

Monitoraggio dei cambiamenti degli ecosistemi forestali in Italia



Informazioni legali

L'Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), insieme alle 21 Agenzie Regionali (ARPA) e Provinciali (APPA) per la protezione dell'ambiente, a partire dal 14 gennaio 2017 fa parte del Sistema Nazionale a rete per la Protezione dell'Ambiente (SNPA), istituito con la Legge 28 giugno 2016, n.132.

Le persone che agiscono per conto dell'Istituto non sono responsabili per l'uso che può essere fatto delle informazioni contenute in questo manuale.

ISPRA - Istituto Superiore per la Protezione e la Ricerca Ambientale
Via Vitaliano Brancati, 48 – 00144 Roma
www.isprambiente.gov.it

Riproduzione autorizzata citando la fonte

Il Responsabile del Centro Operativo di Sorveglianza Ambientale

Il Responsabile del Centro per le Emergenze, le Crisi ambientali e il Danno

Autori

- Emiliano Agrillo - introduzione, analisi risultati e procedure di verifica dei prodotti;
- Federico Filipponi - metodologia, analisi dati ed elaborazione grafica dei prodotti (mappatura);
- Roberto Inghilesi - supervisione ed editing;
- Alessandro Mercatini - verifica e correzione del documento;
- Alice Pezzarossa - metodologia di verifica del modello e dei risultati, coordinamento ed elaborazione del documento

Autori delle fotografie

figura [1.1](#): Effetti della tempesta Vaia del 2018 - autore Gianmaria Bonari

figura [1.2](#) - Effetto dei disseccamenti su una piantagione di pini interessati da fitopatie da insetti fitomizi - autore Emiliano Agrillo

figura [1.3](#) - Piantagione di Pino Nero colpita da incendio nella stagione estiva 2021- autore Emiliano Agrillo

figura [3.1](#) - Esempi delle quattro classi di ecosistemi forestali elaborate dal modello Ecosystems Classification Model - Forest4 ([ECM-F4](#)) - autore Emiliano Agrillo

Copertina: Federico Filipponi

Indice

1	Monitoraggio degli ecosistemi terrestri	1
1.1	Scopo del documento	1
1.2	Sorveglianza ambientale	1
1.3	Disturbi	3
1.4	Changes in Italian Terrestrial Ecosystems (CITE)	6
2	ECM-F4	8
2.1	Metodologia	8
2.1.1	Raccolta dati	8
2.1.2	Variabili di risposta (o Target)	12
2.1.3	Variabili Predittive e loro selezione	13
2.1.4	Calibrazione e validazione del modello statistico	16
3	Mappatura delle classi forestali italiane	22
3.1	Risultati	22
3.1.1	Copertura nazionale ECM-F4	22
3.1.2	Il Lazio	26
3.1.3	Riserva Naturale Statale "Tenuta Presidenziale di Castelporziano"	26
4	Applicazioni	30
4.1	Changes in Italian Terrestrial Ecosystems -Forest Fires (CITE-FF)	30
4.2	Piano di lavoro	32
A	Appendice	33
A.1	Lista degli acronimi ricorrenti	33
A.2	Acronimi variabili predittive modello ECM	34
A.3	Riferimenti WWW	35
	Bibliografia	36

Monitoraggio degli ecosistemi terrestri

1.1 Scopo del documento

Questo documento costituisce un riferimento per interpretare una serie di report periodici e di prodotti di mappatura realizzati dal Centro Operativo di Sorveglianza Ambientale ISPRA dedicati alla sorveglianza dei cambiamenti degli ecosistemi terrestri, in particolare le foreste. I metodi impiegati, basati su analisi dati satellitare e modelli statistici inferenziali (*machine learning*), sono di tipo quantitativo e forniscono un quadro conoscitivo di cui è importante capire le metodologie ed i limiti. Il primo passo del processo è la mappatura degli ecosistemi, il secondo l'analisi dei cambiamenti nel tempo. Qui viene descritto in modo esaustivo il metodo di classificazione dei tipi forestali alla base della mappatura e viene presentato il risultato in termini di distribuzione spaziale degli ecosistemi. La mappatura viene impiegata ad esempio nei Report annuale sugli incendi boschivi in Italia per definire in modo accurato gli effetti degli incendi sulle specifiche classi di ecosistemi.

1.2 Cambiamenti e metodi di sorveglianza ambientale

Le dinamiche ambientali sono alla base dei processi ecologici. Ad esempio, i dinamismi endogeni della terra generano continui cambiamenti della superficie terrestre, originando, come le fluttuazioni climatiche a scala millenaria, continui mutamenti nelle coperture vegetali [Shugart, 1998]. Quest'ultime possono modificarsi repentinamente se soggette a disequilibri ambientali di origine antropica come ad esempio nel caso degli effetti dovuti ai cambiamenti globali [Rockström et al., 2009]. Per tali motivi è necessario sviluppare ambiti di ricerca multidisciplinare finalizzati allo sviluppo di servizi di sorveglianza e monitoraggio ambientale per la creazione di prodotti utili alla gestione del territorio e alla conservazione del patrimonio naturale [Lovett et al., 2007].

Ad oggi, vedi Appendice A.3, i sistemi di monitoraggio e sorveglianza ambientale messi a servizio dalle agenzie nazionali e regionali, tra le quali Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Sistema Nazionale per la Protezione dell'Ambiente (SNPA) ed Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA) sono rivolti esclusivamente alle componenti aria, acqua

e suolo. Anche sistemi internazionalmente riconosciuti, come U.S. Environmental Protection Agency (EPA), U.K. Environment Agency (EA) ed Agencie de la transition écologique (ADEME), definiscono tipologie di monitoraggio prevalentemente improntate al tracciamento chimico e biochimico di contaminanti e finalizzate prevalentemente a prevenire gli effetti degli inquinanti sulla salute umana, con una ridotta capacità di adattamento alle sfide ambientali globali [Mol, 2016]. Inoltre, non risultano al momento servizi sviluppati specificamente al monitoraggio delle coperture vegetazionali (in termini, ad esempio, di superficie occupata) e del relativo “stato di salute” (es. vigoria e biomassa) delle componenti naturali (ecosistemi). Tali servizi, ove sviluppati, sarebbero in grado di fornire una tempestiva valutazione dei cambiamenti in atto (sia reversibili che irreversibili) mettendo in relazione possibili cause ed effetti quantificabili di degrado ambientale.

A partire dal 1990 sono state avviate dall’EPA alcune iniziative volte alla realizzazione di strumenti necessari al monitoraggio e alla valutazione dello *status* e *trend* delle risorse naturali su scala nazionale. In particolare il progetto Environmental Monitoring and Assessment Program (EMAP) aveva il compito di sviluppare ricerche applicate nel campo del monitoraggio ambientale al fine di tradurre i dati da molteplici scale, spaziali e temporali, in valutazioni sulle condizioni ecologiche attuali e previsioni di rischi futuri per le risorse naturali. Dal 2000 in poi la concettualizzazione e la progettazione di sistemi di monitoraggio integrato, e/o reti associate, sulle tematiche ambientali hanno ricevuto una crescente attenzione sia da parte dei ricercatori che degli amministratori [Wiersma and Bruns, 2004], nella speranza di poter essere in grado di affrontare in futuro le sfide ambientali sia da una prospettiva locale che globale [Hallgren et al., 2016, Hampton et al., 2013, Runting et al., 2020].

Nell’ultimo decennio, a seguito della spinta dei progressi tecnologici e metodologici nel settore del *Remote Sensing*, i dati di Earth Observation (EO) acquisiti dai sensori satellitari hanno offerto nuove opportunità per le scienze ambientali e stanno rivoluzionando le metodologie applicate in modo multidisciplinare, utilizzando conoscenze che vanno dalla Fisica teorica alle Scienze computazionali [Guo et al., 2016]. Un esempio di nuovo strumento per la valutazione dello stato di salute e per il monitoraggio è certamente la capacità di individuare e mappare ecosistemi ed habitat con alta risoluzione spaziale (inferiore a 100 m) e temporale [Adamo et al., 2020, Alvarez-Martínez et al., 2018, Corbane et al., 2015]. Recentemente, numerosi progetti e studi scientifici basati su prodotti EO si sono concentrati sull’identificazione ed il monitoraggio degli ecosistemi e del loro stato di conservazione, come ad esempio:

- la mappatura delle praterie utilizzando dati ottici e multispettrali multitemporali ad alta risoluzione;
- i consorzi forestali utilizzando serie temporali Normalized Difference Vegetation Index (NDVI) ;
- habitat costieri utilizzando metodologie di riclassificazione delle immagini spettrali;
- analisi fenologiche utilizzando indici di vegetazione calcolati su serie temporali;

- la valutazione dei tipi forestali utilizzando dati di tipo radar *backscatter*

[De Simone et al., 2021, Marzialetti et al., 2019, Pesaresi et al., 2020, Rüetschi et al., 2018, Spadoni et al., 2020].

Attualmente, a scala nazionale, i prodotti riguardanti la stima di molte tipologie di copertura del suolo, e quindi anche la relativa distribuzione degli ecosistemi terrestri, sono cartografie tematiche generate da ISPRA nell'ambito dei progetti Corine Land Cover (CLC) [Congedo et al., 2016], Carta di Copertura e Uso del Suolo (CCUS) e acCDN (vedi Appendice A.3). Il prodotto CLC viene aggiornato ogni sei anni, la CCUS annualmente e la Carta della Natura (CDN) non prevede aggiornamenti con specifiche scadenze temporali. I suddetti prodotti sono inoltre caratterizzati da specifiche caratteristiche cartografiche, in particolare si distinguono per una diversa risoluzione spaziale dei prodotti di mappatura che varia dai 10 metri per la CCUS, i 250 metri per il CLC e 1 ettaro per la CDN. Alcuni di questi prodotti forniscono informazioni riguardanti l'analisi dei cambiamenti, in dettaglio il CLC fa un confronto con le versioni precedenti a una risoluzione temporale pluriennale (ad oggi è disponibile il confronto CLC 2000 - 2018, in linea con quanto richiesto in ambito europeo (<https://www.eea.europa.eu/data-and-maps/dashboards/land-cover-and-change-statistics>)), con una risoluzione spaziale di 5 ha. ISPRA fornisce anche, a cadenza annuale, un servizio di mappatura del consumo di suolo, restituendo un quadro aggiornato dei processi di trasformazione del nostro territorio. Anche questo prodotto è in linea con il progetto europeo di Copernicus Land Monitoring Service "Imperviousness", il quale, mediante l'utilizzo di immagini telerilevate (Sentinel 2 – MultiSpectral Instrument (MSI)), definisce i dettagli cartografici a partire da una risoluzione temporale pluriennale (triennale a scala europea e annuale a scala nazionale) e una risoluzione spaziale a 10, 20 e 100 metri. In dettaglio, questo tipo di prodotto è finalizzato alla valutazione del cambiamento di una copertura naturale del suolo (prato, alberature, terreno nudo) con una copertura artificiale (edilizia, strade asfaltate, etc.), sia di tipo permanente che temporanea [Strollo et al., 2020]. Questo tipo di analisi è prevalentemente focalizzata ai contesti territoriali urbani, peri-urbani e agricolo intensivi dove sono più frequenti i cambi di destinazione/uso del suolo (sia legale che illegale).

1.3 Disturbi agli ecosistemi e fragilità ecologica del territorio

Nell'ambito della pianificazione e della gestione territoriale, i disturbi alle componenti naturali sono definiti "detrattori ambientali", ossia elementi potenzialmente responsabili della "fragilità ecologica" del territorio (es. frammentazione del paesaggio) [Farina, 2008, Wiens et al., 1993]. Pertanto, l'individuazione dei detrattori ambientali è un elemento essenziale per la corretta pianificazione territoriale mirata al restauro di aree degradate. Per "agente di cambiamento" o "detrattore" ambientale si intende quindi l'insieme di forzanti esterne che potenzialmente sono in grado di introdurre un disturbo del livello sia qualitativo che quantitativo degli ecosistemi, con una conseguente variazione del

valore economico. Tra i principali agenti di cambiamento degli ecosistemi naturali sono certamente da considerare gli incendi boschivi. Gli ecosistemi forestali sono ritenuti tra



Figura 1.1: *Effetti della tempesta Vaia del 2018 (Val di Fiemme, Trento; autore foto Gianmaria Bonari).*

i più vulnerabili [Lindner et al., 2010] e l'Unione Europea sta impegnando risorse per proteggere le nostre foreste e renderle più resilienti contro i rischi dettati dagli incendi, potenzialmente aggravati da altri agenti esterni quali tempeste con vento forte (ad esempio la tempesta Vaia del 2018, vedi Figura 1.1), siccità [Olmo et al., 2021] e attacchi da fitopatogeni [Schelhaas et al., 2003]. Pertanto, le informazioni relative alla distribuzione delle foreste, specialmente per quelle aree ritenute ad elevato rischio di incendio sono dunque essenziali per il monitoraggio e la gestione del territorio.

Il legame tra cambiamenti climatici ed alterazioni degli ecosistemi è complesso, cambiando le condizioni meteorologiche, cambiano ad esempio anche le intensità degli attacchi degli insetti e relative fitopatie, che rendono le piante più vulnerabili e le coperture arboree più suscettibili ad incendiarsi (un esempio è mostrato in Figura 1.2). A tal proposito, alcune piantagioni arboree come il Pino nero, presenti nel territorio nell'Appennino centrale e meridionale, risultano da tempo intaccate da insetti quali la processionaria del pino (le più diffuse in Italia sono, la *Thaumetopoea pityocampa* e la *Thaumetopoea processionea*), da cui risultano disseccamenti e defogliazioni anche di forte entità che hanno portato e stanno portando alla morte degli individui arborei infetti. Pertanto, siccome il ciclo biologico e quindi la relativa diffusione della specie parassita viene alterata dalle persistenti condizioni climatiche 'miti', in breve si potrebbe determinare un aumento di materiale comburente (piante secche) e di conseguenza anche un aumento del rischio incendi. In Figura 1.3 sono mostrate le conseguenze dell'evento di incendio avvenuto nell'estate del 2021 in una zona dell'Appennino centrale (Valle del Cicolano, Rieti), dove sono evidenti da tempo gli effetti causati dalla presenza di piantagioni monospecifiche



Figura 1.2: *Effetto dei disseccamenti su una piantagione di pini interessati da fitopatie da insetti fitomizi (Lazio, Roma).*

di Pino Nero di origine antropica affette dalla presenza di fitopatie causate da insetti [Cantiani et al., 2018].



Figura 1.3: *Piantagione di Pino Nero colpita da incendio nella stagione estiva 2021 (Valle del Cicolano, Rieti).*

1.4 Changes in Italian Terrestrial Ecosystems (CITE)

Al fine di poter valutare gli effetti causati da fattori sia biotici che abiotici sulle componenti naturali e semi-naturali degli ecosistemi terrestri sulla scala nazionale, regionale e locale, si è arrivati allo sviluppo del sistema di monitoraggio operativo ‘Changes in Italian Terrestrial Ecosystems (CITE)’. Nello specifico, il sistema CITE si propone come uno strumento per il monitoraggio e la sorveglianza mirato a specifici target ambientali, in grado di integrare, usando metodi tipo machine learning, informazioni derivanti da:

- la distribuzione di ecosistemi terrestri ad elevata risoluzione spaziale;
- gli agenti di cambiamento (sviluppati ad hoc o acquisiti da fonti certificate);
- lo studio delle modifiche da essi causati tramite la variazione nel tempo di variabili bio-geofisiche stimate da dati (satellitari) di osservazione della Terra.

Il servizio erogato dal sistema si basa su due prodotti di input (vedi livello Input della Figura 1.4):

- un prodotto spazializzato di classificazione probabilistica (i.e. modello di tipo *Machine Learning*) al fine di ottenere la distribuzione di tutti i pattern di vegetazione riferiti alla classificazione European Nature Information System (EUNIS);
- l’identificazione geografica di diversi agenti di cambiamento potenzialmente responsabili di effetti di degrado a breve e lungo termine.

I due elementi vengono analizzati insieme per ottenere una possibile identificazione delle componenti ambientali comprese nelle superfici soggette a cambiamento (vedi livello Output della Figura 1.4). Il calcolo della superficie è necessario all’analisi dei cambiamenti a breve e a lungo termine (vedi livello Analisi della Figura 1.4). Il cambiamento a breve termine, basato su un dato di variazione annuale permette la determinazione dei disturbi potenziali agli ecosistemi terrestri, mentre il cambiamento a lungo termine, basato su un dato di variazione pluriennale, permette l’analisi del cambiamento permanente dei pattern di distribuzione spaziale e dei cambiamenti fisionomici strutturali avvenuti nei diversi ecosistemi, con successiva individuazione delle aree degradate.

Il contributo innovativo nello sviluppo dell’intero sistema CITE è soprattutto dovuto a:

- la messa a sistema di procedure automatizzate e semi-automatizzate per la realizzazione dei prodotti spazializzati di classificazione probabilistica attraverso modelli inferenziali non parametrici che sfruttano variabili predittive derivate da dati di EO e dati ambientali;
- la risoluzione spaziale dei prodotti generati a 20 metri;
- la capacità di aggiornamento dei diversi prodotti con cadenza annuale e pluriennale.

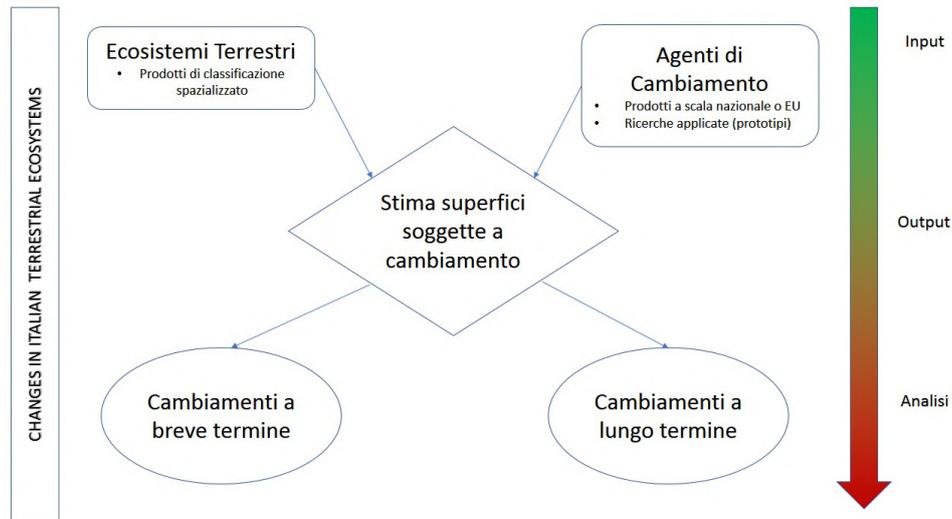


Figura 1.4: *Flowchart Changes in Italian Terrestrial Ecosystems* (**CITE**)

Caratteristiche rilevanti del sistema sono la versatilità dei prodotti per l'impiego su specifici ambiti locali e/o tematici e la rapidità di impiego operativo a supporto della sorveglianza attiva e/o dei rischi ambientali.

Nel capitolo 2 si spiega la metodologia **ECM-F4** impiegata per la classificazione degli ecosistemi forestali sviluppata all'interno del sistema **CITE** e nel capitolo 3 sono descritti i risultati ottenuti con il modello di classificazione **ECM-F4**. L'**ECM-F4** è stato impiegato per realizzare il rapporto sullo studio degli eventi di incendio boschivo avvenuti nel 2021 in Italia: **CITE-FF**.

Il modello Ecosystems Classification Model - Forest4 (**ECM-F4**) per la classificazione degli ecosistemi forestali

2.1 Metodologia

Il prodotto Ecosystems Classification Model - Forest4 (**ECM-F4**), impiegato nel sistema Changes in Italian Terrestrial Ecosystems (**CITE**) per la classificazione degli ecosistemi forestali, è basato su un protocollo standard di impiego di modelli statistici non-lineari di classificazione supervisionata. Il procedimento passa per diverse fasi successive:

1. la raccolta e selezione delle informazioni (denominate *target* o *variabili risposta* necessarie per definire le classi,
2. l'elaborazione delle variabili predittive corrispondenti ai target,
3. la determinazione del modello statistico,
4. la validazione del modello,
5. la proiezione del modello per ottenere la copertura sul territorio nazionale,
6. la verifica del prodotto (mappa) di classificazione.

Il processo è descritto nelle seguenti sotto-sezioni.

2.1.1 Raccolta dati

La fase di preparazione dei dati target è fondamentale nei sistemi che usano modelli statistici di tipo supervisionato. Sono state raccolte informazioni geografiche di tipo spaziale e puntuale relativamente alla copertura forestale a scala nazionale (variabile di risposta) sia dati necessari al calcolo dei principali driver (variabili predittive) in grado di descrivere i confini ecologici di tale copertura.

I criteri essenziali per la selezione dei dati inclusi nella ricerca hanno previsto che ciascun dato puntuale fosse: georeferito, accurato nella georeferenziazione, contenesse informazioni sulla fonte del dato, sulla data di rilievo e sull'occorrenza o di una comunità, con

relativa metodologia di determinazione, o di singola specie.

Per i driver, i criteri di selezione dei dati (di origine esterna ad ISPRA) sono la metadattazione, la presenza di documentazione European Space Agency (**ESA**) come Algorithm Theoretical Basis Document (**ATBD**) Product Specification Document (**PSD**), la fonte del dato, la data di rilievo del dato, la risoluzione spaziale, la risoluzione temporale e informazioni relative ad accuratezza ed errore della stima del dato.

Il dataset finale dei dati collezionati relativo alle coperture forestali (target o record) è composto da 11 160 presenze e quello relativo ai driver è di 163 variabili predittive (colonne o features nel gergo machine learning). Il metodo di classificazione selezionato rientra nella categoria dei modelli di distribuzione degli habitat, che mettono in relazione i dati sulla distribuzione degli ecosistemi con le condizioni abiotiche e biotiche, tra cui quelle derivate nell'ambito dei dati di osservazione della Terra, i quali consentono di osservare e quantificare la risultante dei processi ecologici e ambientali. Metodi di statistica inferenziale poi, come ad esempio i modelli Random Forests, vengono utilizzati per classificare i diversi ecosistemi e stimarne la distribuzione nello spazio come funzione delle diverse variabili predittive.

Il primo prodotto di classificazione sviluppato riguarda gli ecosistemi terrestri forestali, è elaborato a partire da un dataset composto da variabili di risposta e variabili predittive, entrambe certificate e meta-datate.

Il prodotto viene generato attraverso la procedura metodologica schematizzata nei seguenti flowchart (Figura **2.1** e Figura **2.2**). In Figura **2.1**, troviamo schematizzati i passaggi relativi ai diversi step di *data processing* dei dati di vegetazione, di quelli ancillari e di quelli ottenuti da **EO** fino all'ottenimento della variabile di risposta e delle diverse variabili predittive; in Figura **2.2**, troviamo schematizzati i diversi passaggi relativi sia alla calibrazione che alla validazione del modello di classificazione.

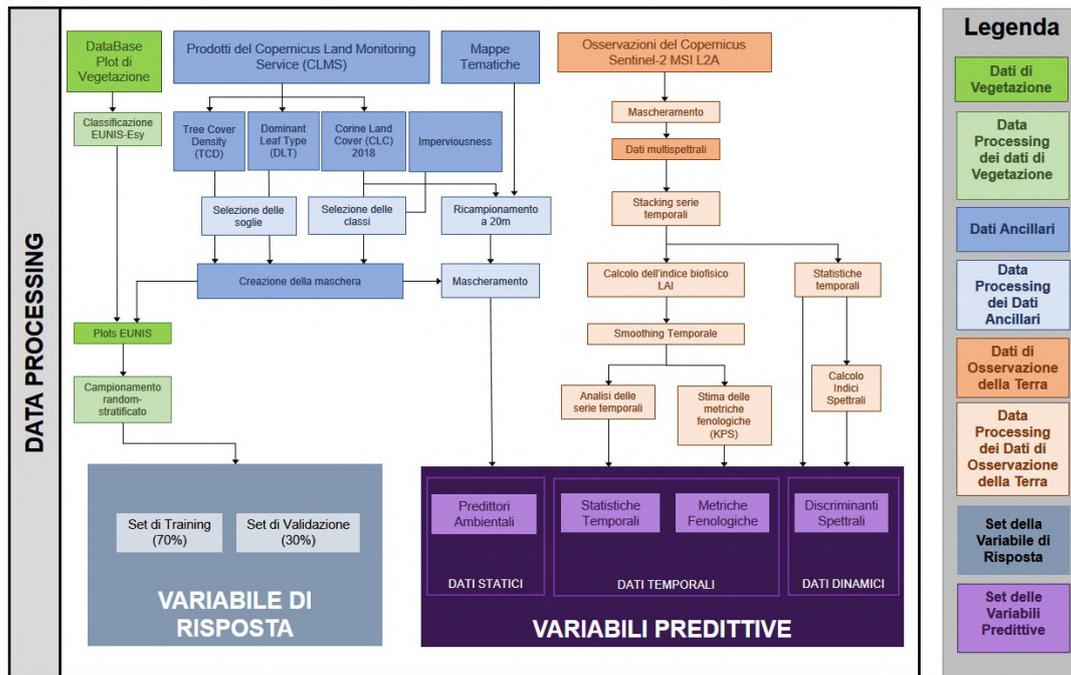


Figura 2.1: Flowchart della fase di data processing relativa al modello di distribuzione degli ecosistemi terrestri

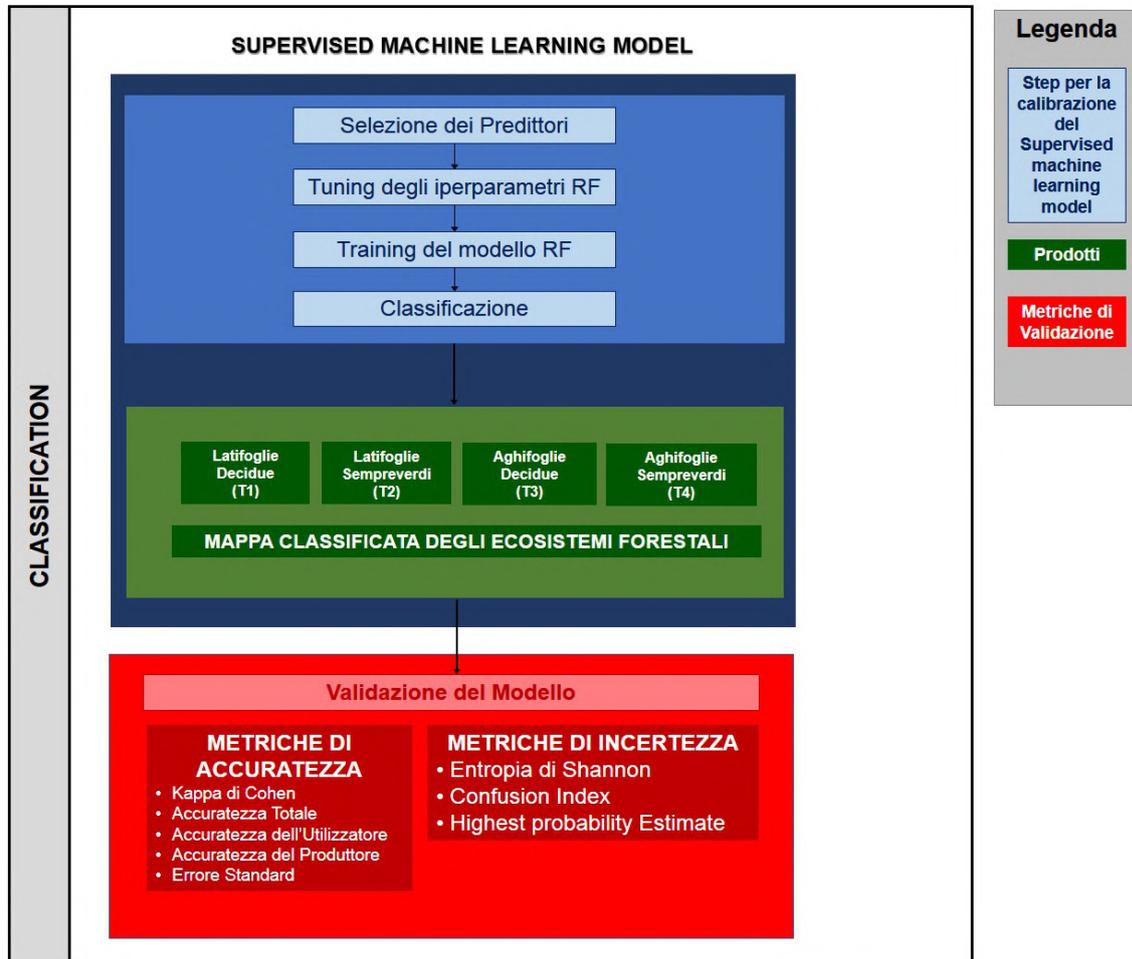


Figura 2.2: *Flowchart della fase di classification relativa al modello di distribuzione degli ecosistemi terrestri*

2.1.2 Variabili di risposta (o Target)

La fonte primaria di dati per la produzione della classificazione degli ecosistemi terrestri forestali è un dataset estratto dall'Archivio Europeo della Vegetazione European Vegetation Archive (**EVA**). L'archivio colleziona dati di plot di vegetazione che includono: una lista completa di piante vascolari presenti; la stima per ciascuna specie dell'abbondanza della copertura; la posizione georeferenziata; e, informazioni aggiuntive sulla struttura vegetativa e su caratteristiche ambientali del plot. I plot con una dimensione maggiore di 200 m² e che riportavano esclusivamente la composizione delle specie senza informazioni sull'abbondanza della copertura sono stati esclusi.

Il dataset finale relativo alla copertura forestale dei punti target è stato ottenuto utilizzando la seguente procedura:

- Ricampionamento dei dati: per effettuare la classificazione sulle sole aree boschive è stata generata una maschera per ricampionare i plot di vegetazione. La maschera è stata ottenuta combinando i prodotti del Copernicus Land Monitoring Service, nello specifico gli High Layer Resolution Land Monitoring Service Copernicus (**HLR**) Tree Cover Density (**TCD**) e l'*Imperviousness*, e il **CLC** nazionale. Tutti i record che si trovavano al di fuori della maschera sono stati esclusi dall'analisi;
- Clustering dei dati: dall'intero dataset vegetazionale sono stati selezionati i soli ecosistemi forestali, i quali sono stati clusterizzati secondo la classificazione della nomenclatura gerarchica **EUNIS**, seguendo le definizioni EUNIS-ES;
- Filtro dei dati: tutti i record che risultavano assegnati a più di un codice **EUNIS** (e.g. bias nell'assegnazione del plot a causa di un'elevata eterogeneità boschiva), e tutti i record duplicati sono stati esclusi. Infine, un'ultimo test a video ha permesso di identificare e correggere eventuali errori di misclassificazione e di posizionamento geografico.

Dall'insieme di dati così ottenuto, in seguito alle procedure effettuate di data cleaning sono stati selezionati 5906 punti corrispondenti alla presenza di variabili target. Ciascun punto viene identificato in modo univoco in base alla classe determinata mediante una legenda EUNIS al II livello gerarchico modificata per esaltare i caratteri funzionali delle più estese categorie forestali nazionali. Nel dettaglio, le quattro classi considerate per la classificazione (con il numero delle rispettive presenze nel database) sono:

- classe ECM-F4 latifoglie decidue (**T1**) - 3 390 presenze;
- classe ECM-F4 latifoglie sempreverdi (**T2**) - 1 623 presenze;
- classe ECM-F4 aghifoglie sempreverdi (**T3**) - 836 presenze);
- classe ECM-F4 aghifoglie decidue (**T3X**) - 57 presenze).

La proporzione di ciascuna classe all'interno del database è rappresentativo della copertura nazionale di ciascun consorzio forestale considerato.

2.1.3 Variabili Predittive e loro selezione

Per ottenere le mappe di distribuzione degli ecosistemi forestali è stato generato un dataset spazialmente esplicito di variabili predittive.

Il dataset include diverse categorie di variabili:

- dati statici: si tratta di variabili geografiche, topografiche, e di proprietà del suolo;
- dati spettrali: si tratta di variabili calcolate da dati di osservazione della terra rilevati dallo strumento Multi-Spectral Instrument (MSI) a bordo del satellite europeo Sentinel-2 in orbita polare. Tra queste sono state calcolate :
 - valori medi mensili di riflettanza relativi alle bande (λ): 2 (492.4 nm), 3 (559.8 nm), 4 (664.6 nm), 5 (704.14 nm), 6 (740.5 nm), 7 (782.8 nm), 8 (832.8 nm), 8a (864.7 nm), 11 (1613.7 nm), e 12 (2202.4 nm);
 - valori medi mensili degli indici spettrali EVI, NDYI, RI, CRI e delle stime dell'indice biofisico LAI;
- dati da serie temporali spettrali: si tratta di variabili che rappresentano l'andamento delle serie temporali di metriche fenologiche ed indici biofisici, sempre ottenute a partire da dati MSI- Sentinel-2, ad esempio LAI annuale, statistiche trimestrali del LAI, deviazione standard del LAI, variazione annuale LAI .

L'indice di area fogliare Leaf Area Index (**LAI**) è una misura di tipo biofisico che rappresenta la superficie fogliare, definita come metà dell'area di intercettazione della radiazione solare, per unità di superficie del suolo [Chen and Black, 1992]. Diversamente dagli indici spettrali comunemente utilizzati nelle procedure di analisi di dati di osservazione della terra, il **LAI** è poco sensibile alla saturazione del segnale radiometrico in aree con copertura della vegetazione densa, e rappresenta una misura verificabile in campo. La serie temporale di stime di **LAI**, realizzata a partire dai dati satellitari di tipo ottico multi-spettrale utilizzando un algoritmo standard, è stata utilizzata per il calcolo delle statistiche temporali: valore medio annuale (**LAI**), deviazione standard LAI rispetto a **LAI** (**LAI_{SD}**), LAI medio nel periodo dicembre, gennaio, febbraio (**LAI_{DJF}**), variazione massima annuale **LAI** (**Δ LAI**) [Filipponi et al., 2022]. Tali statistiche, in particolare, consentono di discriminare gli ecosistemi dominati da vegetazione di tipo sempreverde e caducifoglia. Tutte le variabili sono state collezionate o calcolate ex novo a partire da dataset open access e calcolate o ricampionate ad una risoluzione spaziale di 20 m. I sistemi di riferimento delle coordinate impiegati sono WGS84/UTM zone 32N e WGS84/UTM zone 33N. Infine tutte le procedure utilizzate per costruire le variabili sono state svolte utilizzando GRASS GIS per i dati statici, SNAP per il processamento dei dati Sentinel-2 MSI, e il software **R** per il processamento delle serie temporali (per il link agli strumenti di analisi vedi Appendice **A.3**). Per informazioni più approfondite sulle fonti di dati e sul calcolo delle variabili predittive si può fare riferimento ad [Agrillo et al., 2021].

Il dataset completo delle variabili predittive, risulta così composto:

- 13 variabili statiche-ambientali, di cui: 3 variabili geografiche, 4 geomorfologiche, 3 climatiche e 3 di proprietà del suolo;

124 variabili dinamiche spettrali, di cui 80 derivanti dalle singole bande spettrali e 44 indici spettrali;

26 variabili da serie temporali spettrali, di cui 10 statistiche di serie temporali e 16 metriche di fenologia.

Al fine di ridurre l'elevato numero di variabili, migliorare l'accuratezza modellistica e ridurre la ridondanza causata da una potenziale collinearità tra le variabili, sono stati calcolati il coefficiente di correlazione di Pearson e il Variance Inflation Factor (VIF) tra tutte le variabili. Tutti i predittori con coefficiente di Pearson maggiore di 0.7 e un JacVIF maggiore di 10 sono stati rimossi [Zuur et al., 2010]. L'ultimo passaggio è stato l'utilizzo del pacchetto **R** *Boruta* [Kursa and Rudniki, 2010], che determina l'importanza delle variabili confrontandone la rilevanza reale rispetto a quella di alcune variabili ombra, o spurie (*shadow*). Alcune variabili, che avrebbero dovuto essere eliminate secondo la procedura descritta, sono state comunque reinserite nel dataset seguendo il giudizio esperto, in quanto ritenute ecologicamente significative ai fini della classificazione degli ecosistemi considerati. Al termine della procedura di selezione sono state selezionate per la calibrazione del modello 22 variabili predittive:

8 ambientali, di cui 2 geografiche, 1 geomorfologica, 3 climatiche e 2 relative alle proprietà del suolo;

10 spettrali, di cui 6 bande spettrali e 4 indici spettrali;

4 variabili temporali.

In Tabella 2.1 sono indicate le variabili predittive selezionate, ordinate per indice di importanza decrescente (vedi Appendice A.2 per una spiegazione degli acronimi usati come nomi delle variabili). Per valutare l'indice di importanza è stato considerato l'Indice di Gini (GI) (per maggiori informazioni vedi, tra i possibili riferimenti, [Breiman et al., 1984, Krzywinski and Altman, 2017, Raileanu and Stoffel, 2004]). Le variabili il cui valore di importanza eccede il 75° percentile dell'GI, ossia i predittori che maggiormente contribuiscono alla corretta classificazione degli ecosistemi forestali, sono cinque (la distribuzione dei primi quattro è rappresentata in Figura 2.3), di questi, due sono nella categoria delle variabili spettrali: Dominant Leaf Type (DLT) e LAI_{SD}; e tre nell'ambito delle variabili ambientali: altitudine, distanza dalla linea di costa e le regioni biogeografiche ottenute su base bioclimatica. Il potere predittivo del DLT è risultato essere in assoluto la variabile maggiormente discriminate nella separazione dei due domini forestali a scala nazionale. Questo ha permesso di dettagliare la classificazione da due a quattro classi considerando il principale carattere funzionale delle specie forestali rappresentato dalla caduta stagionale dell'apparato fogliare.

I risultati inoltre mostrano come la variabilità dei pattern climatici e geografici giochino un ruolo cruciale sul territorio nazionale nel descrivere la spazialità delle risposte dei tratti delle comunità forestali.

Variabile	Indice di Gini (GI)
Dominant Leaf Type (DLT)	2226.66
Altitudine	1692.65
deviazione standard LAI rispetto a \overline{LAI} (LAI_{SD})	617.03
Distanza planimetrica dalla costa (log)	533.83
Regione biogeografica	487.67
Redness Index (valore medio marzo) (RI_{Mar})	453.75
Riflettanza banda 11 MSI (valore medio luglio) (B11_{Jul})	440.26
LAI medio nel periodo dicembre, gennaio, febbraio (LAI_{DJF})	396.06
Latitudine	385.15
Enhanced Vegetation Index (valore medio marzo) (EVI_{Mar})	304.03
variazione massima annuale LAI (ΔLAI)	294.84
Riflettanza banda 11 MSI (valore medio agosto) (B11_{Aug})	287.87
Riflettanza banda 8a MSI – valore medio luglio (B8A_{Jul})	232.98
Riflettanza banda 12 MSI (valore medio marzo) (B12_{Mar})	229.74
Precipitazioni annuali cumulate	212.48
Absolute depth to bedrock (BDTICM)	201.41
Riflettanza banda 11 MSI (valore medio maggio) (B11_{May})	185.60
pH calcolato in soluzione acquosa (PHIHOX)	184.94
Riflettanza banda 11 MSI (valore medio marzo) (B11_{Mar})	162.04
Temperatura annuale media normalizzata	138.77
Normalized Difference Yellow Index (valore medio ottobre) (NDYI_{Oct})	126.71
Carotenoid Reflectance Index (valore medio ottobre) (CRI_{Oct})	123.90

Tabella 2.1: Variabili utilizzate per calibrare il modello RF e il corrispettivo Indice di Gini per la valutazione dell'importanza. Più alto l'Indice di Gini, maggiore il contributo della variabile nella discriminazione della classificazione. Vedi Appendice **A.2** per la legenda delle variabili spettrali)

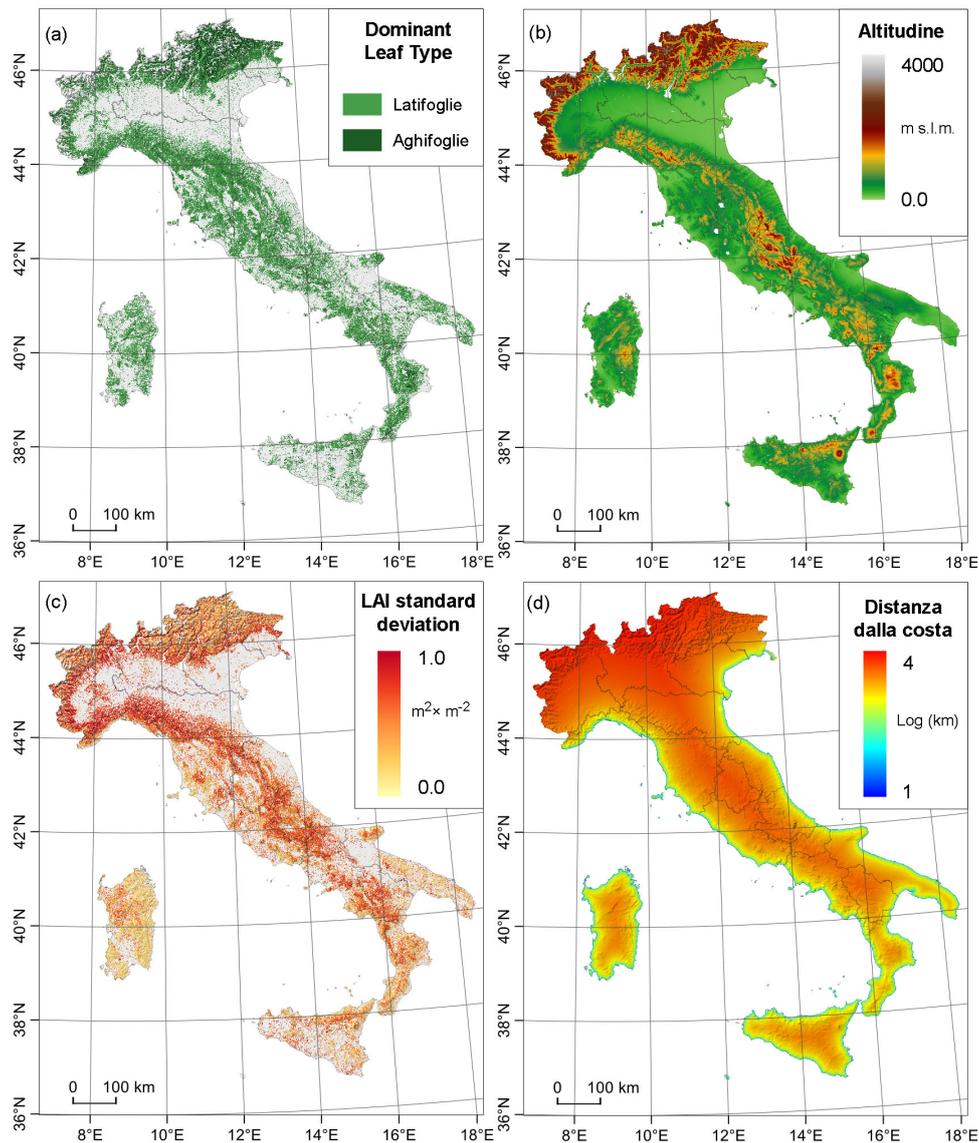


Figura 2.3: Le quattro variabili predittive con i valori più alti dell'Indice di Gini (**GI**).
 a. DLT; b. Altitudine; c. LAI_{sd} ; d. Distanza dalla linea di costa.

2.1.4 Calibrazione e validazione del modello statistico

Per classificare gli ecosistemi forestali è stato utilizzato un modello di apprendimento automatico supervisionato con l'impiego dell'algoritmo Random Forests (**RF**) **Breiman, 2001**], un algoritmo ad oggi applicato con successo ai modelli di classificazione di vegetazione a partire da dati di osservazione della terra (ad esempio **Kattenborn et al. 2019**]). Il modello è stato calibrato utilizzando solo il 70% del dataset di variabile di risposta

(training set, vedi sezione 2.1.2), lasciando il restante 30% del dataset per la successiva verifica del modello. È stato verificato che i dati esclusi dalla calibrazione fossero opportunamente stratificati. Per costruire la foresta di alberi l'algoritmo utilizza una serie di iperparametri che controllano sia la struttura dei singoli alberi (es. *minimum nodesize*) che la struttura e le dimensioni dell'intera foresta (es. *mtry*, *ntree*). Nonostante il **RF** possa in molti casi funzionare sufficientemente bene anche con gli iperparametri di default forniti dai singoli pacchetti, è stato dimostrato che il loro tuning può migliorare la performance dell'algoritmo (Probst et al., 2019). È stato quindi eseguito il tuning dei seguenti parametri **RF** nei seguenti range: *mtry* dalla radice quadrata del numero di predittori al numero di predittori meno 1; *min node size* da 1 a 3; e *ntree* da 401 a 1001. Il miglior set è stato infine selezionato scegliendo quello con il valore del coefficiente kappa di Cohen più alto, a sua volta calcolato tramite una 5-fold cross validation con 20 ripetizioni. Sono stati selezionati i seguenti iperparametri: *mtry* = 8; num. *trees* = 767; *min node size* = 2. A seguire, è stato calibrato il modello **RF**. Le analisi descritte sono state eseguite tramite i pacchetti *mlr* e *ranger* di R (Bischl et al., 2016; Wright and Ziegler, 2017).

Validazione del modello

Il prodotto di classificazione di ecosistemi forestali, ottenuto attraverso i metodi di statistica inferenziale sopra descritti, per essere ritenuto affidabile e replicabile, deve seguire la procedura di controllo di qualità che nell'ambito della modellistica spaziale prende il nome di validazione del modello. La validazione è stata effettuata sul 30% del dataset non impiegato nella fase di calibrazione (*validation set*). Trattandosi di un algoritmo di classificazione **RF**, l'output della procedura di validazione è una matrice di accuratezza o confusione o errore (vedi Tabella 2.2), che confronta i dati di copertura ottenuti dal modello (Map) con i dati di verità a terra del validation set (Reference).

Map	1	2	3	4	Total	User's accuracy
1	P11	P12	P13	P14	P1+	P11/P1+
2	P21	P22	P23	P24	P2+	P22/P2+
3	P31	P32	P33	P34	P3+	P33/P3+
4	P41	P42	P43	P44	P4+	P44/P4+
Totale	P+1	P+2	P+3	P+4	1	
Producer's accuracy	P11/P+1	P22/P+2	P33/P+3	P44/P+4		

Tabella 2.2: Esempio di una matrice di confusione per una classificazione con quattro classi (1-4), dove le righe (*i*) rappresentano la classificazione della mappa e le colonne (*j*) rappresentano la classificazione di riferimento; p_{ij} è la proporzione dell'area con classe di mappa (*i*) e classe di reference (*j*). Il totale della riga p_{i+} e della colonna p_{+j} sono la somma de p_{ij} valori in ciascuna riga e colonna.

Dal confronto si ottengono alcuni indici di accuratezza tra cui l'accuratezza globale (Eq. 2.1),

$$O = \sum_{i=1}^c p_{ji} \quad (2.1)$$

l'accuratezza del produttore (Eq. 2.2)

$$\hat{P}_j = \frac{\hat{p}_{jj}}{\hat{p}_{+j}} \quad (2.2)$$

e l'accuratezza dell'utilizzatore (Eq. 2.3)

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_{i+}} \quad (2.3)$$

(vedi Stehman and Foody [2019]). L'accuratezza globale Overall Accuracy (OA) è calcolata come il numero di pixel correttamente classificati (ottenuti tramite la somma della diagonale della matrice di errore) diviso il numero totale dei pixel.

L'accuratezza del produttore Producer's Accuracy (PA), calcolata per ciascuna classe, rappresenta l'accuratezza condizionata dell'area di riferimento, quindi quanti pixel appartenenti davvero a quella classe vengono classificati come tali; il complementare della PA non è altro che l'errore di omissione.

L'accuratezza dell'utilizzatore User's Accuracy (UA), calcolata anch'essa per ciascuna classe, definisce infine l'accuratezza condizionata dell'area mappata, quindi quanti pixel sono correttamente classificati rispetto al totale di ciascuna classe; il complementare della UA è quindi l'errore di commissione.

Per rendere la procedura di validazione ancora più efficace, sono stati calcolati gli errori standard di ciascun indice di accuratezza calcolato. Questo ha previsto una prima modifica degli indici, che sono stati ricalcolati a partire da una matrice di errore stimata, quindi pesata sulle superfici ottenute dall'output di mappatura per ciascuna classe. In seguito sono state applicate le formule secondo Stehman and Foody [2019].

Gli Standard Error (SE) sono calcolati come la radice quadrata della varianza rispettivamente della OA:

$$\hat{V}(\hat{O}) = \sum_{i=1}^c \frac{W_i^2 \hat{U}_i (1 - \hat{U}_i)}{n_{i+} - 1} \quad (2.4)$$

della PA:

$$\hat{V}(\hat{P}_j) = \frac{1}{\hat{N}_{+j}^2} \left[\frac{\hat{N}_{+j}^2 (1 - \hat{P}_j)^2 \hat{U}_j (1 - \hat{U}_j)}{n_{j+} - 1} + \hat{P}_j^2 \sum_{i \neq j}^c \hat{N}_{i+}^2 \frac{n_{ij}}{n_{i+}} \frac{\left(1 - \frac{n_{ij}}{n_{i+}}\right)}{n_{j+} - 1} \right] \quad (2.5)$$

della UA:

$$\hat{V}(\hat{U}_i) = \frac{\hat{U}_i (1 - \hat{U}_i)}{n_{i+} - 1} \quad (2.6)$$

Questo è stato possibile perché il validation set è stato selezionato tramite un campionamento randomico stratificato, ossia selezionando i punti in maniera casuale ma rispettando la proporzionalità delle classi, ossia avendo un numero di punti maggiori per le classi più abbondanti e viceversa.

Infine sono stati calcolati come ulteriori indici di accuratezza l'Out Of Bag Error (OOB), una procedura interna dell'algoritmo RF che misura l'errore di previsione, e la Cohen's Kappa, un coefficiente statistico che rappresenta il grado di accuratezza e affidabilità di una classificazione statistica.

L'ultimo passaggio della procedura di verifica dei prodotti di classificazione degli habitat terrestri ha permesso di identificare la componente spaziale dell'accuratezza della classificazione. Gli indici calcolati in precedenza infatti hanno il limite di riferirsi o all'accuratezza globale o a quella delle singole classi, ma non danno informazioni su come l'incertezza associata sia distribuita nello spazio. Per rimediare a questo limite sono stati calcolati alcuni indici per ciascun pixel del prodotto di mappatura ottenuto. Gli indici sono la Shannon Entropy, l'Highest Probability Estimate e il Confusion Index (Wegmann et al., 2016). Ciascuno confronta le probabilità di appartenenza alle diverse classi per ciascun pixel dando una stima dell'incertezza della classificazione (vedi Figura 2.4).

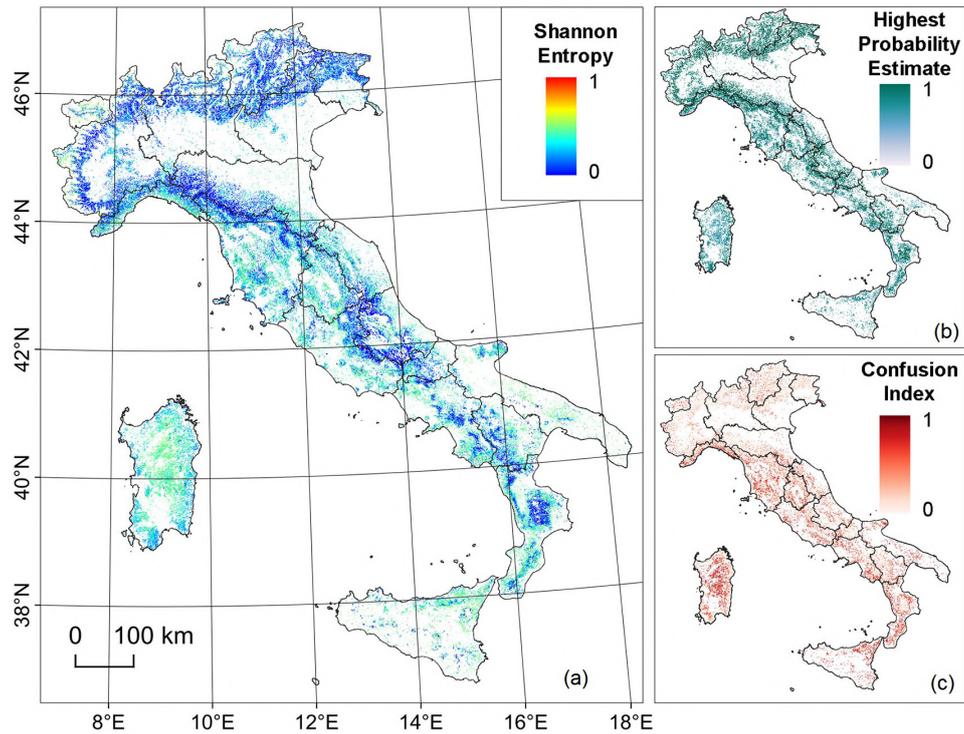


Figura 2.4: Risultato della spazializzazione dell'incertezza ottenuta tramite: (a) l'indice di Shannon Entropy; (b) l'Highest Probability Estimate; (c) il Confusion Index.

La classificazione degli ecosistemi terrestri forestali ha raggiunto un'accuratezza generale **OA** pari a 90.35% ($\kappa = 0.83$), e nell'intervallo 66.7-96.5% e 47.1-98.4% per quanto riguarda le accuratezze del produttore **PA** e dell'utilizzatore **UA** rispettivamente (vedi Tabella **2.3**).

Map Class	Reference Class				Somma	UA	SE(UA)
	T1	T2	T3	T3X			
T1	938	79	0	0	1017	92.2	0.008
T2	79	408	0	0	487	83.8	0.017
T3	0	0	247	9	256	98.4	0.012
T3X	0	0	4	8	12	47.1	0.142
Somma	1017	487	251	17	OA	90.35	0.67
PA	92.2	83.8	96.5	66.7			
SE(PA)	0.976	0.869	0.996	0.555		Kappa	0.83

Tabella 2.3: *Matrice di confusione dove vengono mostrati gli errori di classificazione eseguiti sul testing set della variabile di risposta e la performance del modello predittivo di classificazione. User's Accuracy (UA); Producer's Accuracy (PA); Overall Accuracy (OA); Errore Standard (ES).*

Il modello finale validato, è stato poi proiettato sull'intera area di studio al fine di ottenere la distribuzione degli ecosistemi forestali nazionali.

Controllo di qualità del prodotto

La validazione del prodotto prevede sia una valutazione comparativa dei risultati ottenuti nell'ambito delle stime di superficie con prodotti preesistenti esterni, sia un confronto a video da parte di un operatore del prodotto relativo alle diverse coperture degli ecosistemi forestali considerati, attraverso la sovrapposizione con dati di osservazione della terra ad alta risoluzione. La scelta dei prodotti per la valutazione comparativa è ricaduta sugli unici prodotti a scala nazionale disponibili in termini di stima delle coperture dei consorzi forestali. Nel dettaglio si tratta dell'Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio (INFC), attualmente aggiornato al 2005 (<https://www.sian.it/inventarioforestale/jsp/documentazione.jsp>), e i prodotti in formato vettoriale di riferimento presenti nella Copertura del suolo 2012-dati ISPRA-regionali-Copernicus (CCUS12) (vedi Appendice A.3). Il confronto a video è stato effettuato da operatori specializzati su porzioni di territorio selezionate sulla base degli indici di accuratezza/incertezza ottenuti dalle analisi precedentemente descritte.

Mappatura delle classi forestali italiane

3.1 Risultati

La classificazione degli ecosistemi forestali prodotta dal sistema CITE consiste in una mappa nazionale delle coperture forestali secondo la classificazione EUNIS al II livello gerarchico e ad una risoluzione spaziale fine (20 m). Come precedentemente accennato, ad oggi risultano diversi progetti e/o servizi di mappatura degli habitat forestali che hanno utilizzato i dati satellitari per la classificazione, ma con risoluzioni spaziali non elevate, e/o su ristrette porzioni di territorio (aree a scala subregionale e/o locale) [Pesaresi et al., 2020], [Pignatti et al., 2014], [Spadoni et al., 2020] e/o concentrandosi esclusivamente solo su singole coperture di specie arboree [Chirici et al., 2020]. Una caratteristica primaria del presente approccio riguarda in particolare l'aggiornabilità del prodotto di classificazione, che consente di rinnovare con cadenza annuale la mappatura degli ecosistemi forestali, fornendo così un contributo dinamico al loro monitoraggio e alla valutazione del loro stato ecologico. Inoltre, considerando la programmazione dei piani di acquisizione da parte della costellazione satellitare Sentinel (vd. ESA Sentinel Program Appendice A.3), l'approccio proposto può essere riprodotto ed esteso a tutte le aree e/o regioni di interesse, per specifiche valutazioni sullo stato di salute delle coperture vegetazionali. Nel complesso, il risultato ottenuto dimostra come il supporto delle informazioni derivate dai satelliti Sentinel-2, risulti molto efficace per il riconoscimento e la discriminabilità dei diversi tipi di ecosistemi forestali finalizzata alla stima delle superfici di copertura [Dostálová et al., 2018], [Grabska et al., 2019].

3.1.1 Copertura nazionale ECM-F4

La copertura degli ecosistemi forestali classificati nel 2020 in Ecosystems Classification Model - Forest4 (ECM-F4) complessivamente ammonta a 8 516 994 ha, superficie che costituisce il 95% della superficie complessiva boschiva prodotta dal Copernicus Land Monitoring Service. Il 5% rimasto non classificato è limitato ad aree in cui si sono verificati problemi nella stima di alcune variabili predittive, e quindi il sistema non ha potuto operare nella fase di proiezione sul territorio nazionale. Tali zone sono generalmente aree in cui sono presenti ombre topografiche o ombre dovute a copertura nuvolosa persistente.

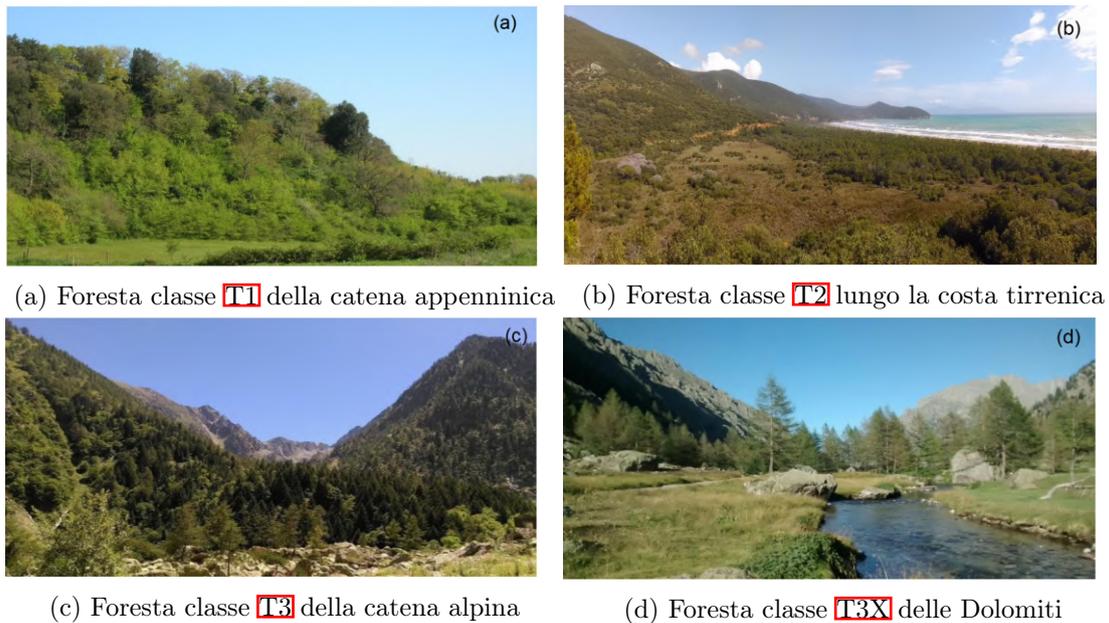


Figura 3.1: Esempi delle quattro classi di ecosistemi forestali selezionate per il prodotto Ecosystems Classification Model - Forest4 (ECM-F4): (a) foresta classe ECM-F4 latifoglie decidue (T1); (b) foresta classe ECM-F4 latifoglie sempreverdi (T2); (c) foresta classe ECM-F4 aghifoglie sempreverdi (T3); (d) foresta classe ECM-F4 aghifoglie decidue (T3X)

In dettaglio, gli ecosistemi forestali classificati sono suddivisibili in quattro classi EUNIS II livello:

- 5 724 481 ha risultano appartenere alla classe ECM-F4 latifoglie decidue (T1);
- 1 633 235 ha risultano appartenere alla classe ECM-F4 latifoglie sempreverdi (T2);
- 1 113 625 ha risultano appartenere alla classe ECM-F4 aghifoglie sempreverdi (T3);
- 26 117 ha risultano appartenere alla classe ECM-F4 aghifoglie decidue (T3X) e/o boschi di Larice

(vedi Figura 3.1 per una rappresentazione fotografica degli ecosistemi forestali corrispondenti alle quattro categorie).

In relazione al risultato ottenuto, la matrice di confusione (vd. Tabella 2.3) oltre a permettere il calcolo di indici di accuratezza, ha evidenziato come i due complessi forestali maggiormente rappresentativi a scala nazionale relativi alle latifoglie e alle aghifoglie non prevedano né errori di commissione (falsi positivi) né errori di omissione (falsi negativi). Pertanto, il risultato ottenuto dal modello ben definisce la distribuzione delle due classi. Questo anche a conferma del peso dell'importanza nel risultato ottenuto dal modello dall'uso della variabile predittrice Dominant Leaf Type (DLT) (vd. Tabella 2.1), favorendo così il bilanciamento tra variabili di risposta e predittori, in particolar modo per la spazializzazione dei boschi di conifere del meridione (in gran parte di origine artificiale, ossia piantagioni). Per quanto concerne le coperture di sempreverdi e caducifoglie, nell'ambito dei boschi di latifoglie, i risultati ottenuti presentano un minimo errore di omissione presumibilmente giustificato dalla effettiva presenza di ampie fasce altitudinali di bosco misto tra il dominio delle leccete e quello delle querce caducifoglie (es. Cerro, Roverella), in particolar modo nell'area appenninica e nella porzione meridionale di alcuni settori alpini (vd. sponde del Lago di Garda e Trentino). Infine si osserva che l'accuratezza del risultato ottenuto per la classe relativa ai boschi di larice classe ECM-F4 aghifoglie decidue (13X) sia circa la metà di quella delle altre classi, e, soprattutto l'errore standard associato sia intorno al 14%, circa 10 volte superiore a quanto ottenuto per le altre classi. Questo indica che il modello mostrerà in fase di proiezione una maggiore difficoltà nel riuscire a distinguere il carattere funzionale che contraddistingue l'unica specie di conifera con il carattere di caducifolia. Le cause che originano questa minore accuratezza sono probabilmente legate a due fattori: il numero di occorrenze troppo basso nel dataset delle variabili di risposta, con conseguente scarso potere predittivo da parte del modello, e la compresenza nel consorzio forestale dominante di altre specie di conifere (es. Pino cembro) con caratteristiche di sempreverde.

In Figura 3.2 è illustrata la mappatura degli habitat forestali a livello nazionale.

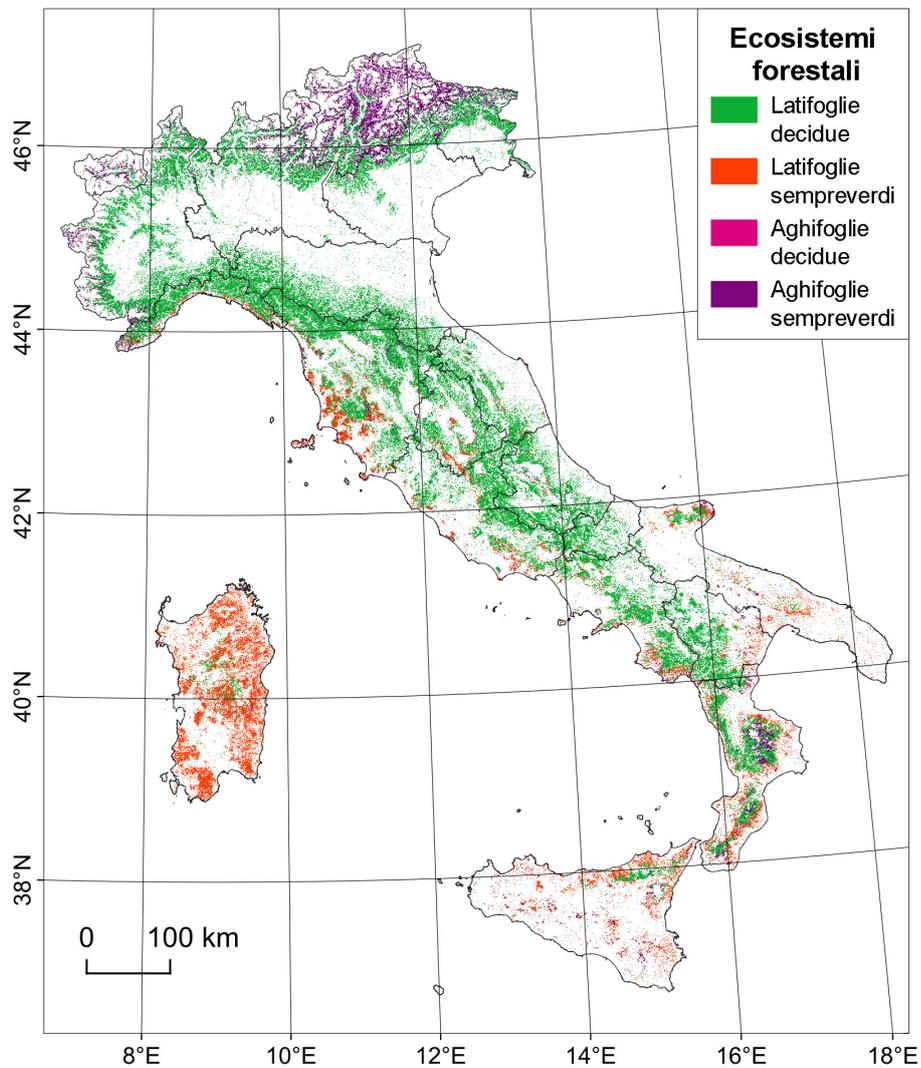


Figura 3.2: Classificazione degli ecosistemi forestali italiani ottenuta con il Ecosystems Classification Model - Forest4 (ECM-F4) relativo all'anno 2020.

I risultati (vd. Figura 3.2) mostrano come la variabilità dell'andamento bioclimatico e geografico del territorio italiano svolga un ruolo fondamentale nell'influenzare la distribuzione delle comunità forestali [Pesaresi et al., 2014], come la presenza delle conifere (Abetaie e Peccete) e conifere caducifoglie (Lariceti) rappresentano il paesaggio montano dell'arco alpino, mentre i boschi di latifoglie decidue (Faggete e Quercete) e boschi di sempreverdi (Leccete e Sugherete) interessano in prevalenza l'Appennino e le isole maggiori.

3.1.2 Il Lazio

Dalla Figura 3.3, dove viene rappresentato il dettaglio della classificazione nella regione Lazio, grazie alla notevole risoluzione spaziale del prodotto di classificazione ottenuto, emerge una chiara frammentazione delle coperture forestali in particolare nelle porzioni prossime alle pianure fluvio-alluvionali attualmente a vocazione agricola e intensamente urbanizzate (Pianura Pontina, Maremma Viterbese, la bassa Valle del Tevere e il Delta Tiberino). Il risultato della mappatura mostra di poter contribuire a superare, almeno in parte, la mancanza di aggiornabilità e potrebbe garantire il miglioramento della risoluzione spaziale dei prodotti a scala regionale ad oggi consultabili, come la Carta della Natura del Lazio aggiornata al 2008 a scala 1:50 000 (vd. <https://www.isprambiente.gov.it/contentfiles/00011500/11526-metadati-lazio.pdf>) e la carta delle formazioni naturali e semi-naturali e forestali della regione Lazio aggiornata al 2006 a scala 1:25 000 [Chirici et al., 2014]. I prodotti sopraindicati sono stati in parte realizzati con tecniche di telerilevamento ed uso di prodotti satellitari (es. immagini Landsat 7 Enhanced Thematic Mapper Plus (ETM+) e SPOT 5 High Resolution Geometric (HRG)), e per tale motivo sono stati scelti come documenti di confronto per un giudizio, comparativo semi-quantitativo da esperto, del risultato ad oggi ottenuto dal prodotto di classificazione ECM-F4.

3.1.3 Riserva Naturale Statale "Tenuta Presidenziale di Castelporziano"

Scendendo in un dettaglio ancora maggiore, si riporta in Figura 3.4 l'esempio di un'area protetta di pregio del settore peri-costiero Laziale. L'area prescelta è la Riserva Naturale Statale "Tenuta Presidenziale di Castelporziano", sito dove vengono tutelati alcuni lembi di territorio intatto delle pianure alluvionali e peri-costiere della penisola Italiana e per questo motivo ad oggi riconosciuti anche nei siti europei della rete Natura 2000. Anche in questo caso emerge l'elevata capacità predittiva della mappa prodotta da ECM-F4 nel rappresentare fedelmente il paesaggio attuale del territorio esaminato. Il risultato è stato confrontato, in base al giudizio esperto, in termini di attinenza geometrica e tematica dei poligoni mappati, con un prodotto cartografico, innovativo per l'epoca della pubblicazione, derivato da ricerche applicative per la mappatura degli habitat tramite tecniche e prodotti da remoto [DellaRocca et al., 2001] (i risultati del confronto qualitativo non sono qui esposti).

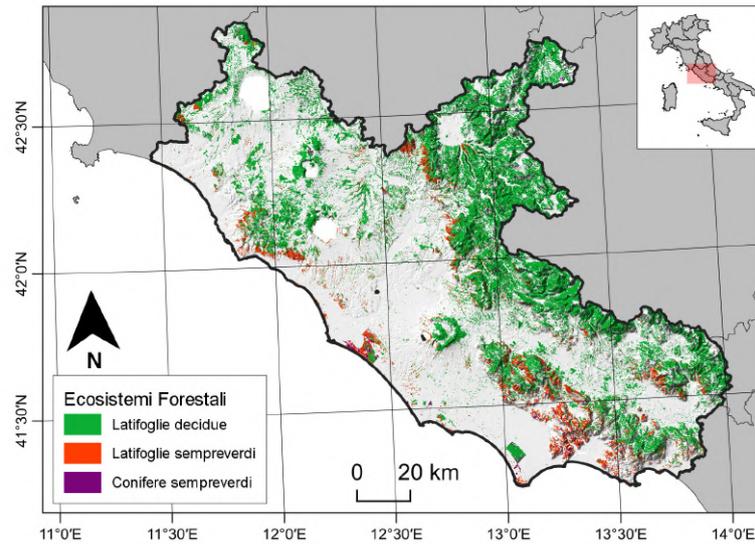


Figura 3.3: Classificazione degli ecosistemi forestali della regione Lazio ottenuta con Ecosystems Classification Model - Forest4 (ECM-F4) relativo all'anno 2020.

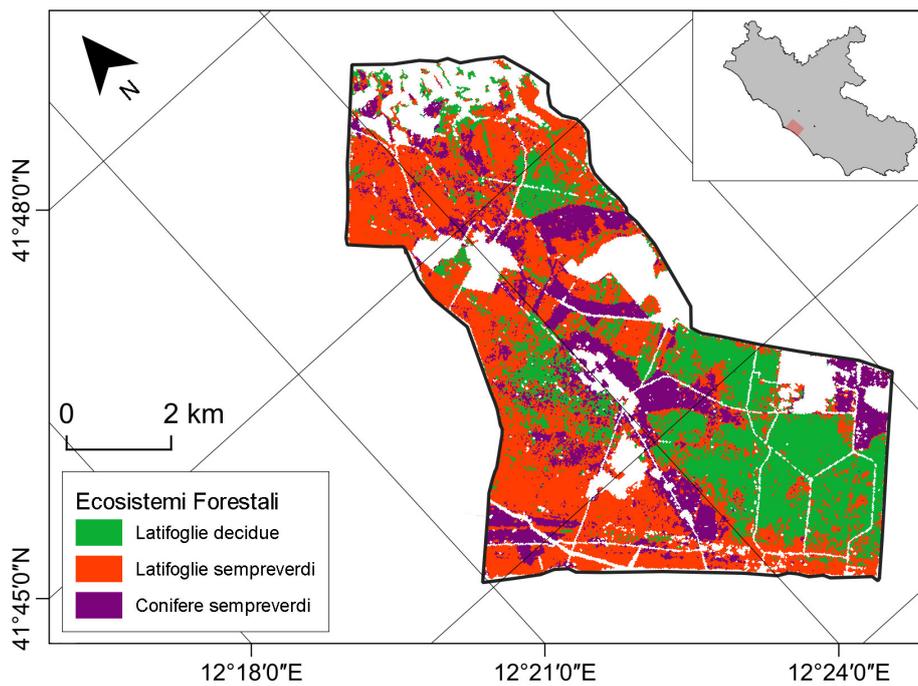


Figura 3.4: Classificazione degli ecosistemi forestali della Riserva Naturale Statale "Tenuta Presidenziale di Castelporziano" situata lungo la costa tirrenica della regione Lazio ottenuta con Ecosystems Classification Model - Forest4 (ECM-F4) relativo all'anno 2020.

CLASS	ECM-F4 [ha]	INFC [ha]	CCUS [ha]
T1	5 724 482	5 994 777	6 444 360
T2	1 663 236	873 632	1 137 659
T3	1 113 625	1 332 188	1 771 110
T3X	26 117	382 372	396 350
Totale	8 582 969	8 497 460	9 749 479

Tabella 3.1: Tabella di confronto tra gli ha ottenuti dalla classificazione di Ecosystems Classification Model - Forest4 (ECM-F4) 2020 e quelli stimati dall'Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio (INFC) 2005 e Copertura del suolo 2012-dati ISPRA-regionali-Copernicus (CCUS12) per le classi forestali considerate: classe ECM-F4 latifoglie decidue (T1), classe ECM-F4 latifoglie sempreverdi (T2), classe ECM-F4 aghifoglie sempreverdi (T3) e classe ECM-F4 aghifoglie decidue (T3X).

Come anticipato nella sezione Controllo di qualità del prodotto, la valutazione quantitativa dell'affidabilità del prodotto ottenuto da Ecosystems Classification Model - Forest4 (ECM-F4) a scala nazionale è stata effettuata comparando i prodotti disponibili sul tema coperture forestali, indipendentemente dalla metodologia usata per la definizione delle stime superficiali di riferimento per ogni singola componente individuata (i.e. consorzi forestali INFC).

Nella Tabella 3.1 sono mostrati i risultati ottenuti da un primo confronto con l'inventario forestale INFC 2005 e CLC 2018. Ribadendo che le metodologie di sviluppo dei due prodotti sono fondamentalmente diverse, gli 85 509 ha stimati in più dal prodotto Ecosystems Classification Model - Forest4 (ECM-F4) rispetto all'INFC potrebbero anche essere frutto di un reale accrescimento dei consorzi forestali negli ultimi 15 anni. Infatti come documentato a livello nazionale, a partire dalla seconda guerra mondiale le superfici boschive e arbustive sono in continua espansione, anche dovuta all'abbandono delle aree agricole in ambito sub-montano.

Nonostante la differenza di copertura non sia particolarmente significativa a livello complessivo, è bene notare come esista un sensibile sbilanciamento tra le singole classi. Un caso riguarda le latifoglie sempreverdi per le quali l'Ecosystems Classification Model - Forest4 (ECM-F4) 2020 stima il doppio degli ettari rispetto all'Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio (INFC) 2005. Questo potrebbe essere dovuto ai differenti disegni sperimentali dietro i due elaborati, in cui INFC si basa su un campionamento ad hoc per la stima delle superfici e delle biomasse, mentre l'altro, Ecosystems Classification Model - Forest4 (ECM-F4) è basato su una raccolta dati preferenziale, mirata all'individuazione di tutti gli ecosistemi forestali esistenti. Il caso opposto, in cui è presente una sottostima della classe classe ECM-F4 aghifoglie decidue (T3X) da parte di Ecosystems Classification Model - Forest4 (ECM-F4), ossia del larice, è probabilmente dovuto ad un numero troppo piccolo di record scelto per i target. Il risultato è consistente con quanto discusso nella sezione Calibrazione e validazione del modello statistico, in

particolare la discussione della matrice di confusione in [2.3](#)).

Infine, anche il risultato della valutazione a video del prodotto in aree campione ha confermato i risultati ottenuti dalla matrice di confusione riguardo il mismatch tra le varie classi, e conferma la presenza delle sequenze altitudinali dei cingoli forestali lungo le catene alpine e appenniniche.

Applicazioni **ECM-F4**, e piano di sviluppo del CITE

4.1 Valutazione delle aree boschive annualmente bruciate in Italia

Dall'inizio degli anni 2000, lo sviluppo tecnologico legato all'osservazione della Terra dallo spazio ha permesso lo sviluppo del sistema europeo d'informazione sugli incendi boschivi European Forest Fire Information System (**EFFIS**). Il programma, condotto in collaborazione con gli Stati membri dell'Unione Europea (**EU**), è nato per fornire una piattaforma adeguata ad uno scambio ampio ed intensivo di dati e informazioni relative al monitoraggio e alla mappatura degli incendi boschivi, e ai loro effetti sull'ambiente. Dal 2015 **EFFIS** è una delle tre componenti del programma **EU** Copernicus Emergency. **EFFIS** produce e distribuisce annualmente la cartografia relativa alle aree bruciate attraverso il processamento dei dati satellitari multispettrali. Il sistema **EFFIS** effettua stime annuali di superficie bruciate, utilizzando i dati **HLR** relativa alle coperture del suolo. Il dataset spazializzato relativo alle aree percorse da incendio può essere ottenuto dal portale del servizio **EFFIS** <https://effis.jrc.ec.europa.eu>. I poligoni **EFFIS** delle aree percorse da incendio nel corso dell'anno possono essere impiegati, insieme altri dataset spazializzati per calcolare statistiche relative ai diversi tematismi di tipo ambientale. In particolare, nel report annuale ISPRA sugli incendi boschivi **CITE-FF** i dati di aree bruciate ottenuti da **EFFIS** sono impiegati insieme alla classificazione ottenuta da **ECM-F4** per valutare quali fossero in origine le coperture delle aree percorse da incendio. I dataset utilizzati sono il prodotto **ECM-F4**, i confini delle aree naturali protette (parchi nazionali, parchi regionali, riserve statali, riserve regionali, siti Natura 2000, Aree importanti per l'avifauna e la biodiversità, aree umide ai sensi della Convenzione di Ramsar), i poligoni delle aree a pericolosità da frana e idraulica, e i confini delle unità amministrative (comuni, province, regioni). Da questa elaborazione possono essere stimate quantitativamente le aree bruciate per ogni tipologia forestale considerata (classe ECM-F4 latifoglie decidue (**T1**), classe ECM-F4 latifoglie sempreverdi (**T2**), classe ECM-F4 aghifoglie sempreverdi (**T3**), classe ECM-F4 latifoglie decidue (**T1**), classe ECM-F4 aghifoglie decidue (**T3X**) a livello nazionale, regionale e l'impatto sulle aree protette nazionali. Un esempio dell'applicazione del metodo è il caso della Regione Calabria rispetto agli incendi occorsi nell'estate 2021, si veda in particolare [Agrillo et al. \[2022\]](#).

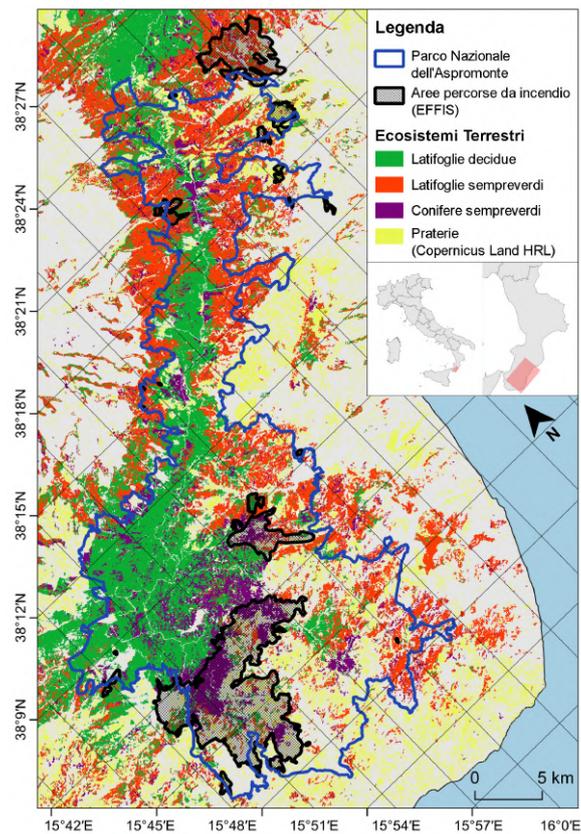


Figura 4.1: Distribuzione delle aree percorse da incendio nell'anno 2021 e rilasciate da European Forest Fire Information System (**EFFIS**) (poligoni neri) per il Parco Nazionale dell'Aspromonte (poligono blu) nella Regione Calabria. Viene riportata anche la classificazione degli ecosistemi forestali ottenuta grazie al Ecosystems Classification Model - Forest4 (**ECM-F4**) relativo all'anno 2020 e la classe prateria proveniente dal layer High Layer Resolution Land Monitoring Service Copernicus (**HLR**).

4.2 Piano di lavoro e sviluppo dei prodotti Changes in Italian Terrestrial Ecosystems (CITE)

Il sistema Changes in Italian Terrestrial Ecosystems (CITE) si pone l'obiettivo di coprire il gap informativo esistente in Italia riguardo alla mappatura dei cambiamenti degli ecosistemi terrestri a breve e lungo termine in risposta a fattori di disturbo ambientale, come per esempio gli incendi. La messa a punto del primo segmento del sistema Ecosystems Classification Model - Forest4 (ECM-F4), riguardante gli ecosistemi forestali, costituisce un importante strumento alla base di una serie di possibili servizi di monitoraggio ambientale. Il sistema è costituito da una serie di elementi che rendono possibile la calibrazione e verifica di modelli di classificazione basati su elaborazione di tipo Machine Learning di informazioni ambientali ed Osservazione della Terra (OT), in particolare dati MultiSpectral Instrument (MSI) - Sentinel-2 Copernicus. Il prodotto di mappatura ottenuto dalla proiezione del modello su tutto il territorio italiano ha un elevato tasso di copertura, 95% della superficie forestale, una altissima risoluzione spaziale e un livello elevato di affidabilità nella classificazione di quasi tutte le classi forestali individuate. La verifica del prodotto ha mostrato la validità della classificazione a scala nazionale, regionale e locale. Il programma di sviluppo del sistema CITE prevede diverse linee evolutive: in primo luogo l'estensione del sistema alla classificazione di altre categorie di ecosistemi terrestri, come CITE-G per le grassland e CITE-W per le wetland. Il prodotto per le classi forestali CITE-F verrà periodicamente aggiornato ed è in corso lo studio di come aumentare il numero delle classi individuabili mantenendo un alto livello di affidabilità delle mappature. Al momento è stato attivato un servizio di monitoraggio annuale dei cambiamenti a breve e lungo termine causati dagli incendi basato su prodotti Copernicus e ECM-F4: CITE-FF. Sono in fase di ricerca servizi orientati al monitoraggio dei cambiamenti dovuti a diversi detrattori ambientali. Il report CITE-FF è basato sui dati delle aree bruciate forniti da Copernicus EFFIS, e permette di stimare la quota parte di aree boschive bruciate annualmente, distinguendo le diverse tipologie forestali individuate da ECM-F4 a scala nazionale, regionale e nelle diverse aree protette italiane. Il caso degli incendi occorsi nell'estate 2021 ha mostrato la validità del prodotto ottenuto da CITE-FF nel definire l'impatto degli incendi superiori a 4 ha sugli ecosistemi forestali italiani. Nel tempo, sarà possibile verificare con ECM-F4 se gli effetti degli incendi saranno di breve periodo o se rischiano di diventare un fattore di degrado del patrimonio forestale italiano.

A.1 Lista degli acronimi ricorrenti

ADEME	Agence de la transition écologique
ATBD	Algorithm Theoretical Basis Document
CCUS	Carta di Copertura e Uso del Suolo
CCUS12	Copertura del suolo 2012-dati ISPRA-regionali-Copernicus
CDN	Carta della Natura
CITE	Changes in Italian Terrestrial Ecosystems
CITE-FF	Changes in Italian Terrestrial Ecosystems -Forest Fires
CLC	Corine Land Cover
CN-CRE	Centro Nazionale per le Emergenze, le Crisi Ambientali e il Danno
CSA	Centro Operativo per la Sorveglianza Ambientale
CUTFAA	Comando Unità per la Tutela Forestale, Ambientale e Agroalimentare
EA	U.K. Environment Agency
ECM-F4	Ecosystems Classification Model - Forest4
EFFIS	European Forest Fire Information System
EMAP	Environmental Monitoring and Assessment Program
EMS	Emergency Management Services
ENEA	Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile
EO	Earth Observation
EOS	Earth Observing System
EPA	U.S. Environmental Protection Agency
ES	Errore Standard
ESA	European Space Agency
ETM+	Enhanced Thematic Mapper Plus
EU	Unione Europea
EUNIS	EUropean Nature Information System
EVA	European Vegetation Archive
HDM	Habitat Distribution Models
HLR	High Layer Resolution Land Monitoring Service Copernicus
INFC	Inventario Nazionale delle Foreste e dei Serbatoi Forestali di Carbonio
ISPRA	Istituto Superiore per la Protezione e la Ricerca Ambientale

HRG	High Resolution Geometric
JPSS	Joint Polar Satellite System
JRC	Joint Research Center
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	MultiSpectral Instrument
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
OA	Overall Accuracy
OT	Osservazione della Terra
PA	Producer's Accuracy
PSD	Product Specification Document
RF	Random Forests
SE	Standard Error
SNPA	Sistema Nazionale per la Protezione dell'Ambiente
T1	classe ECM-F4 latifoglie decidue
T2	classe ECM-F4 latifoglie sempreverdi
T3	classe ECM-F4 aghifoglie sempreverdi
T3X	classe ECM-F4 aghifoglie decidue
UA	User's Accuracy
VIIRS	Visible Infrared Imaging Radiometer Suite

A.2 Acronimi variabili predittive modello ECM

LAI	valore medio annuale
LAI_{SD}	deviazione standard LAI rispetto a LAI
ΔLAI	variazione massima annuale LAI
LAI_{DJF}	LAI medio nel periodo dicembre, gennaio, febbraio
B8A_{Jul}	Riflettanza banda 8a MSI – valore medio luglio
B11_{Mar}	Riflettanza banda 11 MSI (valore medio marzo)
B11_{May}	Riflettanza banda 11 MSI (valore medio maggio)
B11_{Jul}	Riflettanza banda 11 MSI (valore medio luglio)
B11_{Aug}	Riflettanza banda 11 MSI (valore medio agosto)
B12_{Mar}	Riflettanza banda 12 MSI (valore medio marzo)
VIF	Variance Inflation Factor
DLT	Dominant Leaf Type
LAI	Leaf Area Index
EVI_{Mar}	Enhanced Vegetation Index (valore medio marzo)
GI	Indice di Gini
RI_{Mar}	Redness Index (valore medio marzo) Normalized Difference Red/Green
BDTICM	Absolute depth to bedrock
PHIHOX	pH calcolato in soluzione acquosa
NDYI_{Oct}	Normalized Difference Yellow Index (valore medio ottobre)
CRI_{Oct}	Carotenoid Reflectance Index (valore medio ottobre)

OOB	Out Of Bag Error
TCD	Tree Cover Density
NDVI	Normalized Difference Vegetation Index

A.3 Riferimenti WWW

ADEME: <https://www.ademe.fr/en/about-ademe>

Copernicus Land monitoring service “Imperviousness” : <https://land.copernicus.eu/pan-european/high-resolution-layers/imperviousness>

CCUS: <https://www.isprambiente.gov.it/it/attivita/suolo-e-territorio/copertura-del-suolo/carta-nazionale-di-copertura-del-suolo>

CCUS12: <https://groupware.sinanet.isprambiente.it/uso-copertura-e-consumo-di-suolo/library/copertura-del-suolo/carta-di-copertura-del-suolo/copertura-del-suolo-2012-da>

CDN: <https://www.isprambiente.gov.it/it/pubblicazioni/manuali-e-linee-guida/Il-Progetto-Carta-della-Natura-alla-scala-1-250.000>

ESA: <https://www.esa.int>

ESA Sentinel Program: <https://sentinel.esa.int/web/sentinel/missions>

ENEA: https://www.enea.it/it/Ricerca_sviluppo/sicurezza-e-salute

EMAP: <https://archive.epa.gov/emap/archive-emap/web/html/>

EPA: <https://www.epa.gov/>

EUNIS: <https://eunis.eea.europa.eu/>

Sentinel 2 MSI: <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/overview>

GRASS: <https://grass.osgeo.org>

R: <https://www.r-project.org>

SNAP: <https://step.esa.int/main/toolboxes/snap/>

Landsat 7: <https://directory.eoportal.org/web/eoportal/satellite-missions/1/landsat-7>

SPOT: <https://directory.eoportal.org/web/eoportal/satellite-missions/s/spot-5>

U.K. Environment Agency: <https://www.gov.uk/government/organisations/environment-agency>

Bibliografia

- M. Adamo, V. Tomaselli, C. Tarantino, S. Vicario, G. Veronico, R. Lucas, and P. Blonda. Knowledge-based classification of grassland ecosystem based on multi-temporal worldview-2 data and fao-lccs taxonomy. *Remote Sens.*, 12: <https://doi.org/10.3390/rs12091447>, 2020.
- E. Agrillo, F. Filippini, A. Pezzarossa, L. Casella, D. Smiraglia, A. Orasi, F. Attorre, and A. Taramelli. Earth observation and biodiversity big data for forest habitat types classification and mapping. *Remote Sensing*, 13(7), 2021.
- E. Agrillo, F. Filippini, R. Inghilesi, A. Mercatini, and A. Pezzarossa. Gli incendi boschivi in italia nel 2021. Technical report, Istituto Superiore per la Protezione e la Ricerca Ambientale, 2022.
- J. Álvarez-Martínez, B. Jiménez-Alfaro, J. Barquín, B. Ondiviela, M. Recio, A. Silió-Calzada, and J. Juanes. Modelling the area of occupancy of habitat types with remote sensing. *Methods Ecol. Evol.*, 9:580–593, 2018.
- B. Bischl, M. Lang, L. Kotthoff, J. Schiffner, J. Richter, E. Studerus, G. Casalicchio, and Z. M. Jones. mlr: Machine learning in r. *Journal of Machine Learning Research*, 17: 1–5, 2016.
- L. Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.
- L. Breiman, J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. DOI:<https://doi.org/10.1201/9781315139470>. Wadsworth International Group, 1984.
- P. Cantiani, U. D. Salvatore, and R. Romano. La selvicoltura delle pinete artificiali di pino nero: analisi delle legislazioni regionali italiane. *Forest@ - Rivista di Selvicoltura ed Ecologia Forestale*, 15(6):99–111, 2018.
- J. Chen and T. Black. Defining leaf-area index for non-flat leaves. *Plant Cell. Environ.*, 15:421–429, 1992.
- G. Chirici, C. Fattori, N. Cutolo, M. Tufano, P. Corona, A. Barbati, C. Blasi, R. Copiz, L. Rossi, D. Biscontini, A. Ribera, L. Morgante, and M. Marchetti. La realizzazione della carta delle formazioni naturali e semi-naturali e della carta forestale su basi tipologiche della regione Lazio. *Forest@ - Rivista di Selvicoltura ed Ecologia Forestale*, (2):65–71, 2014. doi: 10.3832/efor1204-011. URL <http://foresta.sisef.org/contents/?id=efor1204-011>.

- G. Chirici, F. Giannetti, R. E. McRoberts, D. Travaglini, M. Pecchi, F. Maselli, M. Chiesi, and P. Corona. Wall-to-wall spatial prediction of growing stock volume based on italian national forest inventory plots and remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation*, 84:101959, 2020. ISSN 0303-2434. doi: <https://doi.org/10.1016/j.jag.2019.101959>. URL <https://www.sciencedirect.com/science/article/pii/S0303243419305306>.
- L. Congedo, L. Sallustio, M. Munafò, M. Ottaviano, D. Tonti, and M. Marchetti. Copernicus high-resolution layers for land cover classification in italy. *Journal of Maps*, 12(5):1195–1205, 2016. doi: 10.1080/17445647.2016.1145151. URL <https://doi.org/10.1080/17445647.2016.1145151>.
- C. Corbane, S. Lang, K. Pipkins, S. Alleaume, M. Deshayes, V. Millán, T. Strasser, J. Borre, S. Toon, and F. Michael. Remote sensing for mapping natural habitats and their conservation status—new opportunities and challenges. *Int. J. Appl. Earth Obs.*, 37:7–16, 2015.
- W. De Simone, M. Allegrezza, A. R. Frattaroli, S. Montecchiari, G. Tesei, V. Zuccarello, and M. Di Musciano. From remote sensing to species distribution modelling: An integrated workflow to monitor spreading species in key grassland habitats. *Remote Sensing*, 13(10), 2021. ISSN 2072-4292. doi: 10.3390/rs13101904. URL <https://www.mdpi.com/2072-4292/13/10/1904>.
- A. Della Rocca, S. Pignatti, and S. P. S. Mugnoli. La carta della vegetazione della tenuta di castelporziano. In *Progetto DEDALUS ENEA*, volume 26 of *Scritti e Documenti*, pages 709–748. Accademia Nazionale delle Scienze detta dei XL, 2001.
- A. Dostálová, W. Wagner, M. Milenkovic, and M. Hollaus. Annual seasonality in sentinel-1 signal for forest mapping and forest type classification. *International Journal of Remote Sensing*, 39:7738 – 7760, 2018.
- F. Filipponi, D. Smiraglia, and E. Agrillo. Earth observation for phenological metrics (eo4pm): Temporal discriminant to characterize forest ecosystems. *Remote Sensing*, 14(3)(doi:10.3390/rs14030721), 2022.
- E. Grabska, P. Hostert, D. Pflugmacher, and K. Ostapowicz. Forest stand species mapping using the sentinel-2 time series. *Remote. Sens.*, 11:1197, 2019.
- H. Guo, L. Wang, and D. Liang. Big earth data from space: A new engine for earth science. *Sci. Bull.*, 61:505–513, 2016.
- W. Hallgren, L. Beaumont, A. Bowness, L. Chambers, E. Graham, H. Holewa, S. Laffan, B. Mackey, H. Nix, and J. Price. The biodiversity and climate change virtual laboratory: Where ecology meets big data. *Environ. Model. Softw.*, 76:182–186, 2016.
- S. Hampton, C. Strasser, J. Tewksbury, W. Gram, B. A. A. Batcheller, C. Duke, and J. Porter. Big data and the future of ecology. *Front. Ecol. Environ.*, 11:156–162, 2013.

- T. Kattenborn, J. Lopatin, M. Förster, A. C. Braun, and F. E. Fassnacht. Uav data as alternative to field sampling to map woody invasive species based on combined sentinel-1 and sentinel-2 data. *Remote Sensing of Environment*, 227:61–73, 2019.
- M. Krzywinski and N. Altman. Classification and classification and regression trees. *Nature Methods*, 14(8):757–758, 2017.
- M. B. Kursa and W. Rudniki. Feature selection with boruta package. *Journal of Statistical Software*, 36(11):1–13, 2010.
- M. Lindner, M. Maroschek, S. Netherer, A. Kremer, A. Barbati, J. Garcia-Gonzalo, R. Seidl, S. Delzon, P. Corona, M. Kolström, M. J. Lexer, and M. Marchetti. Climate change impacts, adaptive capacity, and vulnerability of european forest ecosystems. *Forest Ecology and Management*, 259(4):698–709, 2010. ISSN 0378-1127. doi: <https://doi.org/10.1016/j.foreco.2009.09.023>. URL <https://www.sciencedirect.com/science/article/pii/S0378112709006604>. Adaptation of Forests and Forest Management to Changing Climate.
- G. Lovett, D. Burns, C. Driscoll, J. Jenkins, M. Mitchell, L. Rustad, J. Shanley, G. Likens, and R. Haeuber. Who needs environmental monitoring? *Frontiers in Ecology and the Environment*, 5:253–260, 05 2007. doi: 10.1890/1540-9295(2007)5[253:WNEM]2.0.CO;2.
- F. Marzialetti, S. Giulio, M. Malavasi, M.G.Sperandii, A. Acosta, and M. Carranza. Capturing coastal dune natural vegetation types using a phenology-based mapping approach: The potential of sentinel-2. *Remote Sens.*, 11:1506, 2019.
- A. P. Mol. The environmental nation state in decline. *Environmental politics*, 1(25):48–68, 2016.
- V. Olmo, E. Tordoni, F. Petruzzellis, G. Bacaro, and A. Altobelli. Use of sentinel-2 satellite data for windthrows monitoring and delimiting: The case of “vaia” storm in friuli venezia giulia region (north-eastern italy). *Remote Sensing*, 13, 04 2021. doi: 10.3390/rs13081530.
- S. Pesaresi, D. Galdenzi, E. Biondi, and S. Casavecchia. Bioclimate of italy: application of the worldwide bioclimatic classification system. *Journal of Maps*, 10:538 – 553, 2014.
- S. Pesaresi, A. Mancini, G. Quattrini, and S. Casavecchia. Mapping mediterranean forest plant associations and habitats with functional principal component analysis using landsat 8 ndvi time series. *Remote Sens.*, 12:DOI:10.3390/RS12071132, 2020.
- S. Pignatti, R. Cavalli, V. Cuomo, L. Fusilli, S. Pascucci, P. Maurizio, and F. Santini. Evaluating hyperion capability for land cover mapping in a fragmented ecosystem: Pollino national park, italy, 01 2014.

- P. Probst, M. N. Wright, and A.-L. Boulesteix. Hyperparameters and tuning strategies for random forest. *Wires*, 9(3), 2019.
- L. Raileanu and K. Stoffel. Theoretical comparison between the gini index and information gain criteria. *Annals of Mathematics and Artificial Intelligence*, 41:77–93, 2004.
- J. Rockström, W. Steffen, K. Noone, A. Persson, F. S. Chapin, E. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sörlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. Foley. Planetary boundaries: exploring the safe operating space for humanity. *Ecology and Society*, 14(2):32, 2009.
- M. Rüetschi, M. E. Schaepman, and D. Small. Using multitemporal sentinel-1 c-band backscatter to monitor phenology and classify deciduous and coniferous forests in northern switzerland. *Remote Sensing*, 10(1), 2018. ISSN 2072-4292. doi: 10.3390/rs10010055. URL <https://www.mdpi.com/2072-4292/10/1/55>.
- R. Runting, S. Phinn, Z. Xie, O. Venter, and J. Watson. Opportunities for big data in conservation and sustainability. *Nat. Commun.*, 11(1): <https://doi.org/10.1038/s41467-020-15870-0>, 2020.
- M.-J. Schelhaas, G.-J. Nabuurs, and A. Schuck. Natural disturbances in the European forests in the 19th and 20th centuries. *Global Change Biology*, 9(11):1620–1633, Nov. 2003. doi: 10.1046/j.1365-2486.2003.00684.x.
- H. H. Shugart. *Terrestrial ecosystems in changing environments*. Cambridge Univ. Press, 1998.
- G. L. Spadoni, A. Cavalli, L. Congedo, and M. Munafò. Analysis of normalized difference vegetation index (ndvi) multi-temporal series for the production of forest cartography. *Remote Sensing Applications: Society and Environment*, 20:100419, 2020. ISSN 2352-9385. doi: <https://doi.org/10.1016/j.rsase.2020.100419>. URL <https://www.sciencedirect.com/science/article/pii/S2352938520302858>.
- S. V. Stehman and G. M. Foody. Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, 231, 2019.
- A. Strollo, D. Smiraglia, R. Bruno, F. Assennato, L. Congedo, P. D. Fioravante, C. Giuliani, I. Marinosci, N. Riitano, and M. Munafò. Land consumption in italy. *Journal of Maps*, 16(1):113–123, 2020. doi: 10.1080/17445647.2020.1758808. URL <https://doi.org/10.1080/17445647.2020.1758808>.
- M. Wegmann, B. Leutner, and S. Dech, editors. *Remote Sensing and GIS for Ecologists: Using Open Source Software*. Pelagic publishing, 2016.

-
- G. Wiersma and D. Bruns. *Conceptual Basis of Environmental Monitoring Systems: A Geospatial Perspective*, pages 1–35. 04 2004. ISBN 978-1-56670-641-4. doi: 10.1201/9780203495476.ch1.
- M. N. Wright and A. Ziegler. ranger: A fast implementation of random forests for high dimensional data in c++ and r. *Journal of Statistical Software*, 77(1):1–17, 2017.
- A. F. Zuur, E. N. Leno, and C. S. Elphick. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1:3–14, 2010.