro di trovare la forza di ricominciare. Chi mi conosce sa quanto sia legata a loro e quanto per me rappresentiate un esempio da seguire.

Ai miei nonni paterni, Renato e Maria, che non ci sono più, ma che sono sicura sarebbero stati orgogliosi del traguardo raggiunto.

Al mio fidanzato Christian, che mi supporta e sopporta ogni giorno e che ringrazio perchè mi ha insegnato che, anche nei momenti più bui, bisogna trovare la forza dentro di sé per reagire sempre con un sorriso.

Ai miei zii e cugini, che non mi hanno mai fatto mancare il loro affetto e con i quali ho vissuto i momenti più spensierati a casa dei nonni.

Ai miei amici, che mi sono sempre stati vicini in questi anni di studio e che mi hanno dato l'opportunità di confrontarmi e di crescere insieme a loro.

Infine, desidero estendere un sincero ringraziamento a tutte le persone che, pur non essendo state menzionate, hanno contribuito in modo significativo al mio percorso. Grazie di cuore a tutti.

Ancona, Dicembre 2024

Sofia Vignoli

Abstract

In this study, an innovative methodology for the classification and mapping of vegetation is tested. The main objective is to create a system that enables the automatic and rapid production of maps of unexplored areas. The approach used to process the satellite and drone data combines machine learning techniques and functional data analysis. The results show a high reliability of the model used to distinguish different vegetation types on the basis of the seasonal spectral variations detected. The development of this system would therefore contribute to more efficient monitoring, resource management and intervention planning in a given area.

Sommario

In questo studio, viene sperimentata un'innovativa metodologia per la classificazione e mappatura della vegetazione. L'obiettivo principale è creare un sistema che consenta di rendere automatico e rapido il processo di produzione delle mappe di aree inesplorate. L'approccio utilizzato per elaborare i dati satellitari e da drone, combina tecniche di apprendimento automatico e l'analisi funzionale dei dati. Dai risultati emerge un'elevata affidabilità del modello utilizzato per distinguere i differenti tipi vegetazionali sulla base delle variazioni spettrali stagionali rilevate. Lo sviluppo di questo sistema contribuirebbe quindi a rendere più efficiente le attività di monitoraggio, di gestione delle risorse e della pianificazione degli interventi in un determinato territorio.

Indice

Acronimi e Abbreviazioni	xix
1 Introduzione	1
2 Obiettivi	3
3 Materiali e Metodi	5
3.1 Area di studio	5
3.1.1 ZSC Gola di Frasassi	5
3.2 Reference Data	6
3.2.1 Dati topografici	8
3.3 Classificazione supervisionata delle serie temporali per la mappatura della vegetazione	9
3.3.1 Caratteristiche delle Sentinel-2	9
3.3.2 Collezionamento e pre-processamento delle immagini satellitari Sentinel-2	10
3.3.3 <i>Functional Data Analysis</i> (FDA)	11
3.3.4 <i>FPCA e MFPCA</i>	11
3.3.5 <i>Random Forest</i> (RF)	11
3.3.6 Classificazione supervisionata (Supervised Classification) . . .	12
3.3.7 Accuratezza della Mappatura (<i>Ten Fold Cross Validation</i>) . .	12
3.3.8 Software utilizzati	13
4 Risultati	15
4.1 Selezione degli Indici di Vegetazione	15
4.2 Mappa della Gola di Frasassi	15
4.3 Matrice di confusione	18
5 Discussioni	21
5.1 Punti di forza e dominio delle applicazioni	21
5.2 Limiti e futuri sviluppi	22
6 Conclusioni	29
7 Appendice	31
7.1 Armonizzazione dataset	31
7.2 Profili stagionali medi	33

Elenco delle figure

3.1	Inquadramento dell'area di studio su scala nazionale e regionale.	6
3.2	Panoramica della Gola di Frasassi. A sinistra Monte Valmontagnana (931 m slm) e a destra Monte Frasassi (709 m slm). La foto è stata scattata il 1° settembre 2024 durante un'escursione in loco. Foto di Sofia Vignoli.	6
3.3	Distribuzione spaziale dei Reference Data nelle aree del Parco Naturale del Conero (A), della Riserva Naturale Statale Gola del Furlo (B) e del Parco Interregionale del Sasso Simone e Simoncello (C) rispetto all'area target (D).	7
3.4	Sentinel-2 in orbita.	9
3.5	Esempio di un'immagine Sentinel-2 (risalente al 24 luglio 2022) di cui è stato fatto il download e le operazioni di pre-processamento.	10
3.6	Logo di R.	13
3.7	Logo di QGIS.	13
4.1	Mappa della Gola di Frasassi. Le etichette presenti in legenda fanno riferimento alla tabella 3.1 del Reference Data.	16
4.2	Confronto visivo tra la mappa ottenuta in questo lavoro (a) e la mappa prodotta nell'articolo di[Pesaresi et al., 2022] (b). In nero sono cerchiare quelle aree dove c'è maggiore discordanza tra le due mappe e nelle quali è stato deciso di effettuare dei sorvoli con il drone.	17
4.3	Aree sorvolate con il drone (punti in colore nero) per ispezionare le differenze tra le differenze indicate in figura 4.2. L'autorizzazione è avvenuta attraverso il rilascio del nulla osta (Prot. n. 0253749 del 31/10/2024) da parte del Parco Naturale Regionale della Gola della Rossa e di Frasassi.	17
4.4	Mappa della Gola di Frasassi. Le etichette presenti in legenda fanno riferimento alla tabella 3.1 del Reference Data. I punti colorati in nero costituiscono invece il validation set.	18
5.1	Esempio di un'immagine di bosco di roverella (<i>Quercus pubescens</i> L.) scattata con il drone durante un sopralluogo effettuato in campo il 6 novembre 2024.	22

Elenco delle figure

5.2	Attraverso uno zoom ottico, è possibile ottenere immagini che consentono l'identificazione delle specie vegetali e quindi stabilire la specie fisionomizzante che, in questo caso, è rappresentata dalla roverella (Q) (<i>Quercus pubescens L.</i>).	22
5.3	Vista dall'alto di un prebosco ad orniello. La foto è stata scattata il 6 novembre 2024.	24
5.4	Zoom 16x sul prebosco ad orniello.	24
5.5	Vista dall'alto di una delle aree in cui è presente una codominanza tra <i>Quercus pubescens L.</i> ed <i>Ostrya carpinifolia Scop.</i> . La foto è stata scattata il 6 novembre 2024.	25
5.6	Dettaglio della foto precedente. In primo piano la roverella.	25
5.7	Vista dall'alto di una delle aree in cui domina il cerro (<i>Quercus cerris L.</i>). La foto è stata scattata il 6 novembre 2024.	26
5.8	Dettaglio del bosco di cerro ottenuto con lo zoom 16x con cui è equipaggiato il drone Matrice 30 DJI.	26
6.1	Monte Valmontagnana. La foto è stata scattata il 6 novembre 2024.	30
7.1	Profili stagionali multispettrali medi delle categorie relative alla tabella 3.1 del Reference Data.	36

Elenco delle tabelle

3.1	Classi target e Reference data.	8
3.2	Bande spettrali e risoluzioni spaziali del Sentinel-2.	10
4.1	Le 7 migliori serie temporali selezionate dalla <i>Forward Selection</i> con formula F03, non standardizzata e classificatore RF.	15
4.2	Elenco delle categorie escluse dal Reference Data. Le etichette, la fisionomia e l'habitat sono riportati nella tabella 3.1 del Reference Data.	16
4.3	Matrice di confusione.	18
7.1	Processo di armonizzazione tra i due dataset di confronto.	31

Acronimi e Abbreviazioni

DEM DIGITAL ELEVATION MODEL

FDA FUNCTIONAL DATA ANALYSIS

FPCA FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

GAM GENERALIZED ADDITIVE MODELS

MFPCA MULTIDIMENSIONAL FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

RF RANDOM FOREST

ZPS ZONA DI PROTEZIONE SPECIALE

ZSC ZONA SPECIALE DI CONSERVAZIONE

Capitolo 1

Introduzione

Individuare, conoscere e comprendere i fattori ambientali (biotici e abiotici) che determinano la distribuzione della vegetazione nello spazio in un determinato territorio, rappresenta ancora ad oggi, una vera e propria sfida a causa della complessità delle relazioni che si instaurano all'interno degli ecosistemi. Il metodo di analisi che consente di mettere in evidenza queste relazioni, è conosciuto come metodo fitosociologico, la cui unità fondamentale è l'associazione vegetale [Ullah et al., 2024]. Secondo la definizione fornita da Braun-Blanquet (1932), fondatore della fitosociologia, l'associazione vegetale è una combinazione di specie che si ripete statisticamente nello spazio per l'azione di un'insieme di fattori ambientali e che presenta particolari caratteristiche di tipo ecologico e biogeografico [Biondi, 2011]. La biodiversità o diversità vegetale che viene rilevata può quindi essere rappresentata graficamente mediante la realizzazione di mappe della vegetazione, ormai diventate uno strumento fondamentale nella pianificazione, gestione, conservazione e monitoraggio di un territorio. Per contrastare la perdita di biodiversità che si sta registrando negli ultimi decenni, nell'ambito dell'Unione Europea, ciascuno Stato Membro è chiamato ad impegnarsi a rispettare misure specifiche stabilite dalla Direttiva Habitat (92/43/CEE) con lo scopo di garantire il ripristino e la conservazione delle specie e degli habitat naturali e seminaturali che ospitano www.consilium.europa.eu/it/policies/biodiversity/.

Tradizionalmente l'approccio con cui ancora oggi vengono realizzate le mappe fitosociologiche si basa principalmente sulla raccolta e classificazione di dati in campo in categorie vegetazionali e sull'analisi visiva di foto aeree ed ortofoto. Questo metodo è particolarmente efficace su aree di piccole e medie dimensioni e di cui si conoscono le caratteristiche ambientali e climatiche. Non è invece applicabile su aree molto estese perchè richiederebbe tempi lunghi ed un notevole impiego di risorse [Shanmuga Priya and Vani, 2024]. Ulteriori limiti di questo approccio riguardano ad esempio la soggettività nell'interpretazione delle immagini e la necessità di effettuare dei rilievi aggiuntivi per avere una stima dell'accuratezza della mappa finale. Risulta quindi difficile garantire una produzione continua delle mappature.

L'enorme quantità di dati che viene acquisita ogni giorno da piattaforme satellitari come Landsat e Sentinel, rappresenta una risorsa estremamente importante per le analisi multitemporali e di serie temporali finalizzate a distinguere e a classificare le differenti tipologie vegetazionali [Cenci et al., 2018]. Una volta raccolti, questi

dati possono essere analizzati attraverso l'applicazione di modelli di apprendimento automatico. L'efficacia di un modello dipende dalla quantità e dalla qualità dei dati utilizzati per l'addestramento. Qualora non si abbia la possibilità di disporre di un set di dati denso, ci si può avvalere dell'utilizzo di dati derivati dall'interpretazione visiva di immagini RGB ad alta risoluzione realizzati da drone [Kattenborn et al., 2019]. Nel momento in cui invece si ha a disposizione un'adeguata e accurata quantità di dati di riferimento, è importante saperli valorizzare poichè diventa strategico per produrre rapidamente ed in continuo nel tempo cartografie della vegetazione costituendo di fatto un importante strumento di monitoraggio.

Alla luce di queste considerazioni, questa tesi si inserisce all'interno di un più ampio progetto di ricerca che ha avuto inizio dal lavoro di [Quattrini et al., 2017] ed è incentrato sulla sperimentazione di un'innovativa metodologia di mappatura della vegetazione proposta in [Pesaresi et al., 2020]. La strategia di "Forward Selection" introdotta in [Vignoli et al., 2021] e pubblicata in [Pesaresi et al., 2024] ha permesso di facilitare e velocizzare l'individuazione delle migliori serie temporali pertanto viene nuovamente riproposta ma anzichè essere applicata su dati raster, sono stati presi in considerazione dati puntuali. Di conseguenza, l'intento è quello di costruire un sistema esperto all'interno del quale sono racchiusi tutti i reference data collezionati nel corso delle varie pubblicazioni e tesi per poter riuscire a mappare aree ancora inesplorate e di stimarne l'estensione.

Costruire un sistema esperto automatico di mappatura della vegetazione consentirebbe di rendere il monitoraggio efficiente al fine di identificare lo stato di salute e le dinamiche della vegetazione e degli habitat in continuum nel tempo. Realtà come Parchi e Riserve naturali possono trarre vantaggio dallo sviluppo di questo sistema nell'ambito della ricerca perchè permetterebbe di garantire una gestione più efficiente delle risorse e di agevolare la pianificazione degli interventi.

Capitolo 2

Obiettivi

L'obiettivo di questa tesi è utilizzare le mappature della vegetazione realizzate nelle aree del Conero [Casavecchia et al., 2023a], Carpegna [Casavecchia et al., 2023b] e Furlo [Casavecchia et al., 2023c] per creare un sistema esperto di classificazione supervisionata delle loro variazioni spettrali stagionali (fenologia telerilevata), al fine di ottenere in tempi brevi mappature accurate della vegetazione per nuove aree target. I dati di riferimento raccolti nelle aree del Conero, Carpegna e Furlo verranno combinati e utilizzati per addestrare un classificatore di machine learning, in grado di riconoscere e mappare diverse tipologie di vegetazione e habitat basandosi sulle variazioni spettrali stagionali rilevate da remoto.

Questo studio si propone di rispondere al seguente quesito: è possibile costruire un sistema esperto di classificazione e mappatura automatica della vegetazione, in grado di realizzare una cartografia accurata e in tempi brevi di un'area 'inesplorata'?

Il sistema è addestrato per riconoscere diverse categorie di vegetazione basandosi su profili stagionali multispettrali. Per ogni tipo di vegetazione presente nei dati di riferimento, il sistema conosce le variazioni stagionali nelle bande spettrali, che fungono da "prototipi" di riferimento. Questi prototipi includono i momenti stagionali più distintivi per ciascuna categoria. Quindi, se forniamo al classificatore un profilo stagionale nuovo (anche da un'area non presente nei dati di riferimento), il sistema assegnerà quel profilo alla categoria di vegetazione più simile a uno dei prototipi. Il sistema sarà quindi applicato per generare una mappa della vegetazione in un'area esterna, non inclusa nei dati di riferimento utilizzati per l'addestramento. In particolare, il modello esperto addestrato con i dati delle aree di Carpegna, Furlo e Conero verrà impiegato per mappare la vegetazione della Zona Speciale di Conservazione (ZSC) "Gola di Frasassi" IT5320003.

Una volta prodotta, la mappa sarà validata e l'accuratezza stimata utilizzando un set di dati di validazione ottenuti attraverso attività di ground truthing e drone truthing.

Capitolo 3

Materiali e Metodi

In questo capitolo verranno illustrate, in sequenza, tutte le fasi di lavoro necessarie a realizzare una mappa della vegetazione per l'area di interesse basata sulla classificazione supervisionata delle variazioni spettrali stagionali rilevate da punti esterni. Dapprima si è iniziato con il collezionamento delle serie temporali Sentinel-2 a cui è seguita una fase durante la quale è stato possibile selezionare un set di indici di vegetazione adatti all'area di studio. Gli indici sono poi stati trasformati in funzioni continue utilizzando alcuni modelli conosciuti come *Generalized Additive Models* (GAM). Successivamente le funzioni sono state analizzate attraverso la *Multivariate Functional Principal Component Analysis* (MFPCA). Gli *scores* della MFPCA sono stati impiegati come dati di input nei modelli di apprendimento automatico (per questo lavoro è stato scelto il *Random Forest*) per eseguire la classificazione supervisionata ed ottenere la mappa finale. I dati raccolti in campo (ground truthing) e da drone (drone truthing), usati per addestrare il modello, riguardano aree che non coincidono con quella di nostro interesse.

3.1 Area di studio

3.1.1 ZSC Gola di Frasassi

L'area oggetto di studio è la Zona Speciale di Conservazione (ZSC) "Gola di Frasassi" IT5320003. E' uno dei siti di maggiore biodiversità vegetale e animale di tutta la regione Marche ed è racchiuso all'interno del Parco Naturale Regionale della Gola della Rossa e di Frasassi (L.R. n.57 del 02.09.1997) in provincia di Ancona, amministrato dai comuni di Genga e Fabriano. La superficie complessiva della ZSC è di circa 728 ha ed è compresa nella più ampia Zona di Protezione Speciale (ZPS) "Gola della Rossa e di Frasassi" IT5320017. L'altitudine massima che si raggiunge è di 931 metri slm (M. Valmontagnana) ed il tipo di roccia che affiora è il Calcarea Massiccio, appartenente alla successione umbro-marchigiana. Il territorio è conosciuto soprattutto per la presenza di fenomeni carsici che hanno dato origine al complesso ipogeo delle Grotte di Frasassi che ogni anno accoglie migliaia di visitatori. Dall'analisi dei dati climatici raccolti per l'area di studio risulta che la precipitazione media annua è di circa 1.132 mm mentre la temperatura media annua è pari a 12.1°C.



Figura 3.1: Inquadramento dell'area di studio su scala nazionale e regionale.

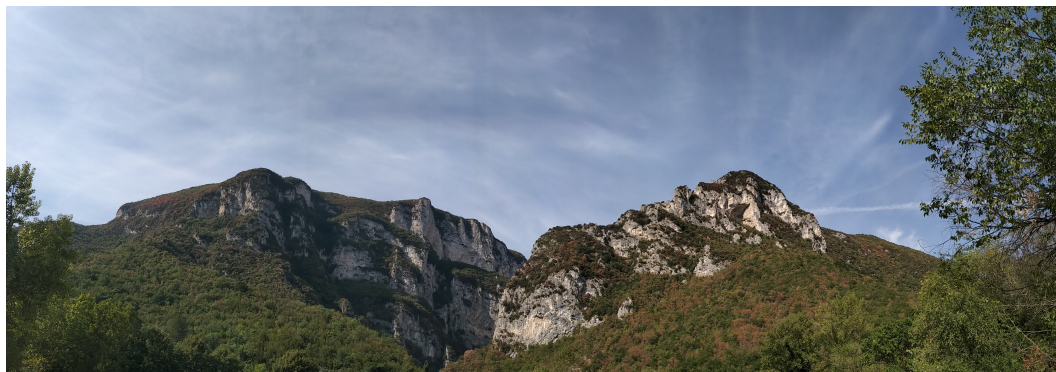


Figura 3.2: Panoramica della Gola di Frasassi. A sinistra Monte Valmontagnana (931 m slm) e a destra Monte Frasassi (709 m slm). La foto è stata scattata il 1° settembre 2024 durante un'escursione in loco. Foto di Sofia Vignoli.

Secondo la classificazione di Rivas-Martinez, l'area appartiene al macrobioclima temperato, nella variante submediterranea, caratterizzata da un breve periodo di aridità durante la stagione estiva [Rivas-Martínez, 2011].

3.2 Reference Data

In accordo con quanto riportato negli obiettivi, si è reso necessario realizzare un unico Reference data costituito da punti esterni rispetto all'area indagata, raccolti tra il 2019 ed il 2022 all'interno del Parco del Conero [Casavecchia et al., 2023a], del Parco Interregionale del Sasso Simone e Simoncello [Casavecchia et al., 2023b] e della Riserva Statale del Furlo [Casavecchia et al., 2023c] e sui quali verrà applicata la metodologia di analisi messa a punto da [Pesaresi et al., 2020] (Vedi tabella 3.1).

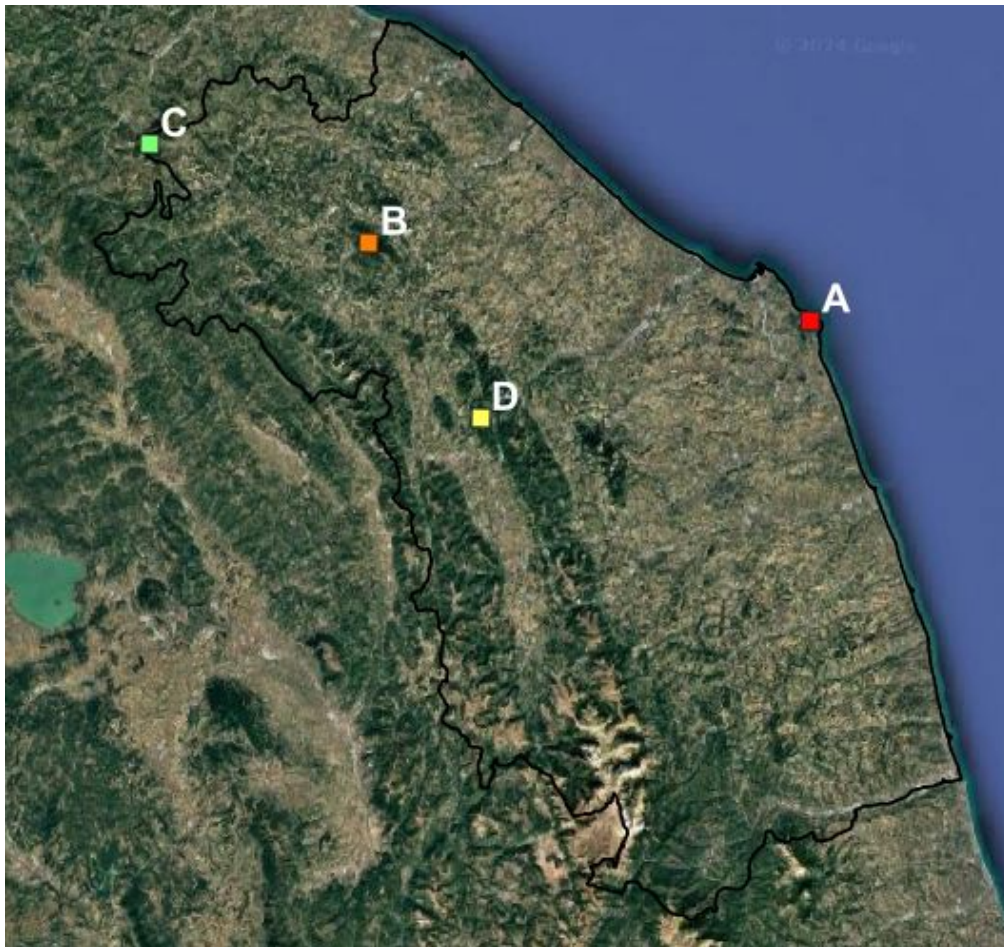


Figura 3.3: Distribuzione spaziale dei Reference Data nelle aree del Parco Naturale del Conero (A), della Riserva Naturale Statale Gola del Furlo (B) e del Parco Interregionale del Sasso Simone e Simoncello (C) rispetto all'area target (D).

Confrontando le caratteristiche morfologiche, litologiche, climatiche e vegetazionali dell'area target con quelle del Conero, del Furlo e di Carpegna, è emerso che le prime due presentano una maggiore affinità con la Gola di Frasassi, mentre Carpegna si distingue maggiormente. L'acquisizione dei punti e la loro relativa associazione con le categorie vegetazionali presenti è avvenuta mediante sopralluoghi in campo. Per le aree di difficile accesso si è dovuto ricorrere all'interpretazione di immagini satellitari oppure a foto scattate da drone (il modello utilizzato in questo lavoro è Matrice 30 DJI ed è dotato di uno zoom 16x). La fase più delicata è stata quella di armonizzare e rendere omogenee le etichette assegnate ad ogni categoria vegetazionale presente nelle tre aree precedentemente indicate. Alcune di queste sono state escluse perchè rappresentate da pochi punti che non riuscirebbero ad addestrare il sistema di classificazione ed altre invece sono state aggregate. In seguito è stato possibile assemblare in un unico file i tre dataset tramite l'operazione di join. I dati di riferimento che sono stati raccolti per la Gola di Frasassi, verranno invece utilizzati

per la validazione della mappa.

Tabella 3.1: Classi target e Reference data.

REFERENCE DATA			
LABEL	FISIONOMIA	HABITAT	TOT
Boschi			
AQ	a dominanza di <i>Quercus cerris</i>	91M0	20
F	a dominanza di <i>Fagus sylvatica</i>	9210*[6430]	61
FR	a dominanza di <i>Fraxinus excelsior</i>	9180*	21
QC	a dominanza di <i>Quercus cerris</i> e <i>Carpinus betulus</i>	91L0 [6430]	70
Q	a dominanza di <i>Quercus pubescens</i>	91AA*	231
L	a dominanza di <i>Quercus ilex</i>	9340	236
O	a dominanza di <i>Ostrya carpinifolia</i>		345
PN	a dominanza di <i>Populus nigra</i> e <i>Salix alba</i>	92A0	56
R	Rimboschimenti (perlopiù di conifere)		304
Arbusteti			
AR	Ripariali (<i>Salix eleagnos</i>)	3270,3280, 6430, 92A0	26
G	a dominanza di <i>Spartium junceum</i>		118
JC	a dominanza di <i>Juniperus communis</i>	5130	47
JO	a dominanza di <i>Juniperus oxycedrus</i>		20
RO	a dominanza di <i>Rosa canina</i> e <i>Prunus spinosa</i>		87
VT	Arbusteti sempreverdi prevalentemente sclerofillici		39
Praterie			
AM	a dominanza di <i>Ampelodesmos mauritanicus</i>	330 [6110*,6220*, 5320]	21
AP	a dominanza di <i>Arundo plinii</i>		50
BE	a dominanza di <i>Bromus erectus</i>	6210*, [6110*,6220*]	225
CY	a dominanza di <i>Bromus erectus</i> e <i>Cynosurus cristatus</i>	6210*	25
DA	a dominanza di <i>Arrhenatherum elatius</i>	6510	26
IV	post coltivo a dominanza di <i>Inula viscosa</i>		50
Garighe e vegetazione casmofitica			
GA	Garighe a dominanza di <i>Artemisia alba</i> e <i>Satureja montana</i>	[6110*,6220*]	57
Vegetazione delle superfici erose			
IN	a dominanza di <i>Sulla coronaria</i>		23
SI	a dominanza di <i>Sesleria italica</i>	6210*, [6110*,6220*]	20
Vegetazione delle spiagge e scogliere			
CM	vegetazione delle Spiagge e delle Scogliere	[1210,1240]	25
Vegetazione rupestre delle pareti			
MP	pareti rocciose scarsamente vegetate	[5330,5320,1240]	93
Altro			
AV	alveo fluviale e/o specchi d'acqua	[3270,3280]	24
CO	aree coltivate e/o vegetazione post coltivo		310
IA	Impianti arborei a dominanza di <i>Prunus avium</i>		22
LP	Laghi e pozze	[3140,3150]	42
MA	Mare		67
OE	oliveti		32
U	aree urbane		103
WV	vigneti		31
TOT			2927

3.2.1 Dati topografici

Alcune categorie vegetazionali, più di altre, come nel caso della vegetazione ripariale che si sviluppa all'interno della Gola, possono essere condizionate dalla conformazione del territorio in cui si trovano. Lo strumento con cui è stato possibile estrapolare i dati topografici è il DEM (Digital Elevation Model) 'TINITALY' avente una risoluzione di 10 metri. Gli indici calcolati con il software **SAGA GIS** [SAGA Development Team, 2024] sono stati i seguenti: *Slope*; *Topographic Position Index*; *Topographic Wetness Index*; *Total Insolation*; *Wind Exposition*.

3.3 **Classificazione supervisionata delle serie temporali per la mappatura della vegetazione**

3.3.1 **Caratteristiche delle Sentinel-2**

Le immagini satellitari analizzate in questo lavoro di tesi sono state acquisite dai satelliti Sentinel-2A e 2B, lanciati in orbita rispettivamente nel 2015 e nel 2017 ed appartenenti a Copernicus, il programma di osservazione della Terra dell'Unione Europea.



Figura 3.4: Sentinel-2 in orbita.

Ciascuno di questi satelliti è dotato di uno strumento multispettrale (MSI - Multi-Spectral Instrument) ad alta risoluzione in grado di campionare tredici bande spettrali. Le informazioni contenute nelle bande coastal/aerosol (banda 1), vegetation red edge (banda 8A), water vapour (banda 9) e SWIR-Cirrus (banda 10) non sono rilevanti ai fini di questo lavoro, pertanto è stato deciso di escluderle a priori dalle analisi. La risoluzione spaziale per le bande del visibile e del vicino infrarosso (NIR) è di 10 metri, per quelle del Red edge e dell'infrarosso medio (SWIR) è di 20 metri ed infine per le bande atmosferiche è di 60 metri. La frequenza con cui i satelliti "visitano" i territori è di 5 giorni. Ciò è dovuto al fatto che i due satelliti, pur viaggiando sulla stessa orbita, risultano sfasati l'uno dall'altro di 180°. Infine la risoluzione radiometrica dello strumento MSI è di 12 bit che permette l'acquisizione di immagini in un intervallo di valori di density number compreso tra 0 e 4096[Agency, 2024].

Tabella 3.2: Bande spettrali e risoluzioni spaziali del Sentinel-2.

Sentinel-2 bands	Central wavelength (μm)	Resolution (m)
Band 1 – Coastal aerosol	0.443	60
Band 2 – Blue	0.490	10
Band 3 – Green	0.560	10
Band 4 – Red	0.665	10
Band 5 – Vegetation red edge	0.705	20
Band 6 – Vegetation red edge	0.740	20
Band 7 – Vegetation red edge	0.783	20
Band 8 – NIR	0.842	10
Band 8A – Vegetation red edge	0.865	20
Band 9 – Water vapour	0.945	60
Band 10 – SWIR – Cirrus	1.375	60
Band 11 – SWIR	1.610	20
Band 12 – SWIR	2.190	20

3.3.2 Collezionamento e pre-processamento delle immagini satellitari Sentinel-2

Per l'area di indagine, è stato possibile collezionare 94 immagini satellitari Harmonized Sentinel-2 tramite l'utilizzo della piattaforma **Google Earth Engine** <https://earthengine.google.com/> che consente di eseguire analisi geospaziali [Gorelick et al., 2017]. Per agevolare la ricerca delle scene che ricadono all'interno dell'area di interesse, è stato realizzato un buffer di 100 metri. Le immagini ottenute, sono racchiuse in un arco temporale di tre anni, da gennaio 2021 a dicembre 2023. Dopo aver mascherato le nuvole e le relative ombre, è stata fatta una valutazione visiva delle immagini ritenute processabili. In seguito, è stata individuata la scena master per l'operazione di co-registrazione (sovrapposizione ed allineamento delle immagini) che in questo lavoro, risale al 4 luglio 2022. Generalmente a subire maggiori spostamenti sono le immagini invernali.

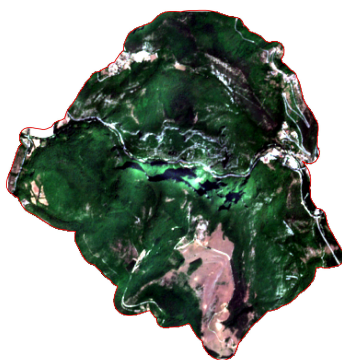


Figura 3.5: Esempio di un'immagine Sentinel-2 (risalente al 24 luglio 2022) di cui è stato fatto il download e le operazioni di pre-processamento.

3.3 Classificazione supervisionata delle serie temporali per la mappatura della vegetazione

Successivamente dalle 94 immagini Sentinel-2, sono state estratte le singole bande che risulteranno poi utili nella costruzione delle serie temporali.

3.3.3 *Functional Data Analysis (FDA)*

Questa tipologia di dati che viene registrata in modo continuo nel tempo può essere rappresentata da curve o funzioni e i metodi statistici per analizzarle sono conosciuti come *Functional Data Analysis (FDA)*. Questo termine è stato coniato per la prima volta da [Ramsay, 1982] e da [Ramsay and Dalzell, 1991]. Per loro natura, i dati funzionali sono infiniti dal punto di vista dimensionale e questa caratteristica comporta non poche difficoltà durante la loro elaborazione le quali, dipendono dalla modalità di campionamento. Allo stesso tempo l'elevata dimensionalità rappresenta una ricca fonte di informazioni, che offre numerose opportunità per la ricerca e l'analisi dei dati [Wang et al., 2016]. La principale caratteristica di FDA è che non considera le funzioni dei dati osservati in maniera separata ma bensì come singole entità unitarie statistiche.

3.3.4 *FPCA e MFPCA*

Una delle tecniche più conosciute ed utilizzate per la rappresentazione e l'interpretazione dei dati funzionali è la *Functional Principal Component Analysis (FPCA)*. Questa procedura infatti, permette di convertire questa tipologia di dati (di dimensioni infinite) in un vettore (di dimensione finita) di *punteggi* o *scores* e di individuare le principali modalità di variazione che vengono proiettate su di uno spazio di ridotte dimensioni [Wang et al., 2016]. Tuttavia, la FPCA non consente di analizzare dati funzionali multivariati o che dipendono da più variabili perchè considera ogni curva come singola osservazione. Esiste però una versione più estesa della FPCA, conosciuta come MFPCA (*Multidimensional- Functional Principal Component Analysis*) che è in grado di individuare le variazioni tra le funzioni, scomponendo i dati in un insieme di componenti principali funzionali multivariate ortogonali o modi di variazione delle funzioni (*eigenfunctions*) a cui si aggiungono gli *eigenvalues* (misurano la variazione spiegata da ogni dimensione) e gli *scores*. In questo modo, si riesce ad avere una rappresentazione sintetica e chiara dei dati, con un punteggio per ogni componente multivariata assegnato a ciascuna osservazione.

In questo lavoro, la MFPCA verrà applicata sia per esplorare le curve che sono nei punti di Conero, Carpegna e Furlo che nei pixel della ZSC Gola di Frasassi.

3.3.5 *Random Forest (RF)*

Il *Random Forest (RF)* [Breiman, 2001] è un algoritmo di apprendimento supervisionato in grado di costruire un grande numero di alberi decisionali utilizzando, durante il processo di training, un sottoinsieme di campioni e variabili selezionati casualmente. Per suddividere le varie classi, l'algoritmo stabilisce un intervallo soglia in base a par-

ticolari caratteristiche riscontrate tra i dati. E' molto apprezzato nel telerilevamento, in particolare per la classificazione e per la regressione, perchè consente di gestire in maniera flessibile, l'elevata dimensionalità dei dati e la multicollinearità. Inoltre, non è sensibile al rumore e all'overfitting [Belgiu and Drăguţ, 2016].

3.3.6 Classificazione supervisionata (Supervised Classification)

Nella classificazione supervisionata sono stati utilizzati come dati di input gli scores della MFPCA e come classificatore il Random Forest. Al fine di ottenere un modello robusto ed accurato, è stato necessario procedere con una pre-elaborazione delle informazioni estrapolate nelle precedenti fasi. Infatti, spesso può accadere, quando si ha a disposizione un elevato quantitativo di dati, che questi siano sbilanciati e ciò comporterebbe ad una sovrastima nei risultati.

Per agevolare la ricerca degli indici di vegetazione, che meglio individuano e discriminano le diverse tipologie vegetazionali associate ai punti esterni e garantiscono un'elevata accuratezza, MFPCA e RF sono state racchiuse all'interno di una strategia di selezione in avanti meglio conosciuta come "*Forward Selection*". Le principali ragioni per cui è stata introdotta questa procedura è che ancora oggi, la selezione si basa sull'esperienza umana ed è anche influenzata dall'area di studio. Inoltre, non tutti gli indici possono essere rilevanti ai fini della classificazione. La formula che è stata adottata per calcolare gli indici è la seguente:

- F03: $(A/B) / (A+B)$

Gli indici che verranno selezionati con la strategia di *Forward Selection* applicata sui punti (profili stagionali) dei Reference Data saranno poi testati nell'area della Gola di Frasassi.

3.3.7 Accuratezza della Mappatura (*Ten Fold Cross Validation*)

Per valutare il modello predittivo che è stato costruito, è stata utilizzata come tecnica di ri-campionamento la *Ten Fold Cross Validation*. Uno degli strumenti più comuni che può essere adottato per esprimere l'accuratezza della classificazione è la cosiddetta *Confusion Matrix* o *Error Matrix* e da cui è possibile ricavare diverse metriche quali ad esempio: *Overall Accuracy (OA)*, *Producer Accuracy (PA)*, *User Accuracy (UA)* ed il *K coefficient*. La OA fornisce una stima globale in termini percentuali dell'accuratezza della mappa finale e si ottiene dalla somma di tutti i valori che si trovano sulla diagonale della matrice. Se si raggiunge un valore uguale o maggiore all'80%, si ottiene un risultato più che soddisfacente. La PA è l'accuratezza della mappa dal punto di vista del produttore. Si calcola dividendo il numero di pixel classificati correttamente per ciascuna categoria (sulla diagonale) per il totale della colonna. La UA è invece la probabilità con cui un pixel classificato in una determinata categoria rappresenti effettivamente la realtà. Si ottiene dividendo il numero di pixel classificati correttamente per ogni categoria per il totale della riga.

3.3 Classificazione supervisionata delle serie temporali per la mappatura della vegetazione

Infine il *K coefficient* permette di valutare se la classificazione sia stata migliore rispetto ad una casuale e può assumere valori tra -1 e 1.

3.3.8 Software utilizzati

Per l'elaborazione dei dati ci si è avvalsi dell'utilizzo del software gratuito ed open source **R** [R Core Team, 2024] e, nello specifico, delle seguenti librerie (o packages):

1. **raster** [Hijmans, 2020] per il processamento delle immagini selezionate;
2. **forecast** [Hyndman and Khandakar, 2008] e **mgcv**[Wood, 2011] per la costruzione delle serie temporali;
3. **fdaPace** [Yao et al., 2005] ed **MFPCA** [Happ and Greven, 2018] per la scomposizione e riduzione dello spazio funzionale multivariato;
4. **randomForest** [Liaw and Wiener, 2002] per la classificazione supervisionata.



Figura 3.6: Logo di R.

Tutte le fasi sopra descritte sono state effettuate con i codici di programmazione in **R** e **Google Earth Engine** sviluppati dal Dott. Simone Pesaresi in collaborazione con il Dott. Giacomo Quattrini ed il Prof. Adriano Mancini frutto di una collaborazione tra l'area botanica del Dipartimento D3A e DII dell'Università Politecnica delle Marche.

Per la visualizzazione, interpretazione e tematizzazione delle mappe ottenute invece è stato usato il software **QGIS versione 3.34 Prizren**.



Figura 3.7: Logo di QGIS.

Capitolo 3 Materiali e Metodi

Il codice per la creazione della matrice di confusione è disponibile sul seguente sito <https://blogs.fu-berlin.de/reseda/accuracy-statistics-in-r/>.

Capitolo 4

Risultati

4.1 Selezione degli Indici di Vegetazione

Dalla procedura di *Forward Selection* applicata alle serie temporali nei punti del Reference Data che contiene ben 34 distinte categorie di vegetazione, sono state selezionate le migliori serie temporali attraverso la formula F03 con valori non standardizzati e raggiungendo, con il *Random Forest* come classificatore, oltre l'82% di accuratezza.

Tabella 4.1: Le 7 migliori serie temporali selezionate dalla *Forward Selection* con formula F03, non standardizzata e classificatore RF.

"x"	
"1"	"S2_F03_11_06_00_00"
"2"	"S2_F03_11_05_00_00"
"3"	"S2_F03_04_03_00_00"
"4"	"S2_F03_03_02_00_00"
"5"	"S2_F03_07_05_00_00"
"6"	"S2_F03_11_10_00_00"
"7"	"S2_F03_07_06_00_00"

Analizzando la tabella 4.1, è possibile notare che la banda più ricorrente è la banda 11 (SWIR) a cui seguono le bande del visibile che quindi potrebbero ricoprire un ruolo importante nella discriminazione della vegetazione.

4.2 Mappa della Gola di Frasassi

La mappa finale della Gola di Frasassi è stata ottenuta dall'addestramento del classificatore *Random Forest* con i dati esterni, sulla base degli indici di vegetazione selezionati dalla procedura di *Forward Selection*. Inoltre, è importante sottolineare che il classificatore non si è allenato su tutte le categorie riferite alle aree esterne. Alcune di queste sono state escluse e racchiuse all'interno della seguente tabella:

Tabella 4.2: Elenco delle categorie escluse dal Reference Data. Le etichette, la fisionomia e l'habitat sono riportati nella tabella 3.1 del Reference Data.

LABEL
Boschi
AQ
FR
Arbusteti
AR
VT
Praterie
AM
AP
CY
DA
Vegetazione delle superfici erose
IN
Vegetazione delle spiagge e scogliere
CM
Altro
IA
MA

Il principale motivo per cui queste categorie sono state scartate, è perchè sono specifiche delle aree esterne. Un esempio è quello della vegetazione delle spiagge e delle scogliere presente solo nel Parco del Conero e non in quello della Gola di Frasassi. Altro caso simile è quello delle cerrete, le quali, stando agli studi condotti nell'area di Frasassi, non erano finora mai state riscontrate mentre lo erano a Carpegna.

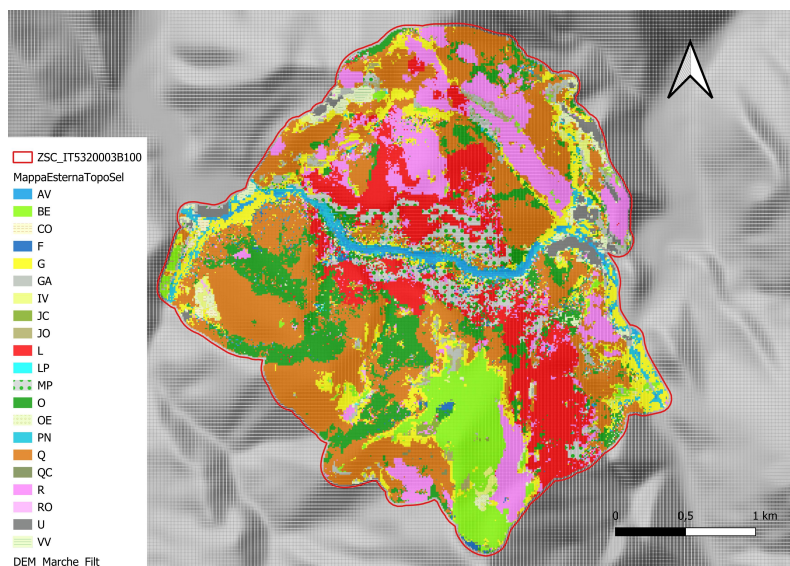


Figura 4.1: Mappa della Gola di Frasassi. Le etichette presenti in legenda fanno riferimento alla tabella 3.1 del Reference Data.

Dall'osservazione della mappa è possibile distinguere la distribuzione o meglio, il pattern spaziale, delle singole categorie vegetazionali. Le più diffuse sono soprattutto

4.2 Mappa della Gola di Frasassi

le formazioni boschive come la lecceta, l'ostrieto, il querceto ed i rimboschimenti. Meno diffuse invece sono le formazioni prative ed arbustive che si concentrano soprattutto su Monte Valmontagnana e le formazioni ripariali. Quest'ultime, in particolare trovano posto all'interno della Gola di Frasassi.

Qui di seguito, è stato invece possibile mettere a confronto in maniera visiva la mappa ottenuta in questo lavoro con quella dell'articolo di [Pesaresi et al., 2022]. Quest'ultima è stata realizzata con un classificatore addestrato con punti ricadenti esclusivamente all'interno dell'area di studio.

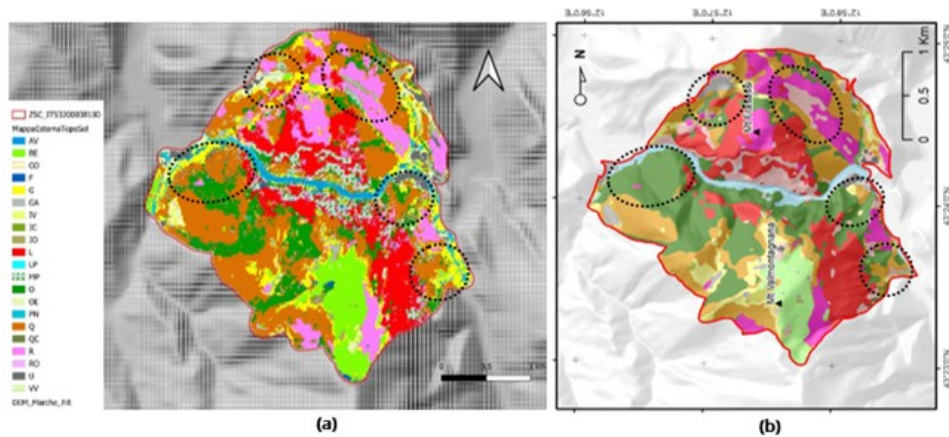


Figura 4.2: Confronto visivo tra la mappa ottenuta in questo lavoro (a) e la mappa prodotta nell'articolo di [Pesaresi et al., 2022] (b). In nero sono cerchiato quelle aree dove c'è maggiore discordanza tra le due mappe e nelle quali è stato deciso di effettuare dei sorvoli con il drone.

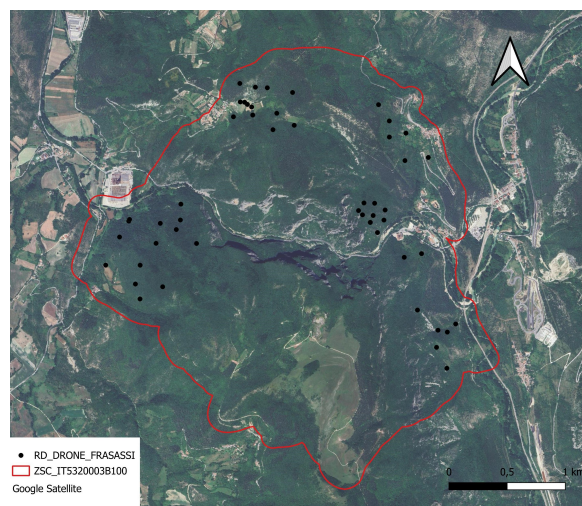


Figura 4.3: Aree sorvolate con il drone (punti in colore nero) per ispezionare le differenze tra le differenze indicate in figura 4.2. L'autorizzazione è avvenuta attraverso il rilascio del nulla osta (Prot. n. 0253749 del 31/10/2024) da parte del Parco Naturale Regionale della Gola della Rossa e di Frasassi.

Capitolo 4 Risultati

Le incongruenze sono state ulteriormente confermate dalla sovrapposizione dei punti di validazione alla mappa ragione per cui si è ritenuto necessario effettuare delle verifiche in campo. L'utilizzo del drone è stato fondamentale per raggiungere le aree di difficile accesso a causa della conformazione del terreno.

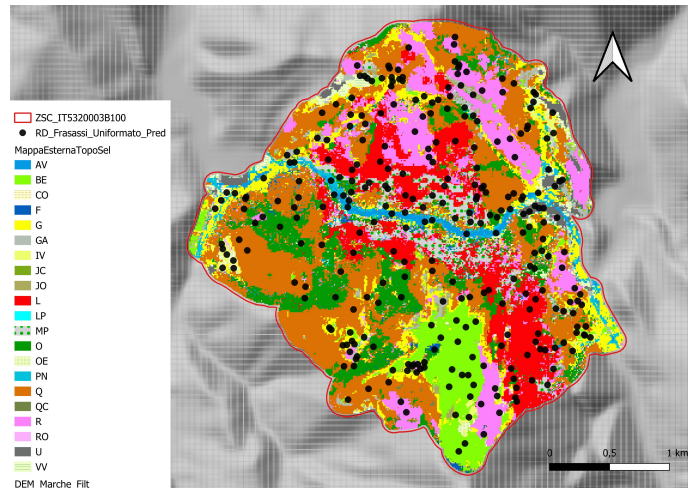


Figura 4.4: Mappa della Gola di Frasassi. Le etichette presenti in legenda fanno riferimento alla tabella 3.1 del Reference Data. I punti colorati in nero costituiscono invece il validation set.

4.3 Matrice di confusione

La matrice di confusione ha permesso di valutare le prestazioni del modello di classificazione adottato. In particolare, sono state confrontate le categorie mappate con quelle predette dal modello. Tra le metriche già citate che sono state calcolate vi sono: *Overall Accuracy (OA)*, *Producer Accuracy (PA)*, *User Accuracy (UA)* ed il *K coefficient*. Il codice modificato, da cui è stata ricavata la matrice, è disponibile in <https://blogs.fuberlin.de/reseda/accuracy-statistics-in-r/>.

Tabella 4.3: Matrice di confusione.

		Reference														
		AV	BE	CO	G	GA	IV	J	L	MP	O	PN	Q	R	U	UA
Prediction	AV	7	0	0	0	0	0	0	0	1	1	0	0	0	0	77.78
	BE	0	14	2	0	0	0	0	1	4	0	0	0	0	0	66.67
	CO	0	0	8	0	0	0	0	0	0	0	0	0	0	0	100.00
	G	0	0	2	16	0	0	3	0	1	0	3	1	0	0	61.54
	GA	0	0	0	0	13	0	0	0	0	0	0	0	0	0	100.00
	IV	0	2	0	0	0	1	0	0	0	0	0	0	0	0	33.33
	J	0	0	0	0	0	0	11	0	0	0	0	0	0	0	100.00
	L	0	0	0	0	0	0	0	35	0	0	0	0	0	0	100.00
	MP	0	0	0	0	0	0	0	0	21	1	0	0	0	0	95.45
	O	0	0	0	0	0	0	0	1	0	31	0	0	0	0	96.88
	PN	0	0	0	0	0	0	0	0	0	1	5	0	0	0	83.33
	Q	0	0	0	1	0	0	0	0	0	19	0	56	1	0	72.73
	R	0	0	0	0	0	0	0	0	0	5	0	0	31	0	86.11
	U	0	0	0	0	0	0	0	0	0	0	0	0	0	14	100.00
	PA	100	87.5	66.67	94.12	100	100	78.57	94.59	77.78	53.45	62.5	98.25	96.88	100	

Dalla somma di tutti i valori posti sulla diagonale, è stata ottenuta un' *Overall Accuracy* del 84 %. Se si vanno ad analizzare nel dettaglio le restanti metriche, è possibile notare che le *Producer Accuracy* sono tendenzialmente alte per tutte le categorie ad eccezione dei coltivi, degli ostrieti e dei pioppeti. Per quanto riguarda invece le *User Accuracy* le più basse si registrano per i brometi, per i ginestreti ed infine per le praterie post-coltivo (es. quelle a dominanza di *Inula viscosa*). Generalmente se si ha una *PA* elevata ma una bassa *UA* significa che il modello è riuscito a rispettare l'idea di chi ha prodotto la mappa, classificando correttamente i campioni nelle rispettive classi ma, al tempo stesso, molti di questi campioni non appartengono effettivamente a quella determinata classe. In questo caso viene detto che la classificazione tende a sovrastimare. Al contrario, quando si ha una *UA* alta ma una *PA* bassa, significa che, seppur il modello non riesca a classificare correttamente una determinata classe, l'utente finale che dispone della mappa ha la certezza che ci sia una effettiva corrispondenza tra la carta e la realtà anche se la categoria risulta probabilmente sottostimata o sottoprodotta. Infine ci sono categorie quali ad esempio le garighe oppure l'urbano che hanno sia una *PA* che una *UA* pari al 100 %.

Capitolo 5

Discussioni

5.1 Punti di forza e dominio delle applicazioni

La realizzazione di un sistema esperto automatico di mappatura della vegetazione, ha reso ancora più efficiente la metodologia avviata in [Pesaresi et al., 2020]. Per quanto riguarda le tempistiche di produzione della mappa della Gola di Frasassi, per il download delle immagini, per la loro elaborazione e per la pianificazione e successiva verifica dei sorvoli effettuati con il drone in campo, è stato impiegato all'incirca un mese. Si tratta quindi di un tempo relativamente breve nel quale è possibile ottenere mappe della vegetazione accurate, ripetibili e aggiornabili anche "annualmente", rispettando ampiamente l'arco temporale imposto dalla Direttiva Habitat 92/43/CEE. Quest'ultima infatti obbliga ciascuno Stato Membro dell'Unione Europea ad effettuare, ogni 6 anni, attività di monitoraggio i cui esiti vengono riportati all'interno di un report redatto da tecnici che hanno preso parte a tale attività al fine di valutare lo stato di conservazione degli ecosistemi e degli habitat [Rapinel et al., 2020]. Nonostante ciò, la maggior parte delle cartografie della regione Marche già esistenti risale ai primi anni del 2000, inclusa quella della Gola di Frasassi e sono state realizzate perlopiù con la metodologia tradizionale.

Poichè all'interno dell'area di studio sono presenti categorie riconosciute dalla Direttiva nell'allegato I come habitat, alcuni dei quali prioritari, avere a disposizione mappe dettagliate ed aggiornate sicuramente potrebbe favorire una gestione più attiva delle aree protette da parte degli enti competenti. Inoltre, a differenza delle cartografie tradizionali, le mappe ottenute con la metodologia proposta in questo lavoro, forniscono sia il livello di accuratezza globale che quello di ciascuna categoria. Ciò facilita non solo la loro comprensione, ma consente anche di mettere a confronto tali mappe con futuri lavori.

L'impiego del drone, nel corso di questo lavoro di tesi, ha consentito di incrementare il quantitativo di dati utili a valutare la classificazione, oltre a verificare gli eventuali punti di incoerenza con la mappa ottenuta e ad offrire la possibilità di osservare la vegetazione da un'altra prospettiva. La visione dall'alto infatti facilita l'individuazione dell'estensione di una comunità. Inoltre, tutte le foto che sono state scattate con il drone possono essere raccolte in archivi consultati liberamente da chiunque.



Figura 5.1: Esempio di un'immagine di bosco di roverella (*Quercus pubescens* L.) scattata con il drone durante un sopralluogo effettuato in campo il 6 novembre 2024.

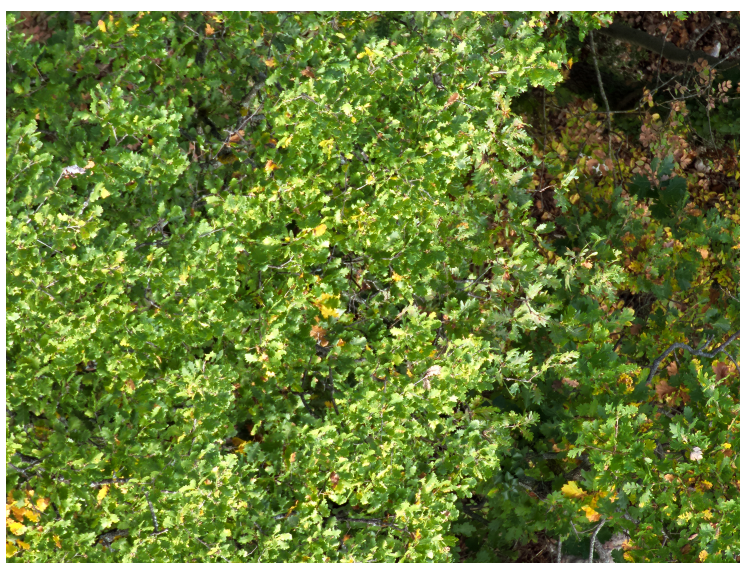


Figura 5.2: Attraverso uno zoom ottico, è possibile ottenere immagini che consentono l'identificazione delle specie vegetali e quindi stabilire la specie fisionomizzante che, in questo caso, è rappresentata dalla roverella (Q) (*Quercus pubescens* L.).

5.2 Limiti e futuri sviluppi

Nonostante i promettenti risultati, è necessario fare alcune considerazioni in merito non solo all'area target che è stata scelta ma anche alla quantità e alla qualità dei dati raccolti.

Per quanto riguarda il Reference Data, averne costruito uno costituito internamente da punti esterni si è dimostrato essere una sfida impegnativa perchè ha richiesto di armonizzare le etichette associate alle diverse categorie vegetazionali riferibili alle aree del Conero, del Furlo e di Carpegna. Risulta importante infatti, avere a disposizione dei dati che siano ben bilanciati perchè altrimenti possono diventare essi stessi fonte di errore. L'operazione di armonizzazione invece non viene richiesta nel caso in cui si debba costruire una banca dati per una specifica area di studio. Ad ogni modo, lo sforzo compiuto per realizzare un unico dataset di riferimento potrebbe comunque rivelarsi estremamente utile in futuro perchè permetterebbe di recarsi in aree nuove senza avere la necessità di costruirne uno ex-novo.

Andando però a confrontare i dati di riferimento con cui è stata realizzata la mappa finale e quelli utilizzati per validare il modello, è emerso che i primi hanno una qualità ed un livello di dettaglio nettamente superiori poichè l'utilizzo del drone consente di ottenere dati di maggiore precisione (centimetrica) rispetto al dispositivo GPS utilizzato in precedenza.

Un ulteriore aspetto da considerare nei lavori futuri è l'utilizzo del grado di incertezza, dettagliato nella matrice di confusione, per calcolare l'incertezza in ettari relativa alle superfici delle categorie mappate. Un esempio di codice R per eseguire queste computazioni è disponibile a questo link: <https://blogs.fu-berlin.de/reseda/area-adjusted-accuracies/>. Le superfici degli habitat, corredate dall'incertezza attribuibile al modello utilizzato, rappresenterebbero un'informazione aggiuntiva rispetto a quelle solitamente riportate nei Formulari Standard. Questi dati sarebbero particolarmente utili per gli enti gestori nella pianificazione e gestione delle risorse.

Dal confronto tra la mappa ottenuta in questo lavoro e quella relativa all'articolo di [Pesaresi et al., 2022] è risultato che tra le categorie particolarmente sovrastimate figura quella dei ginestreti. Infatti, nonostante la *Producer Accuracy* sia relativamente alta, la *User Accuracy* è pari al 61,5%. Ciò significa che, seppure i ginestreti vengano classificati correttamente, essi possono contenere altro al loro interno ed in genere si tratta di categorie poco rappresentate che difficilmente vengono riconosciute dal sistema esperto. E' il caso, ad esempio, dei boschi ripariali. Quelli presenti nelle tre aree esterne prese in considerazione, non sono in grado di allenare correttamente il sistema. Per ovviare a questo limite sarebbe opportuno apportare modifiche al Reference Data di partenza, prendendo in analisi aree di pertinenza prettamente fluviale. Un'altra categoria che viene spesso inclusa nei ginestreti, ma anche nei querceti di roverella, è rappresentata dai preboschi ad orniello che non sono mai stati considerati nelle precedenti mappature. L'orniello (*Fraxinus ornus L.*) rappresenta una delle popolazioni più attive che dominano le fasi della transizione vegetazionale, da prateria a bosco [Ballerini and Biondi, 2002] (Vedi figure 5.3 e 5.4). In futuro, sarebbe indicato creare una categoria apposita, aggiornato il Reference Data, e quindi il sistema esperto rendendolo capace di mappare anche questa categoria.



Figura 5.3: Vista dall'alto di un prebosco ad orniello. La foto è stata scattata il 6 novembre 2024.



Figura 5.4: Zoom 16x sul prebosco ad orniello.

Come già riscontrato nei precedenti lavori ([Vignoli et al., 2021]; [Quattrini et al., 2020]), anche in questa tesi è emerso che le categorie maggiormente soggette a confusione sono gli ostrieti e i querceti. Tuttavia, in altri studi, come quelli relativi alle aree del Carpegna [Casavecchia et al., 2023b], del Conero [Casavecchia et al., 2023a] e del Furlo [Casavecchia et al., 2023c], queste due formazioni erano state accuratamente discriminate. Per approfondire la questione, si è deciso di effettuare ulteriori sorvoli con il drone.

L'analisi delle foto acquisite ha rivelato che, in diversi casi, risulta difficile attribuire una fisionomia chiara ai luoghi osservati. Questo perché carpino nero (*Ostrya carpini-*

folia Scop.) e roverella (*Quercus pubescens L.*) tendono spesso a codominare, rendendo complesso stabilire con certezza quale specie sia dominante e fisionomizzante (Vedi figure 5.5 e 5.6).



Figura 5.5: Vista dall'alto di una delle aree in cui è presente una codominanza tra *Quercus pubescens L.* ed *Ostrya carpinifolia Scop.*. La foto è stata scattata il 6 novembre 2024.

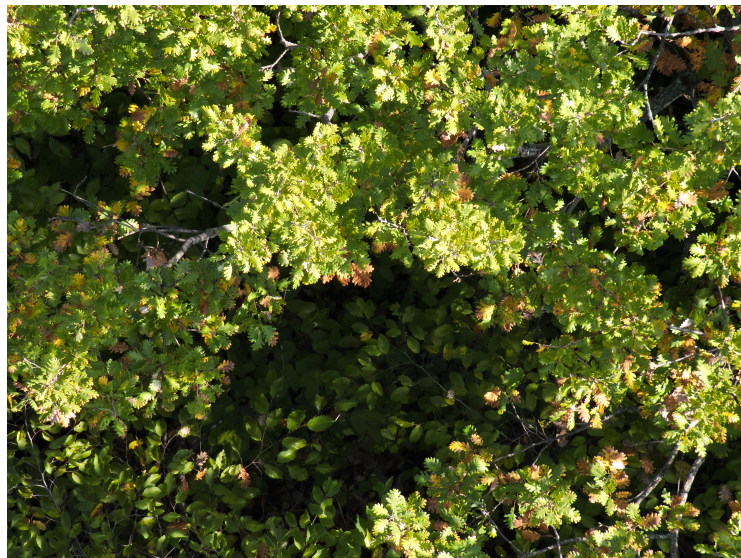


Figura 5.6: Dettaglio della foto precedente. In primo piano la roverella.

Ad ovest della Gola, mentre stavamo verificando la discordanza tra ostrieti e querceti, ci siamo accorti che in realtà era presente una categoria non contenuta nel sistema esperto, perchè tra quelle scartate prima di procedere con la classificazione,

nello specifico quella delle cerrete. Per questo motivo sarebbe quindi opportuno pensare di considerare nei prossimi lavori anche questa categoria.



Figura 5.7: Vista dall'alto di una delle aree in cui domina il cerro (*Quercus cerris* L.). La foto è stata scattata il 6 novembre 2024.

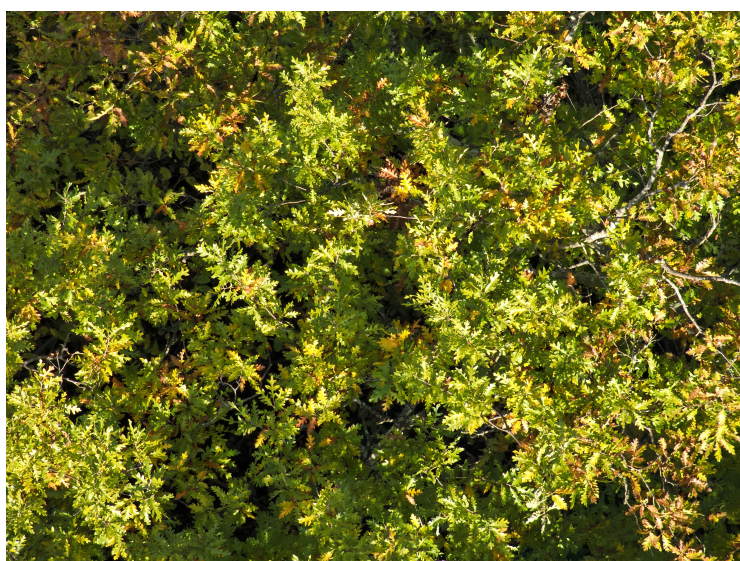


Figura 5.8: Dettaglio del bosco di cerro ottenuto con lo zoom 16x con cui è equipaggiato il drone Matrice 30 DJI.

In un'altra area caratterizzata da discrepanze tra ostrieti e querceti di roverella, i sorvoli hanno rivelato la presenza di formazioni dominate dal cerro (*Quercus cerris* L.) (Vedi figure 5.7 e 5.8). Questa categoria non era inclusa nel sistema esperto, poiché era stata scartata durante la fase preliminare di classificazione. Per questo motivo, sarebbe opportuno considerare l'inclusione del cerro nei prossimi lavori.

5.2 Limiti e futuri sviluppi

Per migliorare l'accuratezza della mappatura nella Gola di Frasassi, attualmente basata esclusivamente su un addestramento fatto con un Reference Data costituito di punti esterni all'area target, si potrebbe riassegnare parte del validation set (Vedi figura 4.4) come training data per affinare ulteriormente i risultati. Il sistema esperto automatico potrebbe essere esteso per esplorare nuove aree e ampliare l'analisi spaziale. Un'ulteriore sfida interessante sarebbe quella di effettuare analisi retrospettive, in grado di rilevare eventuali cambiamenti nella distribuzione spaziale della vegetazione nel tempo, offrendo così una comprensione più approfondita delle dinamiche vegetazionali del territorio.

Capitolo 6

Conclusioni

In conclusione, i risultati di questo lavoro hanno dimostrato che è possibile sviluppare un sistema esperto di classificazione basato sulle variazioni stagionali spettrali caratteristiche della vegetazione, capace di effettuare una mappatura automatica accurata e rapida di un'area target. L'utilizzo di un ampio set di dati esterni e di un sistema esperto addestrato su numerose aree e categorie ha consentito di mappare anche categorie vegetazionali poco rappresentate nell'area target (es. ginepri). Inoltre, ha permesso di individuare la possibile presenza di categorie vegetazionali precedentemente non mappate, come i boschi di cerro, che richiedono ulteriori approfondimenti. L'integrazione dei dati acquisiti con il drone a quelli raccolti nel corso degli anni durante i vari sopralluoghi ha comportato numerosi vantaggi che, nel loro complesso, hanno reso ancora più efficiente la metodologia di mappatura sviluppata da [Pesaresi et al., 2020]. L'applicazione di questa metodologia offrirebbe notevoli vantaggi per il monitoraggio ambientale, consentendo una mappatura continua della vegetazione nel tempo e il monitoraggio delle dinamiche e dei cambiamenti ad essa associati. Inoltre, l'integrazione con foto geolocalizzate acquisite tramite drone permetterebbe di tracciare in modo preciso lo stato degli habitat nel corso degli anni. Questa consapevolezza risulterebbe particolarmente utile per la recente legge approvata dall'Unione Europea, la *Nature Restoration Law*, https://environment.ec.europa.eu/topics/nature-and-biodiversity/nature-restoration-law_en. Nella legge vengono sollecitati tutti gli stati membri a mobilitarsi al fine di ripristinare gli ecosistemi, gli habitat e le specie nelle aree terrestri e marine. E' stato infatti riscontrato che oltre l'80 % degli habitat in Europa risulta degradato. In particolare, deve essere posta maggiore attenzione su quelli che hanno una maggiore capacità di catturare ed immagazzinare carbonio al fine di prevenire e ridurre l'impatto dei disastri naturali. Uno degli obiettivi principali di questa legge è quello di riuscire a ripristinare almeno il 20% degli habitat degradati entro il 2030. Risulta quindi fondamentale tenere in considerazione il fattore tempo. Il successo del ripristino, dipende dalla capacità di agire tempestivamente al fine di limitare il progredire del degrado. Allo stesso tempo, è necessario essere consapevoli che per consentire il completo recupero degli ecosistemi è necessario un tempo relativamente lungo.

In questo studio, è stato possibile dimostrare come l'utilizzo simultaneo di immagini satellitari, riprese da drone e cartografie possa fare la differenza nel garantire una

Capitolo 6 Conclusioni

migliore comprensione del territorio oggetto di analisi.

I limiti che sono stati riscontrati diventano invece delle sfide per migliorare ulteriormente il sistema di classificazione adottato.



Figura 6.1: Monte Valmontagnana. La foto è stata scattata il 6 novembre 2024.

Capitolo 7

Appendice

7.1 Armonizzazione dataset

Tabella 7.1: Processo di armonizzazione tra i due dataset di confronto.

MAPPA	VD	IDS	VERIFICA	MOD	NOTE
AV	O	300	Non è da modificare (Effetto sale-pepe)		
AV	PN	206	PN sostituiti con AV		
AV	PN	208	PN sostituiti con AV		
AV	PN	209	PN sostituiti con AV		
AV	PN	210	PN sostituiti con AV		
AV	PN	211	PN sostituiti con AV		
AV	PN	212	PN sostituiti con AV		
AV	PN	213	PN sostituiti con AV		
BE	CO	196	Verificare	Diretta	Rosenga
BE	CO	168	Verificare	Diretta	M. Valmontagnana
BE	JC	133	Eliminato (E' in zona di transizione)		
BE	L	48	Non è da modificare (Effetto sale-pepe)		
BE	MP	288	Verificare	Drone	M. Valmontagnana
BE	MP	260	Non è da modificare (Effetto sale-pepe)		
BE	MP	265	Verificare	Diretta	M. Frasassi
CO					
F	O	52	Non è da modificare (Effetto sale-pepe)		
G	CO	169	Non concorde (Da non verificare)		
G	CO	200	Non concorde (Da non verificare)		
G	JC	173	Verificare	Drone	M. Valmontagnana
G	JC	243	Verificare	Drone	M. Valmontagnana
G	JC	124	Verificare	Drone	M. Valmontagnana
G	MP	266	Eliminato (Dimensioni ridotte)		
G	MP	267	Non concorde (Da non verificare)		
G	MP	286	Eliminato (Situaz. Eterogenea)		
G	PN	249	Eliminato		
G	PN	203	Non concorde (Da non verificare)		
G	PN	214	Non concorde (Da non verificare)		
G	PN	304	Non concorde (Da non verificare)		
G	PN	215	Spostato		
G	Q	102	Spostato		
G	Q	296	Non concorde (Da non verificare)		
GA	MP	83	MP sostituiti con GA		
GA	MP	100	Eliminato (Dimensioni ridotte)		
GA	MP	105	MP sostituiti con GA		
GA	MP	252	MP sostituiti con GA		
GA	MP	268	MP sostituiti con GA		
GA	MP	269	MP sostituiti con GA		
GA	MP	270	MP sostituiti con GA		
GA	MP	271	MP sostituiti con GA		

Capitolo 7 Appendice

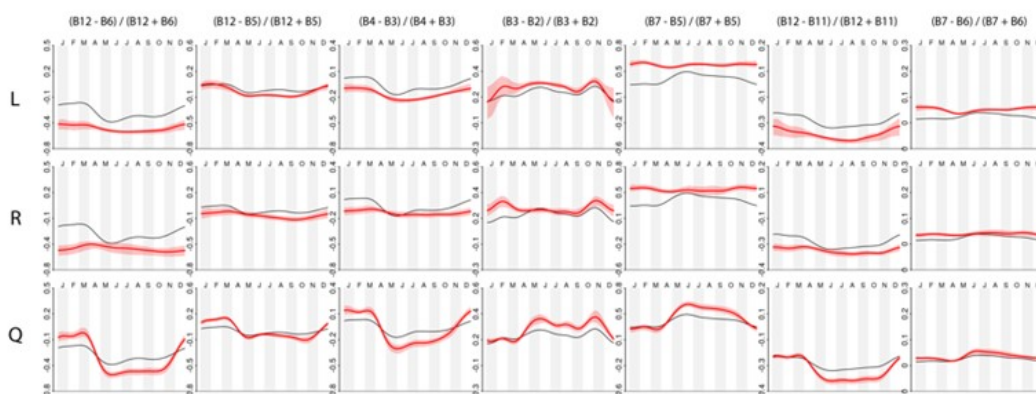
MAPPA	VD	IDS	VERIFICA	MOD	NOTE
GA	MP	272	MP sostituiti con GA		
GA	MP	282	MP sostituiti con GA		
GA	MP	283	MP sostituiti con GA		
GA	MP	284	MP sostituiti con GA		
GA	MP	305	MP sostituiti con GA		
GA	MP	306	MP sostituiti con GA		
GA	U	228	Eliminato		
GA	U	230	Spostato		
GA	U	233	Spostato		
GA	U	235	Spostato		
IV	BE	75	Verificare	Diretta	Pierosara
IV	BE	145	Verificare	Diretta	M. Valmontagnana
IV	CO	182	Spostato		
IV	CO	202	Verificare	Drone	Spineto
IV	MP	278	Spostato		
IV	MP	307	Eliminato		
JC					
JO	MP	281	Eliminato		
L	MP	254	Eliminato		
L	O	49	Eliminato		
L	R	191	Spostato		
LP	CO	198	Eliminare LP dalla mappa finale		
MP	L	46	Spostato		
MP	L	86	Spostato		
MP	L	87	Spostato		
MP	O	38	Spostato		
MP	O	94	Spostato		
MP	O	98	Non concorde (Da non modificare)		
O	L	142	Verificare	Drone	M. Valmontagnana
O	MP	285	Spostato		
O	Q	39	Spostato		
O	R	108	Spostato		
PN	O	47	Non concorde (Da non modificare)		
Q	G	119	Verificare	Drone	M. Valmontagnana
Q	G	237	Spostato		
Q	MP	92	Verificare	Drone	
Q	MP	95	Eliminato		
Q	MP	264	Eliminato		
Q	MP	279	Spostato		
Q	O	111	Verificare	Drone	Cerqueto
Q	O	109	Spostato		
Q	O	114	Verificare	Drone	Pierosara
Q	O	78	Verificare	Diretta	Pierosara
Q	O	90	Verificare	Drone	M. Frasassi
Q	O	88	Verificare	Drone	M. Frasassi
Q	O	96	Verificare	Drone	M. Frasassi
Q	O	152	Verificare	Diretta	M. Valmontagnana
Q	O	29	Verificare	Drone	M. Valmontagnana

7.2 Profili stagionali medi

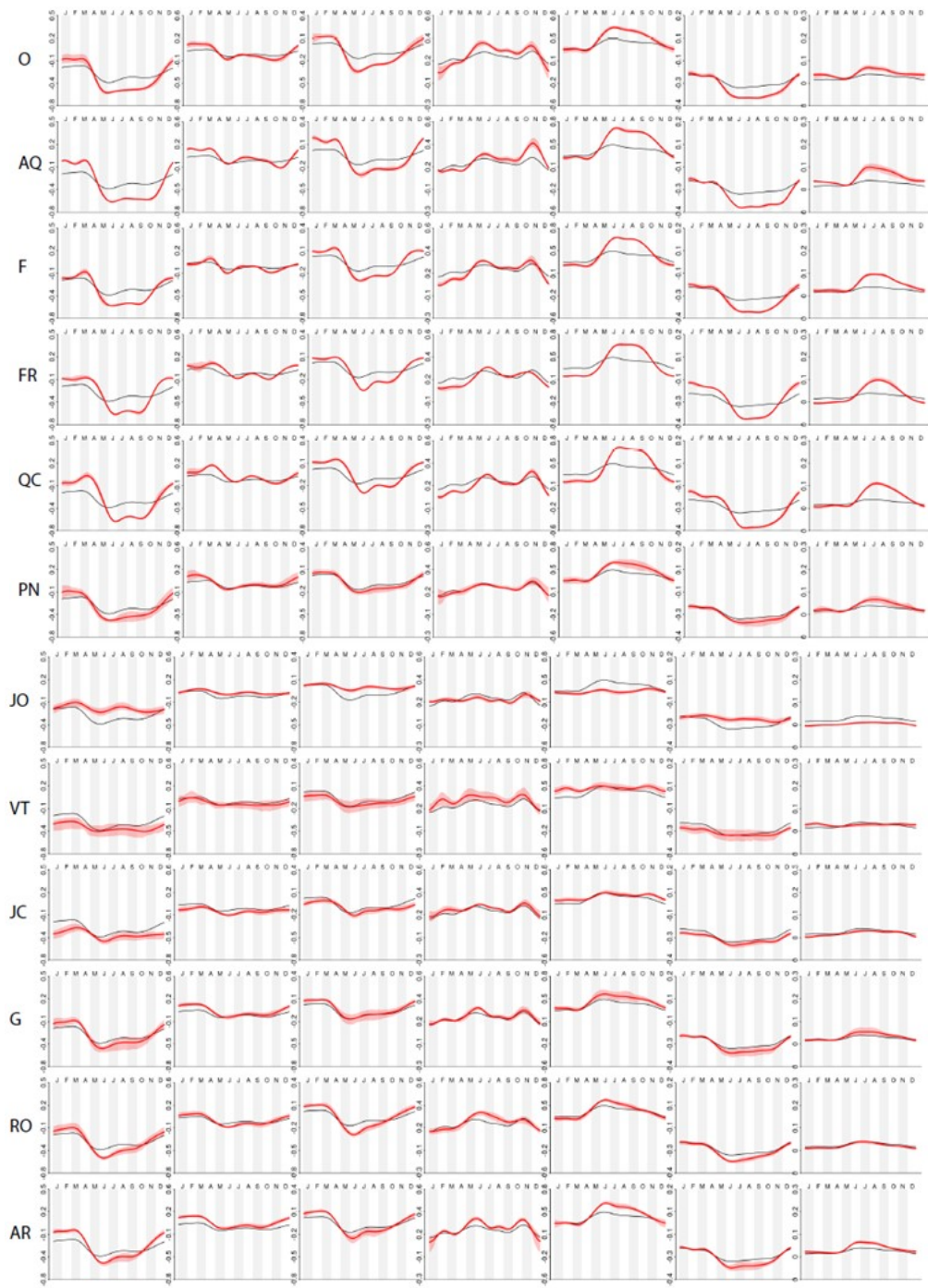
MAPPA	VD	IDS	VERIFICA	MOD	NOTE
Q	O	6	Verificare		
Q	O	21	Verificare	Drone	Mutilog spa Genga
Q	O	20	Verificare	Drone	Mutilog spa Genga
Q	O	25	Verificare		
Q	O	292	Verificare	Drone	M. Frasassi
Q	O	161	Spostato		
Q	O	160	Verificare	Drone	M. Valmontagnana
Q	O	217	Verificare	Diretta	Via Palombare di Ferro
Q	O	216	Verificare	Drone	Via Palombare di Ferro
Q	O	293	Verificare	Drone	Via Palombare di Ferro
Q	O	218	Verificare		
Q	O	15	Spostato		
Q	O	4	Spostato		
Q	O	34	Verificare	Drone	Mutilog spa Genga
Q	O	33	Verificare		
Q	O	19	Verificare	Drone	Mutilog spa Genga
Q	O	18	Verificare	Drone	Mutilog spa Genga
R	JC	125	Spostato		
R	JC	127	Spostato		
R	L	104	Spostato		
R	L	221	Spostato		
R	MP	253	Verificare		
R	MP	289	Spostato		
R	O	41	Non è da modificare (Effetto sale-pepe)		
R	O	42	Non è da modificare (Effetto sale-pepe)		
R	O	62	Verificare (circondate da rimboscimento)	Drone	Rosenga
R	O	68	Verificare (circondate da rimboscimento)	Drone	Rosenga
R	O	164	Verificare (circondate da rimboscimento)	Drone	M. Valmontagnana
U	G	163	Spostato		
VV	CO	192	Verificare	Drone	Rosenga
VV	CO	193	Verificare	Drone	Rosenga
VV	CO	195	Verificare	Drone	Rosenga
VV	CO	197	Eliminato		
VV	Q	37	Spostato		

7.2 Profili stagionali medi

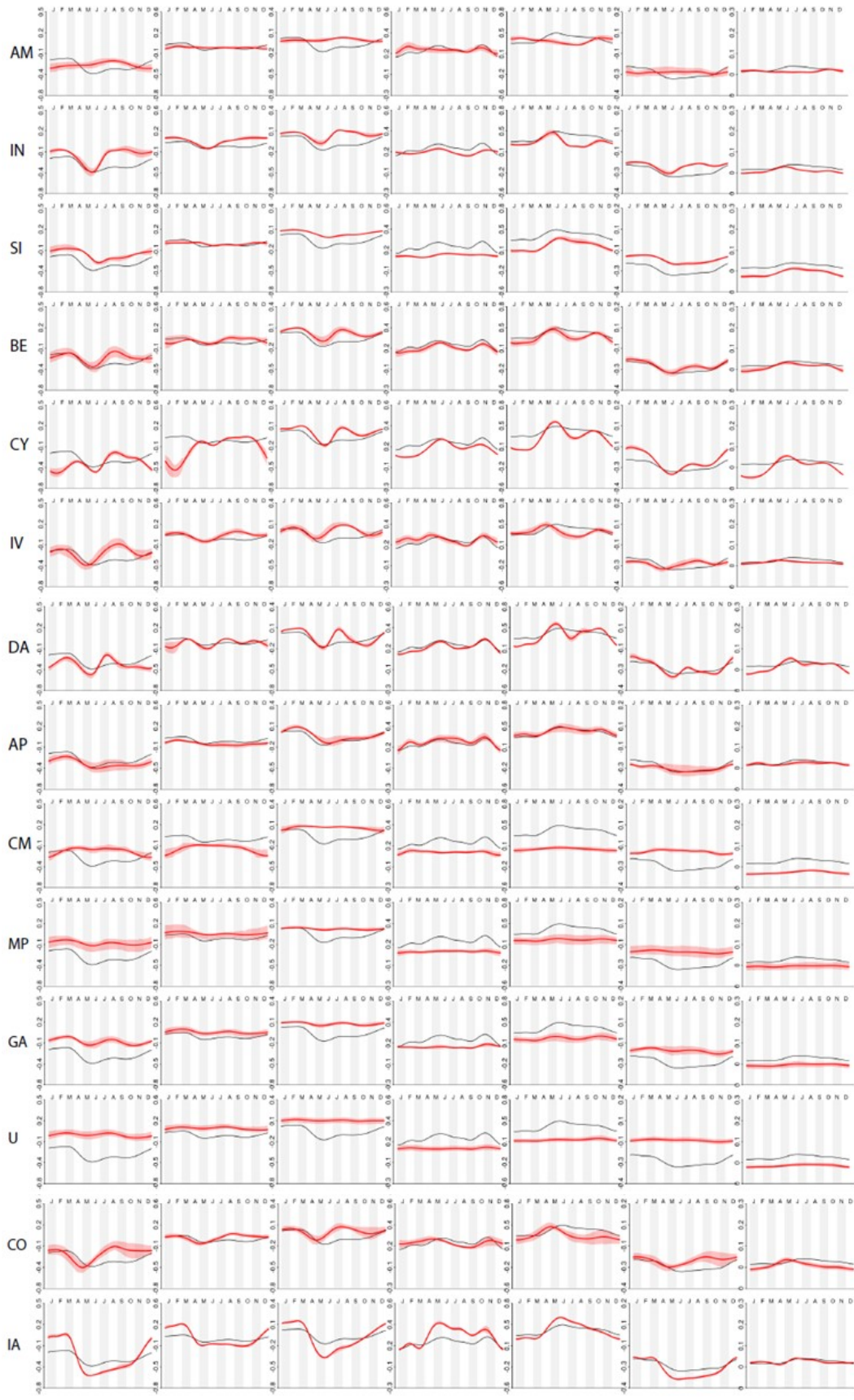
Qui di seguito vengono riportate le variazioni stagionali medie delle 34 categorie dei 7 indici selezionati.



Capitolo 7 Appendice



7.2 Profili stagionali medi



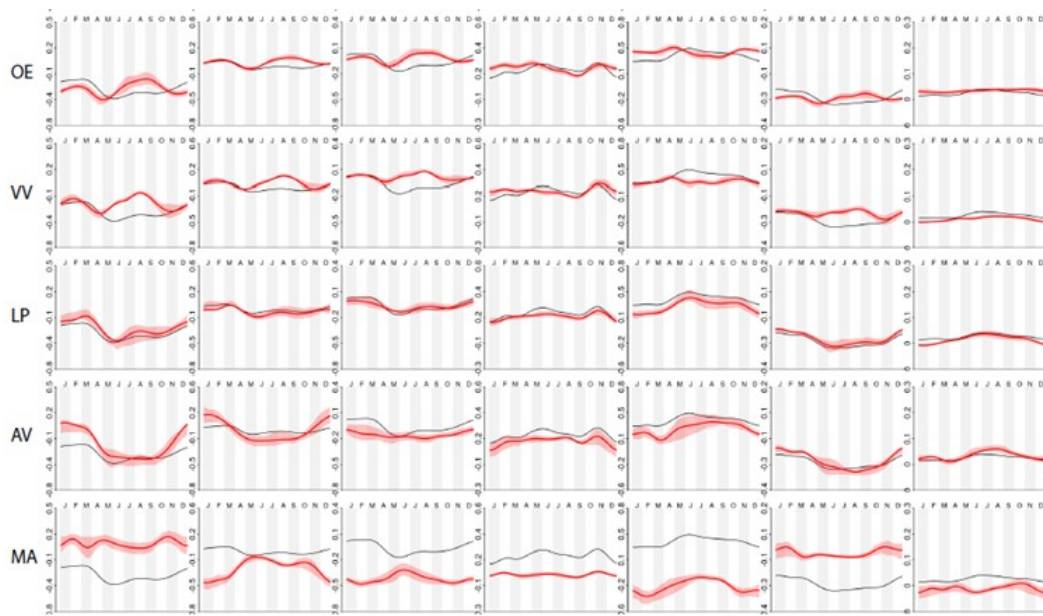


Figura 7.1: Profili stagionali multispettrali medi delle categorie relative alla tabella 3.1 del Reference Data.

Bibliografia

- [Agency, 2024] Agency, E. S. (2024). *Sentinel-2*.
- [Ballerini and Biondi, 2002] Ballerini, V. and Biondi, E. (2002). Dinamica di popolazioni arbustive e preforestali nell’appenino umbro-marchigiano (italia centrale). *Fitosociologia*, 39:175–183.
- [Belgiu and Drăguț, 2016] Belgiu, M. and Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114:24–31.
- [Biondi, 2011] Biondi, E. (2011). Phytosociology today: Methodological and conceptual evolution. *Plant Biosystems*.
- [Breiman, 2001] Breiman, L. (2001). Random forests. Technical report, Berkeley: Statistics Department University of California.
- [Casavecchia et al., 2023a] Casavecchia, S., Quattrini, G., Hofmann, N., and Pesaresi, S. (2023a). Relazione finale della convenzione di ricerca verifica e aggiornamento dell’attuale carta regionale degli habitat 1:10000 per le aree della rete natura 2000 del parco del conero utile per i monitoraggi degli habitat. Technical report, Dipartimento di Scienze Agrarie, Alimentari ed Ambientali.
- [Casavecchia et al., 2023b] Casavecchia, S., Quattrini, G., Hofmann, N., and Pesaresi, S. (2023b). Relazione finale della convenzione di ricerca verifica e aggiornamento dell’attuale carta regionale degli habitat 1:10000 per le aree della rete natura 2000 del parco interregionale del sasso simone e simoncello utile per i monitoraggi degli habitat. Technical report, Dipartimento di Scienze Agrarie, Alimentari ed Ambientali.
- [Casavecchia et al., 2023c] Casavecchia, S., Quattrini, G., Hofmann, N., and Pesaresi, S. (2023c). Relazione finale della convenzione di ricerca “verifica e aggiornamento dell’attuale carta regionale degli habitat 1:10000 per le aree della rete natura 2000 e della riserva statale del furlo utile per i monitoraggi degli habitat”. Technical report, Dipartimento di Scienze Agrarie, Alimentari ed Ambientali.
- [Cenci et al., 2018] Cenci, L., Disperati, L., Persichillo, M. G., Oliveira, E. R., Alves, F. L., and Phillips, M. (2018). Integrating remote sensing and gis techniques for monitoring and modeling shoreline evolution to support coastal risk management. *GIScience Remote Sensing*.

Bibliografia

- [Gorelick et al., 2017] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*.
- [Happ and Greven, 2018] Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, 113(522):649–659.
- [Hijmans, 2020] Hijmans, R. J. (2020). *raster: Geographic Data Analysis and Modeling*. R package version 3.4-5.
- [Hyndman and Khandakar, 2008] Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22.
- [Kattenborn et al., 2019] Kattenborn, T., Eichel, J., and Fassnacht, F. E. (2019). Convolutional neural networks enable efficient, accurate and finegrained segmentation of plant species and communities from high-resolution uav imagery. *Nature*.
- [Liaw and Wiener, 2002] Liaw, A. and Wiener, M. (2002). Classification and regression by randomforest. *R News*, 2(3):18–22.
- [Pesaresi et al., 2020] Pesaresi, S., Mancini, A., Quattrini, G., and Casavecchia, S. (2020). Mapping mediterranean forest plant associations and habitats with functional principal component analysis using landsat 8 ndvi time series. *Remote Sensing*, 12(7).
- [Pesaresi et al., 2022] Pesaresi, S., Mancini, A., Quattrini, G., and Casavecchia, S. (2022). Functional analysis for habitat mapping in a special area of conservation using sentinel-2 time-series data. *Remote Sensing*, 14(5).
- [Pesaresi et al., 2024] Pesaresi, S., Mancini, A., Quattrini, G., and Casavecchia, S. (2024). Evaluation and selection of multi-spectral indices to classify vegetation using multivariate functional principal component analysis. *Remote Sensing*, 16(7).
- [Quattrini et al., 2017] Quattrini, G., Pesaresi, S., and Casavecchia, S. (2017). *Classificazione e mappatura delle serie temporali NDVI di alcune comunità forestali del Monte Conero*. PhD thesis, Università Politecnica delle Marche.
- [Quattrini et al., 2020] Quattrini, G., Pesaresi, S., Casavecchia, S., and Mancini, A. (2020). *Mappatura della vegetazione e degli habitat tramite l'analisi della fenologia telerilevata: il caso studio della ZSC "IT5320003-Gola di Frasassi"*. PhD thesis, Università Politecnica delle Marche.
- [R Core Team, 2024] R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

- [Ramsay, 1982] Ramsay, J. (1982). When the data are functions. *Psychometrika*, 47:379–396.
- [Ramsay and Dalzell, 1991] Ramsay, J. and Dalzell, C. (1991). Some tools for functional data analysis. *Journal of the Royal Statistical Society. Serie B (Methodological)*, pages 539–572.
- [Rapinel et al., 2020] Rapinel, S., Rozo, C., Delbosc, P., Arvor, D., Thomas, A., Bouzillé, J.-B., Bioret, F., and Hubert-Moy, L. (2020). Mapping the functional dimension of vegetation series in the mediterranean region using multitemporal modis data. *GIScience and Remote Sensing*, 57:60–73.
- [Rivas-Martínez, 2011] Rivas-Martínez, S.; Sáenz, S. P. A. (2011). Worldwide bioclimatic classification system. *Global Geobotany*.
- [SAGA Development Team, 2024] SAGA Development Team (2024). *SAGA GIS: System for Automated Geoscientific Analyses*.
- [Shanmuga Priya and Vani, 2024] Shanmuga Priya, R. and Vani, K. (2024). Vegetation change detection and recovery assessment based on post-fire satellite imagery using deep learning. *Nature*.
- [Ullah et al., 2024] Ullah, T., Muhammad, Z., Ali Shah, I., Bourhia, M., Nafidi, H.-A., Salamatullah, A. M., and Younous, Y. A. (2024). Multivariate analysis of the summer herbaceous vegetation and environmental factors of the sub-tropical region. *Nature*.
- [Vignoli et al., 2021] Vignoli, S., Pesaresi, S., Casavecchia, S., and Mancini, A. (2021). *L'analisi funzionale delle serie temporali Landsat 8 per la classificazione della vegetazione*. PhD thesis, Università Politecnica delle Marche.
- [Wang et al., 2016] Wang, L.-J., Chiou, M.-J., and Müller, G.-H. (2016). Functional data analysis. *Annual Reviews*.
- [Wood, 2011] Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(1):3–36.
- [Yao et al., 2005] Yao, F., Müller, H.-G., and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, 100(470):577–590.