Relazioni tra clima, fenologia e produttività delle principali specie forestali italiane da prodotti Copernicus HR-VPP

Elia Vangi[1,2], Giovanni D'Amico[1], Vincenzo Saponaro[2,3], Mattia Niccoli[1], Gioele Tiberi[4], Saverio Francini[5], Costanza Borghi[1], Alessio Collalti[2], Francesco Parisi[6,7] e Gherardo Chirici[1,8]

[1] geoLAB, Laboratorio di Geomatica Forestale, Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali (DAGRI), Università degli Studi di Firenze, Via San Bonaventura 13, 50145 Firenze
[2] Forest Modelling Laboratory, Istituto per i Sistemi Agricoli e Forestali del Mediterraneo, Consiglio Nazionale delle Ricerche (CNR-ISAFOM), Via Madonna Alta 128, 06128 Perugia
[3] Facoltà di Scienze Agrarie, Ambientali e Alimentari, Libera Università di Bolzano, Piazza Università, 39100 Bolzano
[4] Dipartimento di Architettura (DIDA), Università degli Studi di Firenze, Via della Mattonaia 8, 50121 Firenze
[5] Dipartimento di Scienze e Tecnologie Agro-Alimentari (DISTAL), Università di Bologna, 40126 Bologna
[6] Dipartimento di Bioscienze e Territorio, Università degli Studi del Molise, Contrada Fonte Lappone, 86090 Pesche (IS)
[7] NBFC, National Biodiversity Future Center, 90133 Palermo
[8] Fondazione per il Futuro delle Città, Firenze
elia.vangi@unifi.it

Abstract. In questo studio, analizziamo le tempistiche fenologiche e la produttività stagionale complessiva di diverse specie forestali italiane distribuite in ambienti mediterranei, temperati e montani, avvalendoci del nuovo prodotto ad alta risoluzione sulla produttività e sulla fenologia della vegetazione del Copernicus Land Monitoring Service, di modelli di machine learning (random forest) e di analisi basate sull'intelligenza artificiale spiegabile (SHAP). I risultati indicano una tendenza generale all'allungamento della stagione vegetativa, guidata principalmente dall'accumulo di freddo invernale e dalle temperature primaverili. Le condizioni più calde anticipano l'inizio della stagione di 1-10 giorni, a seconda della specie, mentre gli effetti combinati di temperatura, radiazione e umidità possono prolungare la stagione vegetativa di 20-30 giorni. Le dinamiche di fine stagione e la sua durata sono influenzate più dalla disponibilità di luce e di acqua che dalla sola temperatura. In diverse specie mediterranee, la fine della stagione può anticipare fino a 40 giorni a causa della siccità estiva, dell'elevato deficit di pressione di vapore e dell'esposizione del sito. A causa delle variazioni fenologiche, la durata complessiva della stagione risulta spesso disaccoppiata dalla produttività, evidenziando risposte al cambiamento climatico specifiche

per specie e sito. I risultati sottolineano l'importante vantaggio della combinazione di dati di telerilevamento e tecniche di machine learning per approfondire la comprensione della fenologia e della produttività forestale lungo ampi gradienti spaziali e climatici.

Parole chiave: Fenologia, Inventario Forestale Nazionale, Sentinel-2, Copernicus HR-VPP, SHAP.

1 Introduzione

Le foreste italiane si distribuiscono lungo uno dei gradienti ambientali più ampi d'Europa. In poco più di 11° di latitudine si passa dalle formazioni alpine, limitate dalla temperatura, agli ecosistemi mediterranei, limitati dalla disponibilità idrica. Su questo gradiente, il cambiamento climatico agisce in modo disomogeneo e la fenologia della vegetazione è uno degli indicatori più sensibili dei suoi effetti [1]. Nell'Europa centrale e settentrionale l'allungamento della stagione vegetativa è stato spesso associato a un aumento della crescita e dell'assorbimento di carbonio [2]. Studi recenti mostrano, però, che la senescenza fogliare può avvenire indipendentemente dalla fine dell'attività fotosintetica [3, 4]. Nell'Italia meridionale, inoltre, la capacità di assorbimento del carbonio risulta già disaccoppiata dall'incremento di area basimetrica [5]. Resta quindi aperta la domanda se una stagione più lunga corrisponda davvero a una foresta più produttiva, soprattutto dove l'acqua e la radiazione diventano fattori limitanti.

Rispondere a questa domanda su scala nazionale richiede osservazioni fenologiche dense e a una risoluzione compatibile con quella dei popolamenti. Le osservazioni di campo sono specie-specifiche ma costose e spazialmente limitate. I sensori a bassa risoluzione come MODIS, d'altra parte, aggregano nello stesso pixel specie, quote e usi del suolo diversi, un limite particolarmente evidente nei paesaggi eterogenei del Mediterraneo. Dal 2023 il Copernicus Land Monitoring Service (CLMS) distribuisce la suite High-Resolution Vegetation Phenology and Productivity (HR-VPP), che deriva fenologie annuali a 10 m dalle serie temporali del Plant Phenology Index (PPI) di Sentinel-2 [6, 7]. Il PPI è un indice basato su misure fisiche, linearmente correlato al Leaf Area Index (LAI) e poco sensibile alla copertura nevosa, e per questo si presta bene a seguire le fasi della chioma.

Le relazioni tra fenologia e clima non sono lineari e le variabili climatiche sono fortemente correlate tra loro. Gli algoritmi di machine learning (ML) gestiscono bene queste condizioni, ma producono modelli difficili da interpretare. I metodi di explainable artificial intelligence (XAI), e in particolare le SHapley Additive exPlanations (SHAP) [8], scompongono ogni previsione nel contributo delle singole variabili, espresso nella stessa unità di misura della variabile di risposta (giorni o valori di PPI).

Questo contributo sintetizza i risultati di Vangi et al. [9], uno dei pochi studi, a nostra conoscenza, che combina le fenometriche HR-VPP, i dati dell'Inventario Nazionale delle Foreste e dei serbatoi forestali di Carbonio (INFC) e l'XAI per analizzare fenologia e produttività di più specie forestali a scala nazionale. Gli obiettivi erano due: verificare se un approccio ML automatizzato riproduce i pattern macroecologici noti

per dieci specie, e testare l'ipotesi che l'allungamento della stagione indotto dal riscaldamento sia disaccoppiato dalla produttività negli ambienti mediterranei.

2 Materiali e metodi

2.1 Area di studio e dati inventariali

Lo studio copre l'intero territorio nazionale, dove le foreste occupano oltre il 35% della superficie (circa 11,05 milioni di ettari) e sono dominate dalle latifoglie, che rappresentano il 68% della superficie forestale [10]. Le aree di saggio provengono dalla terza fase dell'INFC2015, in cui i dati quantitativi sono rilevati in aree circolari di raggio 13 m su un campione di 6174 punti [10]. Per ridurre il rumore dovuto ai popolamenti misti sono state selezionate solo le aree in cui la specie dominante rappresentava almeno l'80% dell'area basimetrica. Tra queste sono state mantenute le specie presenti in almeno 100 aree. Il campione finale comprende 2453 aree e dieci specie (Fig. 1): *Castanea sativa* Mill., *Fagus sylvatica* L., *Larix decidua* Mill., *Picea abies* (L.) H. Karst., *Pinus halepensis* Mill., *Pinus nigra* J.F. Arnold, *Quercus cerris* L., *Quercus ilex* L., *Quercus pubescens* Willd. e *Quercus suber* L.

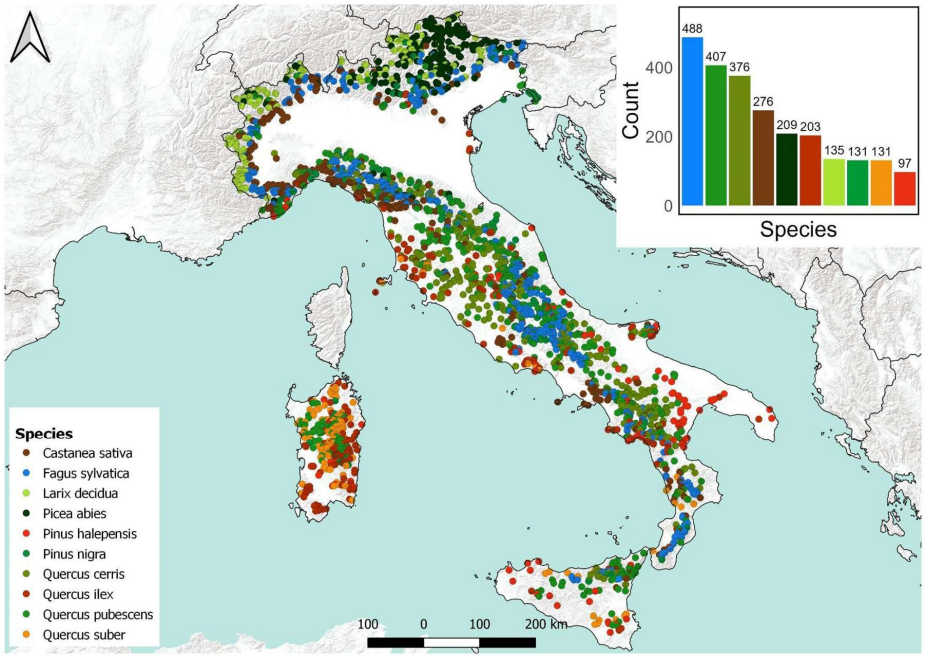

Fig. 1. Distribuzione delle 2453 aree INFC analizzate. Il grafico in alto a destra riporta il numero di aree per specie (da [9], licenza CC BY 4.0).

2.2 Fenometriche e variabili predittive

Dalla suite HR-VPP sono state scaricate, tramite le API del catalogo WEkEO, quattro fenometriche per il periodo 2017–2023: data di inizio (SOSD) e di fine (EOSD) della stagione vegetativa, durata della stagione (LOS) e produttività stagionale totale (TPROD), cioè l'integrale del PPI tra SOSD ed EOSD. Il dataset completo comprende 3584 tile, per circa 400 GB. Le fenometriche sono generate con TIMESAT [7], che filtra nuvole, ombre e outlier dalle serie PPI e adatta una funzione logistica doppia a ciascuna stagione. SOSD ed EOSD corrispondono ai giorni in cui il PPI raggiunge, rispettivamente, il 25% e il 15% dell'ampiezza stagionale. Per garantire un allineamento coerente con i predittori climatici annuali è stata considerata solo la prima stagione di ogni anno, purché iniziata e conclusa nello stesso anno solare.

Le variabili climatiche derivano dalla rianalisi ERA5 regionalizzata sull'Italia a 2,2 km con il modello COSMO-CLM [11]. Dai dati orari di temperatura, radiazione solare, precipitazione e deficit di pressione di vapore (VPD) sono state calcolate, per ogni anno, 176 metriche mensili e stagionali. Queste descrivono la forzante termica primaverile, l'accumulo di freddo invernale (chilling), la radiazione e lo stress idrico. Quota, pendenza ed esposizione sono state ricavate dal modello digitale del terreno TINITALY a 10 m [12], a cui si aggiungono la latitudine, la distanza dal mare e l'età del popolamento. Tutte le variabili sono state estratte come media (moda per le date) in un intorno di 13 m dal centro di ciascuna area. Il database finale conta 17.171 osservazioni (2453 aree × 7 anni) e 185 variabili.

2.3 Selezione delle variabili, modellizzazione e interpretazione

Data la brevità della serie HR-VPP, le fenometriche e i predittori sono stati mediati per area nel periodo 2017–2023. L'analisi segue quindi un approccio di sostituzione spazio-temporale, che sfrutta l'ampio gradiente ambientale occupato da ciascuna specie. Per ciascuna delle 40 combinazioni specie-fenometriche, le variabili sono state selezionate con l'algoritmo VSURF [13], che individua, in tre passaggi, il sottoinsieme minimo di predittori con una capacità predittiva accettabile, limitando gli effetti della collinearità. Con le variabili selezionate è stato addestrato un random forest condizionale (pacchetto R *party* [14]), valutato tramite l'errore quadratico medio (RMSE) out-of-bag.

Le previsioni di ciascun modello sono state interpretate con l'approccio SHAP condizionale basato su alberi di inferenza condizionale [15], implementato nel pacchetto R *shapr* e adatto a predittori dipendenti. Un valore SHAP positivo indica un ritardo di SOSD ed EOSD, un allungamento di LOS o un aumento di TPROD, mentre un valore negativo indica l'effetto opposto. Per la visualizzazione le variabili sono state raggruppate in cinque classi (temperatura invernale e chilling, temperatura primaverile, disponibilità idrica, fotoperiodo e topografia), sommando i valori SHAP grazie alla loro additività. Infine, per verificare il disaccoppiamento tra la durata della stagione e la produttività, le anomalie interannuali standardizzate (z-score) di TPROD sono state regredite su quelle di LOS. Il coefficiente di determinazione (R^2) di questa regressione misura il grado di correlazione tra le due grandezze.

3 Risultati

3.1 Prestazioni dei modelli e variabili selezionate

VSURF ha selezionato, in media, sei variabili per combinazione, da un minimo di due (TPROD in *P. nigra*) a un massimo di dieci (LOS in *Q. pubescens*). Le metriche di temperatura e precipitazione dominano la selezione per quasi tutte le fenometriche. Quota e pendenza sono comunque spesso tra le variabili più importanti, in particolare per *C. sativa*, *F. sylvatica* e *L. decidua*. L'inizio della stagione presenta un RMSE medio di 11,5 giorni, mentre la fine della stagione è molto più difficile da modellare, con un RMSE medio di 28 giorni (Table 1). Gli errori più elevati si concentrano nelle specie mediterranee: in *P. halepensis*, *Q. pubescens* e *Q. suber* l'RMSE di EOSD varia tra 39 e 56 giorni.

Table 1. RMSE out-of-bag dei modelli per specie e fenometrica (da [9]). SOSD, EOSD e LOS sono espressi in giorni, TPROD in unità di PPI.

Specie	SOSD	EOSD	LOS	TPROD
Castanea sativa	9	15	16	60
Fagus sylvatica	8	13	16	72
Larix decidua	10	15	19	51
Picea abies	14	14	22	57
Pinus halepensis	15	56	50	52
Pinus nigra	13	25	28	71
Quercus cerris	11	28	26	63
Quercus ilex	11	20	23	49
Quercus pubescens	12	42	33	57
Quercus suber	10	39	34	46

3.2 Inizio, fine e durata della stagione

Le temperature minime e massime di gennaio, febbraio e aprile sono i principali determinanti di SOSD per la maggior parte delle specie. Nelle specie mediterranee l'inizio della stagione è più sensibile al chilling invernale, con un anticipo di circa 0,5 giorni per grado contro circa 0,3 giorni nelle specie montane. Queste ultime rispondono invece maggiormente alla temperatura primaverile, con un anticipo fino a 0,75 giorni per grado. Nel complesso, condizioni più calde anticipano l'inizio della stagione di 1–10 giorni e l'effetto combinato di temperatura, radiazione e umidità può allungarla di 20–30 giorni.

La fine della stagione e la sua durata dipendono più da luce e disponibilità idrica che dalla sola temperatura. In *P. halepensis*, *Q. ilex* e *Q. suber*, la radiazione e il VPD primaverili ed estivi, insieme alle precipitazioni, possono anticipare l'EOSD fino a 40 giorni. Nelle specie montane (*L. decidua* e *P. abies*) prevalgono la quota e la radiazione estiva, mentre la pendenza riduce la durata della stagione in *P. nigra*, *Q. cerris* e *Q. pubescens*. La Fig. 2 riassume il contributo dei cinque gruppi di variabili a LOS. La

durata della stagione varia complessivamente fino a 20 giorni, con una tendenza all'accorciamento nelle specie mediterranee e all'allungamento in quelle montane.

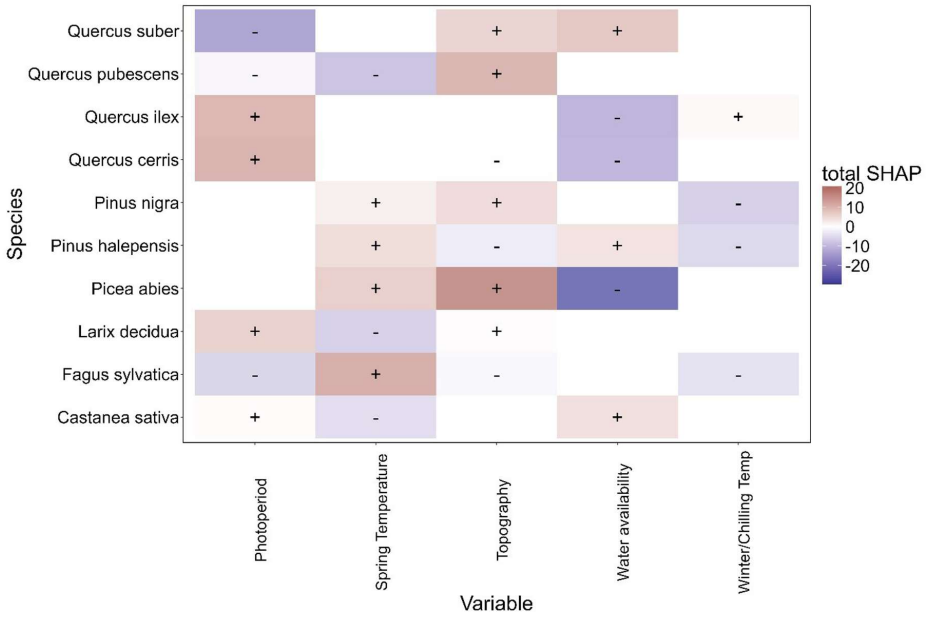

Fig. 2. Valore SHAP totale di LOS per gruppo di variabili (asse x) e specie (asse y), espresso in giorni rispetto alla previsione media. I segni "+" e "–" indicano il verso del contributo (da [9], licenza CC BY 4.0).

3.3 Produttività

La radiazione solare primaverile e autunnale è il principale driver di TPROD in *P. nigra* e *C. sativa*, mentre nelle specie montane domina ancora la quota. Nelle specie mediterranee, la produttività risponde in modo non lineare ai predittori autunnali, con evidenti effetti a soglia. La variazione della radiazione in autunno, ad esempio, ha un effetto positivo solo quando la riduzione stagionale è moderata (pendenza della regressione lineare > −2). L'esposizione influenza TPROD in *F. sylvatica*, *Q. cerris* e *Q. pubescens*, con contributi positivi tra 0° e 90° e tra 250° e 360°. Il legame interannuale tra anomalie di LOS e di TPROD è debole per tutte le specie. R^2 resta sempre inferiore a 0,2, con un massimo di 0,14 in *L. decidua* e valori inferiori a 0,01 per tutte le querce.

4 Discussione

I risultati confermano, con un approccio interamente guidato dai dati, il comportamento atteso per le specie forestali italiane: inverni e primavere più caldi anticipano l'inizio della stagione. Il peso delle temperature minime invernali e delle massime di aprile

suggerisce che contino sia l'accumulo di freddo sia la forzante primaverile, in accordo con il ruolo del chilling nella quiescenza e nel germogliamento [16]. Nelle specie alpine la quota sintetizza una buona parte dell'informazione climatica, molto probabilmente perché il modello del terreno a 10 m descrive gradienti microclimatici che la rianalisi a 2,2 km non riesce a risolvere.

La fine della stagione resta la fase più difficile da prevedere, in linea con l'assenza di consenso nella letteratura sui suoi determinanti. Le specie mediterranee rispondono in modo diverso alla siccità estiva, verosimilmente in base alla loro strategia idraulica. Nelle specie più isoidriche (*P. halepensis*, *Q. ilex*, *Q. suber*) la chiusura stomatica e la perdita fogliare, che limitano il rischio di cavitazione, anticipano la fine della stagione [17]. Nelle specie più tolleranti alla siccità, come *Q. pubescens* e *Q. cerris*, la crescita si arresta durante l'estate e riprende con le piogge autunnali, ritardando la senescenza. Questi spostamenti compensativi tra l'inizio e la fine mantengono la durata della stagione relativamente stabile. Nelle specie montane, al contrario, un inizio ritardato si traduce più direttamente in una stagione più breve.

È interessante notare che anche la temperatura primaverile contribuisce a spiegare la senescenza, come ipotizzato da Zani et al. [3] e Zohner et al. [4]. Una primavera più calda potrebbe esaurire prima le riserve idriche del suolo, oppure saturare prematuramente la capacità di sink del carbonio. La selezione della radiazione per nove specie su dieci è coerente soprattutto con questo secondo meccanismo, più che con un decadimento passivo guidato unicamente dalla temperatura.

Il risultato principale riguarda il legame tra durata della stagione e produttività. L'assunto, diffuso negli studi sul cambiamento climatico, secondo cui una stagione più lunga porti a una maggiore produttività [1, 2] non trova riscontro nei dati. Valori di R^2 inferiori a 0,2 per tutte le specie, e prossimi a zero per le querce, indicano un forte disaccoppiamento. La produttività sembra essere regolata soprattutto dalla disponibilità di energia e di acqua, con effetti a soglia e un ruolo dell'esposizione nel modulare il microclima. In ambiente mediterraneo, quindi, un allungamento della stagione può segnalare un periodo di siccità più lungo piuttosto che un guadagno di carbonio.

Alcuni limiti vanno considerati nell'interpretare questi risultati. Gli indici spettrali osservano le variazioni ottiche della chioma e non direttamente gli eventi fenologici, e la selezione di aree pure riduce ma non elimina la miscelazione spettrale con il sottobosco. L'aggregazione sul periodo 2017–2023 attenua inoltre l'effetto degli eventi estremi, sempre più rilevanti per il funzionamento delle foreste. L'allungamento progressivo della serie HR-VPP permetterà di passare dalla sostituzione spazio-tempo all'analisi diretta delle tendenze temporali.

5 Conclusioni

Questo lavoro mostra che le fenometriche Copernicus HR-VPP, combinate con i dati INFC, il ML e l'analisi SHAP, consentono di studiare la fenologia di singole specie forestali a livello nazionale. L'approccio riproduce i pattern noti all'inizio della stagione e, al contempo, evidenzia risposte specie- e sito-specifiche alla fine della stagione e sulla produttività, soprattutto negli ambienti limitati dall'acqua. Il

riscaldamento futuro probabilmente continuerà ad anticipare l'inizio della stagione, ma i suoi effetti sulla produttività non potranno essere dedotti dalla sola durata della stagione. Questo ha implicazioni dirette sulla parametrizzazione dei modelli di vegetazione e sulle stime del bilancio del carbonio forestale. Le nove fenometriche HR-VPP non analizzate in questo studio offrono ulteriori opportunità per il monitoraggio operativo degli ecosistemi forestali.

Ringraziamenti. E.V. ringrazia il progetto NextGenCarbon (Horizon Europe, n. 101184989) e il progetto Space It Up!, finanziato dall'Agenzia Spaziale Italiana e dal Ministero dell'Università e della Ricerca (contratto n. 2024-5-E.0, CUP I53D24000060005). A.C. ringrazia il progetto FORESTNAVIGATOR (Horizon Europe, n. 101056875).

Riferimenti bibliografici

1. Piao, S., Liu, Q., Chen, A., Janssens, I.A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., Zhu, X.: Plant phenology and global climate change: current progresses and challenges. Global Change Biology 25, 1922–1940 (2019).
2. Keenan, T.F., Gray, J., Friedl, M.A., Toomey, M., Bohrer, G., Hollinger, D.Y., et al.: Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nature Climate Change 4(7), 598–604 (2014).
3. Zani, D., Crowther, T.W., Mo, L., Renner, S.S., Zohner, C.M.: Increased growing-season productivity drives earlier autumn leaf senescence in temperate trees. Science 370(6520), 1066–1071 (2020).
4. Zohner, C.M., Mirzagholi, L., Renner, S.S., Mo, L., Rebindaine, D., Bucher, R., et al.: Effect of climate warming on the timing of autumn leaf senescence reverses after the summer solstice. Science 381, eadf5098 (2023).
5. Dalmonech, D., Vangi, E., Chacón, D.Q., Collalti, A.: Decadal climate-driven decoupling between gross primary productivity and tree growth in Mediterranean forests. Forest Ecology and Management 615, 123843 (2026).
6. Jin, H., Eklundh, L.: A physically based vegetation index for improved monitoring of plant phenology. Remote Sensing of Environment 152, 512–525 (2014).
7. Jönsson, P., Eklundh, L.: TIMESAT, a program for analyzing time-series of satellite sensor data. Computers & Geosciences 30(8), 833–845 (2004).
8. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Advances in Neural Information Processing Systems 30, pp. 4765–4774. Curran Associates, Red Hook (2017).
9. Vangi, E., D'Amico, G., Saponaro, V., Niccoli, M., Tiberi, G., Francini, S., Borghi, C., Collalti, A., Parisi, F., Chirici, G.: Investigating climate-phenology relationships among the most common Italian forest species using Sentinel-2-derived vegetation phenology and productivity products. Agricultural and Forest Meteorology 390, 111496 (2026).
10. INFC: Italian forests. Selected results of the third National Forest Inventory (INFC2015). Arma dei Carabinieri, CREA Centro di ricerca Foreste e Legno. Tipografia Supernova, Trento (2021).
11. Raffa, M., Reder, A., Marras, G.F., Mancini, M., Scipione, G., Santini, M., Mercogliano, P.: VHR-REA_IT dataset: very high resolution dynamical downscaling of ERA5 reanalysis over Italy by COSMO-CLM. Data 6(8), 88 (2021).

12. Tarquini, S., Isola, I., Favalli, M., Mazzarini, F., Bisson, M., Pareschi, M.T., Boschi, E.: TINITALY/01: a new triangular irregular network of Italy. Annals of Geophysics 50, 407–425 (2007).
13. Genuer, R., Poggi, J.-M., Tuleau-Malot, C.: VSURF: an R package for variable selection using random forests. The R Journal 7(2), 19–33 (2015).
14. Hothorn, T., Hornik, K., Zeileis, A.: Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics 15(3), 651–674 (2006).
15. Redelmeier, A., Jullum, M., Aas, K.: Explaining predictive models with mixed features using Shapley values and conditional inference trees. In: Holzinger, A., Kieseberg, P., Tjoa, A.M., Weippl, E. (a cura di) CD-MAKE 2020, LNCS, vol. 12279, pp. 117–137. Springer, Cham (2020).
16. Hänninen, H.: Boreal and temperate trees in a changing climate. Biometeorology, vol. 3. Springer, Dordrecht (2016).
17. McDowell, N., Pockman, W.T., Allen, C.D., Breshears, D.D., Cobb, N., Kolb, T., et al.: Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? New Phytologist 178, 719–739 (2008).