ORIGINAL ARTICLE

Uncovering Early Traces of Bark Beetle Induced Forest Stress via Semantically Enriched Sentinel-2 Data and Spectral Indices

Lisa Mandl[1,2] · Stefan Lang[3]

Abstract

Forest ecosystems are shaped by both abiotic and biotic disturbances. Unlike sudden disturbance agents, such as wind, avalanches and fire, bark beetle infestation progresses gradually. By the time infestation is observable by the human eye, trees are already in the final stages of infestation—the red- and grey-attack. In the relevant phase—the green-attack—biochemical and biophysical processes take place, which, however, are not or hardly visible. In this study, we applied a time series analysis based on semantically enriched Sentinel-2 data and spectral vegetation indices (SVIs) to detect early traces of bark beetle infestation in the Berchtesgaden National Park, Germany. Our approach used a stratified and hierarchical hybrid remote sensing image understanding system for pre-selecting candidate pixels, followed by the use of SVIs to confirm or refute the initial selection, heading towards a 'convergence of evidence approach'. Our results revealed that the near-infrared (NIR) and short-wave-infrared (SWIR) parts of the electromagnetic spectrum provided the best separability between pixels classified as healthy and early infested. Referring to vegetation indices, we found that those related to water stress have proven to be most sensitive. Compared to a SVI-only model that did not incorporate the concept of candidate pixels, our approach achieved distinctively higher producer's accuracy (76% vs. 63%) and user's accuracy (61% vs. 42%). The temporal accuracy of our method depends on the availability of satellite data and varies up to 3 weeks before or after the first ground-based detection in the field. Nonetheless, our method offers valuable early detection capabilities that can aid in implementing timely interventions to address bark beetle infestations in the early stage.

Keywords Image understanding · Pre-classification · Expert systems · Disturbances

1 Introduction

Environmental data have become increasingly available over the last decade. However, there remains a significant challenge in achieving constant monitoring that provides a comprehensive understanding of ecosystem dynamics and forest disturbance mechanisms (Fer et al. 2021). Satellite Earth observation (EO) provides to the ecological community a strong tool for understanding ecosystem dynamics, filling an important technological niche by serving spatially explicit information at broad spatial and long temporal scales (Senf 2022). The launch of the Sentinel-2 satellites as part of the European Copernicus programme has marked a paradigm shift in EO-based monitoring capabilities, making a significant contribution to ecosystem-related research (Fer et al. 2021). The ultimate goal of space assets such as EO is to convert "big Earth data" to valuable information that contributes to the understanding of—inter alia—natural processes (Sudmanns et al. 2020).

This study focuses on the detection of one of the most prominent forest ecosystem drivers in Central Europe, the European spruce bark beetle (*Ips typographus* L.), whose host tree is Norway spruce (*Picea abies* L.). This disturbance agent can have a profound impact on landscapes by initiating canopy gap dynamics, which may lead to significant changes in ecosystem functioning (Stritih et al. 2021). Although such insect disturbances lead to disrupted structures and alter the composition of the ecosystem, they

✉ Lisa Mandl
  lisa.mandl@tum.de

1 Berchtesgaden National Park, Berchtesgaden, Germany

2 Technical University of Munich, School of Life Sciences, Ecosystem Dynamics and Forest Management Group, Munich, Germany

3 Department of Geoinformatics, Faculty of Digital and Analytical Sciences, University of Salzburg, Salzburg, Austria

also foster diversity on the landscape (Thom and Seidl 2016). Historically, heavy storms and snow breakages have resulted in a large surplus of suitable breeding material for insects, particularly for bark beetles (Giunta et al. 2016; Seidl et al. 2017). Nowadays, climate change plays a major role in the increasing frequency and severity of bark beetle outbreaks (Thom and Seidl 2016). Warmer temperatures can lead to a more frequent occurrence of drought stress and increased severity of storms and windthrow, which in turn negatively impacts forest health (Filchev 2012; Marini et al. 2013) and results in an increase of annual bark beetle breeding generations (Jakoby et al. 2019). The damage mechanism of bark beetle infestation is a typical tree-killing strategy. The physiological response of trees to infestation is characterized by xylem embolism, which disrupts the flow of water and nutrients from the roots to the canopy. The damaged phloem tissue, responsible for transporting sugars produced by photosynthesis, reduces the amount of photosynthate available to the tree and thus the tree's photosynthetic activity (Christiansen and Bakke 1988; Edburg et al. 2012; Mikkelson et al. 2013; Wermelinger 2004). These biochemical processes ultimately result in a decreased water content and a reduction in leaf pigments leading to modifications in the spectral signature of a tree's canopy. These modifications make infested trees distinguishable from healthy ones (Ortiz et al. 2013; Sprintsin et al. 2011). Observing the changes in the biochemical and spectral signal over time, the beginning of bark beetle infestation (green-attack) is characterised by the tree's internal suffering described above, whilst still appearing green, i.e. with almost no visual signs of stress. As the process progresses, the tree's chlorophyll pigments decrease, resulting in discolouring whilst there is no needle loss yet, referred to as red-attack. The final stage (grey-attack) marks dead trees, which may have already lost their needles and which are clearly visible from both ground and space (Coulson 1985; Dalponte et al. 2022).

Early detection during the initial green-attack phase of bark beetle infestation, whilst essential for proactive actions, remains a challenging endeavour. Both ground-based and remote sensing-driven detection have human and technical limitations since early infested trees lack unambiguous symptoms clearly observable by the human eye (Abdullah et al. 2019a). Ground-based detection includes visual examination of bark beetle signs, like sawdust around the bore hole whilst trees are appearing green, and pheromone beetle traps (Sukovata et al. 2021). Although field surveys are still the state-of-the-art method for monitoring early signs of bark beetle calamities, they are time-consuming and insufficient for larger areas, and their lack of automation limits their practicality (Fernandez-Carrillo et al. 2020; Lang et al. 2006). Besides, field surveys are not straightforward either,

as it might be difficult to detect the small boreholes caused by the beetles (Wulder et al. 2006).

With the advent of aerial photographs in the 1930s, aerial imageries have been used for the first time for bark beetle detection (Coppin and Bauer 1996). A few decades later, remote sensing techniques gradually supplement the manual field work (Koch 2010; McRoberts and Tomppo 2007) enabling the investigation of ecosystem processes like disturbances at a broader scale (Turner et al. 1995). Relying on spectral changes of the forests canopy, remote sensing is able to capture forest traits like biochemical, functional, structural and geometrical characteristics (Hellwig et al. 2021; Lausch et al. 2016) and hence provide information about forests health. Some studies rely on chlorophyll and water stress-based indices to represent the biophysical and biochemical changes during bark beetle infestation (Verrelst et al. 2010; Wulder et al. 2006), whilst others employ forest structure information for the detection (Stereńczak et al. 2019).

The selection of the indices, bands and forest attributes for detection is closely linked to the platform (UAVs, airborne, and spaceborne) and their specific characteristics. Especially the use of unmanned aerial vehicles (UAVs) has become increasingly popular in early bark beetle detection (Alvarez-Vanhard et al. 2021). The ability for end-user customisation, such as mounting different sensors with varying spatial resolutions and acquisition intervals, and the resulting greater flexibility in application compared to space- and airborne sensors, contributes to promising results. This is reflected in high accuracies, often > 90%, and a better differentiation between different stages of bark beetle infestation (e.g. green-attack, red-attack, and grey-attack) in comparison to air- and spaceborne sensors. Whilst several studies (e.g. Hellwig et al. (2021), Honkavaara et al. (2020), Klouček et al. (2019), Paczkowski et al. (2020) or Schaeffer et al. (2021)) have highlighted the potential of this emerging technology, the use of UAVs is only suitable for small-scale monitoring as not viable for vast geographic regions (Alvarez-Vanhard et al. 2021).

Sensors mounted on aircrafts bridge the gap between the use of high-resolution UAVs and satellite data covering large areas. Most of these studies employed multi- or hyperspectral airborne sensors (e.g. Abdullah et al. (2019a), Fassnacht et al. (2014), Lausch et al. (2013)). However, despite the high number of spectral bands for hyperspectral sensors (e.g. 125 bands for HyMap), the potential of airborne hyperspectral data to distinguish early infested from healthy trees has been ambiguous, as shown by studies such as Fassnacht et al. (2012) or Wulder et al. (2006).

Most of the studies, however, employ commercial and/or non-commercial satellite data. Latifi et al. (2014a, 2014b) combined Landsat data with higher-resolution SPOT data whilst Latifi et al. (2018) used RapidEye

and MODIS data. Dense Landsat time series data were used by Trubin et al. (2022). Since the launch of the Sentinel-2 mission, an increasing number of studies have been published using these freely available data sources with a spatial resolution of 10 m, yet there is still some unused potential to be uncovered. Most of these approaches focus on spectral vegetation indices (SVIs) which are used as proxies for biophysical and biochemical changes of the canopy. Abdullah et al. (2019b) evaluated the potential of SVIs derived from Landsat and Sentinel-2 and found a substantial higher accuracy for Sentinel-2 data (67% of matching pixels compared to 36% for Landsat). Fernandez-Carrillo et al. (2020) used a change detection approach and bi-temporal regression to estimate forest vitality and reached high accuracies (> 0.8.) for the late phases of bark beetle infestation, yet the minor damage class (= early infestation stage) was often confused with the no damage class (= healthy). Huo et al. (2021) presented another approach combining Sentinel-1 and Sentinel-2 data, using SVIs as well as a combination of multiple bands in a non-parametric model. Seasonal and spectral trajectories of SVIs derived from Sentinel-2 data were investigated by Bárta et al. (2021). Dalponte et al. (2022) applied a support vector machine classifier to Sentinel-2 and LiDAR data and were able to differentiate between early and late infestation with an accuracy of 83%. Regardless of the exact approach, a remote sensing-based framework for early bark beetle detection requires undeniably a high temporal availability of satellite data (Bárta et al. 2021) as proven by most of the studies mentioned.

In this study, we propose a novel method to detect early infestation stages of the European Spruce Bark Beetle combining semantically enriched Sentinel-2 time series with SVIs in a mountain landscape. Our aim is to assess the benefit of pre-selecting candidate pixels from semantic Sentinel-2 data compared to a SVI-only approach. To address this objective, we investigated temporal response patterns and spectral category trajectories from the semantically enriched Sentinel-2 data and defined candidate pixels. We performed two model runs, one combining candidate pixels + SVIs and one with SVIs only and compared the results. Further, we studied the sensitivity of spectral bands and indices for detecting early stages of bark beetle detection. We evaluated our method not only in terms of the benefits of a preselection, but also at two different scales: pixel level and plot level. The plot level provided information about the general potential of our method to detect larger infestation areas. The pixel level, however, shows whether the method also detects smaller infested

areas, such as individual trees. The study was conducted in a national park, characterized by near-natural forests, yet with strong past land use legacies.

2 Materials and Methods

2.1 Study Site

The Berchtesgaden National Park, located in the southeastern part of Germany (Fig. 1), was chosen as the study site for this research. Covering an area of 20,808 ha, this region boasts a wide range in altitude, spanning over 2000 m from Lake Koenigssee (603 m) to Mount Watzmann (2713 m) (National Park Administration Berchtesgaden 2001). With a rich fundus of reference data, this area is well-suited for our research purpose. Of the total area, 11,835 ha (equivalent to 57%) are covered by forest, as defined by the FAO forest definition. Within this forested area, 8154 ha (81.5%) are dominated by conifers, while 1851 ha (18.5%) are dominated by deciduous trees (Mandl and Oravec 2020).

2.2 Satellite Datasets

The satellite data used in this study consisted of Copernicus Sentinel-2 level 2 and level 3 data. The latter corresponds to best-available pixel composites computed by the Weighted Average Synthesis Processor (WASP) (German Aerospace Center (DLR) 2020). Both level 2 and level 3 data were downloaded from the German Copernicus exploitation platform code-de.org. With its 13 channels optimized for land surface observations and a high resolution of up to 10 m and swath width of 290 km, Sentinel-2 is ideal for detecting changes in vegetation. For this study, ten Sentinel-2 scenes from March to October 2020 were used (for detailed dates, see Table 1). For the months of March and October, where illumination conditions and snow/shadow cover for mono-temporal imagery were unfavourable, we used level 3 best-available pixel composites. These composites allow for (almost) cloudless monthly synthesis of land surface reflectance compiled from cloud-free pixels within the respective month. When atmospheric conditions permitted, two images per month were acquired. For level 2 products, we used the FORCE data cube analysis-ready collection, provided by the German Copernicus platform code-de.org (CODE-de. org 2023; Frantz 2019). FORCE pre-processed data are
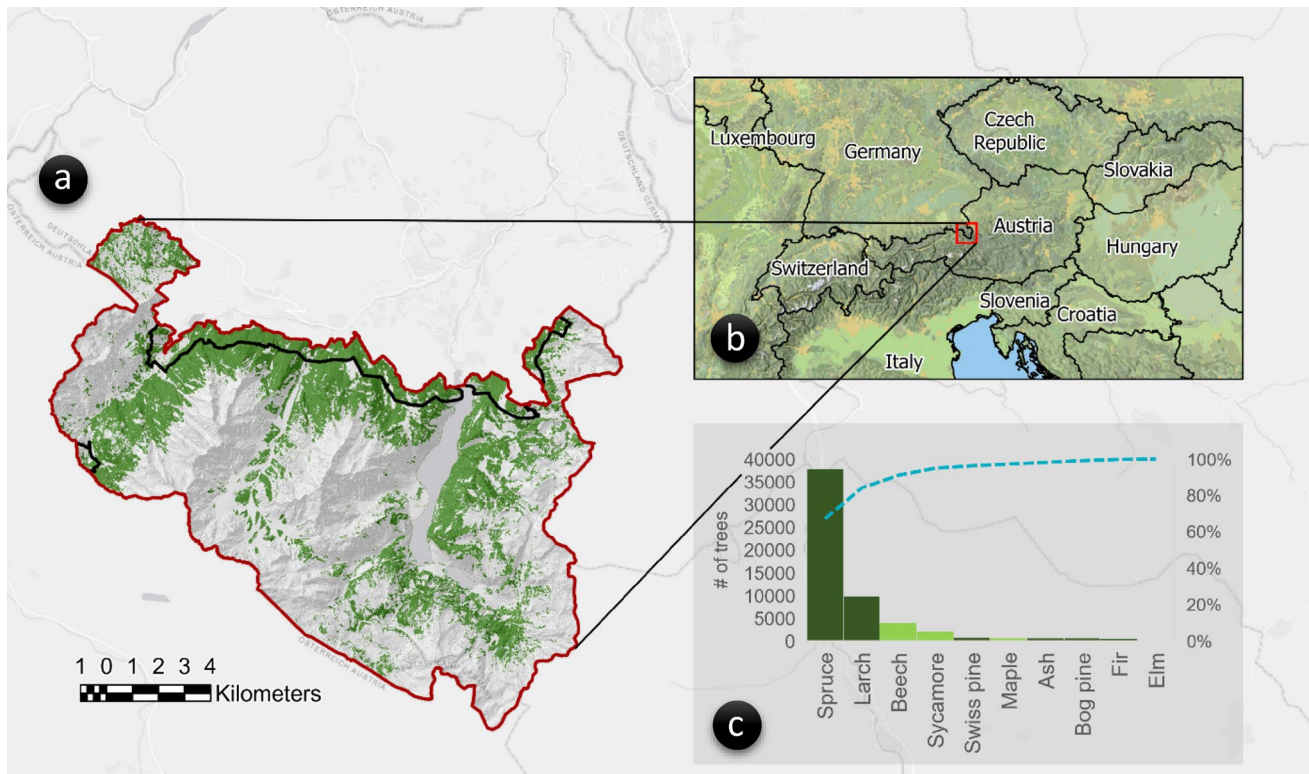
Fig. 1 The study site of Berchtesgaden National Park with the SIAM™ coniferous mask and the bark beetle management zone (black outlined area) (**a**), the location of the study site in Central Europe (**b**) and the occurrence of the top 10 tree species in the study site (**c**) indicating that the host tree spruce accounts for nearly 80% of the tree species within the National Park. In dark green, coniferous tree species, in light green deciduous ones. The blue line represents the cumulative area [%] of tree species in the National Park

Table 1 Sentinel-2 data used for the study with the exact date, the DOY, the processing level and according cloud coverage

Date	Day of year (DOY)	Processing level	Cloud coverage [%]
2020–03	–	Level 3-WASP (BAP)	0
2020–04–23	114	Level 2A-FORCE	4
2020–05–08	129	Level 2A-FORCE	20
2020–05–18	139	Level 2A-FORCE	16
2020–06–12	164	Level 2A-FORCE	0
2020–06–27	179	Level 2A-FORCE	14
2020–07–27	209	Level 2A-FORCE	9
2020–08–21	234	Level 2A-FORCE	0
2020–09–15	259	Level 2A-FORCE	0
2020–10	–	Level 3-WASP (BAP)	0

topographically and atmospherically corrected and feature advanced cloud and shadow detection, correction of adjacent effects, and a bidirectional reflectance distribution function, and deliver quality flags (Frantz 2019). FORCE considers all data with a cloud cover of less than 70% (according to the ESA/USGS metadata) for image acquisition. Image chips for which the FORCE internal cloud masking algorithm determined a cloud cover of more than 90% were discarded. To further reduce cloudy scenes, we applied a filter that excluded scenes with a cloud cover greater than 20%. Additionally, Sentinel-2 data were co-registered with the Landsat data, making FORCE data suitable for topographically challenging terrain, such as the Berchtesgaden National Park.

Based on pre-processed Sentinel-2 data, we computed SIAM™ (Satellite Image Automatic Mapper) spectral categories (see Sect. 2.4) and spectral vegetation indices (see Sect. 2.5) used in subsequent analysis. Based on temporal compositions of according spectral categories, we derived a coniferous forest dataset as a processing mask (for more details refer to Sect. 2.4).

Fig. 2 Grey-attack plot in a colour-infrared aerial imagery (left) and classification result of this infested patch (right)
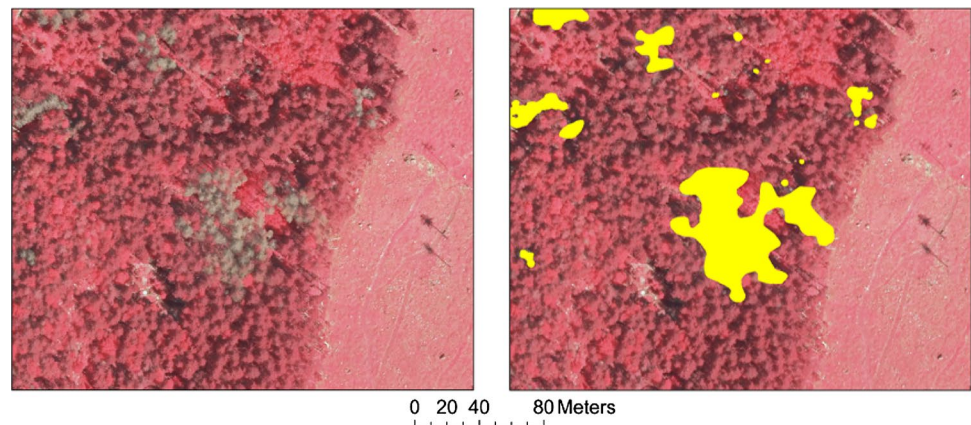
2.3 Reference Datasets

As reference data, we used a colour-infrared aerial imagery from the maintenance zone of Berchtesgaden National Park with a spatial resolution of 0.2 m dated September 2020. Grey-attack stages were derived by a supervised machine learning approach using a random forest ensemble model (Breiman 2001) in R's machine learning package *SuperML* (v0.5.3; [Saraswat 2020]*).* The resulting dataset, referred to as "grey-attack plots," represents the grey-attack state at the end of 2020 (see Fig. 2). We used this dataset to derive candidate pixels and for spatial accuracy assessment (see Sects. 2.6 and 2.9 for details).

To assess the temporal accuracy, we incorporated ground reference data collected in the field. Trained field workers systematically searched for early signs of bark beetle infestation, such as dry dust and holes in the bark, whilst the needles were still green. The bark beetle management zone was walked on a 4-weekly basis, and all infested areas were recorded using a mobile application.

2.4 Computation of Spectral Categories

We utilized SIAM™, a software that enables a stratified-hierarchical, physical and model-driven pre-classification (semantic enrichment) of Sentinel-2 data. The software incorporates the bands Blue, Green, Red, near-infrared (NIR), short-wave infrared-1 (SWIR-1) and short-wave-infrared-2 (SWIR-2) and is capable of categorizing satellite data into different spectral categories with varying degrees of granularity (Baraldi et al. 2006). This software — as opposed to the indices-based approach — considers the entire feature space of an image and divides it into spectral categories (SC). For this study, we used L-SIAM, which pre-classifies a Sentinel-2 scene into 96 spectral categories (Baraldi 2011; Baraldi et al. 2006, 2022). SIAM™ is a decision tree-based software that utilizes a priori knowledge of physical-spectral models to categorize satellite data into different granularities. This process is applied to each pixel in a satellite scene, resulting in a stable, sensor-agnostic multispectral colour name for each pixel. The software produces a discrete finite set of mutually exclusive and totally exhaustive spectral categories, as shown in Fig. 3 (Baraldi 2011; Baraldi et al. 2022; Sudmanns et al. 2021). The spectral categories are encoded by a spectral type ID, which is used to compute change values between two time periods. The change values emerge from the alterations of the spectral categories within the grey-attack plots of the bark beetle management zone and represent—inter alia—changes in the tree's caused by bark beetles.

Each of these spectral categories belongs to a broad set of land cover classes. These so-called parent classes consist of "clouds, "either snow or ice", "either water or shadow", "vegetation", equivalent to the land cover class set, "either

Fig. 3 Spectral endmembers of the SIAM™ output (modified after Baraldi et al. (2006) and a pre-classified Sentinel-2 scene of the study site
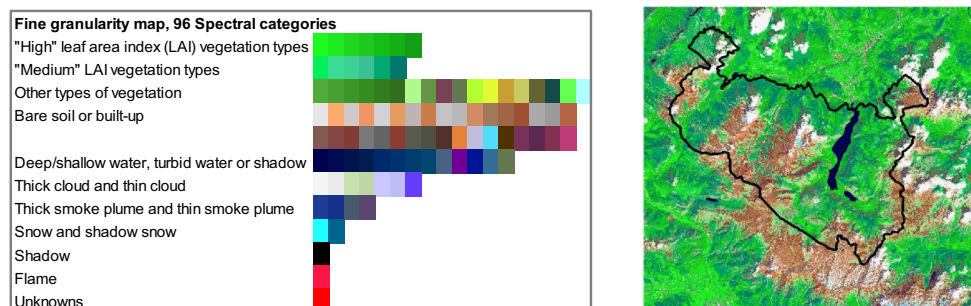
woody vegetation or cropland or grassland (herbaceous vegetation) or (shrub and brush) rangeland", "either bare soil or built-up" and "outliers". However, these classes do not claim absolute uniqueness; several semantic categories may be valid options derived from one and the same pixel, like water or shadow. Hence, the output of the semantic enrichment is to be regarded from the semantic context below a land cover classification (Baraldi et al. 2006). Still, in this way, even subtle differences or changes in the "greenness" of vegetation are detected in a spatially and spectrally discrete manner. In the context of this study, we used semantically enriched Sentinel-2 data to pre-select potentially infested pixels (candidate pixels), which are confirmed/refuted in the subsequent combination with SVIs. In this way, we account not only for subtle spectral changes, but also for canopy biophysical and biochemical changes during infestation. Besides using SIAM™ to select candidate pixels, we computed a coniferous mask by creating a SIAM™ spectral and temporal composite containing the spectral categories *strong (dense) vegetation with low NIR leaf SC* (spectral category 6), *strong (dense) vegetation with very low NIR leaf SC* (spectral category 7), *average (dense) vegetation with low NIR leaf SC* (spectral category 12) *and average (dense) vegetation with very low NIR leaf SC* (spectral category 13; for all spectral categories refer to supplementary material 5), which all are related to a coniferous land cover class as identified by the USGS (Baraldi et al. 2006). To enhance the robustness of the mask, we computed a multi-year forest mask consisting of all coniferous-related SIAM™ spectral categories for the leaf-on season (April–September) and for the years 2018–2020. In this way, we account for phenological effects and varying illumination conditions and thus

avoid excluding potential coniferous pixels, what is likely to happen when using a single-year coniferous mask.

As input, the software expects at least radiometrically calibrated and top-of-atmosphere (ToA) computed data (Baraldi 2011). For Sentinel-2 level 2 data exhibiting bottom-of-atmosphere (BoA) correction, this requirement was already given. In order to execute the SIAM™ software, Sentinel-2 level 2 data were pre-processed including spatial resampling, image stacking, conversion to 8-bit ENVI format and the creation of noData mask. In case thermal infrared information is available (as in Landsat sensors), this dimension is considered in the model as well. Finally, the semantic enrichment was conducted by calculating the spectral categories of input data in batch mode.

2.5 Selection and Computation of Spectral Vegetation Indices

We developed a comprehensive set of spectral vegetation indices by applying a principal component analysis (PCA) on eight both uni- and multidimensional indices identified in a literature research. Tasseled Cap components (Tasseled Cap Greenness, Tasseled Cap Brightness, Tasseled Cap Wetness) were derived by applying Sentinel-2 specific coefficients from a principal component-based procrustes analysis developed by Shi and Xu (2019). For the unidimensional indices, we included both chlorophyll/cell structure-related indices and water stress-related indices, which were used as proxies for biophysical and biochemical processes during infestation.

A total of 1000 randomly selected pixels were employed and the respective index values were determined in order to investigate indices which explain a substantial proportion

Table 2 Selected spectral indices, their abbreviation, grouping into water-stress- and chlorophyll-related indices, their formula as well as the reference

Spectral index	Abbrev	Type	Formula	References
Disease Water Stress Index	DSWI	Water stress index	(NIR + GREEN)/(SWIR1 + RED)	Galvão et al. (2005)
Normalized Difference Water Index	NDWI	Water stress index	(GREEN−NIR)/(GREEN + NIR)	McFeeters (1996)
Normalized Difference Moisture Index	NDMI	Water stress index	(NIR-SWIR) / (NIR + SWIR)	Buschmann (1993)
Tasseled Cap Wetness (for Sentinel-2 MSI)	TCW	Water stress index	0.1763*BLUE + 0.1615*GREEN + 0.0486*RED − 0.0755*NIR-1 − 0.7701*SWIR1-0.5293*SWIR2	Shi and Xu (2019)
Normalized Difference Green Red Index	NGRDI	Chlorophyll/cell structure index	(GREEN–RED)/(GREEN + RED)	Hunt et al. (2011)
Normalized Difference Red-Edge-3 Index	NDRE-3	Chlorophyll//cell structure index	(NIR-RED EDGE-2) / (NIR + RED EDGE-2)	Abdullah et al. (2019b)
Normalized Difference Index Bands 4/5	NDI45	Chlorophyll//cell structure index	(RED EDGE1 -RED) / (RED EDGE1 + RED)	Delegido et al. (2011)
Tasseled Cap Greenness (for Sentinel-2 MSI)	TCG	Chlorophyll//cell structure index	− 0.2818*BLUE − 0.3020*GREEN − 0.4283*RED + 0.4261*NIR-2 + 0.1341*SWIR1 − 0.2538*SWIR2	Shi and Xu (2019)

of the variance (Wickham and Grolemund 2017). The final set of indices was subsequently used to confirm or refute, respectively, the pre-selected candidate pixels determined by the semantic enrichment approach (see Sect. 2.4). The indices-based computations (see Sect. 2.5, Table 2) and the feature engineering for determining index-based feature importance (see Sect. 2.8) were both conducted using the R packages sen2r and SuperML (R Core Team 2021; Ranghetti et al. 2020; Saraswat 2020).

2.6 Combining Candidate Pixels with SVIs – Identifying Early Traces of Bark Beetle Infestation

We computed change detection raster stacks of SIAM™ spectral categories and SVIs consisting of nine bands, where each band corresponds to one epoch. We defined an epoch as a set of two consecutive satellite scenes and their corresponding semantically enriched Sentinel-2 data and derived SVIs. Within the grey-attack plots, we observed the change values of spectral categories following the approach outlined in Mandl and Lang (2022). The grey-attack plots

0 = no candidate) and intersected all candidate pixels with the SVI change detection grids. By extracting the change values from the change detection stacks per SVI and epoch at the candidate's pixel locations, we obtained a set of descriptive statistics per SVI and epoch. Based on these statistics, we tested different thresholds for reclassifying SVIs as either 1 (early infested) or 0 (healthy). We checked following combinations for the lower and upper thresholds of indices: (1) lower: min, upper: max, (2) lower: 3rd quartile, upper: 1st quartile, (3) lower: min, upper: 3rd quartile. We found in this preliminary analysis, that the minimum value serves best as lower threshold, whilst the 3rd quartile led to best results for the upper threshold (for detailed descriptive statistics and used thresholds per index and epoch, please refer to supplementary material 1a and 1b). Based on these thresholds, we reclassified the SVIs. To address only potential host trees, we used a coniferous mask derived from the SIAM™ spectral categories composition explained in Sect. 2.4. Just like the candidate pixels and SVIs, binary reclassification was applied to the coniferous mask (1 = coniferous, 0 = other). An expert and rule-based conditional statement based on reclassified SIAM™ indicators, reclassified SVIs and the reclassified coniferous mask determines a pixel as *early infested*, applying following rule set:

$$\text{Early infested} = \begin{cases} 1 & \text{if SIAM}^{\text{TM}} \text{ candidates } == 1 \wedge \text{ all indices } == 1 \wedge \text{ coniferous mask } == 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

were used to initially determine candidate pixel values as it is known that there will be bark beetle attack sometime during the year, yet there is no exact temporal information, that is when infestation took place. Specifically, we considered the change value of each pixel classified as one of the SIAM™ end members (high leaf area index (LAI) vegetation types, medium LAI vegetation types, and other types of vegetation) within the grey-attack plots as a candidate pixel. Our approach is based on two key assumptions: (1) Changes in the spectral categories of SIAM™ within the grey-attack plots indicate bark beetle infestation at an early stage and hence, the changes in these categories act as indicators, referred to as candidate pixels. Initially, all spectral category changes are considered as candidates. Vegetation indices determine if changes are due to bark beetle infestation or other factors like phenology or non-stationary surface properties (Baraldi 2011), confirming or refuting the candidates. The interaction between the SIAM™ candidate pixels and the SVIs is the core of our method. Using the SVIs to check the causality of the pre-selected candidate pixels, we adopted a "convergence of evidence" approach.

In detail, after defining SIAM™ pixel candidate values, we binary-reclassified all raster cells featuring the previously determined candidate values (1 = SIAM™ pixel candidate,

Each pixel must be classified as a SIAM-derived candidate, must be a pixel of the class coniferous and must be within the defined threshold range of selected indices in order to be included in the infestation grid. The infestation grid consists of two values where the value 1 indicates an infested pixel and 0 a healthy pixel. Figure 4 illustrates the workflow and the interrelations of its methodological components.

The workflow was applied to every time step, leading to one infestation grid per epoch. We also implemented a double-check mechanism to minimize the potential error that may result from pixel shift or adverse illumination conditions. That means that all pixels classified as *early infested* in a first iteration were re-checked in a second one. If both iterations classified the according pixel as *early infested*, the pixel was considered as infested, otherwise it was reset to *healthy*.

2.7 Comparison Between a SVI-Only Approach and the Candidate Pixel Approach

To evaluate the effectiveness of our candidates approach compared to SVI-only approach, we investigated the spectral characteristics and thus the discriminability between healthy
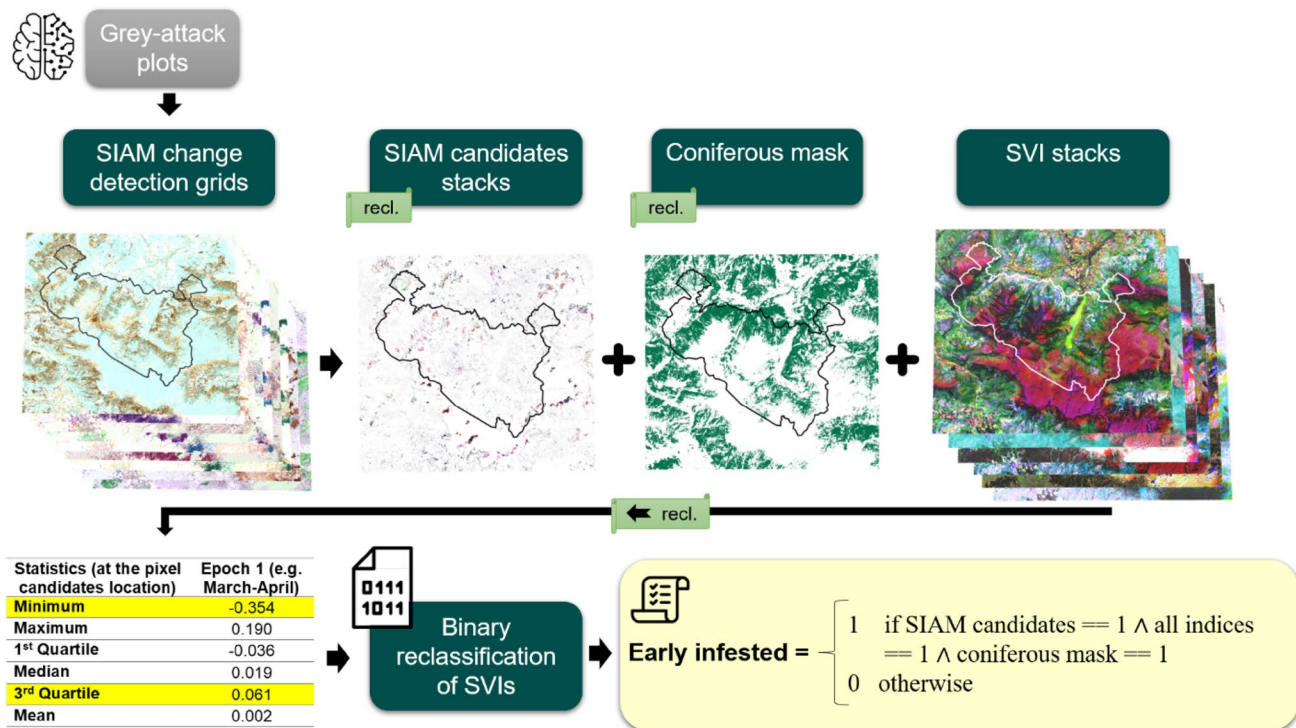
Fig. 4 Visualization of the workflow, starting with the intersection of grey-attack plots with the SIAM™ change detection grids, the subsequent derivation of SIAM™ candidates and their reclassification, the application of a coniferous mask and finally the determination of SVI thresholds via statistics. The computation of infestation grids is done by applying an expert-based rule set

and early infested pixels for our combined model and for a SVI-only model. We performed a second run of our algorithm using only the previously determined SVI thresholds and the forest mask, without using the candidate pixels. We then plotted the spectral reflectance of early infested pixels per band for both the SVI-only and combined (SVI + candidate pixel) models, and compared the results to the baseline—the median reflectance of healthy pixels. A larger difference from the baseline indicates a better differentiation between the two classes and thus a more accurate classification of early infested pixels. To avoid bias introduced by the grey-attack plots, we excluded them in this analysis.

2.8 Statistical Analysis

After computing the infestation grids for each epoch, we used these grids and generated a representative dataset for the statistical analysis. We selected pixels from the class *early infested*, which are centre pixels of an infestation plot to ensure a high likelihood, that these pixels really represent the early infested class. This results in a total of 350 pixels used for subsequent analysis. We also randomly sampled 350 pixels each from the healthy and grey-attack (from the ML approach) classes and investigated the spectral

separability and temporal variation of canopy reflectance under bark beetle infestation, both for the complete year and stratified by time steps. To evaluate the significance of the spectral differences between the healthy and early infested class for selected bands and indices, we applied a t-test. We also determined the temporal response of vegetation indices under bark beetle infestation using the same dataset as for the assessment of the spectral feature space.

Further, we assessed the impact of indices by fitting two random forest models. The first RF model determined feature importance of SVIs for the entire year by including the SVI change detection values at the location of the 350 early infested pixels and the 350 healthy pixels mentioned above. The second RF model used the same dataset as the first, but the model was stratified by epoch to investigate intra-annual changes in feature importance of SVIs. The relative rank of SVIs with respect to the predictability to the target variable "early infested" was computed by the impurity-based feature importance (Gini importance).

2.9 Validation

A reliable validation in terms of early bark beetle detection via remote sensing needs to consider both, the spatial and the temporal dimension. Reference data for the spatial

accuracy assessment were only available for the maintenance zone (~ 5300 ha) and for the temporal accuracy only for the bark beetle management zone (~ 1800 ha) of Berchtesgaden National Park. Thus, the number of predicted bark beetle plots being included in the validation is somewhat limited. The detection algorithm, however, was applied to the entire national park (for a heatmap of infestation, refer to Material 2 in the supplementary material). The spatial accuracy was evaluated based on the ML approach detecting grey-attack stages of bark beetle attack whilst ground truth data was used for the temporal assessment.

For the spatial accuracy assessment, we converted the grey-attack vector dataset into a grid with 10 m cell size to match the spatial resolution of Sentinel-2 data. This resulted in a total of 238 reference pixels, which were either single pixels ($n = 41$) or clustered into plots ($n = 31$, $\varnothing$ number of pixels per plot $= 6$) where a plot referred to a larger group of infested trees, i.e. at least two 10 m pixels. For the subsequent accuracy assessment, we kept the subdivision of our grey-attack reference dataset into these two different scales, the pixel- and the plot-level. Thereby, we checked the potential of our method to also detect small-scale infested trees and not just larger plots. On the plot-level, a plot was considered to be detected if the infestation grid indicated at least 50% of the reference pixels as infested.

For the temporal validation, we first computed a 20 m radius around the location to account for potential geolocation errors due to weak GNSS signal in forests. We then intersected our detected early infested pixels with the ground truth data and compared their date of recording with the date of first detection by the early detection algorithm.

To assess the use of the combined model not only in terms of spectral characteristics but also in terms of spatial and temporal accuracy, we further compared the results of the combined model with the SVI-only model using the same datasets and methods.

3 Results

3.1 Determining Candidate Pixels via Spectral Categories

The intersection of semantically enriched Sentinel-2 change detection grids with the grey-attack plots revealed the most common changes in the SIAM™ spectral categories. Figure 5 depicts these changes in a Sankey plot. The plot shows the shifting proportions and change values between spectral categories when comparing time step n and time step $n + 1$. However, the change values do not represent real quantities, but are an indirect measure of the (semantic) difference between the spectral categories. The most frequent changes occur from the spectral categories 5, 12, 13 and 24 to the spectral categories 11, 12, 13, 24 and 32. For detailed description of these spectral categories, refer to Table 3. To define the final set of candidate pixel values, we used the resulting values consisting of the change values − 11, − 8, − 1, 1, 6, 7, 8, 11, 12 and 70.

Fig. 5 Sankey plot showing the changes in spectral categories (expressed by their index ID) from one time step to the next time step. The right nodes show the change values resulting from substracting the first time step (time step n) by the subsequent time step ($n + 1$). The ten most frequent change values were used for creating the final set of spectral category-based candidate pixels. The plot contains changes from the observation period March – October. For more details about the spectral categories, refer to Table 3 or supplementary material 5
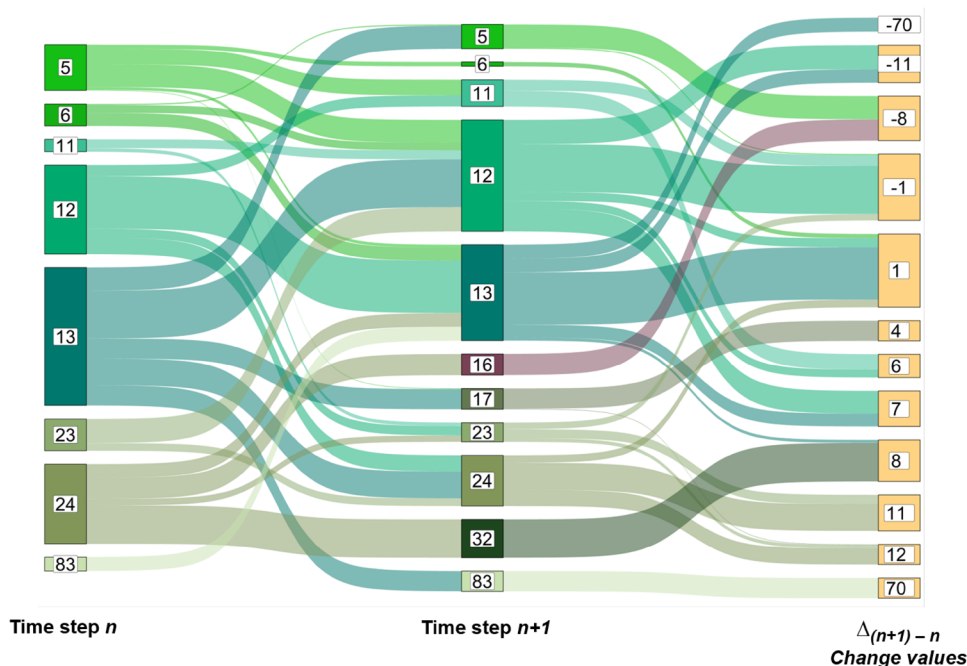
Table 3 Description of spectral categories (SC) shown in Fig. 5. For a list of all SIAM™ spectral classes, refer to supplementary material 5

Fine semantic granularity, 96 spectral types (Large set)	(L)SC name (spectral-based semi-concept)	Land cover classes (identified by the USGS index, if any) likely to overlap with that (L)SC	Pseudo-color
5	Strong (Dense) Vegetation with Medium NIR Leaf SC	(Dense) evergreen forest land, especially broadleaved forests (dense) deciduous forest land (41).	
6	Strong (Dense) Vegetation with Low NIR Leaf SC	Evergreen forests, especially coniferous forests.	
11	Strong (Dense) Vegetation with Medium NIR Leaf SC	(Dense) evergreen forest land, especially broadleaved forests, (dense) deciduous forest land.	
12	Average (Dense) Vegetation with Low NIR Leaf SC	Evergreen forests, especially coniferous forests.	
13	Average (Dense) Vegetation with Very Low NIR Leaf SC	Evergreen forests, especially coniferous forests.	
16	Rangeland (Sparse vegetation) or Wetland in Shadow or Turbid water or Shadow Leaf SC	Non-forested wetland.	
17	(Dense) Vegetation in shadow Leaf SC	Shadow areas with (Dense) Vegetation.	
23	Average Shrub Rangeland (Sparse vegetation) with Low NIR Leaf SC	Mixed (open) forests, olive groves, herbaceous rangeland.	
24	Average Shrub Rangeland (Sparse vegetation) with Very Low NIR Leaf SC	Mixed (open) forests, olive groves, herbaceous rangeland.	
32	(Dense) Vegetation in water or shadow (including algae) Leaf SC	Vegetation in water or shadow	
83	Thin Cloud on Vegetated areas or Shrub Rangeland (Sparse vegetation) Leaf SC	Thin clouds (e.g., "popcorn" clouds, stratus, edges of cumulus) upon vegetated areas.	
84	Thin Cloud on Vegetated areas or Herbaceous Rangeland (Sparse vegetation) Leaf SC	Thin clouds (e.g., "popcorn" clouds, stratus, edges of cumulus) upon vegetated areas.	

3.2 Selection of Suitable Indices

The results of the PCA analysis revealed that Tasseled Cap Greenness (TCG) and NDRE3, as well as DSWI and NDMI, are strongly interrelated (Fig. 6). The loading plot depicts each input variable (SVI) as a vector and plots its coefficients for the first and second components. The length of the vector represents the corresponding loading values per principal component. This plot helped to identify which SVIs have the strongest effect on each component. The PCA results indicated a strong positive correlation between DSWI and NDMI, as well as between TCG and NDRE3, suggesting spectral dependence. Consequently, we excluded TCG and NDMI from further analysis, as these two indices have shorter vectors compared to their correlated indices, indicating that they explain less of the variance.

The final set of indices, based on the PCA, consisted of NDWI, TCW and DSWI for water stress-related indices and NDRE3, NGRDI and NDI45 representing the group of chlorophyll and cell structure-related indices.
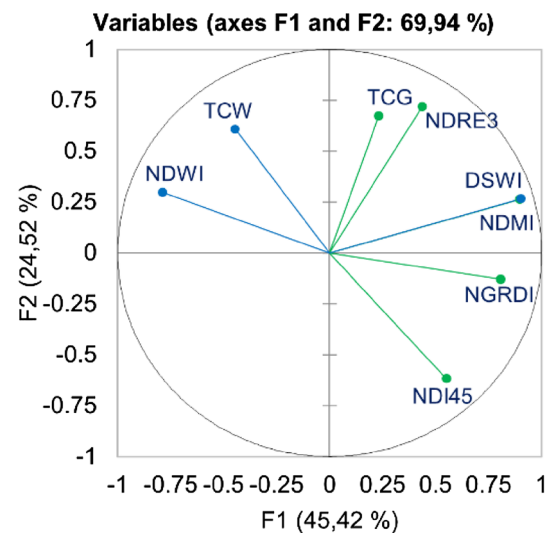

Variables (axes F1 and F2: 69,94 %)

Fig. 6 Loading plot resulted from PCA. Blue vectors represent water stress-related indices, green vector chlorophyll- and structure-related ones

Fig. 7 Mean reflectance of pixels classified as healthy and corresponding confidence intervals. The blue squares indicate reflectance of pixels classified as early infested for the combined model (SIAM™ candidates + SVIs) and their corresponding error bars, whereas green triangles show reflectance of early infested pixels for the SVI-only model and their error bars
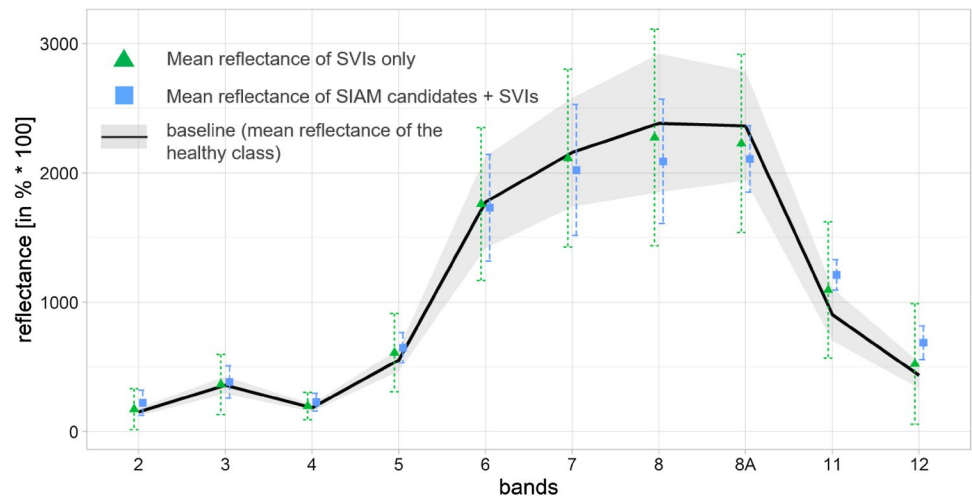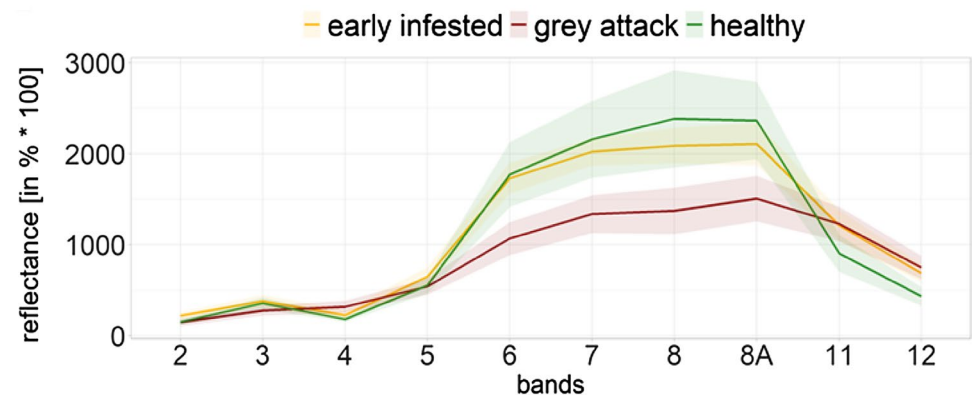
Fig. 8 Spectral behaviour of pixels of the classes *early infested*, *grey-attack* and *healthy* in the electromagnetic spectrum. The ribbons show confidence intervals

3.3 Comparison of Spectral Characteristics for the Combined and the SVI-Only Model

Comparing the spectral characteristics of the SIAM™ candidates + SVI model (combined model) to the SVI-only model, we consistently observed larger differences from the baseline, which is the median reflectance of pixels from the healthy class. The increased differences suggested that the combined model had a greater potential for distinguishing pixels from the healthy and early infested classes, as the spectral differences became more apparent. Whilst we found only slight differences in the visible (bands 2–4) and red-edge (bands 5–7) parts of the electromagnetic spectrum, we observed stronger differences in the NIR and SWIR regions (Fig. 7). Detailed differences in reflectance per band of the SVI-only model and the combined model can be found in the supplementary material 3. In contrast to the healthy class, the early attack class demonstrated higher standard deviations in both models. However, the combined model consistently exhibited lower standard deviations than the SVI-only model. Based on these results, we performed

further analysis only for the combined model and not for the SVI-only model.

3.4 Spectral Separability and Temporal Variation of Canopy Reflectance Under Infestation

Our analysis of the combined model revealed distinct differences in pixel spectra between those classified as healthy and grey-attack pixels (Fig. 8). These differences were particularly prominent in the red-edge and NIR regions of the electromagnetic spectrum, spanning from band 5 to 8, but less pronounced in the visible (bands 2–4) and SWIR (bands 11 and 12) range. The comparison between pixels of the healthy and early infested class showed significantly smaller differences across all spectral bands. We found the best separability between the two classes in bands 8 and 8A (NIR-1 and NIR-2), where pixels classified as early infested showed lower reflectance, and in bands 11 and 12 (SWIR-1 and SWIR-2), where we observed higher reflectance compared to the healthy class.

The stratification by the single time steps of available Sentinel-2 data (Fig. 9) revealed intra-annual differences in

Fig. 9 Reflectance for pixels of the class healthy and early-infested per band and stratified by time steps. The black horizontal lines indicate the mean, the boxes the interquartile range (limited by the 25th and 75th quantile). Black points show outliers (> 1.5 interquartile range below/above the 25th and 75th quantile)
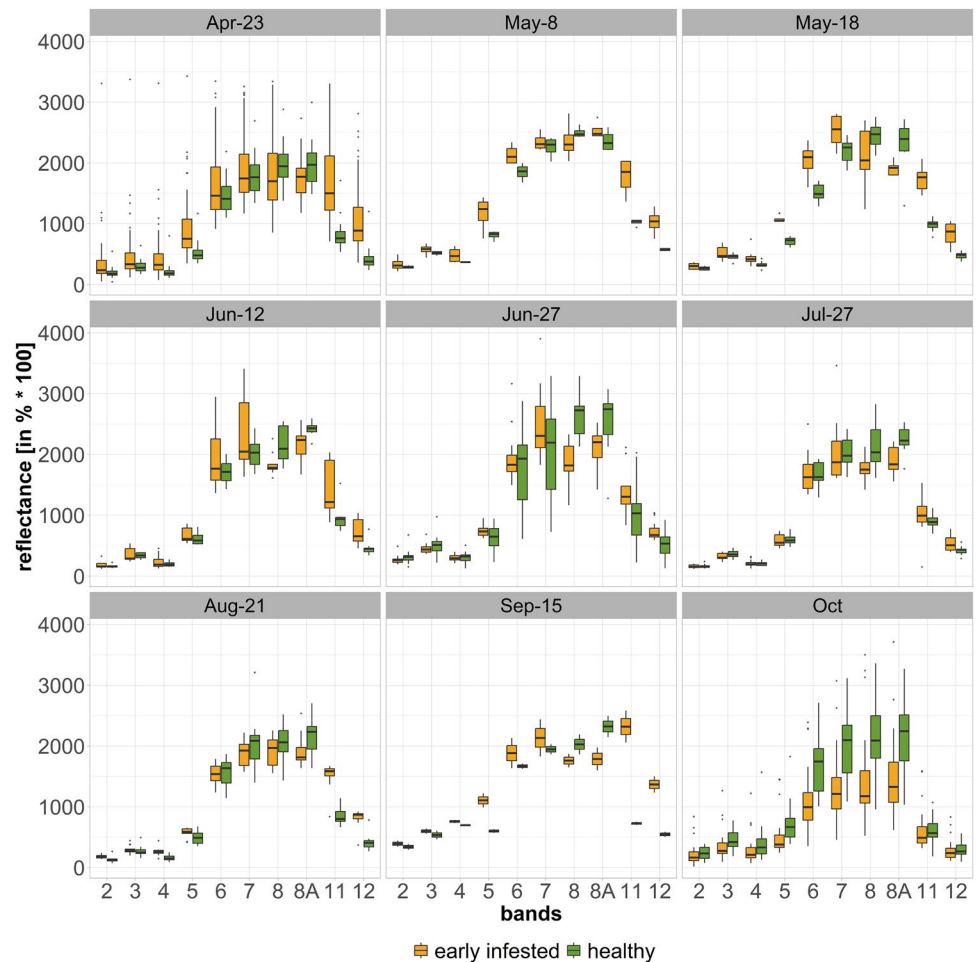

Table 4 Significance levels of the single epochs for the bands 8 (NIR-1), 8A (NIR-2), 11 (SWIR-1) and 12 (SWIR-2) as well as for the indices *p*-values were computed by applying a *t* test

| Epochs	B8	B8A	B11	B12	DSWI	NDWI	TCW	NDI45	NDRE3	NGRDI
2020–03/2020–04-23	0.491	**0.001**	**< 0.001**	**< 0.001**	0.051	0.188	0.292	0.103	**0.032**	0.218
2020–04-23/2020–05-08	0.211	0.102	**< 0.001**	**< 0.001**	**< 0.001**	0.164	**0.037**	**0.021**	0.125	0.093
2020–05-08/2020–05-18	0.101	**0.034**	**< 0.001**	**0.002**	**< 0.001**	0.235	0.204	**0.042**	0.072	0.113
2020–05-18/2020–06-12	**0.02**	**0.026**	**0.013**	**0.004**	**< 0.001**	**0.046**	**0.048**	**0.022**	0.094	0.088
2020–06-12/2020–06-27	**< 0.001**	**0.003**	0.065	**0.002**	**0.009**	**0.002**	**0.045**	0.098	0.172	**0.048**
2020–06-27/2020–07-27	**0.005**	**0.001**	0.205	**0.004**	**< 0.001**	0.058	0.062	0.191	**0.021**	0.101
2020–07-27/2020–08-21	0.195	**0.02**	**< 0.001**	**< 0.001**	**< 0.001**	**0.022**	**0.047**	**0.039**	0.095	**0.042**
2020–08-21/2020–09-15	0.306	0.172	**< 0.001**	**< 0.001**	**< 0.001**	**< 0.001**	**0.038**	**0.050**	0.062	0.105
2020–09-15/2020–10	**0.003**	**0.004**	0.667	0.938	**< 0.001**	**< 0.001**	0.053	0.133	**0.047**	0.182

In bold: values with *p* < 0.05

the spectral behaviour of pixels classified as healthy and early infested, respectively.

We found that pixels classified as early infested generally had a larger interquartile range (IQR) for most time steps, indicating greater signal heterogeneity compared to the healthy class. As the year progressed, we observed a trend towards more distinct signals from pixels classified as early infested, as evidenced by a decreasing interquartile range (IQR) and more pronounced differences between the two groups. Outliers were observed in April and October, such as high reflectance in the SWIR range (bands 11 and 12) for the early infested class in April. The visible range

Fig. 10 Temporal response of indices for healthy (green) and early infested (yellow) plots. The first row shows water stress-related indices, the second one chlorophyll-related indices. The vertical error bars indicate the standard deviation per index and time stamp. Pixels of the class healthy show lower standard deviation whilst pixels of the class early infested exhibit stronger deviation from the mean
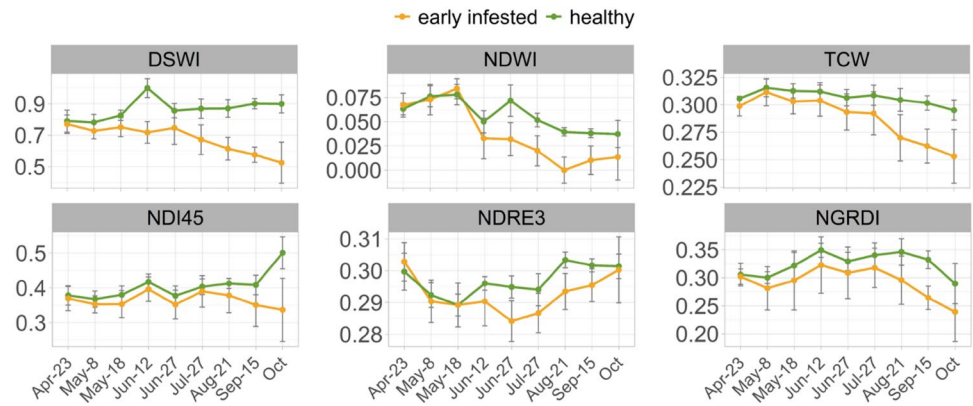
of the EM spectrum (bands 2–4) had low intra-class and inter-class variability. Whilst the red-edge bands (bands 5–7) showed slightly stronger inter-class differences compared to April, these differences became even more distinctive for the near- and shortwave infrared ranges (bands 8–12), indicating better separability between the healthy and early infested classes. This is consistent with the overall evaluation of separability throughout the entire bark beetle season (see Fig. 8). From May to August, our results consistently showed lower reflectance for the early infested class in the NIR bands and higher reflectance in the SWIR bands. Differences in the SWIR range equalized at the end of the growing season (October). The t-test performed for each time step for the NIR and SWIR bands revealed statistical significance in separating early infested versus healthy pixel spectra, particularly for the SWIR range. These wavelength ranges indicated significant differences between the healthy

and early infested classes at an earlier stage compared to the NIR bands (see Table 4). However, the NIR bands showed good sensitivity from June onwards.

3.5 Temporal response of vegetation indices

Moving from the entire reflectance spectrum to proxies for biophysical and biochemical properties represented by vegetation indices, we found that water stress-related indices were more sensitive to early bark beetle infestation than chlorophyll- and cell structure-related ones are (Fig. 10). This was also confirmed by the t-test (Table 4), which revealed a much more frequent significant difference between the healthy and early infested class for water stress indices.

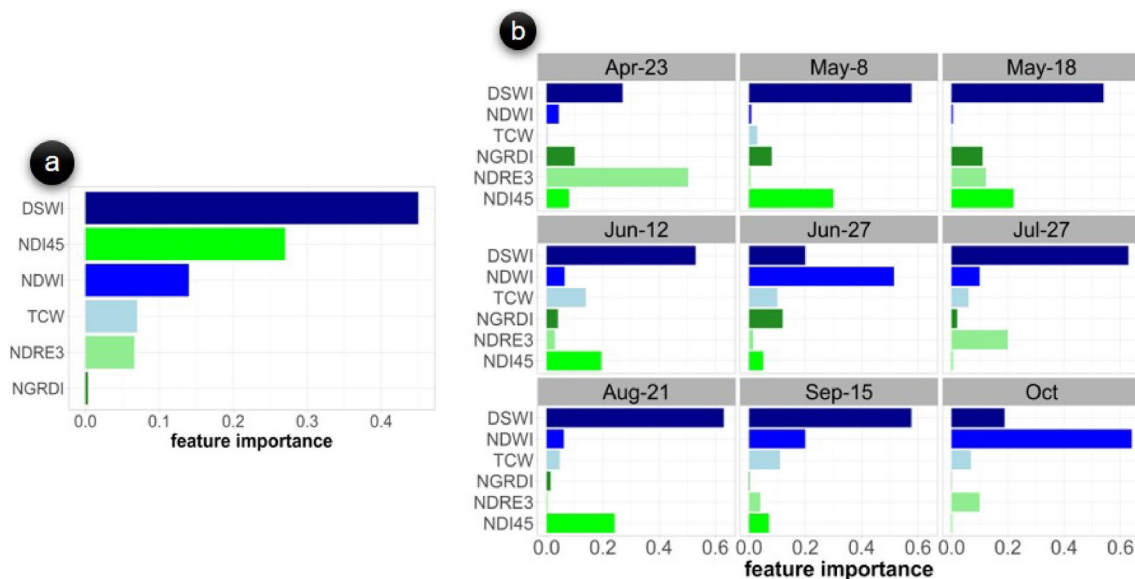At the beginning of the season, differences between the classes healthy and early infested were less pronounced for

Fig. 11 Feature importance for the entire year (**a**) and for all single time steps (**b**). Blue colours indicate water stress-related indices, green colours chlorophyll-related ones

Table 5 Accuracy of the random forest model for determining the epochs feature importance of SVIs

Epochs	Accuracy	OOB error
2020–03/2020–04-23	36.22%	0.67
2020–04–23/2020–05–08	54.67%	0.23
2020–05–08/2020–05–18	72.00%	0.16
2020–05–18/2020–06–12	62.50%	0.27
2020–06–12/2020–06–27	81.67%	0.11
2020–06–27/2020–07–27	66.67%	0.24
2020–07–27/2020–08–21	70.59%	0.14
2020–08–21/2020–09–15	78.29%	0.11
2020–09–15/2020–10	37.99%	0.68

both index groups, yet DSWI showed the largest differences from the pixels classified as non-infested, both in the t-test and when investigating the trajectories in the seasonal course (Fig. 10). The gap between the healthy and early infested class became more pronounced earlier in the year for water stress-related indices (DSWI, NDWI, and TCW) compared to chlorophyll- and cell structure-related indices. Chlorophyll- and structure-related indices (NDI45, NDRE 3, and NGRDI), however, showed stronger parallel trajectories for the two classes healthy and early infested than water stress-related indices did, indicating less sensitivity to predict bark beetle infestation. The results from the t-test for these indices confirmed this observation (Table 4). As the year progressed, the chlorophyll and structure indices showed increased relevance, with more frequent significant differences between healthy and early infested classes from July onwards. However, water stress indices still outperformed chlorophyll and structure indices in discriminating between the two classes.

3.6 Impact of Indices on the Classification Result Based on Impurity-Based Feature Importance

The importance of different vegetation indices in predicting early bark beetle infestation was evaluated by assessing feature importance, both for the entire season (Fig. 11a)

and for each time step (Fig. 11b). DSWI was found to be the most important index for the entire year, followed by NDI45 and NDWI, whereas the impact of TCW, NDRE3 and NGRDI was relatively low. The intra-annual comparison showed substantial variation in index ranking, accuracy and out-of-bag (OOB) error (Table 5). DSWI consistently ranked first or second in all time steps, whilst the importance of chlorophyll-related indices, such as NGRDI, NDRE3, and NDI45, remained moderate throughout the season. For water stress-related indices, DSWI had consistently high feature importance, whilst the importance of other water stress indices increased significantly from June onwards.

The RF model had high accuracy (> 70%) and low OOB errors for time steps in May, June, August and September. However, the accuracy was weak and OOB errors were high for April and October, which represent the beginning and end of the bark beetle season. June 12th and July 27th showed relatively lower accuracy despite being in the middle of the season.
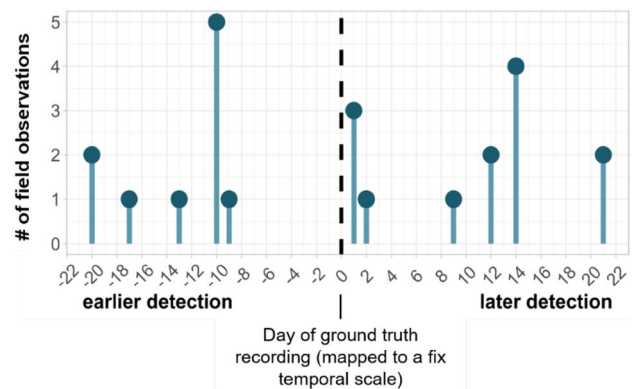
Fig. 12 Graphical representation of the aggregated temporal accuracy. We mapped all predictions and associated field recordings to a fix point in time, which is shown as a dashed line. Deviations from these reference dates are distinguished between earlier detection (left) and later detection (right)

Table 6 Site-specific accuracy metrics for the pixel- and the plot-level. PA = Producer's accuracy, UA = User's accuracy

	# of early infested (prediction)	# of grey-attack (reference)	# of matching pixels	PA (%)	UA (%)	Error of omission (%)	Error of commission (%)
Combined model							
Pixel-level	103	238	41	17	40	83	61
Plot-level	57	46	35	76	61	24	39
SVI-only model							
Pixel-level	312	238	54	23	17	77	83
Plot-level	69	46	29	63	42	39	58

3.7 Validation

We employed a validation approach that incorporated both spatial and temporal dimensions using the grey-attack (spatial validation) and ground reference dataset (temporal validation) as reference data, as outlined in Sect. 2.3. The validation process assessed accuracy at both the pixel and plot levels and included evaluation of the accuracy metrics for both our combined model and the SVI-only model.

Starting with the spatial accuracy assessment of the combined model, we found that the site-specific pixel-level validation yielded a low producer's accuracy of only 17% and a user's accuracy of 40% (see Table 6). The error of omission was high, indicating that the majority of infested grey-attack pixels were not matched by the prediction, whilst the error of commission was 61%, meaning that pixels were incorrectly assigned to the wrong class. On the other hand, for the plot-level validation, the producer's accuracy was 76%, and the user's accuracy was 61%. Both plot validation metrics showed distinctively higher accuracies compared to the pixel-level assessment. The error of omission at the plot-level was 24%, whilst the error of commission was 39%.

In contrast, for the SVI-only model, we found substantially lower producer's and user's accuracy, with values of 23% and 17%, respectively, at the pixel-level. Accordingly, the errors of omission and commission were high, at 77% and 83%, respectively. Even though the differences in the plot-level accuracy metrics between the two models were smaller compared to the pixel-level, the combined model still revealed distinctively higher accuracy.

From a temporal perspective, earlier and later detection, respectively, scattered from 3 weeks earlier until 3 weeks later compared to ground truth data (Fig. 12). The results of the temporal accuracy assessment revealed no clear trends in temporal accuracy between months. The share of pixels classified before field sighting was fairly balanced with those representing the period afterwards. For detailed dates about ground truth recording and related remote sensing-based detection, please refer to supplementary material 4.

4 Discussion

4.1 Methodological Perspective

We have demonstrated that our proposed method is able to detect early stages of bark beetle infestation through the use of semantically enriched Sentinel-2 data and the candidate pixels concept, coupled with spectral vegetation indices (SVIs) and their interactions. We have shown that the combined model, e.g. the SIAM™ candidates + SVI model outperforms the SVI-only model in terms of spectral separability and accuracy. Despite inherent limitations in pixel-level

prediction of early bark beetle infestation, our method offers several advantages. First and most important, the introduction of candidate pixels by applying semantic enrichment of Sentinel-2 data distinctively improved the accuracy of the model and reduced the error of commission, meaning false-positive pixels. Although the input parameters for the rule set were meticulously prepared through feature engineering, the proportion of false positives in the SVI-only model remained considerably high. Specifically, the combined model had a 39% share of false positives, whilst the SVI-only model had an even higher rate of 58% at the plot level. In complex topographies such as Berchtesgaden National Park, where the large height gradient, steep terrain, and resulting fragmentation of canopy cover lead to disconnected and patchy infestation plots, pre-selection of candidate pixels seems to be essential for achieving a relatively high success rate. Moreover, shadowing, snow cover fragments, and frequent cloud cover necessitate novel approaches beyond SVIs for early detection. In contrast to the study sites of most other studies (e.g. low mountain range in Abdullah et al. (2019b) or Bárta et al. (2021) and lowlands in Huo et al. (2021)), the topography in Berchtesgaden National Park hampered the remote sensing-based detection of bark beetle infestation. Therefore, the pre-selection of candidate pixels is crucial in challenging terrains, although the exclusive use of SVIs may prove sufficient in topographically simpler areas.

The SVI-only model classified a substantially larger number of pixels as early infested than the combined model. Nevertheless, the number of matching pixels was lower, resulting in lower accuracy. The larger number of pixels classified as early infested could be attributable to other stress indicators such as drought-induced stress, to microclimatic effects in mountain forests (Hofmeister et al. 2019) or to a general lack of sensitivity of SVIs in highly fragmented and diverse landscapes. SVIs alone did not seem to be informative enough to distinguish such biophysical and biochemical changes caused by biotic disturbance agents like bark beetles. However, in interaction with the changes in the spectral categories of SIAM™, many of the pixels classified as false positives were filtered out. Whilst the spectral reflectance differences between pixels classified as early infested by the combined model and the SVI-only model were relatively small, these differences had a significant impact on the higher accuracy of the combined model. This is likely due to the naturally subtle spectral variations between pixels classified as early infested and those classified as healthy.

Besides the higher accuracy and the reduction of false positives of the combined model, another advantage of our approach is the use of dynamic thresholds (compare supplementary material 1), which were adapted to the respective atmospheric and phenological conditions and not determined by trial-and-error, as e.g. in Zimmermann and

Hoffmann (2020). This allows the thresholds to be adjusted if the results of the model are not considered satisfactory.

The proposed method's usability is enhanced by the use of freely available Sentinel-2 data, which can be employed by everyone. There is no need for commercial VHR satellite datalike SPOT-5 or RapidEye data (as e.g. in Abdullah et al. 2019a), WorldView data (as e.g. in Immitzer and Atzberger (2014)) or hyperspectral data (as e.g. in Einzmann et al. (2021) or Hellwig et al. (2021)). Moreover, the semantic enrichment and computation of Sentinel-2 derivates can be executed in a highly automated and efficient manner, allowing for quick and easy implementation of this approach in operational settings.

Finally, the spectral categories resulting from the semantic enrichment allow for computing a reliable coniferous mask, representing the forest type targeted by bark beetles. Although the coniferous mask does not claim to enhance detection per se, it still helps to narrow down the potentially infested pixels and thus increases the computational efficiency of the rule set. Despite a more detailed forest dataset at the species level might improve the results, such a dataset is often not available, which restricts the application in data-poor regions. Given that most other studies (amongst others Abdullah et al. (2019b), Bárta et al. (2021) and Huo et al. (2021)) make use of very detailed forest stand delineation or even species compositions maps, our approach's accuracy can be deemed satisfactory.

With these advantages in mind, we want to emphasize several factors important to consider when applying this method. First, the workflow relies on alterations in SIAM™ spectral categories. Undoubtedly, the semantic enrichment can only indicate changes that are related to the spectral behaviour of vegetation and does not incorporate sophisticated concepts of biophysical or biochemical changes. Therefore, the term green-attack may be somewhat misleading here, which is why we used the term *early detection* in this study. Moreover, bark beetle attacks are a continuous process, meaning that various environmental factors drive this particular ecosystem dynamic in time and space (Bárta et al. 2021), which further impedes the reliable detection and evaluation. That said, the temporal accuracy assessment should only be considered as a first indicator but should not be over-interpreted. The fairly wide-scattering of ± 3 weeks from the day of ground truth recording is assumed to be mainly related to the number of satellite scenes per month and thus to data availability. Earlier detection compared to ground truth data might indicate previously acting stressors like drought or nutrient deficiency which leads to reduced tree vitality. Another influencing factor is certainly the interval of the ground surveys. Within the study site, the areas of the bark beetle management zone are walked every 4 weeks by field workers. The relatively large interval might lead to biases of the temporal accuracy assessment

(in both directions). This depends primarily on how well the acquisition date of the Sentinel-2 data matches the field data recording.

From a spatial perspective, the high error of commission (39%) at the plot level in the spatial accuracy assessment might be attributed to the relatively early image flight campaign (early September), to which the grey-attack machine learning approach is applied to and from which the grey-attack reference dataset results. Especially from July onwards, the accuracy assessment is subject to uncertainties, as it is highly likely that trees will not have discoloured when reference data was acquired and thus will not be detected by the grey-attack algorithm, even if these trees are infested. All pixels classified as *early infested* from July and later, but not appearing in the grey-attack dataset, would thus require re-analysis with data from subsequent years. Moreover, field recording of early infestation is not always "obvious", meaning that even field sights are error-prone. Thus, the absence of reference data on a spot classified as early infested does not necessarily indicate an error of omission, but may equally be due to difficulty in locating infested spruce trees, especially in a challenging terrain such as in the Berchtesgaden National Park.

In addition, indices used as proxies need to be viewed critically, regardless of whether they are multi-dimensional indices such as the Tasseled Cap components or uni-dimensional indices. The "lack of confidence" is attributable to the difficult understanding and interpretation of them. Of course, they indicate processes like a decrease in the water availability. However, they also reflect atmospheric and landscape variations that are hard to filter out or to map to fixed points on the index scale (Moffiet et al. 2006). Hence, there is no guarantee that these proxies actually represent what they are intended to be used for and their information content is therefore limited. The double check with representative SIAM™ candidate pixels tries to counteract this with a 'convergence of evidence' strategy, but a reliable and space-based determination of the water or chlorophyll content of trees is not yet possible via satellite remote sensing (Van Leeuwen and Orr 2006). These uncertainties are also reflected in the high error of commission in the SVI-only model and in the mismatch between the spectral separability and the accuracy of the RF model of SVIs, which we discuss in more detail in chapter 4.2. That said, approaches incorporating SVIs or any other proxies for biochemical and biophysical changes in the canopy are subject to uncertainties whose investigation is still not fully explored. Even though the use of indices has proven its great potential over decades and in a variety of application (Huo et al. 2020, 2021; Immitzer and Atzberger 2014; Kefalas et al. 2018; Potterf et al. 2015), one should be aware of the ambiguities associated with the use of SVIs.

Regarding our assessment of feature importance of SVIs using RF, we found strong fluctuations of accuracy and OOB error for some epochs. The results suggested low accuracy and hence high OOB errors for the epoch March–April and September–October, which indicates a less reliable determination of feature importance in these months. The low accuracy might be attributable to the use of best-available-pixel-composites in March and October for deriving SVIs.

4.2 Impact of Early Infestation on Spectral Features

The results of the spectral separability between pixels classified as healthy and early infested for the entire year (Fig. 8) revealed best separability for the NIR and SWIR range of the EM spectrum. Both, bands and SVIs showed varying sensitivity for differentiating between pixels classified as early infested and healthy in the course of the year. Whilst the SWIR bands uncovered consistently good separability (see Table 4) in the early months of the year, the NIR bands became increasingly important from June onwards. The SWIR ranges are particularly useful in detecting changes in leaf water content, which can be an early indicator of tree stress (Bowman 1989). During infestation, bark beetles bore into the bark and create galleries in the phloem tissue, which is responsible for transporting sugars and other nutrients throughout the tree. As a result, the tree's ability to transport water and nutrients is compromised, leading to changes in its physiological and biochemical properties. Besides damaging the phloem, bark beetles also introduce a fungus into the tree, which can further damage the vascular system. The fungus can cause discoloration and decay of the wood, which leads to a reduced structural integrity of the tree (Ayres and Lombardero 2000). These changes can be detected by analyzing the reflectance spectra of the tree in the SWIR range, which is sensitive to changes in the water content and biochemical composition of plant tissues. The reflectance spectra of infested trees show a decrease in water content and an increase in the absorption of light by organic compounds such as lignin and tannins (Abdullah et al. 2019b; Raffa et al. 2008). As the year progresses, the relevance of the NIR ranges in detecting the physiological changes caused by bark beetle infestation increased. One common response of trees to bark beetle attacks is a reduction in photosynthetic activity, which can cause a decrease in NIR reflectance (Coops et al. 2009). Additionally, the disruption of the flow of sap and water through the tree lead to changes in the structure of the bark and tissues. These changes can affect the interaction between NIR radiation and the tree, leading to alterations in the reflectance signal that can be detected by remote sensing instruments (Raffa et al. 2008).

Regarding the SVIs, DSWI showed by far the highest potential for separability, independent from time (see Table 4). The other two water stress-related indices (NDWI and TCW) revealed less clear temporal patterns, yet they still proved medium to high separability for most of the time steps observed as shown by the t-test. The separability of spectral bands may not be reflected in SVIs (e.g. water stress-related indices are assumed to be related to SWIR-1/2) due to band aggregation and offsetting with other bands (see Table 2). The DSWI, for example, includes NIR, green and red in addition to SWIR, and the NDWI includes green and NIR.

The partially observable mismatch between the accuracy of the RF model for determining feature importance for some epochs (e.g. April 23rd, May 8th and June 12th) and the nevertheless high separability in the SWIR range (Fig. 9) can be attributed to the potential information loss when aggregating spectral values to a scalar variable (SVIs). In the specific case of May 8th and June 12th, it appears that relevant information carried by certain wavelength regions was lost when computing according water stress-related SVIs. The apparently good spectral separability of the classes early infested and healthy in April (see Fig. 9), especially in the SWIR ranges might additionally be driven by the use of satellite composites. That said, it is rather unlikely that enough trees are already infested between March and April, leading to such a clear signal as it is the case here. The recorded mean daily temperatures of 12.8 °C at 1000 m and 11.1 °C at 1400 m during April, as obtained from two weather stations in the national park, confirmed our assumption. These temperatures were below the 16.5 °C threshold required for bark beetle swarm activity (Wermelinger 2004). Besides the use of best-available-pixel composites which might introduce some bias, it is cloud (and partly snow) coverage, that affects the separability between pixels classified as healthy and early infested.

Referring to chlorophyll- and cell structure-related indices, NDI45 showed good separability in May and June, whilst NDRE3 and NGRDI exposed higher separability in the further course of the year. The overall clearer and stronger signal of the SWIR bands and the water stress dependent indices in terms of separability of pixels classified as healthy and early infested fits well with the results in Figs. 9 and 11, which show the intra-annual differences in the reflectance of the bands and the feature importance of the individual indices. For most time steps, the variation of SWIR reflectance was smaller than for the NIR bands (Fig. 9) and feature importance of water stress-related indices was consistently high (Fig. 11). This corresponds with findings in Huo et al. (2022), who found larger variations for early infested pixels in the NIR ranges.

4.3 Spatial and Radiometric Resolution of Sentinel-2 Data

A common phenomenon in remote sensing of ecosystems is the mismatch between the resolution of the sensor and the ecological process being observed (Senf 2022). The same applies for our study, where mixed pixels are expected due to Sentinel-2's spatial resolution of 10 m. Detection methods—no matter if early detection or grey-attack detection—may come to its limit when there are pixels present that do not purely represent spruce stands (Meddens et al. 2013). An assessment of spectral mixture effects would require very high resolution (VHR) data and the development of a representative spectral library to apply a spectral unmixing approach and hence to disentangle the fractional cover types contained in one pixel (Okujeni et al. 2017). However, further research is needed to assess the potential of spectral unmixing approaches for early detection of bark beetle infestation. Commonly, high spatial resolution is necessary when bark beetle infestations occur at endemic levels, as the infested areas are usually smaller, isolated patches. In contrast, the relevance of spatial resolution decreases when the infestation is epidemic, since the infested areas are then mostly large, contiguous patches (Fernandez-Carrillo et al. 2020; Kautz et al. 2011). As shown by our results, the spatial accuracy at the pixel level is insufficient. Both the error of omission and commission is high, at 83% and 61%, respectively, indicating that it is not possible to make reliable predictions at the pixel level. This can be attributed to the spatial resolution of Sentinel-2 data and the challenges associated with mixed pixels. Nevertheless, when using plot-level (defined as $> = 2$ pixels) accuracy metrics, the accuracy distinctively improves, highlighting the potential of freely available medium-resolution satellite data in detecting epidemic bark beetle infestations.

5 Conclusion and Outlook

Our proposed method provides a novel prototypical tool to detect forest stress caused by bark beetles leveraging spatial, temporal and spectral characteristics of remotely sensed data. The method can be considered a form of "hybrid AI" (Gevaert 2022), i.e. combining a deductive, teaching-by-rules approach with an inductive, learning-from-samples method. The distinctive feature of the method is the use of semantically enriched Sentinel-2 data via the highly advanced, knowledge-based software SIAM™ and the derivation of indicator values, which contributed significantly to the success of this study. Unlike the use of vegetation indices alone, which aggregate spectral values to a scalar variable, SIAM's spectral categories consider the full feature space. This unique research design gets along with a minimum number of additional data, e.g. there is no need for a detailed species composition map or commercial high-resolution satellite data, promoting broad application options. Further advantages of using SIAM™ is the complete automation of the software, which does not require any parameters from users or samples. In addition to Sentinel-2 data, all other at least top-of-atmosphere reflectance optical satellite data can be potentially used as input, including VHR data exhibiting infrared bands. Thus, the workflow can also be adapted for other scale levels.

Results from our approach proved that especially water stress-related indices as well as the SWIR ranges of the EM spectrum show high sensitivity for early bark beetle detection. Future development shall focus on building up a samples database in order to apply more sophisticated and self-adaptive learning algorithms, which are currently state-of-the-art in recognition performance. This allows also to run the tool in real-time, i.e. without a priori data (here the grey-attack dataset). Alternatively, this approach can be repeated for several years to derive more robust index-specific thresholds, thus staying with a purely index- and semantic enrichment-based method. Regardless of the choice of the methodology, a monitoring system that detects bark beetle infestations at an early stage makes an important contribution both to the prevention of large-scale tree mortality and to ecological research. The interaction of modern, satellite-based technologies and the on-site knowledge of forest experts has proven to be very beneficial and shows that ecological field research and remote sensing-based monitoring complement each other very well and profitably.

Declarations

References

Abdullah H, Skidmore A, Darvishzadeh R, Heurich M (2019a) Timing of red-edge and shortwave infrared reflectance critical for early stress detection induced by bark beetle (*Ips typographus*, L.) attack. Int J Appl Earth Observ Geoinf 82:101

Abdullah H, Skidmore AK, Darvishzadeh R, Heurich M (2019b) Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (Ips typographus, L.) compared with Landsat-8. Remote Sens Ecol Conserv 5:87–106

Alvarez-Vanhard E, Corpetti T, Houet T (2021) UAV & satellite synergies for optical remote sensing applications: A literature review. Science of Remote Sensing 3:100019

Ayres MP, Lombardero MJ (2000) Assessing the consequences of global change for forest disturbance from herbivores and pathogens. Sci Total Environ 262:263–286

Baraldi A (2011) Satellite image automatic Mapper™ (SIAM™) - a turnkey software executable for automatic near real-time multisensor multi-resolution spectral rule-based preliminary classification of spaceborne multi- spectral images. Recent Patents Space Technol 1:81–106

Baraldi A, Puzzolo V, Blonda P, Bruzzone L, Tarantino C (2006) Automatic spectral rule-based preliminary mapping of calibrated landsat TM and ETM+ images. IEEE Transactions Geosci Remote Sens 44:2563–2586

Baraldi A, Sapia LD, Tiede D, Sudmanns M, Augustin HL, Lang S (2022) Innovative Analysis Ready Data (ARD) product and process requirements, software system design, algorithms and implementation at the midstream as necessary-but-not-sufficient precondition of the downstream in a new notion of Space Economy 4.0 - Part 1: Problem background in Artificial General Intelligence

(AGI). Big Earth Data. https://doi.org/10.1080/20964471.2021.2017582

Bárta V, Lukeš P, Homolová L (2021) Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. Int J Appl Earth Observ Geoinf 100:102335

Bowman WD (1989) The relationship between leaf water status, gas exchange, and spectral reflectance in cotton leaves. Remote Sens Environ 30:249–255

Breiman L (2001) Random forests. Mach Learn 45:5–32

Buschmann C (1993) Fernerkundung von Pflanzen - Ausbreitung, Gesundheitszustand und Produktivitaet. Die Naturwiss 80(1993):439–453

Christiansen E, Bakke A (1988) The Spruce Bark Beetle of Eurasia. In: Berryman AA (ed) Dynamics of forest insect populations: patterns, causes, implications. Springer, Boston, pp 479–503

CODE-de.org (2023) DataCube-Funktionalität für CODE-DE. https://code-de.org/en/news/datacube-functionality-on-code-de/. Accessed 19 Feb 2023

Coops N, Waring R, Wulder M, White J (2009) Prediction and assessment of bark beetle-induced mortality of lodgepole pine using estimates of stand vigor derived from remotely sensed data. Remote Sens Environ 113:1058–1066

Coppin PR, Bauer ME (1996) Digital change detection in forest ecosystems with remote sensing imagery. Remote Sens Rev 13:207–234

Coulson RN (1985) Forest bark beetle interactions: bark beetle population dynamics. Integrated pest management in pine-bark beetle ecosystems. Springer, New York, pp 61–80

Dalponte M, Solano-Correa YT, Frizzera L, Gianelle D (2022) Mapping a European spruce bark beetle outbreak using sentinel-2 remote sensing data. Remote Sensing 14:3135

Delegido J, Verrelst J, Alonso L, Moreno J (2011) Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. Sensors 11:7063

Edburg SL, Hicke JA, Brooks PD, Pendall EG, Ewers BE, Norton U, Gochis D, Gutmann ED, Meddens AJH (2012) Cascading impacts of bark beetle-caused tree mortality on coupled biogeophysical and biogeochemical processes. Front Ecol Environ 10:416–424

Einzmann K, Atzberger C, Pinnel N, Glas C, Böck S, Seitz R, Immitzer M (2021) Early detection of spruce vitality loss with hyperspectral data: results of an experimental study in Bavaria. Germany Remote Sens Environ 266:112676

Fassnacht FE, Latifi H, Koch B (2012) An angular vegetation index for imaging spectroscopy data—Preliminary results on forest damage detection in the Bavarian National Park, Germany. Int J Appl Earth Obs Geoinf 19:308–321

Fassnacht FE, Latifi H, Ghosh A, Joshi PK, Koch B (2014) Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. Remote Sens Environ 140:533–548

Fer I, Gardella AK, Shiklomanov AN, Campbell EE, Cowdery EM, De Kauwe MG, Desai A, Duveneck MJ, Fisher JB, Haynes KD, Hoffman FM, Johnston MR, Kooper R, LeBauer DS, Mantooth J, Parton WJ, Poulter B, Quaife T, Raiho A, Schaefer K, Serbin SP, Simkins J, Wilcox KR, Viskari T, Dietze MC (2021) Beyond ecosystem modeling: a roadmap to community cyberinfrastructure for ecological data-model integration. Glob Change Biol 27:13–26

Fernandez-Carrillo A, Patočka Z, Dobrovolný L, Franco-Nieto A, Revilla-Romero B (2020) Monitoring bark beetle forest damage in central europe. A remote sensing approach validated with field data. Remote Sens 12:1–19

Filchev L (2012) An assessment of European spruce bark beetle infestation using worldview-2 satellite data. Conference proceedings of European SCGIS Conference, Sofia, Bulgaria. https://doi.org/10.13140/2.1.3005.2647

Frantz D (2019) FORCE—Landsat + Sentinel-2 analysis ready data and beyond. Remote Sens 11:1124

Galvão LS, Formaggio AR, Tisot DA (2005) Discrimination of sugarcane varieties in Southeastern Brazil with EO-1 hyperion data. Remote Sens Environ 94:523–534

German Aerospace Center (DLR) (2020) Sentinel-2 MSI - Level 3A (MAJA/WASP Tiles)-Germany. German Aerospace Center

Gevaert CM (2022) Explainable AI for earth observation: A review including societal and regulatory perspectives. Int J Appl Earth Obs Geoinf 112:102869

Giunta AD, Jenkins MJ, Hebertson EG, Munson AS (2016) Disturbance agents and their associated effects on the health of interior douglas-fir forests in the central rocky mountains. Forests. https://doi.org/10.3390/f7040080

Hellwig FM, Stelmaszczuk-Górska MA, Dubois C, Wolsza M, Truckenbrodt SC, Sagichewski H, Chmara S, Bannehr L, Lausch A, Schmullius C (2021) Mapping European spruce bark beetle infestation at its early phase using gyrocopter-mounted hyperspectral data and field measurements. Remote Sens. https://doi.org/10.3390/rs13224659

Hofmeister J, Hošek J, Brabec M, Střalková R, Mýlová P, Bouda M, Pettit JL, Rydval M, Svoboda M (2019) Microclimate edge effect in small fragments of temperate forests in the context of climate change. For Ecol Manag 448:48–56

Honkavaara E, Näsi R, Alves de Oliveira R, Viljanen N, Suomalainen J, Khoramshahi E, Hakala T, Nevalainen O, Markelin L, Vuorinen M, Kankaanhuhta V, Paivi L-S, Haataja L (2020) Using multitemporal hyper- and multispectral uav imaging for detecting bark beetle infestation on norway spruce. In: ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B3–2020, pp 429–434

Hunt E, Daughtry C, Eitel J, Long D (2011) Remote Sensing Leaf Chlorophyll Content Using a Visible Band Index. Agron J 103:1090

Huo L, Persson HJ, 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 Sens Environ 255:112240

Huo L, Lindberg E, Persson H (2020) Normalized projected red & SWIR (NPRS): a new vegetation index for forest health estimation and its application on spruce bark beetle attack detection. In: IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, pp 4618–4621

Huo L, Lindberg E, Fransson JES, Persson HJ (2022). Comparing Spectral Differences Between Healthy and Early Infested Spruce Forests Caused by Bark Beetle Attacks using Satellite Images. In: International Geoscience and Remote Sensing Symposium (IGARSS), pp 7709–7712

Immitzer M, Atzberger C (2014) Early detection of bark beetle infestation in Norway Spruce (*Picea abies*, L.) using worldview-2 data. Photogramm Fernerkund Geoinf. https://doi.org/10.1127/1432-8364/2014/0229

Jakoby O, Lischke H, Wermelinger B (2019) Climate change alters elevational phenology patterns of the European spruce bark beetle (Ips typographus). Glob Chang Biol 25:4048–4063

Kautz M, Dworschak K, Gruppe A, Schopf R (2011) Quantifying spatio-temporal dispersion of bark beetle infestations in epidemic and non-epidemic conditions. For Ecol Manag 262:598–608

Kefalas G, Lattas P, Xofis P, Lorilla R, Martinis A, Poirazidis K (2018) The use of vegetation indices and change detection techniques as a tool for monitoring ecosystem and biodiversity integrity. Int J Sustain Agric Manag Inform 4:47

Klouček T, Komárek J, Surový P, Hrach K, Janata P, Vašíček B (2019) The Use of UAV Mounted Sensors for Precise Detection of Bark Beetle Infestation. Remote Sens. https://doi.org/10.3390/rs11131561

Koch B (2010) Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment. ISPRS J Photogramm Remote Sens 65:581–590

Lang S, Tiede D, Maier B, Blaschke T (2006) 3D Forest structure analysis from optical and LIDAR data. AMBIÊNCIA 2:95–110

Latifi H, Fassnacht FE, Schumann B, Dech S (2014a) Object-based extraction of bark beetle (Ips typographus L.) infestations using multi-date LANDSAT and SPOT satellite imagery. Progr Phys Geogr 38:755–785

Latifi H, Schumann B, Kautz M, Dech S (2014b) Spatial characterization of bark beetle infestations by a multidate synergy of SPOT and Landsat imagery. Environ Monit Assess 186:441–456

Latifi H, Dahms T, Beudert B, Heurich M, Kübert C, Dech S (2018) Synthetic RapidEye data used for the detection of area-based spruce tree mortality induced by bark beetles. Gisci Remote Sens 55:839–859

Lausch A, Heurich M, Gordalla D, Dobner HJ, Gwillym-Margianto S, Salbach C (2013) Forecasting potential bark beetle outbreaks based on spruce forest vitality using hyperspectral remote-sensing techniques at different scales. For Ecol Manag 308:76–89

Lausch A, Erasmi S, King DJ, Magdon P, Heurich M (2016) Understanding forest health with remote sensing-part I—a review of spectral traits, processes and remote-sensing characteristics. Remote Sens 8:1029

Mandl L, Lang S (2022) Early detection of bark beetle triggered forest stress using Sentinel-2 data. Wissenschaftlich-Technische Jahrestagung der DGPF. Dresden, pp 342–354

Mandl L, Oravec A (2020) Standardized derivation of forest stands using LiDAR data. A case study for the Berchtesgaden National Park. Project documentation National Park administration Berchtesgaden

Marini L, Lindelöw Å, Jönsson AM, Wulff S, Schroeder LM (2013) Population dynamics of the spruce bark beetle: a long-term study. Oikos 122:1768–1776

McFeeters SK (1996) The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int J Remote Sens 17:1425–1432

McRoberts RE, Tomppo EO (2007) Remote sensing support for national forest inventories. Remote Sens Environ 110:412–419

Meddens AJH, Hicke JA, Vierling LA, Hudak AT (2013) Evaluating methods to detect bark beetle-caused tree mortality using single-date and multi-date Landsat imagery. Remote Sens Environ 132:49–58

Mikkelson KM, Bearup LA, Maxwell RM, Stednick JD, McCray JE, Sharp JO (2013) Bark beetle infestation impacts on nutrient cycling, water quality and interdependent hydrological effects. Biogeochemistry 115:1–21

Moffiet T, Mengersen K, King, R, Armston J, Witte C (2006) Modelling of foliage projected cover using Landsat-7 spectral imagery: Spectral indices for greenness and brightness. Conference proceedings of 13th Australian Remote Sensing and Photogrammetry Conference, Canberra, Australia. https://doi.org/10.13140/2.1.4855.0726

National Park Administration Berchtesgaden (2023) Nationalparkplan. https://www.nationalpark-berchtesgaden.bayern.de/service/publikationen/veroeffentlichungen/doc/nationalparkplan_masterplan.pdf. Accessed 28 Mar 2023

Okujeni A, Linden SVD, Suess S, Hostert P (2017) Ensemble learning from synthetically mixed training data for quantifying urban land cover with support vector regression. IEEE J Sel Top Appl Earth Observ Remote Sens 10:1640–1650

Ortiz S, Breidenbach J, Kändler G (2013) Early detection of Bark beetle green attack using TerraSAR-X and RapidEye Data. Remote Sens 5(4):1912–1931

Paczkowsk S, Datta P, Pelz S, Jaeger D (2020) Project PROTECT[FOREST]©: Early detection of bark beetle infestation by

drone-based monoterpene detection. Conference proceedings of The 1st International Electronic Conference on Forests — Forests for a Better Future: Sustainability, Innovation, Interdisciplinarity online. https://doi.org/10.3390/IECF2020-08473

Potterf M, Bucha T, Ferenčík J, Jakuš R (2015) Applicability of a vegetation indices-based method to map bark beetle outbreaks in the High Tatra Mountains. Ann Forest Res. https://doi.org/10.15287/afr.2015.388

R Core Team (2021) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed 22 Mar 2023

Raffa KF, Aukema BH, Bentz BJ, Carroll AL, Hicke JA, Turner MG, Romme WH (2008) Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions. Bioscience 58:501–517

Ranghetti L, Boschetti M, Nutini F, Busetto L (2020) sen2r: an R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data. Comput Geosci 139:1044

Saraswat M (2020) superml: build machine learning models like using python's scikit-learn library in R. https://github.com/saraswatmks/superml. Accessed 22 Mar 2023

Schaeffer SE, Manuel J-L, Sara VR-S, Gerardo C-R, Oscar AA-C, Angel MR-G, Alan E (2021) Detection of bark beetle infestation in drone imagery via thresholding cellular automata. J Appl Remote Sens 15:016518

Seidl R, Thom D, Kautz M, Martin-Benito D, Peltoniemi M, Vacchiano G, Wild J, Ascoli D, Petr M, Honkaniemi J, Lexer MJ, Trotsiuk V, Mairota P, Svoboda M, Fabrika M, Nagel TA, Reyer CPO (2017) Forest disturbances under climate change. Nat Clim Chang 7:395–402

Senf C (2022) Seeing the system from above: the use and potential of remote sensing for studying ecosystem dynamics. Ecosystems. https://doi.org/10.1007/s10021-022-00777-2

Shi T, Xu H (2019) Derivation of tasseled cap transformation coefficients for sentinel-2 MSI at-sensor reflectance data. IEEE J Sel Top Appl Earth Observ Remote Sens. https://doi.org/10.1109/JSTARS.2019.2938388

Sprintsin M, Chen J, Czurylowicz P (2011) Combining land surface temperature and shortwave infrared reflectance for early detection of mountain pine beetle infestations in western Canada. J Appl Remote Sens 5:0566

Stereńczak K, Mielcarek M, Modzelewska A, Kraszewski B, Fassnacht FE, Hilszczański J (2019) Intra-annual Ips typographus outbreak monitoring using a multi-temporal GIS analysis based on hyperspectral and ALS data in the Białowieża Forests. For Ecol Manag 442:105–116

Stritih A, Bebi P, Rossi C, Grêt-Regamey A (2021) Addressing disturbance risk to mountain forest ecosystem services. J Environ Manag 296:113188

Sudmanns M, Tiede D, Lang S, Bergstedt H, Trost G, Augustin H, Baraldi A, Blaschke T (2020) Big Earth data: disruptive changes in Earth observation data management and analysis? Int J Digit Earth 13:832–850

Sudmanns M, Augustin H, van der Meer L, Baraldi A, Tiede D (2021) The Austrian semantic EO data cube infrastructure. Remote Sens 13:4807

Sukovata L, Jaworski T, Plewa R (2021) Effectiveness of different lures for attracting Ips acuminatus (Coleoptera: Curculionidae: Scolytinae). Agric for Entomol 23:154–162

Thom D, Seidl R (2016) Natural disturbance impacts on ecosystem services and biodiversity in temperate and boreal forests: disturbance impacts on biodiversity and services. Biol Rev 91:760

Trubin A, Mezei P, Zabihi K, Surový P, Jakuš R (2022) Northernmost European spruce bark beetle Ips typographus outbreak: modelling tree mortality using remote sensing and climate data. For Ecol Manage 505:119829

Turner MG, Gardner RH, O'Neill RV (1995) Ecological dynamics at broad scales. Bioscience 45:S29–S35

Van Leeuwen W, Orr BJ (2006) Spectral vegetation indices and uncertainty: Insights from a user's perspective. IEEE Trans Geosci Remote Sens 44:1931–1933

Verrelst J, Schaepman ME, Malenovský Z, Clevers JGPW (2010) Effects of woody elements on simulated canopy reflectance: implications for forest chlorophyll content retrieval. Remote Sens Environ 114:647–656

Wermelinger B (2004) Ecology and management of the spruce bark beetle Ips typographus: a review of recent research. For Ecol Manag 202:67–82

Wickham H, Grolemund G (2017) R for data science: import, tidy, transform, visualize, and model data. Reilly Media Inc

Wulder M, Dymond C, White J, Leckie D, Carroll A (2006) Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. For Ecol Manage 221:27–41

Zimmermann S, Hoffmann K (2020) Evaluating the capabilities of Sentinel-2 data for large-area detection of bark beetle infestation in the Central German Uplands. J Appl Remote Sens 14:1–12