

Faculté des bioingénieurs

Multispectral Imagery for Forest Conditions Monitoring of a Pilot Watershed in the Vosges, France

Auteur: Janathi Wathsala Rathnayaka Wijesingha Rathnayaka Appuhamilage

Promoteurs : Pr. Pierre Defourny

Co-Promoteurs : Dr. Maxime Thomas, Tom Kenda

Lecteurs : Dr. Christian Piedallu

Ir. Baptiste Delhez

Année académique : 2025-2026

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Master : SAIV -
MSc Erasmus+ GEM : Geo-Information Science and Earth Observation for
Environmental Modelling and Management

Acknowledgements

First and foremost, I would like to thank the consortium of the Geoinformation Science and Earth Observation for Environmental Modelling and Management (GEM) programme for providing me with the opportunity to be part of this unique and enriching academic experience. The knowledge, skills, and international exposure gained through the programme have been invaluable and have greatly contributed to both my academic and personal development.

I would like to express my sincere gratitude to my supervisor, Professor Pierre Defourny, for his guidance, support, and encouragement throughout this research. His expertise and advice provided valuable direction and helped shape this work.

I am especially grateful to my co-supervisors, Maxime Thomas and Tom Kenda, for their continuous support, constructive feedback, and guidance at every stage of this thesis. Their insightful comments, encouragement, and willingness to share their knowledge were invaluable in developing and strengthening this research. I am truly appreciative of the time and effort they dedicated to supporting me throughout this process.

I would also like to extend my sincere thanks to Laurent Saint-André (INRAE), Christian Piedallu (AgroParisTech), Violette Gautier (AgroParisTech), and the PHENET Project Work Package 3 team. Their contribution in providing the field dataset, which formed a critical foundation for this study, was essential to the successful completion of this research. I am equally grateful for their guidance, expertise, and support during the course of this work.

Finally, I would like to thank my family and friends for their constant encouragement, understanding, and support throughout my studies. Their belief in me, patience, and unwavering support helped me overcome challenges and stay motivated throughout this journey.

To all those who contributed directly or indirectly to this thesis, thank you.

Statement of Contribution and Use of Generative AI

Stage	Author Janathi Rathnayaka	Other Contributor(s) Maxime Thomas Tom Kenda	Artificial Intelligence ChatGPT Gemini
Conceptualization	Developed research objectives and study design	Scientific guidance and feedback	Literature exploration
Methodology	Developed and implemented methodology	Methodological advice	-
Investigation	Conducted data processing and analysis	INRAE, AgroParisTech, PHENET: field data collection	-
Formal Analysis	Performed statistical and machine learning analyses	Discussion and interpretation support	Python coding assistance
Validation	Conducted model validation	Review of results	-
Writing – original draft	Wrote the manuscript	-	Language and grammar support
Writing – review & editing	Revised and edited the manuscript	Feedback and corrections	Text refinement
Visualization	Created maps, figures, and tables	Suggestions	-
Supervision	-	Scientific supervision	-

This thesis was conducted by the author, Janathi Wathsala Rathnayaka, under the co-supervision of Maxime Thomas and Tom Kenda. The research objectives, methodological framework, analysis strategy, and interpretation of results were developed through regular discussions with the co-supervisors. The author carried out all data processing, calculations, statistical analyses, machine learning model development, validation procedures, map production, and thesis writing. ChatGPT and Gemini were used as supporting AI tools for literature exploration, Python programming assistance, code debugging, methodological improvements, language editing, grammar correction, and text reformulation. All AI-generated outputs were critically reviewed and verified by the author.

Table of Contents

1. Introduction.....	6
2. Literature Review.....	8
2.1 Forest Ecosystems and Drivers of Forest Degradation.....	8
2.2 Dynamics of Forest Dieback in a Changing Climate.....	9
2.3 Remote Sensing in Forest Dieback Detection	11
2.4 Potential of Multispectral Remote Sensing for Forest Dieback Detection	14
2.5 Artificial Intelligence (AI) for Forest Health Monitoring.....	15
2.5.1 Machine Learning Approaches in Forest Dieback Detection	17
2.6 Challenges in Remote Sensing and ML-based Forest Dieback Detection	18
3. Study Objectives	21
4. Methodology	22
4.1 Study Area.....	22
4.2 Data Description	23
4.2.1 Field Data.....	23
4.2.2 Satellite Data	26
4.3 Computation of Vegetation Indices.....	28
4.4 Time Series Extraction.....	30
4.5 Spectral Characterization of Forest Dieback Severity Levels	31
4.5.1 Statistical Analyses	32
4.6 Model Development and Machine Learning Framework.....	32
4.6.1 Random Forest Classifier.....	33
4.6.2 Hyperparameter Tuning and Model Optimization.....	34
4.6.3 Feature Importance Assessment.....	34
4.6.4 Accuracy Assessment.....	35
4.7 Dieback Prediction Map Development for the Doller Valley	36
5. Results.....	39
5.1 Overview of WDB Score Variations Among Plots	39
5.2 Time Series Analysis.....	40
5.3 Spectral Performance of Indices Across Forest Dieback Severity Levels	42
5.3.1 Spectral Differentiation Between Healthy and Severe Dieback Classes	42

5.3.2 Spectral Differentiation Among Healthy, Moderate, and Severe Dieback Classes	44
5.3.3 Spectral Responses Across Forest Types	46
5.4 Temporal Evolution of Forest Dieback Using Orthophotos and Spectral Indices	47
5.5 RF Dieback Detection Model	50
5.5.1 Initial Hyperparameter Optimization and Model Performance	50
5.5.2 Recursive Feature Elimination and Permutation-Based Feature Selection	51
5.6 Spatial Prediction for Forest Dieback and Temporal Evolution	53
5.6.1 Error Adjusted Area Calculations	56
6. Discussion	57
6.1 Spectral Indicators And Their Effectiveness For Dieback Detection	57
6.1.1 Temporal Behaviour of Spectral Indices.....	57
6.1.2 Spectral Separability of Dieback Classes	58
6.1.3 Influence of Understory Vegetation on Spectral Signals	59
6.2 Random Forest Model Performance	60
6.3 Random Forest Performance Improvement with RFE and Permutation Importance	62
6.4 Error Matrix and Area Prediction Correction	63
6.5 Influence of Species Composition on Spectral Response.....	64
7. Conclusion	66
References	68
Appendices.....	79
Appendix 1. Species composition map of Doller Valley	79
Appendix 2. Identification of the suitable smoothing parameter (λ) using a (a) Healthy Plot and a (b) Severe Plot	80
Appendix 3. Initial 21 Predictor features	81
Appendix 4. Histogram showing the distribution of sampled plots across dieback classes	82
Appendix 5. Time series visualizations of vegetation indices (NDWI, IRECI, CR_RE, and CR_SWIR).....	83
Appendix 6. Species composition-specific time series visualizations for the three dieback classes: (a) Conifer plots, (b) Deciduous Plots, and (c) Mixed Plots	84
Appendix 7. Temporal evolution of dieback in a (a) Healthy plot and a (b) Severely affected plot	87
Appendix 8. Annual model confidence variation	88

List of Acronyms

AI: Artificial Intelligence
ANN: Artificial Neural Network
BA: Basal Area
BM: Branch Mortality
BOA: Bottom of Atmosphere
CNN: Convolution Neural Network
CR_RE: Continuum Removed Red Edge
CR_SWIR: Continuum Removed Short Wave Infra-Red
DL: Deep Learning
DT: Decision Trees
IPCC: Intergovernmental Panel on Climate Change
IRECI: Inverted Red Edge Chlorophyll Index
kNN: k-Nearest Neighbors
ML: Machine Learning
NDVI: Normalized Difference Vegetation Index
NDWI: Normalized Difference Water Index
NIR: Near Infra-Red
RF: Random Forest
RFE: Recursive Feature Elimination
RNN: Recurrent Neural Network
SI: Spectral Indices
SMOTE: Synthetic Minority Over-sampling Technique
SVM: Support Vector Machine
SWIR: Short Wave Infra-Red
TOA: Top of Atmosphere
UAV: Unmanned Aerial Vehicles
VIS: Visible Spectrum
WDB: Weighted Dieback

1. Introduction

Forest ecosystems play a critical role in sustaining life on Earth by providing essential ecological and socio-economic services that support both environmental stability and human well-being. Their key ecological functions include biodiversity conservation, regulation of biogeochemical cycles, soil protection, and climate regulation, all of which are fundamental to maintaining ecosystem resilience. In addition to these ecological benefits, forests also support the livelihoods of millions of people worldwide by providing resources such as food, fuel, and raw materials. However, in recent decades, forest ecosystems have been increasingly affected by climate change, particularly through rising temperatures, prolonged droughts, and associated disturbances. These pressures have led to growing concerns about forest health and the accelerating occurrence of forest dieback, a gradual process of canopy degradation and tree decline. Monitoring these changes is essential for understanding ecosystem resilience and supporting sustainable forest management.

This research was conducted within the framework of the PHENET project, a collaborative initiative aimed at developing innovative digital solutions for monitoring forest and agricultural ecosystems through the integration of Earth Observation data, field measurements, and environmental sensor networks. The project focuses on understanding spatial and temporal variability in ecosystem conditions and improving the use of multi-source data for environmental monitoring applications. Within PHENET, field campaigns and in situ data collection play a central role in providing high-quality reference information for the calibration and validation of remote sensing analyses.

The study area of this thesis is the Doller Valley, located in the Vosges Mountains in northeastern France. This region is characterized by a heterogeneous forest landscape composed of both deciduous and coniferous species, pronounced elevational gradients, and strong climatic variability. In recent years, the area has experienced increasing evidence of forest stress and dieback, largely associated with drought events and long-term climatic changes. These conditions, combined with its ecological diversity and well documented forest structure, make the Doller Valley particularly suitable for studying forest decline dynamics and their spectral expression using Earth observation data. The availability of consistent field observations across multiple years further enhances its value as a study site for linking ground based measurements with satellite derived indicators of forest condition.

The field data used in this study were collected through collaboration with French research partners from INRAE and AgroParisTech within the framework of the PHENET project. These partners conducted repeated field campaigns in the study area between 2019 and 2025, providing in situ observations that serve as reference data for forest condition assessment. This information supports the calibration and interpretation of remotely sensed data used in this study.

Building on this collaboration, the main objective of this thesis is to evaluate the potential of Sentinel-2 multispectral time-series imagery for monitoring forest dieback. Specifically, the study focuses on identifying spectral indicators associated with forest decline and analyzing their spatial and temporal patterns using a machine learning based approach over the study period.

Overall, this research contributes to ongoing efforts within the PHENET project to improve the integration of field observations and Earth Observation data for environmental monitoring. By combining in situ measurements with satellite derived information, the study aims to support the development of scalable approaches for assessing forest health under increasing climatic stress and to enhance understanding of forest dieback dynamics in temperate ecosystems.

2. Literature Review

2.1 Forest Ecosystems and Drivers of Forest Degradation

Despite the substantial ecological and socio-economic values of forests, they are increasingly exposed to a wide range of degradation pressures driven by both anthropogenic and environmental factors. Historically, forest degradation has been strongly linked to direct human activities such as deforestation, agricultural expansion, urbanization, and unsustainable resource exploitation, all of which have contributed significantly to the global decline of forest ecosystems and the disruption of their ecological functioning. Mgelwa et al. (2025) reported that global deforestation has continued at an alarming rate of approximately 13 million hectares per year over recent decades, highlighting the persistent scale of forest loss worldwide. In addition, Sylvester et al. (2024) noted that nearly 80% of global forest cover change is attributable to agricultural expansion and intensification, emphasizing the dominant role of land-use conversion in driving forest degradation. However, beyond these direct pressures, broader environmental changes are increasingly intensifying forest vulnerability. Among these emerging pressures, climate change has become a major indirect driver of forest degradation. Changes in global climate systems, influenced by both natural variability and human-induced environmental alterations, have contributed to rising temperatures, shifting precipitation patterns, increased drought frequency, and more frequent extreme weather events. IPCC (2022) further highlights that an increase in global warming (1.5°C - 2°C) will trigger substantial ecosystem transformations and shifts in the geographic distribution of plant species. Moreover, in tropical regions such as Central America, global warming is projected to reduce forest biomass by up to 40%, potentially transforming rainforest ecosystems into more grassland- and savanna-like environments. These climatic changes impose substantial abiotic stress on forest ecosystems by disrupting water availability, increasing thermal stress, and altering ecological processes essential for tree growth and survival. Prolonged exposure to such climatic stress can progressively weaken tree vitality, reduce growth, and impair physiological resilience. As trees become increasingly stressed, their natural defense systems decline, making them more susceptible to secondary disturbances such as pests and pathogens. The combined effects of abiotic stress and subsequent biotic disturbances can accelerate tree mortality and intensify broader patterns of forest decline (Buras et al., 2021; Senf & Seidl, 2020).

2.2 Dynamics of Forest Dieback in a Changing Climate

As climate conditions continue to shift, tree species that lack sufficient adaptive capacity to cope with emerging environmental stresses become increasingly vulnerable, often resulting in local extinctions, shifts in species composition, and profound changes in forest structure, functioning, and resilience (Aitken et al., 2008; Allen, 2007). Within this context, global climate change has emerged as a critical driver of widespread forest decline, with forest dieback representing one of its most severe and ecologically significant consequences (Allen et al., 2010). Intensified by extreme climatic conditions such as prolonged drought, altered precipitation patterns, and the increasing prevalence of pests and pathogens under climate-induced stress, forest dieback reflects a complex process of progressive tree health deterioration that extends beyond normal mortality patterns (Moreno-Fernández et al., 2021). More specifically, forest dieback is understood as a multifactorial and progressive phenomenon involving the general decline of tree vitality and condition, driven by the interaction of abiotic stressors and subsequent biotic infestations, ultimately leading to elevated mortality rates, reduced ecosystem resilience, and large-scale degradation of forest ecosystems (Allen, 2009; Carletti et al., 2025). From an observational perspective, forest dieback is commonly characterized by progressive structural and morphological deterioration of tree canopies, including crown thinning, severe defoliation, and the accumulation of dead branches within the upper canopy layers (Figure 1) (Anwander et al., 2024; Mouret et al., 2024).

Numerous studies worldwide have documented increasing forest dieback associated with changing climatic conditions. In Belgium and northeastern France, Arthur et al. (2024) reported that extreme drought conditions triggered a bark beetle outbreak, resulting in the loss of 12.2% of the region's *Picea abies* forests between year 1 and year 2.

In Europe, analyses of Landsat data conducted by Senf et al. (2018) revealed an average annual increase of 2.4% in forest canopy mortality between 1984 and 2016, indicating a continuous intensification of forest decline across the continent. Similarly, analyses of MODIS data by Neumann et al. (2017) identified several mortality hotspots across both northern and southern Europe. Similarly, in the western United States, van Mantgem et al. (2009) observed a doubling of background tree mortality rates every 17 to 29 years, while Peng et al. (2011) documented an annual increase of 4.9% in tree mortality in Canadian boreal forests from 1963 to 2008. Allen et

al. (2010) further synthesized evidence at a continental scale, showing that extreme heat events and multi-year droughts have substantially increased mortality across diverse forest species and regions.

Rather than occurring as a uniform process, forest dieback functions as a selective ecological pressure that differentially affects species, functional groups, and forest types according to their physiological traits and adaptive strategies (Anderegg et al., 2013). This selective mortality can rapidly reshape forest stand structure, altering size distributions, age classes, and species composition, while redirecting successional pathways (Van Mantgem et al., 2009). For example, severe drought in Queensland, Australia, disproportionately increased mortality among large-stemmed individuals of the *Eucalyptus crebra*–*Eucalyptus xanthoclada* complex, significantly modifying stand structure, whereas co-occurring species showed no comparable response. Similarly, in northern Patagonia, drought-induced canopy gaps promoted the preferential recruitment of *Austrocedrus chilensis* over *Nothofagus dombeyi*, thereby influencing future canopy dominance and long-term forest composition (Suarez & Kitzberger, 2008). Beyond overstory impacts, drought-related dieback can also trigger substantial shifts in understory communities, as demonstrated by persistent compositional changes observed seven years after a major mortality event in the southwestern United States (Kane et al., 2011). Similarly, in the Vosges Mountain Range in France, *Abies alba* and *Picea abies* displayed markedly different dieback rates despite inhabiting the same environmental setting, highlighting substantial interspecific variation in drought susceptibility (Anderegg et al., 2013; Piedallu et al., 2023).

Unlike direct human-driven degradation, which is often immediate and visibly apparent through deforestation or land conversion, forest dieback can unfold gradually and often silently, progressively undermining forest health before visible signs of decline fully emerge. This subtle yet pervasive form of degradation presents an increasingly serious challenge, as its impacts may remain undetected until substantial ecological damage has already occurred. As climate change continues to intensify environmental stressors worldwide, understanding these hidden processes of forest decline has become essential for safeguarding forest ecosystems and developing effective long-term monitoring strategies.

Defoliation	Category	Discoloration
	healthy	
	light damage	
	medium damage	
	heavy damage	

*Figure 1. Representation of different levels of forest dieback with their morphological characteristics.
Source: Anwander et al. (2024)*

2.3 Remote Sensing in Forest Dieback Detection

Early forest monitoring was predominantly based on manual field assessments supported by local ecological knowledge. For decades, forestry professionals and communities jointly undertook observations of forest conditions, often through community based monitoring systems that formally engage local stakeholders in data collection and resource surveillance (Evans &

Gauriguata, 2008). Despite their value, traditional field based methods exhibit substantial limitations. They are time-consuming, labor intensive, and costly to implement over large spatial extents, and the resulting datasets are often sparse, inconsistent, or temporally irregular (Dainelli et al., 2021). These constraints prompted a significant methodological shift toward remote sensing based forest monitoring. In forestry, remote sensing has become an indispensable tool supporting a broad spectrum of applications, including the mapping of forest types, the assessment of vegetation productivity at regional to global scales, and the generation of detailed three-dimensional representations of forest structures (Fassnacht et al., 2024). The availability of remote sensing imagery across various temporal (periodic and seasonal) and spatial scales, like maps and data products representing forest types or tree species information that are derived from time series of freely available passive optical satellite data, coupled with its unique spectral characteristics and integration with advanced geospatial techniques, provides significant opportunities to derive key forest attributes, which are essential for effective and sustainable forest management (Grabska et al., 2020; Hermosilla et al., 2022).

Effective understanding of forest dieback requires continuous and systematic observation of stressed and declining trees throughout the full progression of stress and mortality (Buras et al., 2021; Scharnweber et al., 2020). Traditionally, assessments have relied on ground based inventories that classify dieback severity according to visible indicators such as foliage loss, increased canopy transparency, and the death of upper and peripheral branches (Pontius & Hallett, 2014). Traditional assessments often failed to accurately reflect dieback patterns because mortality can occur in highly localized patches that are easily overlooked during field surveys, and dead trees may be harvested or removed before sampling takes place, resulting in incomplete or biased observations (Carletti et al., 2025). In addition, certain decline patterns can be sparse or spatially diffused, making them difficult to detect through manual observations alone, particularly across large or inaccessible forest areas. These limitations underscore the growing importance of remote sensing technologies for dieback assessment. Spaceborne remote sensing offers the spatial and temporal coverage necessary to monitor forest conditions over broad regions and to observe changes as they evolve over time. By capturing spectral signals associated with physiological stress, such as reductions in pigment content, canopy thinning, or structural degradation, remote sensing enables the detection and quantification of decline symptoms that may not be visible from ground surveys (Anderegg et al., 2013, 2015; Campbell et al., 2020). Moreover, because forest

dieback is often selective, timely identification of decline patterns is crucial for guiding targeted management interventions and mitigating the impacts of biotic and abiotic stressors (Iglhaut et al., 2019; Raparelli & Bajocco, 2019). Despite these advantages, remote sensing based dieback monitoring also presents several important limitations, particularly in advanced applications such as dieback modelling and early stress detection. Mouret et al. (2024) noted that satellite observations alone often lack the intrinsic physical context required to fully distinguish true canopy decline from normal phenological variability or seasonal spectral fluctuations. This challenge becomes especially pronounced in complex, heterogeneous, or mixed species forest canopies, where individual tree level decline signals may be diluted within medium resolution satellite pixels. As a result, different dieback conditions can remain difficult to isolate solely from satellite observations. In such contexts, ground based field data remain essential for calibration and validation purposes, particularly for distinguishing true canopy decline from background spectral variability. High to very-high resolution imagery (e.g., UAV data) can also substantially improve the detection of individual tree level dieback signals and support the validation of regional scale classifications. However, despite their advantages, these data sources are often difficult to acquire consistently over large spatial extents due to financial, logistical, and accessibility constraints (Senf et al., 2018)

When considering the temporal evolution of dieback, one of the first symptoms in affected trees is leaf discoloration caused by reduced chlorophyll content, followed by canopy thinning, defoliation, and browning due to shoot and leaf mortality at later stages. These changes serve as important indicators of stress leading to dieback and can be observed through spectral responses captured by optical remote sensing systems (Anwander et al., 2024; Moreno-Fernández et al., 2021). Spectral indices (SIs) derived from remote sensing data provide a means to quantify these changes and relate dieback to phenological and physiological metrics (Lausch et al., 2017). Two groups of indices are particularly relevant for dieback assessment: vegetation indices, which estimate absorbed photosynthetically active radiation, and wetness indices, which focus on vegetation water content (Marusig et al., 2020). According to Anderegg et al. (2019), analyses of SI trends can be used to identify drought-prone or otherwise vulnerable areas before extensive dieback and mortality occur. Within this context, previous studies have also highlighted that the sensitivity of spectral indices is not uniform across vegetation stress types or species. For instance, Abdullah et al. (2019) showed that red-edge and green-based indices can be particularly effective

for early stress detection, as physiological degradation is often reflected in these spectral regions before structural canopy changes become apparent. Similarly, Ye et al. (2021) demonstrated that the optimal choice of spectral indices is strongly species dependent, with some conifer systems responding more clearly to red-edge based metrics, while others show stronger sensitivity to shortwave infrared based moisture indicators.

Additionally, species specific variations in spectral reflectance under stress indicate that selecting indices tailored to individual species can improve dieback detection performance. For example, forest responses to water stress differ among tree functional types: broadleaf species can draw on water stored in their heartwood during prolonged dry periods, whereas conifers depend primarily on real-time water absorption. As a result, drought leads to declines in photosynthetic activity and canopy greenness, making wetness indices particularly suitable for detecting early-stage dieback in coniferous species (Liu et al., 2021; Querejeta et al., 2007).

2.4 Potential of Multispectral Remote Sensing for Forest Dieback Detection

Multispectral remote sensing is currently the most widely used approach for monitoring and assessing forest health. Sensors such as MODIS, Landsat, Sentinel, Worldview, and RapidEye are commonly applied due to their demonstrated capability and accessibility (Torres et al., 2021). Multispectral sensors capture reflectance across key regions of the electromagnetic spectrum, where vegetation exhibits distinct spectral behaviors. In the visible region (VIS) (0.4–0.7 μm), reflectance is primarily influenced by pigment concentrations, which absorb solar radiation strongly, resulting in relatively low reflectance values. In the near-infrared (NIR) (0.7–1.2 μm), higher reflectance occurs due to leaf cellular structure, which enhances the transmittance and reflectance, further amplified by multiple scattering within the canopy. The red-edge region, located at the boundary between the VIS and NIR bands, is particularly sensitive to changes in chlorophyll and nitrogen content, reflecting subtle shifts in leaf physiology. Finally, the mid-infrared domain is divided into shortwave infrared (SWIR) (1.2–2.5 μm) and mid-infrared (2.5–8 μm), with SWIR providing valuable information on vegetation water content. The reflectance in these spectral regions can vary depending on the stage of dieback and, in some cases, species-specific characteristics (Clevers & Gitelson, 2013; Zarco-Tejada et al., 2018).

Among the spectral regions most widely utilized in forest dieback research, the red-edge, near-infrared, and shortwave infrared bands have emerged as particularly important due to their strong sensitivity to subtle physiological and structural changes in forest canopies. In particular, growing evidence suggests that canopy reflectance within narrow spectral regions, especially the red-edge, provides valuable diagnostic cues for the early detection of vegetation stress and decline processes (Amoah-Nuamah et al., 2025). However, the direct use of raw reflectance is often limited by environmental noise, including topographic, illumination, and atmospheric effects (Matsushita et al., 2007). Spectral indices mitigate these influences and improve the detection of subtle physiological changes associated with early forest decline (Bárta et al., 2021; Carletti et al., 2025). Building on this, several studies have successfully applied Sentinel-2 derived indices to monitor diverse dieback processes. For example, Andresini et al. (2024) applied multispectral indices to detect bark beetle-induced dieback in north-eastern France following the 2018 outbreak, while Migas-Mazur et al. (2021) used similar approaches to map beetle-related forest changes in the Tatra Mountains (Poland and Slovakia) between 2015 and 2018. Abdullah et al. (2019) demonstrated that red-edge and shortwave infrared multispectral information can enable early detection of beetle-induced stress in south-eastern Germany during 2015, while Marusig et al. (2020) employed Sentinel-2 data in 2019 to assess vegetation water stress as an indicator of drought related dieback vulnerability in north-eastern Italy. Collectively, these studies highlight both the versatility of multispectral remote sensing across multiple disturbance contexts and the distinct importance of specific spectral regions in capturing the complex dynamics of forest decline.

2.5 Artificial Intelligence (AI) for Forest Health Monitoring

With advances in remote sensing and computational capabilities, forest health monitoring has shifted from traditional analytical methods toward more scalable, data driven approaches for detecting physiological stress, pest outbreaks, and disease impacts across large spatial scales. This transformation is largely driven by artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), which enable the extraction of complex ecological patterns from multispectral and structural datasets. Following data pre-processing and feature extraction, these algorithms classify forest conditions by analyzing spectral, spatial, and ecological characteristics to differentiate varying levels of forest health.

In forestry research, widely used ML approaches such as Random Forest (RF), Support Vector Machine (SVM), Decision Trees (DT), k-Nearest Neighbours (kNN), and Extreme Gradient Boosting (XGBoost) have become essential for detecting disturbances, monitoring pest and disease dynamics, and assessing forest degradation (Amoah-Nuamah et al., 2025). For instance, Khan et al. (2025) compared DT, SVM, and RF models for classifying forest health status in the Western Himalayas into healthy, moderate, and unhealthy categories using ecological and structural data, achieving accuracies of 66%, 75%, and 83%, respectively. Similarly, Hawryło et al. (2018) applied kNN, RF, and SVM algorithms to detect defoliation caused by *Dendrolimus pini* L. infestations and classify forest condition based on defoliation severity, reporting accuracies of 78% for both SVM and kNN, and 75% for RF. Collectively, these studies highlight the growing role of AI-based methods in enhancing the accuracy, efficiency, and practical application of forest health monitoring in the face of increasing environmental challenges.

Moreover, DL has advanced forest health monitoring by enabling the automated extraction of complex spatial, spectral, and temporal patterns that are often difficult for traditional machine learning approaches to capture (Amoah-Nuamah et al., 2025). Unlike classical algorithms, DL architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and multimodal networks learn hierarchical, multi-scale representations directly from remote sensing data, improving the detection of subtle, early stage forest disturbances with greater precision. For example, Chiang et al. (2020) developed a CNN-based forest health monitoring system using aerial imagery to detect dead trees, achieving an overall accuracy of 80%. Similarly, Ma et al. (2024) applied a multi-step CNN framework with UAV derived hyperspectral and LiDAR data to classify tree species in complex forests, reporting an accuracy of 83%. Beyond standalone models, hybrid ML–DL frameworks have also shown strong potential. For instance, Hou et al. (2025) combined UAV-derived indices with CNN-extracted features to improve early detection of pine wilt disease, achieving approximately 92% accuracy and outperforming single-model approaches. These hybrid approaches integrate DL-based feature extraction with ML-based classification, offering more efficient and adaptable solutions, particularly in data limited forest monitoring contexts.

Despite the substantial progress achieved through AI-based approaches, the widespread operational implementation of these models in forest health monitoring remains constrained by

several persistent challenges. These limitations extend beyond purely computational or algorithmic concerns and are deeply rooted in the ecological complexity and heterogeneity of forest ecosystems. A major obstacle is the limited availability of high quality, labeled training datasets, which are often difficult and costly to obtain across diverse forest environments. Additionally, remote sensing data are frequently affected by environmental noise, including cloud contamination, seasonal phenological variation, and sensor specific inconsistencies, all of which can reduce model reliability and generalizability (Amoah-Nuamah et al., 2025).

2.5.1 Machine Learning Approaches in Forest Dieback Detection

ML approaches remain highly relevant, as they can produce robust and reliable results even with relatively small or region specific datasets, making them particularly suitable for targeted applications such as forest dieback assessment (Amoah-Nuamah et al., 2025). The study conducted by Candotti et al. (2022) investigated spruce bark beetle induced dieback in the context of climate change, focusing on Norway spruce stands across Central and Southern Europe, particularly in countries such as Italy, Austria, and Slovenia. Using Sentinel-2 time-series imagery, the study applied supervised machine learning approaches, including RF, SVM, and KNN, to classify forest conditions based on both individual spectral bands and a range of vegetation indices. The results demonstrated that these models achieved a mean overall accuracy of 89%, effectively distinguishing between healthy and stressed canopy conditions, and highlighting the potential of machine learning for reliable dieback detection at regional scales.

Among the various machine learning approaches applied in forest dieback studies, RF is consistently favored due to its strong performance in both classification and regression tasks, relatively high predictive accuracy, and its ability to provide variable importance measures, which enhance the interpretability of results compared to more opaque models such as ANNs (Torres et al., 2021). Kautz et al. (2024) further emphasizes its widespread use, reporting that RF accounts for a substantial proportion of studies (38% of all studies reviewed) investigating pest induced dieback detection. Its popularity is largely attributed to its robustness in handling high dimensional spectral datasets and its capacity to generate feature importance rankings that help identify wavelengths most sensitive to beetle induced stress. However, despite these advantages, the performance reported across studies remains variable. Kautz et al. (2024) note that, on average, RF based approaches achieved a mean user accuracy (Precision) of 59% and a mean producer

accuracy (Recall) of 63%, values that fall below the ~80% accuracy threshold, often considered necessary for operational deployment in forestry applications. Moreover, RF performance has been shown to depend strongly on the type and quality of input sensor data, indicating that its effectiveness is not uniform across different remote sensing configurations and study contexts.

Moreover, Andresini et al. (2024) evaluated Sentinel-2 time-series data for mapping bark beetle-induced forest dieback in north-eastern France, comparing machine learning methods (RF and XGBoost) with a deep learning based U-Net model. Their results showed that RF achieved the best performance when using the full temporal dataset, reaching the highest accuracy with an F-score of 0.66 for damaged areas, while U-Net performed better on single date imagery but declined when applied to the full time series. Although the study focused on post event mapping rather than early detection, it highlights the importance of using methods capable of handling high dimensional, temporally correlated data. In this context, machine learning approaches remain particularly practical and robust, as they can effectively manage complex feature spaces without requiring large training datasets or highly complex model architectures.

2.6 Challenges in Remote Sensing and ML-based Forest Dieback Detection

Despite major advances in remote sensing technologies, detecting forest dieback remains difficult, especially at its early stages. Camarero et al. (2015) note that certain disturbances can push ecological systems toward abrupt transitions between alternative states. However, identifying when trees first begin to experience biotic and abiotic stress, and when mortality actually starts, often requires long periods of observation. By the time symptoms become clearly visible in remote sensing data, the opportunity for early intervention may already be lost. This limitation is clearly reflected in most dieback studies, where classifications are often simplified into a small number of discrete classes, typically contrasting healthy forests with a single degraded state (Abdullah et al., 2019; Allen et al., 2010). As a result, the continuous nature of forest decline is not fully represented in many analytical frameworks. Consequently, capturing the full transition gradient of dieback, particularly the early warning stages marked by subtle symptoms such as moderate crown defoliation and dead branches, remains both essential and frequently overlooked (Trumbore et al., 2015). This challenge is further compounded by limitations in the spatial and temporal resolution of available remote sensing data, which restrict the consistent observation of these intermediate stages and ultimately create important gaps in early detection capabilities. Kautz et al. (2024)

emphasize that, based on analyses conducted on European spruce forests affected by bark beetle–induced dieback, no study has yet achieved the ~80% accuracy threshold required for operational forestry use. According to Amoah-Nuamah et al. (2025), this limitation is primarily driven by constraints such as the scarcity of labeled datasets, noise introduced by cloud cover and phenological variability, and difficulties in transferring models across different species, ecosystems, and sensor platforms. Kautz et al. (2024) further provides a synthesis of the potential and limitations of different sensor platforms for dieback detection, highlighting the variability in performance across approaches. In addition, they note that the lack of large training datasets (often fewer than 500 severely impacted trees) and the inability of current models to distinguish beetle-specific signatures from general environmental stressors such as drought significantly hinder detection performance.

Another key limitation is spectral overlap, where current machine learning models struggle to separate dieback signals because changes in chlorophyll and water content are often very similar at early stages. In pest-induced dieback, especially during the initial attack phase and depending on tree type, this separation becomes even more difficult. Asner et al. (2015) further highlight that spectral overlap among broadleaf species reduces class separability, particularly for disease and drought detection. In addition to spectral constraints, spatial resolution presents a fundamental limitation for dieback detection. Senf et al. (2018) note that early stage dieback often occurs at very fine spatial scales, such as individual trees or small patch level disturbances, which are easily obscured in medium resolution satellite imagery due to sub-pixel mixing with surrounding healthy vegetation. Similarly, Fassnacht et al. (2024) emphasize that spatial resolution constraints directly affect machine learning performance, as mixed pixels containing combinations of healthy trees, stressed vegetation, and background soil produce diluted spectral signals. This mixing reduces class separability and makes it difficult for classifiers to reliably distinguish early transitional stress stages.

Collectively, the literature highlights that while remote sensing and machine learning have significantly advanced forest dieback detection, key challenges remain, particularly in early stage identification, sub-pixel mixing, and data limitations. These constraints continue to affect the operational reliability of large scale monitoring models. As a result, before broader ecological generalization or model transferability can be achieved across forest ecosystems, robust site

specific calibration against detailed field inventories remains essential. Addressing this, the present study evaluates the integration of Sentinel-2 multispectral remote sensing data and machine learning approaches in a temperate forest ecosystem in the Doller Valley, France. By linking high dimensional spectral information with ground truth observations, the research aims to establish a reliable local framework for mapping the progressive dynamics of canopy decline under changing environmental conditions.

3. Study Objectives

The objectives of this study are designed to evaluate the effectiveness of Sentinel-2 multispectral time series imagery in detecting and monitoring annual forest dieback associated with climate-induced abiotic stress in the Doller Valley, Vosges, France.

The first objective is to identify the spectral properties of Sentinel-2 imagery that are most indicative of forest dieback. Vegetation affected by abiotic stress exhibits changes in spectral reflectance due to alterations in canopy structure, chlorophyll content, and vegetation moisture. By analyzing the spectral response of Sentinel-2 bands and derived vegetation indices, this study aims to determine the most reliable indicators for distinguishing healthy forest areas from those experiencing decline.

The second objective is to develop a spatial model for detecting forest dieback using Sentinel-2 multispectral imagery. The model aims to identify and map spatial patterns of forest decline by analyzing temporal changes in vegetation conditions throughout the study period.

The final objective is to predict the annual rate of forest dieback within the study area during the study period. Temporal analysis of Sentinel-2 imagery enables the assessment of yearly changes in forest condition and the identification of trends in dieback progression. Quantifying the annual rate of forest decline will contribute to a better understanding of the temporal dynamics of climate-related forest stress and support future forest monitoring and management strategies in the Vosges region.

4. Methodology

This chapter describes the methodology used to investigate forest dieback in the Doller Valley. It begins with an overview of the study area and the data sets used, followed by the preprocessing of field and satellite data. The chapter then presents the computation of vegetation indices, time-series extraction, and the characterization of dieback severity. Finally, the machine learning framework, model evaluation procedures, and the development of forest dieback prediction maps are described.

4.1 Study Area

The study was conducted in the Doller Valley, located in the southern part of the Vosges Mountain range in northeastern France. Covering approximately 123 km², the valley is predominantly forested, with more than 70% of its total area occupied by forests (Piedallu et al., 2023). According to Copernicus forest cover layers, these forests are composed of approximately 60% deciduous species and 40% coniferous species, reflecting the ecological diversity of the region.

The forests of the Vosges Mountains are characterized by a range of productive tree species, among which European beech is the most widespread, occurring in 75% of the plots, followed by silver fir (59%), Norway spruce (53%), sessile oak (40%), and Scots pine (25%). Overall, softwood species constitute around 70% of the forest stands, while silver fir and Norway spruce dominate approximately 40% of these stands (Piedallu et al., 2023) (Figure 2).

The Doller Valley also exhibits considerable topographic variation, with elevations ranging from 200 m to 1,300 m above sea level. This elevational gradient is accompanied by distinct climatic variability, as mean annual precipitation is approximately 1,000 mm, while mean annual temperatures range from 2.5°C to 15.0°C, with an overall average of 8.7°C. Together, the valley's ecological diversity, varied climatic conditions, and heterogeneous forest composition make it a highly suitable study area for monitoring forest dieback and assessing ecosystem responses within temperate mountainous environments.

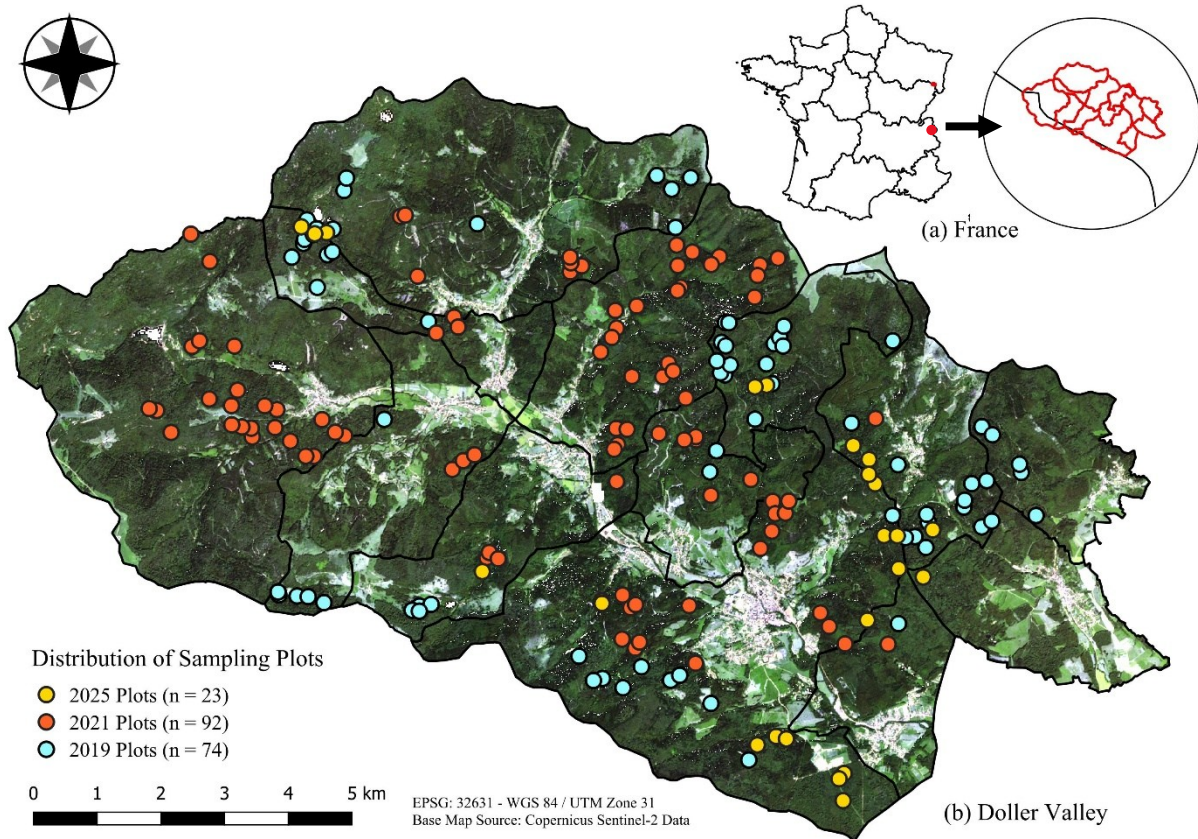


Figure 2. (a) Location of Doller Valley in France, (b) Doller Valley, France, with the distribution of field sampling plots; Basemap source: Copernicus Sentinel-2

4.2 Data Description

4.2.1 Field Data

Field data for the study area were provided through collaboration with the French National Research Institute for Agriculture, Food and Environment (Institut national de recherche pour l'agriculture, l'alimentation et l'environnement – INRAE), AgroParisTech, within the frame of the [PHENET](#) project. The PHENET project (Work Package 3) focuses on developing digital solutions for assessing spatial and temporal variability in forest and agricultural environments through the integration of Earth Observation (EO) technologies, UAV systems, and field based sensors.

Field campaigns were conducted during three separate periods: November–December 2019, November–December 2021, and May 2025. Across these campaigns, data were collected from

multiple locations distributed throughout the entire Doller Valley, ensuring broad spatial coverage of the study area. Data collection was carried out at both plot and tree levels. At the plot level, geographic coordinates and the total number of trees within each plot were recorded, while detailed physiological measurements were collected for individual trees. A concentric plot sampling approach was employed to ensure representative sampling across varying tree size classes. In this method, concentric circular zones were established around a plot center, with each ring assigned a specific radius corresponding to particular tree size categories. This design accounts for the natural variability in tree size distribution within forest stands and ensures that trees from different size classes are systematically and proportionally represented in the dataset. The concentric sampling design is illustrated in Figure 3. At the tree level, physiological measurements included circumference, branch mortality status, tree vitality condition, and sampling date. Tree vitality was recorded using a binary classification system, where dead trees were assigned a value of 1 and living trees a value of 0. Branch mortality (BM) was assessed based on the proportion of dead branches observed in each tree crown, and each tree was subsequently classified into one of six branch mortality classes according to the severity of branch dieback, as detailed in Table 1. In total, 74 plots were sampled in 2019, 92 plots in 2021, and 23 plots in 2025. The complete dataset consists of 189 plots and 3,366 sampled trees across all survey years. Sampling was restricted to trees with a circumference greater than 20 cm, ensuring the inclusion of sufficiently developed individuals for physiological assessment.

In addition to the field dataset, a species type map of the Doller Valley was also provided by the same collaborators (Appendix 1). This map classifies the forested area into seven distinct classes based on species composition and proportional dominance within each forest polygon. The classification distinguishes between conifer stands ($\geq 95\%$ of conifer species), mixed conifer dominant stands ($\geq 50\%$ dominance of conifer), mixed stands ($< 50\%$ conifer), and deciduous stands ($< 5\%$ conifer). This spatial dataset was used to characterize stand composition and canopy dominance across the study area, providing important contextual information for understanding species distribution patterns and supporting forest dieback analysis.

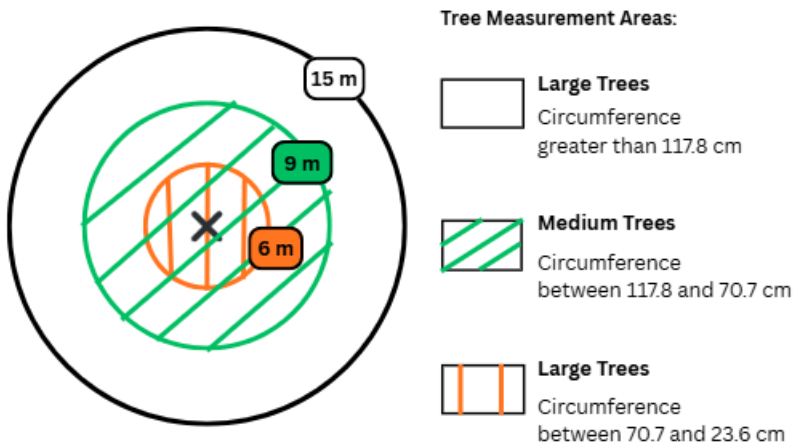


Figure 3. Representation of the concentric circular plot design used for tree sampling. The sampling area is divided into three sub-plots based on tree circumference: 15 m radius for large trees (>117.8 cm), 9 m for medium trees (70.7–117.8 cm), and 6 m for small trees (23.6–70.7 cm).

Table 1. Different levels of branch mortalities present in the field data

Class	Level of Branch Mortality
0	Non to very low
1	Low
2	Medium
3	Strong
4	Very strong
5	Full

Pretreatment of Field Data

The field dataset was first processed to distinguish dominant and suppressed trees within each plot. This classification was based on the ratio between an individual tree's circumference and the mean circumference of all trees within the same plot, following the approach described by Carletti et al. (2025). This step was necessary because, in the context of satellite based forest monitoring, canopy level signals are primarily driven by dominant trees, as suppressed trees typically do not reach the upper canopy and therefore have limited influence on Sentinel-2 reflectance.

Following this classification, a plot level weighted dieback (WDB) score was computed to quantify forest dieback at the plot scale. This metric integrates individual BM levels of dominant trees, weighted by their basal area (BA) relative to the total basal area of the plot (Eq. 1 & Eq. 2). The basal area was derived from tree circumference measurements, ensuring that larger trees exert a greater influence on the final plot-level dieback estimate compared to smaller individuals. Consequently, this approach produces a robust dieback severity indicator for each plot, reflecting both tree size structure and canopy condition in a single integrated metric. The formulation of this WDB indicator was inspired by the plot-level dieback assessment approach proposed by Carletti et al. (2025).

$$BA = (C^2)/(4\pi) \quad [1]$$

C = Circumference of the tree

$$\text{Weighted Dieback Severity} = \frac{\sum_{i=1}^n BM_i \cdot BA_i}{\sum_{i=1}^n BA_i} \quad [2]$$

4.2.2 Satellite Data

Sentinel-2

Multispectral satellite imagery for this study was acquired from the Sentinel-2 mission of the Copernicus Programme in orthorectified Top of Atmosphere (TOA) reflectance format (Level-1C). The Sentinel-2 tile covering the study area (UTM Zone 31 UGP) in northeastern France was selected for analysis for the period August 2016 to December 2025.

Of the 13 spectral bands available from Sentinel-2, 10 bands were utilized for this analysis. These included four bands at 10 m spatial resolution (B2, B3, B4, and B8) and six bands at 20 m spatial resolution (B5, B6, B7, B8A, B11, and B12). Together, these bands cover the VIS, NIR, and SWIR regions of the electromagnetic spectrum. To minimize atmospheric interference and ensure data quality, images with cloud cover exceeding 80% were excluded from the dataset.

Pretreatment of Sentinel-2 Data

The downloaded Sentinel-2 Level-1C (L1C) images were first processed using the MAJA (MACCS-ATCOR Joint Algorithm) processor, a tool developed for atmospheric correction and cloud detection (MAJA – The Multi-Temporal Cloud Screening and Atmospheric Correction Software v.4.5.4, 2022). This processing step converts TOA reflectance into Bottom of Atmosphere (BOA) reflectance, resulting in Sentinel-2 Level-2A (L2A) products suitable for quantitative analysis.

Following atmospheric correction, the Function of Mask (Fmask) was applied to identify and remove clouds and their associated shadows from the imagery. Scenes with missing or incomplete Fmask information were excluded from further processing to ensure data reliability. Subsequently, the 20 m spatial resolution bands were resampled to 10 m resolution (nearest neighbor) in order to generate a consistent multispectral dataset across all selected bands.

To further improve temporal consistency and reduce noise in the reflectance time series, a Whittaker smoothing and gap filling approach was applied (Whittaker, 1927). This method was used to address temporal gaps between Sentinel-2 acquisitions by generating 10-day mean composites, which were then smoothed based on the number of valid observations within each interval. A smoothing parameter (λ) was iteratively tested, and $\lambda = 100$ was identified as optimal for effectively reducing outliers while preserving the spectral signal characteristics (Appendix 2) (Bos et al., 2025). This procedure ultimately resulted in a continuous and temporally regular stack of Sentinel-2 bands at a 10-day temporal resolution, suitable for subsequent analysis.

Orthophotos

In addition to Sentinel-2 imagery, georeferenced high-resolution orthophotos in RGB format with a spatial resolution of 20 cm were obtained from the National Institute of Geographic and Forest Information (IGN), France. These orthophotos were primarily used for plot-level visualization, allowing detailed examination of tree appearance, surrounding forest structure, and the temporal evolution of canopy condition and dieback in relation to field sampling locations. Orthophotos from 2018, 2021, and 2024 were selected to correspond as closely as possible with the field campaign years, although imagery from the exact sampling years was not consistently available for the entire study area.

All pretreatment procedures, including data cleaning, filtering, and preparation, were conducted using Python 3.11.10 (Python Software Foundation, 2023).

4.3 Computation of Vegetation Indices

Using the pretreated Sentinel-2 imagery, several vegetation and water-related spectral indices were derived: Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Inverted Red Edge Chlorophyll Index (IRECI), Continuum Removed Red Edge (CR_RE) index, and Continuum Removed Short Wave Infrared (CR_SWIR) index (Table 02). The selection of these indices is supported by findings from previous forest dieback detection studies, which have consistently demonstrated their strong sensitivity to vegetation stress, canopy degradation, and moisture-related changes associated with dieback processes (Abdullah et al., 2019; Carletti et al., 2025)

NDVI is a widely used spectral index derived from reflectance measurements in the red (600–700 nm) and NIR (700–1300 nm) regions of the electromagnetic spectrum. It is commonly used as an indicator of vegetation greenness, biomass, structural conditions, and phenological dynamics by exploiting the contrast between strong chlorophyll absorption in the red region and high reflectance in the NIR region from healthy vegetation canopies (Huete, 2012). NDVI values generally range from -1 to $+1$, where higher positive values indicate dense and healthy vegetation cover, while lower values correspond to sparse vegetation or non-vegetated surfaces.

NDWI is designed to assess vegetation water content and plant water stress conditions using reflectance measurements from the NIR (700–1300 nm) and SWIR (1300–2100 nm) spectral regions (Gao, 1996). In this formulation, the NIR band is sensitive to internal leaf structure, whereas SWIR reflectance responds strongly to variations in vegetation moisture content. Consequently, NDWI is widely used as an indicator of vegetation water status, with higher values generally representing healthier and well-hydrated vegetation (Ceccato et al., 2002).

IRECI is a red-edge based spectral index developed to estimate canopy chlorophyll content and vegetation physiological condition using Sentinel-2 imagery. The index integrates reflectance information from the red, red-edge (705 nm and 740 nm), and near-infrared spectral regions, making it particularly sensitive to variations in chlorophyll concentration and canopy condition. By relying strongly on red-edge responses, IRECI reduces saturation effects commonly observed

in dense vegetation canopies and improves sensitivity to physiological variability within forest stands (Frampton et al., 2013).

The Continuum Removed (CR) indices, namely CR Red Edge (CR_RE) and CR Short Wave Infrared (CR_SWIR), are derived using the continuum removal technique (Clark & Roush, 1984), which is commonly applied to hyperspectral and multispectral reflectance data to enhance the interpretation of specific absorption features. According to Clark & Roush, (1984) in vegetation spectra, the observed reflectance signal is a compound response influenced both by localized biochemical absorption and by broader background effects such as canopy architecture, illumination geometry, shadowing, and soil reflectance. To isolate the absorption signal, continuum removal applies a geometric normalization procedure in which a convex hull is first constructed by connecting the local reflectance maxima across a selected spectral region. This hull acts as a reference baseline representing the hypothetical reflectance curve in the absence of absorption features. The observed reflectance at each wavelength is then divided by the corresponding continuum baseline value. This transformation standardizes the reflectance spectrum by scaling the continuum baseline to 1.0 and preserving absorption features as localized troughs below the baseline. Consequently, continuum removal reduces the influence of overall brightness and structural variability while enhancing the depth and shape of absorption features associated with vegetation biochemical properties.

CR_RE index is primarily sensitive to chlorophyll content and represents the depth of chlorophyll absorption within vegetation canopies. It exploits the red-edge region, which encompasses the strong chlorophyll absorption feature in the visible red domain and the subsequent transition toward the near-infrared region. By applying continuum removal in this region, the CR_RE index isolates the chlorophyll absorption “pit” in the red wavelengths while reducing the influence of canopy structural effects associated with the red-edge slope, thereby enhancing sensitivity to variations in chlorophyll content. In contrast, the CR_SWIR index is sensitive to vegetation water content and targets absorption features associated with liquid water in plant tissues. By focusing on the SWIR region, this index isolates water-related absorption signals and reduces the confounding effects of overall canopy structure. As a result, CR_SWIR enables the detection of internal physiological water stress even in cases where overall canopy greenness or leaf area has not yet shown visible decline. Together, these CR indices provide complementary information on

key physiological processes, improving the ability to detect early stage vegetation stress and forest dieback dynamics (Verrelst et al., 2015).

Table 2. Used Vegetation Indices for the Analysis (λ indicates the wavelengths)

Vegetation Index	Equation
NDVI	$\frac{B8 - B4}{B8 + B4}$
NDWI	$\frac{B8 - B11}{B8 + B11}$
IRECI	$\frac{B7 - B4}{\frac{B5}{B6}}$
CR_RE	$\frac{B5}{B4 + (\lambda B5 - \lambda B4) \times \frac{B6 - B4}{\lambda B6 - \lambda B4}}$
CR_SWIR	$\frac{B11}{B8A + (\lambda B11 - \lambda B8A) \times \frac{B12 - B8A}{\lambda B12 - \lambda B8A}}$

4.4 Time Series Extraction

Time series were extracted for each calculated spectral index over the period 2016–2025 using the processed Sentinel-2 index layers. By integrating data from all sampled plots across the Doller Valley, these temporal profiles enabled the assessment of overall forest condition, seasonal patterns, and long-term variability within the study area.

To derive representative plot level signals, pixel values were extracted using plot center coordinates with a 10 m buffer radius. The buffer was intersected with the raster grid, and pixels were included based on a centroid-based selection rule, whereby only pixels whose centroids fell within the buffer polygon were extracted. This method was employed to ensure that the extracted spectral information accurately reflected the monitored plots while minimizing the influence of surrounding land cover. The selected buffer size was consistent with the field sampling design, in which tree measurements were conducted within concentric rings extending up to 15 m from the

plot center, thereby ensuring that the satellite derived signals correspond closely to the dominant canopy conditions observed in the field.

4.5 Spectral Characterization of Forest Dieback Severity Levels

Following the development of overall time series profiles, additional analyses were conducted to identify spectral signal differences associated with varying levels of forest dieback. For each spectral index, separate time series plots were generated to examine how spectral responses differed across plot level dieback severity classes. Plot level index values were calculated as the mean of all extracted pixels within each plot buffer, ensuring that each time series represented the average spectral condition of the sampled plot.

Initially, multiple classification schemes with a greater number of dieback categories (e.g., 4 to 5 classes) were explored through an iterative trial-and-error process to determine an optimal threshold structure for distinguishing spectral responses. However, as class intervals became narrower, spectral separability between adjacent dieback levels decreased substantially, reducing the clarity and interpretability of class-specific temporal patterns. Consequently, a simplified 3-classes framework was adopted, with an initial 2-classes analysis developed to assess the contrast between healthy and severely affected plots over the full study period. Plot classification was based on WDB scores, where plots with $WDB \leq 2$ were classified as healthy and plots with $WDB \geq 4$ as severely affected. A 3-classes scheme was then implemented by introducing a moderate dieback category, defined as WDB values between 2 and 4. The final threshold values were determined through visual interpretation of high-resolution orthophotos and iterative evaluation of spectral class separability, with the objective of maximizing observable differences between classes while maintaining consistency with branch mortality classifications derived from field observations.

In addition, the species distribution map was used to further investigate spectral variability across different forest types. This enabled the assessment of spectral signal differences among coniferous, deciduous, and mixed forest stands within the healthy, moderate, and severe dieback classes, providing additional insight into how forest composition influences spectral responses under varying levels of canopy decline.

4.5.1 Statistical Analyses

Statistical analyses were conducted to evaluate the spectral differentiation between forest dieback severity classes and to determine whether observed differences among classes were statistically significant. Initially, the Shapiro–Wilk test was applied to assess the normality of the spectral index data (Shapiro & Wilk, 1965). As the data did not consistently meet normal distribution assumptions, non-parametric statistical methods were employed for subsequent analyses. For the two-class classification (healthy vs. severe dieback plots), the Mann–Whitney U test was used to determine whether significant differences existed between these two independent groups (Mann & Whitney, 1947). This test provided a robust assessment of spectral separability between healthy and severely affected forest conditions. For the 3-classes classification (healthy, moderate, and severe dieback plots), the Kruskal–Wallis H test was applied to assess whether statistically significant differences were present among the three independent dieback categories (Kruskal & Wallis, 1952). Following significant Kruskal–Wallis results, Dunn’s post-hoc test was performed to identify which specific class pairs and spectral indices exhibited statistically significant differences (Dinno, 2015; Dunn, 1964). This additional analysis enabled a more detailed interpretation of pairwise spectral separability among dieback severity levels. Together, these statistical procedures provided quantitative support for evaluating the effectiveness of each spectral index in distinguishing varying forest dieback conditions while also clarifying where significant differences were most pronounced among classes. In the Results section, numerical statistics presented in the text are reported as mean $\pm$ standard deviation (SD).

4.6 Model Development and Machine Learning Framework

To fulfil the primary objective of this study, a machine learning framework was developed to enable spatial detection of forest dieback across the Doller Valley. This framework focused on constructing a classification model capable of distinguishing between different levels of dieback severity using integrated field observations and Sentinel-2-derived spectral information (Figure 4).

A Random Forest (RF) classifier was employed due to its effectiveness in handling complex, non-linear relationships and its strong performance with multidimensional environmental datasets. The model was trained and evaluated using combined field-based dieback classifications and spectral

predictors derived from Sentinel-2 imagery. All stages of data processing, model implementation, and analysis were conducted in Python 3.11.10.

4.6.1 Random Forest Classifier

RF classifier is an ensemble machine learning algorithm that generates predictions by combining the outputs of multiple individual decision trees, thereby improving predictive accuracy and reducing overfitting.

The model input dataset was developed at the pixel level, and the model was trained using individual pixel observations rather than aggregated plot level means. This approach was chosen to preserve within-plot spectral variability associated with heterogeneous forest conditions. In particular, forest dieback is a spatially non-uniform process, often exhibiting fine-scale variation in canopy condition within the same plot. Therefore, aggregating pixel values to plot level mean values could potentially dilute or obscure localized stress signals associated with early or partial dieback stages.

The dataset included spectral information derived from all calculated indices, together with spatial, temporal, and species composition features. Spatial features included the spatial standard deviation (`spatial_std`), calculated from all pixels within the 10 m buffer surrounding each plot center, representing within plot spectral variability. Temporal features included temporal standard deviation (`temp_std`), which quantified the variability of individual pixel values over time. Additionally, a delta feature was calculated to represent the difference between the mean spectral value of the sampling year and that of a baseline reference year. The year 2017 was selected as the baseline reference year, as it was assumed to represent comparatively healthy forest conditions prior to the widespread drought events that affected European forests from 2018 onward, contributing substantially to forest decline (Knutzen et al., 2025; Schuldt et al., 2020). For these calculations, annual mean values were used for each pixel to ensure temporal consistency. In total, the initial model incorporated 21 predictor features (Appendix 3).

To maintain spatial independence and prevent data leakage, the dataset was divided into an 80/20 training and holdout testing split using a ‘Group_Shuffle_Split’ strategy. This ensured that all pixels originating from the same sample plot were assigned exclusively to either the training or testing subset. Initial model development was performed using the full feature set, followed by

hyperparameter tuning. Given the class imbalance within the field dataset, the Synthetic Minority Oversampling Technique (SMOTE) was applied to improve class representation during model training. This approach enhanced the classifier's ability to effectively learn from underrepresented severe dieback cases and improve predictive performance across all dieback severity levels (Chawla et al., 2002).

4.6.2 Hyperparameter Tuning and Model Optimization

Hyperparameter tuning was conducted to optimize the structural configuration of the RF model and ensure its suitability for the characteristics of the dataset. This process was essential for identifying the most effective parameter settings while minimizing the risks of overfitting or underfitting. By systematically evaluating multiple hyperparameter combinations, the model could be calibrated to achieve improved predictive performance and generalization.

To assess candidate hyperparameter configurations, K-fold cross-validation was employed as the internal evaluation strategy. The training dataset was partitioned into five folds, and the model was iteratively trained and validated five times, with each fold serving once as the validation subset while the remaining folds were used for training. This procedure was repeated five times so that each fold served once as the validation set. For each hyperparameter configuration, performance metrics were computed for each fold, and the resulting fold-wise scores were averaged to obtain an overall cross-validation score. This score was then used to compare and select the optimal hyperparameter configuration. This approach reduces sensitivity to any single train-validation split and improves the robustness and reliability of model selection.

4.6.3 Feature Importance Assessment

Feature selection is an important step in machine learning, as models often perform better when trained on a reduced set of informative variables rather than all available predictors (Grabska et al., 2020). Two complementary approaches were employed: Recursive Feature Elimination (RFE) and Permutation Importance. RFE was first applied using all 21 input features, where the model iteratively removed the least important variables based on internal importance rankings. At each iteration, the model was retrained, and this process continued until an optimal subset of features was identified that maximized model performance. To further validate the robustness of the selected features, Permutation Importance was applied. This method evaluates the effect of each

feature on model performance by randomly permuting its values and observing the resulting change in accuracy (Genuer et al., 2010). A significant drop in performance after shuffling indicates that the feature carries strong predictive information, whereas negligible changes suggest limited importance. By combining RFE with Permutation Importance, the feature selection process ensured both model-driven and performance-based validation of variable relevance. The final set of features used to build the model was selected based on those that achieved the highest cross-validation accuracy in combination with permutation importance results.

4.6.4 Accuracy Assessment

The performance of the developed model was evaluated using both a confusion matrix and a classification report. The confusion matrix provides a detailed comparison between the actual ground truth labels and the model's predicted labels. It summarizes the outcomes into four categories: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

The classification report was used to compute key evaluation metrics, including precision, recall, F1-score, and overall accuracy. Precision quantifies the correctness of positive predictions, defined as the proportion of correctly predicted positive observations relative to all predicted positives for a given class (Eq. 3).

$$Precision = \frac{TP}{TP + FP} \quad [3]$$

Recall, on the other hand, measures the model's ability to capture all relevant instances of a class, representing the proportion of correctly predicted positives relative to all actual positives in that class (Eq. 4).

$$Recall = \frac{TP}{TP + FN} \quad [4]$$

The F1-score provides a balanced measure of performance by computing the harmonic mean of precision and recall, thereby accounting for both false positives and false negatives simultaneously (Eq. 5).

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad [5]$$

Finally, overall accuracy represents the proportion of correctly classified samples relative to the total number of validation samples across all classes (Eq. 6).

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad [6]$$

In addition, Cohen's Kappa (k) coefficient was calculated to measure the level of agreement between predicted and observed dieback classes while accounting for agreement occurring by chance (Eq. 7).

$$\text{Kappa}(k) = \frac{p_o - p_e}{1 - p_e} \quad [7]$$

P_o – Observed Agreement, P_e – Expected Agreement

4.7 Dieback Prediction Map Development for the Doller Valley

Following the development of the final optimized model with selected features and tuned hyperparameters, the classifier was applied across the study area to generate spatially explicit forest dieback prediction maps for the Doller Valley. The selected predictor variables were computed for each pixel, enabling full spatial classification of dieback severity. The year 2017 was used as a static baseline, as it represents pre-drought conditions prior to the major European drought events beginning in 2018, providing a stable reference for initial forest conditions. A 3×3 moving window was applied to derive spatial standard deviation (spatial_std), and a 3×3 majority filter was used post-classification to reduce salt and pepper effect and improve spatial consistency. Annual dieback maps were then generated for the period 2018–2025, and class-wise area statistics were extracted to quantify temporal changes in forest condition.

Model confidence was derived from the Random Forest class probabilities, where each tree votes for a class, and the final prediction is based on majority vote. Confidence was defined as the highest class probability for each pixel and used to generate yearly confidence maps and mean annual confidence values.

To account for classification uncertainty in area estimation, an error-adjusted correction was applied following Hay, (1988) using information from the confusion matrix to construct an error probability matrix. This matrix describes the relationship between reference and mapped classes and is used to correct biased area estimates. The inverse correction approach, explained by Czaplewski & Catts (1992) was then applied to adjust class areas, providing more reliable estimates of true dieback extent by reducing systematic misclassification bias.

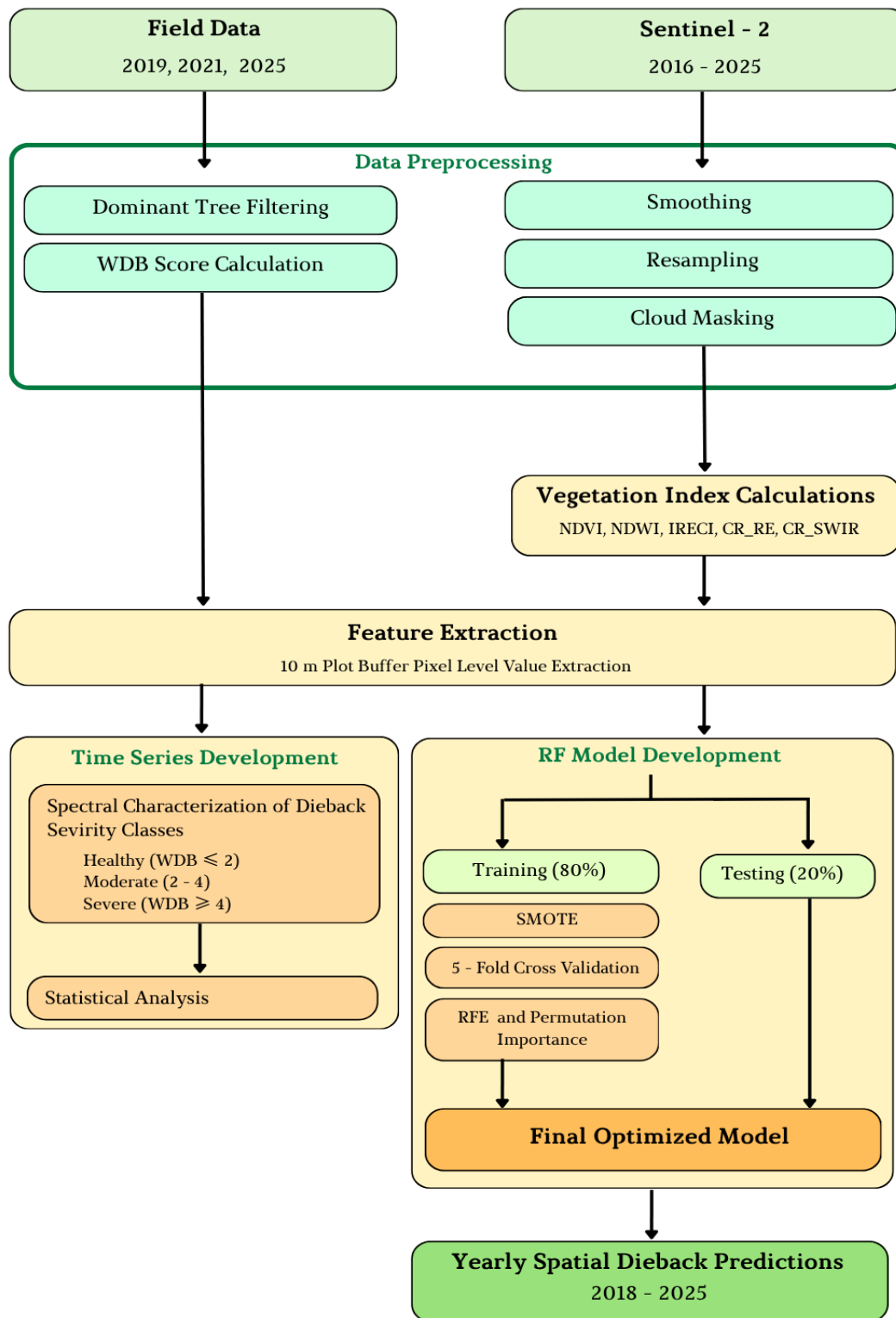


Figure 4. Workflow illustrating the development of the RF model for forest dieback classification using Sentinel-2 time-series imagery and field observations. (WDB = Weighted Dieback Score, RFE = Recursive Feature Elimination)

5. Results

This chapter presents the results obtained from the field observations, spectral analyses, and machine learning modelling. It begins with an assessment of dieback severity across the sampled plots, followed by analyses of spectral and temporal patterns associated with forest condition. The chapter then presents the outcomes of the Random Forest modelling framework and concludes with the spatial prediction and temporal evolution of forest dieback across the study area.

5.1 Overview of WDB Score Variations Among Plots

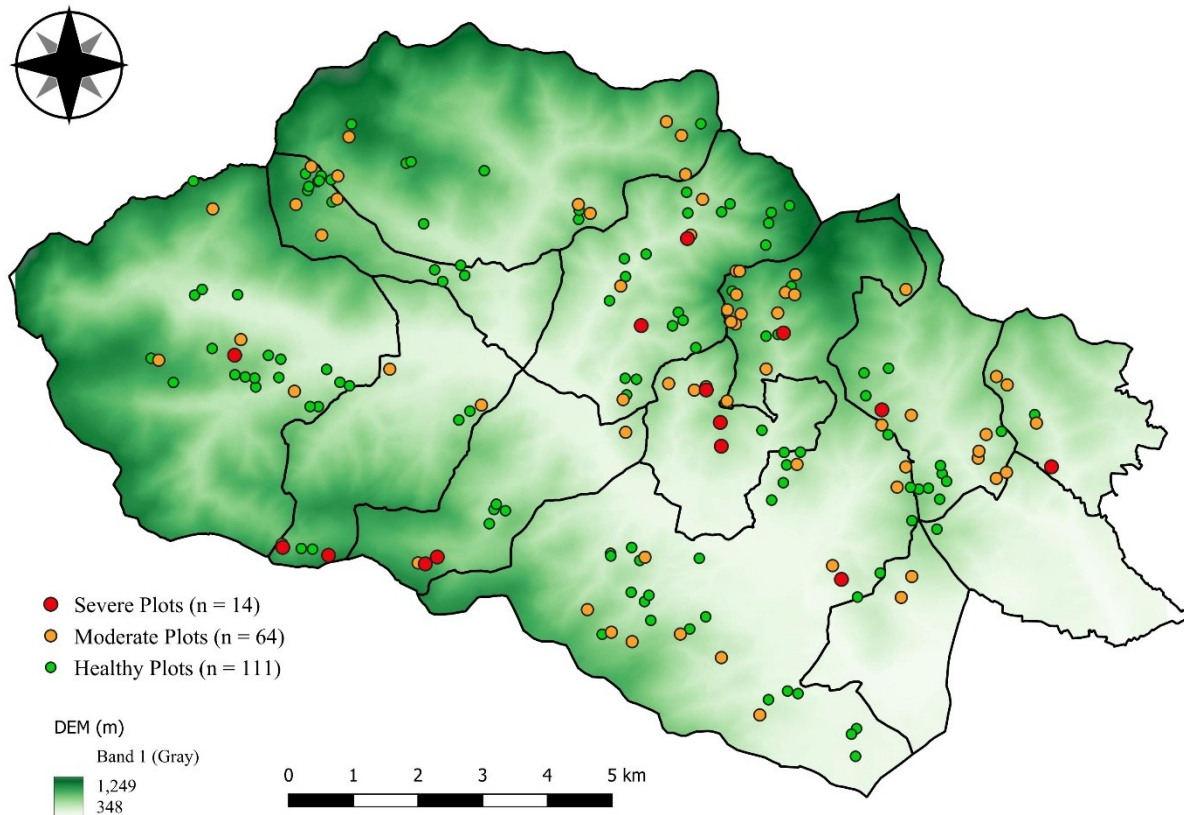


Figure 5. Map of the healthy, moderate, and severe dieback plots (sampled in 2019, 2021, and 2025) across the Doller valley (Base map DEM Source: Copernicus DEM data).

Following field data preprocessing and dominant tree selection, 1,292 dominant canopy trees were retained from the original 3,366 sampled trees across 189 plots for subsequent dieback analysis. Plot-level WDB scores derived from these dominant trees ranged from 0 to 5, reflecting substantial variability in the severity of canopy decline across the study area. Based on the established WDB

classification thresholds, the plots were categorized into three dieback severity classes: healthy ($WDB \leq 2$), moderate ($2 < WDB < 4$), and severe ($WDB \geq 4$) (Appendix 4). This classification resulted in 111 healthy plots, 64 moderate dieback plots, and 14 severe dieback plots, indicating that most sampled plots exhibited relatively low dieback severity, while a smaller proportion represented advanced canopy decline (Figure 5).

5.2 Time Series Analyses

The developed spectral index time series revealed the general temporal dynamics of forest condition across the Doller Valley throughout the study period (2016–2025). For each index, time series plots represented the mean spectral response across all sampled plots, accompanied by ± 1 standard deviation to illustrate spatial variability among plot-level signals. Collectively, these profiles captured both seasonal vegetation dynamics and inter-plot heterogeneity across the study area.

NDVI time series analysis demonstrated a strong and consistent seasonal pattern characteristic of temperate forest ecosystems. Mean NDVI values were consistently highest during the growing season, typically peaking at 0.8 ± 0.1 during the summer months, reflecting maximum canopy greenness and photosynthetic activity. During winter, average NDVI values declined to approximately 0.7 ± 0.1 , corresponding to reduced photosynthetic activity and seasonal leaf senescence associated with deciduous species shedding foliage. Spatial variability was notably greater during winter, as indicated by wider standard deviation ranges, reflecting increased heterogeneity among plots due to differences in species composition, particularly between deciduous and evergreen forest stands (Figure 6).

Similarly, NDWI exhibited a clear seasonal pattern, with mean values generally reaching approximately 0.4 ± 0.1 during summer and declining to winter minima between 0.3 ± 0.2 . These variations reflect seasonal fluctuations in vegetation water content, with lower winter values indicating reduced canopy moisture conditions. IRECI values generally ranged between 0.5 ± 0.2 , with elevated summer values indicating higher canopy chlorophyll content during peak vegetation productivity, followed by reduced values during winter (0.3 ± 0.2) corresponding to seasonal declines in chlorophyll concentration.

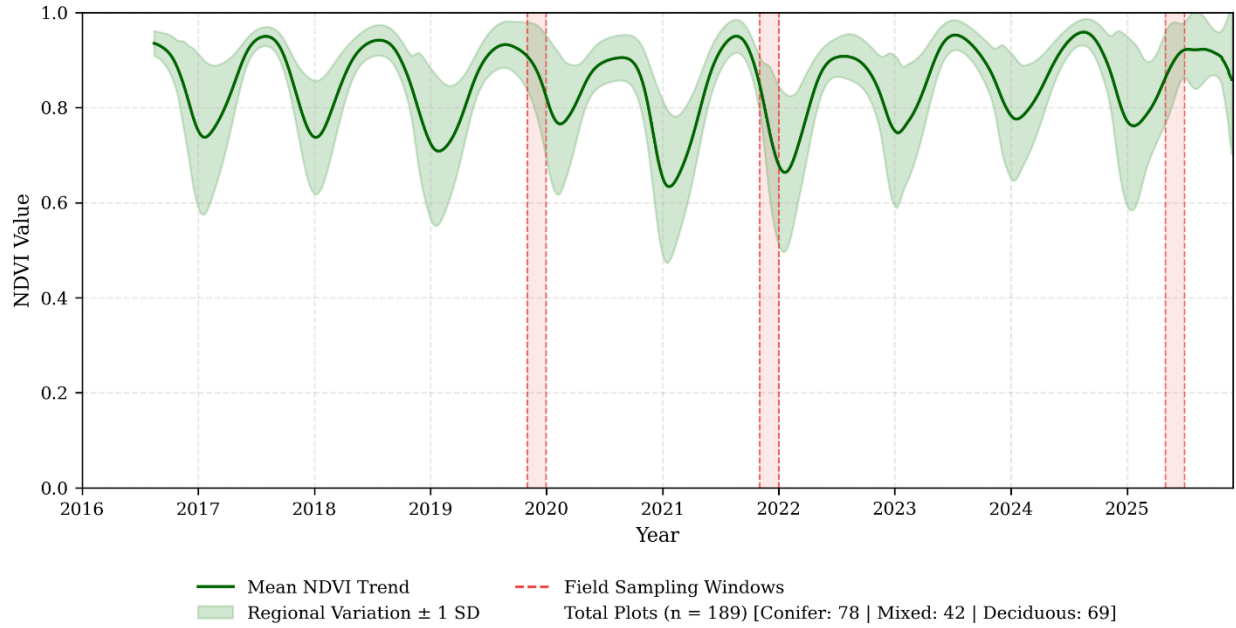


Figure 6. Time series plot of NDVI (2016-2025) showing the general trend of NDVI across all the plots in the study area, with the field sampling windows

The CR_RE index consistently reflected seasonal chlorophyll absorption dynamics, with deeper red-edge absorption features observed during summer (0.2 ± 0.2) and reduced absorption depths during winter (0.4 ± 0.2), corresponding to lower chlorophyll activity. Likewise, CR_SWIR demonstrated seasonal moisture absorption trends, with deeper absorption features during summer (0.3 ± 0.2) associated with higher vegetation water content and reduced absorption during winter months (0.4 ± 0.3). Unlike NDVI, however, NDWI, IRECI, CR_RE, and CR_SWIR all exhibited relatively high spatial heterogeneity throughout both summer and winter periods (Appendix 5).

5.3 Spectral Performance of Indices Across Forest Dieback Severity Levels

5.3.1 Spectral Differentiation Between Healthy and Severe Dieback Classes

Visual inspection of the developed time series plots for the initial two-class classification (healthy vs. severe dieback) revealed clear differences in spectral signatures between the two forest dieback classes, indicating that varying levels of dieback are associated with distinct spectral responses. Across all five spectral indices, healthy and severe plots generally exhibited separable mean temporal profiles. Among these, the continuum removed indices (CR_SWIR and CR_RE) showed the strongest visual distinction between classes, demonstrating greater sensitivity to dieback-related physiological changes compared to more generalized vegetation indices such as NDVI (Figure 7).

These visual patterns were supported by non-parametric Mann–Whitney U (MWU) tests, which confirmed highly significant spectral differences across all evaluated indices ($p < 0.05$), demonstrating that healthy and severe dieback plots were consistently spectrally distinct. The MWU statistic provides a measure of class separation, where lower U values indicate stronger separation and reduced overlap between classes. Consistent with the visual time series results, CR_SWIR and CR_RE exhibited the lowest U statistics among all evaluated indices (Table 3), confirming their superior ability to discriminate between the two classes. This strong separation was also reflected in their relative spectral responses, with CR_SWIR and CR_RE showing the highest mean increases from healthy to severe conditions (77% and 69%, respectively). In contrast, NDWI and IRECI showed intermediate differences, while NDVI displayed the weakest response, with only a 7% reduction between healthy and severe plots.

Table 3. Mann–Whitney U test results comparing healthy and severe dieback classes

Index	Healthy Mean	Severe Mean	U - Statistic	p-value
NDVI	0.9	0.8	1217	< 0.05
NDWI	0.4	0.2	1271	< 0.05
IRECI	0.4	0.3	1127	< 0.05
CR_RE	0.3	0.5	258	< 0.05
CR_SWIR	0.3	0.5	200	< 0.05

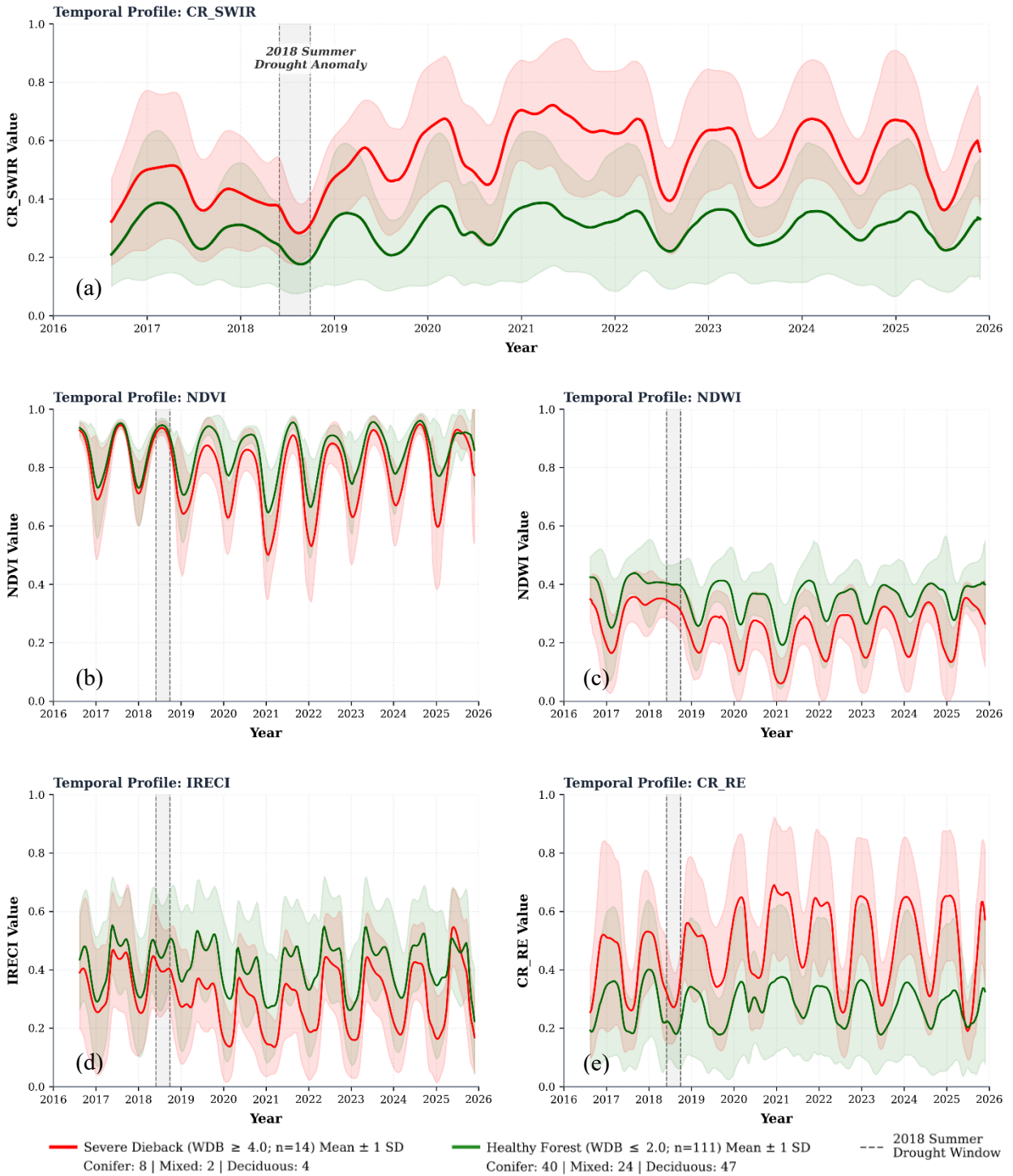


Figure 7. Comparison of spectral signatures between healthy ($WDB \leq 2$) and severe ($WDB \geq 4$) dieback plots across all indices

5.3.2 Spectral Differentiation Among Healthy, Moderate, and Severe Dieback Classes

Within the 3-classes classification framework (healthy, moderate, and severe), visual spectral differentiation was less clear than in the two-class comparison. Although severe dieback plots generally remained distinguishable, the spectral differences between healthy and moderate plots were often inconsistent across indices. Moderate dieback plots frequently showed overlapping temporal patterns with both healthy and severe classes, making class separation less visually distinct. This reduced separability is partly due to the use of plot-level mean spectral values, where averaging all pixels within a plot can dilute localized dieback signals. Since forest dieback is often spatially uneven within plots, this averaging process can mask fine-scale physiological variation and increase overlap among classes (Figure 8).

The Kruskal–Wallis H test confirmed statistically significant differences among the dieback severity classes ($p < 0.05$), further supporting the patterns observed in both the visual and two-class analyses (Table 4). The H statistic reflects differences in median ranks between groups, where higher values indicate stronger separation among classes. In this context, CR_SWIR and CR_RE consistently produced the highest H values, indicating that continuum-removed indices provide the most robust separation of dieback severity levels.

Post-hoc analysis using Dunn’s test further clarified pairwise differences between classes. The results showed that the healthy and moderate classes were not significantly different across all indices ($p > 0.05$), suggesting that moderate dieback conditions retain spectral characteristics that are still very similar to healthy forests. This statistical result supports the overlap observed in the temporal profiles. In contrast, the moderate and severe classes showed statistically significant differences across all indices ($p < 0.05$), with the strongest separation again observed for CR_SWIR and CR_RE. This indicates that the transition from moderate to severe dieback is associated with more pronounced structural and biochemical changes, which are more effectively captured by continuum-removed indices than by traditional vegetation indices.

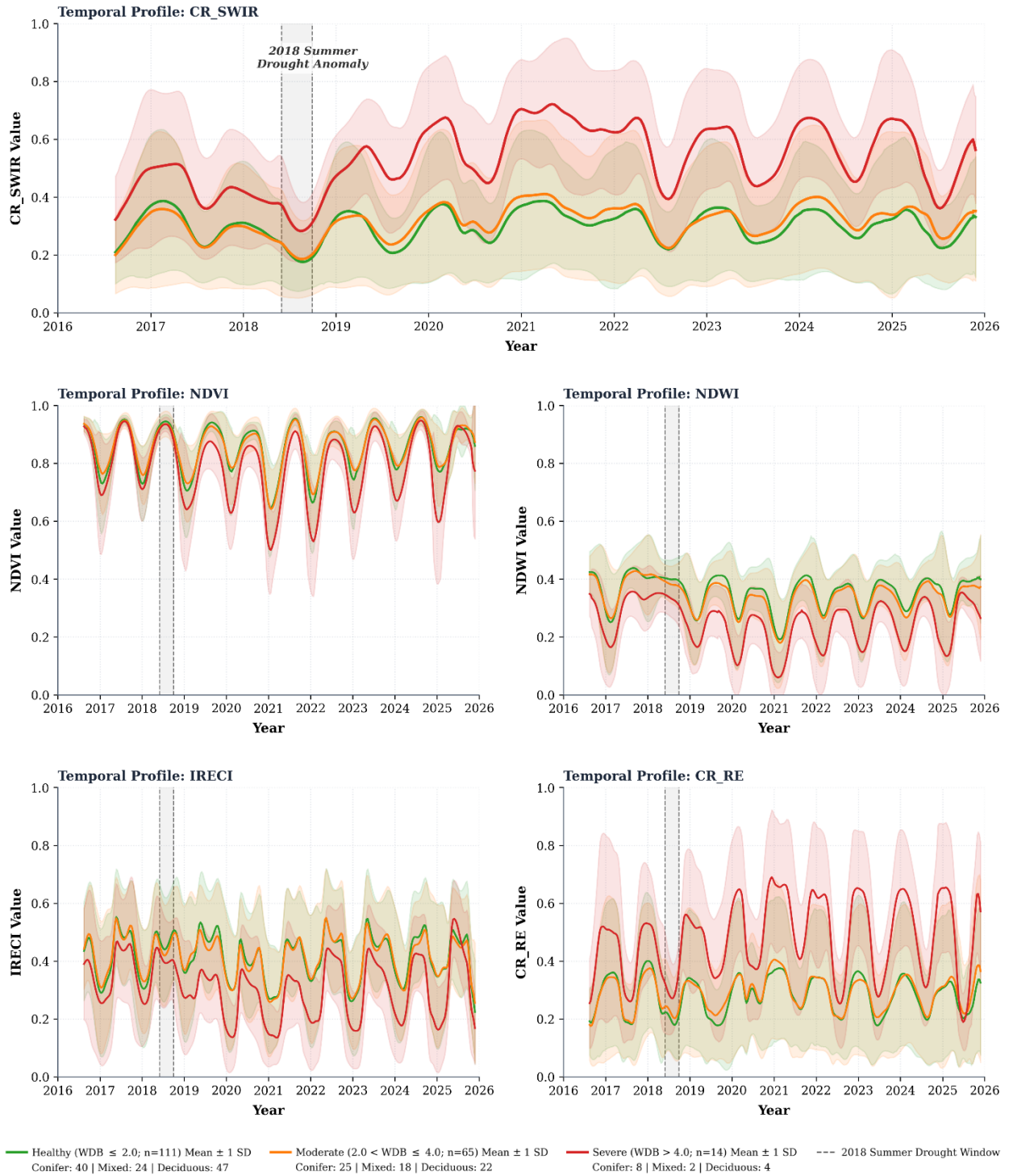


Figure 8. Comparison of spectral signatures between healthy ($WDB \leq 2$), moderate ($2 < WDB < 4$), and severe ($WDB \geq 4$) dieback plots across all indices

Table 4. Kruskal-Wallis H test with Dunn's post-hoc analysis results comparing the healthy, moderate, and severe dieback classes

Index	H - Statistic	Healthy Vs. Moderate (p)	Moderate Vs. Severe (p)	Healthy Vs. Severe (p)
NDVI	14	> 0.05	< 0.05	< 0.05
NDWI	16	> 0.05	< 0.05	< 0.05
IRECI	8	> 0.05	< 0.05	< 0.05
CR_RE	16	> 0.05	< 0.05	< 0.05
CR_SWIR	20	> 0.05	< 0.05	< 0.05

5.3.3 Spectral Responses Across Forest Types

When analyses were further separated by forest type, visual differences in spectral signal ranges were observed among coniferous, mixed, and deciduous forests. Coniferous and mixed forests generally showed similar temporal patterns, although differences in variability and signal range were present. Deciduous forests displayed more visually distinct differences between dieback classes, likely due to stronger seasonal phenological changes such as leaf development and senescence. However, due to the relatively small number of moderate and severe deciduous plots within the dataset, these visual patterns may not fully represent broader deciduous forest dieback dynamics across the study area (Figure 9) (Appendix 6).

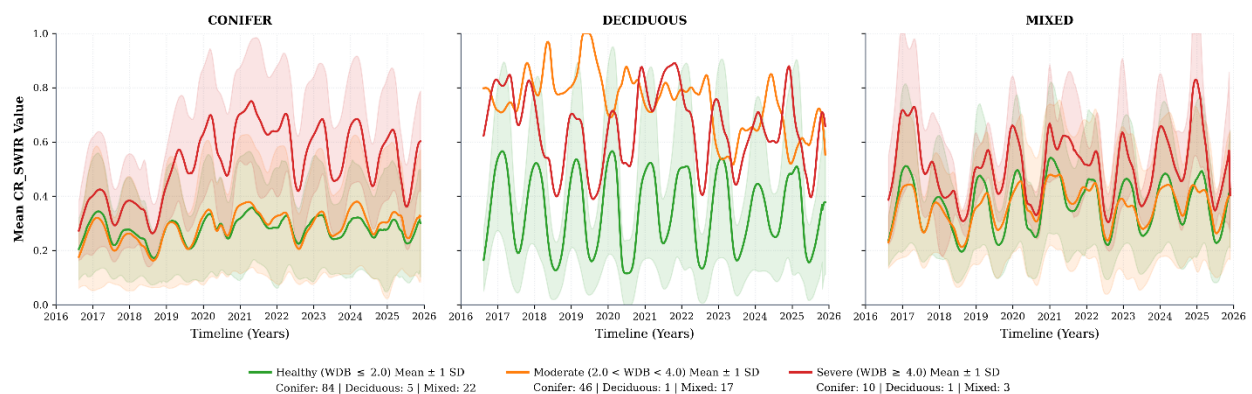


Figure 9. Comparison of CR_SWIR temporal variation between conifer, deciduous, and mixed plots

5.4 Temporal Evolution of Forest Dieback Using Orthophotos and Spectral Indices

The temporal evolution of forest dieback was further examined by integrating spectral time series with high-resolution orthophotos, which provided a visual reference for validating observed index behaviour over time. Overall, the comparison revealed a strong correspondence between spectral dynamics and visible canopy conditions, indicating that the indices effectively capture real changes in forest structure and health.

For a representative healthy plot (sampled in 2021; WDB = 0), NDVI time series from 2017 to 2025 showed a highly consistent seasonal pattern across all years (Figure 10). Each year exhibited a clear summer peak, reflecting peak canopy greenness, with tightly clustered maximum values and stable interannual behaviour. Winter values showed greater variability, likely reflecting differences in species composition within the plot, where mixed deciduous and coniferous stands maintain some level of winter photosynthetic activity due to the presence of evergreen species. Importantly, the orthophotos from 2018, 2021, and 2024 consistently confirmed the presence of a dense and healthy canopy, supporting the stability observed in the NDVI signal. The other indices also showed similarly stable temporal behaviour, with low interannual variability and closely aligned trajectories across the study period (Appendix 7), further confirming the persistence of healthy forest conditions.

In contrast, the severe dieback plot (sampled in 2019; WDB = 5) exhibited pronounced temporal changes in spectral behaviour that closely aligned with observed structural changes in the orthophotos (Figure 11). Prior to 2019, the NDVI signal remained relatively high (2016–2018), indicating that the plot was still largely healthy during this period. This is consistent with the 2018 orthophoto, which shows a still green and relatively intact canopy. From 2019 onwards, a clear decline in NDVI was observed, reaching minimum values around 2021, which corresponds well with orthophoto evidence showing substantial canopy loss and open forest structure in the same year. Following 2021, NDVI values gradually increased again between 2022 and 2024. This trend is also reflected in the 2024 orthophoto, which indicates partial regrowth and understory development, likely contributing to increased reflectance despite the absence of a fully recovered canopy.

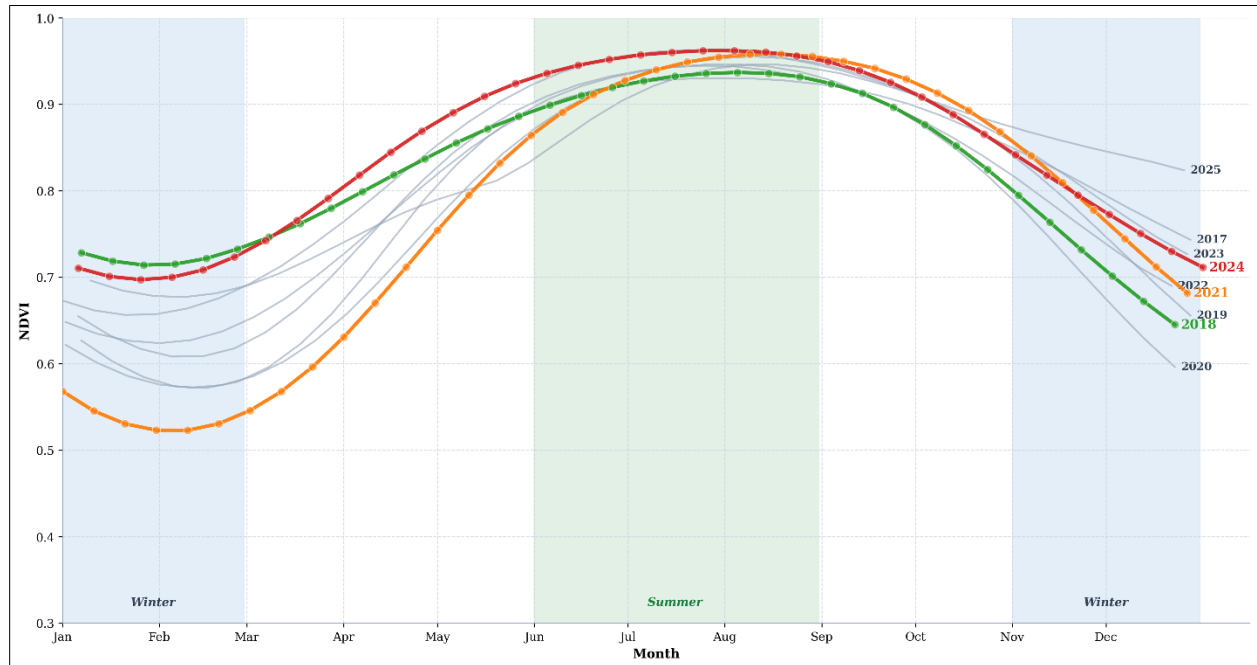


Figure 10. Temporal NDVI profile of a healthy plot matched with multi-temporal orthophotos. Each line indicates the NDVI trend over the years (2017 - 2025), accompanied by orthophotos capturing the canopy condition at the same plot location (within 10m buffer) from 2018 (before field sampling), 2021 (2 years after field sampling and 2025 (6 years after field sampling)

However, the overall signal range in the severe plot remained much more variable across years compared to the healthy plot, highlighting persistent instability in vegetation condition following dieback onset. The remaining indices (NDWI, IRECI, CR_RE, and CR_SWIR) showed similarly consistent temporal responses, with greater interannual variability compared to the healthy plot (Appendix 7). In particular, CR_SWIR captured a pronounced shift between 2018 and 2019, indicating a rapid change in canopy water related properties during the period when dieback became visually apparent. Overall, the spectral trajectories across all indices closely matched the changes observed in the orthophotos, demonstrating that the indices are highly sensitive to both

gradual and abrupt changes in forest condition and provide a reliable representation of dieback progression over time.

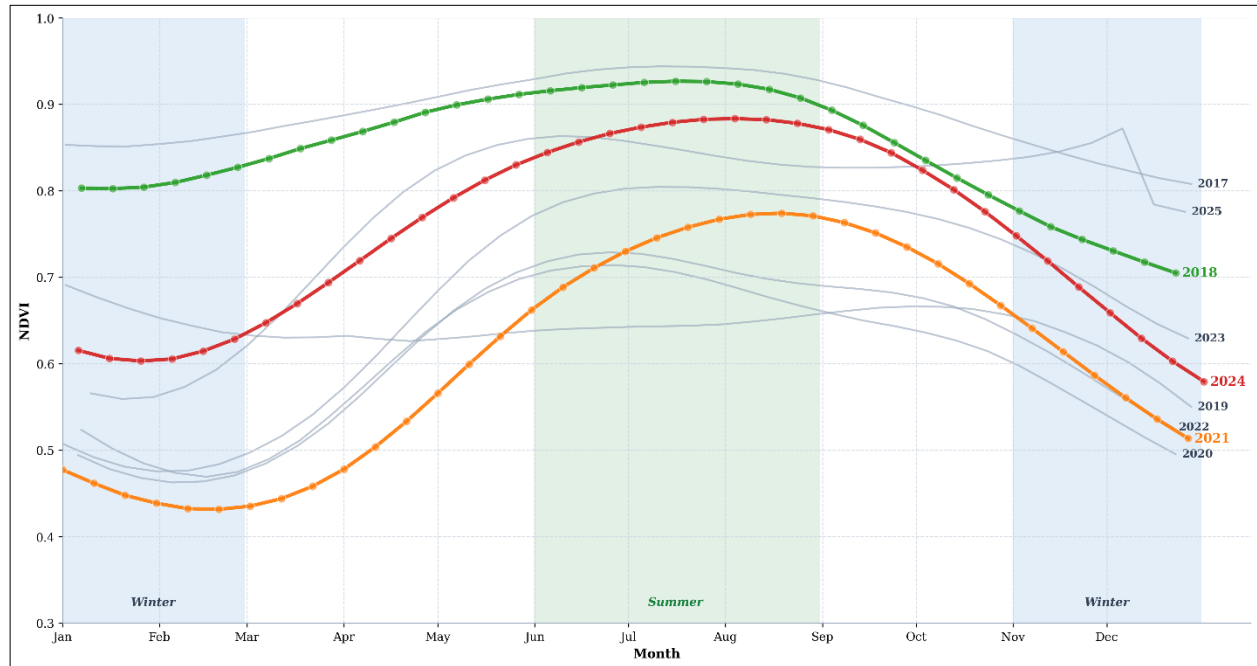


Figure 11. Temporal NDVI profile of a severe plot matched with multi-temporal orthophotos. Each line indicates the NDVI trend over the years (2017 - 2025), accompanied by orthophotos capturing the canopy condition at the same plot location (within 10m buffer) from 2018 (before field sampling), 2021 (2 years after field sampling) and 2025 (6 years after field sampling)

5.5 RF Dieback Detection Model

5.5.1 Initial Hyperparameter Optimization and Model Performance

The initial hyperparameter optimization was conducted using all 21 predictor features (listed in Appendix 3). Through 100 iterations combined with 5-fold cross-validation, the model explored a broad range of parameter configurations to identify an optimal learning structure. The final selected hyperparameters, along with their corresponding search ranges, are presented in Table 5. Using the optimized settings and full feature set, the Random Forest model achieved an overall accuracy of 60%, with a macro-averaged F1-score of 0.61. The class-wise precision, recall, and F1-scores are summarized in Table 6, and the corresponding confusion matrix is shown in Figure 12.

Performance was notably higher for the healthy and severe classes, both of which were classified with relatively good accuracy. In contrast, the moderate class showed considerably lower performance, indicating difficulty in separating this intermediate category from the other two classes. The confusion matrix further highlights this pattern, showing substantial misclassification between healthy and moderate classes, with a considerable proportion of moderate pixels being classified as healthy and, to a lesser extent, healthy pixels being assigned to the moderate class.

Table 5. Hyperparameters for the initial model run with 21 features and the optimized model run with 7 features

Hyperparameter	Search Range	Optimized Values with 21 features	Optimized Values with 7 features
Number of estimators	100 - 1000	317	665
Max depth	5 - 50	18	44
Max features	0.21 – 1.00	0.63	5
Min samples per split	2 – 100	6	7
Min samples per leaf	1 - 100	1	2
Bootstrap	True / False	True	False

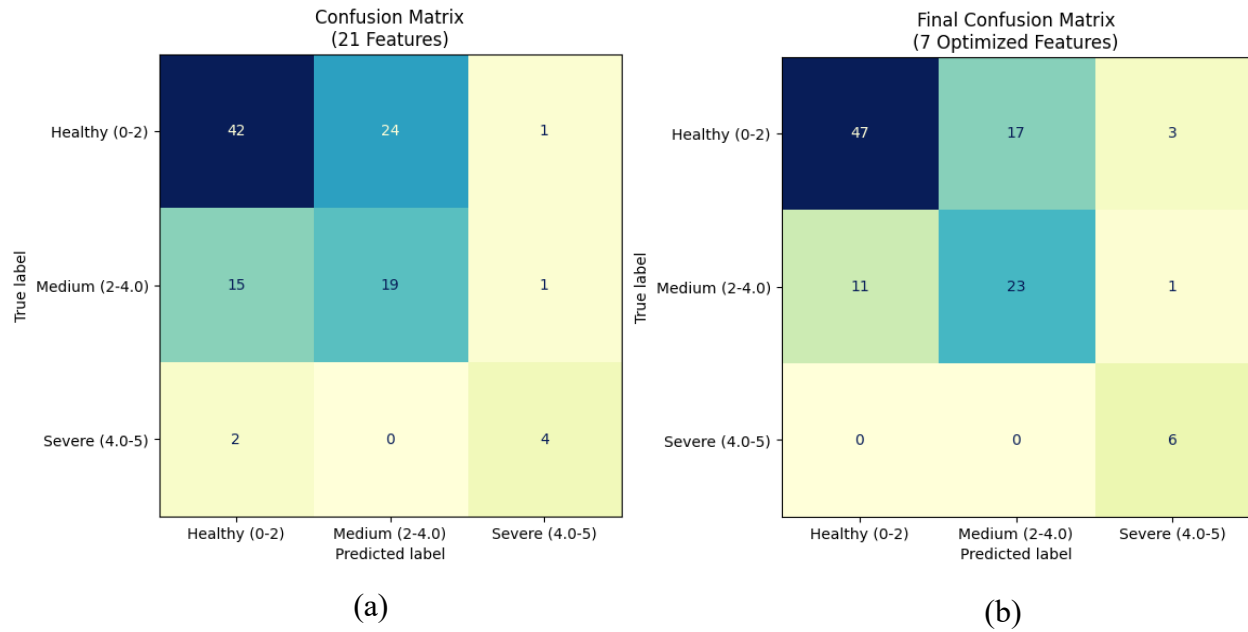


Figure 12. Confusion matrix showing the model performance with (a) 21 features and (b) optimized 7 features

5.5.2 Recursive Feature Elimination and Permutation-Based Feature Selection

Following hyperparameter optimization, RFE and permutation importance were applied to identify the most informative predictor variables. RFE results indicated that the highest predictive performance was achieved using a reduced subset of 7 features. Starting from the full set of 21 variables (Figure 13), mean cross-validation accuracy gradually improved as less informative and redundant predictors were removed, reaching a peak value of 0.53. Beyond this optimal point, further feature elimination resulted in a decline in performance (CV accuracy < 0.53), indicating that the selected seven variables capture the essential spectral and structural information required for distinguishing dieback severity levels (Figure 14).

The selected feature set included species type, CR_SWIR_delta, IRECI_val, NDVI_delta, CR_SWIR_spatial_std, NDWI_delta, and CR_RE_temp_std (detailed definitions in Appendix 3). Using this reduced feature set, the model was re-trained, and hyperparameters were re-tuned to ensure optimal configuration under the updated input space (Table 5). The final optimized model achieved an overall accuracy of 70% with a balanced accuracy of 79% and a macro-averaged F1-

score of 0.71, indicating balanced performance across all dieback classes rather than dominance by the majority class.

Class-wise evaluation (Table 6) shows improved precision, recall, and F1-scores across all categories compared to the full-feature model. The confusion matrix (Figure 12) further confirms that the model correctly classifies the majority of pixels in each class, with reduced confusion particularly between healthy and moderate classes relative to the initial model.

Table 6. Classification accuracies with all 21 features and the top 7 features after RFE

	Precision	Recall	F1 – score
With the initial 21 features			
Healthy ($\text{WDB} \leq 2$)	0.71	0.63	0.67
Moderate ($2 < \text{WDB} < 4$)	0.44	0.54	0.49
Severe ($4 \leq \text{WDB}$)	0.67	0.67	0.67
With the top 7 features			
Healthy ($\text{WDB} \leq 2$)	0.81	0.70	0.75
Moderate ($2 < \text{WDB} < 4$)	0.57	0.66	0.61
Severe ($4 \leq \text{WDB}$)	0.60	1.00	0.75

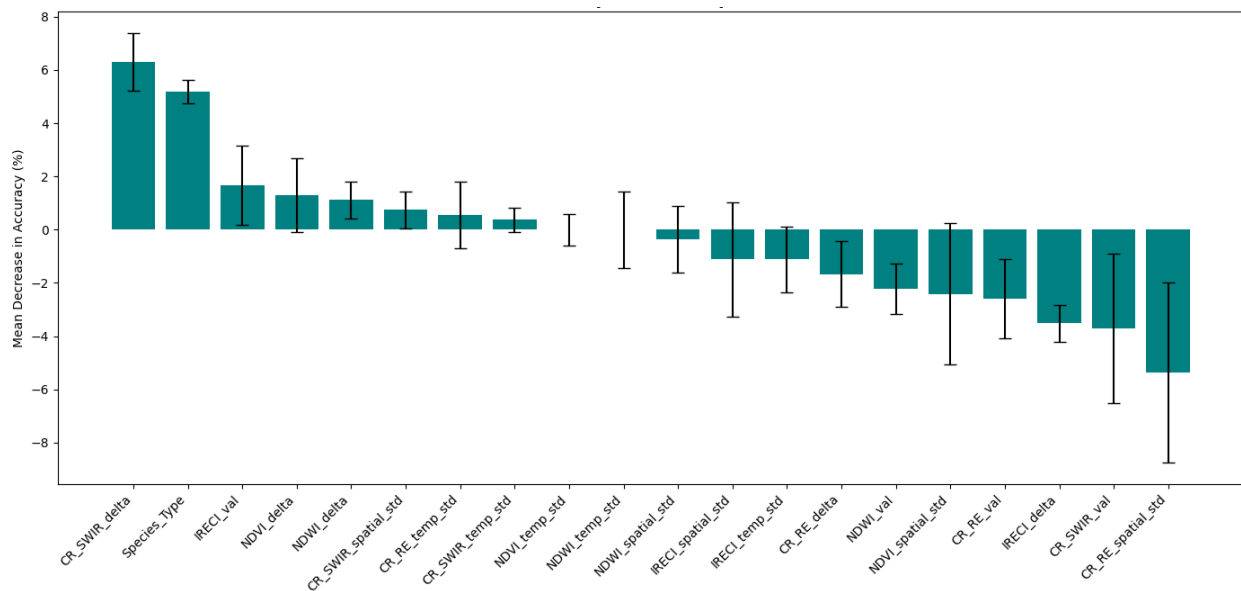


Figure 13. Permutation importance showing the mean decrease in accuracy (%) for all 21 input features

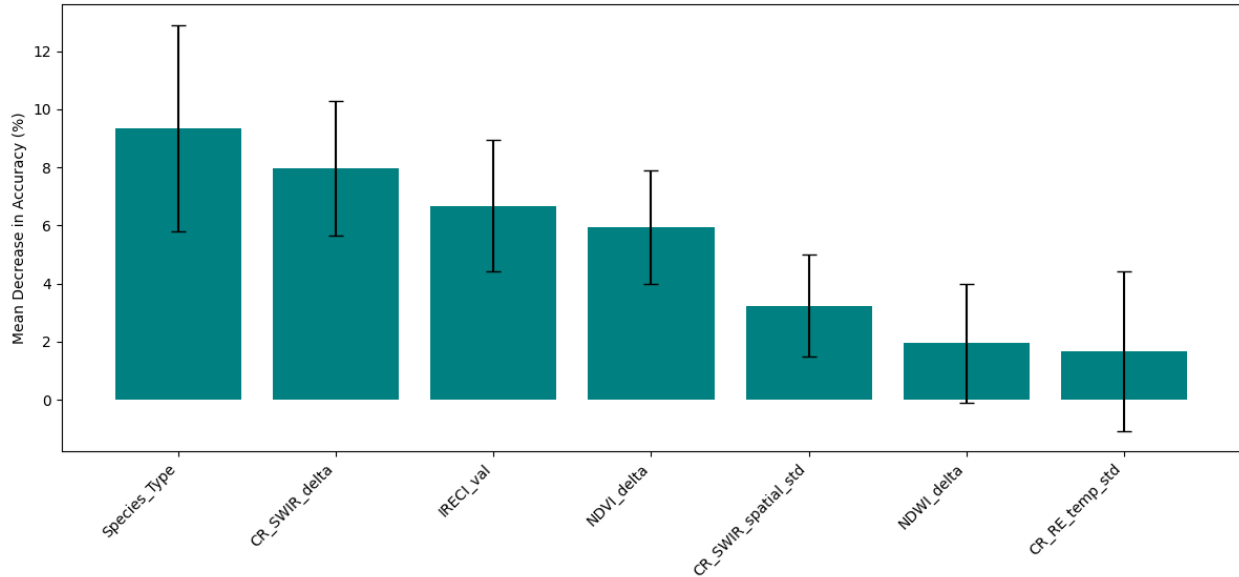


Figure 14. Permutation importance showing the mean decrease in accuracy (%) for the optimized 7 features

Cohen's Kappa coefficient increased from 0.24 in the full-feature model to 0.55 in the reduced-feature model, indicating a shift from slight to moderate agreement. This improvement demonstrates that the reduced feature set not only enhances classification performance but also increases robustness beyond chance agreement, despite the inherent class imbalance in the field dataset. Overall, the combination of feature selection and re-tuned hyperparameters substantially improved the model's ability to distinguish between the three dieback severity classes.

5.6 Spatial Prediction for Forest Dieback and Temporal Evolution

Spatial prediction maps were generated annually for the period 2018–2025, with an overall average model confidence of approximately 68%, indicating consistent predictive performance across the study period (Figure 15). Based on area calculations for the three dieback severity classes, 2018 exhibited the lowest extent of healthy forest (56 km²) and the highest extent of severe dieback (14 km²) observed throughout the study period. Between 2019 and 2022, the spatial predictions indicated a gradual recovery phase, characterized by an approximate 10 km² increase in healthy forest area and a corresponding reduction of approximately 5 km² in severe dieback extent. However, from 2022 to 2023, this recovery trend partially reversed, with reductions observed in both healthy and moderate dieback classes alongside an increase in severe dieback area. This

suggests renewed forest stress or delayed dieback progression during this period. In 2024, the maps indicated a slight recovery, reflected by a reduction in severe dieback extent (Table 7).

Table 7. Annual extent (km²) of forest dieback under healthy, moderate, and severe dieback classes from 2018 to 2025.

Year	Healthy (WDB ≤ 2) (km ²)	Moderate ($2 < \text{WDB} < 4$) (km ²)	Severe (WDB ≥ 4) (km ²)
2018	56	36	14
2019	64	35	8
2020	68	31	8
2021	66	33	7
2022	65	32	10
2023	64	32	11
2024	67	32	8
2025	67	27	12

Comparison with annual model confidence maps further demonstrated that healthy (0.68 ± 0.14) and severe (0.69 ± 0.16) dieback classes were predicted with slightly higher confidence than the moderate class (0.64 ± 0.13) (Appendix 8). This pattern aligns with previous classification results, indicating that moderate dieback remains the most spectrally ambiguous class, whereas healthy and severe forest conditions exhibit more distinct spectral signatures and are therefore classified more reliably.

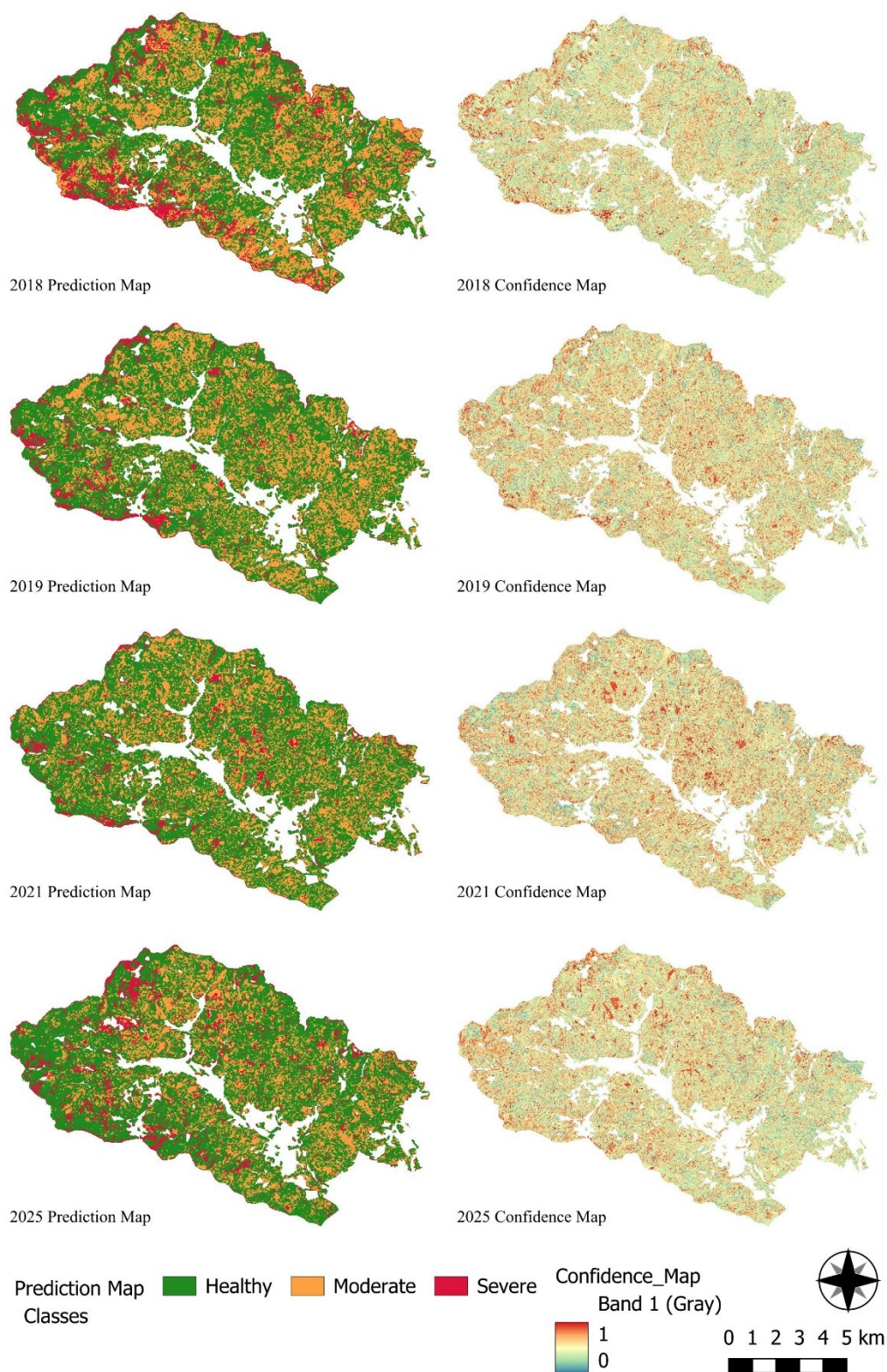


Figure 15. Annual spatial predictions of forest dieback severity and corresponding model confidence maps derived from the RF classification from 2018 to 2025.

5.6.1 Error Adjusted Area Calculations

Application of the Hay (1988) area correction method revealed notable differences between raw classified dieback areas and bias-adjusted estimates. Due to the availability of field validation data for only three survey years, error-adjusted area calculations were conducted exclusively for 2019 and 2021. Although field data were also available for 2025, this year was excluded from the correction analysis due to insufficient representation of severe dieback plots.

For 2019, the raw classification substantially overestimated healthy forest area by approximately $\sim 9 \text{ km}^2$, while moderate dieback area was underestimated by $\sim 10 \text{ km}^2$. Following calibration, healthy forest extent decreased from $\sim 64 \text{ km}^2$ to $\sim 55 \text{ km}^2$, whereas moderate dieback increased from 35 km^2 to 45 km^2 . Severe dieback showed only a minor reduction, from $\sim 8 \text{ km}^2$ to $\sim 7 \text{ km}^2$.

In 2021, bias correction resulted in comparatively smaller adjustments, reflecting improved classification performance. Healthy forest area increased from $\sim 66 \text{ km}^2$ to $\sim 71 \text{ km}^2$, while moderate dieback decreased from $\sim 33 \text{ km}^2$ to $\sim 29 \text{ km}^2$. Severe dieback again required only a minor correction, decreasing from $\sim 7 \text{ km}^2$ to $\sim 6 \text{ km}^2$. Overall, classification errors were notably lower in 2021 than in 2019. These findings indicate that moderate dieback classes were the most susceptible to misclassification, likely due to their intermediate spectral characteristics, whereas severe dieback estimates remained comparatively robust across both assessed years.

6. Discussion

6.1 Spectral Indicators and Their Effectiveness For Dieback Detection

6.1.1 Temporal Behaviour of Spectral Indices

The time-series analysis of all spectral indices revealed clear seasonal dynamics across the study area, reflecting the phenological cycles of forests within the Doller Valley. However, the magnitude of variation differed substantially among the indices. NDVI consistently exhibited the narrowest variation range, with only small deviations around the mean, whereas NDWI, CR_SWIR, CR_RE, and IRECI displayed considerably wider variability over time and across plots (section 5.2 provides mean $\pm$ SD). This contrast suggests that although the forests generally maintained a green healthy canopy appearance, the underlying physiological conditions were far more heterogeneous. Indices associated with canopy moisture content (NDWI and CR_SWIR), chlorophyll concentration, and red-edge region (CR_RE and IRECI) revealed substantial differences in forest condition that were not fully reflected by canopy greenness alone. Consequently, while NDVI effectively represents overall canopy greenness, it may be less sensitive to the physiological changes associated with forest stress and dieback. Similar findings have been reported by Marusig et al. (2020), who observed weaker relationships between NDVI, and physiological drought stress indicators compared to moisture-sensitive indices such as NDWI.

This relatively stable and uniformly high NDVI values can be explained by the saturation effect commonly observed in dense temperate forests. In forests with high Leaf Area Index (LAI), red light is strongly absorbed by chlorophyll, resulting in low red reflectance, while near-infrared reflectance remains consistently high due to leaf internal structure (Huete, 2012; Marusig et al., 2020; Tucker, 1979). Under these conditions, NDVI approaches a plateau where additional physiological variation within the canopy produces only limited changes in index values. As a result, stressed forests can continue to exhibit high NDVI values as long as the canopy remains visually dense and green. This behaviour explains why NDVI showed consistently high values with comparatively low variability throughout the study period, despite clear differences in moisture and chlorophyll-related conditions observed in the other indices. Hence, the results of this study indicate that NDVI mainly captures canopy greenness rather than internal physiological

stress responses. This limits its effectiveness as a standalone indicator for detecting forest dieback in dense temperate forest environments.

6.1.2 Spectral Separability of Dieback Classes

Although all indices showed statistically significant differences among dieback classes, the continuum removal indices (CR_SWIR and CR_RE) consistently provided the strongest separability between healthy, moderate, and severe conditions (Table 3 and Table 4). They outperformed NDVI, NDWI, and IRECI because these indices are more directly linked to physical absorption features within the canopy reflectance spectrum, while continuum removal indices are more sensitive to internal biochemical changes within the canopy. Specifically, CR_SWIR is particularly sensitive to canopy moisture status, whereas CR_RE captures changes associated with chlorophyll degradation and alterations in red-edge structure. By emphasizing the depth and shape of absorption features rather than overall reflectance magnitude, continuum removal enhances sensitivity to internal canopy changes while reducing the masking effects of canopy brightness. Consequently, these indices were better able to distinguish varying levels of forest condition across the dieback gradient. The temporal evolution of class separability further highlights these differences.

During the 2018–2019 period, NDVI showed substantial overlap between healthy and severely affected plots, indicating limited sensitivity to early physiological stress following the 2018 European drought. In contrast, CR_SWIR and CR_RE already exhibited clear separation between health classes from 2018 onwards, suggesting earlier detection of canopy stress. A clearer divergence in NDVI only emerged after 2019, when canopy degradation became more advanced (Figure 7). However, NDVI still captured the overall trajectory of forest decline. A pronounced decrease in NDVI values from 2019 onwards closely corresponded with the onset of visible dieback symptoms following the severe 2018 drought. This indicates that, although NDVI is less effective in distinguishing early-stage physiological stress, it remains valuable for representing broader canopy-level deterioration once dieback becomes more advanced (Figure 7b).

These findings are consistent with previous studies highlighting the importance of red-edge and SWIR spectral regions for forest stress detection. Abdullah et al. (2019) emphasized their role in early drought stress identification, while Ye et al. (2021) demonstrated strong performance of red-

and SWIR-based indices for mapping bark beetle-induced dieback. Similarly, Carletti et al. (2025) identified CR_SWIR, CR_RE, and IRECI as among the most effective indicators for detecting dieback across multiple conifer species. The spectral behaviour observed in this study also aligns with the operational oak dieback monitoring framework proposed by Mouret et al. (2024) where SWIR-based indicators provided the strongest separation between health conditions. Comparable results were reported by Bárta et al. (2021) and Huo et al. (2021) who identified red, red-edge, and SWIR bands as the most reliable spectral domains for distinguishing healthy and stressed forest stands.

Moreover, although mean class level spectral trends showed clear differences in class classifications (Figure 7), substantial overlap was observed among individual pixel values within plots. This pronounced intra plot heterogeneity shows that forest dieback does not occur uniformly across individual stands. Even severely affected plots may contain localized areas exhibiting relatively healthy spectral responses due to resilient canopy patches, species specific variation, or microenvironmental conditions, while healthy plots may also contain isolated stressed or declining areas. Consequently, despite clear visual class separation at the aggregated level, considerable within plot variability was evident across both classification cases.

6.1.3 Influence of Understory Vegetation on Spectral Signals

An important pattern emerged in the NDVI and IRECI time series, where several severely affected plots showed an apparent recovery by 2025. NDVI values in these plots approached levels close to values observed in healthy stands. However, this trend was not reflected in CR_SWIR or CR_RE, which remained consistent with ongoing canopy degradation (Figure 7). Orthophoto comparison confirmed that this apparent recovery was not caused by overstory regeneration but instead resulted from increased understory vegetation following canopy opening (Figure 11). As overstory trees declined, understory vegetation became more exposed and contributed more strongly to the satellite signal. Because this vegetation remained green and photosynthetically active, it artificially increased NDVI and partially affected IRECI values, leading to an overestimation of forest recovery.

Understory influence on canopy reflectance has been widely documented in previous studies. Lu & Zheng (2016) showed that understory vegetation can contribute between 20% and 40% of total canopy reflectance, while Eriksson et al. (2006) reported variations of up to $\pm 18\%$ in the red band

and $\pm 10\%$ in the near-infrared region due to understory effects. Although these studies primarily focused on LAI estimation, the same mechanisms directly affect vegetation indices such as NDVI and IRECI, leading to systematic overestimation of forest health in open canopies. In contrast, the continuum removal based indices demonstrated greater robustness under these conditions. Their reduced sensitivity to overall brightness variations and mixed background reflectance allowed CR_SWIR and CR_RE to maintain stable dieback trajectories despite increasing understory exposure. Similar advantages of continuum removal approaches have been reported by Malenovsky. (2006) who showed that emphasizing absorption feature geometry minimizes background and albedo effects. Comparable findings were also reported by Kokaly & Clark (1999) and Curran et al. (2001) who demonstrated that continuum-removed spectral features improve the detection of subtle biochemical and structural canopy changes in heterogeneous vegetation environments.

6.2 Random Forest Model Performance

The Random Forest model achieved an overall accuracy of 70% and a balanced accuracy of 79%, indicating a generally strong ability to discriminate between forest health classes across the Doller Valley. However, performance varied substantially among classes, with the primary limitation observed in the separation between healthy and moderate dieback conditions, as reflected in the reduced recall for the moderate class (0.66). This suggests systematic difficulty in identifying transitional canopy states which also have been explained by Bárta et al. (2021) and Hawryło et al. (2018).

The results are consistent with the statistical analysis, where Kruskal–Wallis and post-hoc tests indicated no significant differences between healthy and moderate classes across several spectral indices ($p > 0.05$). Although separability between these classes remained limited, the inclusion of a 3-classes framework was necessary to represent the full degradation continuum. Forest dieback is not a binary process, and excluding intermediate conditions would oversimplify gradual physiological transitions within forest canopies.

Despite the statistical overlap observed between healthy and moderate classes, the Random Forest model achieved comparatively strong classification performance within the 3-classes framework. Belgiu & Drăguț (2016) explained that ensemble learning methods such as Random Forest can

achieve strong predictive performance in high dimensional datasets with overlapping class distributions by exploiting complex, non-linear interactions among variables rather than relying solely on strong univariate separability. Michelucci & Venturini (2026) further support this interpretation. In the present study, the relatively high overall accuracy is therefore likely explained by the model's ability to integrate multiple weakly separable spectral features into a robust classification structure.

Most existing studies on forest dieback detection have relied on binary classification schemes, typically distinguishing only between healthy and disturbed conditions. For example, Bárta et al. (2021) achieved an overall accuracy of 78% in discriminating healthy from declining Norway spruce stands using time-series analysis in Central Europe, while Grabska et al. (2020) reported 75% accuracy in detecting Scots pine forest loss using Sentinel-2 time series in Poland. In contrast, relatively few studies have implemented a 3-classes framework. Mouret et al. (2024) provide one of the most comparable examples in oak dieback monitoring, where similar performance levels to this study were reported despite increased classification complexity. Mouret et al. (2024) reported overall accuracies of approximately 65% with balanced accuracies around 61%, with substantial confusion between declining and very declining classes. The results of the present study, therefore, align with the broader pattern observed in literature, where classification performance remains relatively stable across binary studies but becomes increasingly challenging when transitioning from binary to multi-class representations of forest decline.

A key limitation of the model in this study is class imbalance in the field dataset. Despite the application of SMOTE during training, the underlying distribution remained strongly skewed toward healthy plots (111 healthy plots compared to only 14 severe plots). This imbalance was also reflected in the validation dataset, where the 20% test set contained limited representation of severe conditions. As a result, the model exhibited a bias toward the dominant class, reducing sensitivity to severe but ecologically important dieback conditions. This limitation is consistent with Meddens et al. (2013) who demonstrated in the context of bark beetle-induced tree mortality that detection accuracy decreases as the proportion of affected trees within mixed pixels declines, highlighting the difficulty of identifying relatively rare disturbance signals in heterogeneous forest environments.

6.3 Random Forest Performance Improvement with RFE and Permutation Importance

The initial Random Forest model, trained using all 21 features, achieved an overall accuracy of approximately 60%, indicating that the full feature set included redundant and non-informative variables that reduced predictive performance. Following feature selection using RFE, supported by permutation importance analysis, model performance improved across all evaluation metrics, with the optimal configuration achieved using only seven features. This reduction to 7 predictors demonstrates that a substantial proportion of the original variables did not contribute additional discriminative information and instead introduced noise that weakened class separation. After feature selection, the overall accuracy of the model was improved by 10%, confirming that feature reduction substantially improved both classification performance and robustness.

Permutation importance analysis, further showed that 10 out of 21 features exhibited near zero or negative importance values, indicating limited independent contribution to the model. According to Genuer et al. (2010), this behaviour is typically associated with highly correlated or redundant predictors that encode similar information. This is consistent with multi-temporal remote sensing datasets, where vegetation indices often display strong intercorrelation and therefore do not contribute independently to model performance. This also explains why certain continuum removal variables (CR_RE_val and CR_SWIR_val), despite showing strong spectral separability in Section 5.3, appeared weak in permutation importance rankings. As discussed by Cutler et al. (2007), permutation based importance can underestimate the contribution of correlated predictors because their information is partially shared among related variables. Therefore, low or negative permutation importance values cannot be interpreted as lack of relevance, but rather as an indication of redundancy and shared explanatory power among correlated features.

Among all predictors, species composition emerged as the most influential variable, with its permutation causing the largest reduction in classification accuracy. This highlights the importance of forest structural composition in shaping spectral responses through differences in canopy architecture and phenology between coniferous and deciduous stands. Within the spectral variables, CR_SWIR_delta (Difference between the annual mean value and the 2017 baseline mean of CR_SWIR) ranked highest, confirming the strong sensitivity of SWIR-based continuum removal indices that were explained above. IRECI_val (Mean annual value of IRECI for a given

year) followed as the second most important spectral feature, reflecting its sensitivity to chlorophyll content and photosynthetic activity. Although NDVI showed limited sensitivity due to saturation effects in dense temperate forests, NDVI_delta still provided useful information by capturing canopy opening and defoliation during the transition from moderate to severe dieback. CR_SWIR_spatial_std (Standard deviation of all pixels within the plot buffer) ranked fifth, indicating that spatial variability in canopy moisture contributes additional discriminatory power. This reflects the heterogeneous structure of dieback affected stands, where healthy, moderate, and severely stressed trees coexist within short spatial distances. Spatial variability, therefore, captures fragmentation and moisture heterogeneity not represented by mean spectral values alone.

This observed variability is conceptually linked to texture characteristics in remote sensing imagery. Texture measures derived from the grey-level co-occurrence matrix (GLCM) describe the spatial organization of pixel values by quantifying how values co-occur within a defined neighborhood (Haralick et al., 1973; Haralick, 1979; Hall-Beyer, 2017). In this sense, CR_SWIR_spatial_std reflects the degree of heterogeneity but does not explicitly account for its spatial arrangement. The results therefore suggest that additional texture-based descriptors could further refine the characterization of canopy moisture patterns. Such information may improve the interpretation of within stand variability and provide complementary insight into the spatial structure of forest decline.

Other variables, including NDWI_delta and CR_RE temporal variability, showed lower importance due to redundancy with stronger predictors such as CR_SWIR_delta and IRECI_val.

The average model confidence across all predictions was approximately 68%, indicating moderate reliability. Higher confidence was generally observed for healthy and severe classes, while moderate dieback showed greater uncertainty due to spectral overlap between transitional canopy states. Similar patterns were reported by Mouret et al. (2024) who observed strong confusion in intermediate stress classes in multi-class dieback frameworks.

6.4 Error Matrix and Area Prediction Correction

Based on area calculations for the three dieback severity classes, 2018 showed the most pronounced shift in dieback severity, with reduced healthy forest extent and increased severe

dieback. This pattern is consistent with the severe 2018 European drought, which likely induced significant water stress and triggered widespread canopy decline across the study area.

The error-adjusted area estimates highlight the importance of incorporating classification uncertainty when interpreting spatial patterns of forest dieback. In particular, the notable underestimation of moderate dieback in 2019 indicates that transitional forest conditions are more difficult to map accurately, largely due to spectral overlap between healthy and moderately stressed stands. Without error correction, raw classification outputs tend to simplify the degradation gradient by assigning a substantial proportion of moderate conditions to the healthy class. The need for such correction is supported by Foody, (2002) who highlight that area estimates derived from classified maps can be systematically biased in the presence of non-random classification errors, particularly in transitional environments where class boundaries are diffuse.

The relatively small corrections required for severe dieback ($\sim 1\text{-}2\text{ km}^2$) indicate that advanced canopy decline produces stronger and more distinct spectral signals, allowing for more reliable classification. This supports previous findings that severe physiological stress is more readily detectable through remote sensing than early or intermediate decline stages. The improved calibration performance observed in 2021 further suggests that model reliability can vary substantially across years, potentially reflecting differences in environmental conditions, spectral separability, or class representation within the validation dataset.

6.5 Influence of Species Composition on Spectral Response

This study did not explicitly focus on species-level classification; instead, all tree types were integrated within a single modelling framework. At the plot level, the dataset included coniferous, deciduous, and mixed plots. However, exploratory analysis of the three dieback classes indicated that spectral behaviour varied with species composition, particularly in CR_SWIR and CR_RE indices. Conifer and mixed stands exhibited broadly similar temporal trajectories, while deciduous stands showed more distinct class separation. This pattern was evident in the continuum-removed indices, suggesting that species structure influences the expression of dieback-related spectral signals.

It is important to note that the dataset is strongly dominated by coniferous plots, with comparatively few deciduous observations. This imbalance limits the robustness of deciduous-

specific interpretations and may introduce bias into observed trends. While conifer plots are sufficiently represented across all dieback classes and therefore reflect more stable temporal trajectories, the deciduous subset is too small and unevenly distributed to support equally reliable generalizations. Nevertheless, severe deciduous stands still exhibited clear anomalies in CR_SWIR and CR_RE, closely matching patterns observed in conifers, indicating that the spectral response of advanced canopy desiccation is broadly consistent across forest types (Appendix 6).

Carletti et al. (2025) reported that species-specific dieback models in pure stands achieved moderate accuracies (60–70%) and highlighted the difficulty of early stage detection in mixed and heterogeneous forests. Similarly, Mouret et al. (2024) demonstrated that incorporating forest type information stabilizes classification boundaries by reducing confusion caused by structural and seasonal variability from mixed species canopies. However, they also noted that models developed in relatively homogeneous forest conditions are less transferable to mixed forests and edge environments, where spectral mixing is more pronounced.

Although this study did not implement a detailed species-specific modelling framework, the inclusion of a simplified species composition layer (conifer vs. deciduous vs. mixed) improved classification performance and helped stabilize dieback separation across classes. This suggests that even coarse representations of forest composition provide valuable structural context for interpreting spectral signals in multi-species landscapes.

7. Conclusion

This study evaluated the potential of Sentinel-2 multispectral time series imagery combined with machine learning techniques for detecting and monitoring forest dieback in the Doller Valley, located in the Vosges Mountains of northeastern France.

The results demonstrated that Sentinel-2 multispectral time series data provides valuable information for characterizing both temporal and spatial variations in forest condition. Seasonal and interannual variations in vegetation dynamics were successfully captured through spectral index analysis, enabling the identification of distinct temporal patterns associated with different dieback severity levels. Among the evaluated indices, shortwave-infrared based and red-edge indices, particularly CR_SWIR, CR_RE, and IRECI, showed the strongest sensitivity to canopy degradation and vegetation water stress, highlighting the importance of the SWIR, red-edge, and red spectral regions for forest dieback detection. The study further demonstrated that greenness alone can be misleading in dieback monitoring, as understory vegetation development following canopy decline may compensate for spectral signals and mask actual canopy deterioration. Consequently, traditional greenness based indices such as NDVI alone are insufficient for reliable dieback assessment. Statistical analyses further confirmed significant spectral differences between healthy and severely affected plots, although moderate dieback conditions remained more difficult to distinguish due to spectral overlap and transitional canopy characteristics.

The Random Forest classification framework successfully integrated spectral, temporal, and spatial predictor variables to map forest dieback severity across the study area. The developed model, designed to classify healthy, moderate, and severe dieback conditions, achieved an overall accuracy of 70%, with relatively high precision and recall across the three classes. Feature selection approaches, including RFE and permutation importance analysis, improved model efficiency by identifying the most informative predictors for dieback classification. In particular, the inclusion of temporal variability measures substantially enhanced model performance, emphasizing the importance of time series analysis for detecting gradual forest decline processes. The resulting annual dieback prediction maps revealed clear spatial and temporal variability in canopy decline throughout the 2018–2025 period. Overall, the findings highlight the strong potential of freely available Sentinel-2 imagery for operational forest health monitoring in temperate ecosystems.

Despite these contributions, several limitations should be acknowledged. The field dataset exhibited substantial class imbalance, with a high proportion of healthy plots and only a limited number of severe dieback plots. This imbalance reduced the availability of moderate and severe samples for model training and likely introduced bias toward the healthy class. In addition, the dataset was dominated by coniferous forest stands, which may have influenced the spectral behaviour captured by the model and limited the generalizability of the results across broader deciduous or mixed forest ecosystems. The medium spatial resolution of Sentinel-2 imagery also constrained the detection of fine-scale canopy decline and individual-tree stress responses. In heterogeneous forest environments, spectral mixing between healthy and declining vegetation further reduced class separability, particularly for moderate dieback conditions. Furthermore, the framework remained dependent on field-based calibration data for reliable classification and validation.

Future research should focus on integrating higher spatial and spectral resolution datasets, such as hyperspectral and LiDAR observations, to improve the detection of early stage dieback and reduce uncertainties associated with mixed pixels. The incorporation of additional climatic and topographic variables may further improve model robustness and enhance transferability across different forest ecosystems. Overall, this study demonstrates that the integration of Sentinel-2 multispectral time series imagery, field observations, and machine learning techniques provides an effective framework for monitoring forest dieback dynamics in temperate forest ecosystems and offers valuable insights for future remote sensing based forest health assessment approaches.

References

- Abdullah, H., Skidmore, A. K., Darvishzadeh, R., & Heurich, M. (2019). Timing of red-edge and shortwave infrared reflectance critical for early stress detection induced by bark beetle (*Ips typographus*, L.) attack. *International Journal of Applied Earth Observation and Geoinformation*, 82, 101900. <https://doi.org/10.1016/j.jag.2019.101900>
- Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evolutionary Applications*, 1(1), 95–111. <https://doi.org/10.1111/j.1752-4571.2007.00013.x>
- Allen, C. D. (2007). Interactions Across Spatial Scales among Forest Dieback, Fire, and Erosion in Northern New Mexico Landscapes. *Ecosystems*, 10(5), 797–808. <https://doi.org/10.1007/s10021-007-9057-4>
- Allen, C. D. (2009). *Climate-induced forest dieback: An escalating global phenomenon?* *Unasylva*, 60, 43–49.
- Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H. (Ted), Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S. W., Semerci, A., & Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, 259(4), 660–684. <https://doi.org/10.1016/j.foreco.2009.09.001>
- Amoah-Nuamah, J., Child, B., Okyere, E. Y., Adams, O., & Danquah, J. A. (2025). Applications of artificial intelligence in forest health surveillance and management. *Discover Forests*, 1(1), 56. <https://doi.org/10.1007/s44415-025-00061-w>
- Anderegg, W. R. L., Anderegg, L. D. L., & Huang, C. (2019). Testing early warning metrics for drought-induced tree physiological stress and mortality. *Global Change Biology*, 25(7), 2459–2469. <https://doi.org/10.1111/gcb.14655>
- Anderegg, W. R. L., Hicke, J. A., Fisher, R. A., Allen, C. D., Aukema, J., Bentz, B., Hood, S., Lichstein, J. W., Macalady, A. K., McDowell, N., Pan, Y., Raffa, K., Sala, A., Shaw, J. D., Stephenson, N. L., Tague, C., & Zeppel, M. (2015). Tree mortality from drought, insects, and their interactions in a changing climate. *New Phytologist*, 208(3), 674–683. <https://doi.org/10.1111/nph.13477>

- Anderegg, W. R. L., Kane, J. M., & Anderegg, L. D. L. (2013). Consequences of widespread tree mortality triggered by drought and temperature stress. *Nature Climate Change*, 3(1), 30–36. <https://doi.org/10.1038/nclimate1635>
- Andresini, G., Appice, A., & Malerba, D. (2024). Leveraging Sentinel-2 time series for bark beetle-induced forest dieback inventory. *Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing*, 875–882. <https://doi.org/10.1145/3605098.3635908>
- Anwander, J., Brandmeier, M., Paczkowski, S., Neubert, T., & Paczkowska, M. (2024). Evaluating Different Deep Learning Approaches for Tree Health Classification Using High-Resolution Multispectral UAV Data in the Black Forest, Harz Region, and Göttinger Forest. *Remote Sensing*, 16(3), 561. <https://doi.org/10.3390/rs16030561>
- Arthur, G., Jonathan, L., Juliette, C., Nicolas, L., Christian, P., & Hugues, C. (2024). Spatial and remote sensing monitoring shows the end of the bark beetle outbreak on Belgian and north-eastern France Norway spruce (*Picea abies*) stands. *Environmental Monitoring and Assessment*, 196(3), 226. <https://doi.org/10.1007/s10661-024-12372-0>
- Asner, G. P., Martin, R. E., Anderson, C. B., & Knapp, D. E. (2015). Quantifying forest canopy traits: Imaging spectroscopy versus field survey. *Remote Sensing of Environment*, 158, 15–27. <https://doi.org/10.1016/j.rse.2014.11.011>
- Bárta, V., Lukeš, P., & Homolová, L. (2021). Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. *International Journal of Applied Earth Observation and Geoinformation*, 100, 102335. <https://doi.org/10.1016/j.jag.2021.102335>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Bos, A., Lamarche, C., Niro, F., & Defourny, P. (2025). Monitoring biweekly dynamics of pan-tropical industrial plantations over 6 years using 100 m PROBA-V data. *Frontiers in Remote Sensing*, 6, 1575100. <https://doi.org/10.3389/frsen.2025.1575100>
- Buras, A., Rammig, A., & Zang, C. S. (2021). The European Forest Condition Monitor: Using Remotely Sensed Forest Greenness to Identify Hot Spots of Forest Decline. *Frontiers in Plant Science*, 12, 689220. <https://doi.org/10.3389/fpls.2021.689220>

- Camarero, J. J., Gazol, A., Sangüesa-Barreda, G., Oliva, J., & Vicente-Serrano, S. M. (2015). To die or not to die: Early warnings of tree dieback in response to a severe drought. *Journal of Ecology*, 103(1), 44–57. <https://doi.org/10.1111/1365-2745.12295>
- Campbell, M. J., Dennison, P. E., Tune, J. W., Kannenberg, S. A., Kerr, K. L., Coddling, B. F., & Anderegg, W. R. L. (2020). A multi-sensor, multi-scale approach to mapping tree mortality in woodland ecosystems. *Remote Sensing of Environment*, 245, 111853. <https://doi.org/10.1016/j.rse.2020.111853>
- Candotti, A., De Giglio, M., Dubbini, M., & Tomelleri, E. (2022). A Sentinel-2 Based Multi-Temporal Monitoring Framework for Wind and Bark Beetle Detection and Damage Mapping. *Remote Sensing*, 14(23), 6105. <https://doi.org/10.3390/rs14236105>
- Carletti, H., Gégout, J.-C., Dutrieux, R., Féret, J.-B., Vega, C., Belouard, T., Jolly, A., Cansell, J., & Piedallu, C. (2025). *Sentinel-2 Time Series Reveal Species-Specific Responses in Temperate Conifer Dieback*. SSRN. <https://doi.org/10.2139/ssrn.5120221>
- Ceccato, P., Flasse, S., & Gregoire, J.-M. (2002). *Designing a spectral index to estimate vegetation water content from remote sensing data Part 2. Validation and applications*.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Chiang, C.-Y., Barnes, C., Angelov, P., & Jiang, R. (2020). Deep Learning-Based Automated Forest Health Diagnosis From Aerial Images. *IEEE Access*, 8, 144064–144076. <https://doi.org/10.1109/ACCESS.2020.3012417>
- Clark, R. N., & Roush, T. L. (1984). Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research: Solid Earth*, 89(B7), 6329–6340. <https://doi.org/10.1029/JB089iB07p06329>
- Clevers, J. G. P. W., & Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *International Journal of Applied Earth Observation and Geoinformation*, 23, 344–351. <https://doi.org/10.1016/j.jag.2012.10.008>
- Curran, P. J., Dungan, J. L., & Peterson, D. L. (2001). *Estimating the foliar biochemical concentration of leaves with reflectance spectrometry Testing the Kokaly and Clark methodologies*.

- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). RANDOM FORESTS FOR CLASSIFICATION IN ECOLOGY. *Ecology*, 88(11), 2783–2792. <https://doi.org/10.1890/07-0539.1>
- Czaplewski, R. L., & Catts, G. P. (1992). Calibration of remotely sensed proportion or area estimates for misclassification error. *Remote Sensing of Environment*, 39(1), 29–43. [https://doi.org/10.1016/0034-4257\(92\)90138-A](https://doi.org/10.1016/0034-4257(92)90138-A)
- Dainelli, R., Toscano, P., Di Gennaro, S. F., & Matese, A. (2021). Recent Advances in Unmanned Aerial Vehicle Forest Remote Sensing—A Systematic Review. Part I: A General Framework. *Forests*, 12(3), 327. <https://doi.org/10.3390/f12030327>
- Dinno, A. (2015). Nonparametric Pairwise Multiple Comparisons in Independent Groups using Dunn’s Test. *The Stata Journal: Promoting Communications on Statistics and Stata*, 15(1), 292–300. <https://doi.org/10.1177/1536867X1501500117>
- Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. *Technometrics*, 6(3), 241–252. <https://doi.org/10.1080/00401706.1964.10490181>
- Eriksson, H. M., Eklundh, L., Kuusk, A., & Nilson, T. (2006). Impact of understory vegetation on forest canopy reflectance and remotely sensed LAI estimates. *Remote Sensing of Environment*, 103(4), 408–418. <https://doi.org/10.1016/j.rse.2006.04.005>
- Fassnacht, F. E., White, J. C., Wulder, M. A., & Næsset, E. (2024). Remote sensing in forestry: Current challenges, considerations and directions. *Forestry: An International Journal of Forest Research*, 97(1), 11–37. <https://doi.org/10.1093/forestry/cpad024>
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Frampton, W. J., Dash, J., Watmough, G., & Milton, E. J. (2013). Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 83–92. <https://doi.org/10.1016/j.isprsjprs.2013.04.007>
- Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)

- Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, 31(14), 2225–2236.
<https://doi.org/10.1016/j.patrec.2010.03.014>
- Grabska, E., Frantz, D., & Ostapowicz, K. (2020). Evaluation of machine learning algorithms for forest stand species mapping using Sentinel-2 imagery and environmental data in the Polish Carpathians. *Remote Sensing of Environment*, 251, 112103.
<https://doi.org/10.1016/j.rse.2020.112103>
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Haralick, R. (1979). *Statistical and structural approaches to texture. Proceedings of the IEEE*, 67(5), 786–804. <https://doi.org/10.1109/proc.1979.11328>.
- Hawryło, P., Bednarz, B., Wężyk, P., & Szostak, M. (2018). Estimating defoliation of Scots pine stands using machine learning methods and vegetation indices of Sentinel-2. *European Journal of Remote Sensing*, 51(1), 194–204.
<https://doi.org/10.1080/22797254.2017.1417745>
- Hay, A. M. (1988). The derivation of global estimates from a confusion matrix. *International Journal of Remote Sensing*, 9(8), 1395–1398.
<https://doi.org/10.1080/01431168808954945>
- Hermosilla, T., Bastyr, A., Coops, N. C., White, J. C., & Wulder, M. A. (2022). Mapping the presence and distribution of tree species in Canada's forested ecosystems. *Remote Sensing of Environment*, 282, 113276. <https://doi.org/10.1016/j.rse.2022.113276>
- Hou, R., Zhang, B., Fang, G., Yang, S., Guo, L., Huang, W., Yao, J., Jiao, Q., Sun, H., & Yan, J. (2025). Early Detection of Pine Wilt Disease by Combining Pigment and Moisture Content Indices Using UAV-Based Hyperspectral Imagery. *Remote Sensing*, 17(11), 1833. <https://doi.org/10.3390/rs17111833>
- Huete, A. R. (2012). Vegetation Indices, Remote Sensing and Forest Monitoring. *Geography Compass*, 6(9), 513–532. <https://doi.org/10.1111/j.1749-8198.2012.00507.x>
- Huo, L., Persson, H. J., & Lindberg, E. (2021). Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR

- (NDRS). *Remote Sensing of Environment*, 255, 112240.
<https://doi.org/10.1016/j.rse.2020.112240>
- Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., & Rosette, J. (2019). Structure from Motion Photogrammetry in Forestry: A Review. *Current Forestry Reports*, 5(3), 155–168.
<https://doi.org/10.1007/s40725-019-00094-3>
- Ippc. (2022). *Global Warming of 1.5°C: IPCC Special Report on Impacts of Global Warming of 1.5°C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009157940>
- K., E., & M.R., G. (2008). *Participatory monitoring in tropical forest management: A review of tools, concepts and lessons learned*. Center for International Forestry Research (CIFOR).
<https://doi.org/10.17528/cifor/002486>
- Kane, J. M., Meinhardt, K. A., Chang, T., Cardall, B. L., Michalet, R., & Whitham, T. G. (2011). Drought-induced mortality of a foundation species (*Juniperus monosperma*) promotes positive afterlife effects in understory vegetation. *Plant Ecology*, 212(5), 733–741.
<https://doi.org/10.1007/s11258-010-9859-x>
- Kautz, M., Feurer, J., & Adler, P. (2024). Early detection of bark beetle (*Ips typographus*) infestations by remote sensing – A critical review of recent research. *Forest Ecology and Management*, 556, 121595. <https://doi.org/10.1016/j.foreco.2023.121595>
- Khan, R. W. A., Shaheen, H., Islam Dar, M. E. U., Habib, T., Manzoor, M., Gillani, S. W., Al-Andal, A., Ayoola, J. O., & Waheed, M. (2025). A data-driven approach to forest health assessment through multivariate analysis and machine learning techniques. *BMC Plant Biology*, 25(1), 915. <https://doi.org/10.1186/s12870-025-06937-5>
- Knutzen, F., Auerbeck, P., Barrasso, C., Bouwer, L. M., Gardiner, B., Grünzweig, J. M., Hänel, S., Haustein, K., Johannessen, M. R., Kollet, S., Müller, M. M., Pietikäinen, J.-P., Pietras-Couffignal, K., Pinto, J. G., Rechid, D., Rousi, E., Russo, A., Suarez-Gutierrez, L., Veit, S., ... Gliksman, D. (2025). Impacts on and damage to European forests from the 2018–2022 heat and drought events. *Natural Hazards and Earth System Sciences*, 25(1), 77–117. <https://doi.org/10.5194/nhess-25-77-2025>

- Kokaly, R. (1999). Spectroscopic Determination of Leaf Biochemistry Using Band-Depth Analysis of Absorption Features and Stepwise Multiple Linear Regression. *Remote Sensing of Environment*, 67(3), 267–287. [https://doi.org/10.1016/S0034-4257\(98\)00084-4](https://doi.org/10.1016/S0034-4257(98)00084-4)
- Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260), 583–621. <https://doi.org/10.1080/01621459.1952.10483441>
- Lausch, A., Erasmi, S., King, D., Magdon, P., & Heurich, M. (2017). Understanding Forest Health with Remote Sensing-Part II—A Review of Approaches and Data Models. *Remote Sensing*, 9(2), 129. <https://doi.org/10.3390/rs9020129>
- Liu, F., Liu, H., Xu, C., Shi, L., Zhu, X., Qi, Y., & He, W. (2021). Old-growth forests show low canopy resilience to droughts at the southern edge of the taiga. *Global Change Biology*, 27(11), 2392–2402. <https://doi.org/10.1111/gcb.15605>
- Lu, X., & Zheng, G. (2016). Assessing the effects of understory to forest canopy leaf area index by combining moderate resolution data and geometric optical (GO) model in temperate forest. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 1318–1321. <https://doi.org/10.1109/IGARSS.2016.7729335>
- Ma, Y., Zhao, Y., Im, J., Zhao, Y., & Zhen, Z. (2024). A deep-learning-based tree species classification for natural secondary forests using unmanned aerial vehicle hyperspectral images and LiDAR. *Ecological Indicators*, 159, 111608. <https://doi.org/10.1016/j.ecolind.2024.111608>
- Malenovský, Z., Ufer, C., Lhotáková, Z., Clevers, J. G. P. W., Schaepman, M. E., Albrechtová, J., & Cudlín, P. (2006). A new hyperspectral index for chlorophyll estimation of a forest canopy: Area under curve normalised to maximal band depth between 650–725 nm. *Zurich Open Repository and Archive (University of Zurich)*, 5(2), 161–172. <https://doi.org/10.5167/uzh-62112>
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), 50–60. <https://doi.org/10.1214/aoms/1177730491>
- Marusig, D., Petruzzellis, F., Tomasella, M., Napolitano, R., Altobelli, A., & Nardini, A. (2020). Correlation of Field-Measured and Remotely Sensed Plant Water Status as a Tool to

- Monitor the Risk of Drought-Induced Forest Decline. *Forests*, 11(1), 77.
<https://doi.org/10.3390/f11010077>
- Matsushita, B., Yang, W., Chen, J., Onda, Y., & Qiu, G. (2007). Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Case Study in High-density Cypress Forest. *Sensors*, 7(11), 2636–2651. <https://doi.org/10.3390/s7112636>
- Meddens, A. J. H., Hicke, J. A., Vierling, L. A., & Hudak, A. T. (2013). Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. *Remote Sensing of Environment*, 132, 49–58.
<https://doi.org/10.1016/j.rse.2013.01.002>
- Mgelwa, A. S., Ngaba, M. J. Y., Hu, B., Gurmesa, G. A., Mwakaje, A. G., Nyemeck, M. P. B., Zhu, F., Qiu, Q., Song, L., Wang, Y., Fang, Y., & Rennenberg, H. (2025). Meta-analysis of 21st century studies shows that deforestation induces profound changes in soil characteristics, particularly soil organic carbon accumulation. *Forest Ecosystems*, 12, 100257. <https://doi.org/10.1016/j.fecs.2024.100257>
- Michelucci, U., & Venturini, F. (2026). The infinite-dimensional nature of spectroscopy and why models succeed, fail, and mislead. *The Analyst*, 151(10), 2978–2994.
<https://doi.org/10.1039/D6AN00346J>
- Migas-Mazur, R., Kycko, M., Zwijacz-Kozica, T., & Zagajewski, B. (2021). Assessment of Sentinel-2 Images, Support Vector Machines and Change Detection Algorithms for Bark Beetle Outbreaks Mapping in the Tatra Mountains. *Remote Sensing*, 13(16), 3314.
<https://doi.org/10.3390/rs13163314>
- Moreno-Fernández, D., Viana-Soto, A., Camarero, J. J., Zavala, M. A., Tijerín, J., & García, M. (2021). Using spectral indices as early warning signals of forest dieback: The case of drought-prone *Pinus pinaster* forests. *Science of The Total Environment*, 793, 148578.
<https://doi.org/10.1016/j.scitotenv.2021.148578>
- Mouret, F., Morin, D., Martin, H., Planells, M., & Vincent-Barbaroux, C. (2024). Toward an Operational Monitoring of Oak Dieback With Multispectral Satellite Time Series: A Case Study in Centre-Val De Loire Region of France. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 643–659.
<https://doi.org/10.1109/JSTARS.2023.3332420>

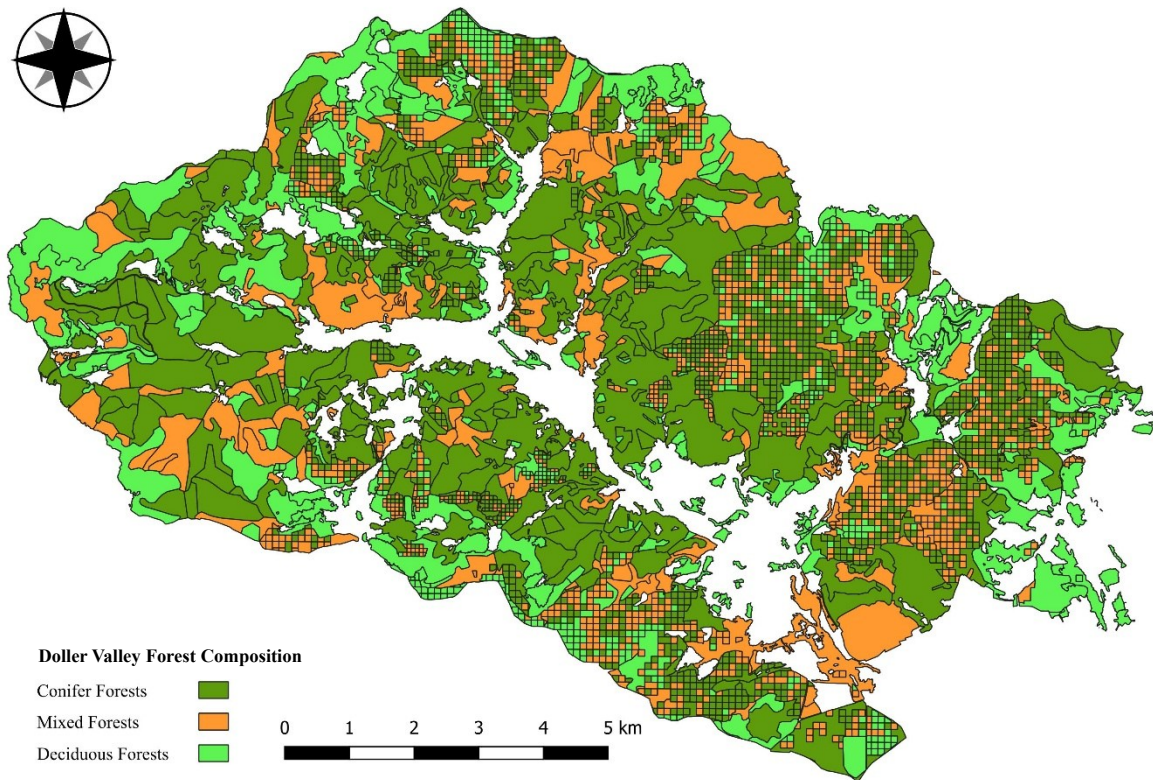
- Neumann, M., Mues, V., Moreno, A., Hasenauer, H., & Seidl, R. (2017). Climate variability drives recent tree mortality in Europe. *Global Change Biology*, 23(11), 4788–4797. <https://doi.org/10.1111/gcb.13724>
- Peng, C., Ma, Z., Lei, X., Zhu, Q., Chen, H., Wang, W., Liu, S., Li, W., Fang, X., & Zhou, X. (2011). A drought-induced pervasive increase in tree mortality across Canada's boreal forests. *Nature Climate Change*, 1(9), 467–471. <https://doi.org/10.1038/nclimate1293>
- Piedallu, C., Dallery, D., Bresson, C., Legay, M., Gégout, J.-C., & Pierrat, R. (2023). Spatial vulnerability assessment of silver fir and Norway spruce dieback driven by climate warming. *Landscape Ecology*, 38(2), 341–361. <https://doi.org/10.1007/s10980-022-01570-1>
- Pontius, J., & Hallett, R. (2014). Comprehensive Methods for Earlier Detection and Monitoring of Forest Decline. *Forest Science*, 60(6), 1156–1163. <https://doi.org/10.5849/forsci.13-121>
- Querejeta, J. I., Estrada-Medina, H., Allen, M. F., & Jiménez-Osornio, J. J. (2007). Water source partitioning among trees growing on shallow karst soils in a seasonally dry tropical climate. *Oecologia*, 152(1), 26–36. <https://doi.org/10.1007/s00442-006-0629-3>
- Raparelli, E., & Bajocco, S. (2019). A bibliometric analysis on the use of unmanned aerial vehicles in agricultural and forestry studies. *International Journal of Remote Sensing*, 40(24), 9070–9083. <https://doi.org/10.1080/01431161.2019.1569793>
- Scharnweber, T., Smiljanic, M., Cruz-García, R., Manthey, M., & Wilmking, M. (2020). Tree growth at the end of the 21st century—The extreme years 2018/19 as template for future growth conditions. *Environmental Research Letters*, 15(7), 074022. <https://doi.org/10.1088/1748-9326/ab865d>
- Schuldt, B., Buras, A., Arend, M., Vitasse, Y., Beierkuhnlein, C., Damm, A., Gharun, M., Grams, T. E. E., Hauck, M., Hajek, P., Hartmann, H., Hiltbrunner, E., Hoch, G., Holloway-Phillips, M., Körner, C., Larysch, E., Lübke, T., Nelson, D. B., Rammig, A., ... Kahmen, A. (2020). A first assessment of the impact of the extreme 2018 summer drought on Central European forests. *Basic and Applied Ecology*, 45, 86–103. <https://doi.org/10.1016/j.baae.2020.04.003>
- Senf, C., Pflugmacher, D., Zhiqiang, Y., Sebal, J., Knorn, J., Neumann, M., Hostert, P., & Seidl, R. (2018). Canopy mortality has doubled in Europe's temperate forests over the last three

- decades. *Nature Communications*, 9(1), 4978. <https://doi.org/10.1038/s41467-018-07539-6>
- Senf, C., & Seidl, R. (2020). Mapping the forest disturbance regimes of Europe. *Nature Sustainability*, 4(1), 63–70. <https://doi.org/10.1038/s41893-020-00609-y>
- Shapiro, S. S., & Wilk, M. B. (1965). *An analysis of variance test for normality (complete samples)*. *Biometrika*, 52(3–4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>.
- Suarez, M. L., & Kitzberger, T. (2008). Recruitment patterns following a severe drought: Long-term compositional shifts in Patagonian forests. *Canadian Journal of Forest Research*, 38(12), 3002–3010. <https://doi.org/10.1139/X08-149>
- Sylvester, J. M., Gutiérrez-Zapata, D. M., Pérez-Marulanda, L., Vanegas-Cubillos, M., Bruun, T. B., Mertz, O., & Castro-Nunez, A. (2024). Analysis of food system drivers of deforestation highlights foreign direct investments and urbanization as threats to tropical forests. *Scientific Reports*, 14(1), 15179. <https://doi.org/10.1038/s41598-024-65397-3>
- Taccoen, A. (2019). *Détermination de l'impact potentiel du changement climatique sur la mortalité des principales essences forestières européennes*.
- Hall-Beyer, M. (2017). *GLCM texture: A tutorial (Version 3.0)*. <https://doi.org/10.13140/RG.2.2.12424.21767>.
- Torres, P., Rodes-Blanco, M., Viana-Soto, A., Nieto, H., & García, M. (2021). The Role of Remote Sensing for the Assessment and Monitoring of Forest Health: A Systematic Evidence Synthesis. *Forests*, 12(8), 1134. <https://doi.org/10.3390/f12081134>
- Trumbore, S., Brando, P., & Hartmann, H. (2015). Forest health and global change. *Science*, 349(6250), 814–818. <https://doi.org/10.1126/science.aac6759>
- Tucker, C. J. (n.d.). *Red and Photographic Infrared Linear Combinations for Monitoring Vegetation*.
- Van Mantgem, P. J., Stephenson, N. L., Byrne, J. C., Daniels, L. D., Franklin, J. F., Fulé, P. Z., Harmon, M. E., Larson, A. J., Smith, J. M., Taylor, A. H., & Veblen, T. T. (2009). Widespread Increase of Tree Mortality Rates in the Western United States. *Science*, 323(5913), 521–524. <https://doi.org/10.1126/science.1165000>
- Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J. P., Veroustraete, F., Clevers, J. G. P. W., & Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation

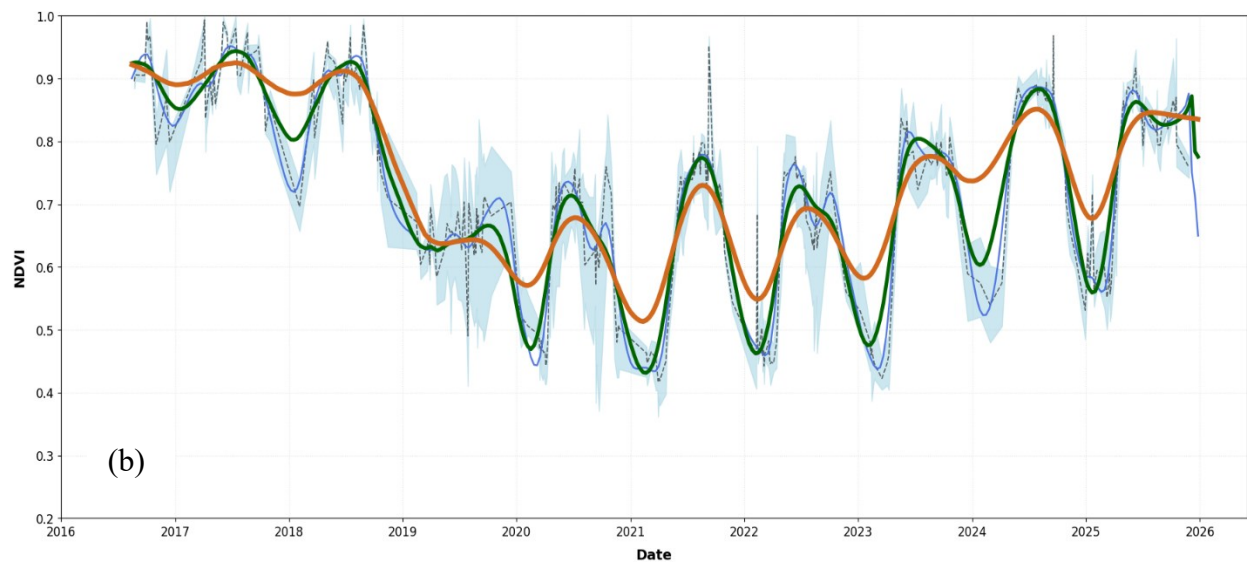
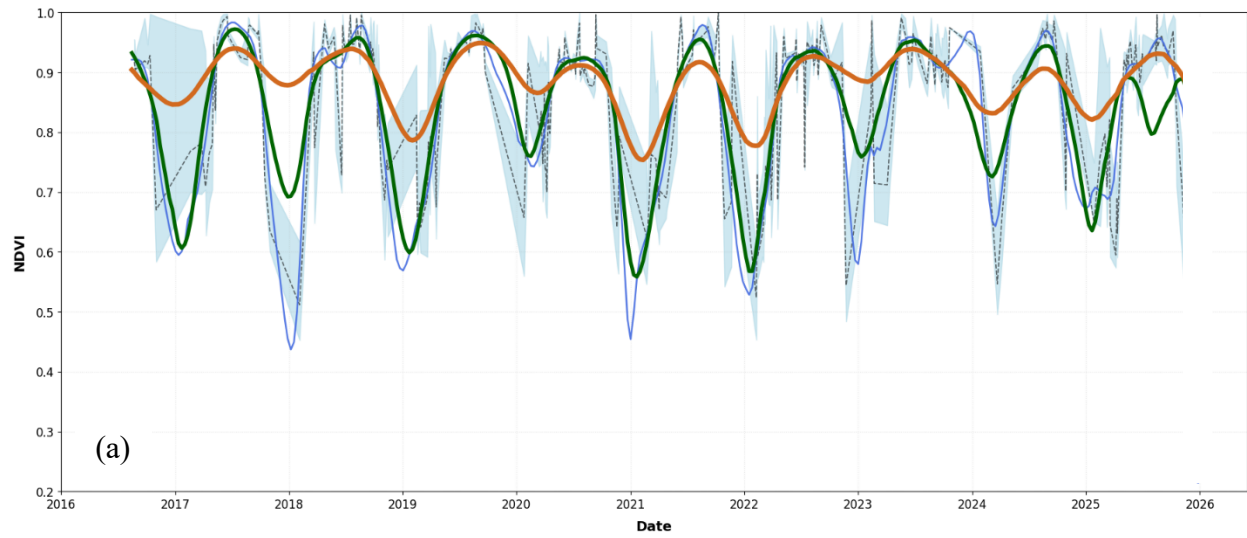
- bio-geophysical properties – A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 108, 273–290. <https://doi.org/10.1016/j.isprsjprs.2015.05.005>
- Whittaker, J. M. (1927). On the Cardinal Function of Interpolation Theory. *Proceedings of the Edinburgh Mathematical Society*, 1(1), 41–46.
<https://doi.org/10.1017/S0013091500007318>
- Ye, S., Rogan, J., Zhu, Z., Hawbaker, T. J., Hart, S. J., Andrus, R. A., Meddens, A. J. H., Hicke, J. A., Eastman, J. R., & Kulakowski, D. (2021). Detecting subtle change from dense Landsat time series: Case studies of mountain pine beetle and spruce beetle disturbance. *Remote Sensing of Environment*, 263, 112560. <https://doi.org/10.1016/j.rse.2021.112560>
- Zarco-Tejada, P. J., Hornero, A., Hernández-Clemente, R., & Beck, P. S. A. (2018). Understanding the temporal dimension of the red-edge spectral region for forest decline detection using high-resolution hyperspectral and Sentinel-2a imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 137, 134–148.
<https://doi.org/10.1016/j.isprsjprs.2018.01.017>

Appendices

Appendix 1. Species composition map of Doller Valley



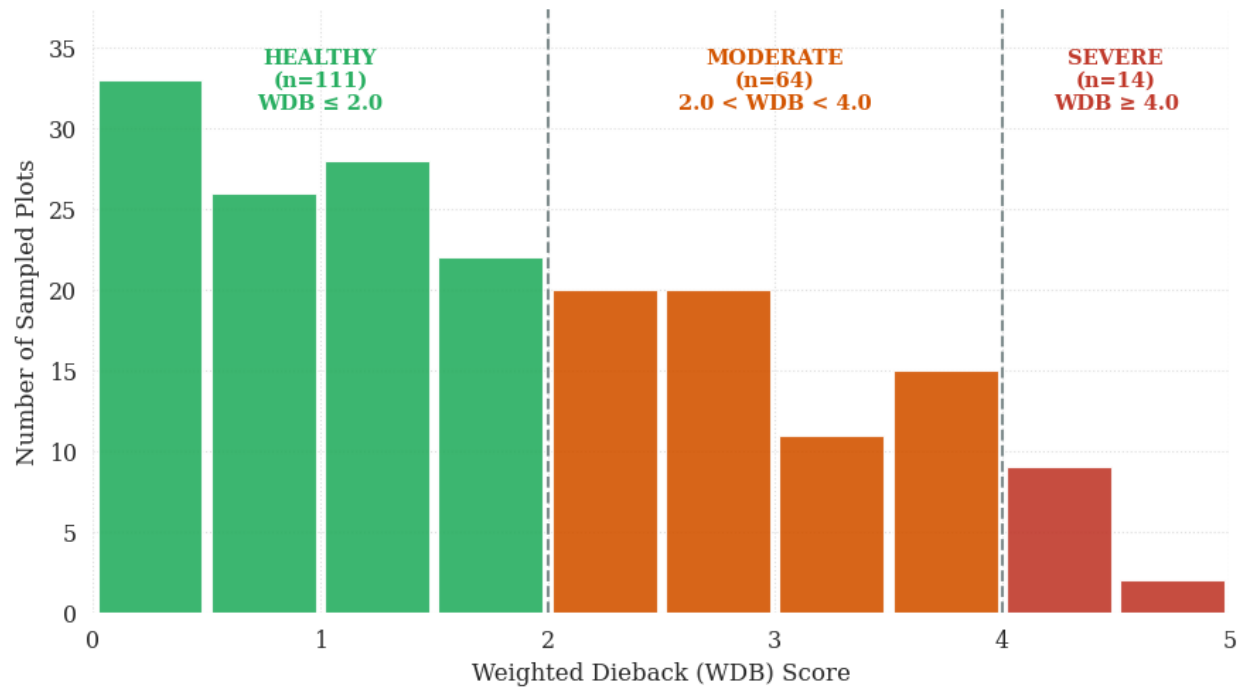
Appendix 2. Identification of the suitable smoothing parameter (λ) using a (a) Healthy Plot and a (b) Severe Plot



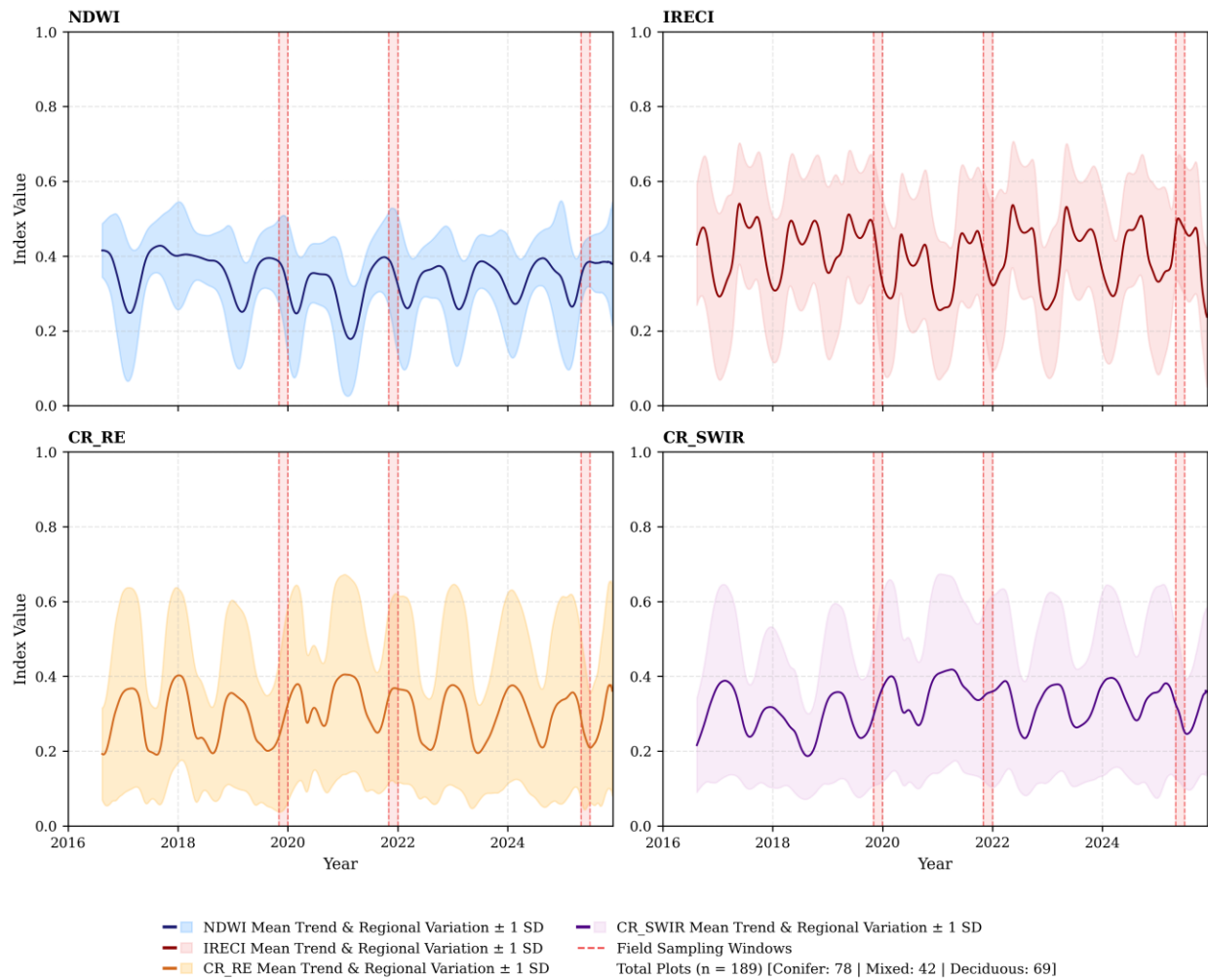
Appendix 3. Initial 21 Predictor features

Predictor Feature	Explanation
1. NDVI_val	Mean annual value of the spectral index for a given year
2. NDWI_val	
3. IRECI_val	
4. CR_RE_val	
5. CR_SWIR_val	
6. NDVI_delta	Difference between the annual mean value and the 2017 baseline mean
7. NDWI_delta	
8. IRECI_delta	
9. CR_RE_delta	
10. CR_SWIR_delta	
11. NDVI_spatial_std	Standard deviation of all pixels within the plot buffer (10m), representing spatial variability
12. NDWI_spatial_std	
13. IRECI_spatial_std	
14. CR_RE_spatial_std	
15. CR_SWIR_spatial_std	
16. NDVI_temp_std	Standard deviation of a pixel's values over years, representing temporal variability
17. NDWI_temp_std	
18. IRECI_temp_std	
19. CR_RE_temp_std	
20. CR_SWIR_temp_std	
21. Species Composition	Categorizing conifer, deciduous, and mixed plots

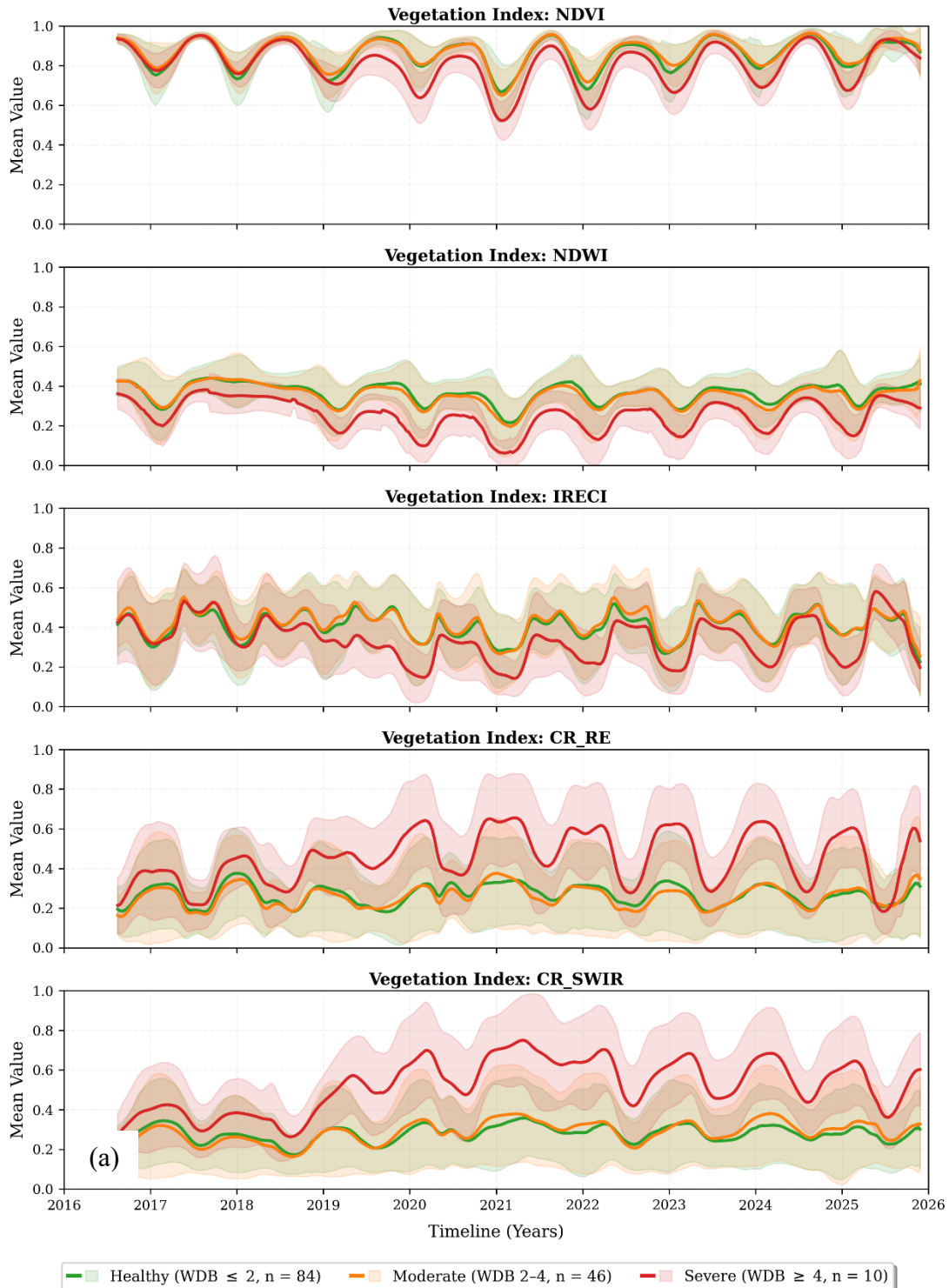
Appendix 4. Histogram showing the distribution of sampled plots across dieback classes

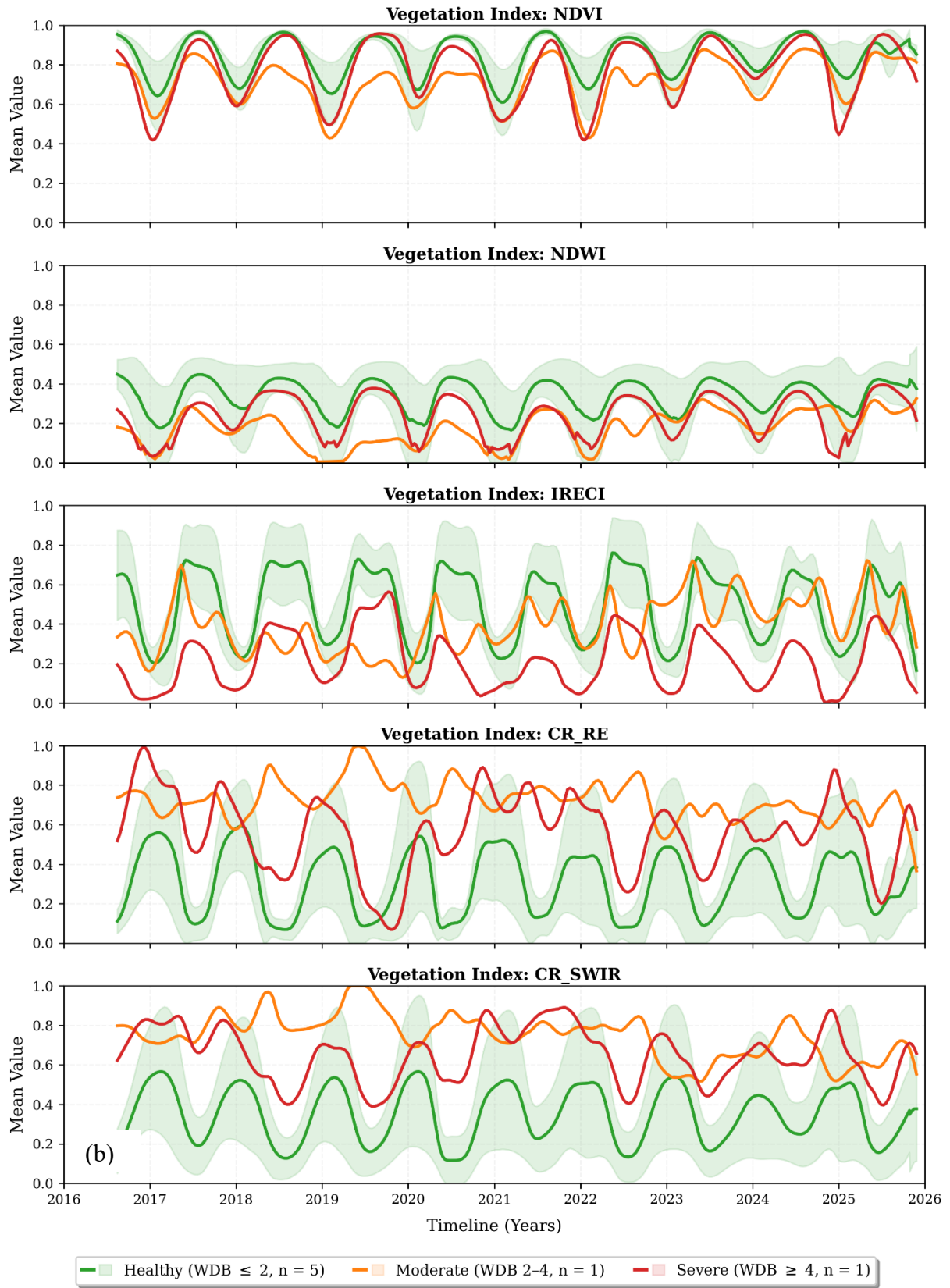


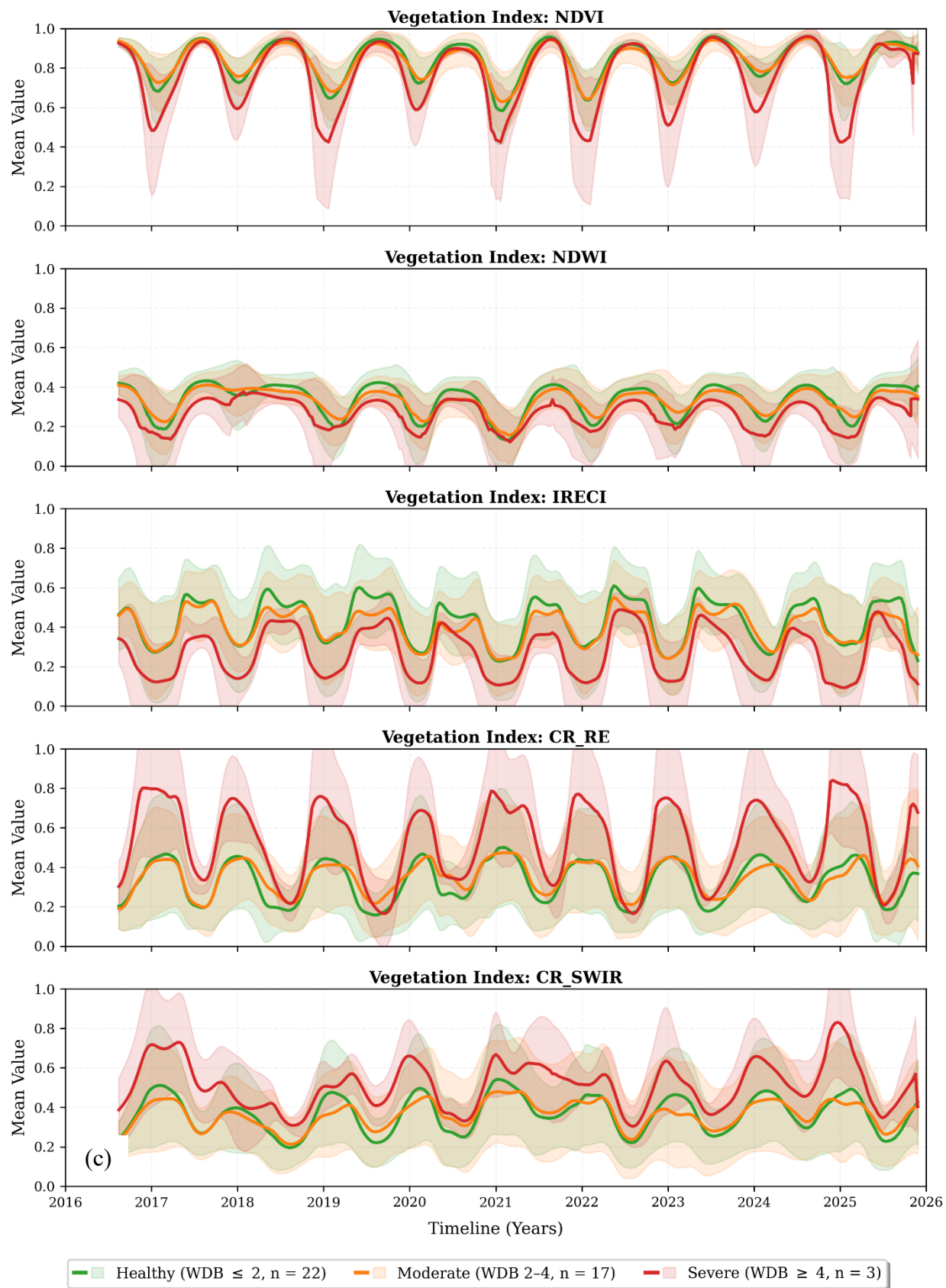
Appendix 5. Time series visualizations of vegetation indices (NDWI, IRECI, CR_RE, and CR_SWIR)



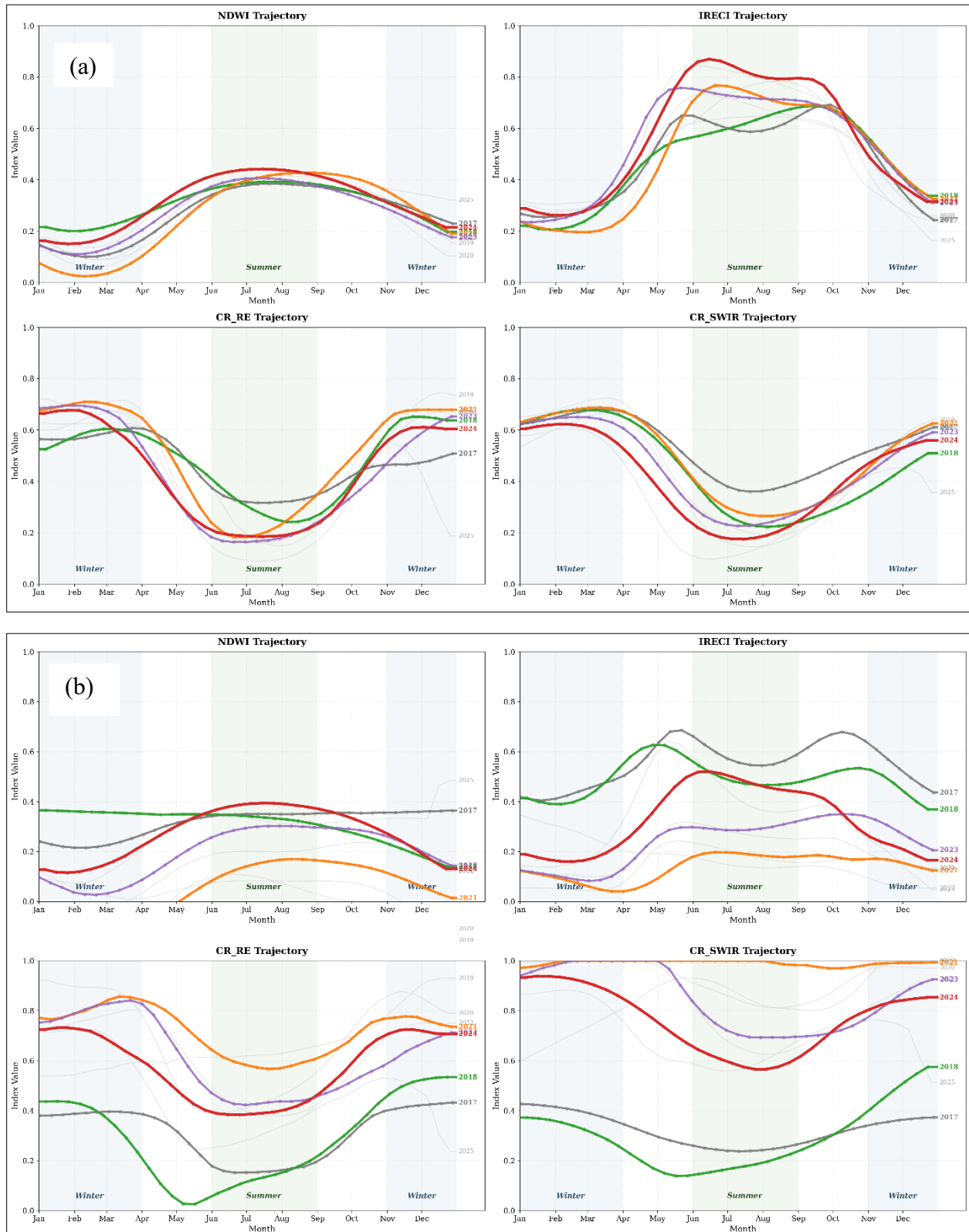
Appendix 6. Species composition-specific time series visualizations for the three dieback classes: (a) Conifer plots, (b) Deciduous Plots, and (c) Mixed Plots



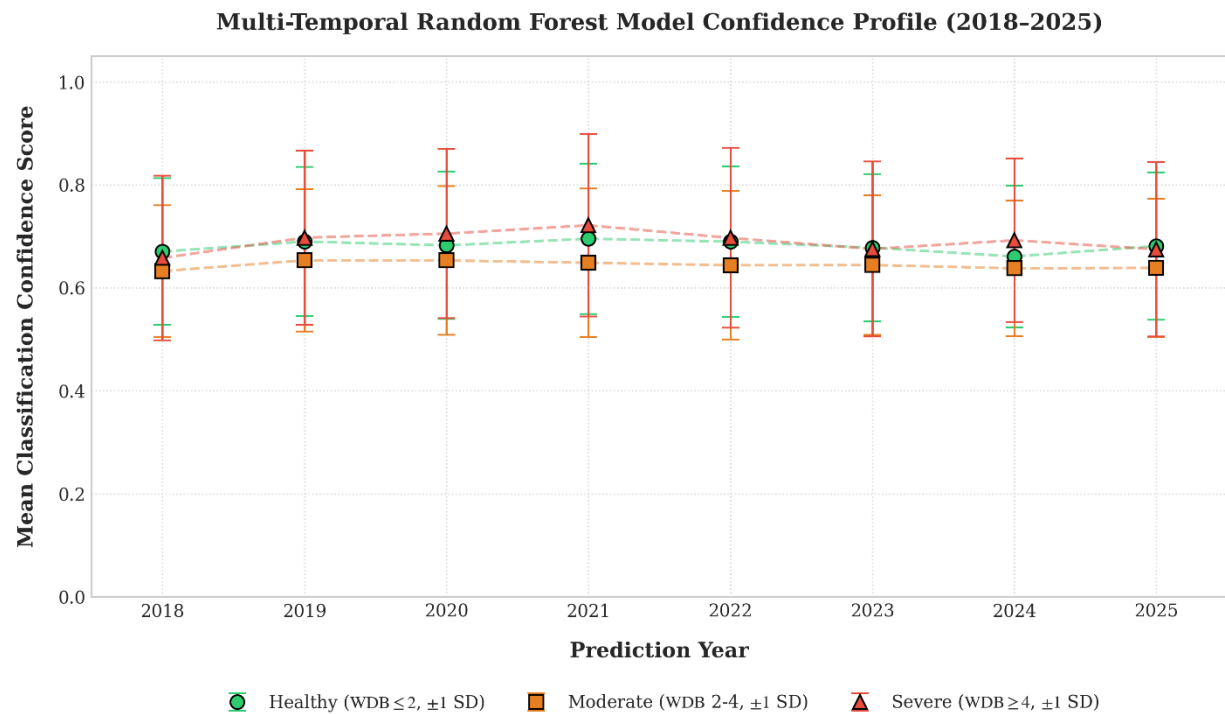




Appendix 7. Temporal evolution of dieback in a (a) Healthy plot and a (b) Severely affected plot



Appendix 8. Annual model confidence variation



Multispectral Imagery for Forest Conditions Monitoring of a Pilot Watershed in the Vosges, France

Janathi Wathsala Rathnayaka

Forest ecosystems are increasingly affected by climate-induced stressors, resulting in progressive forest dieback and declining ecosystem resilience. In this context, efficient large scale monitoring approaches are essential for supporting sustainable forest management and climate adaptation strategies. This study evaluated the potential of Sentinel-2 multispectral time-series imagery combined with machine learning techniques for detecting and monitoring forest dieback in the Doller Valley, France, while also identifying the most informative Sentinel-2 spectral properties associated with dieback dynamics. Sentinel-2 imagery acquired between 2016 and 2025 was integrated with field observations collected from 189 forest plots. A plot-level Weighted Dieback (WDB) score was derived from field measurements to represent canopy decline severity. Vegetation indices; NDVI, NDWI, IRECI, CR_RE, and CR_SWIR were calculated, from which spatial and temporal variability metrics were extracted. These features, together with WDB scores, were incorporated into a Random Forest (RF) classification framework for healthy, moderate, and severe dieback classes. The final RF model achieved an overall accuracy of 70%, with F1-scores of 0.75, 0.61, and 0.75 for healthy, moderate, and severe classes, respectively. The results demonstrated that the shortwave infrared (SWIR), red-edge, and red spectral regions were the most informative for forest dieback detection, with CR_SWIR, CR_RE, and IRECI showing the strongest sensitivity to canopy chlorophyll degradation and vegetation water stress. In contrast, NDVI alone was found insufficient for reliable dieback detection, as understory vegetation could compensate for canopy loss and mask actual forest decline. Statistical analyses revealed clear spectral separability between healthy and severely affected plots, whereas moderate dieback conditions exhibited substantial spectral overlap with healthy forests. However, the inclusion of temporal variability and spatial heterogeneity metrics improved the detection of moderate dieback conditions, highlighting the importance of time-series information for identifying gradual canopy decline processes. Annual dieback prediction maps further revealed pronounced spatial and temporal variability between 2018 and 2025, with the highest severe dieback extent observed in 2018, followed by a gradual reduction toward 2024. Overall, the findings demonstrate that Sentinel-2 imagery combined with machine learning provides an effective and scalable framework for forest dieback monitoring in forest ecosystems. The developed approach can be further improved through the integration of higher-resolution datasets and application across different forest types and ecological regions, contributing toward more generalized and transferable remote sensing based forest health monitoring frameworks.

Keywords: Climate Stress, Forest Dieback, Random Forest, Sentinel-2, Time Series

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Faculté des bioingénieurs

Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/agro