remote sensing

MDPI

Article

Detection of Bark Beetle Attacks Using Time-Aggregated Satellite Data with Machine Learning

Shokoufa Zeinali [1,*], Per-Ola Olsson [2], Ted Kronvall [1], Magnus Wiktorsson [1] and Johan Lindström [1]

[1] Centre for Mathematical Sciences, Lund University, Märkesbacken 4, 223 62 Lund, Sweden; ted.kronvall@matstat.lu.se (T.K.); magnus.wiktorsson@matstat.lu.se (M.W.); johan.lindstrom@matstat.lu.se (J.L.)
[2] Department of Earth and Environmental Sciences, Lund University, Sölvegatan 12, 223 62 Lund, Sweden; per-ola.olsson@mgeo.lu.se
* Correspondence: shokoufa.zeinali@matstat.lu.se

Highlights

What are the main findings?

- Combining weekly Sentinel-2 data with static geo- and forestry features using XGBoost significantly improves early bark beetle detection compared to using either data source alone and outperforms Random Forest and prior studies.
- Key spectral bands (Red-edge 3, blue) and specific weeks after main swarming (19, 25, 31, 33) are most informative; static features (land cover, landforms) enhance performance and shift feature importance.

What are the implications of the main findings?

- Integrating temporal satellite data with static environmental information is critical for accurate early detection of bark beetle attacks.
- Identifying important weeks and features supports the development of more effective, timely early warning and monitoring systems.

Abstract

In this study, we explored how early bark beetle attacks can be detected using weekly aggregated Sentinel-2 data in combination with static data, such as geo- and forestry data. We used an XGBoost classifier, known for its strength and reliability in classification, and compared three sets of data: static data only, satellite data only, and using both together. Having trained the models on cumulative weekly data, we were able to track changes in model performance and feature importance over time, identifying key weeks for the detection of bark beetle attacks. A systematic overview of feature importance identified the Red-edge 3 and blue Sentinel-2 bands as the most important when combined with static data; it also showed changes in feature importance compared to using satellite-only data, e.g., adding static features reduced the importance of red and red-edge 2 bands. Among the static features, land cover and landforms were the most important. Evaluating the temporal features for the combined model highlighted certain weeks as containing key information for detection: week 19, which was the main swarming week; week 25, which is 6 weeks after swarming and just before the second generation is completed; and weeks 31 and 33, more than 3 months after the tree was attacked and well after the new generation has swarmed. The study shows that combining static features with cumulative Sentinel-2, accumulated across weeks, are all important for improving the detection of bark beetle attacks, and that such ideas form an important part of early warning systems.

Check for updates

Academic Editor: Michael Sprintsin

Received: 28 April 2026
Revised: 17 June 2026
Accepted: 24 June 2026
Published: 6 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: detecting insect infestations; XGBoost; static features; time series analysis; forest health monitoring; Sentinel-2

1. Introduction

Bark beetle outbreaks represent one of the most severe disturbances affecting coniferous forests, posing significant threats to forest health, ecosystem services, the economic viability of forestry operations, as well as carbon sequestration [1–4]. The European spruce bark beetle (*Ips typographus*) is a phloem-feeding insect whose life cycle and outbreak dynamics are strongly influenced by temperature. Adult beetles locate and attack suitable Norway spruce (*Picea abies*) trees, mate, and lay eggs under the bark; the larvae feed and develop in the phloem, and the brood adults emerge to colonize new hosts. Successful colonization is facilitated by mass attack and, in many cases, by associated fungi that help overcome host defenses [5]. Temperature-dependent development determines the timing of key life-cycle stages such as swarming, mating, and larval development, making the species highly sensitive to inter-annual climatic variability. In southern Sweden, a new generation of bark beetles usually swarm around eight weeks after a tree was colonized [6].

In Europe, the European spruce bark beetle (*Ips typographus*) is estimated to have accounted for 8% of mortality due to natural disturbances in Norway spruce (*Picea abies*) during the period 1850 to 2000 [7]; this proportion increased after 2000 [8]. Recent bark beetle outbreaks have killed nearly 31 million m^3 of spruce trees in Sweden, 50 million m^3 in Czechia, and 86 million m^3 in Germany [9,10]. Unlike the extensive, contiguous outbreak patterns often reported in Central Europe [11], bark beetle damage in Sweden was more spatially dispersed, with many attacked trees scattered across the landscape.

Several studies have investigated the ability of satellite-based remote sensing data to detect beetle damage and attacked trees. Sentinel-2 gave better results than Landsat for early detection of bark beetle-infested trees in Germany [12] and Czechia [13]. Regarding the utility of different spectral bands, ref. [14] developed a new vegetation index, NDRS (normalized distance Red and SWIR), while [15] obtained the best detection results with the tasseled cap wetness (TCW) index. Various areal-based hyperspectral platforms have been used as an alternative to satellites, e.g., [16–18].

From a forest management perspective, it is important to detect attacked trees as early as possible. Several studies (e.g., [15,19–21]) have investigated early detection and the ability to separate trees in the early (green) and late (red) stages of attack. While evaluating how model performance changes with additional temporal observations, these studies do not systematically evaluate which weeks are the most important for detection, using, e.g., feature importance.

The high accuracies for early detection might reflect the detection of stressed trees rather than the true symptoms of early attacks, since there might have been differences between healthy and green attack pixels before an attack [13,14,20]. This has inspired the use of static forest features to identify susceptible trees [22,23], with promising results. Ref. [24] combined static features with Sentinel-2 to detect trees at risk one year before the attack. Using Random Forest models, authors of [24] found a small (5%) increase when adding Sentinel-2 data to static features. Detection accuracy was the same before and after the attack, indicating that the model identified stressed/susceptible trees. Using only Sentinel-2 from single dates, authors of [24] did not investigate how increased availability of Sentinel-2 during the growing seasons affects detection, or which time periods have the greatest influence on detection.

In this paper, we expand on the single-date analysis in [24] by using time-aggregated satellite data combined with static features (data are described in Section 2). We also replace random forests with the computationally more efficient XGBoost (see [25] and methodology in Section 3). The aim of the analysis (Section 4) is to:

1. Systematically evaluate how the temporal information in the satellite data contributes to detecting bark beetle attacks;
2. Investigate which individual weeks are the most important for detection and relate these weeks to the time of bark-beetle swarming as observed by pheromone traps;
3. Determine when time-aggregated satellite data is informative enough to outperform static features for detection; and
4. Investigate the importance of individual spectral bands and static features for detection.

The methodology of this study is deliberately based on well-established machine learning techniques, such as XGBoost and standard pre-processing strategies, to ensure that observed differences in performance are primarily driven by the structure and temporal availability of the input data rather than model complexity. The main contribution of this work lies in the systematic evaluation of how temporally accumulating multi-spectral satellite observations influence the detectability of bark beetle infestations. In particular, our experimental design enables controlled comparisons across different data configurations, including static-only predictors, satellite-only temporal features, and their combination. To analyze how the predictive performance evolves over time and how early reliable detection can be achieved the number of weekly satellite observations seen by the model is gradually increased.

2. Study Area and Datasets

The study area is 180 × 180 km^2 and located in southeastern Sweden (Figure 1). The annual mean temperature is 6 to 8 °C and the average summer temperature is 15 to 16 °C for the period 1991–2020 [26]. The area is dominated by forests with Norway spruce (*Picea abies* (L.) Karst) and Scots pine (*Pinus sylvestris L.*) as the most common species, with only smaller patches of deciduous forests [27]. The hot drought in 2018 triggered the largest documented bark beetle outbreak in Sweden, with 31 million m^3 of spruce trees affected in the south between 2018 and 2022 [10] and strongly affecting the study area.

2.1. Sentinel-2 Data

The satellite images used in this article are derived from Sentinel-2 with a spatial resolution of 10–60 m and a revisiting period of 2–3 days across the study area [28]. A total of 21 weeks of satellite data were retained, covering calendar weeks 15 to 43 (April to late October) of 2018. The images were downloaded as level 1C and preprocessed with the Framework for Operational Radiometric Correction for Environmental Monitoring [29] (FORCE ver. 3.6.5) to calculate surface reflectance and perform image-to-image registration [30]. The preprocessing followed the same steps as in [24]. Only bands with a spatial resolution of 10 and 20 m were included (see Table 1), and all bands with 20 m resolution were resampled to 10 m with nearest neighbor interpolation. The study area was divided into four tiles as depicted in Figure 1. To ensure consistency across tiles, only the calendar weeks with available data in all tiles were included in the analysis. The distribution of available weeks per tile is shown in Figure 2 and correlation between bands can be found in Figure A1 of Appendix A.

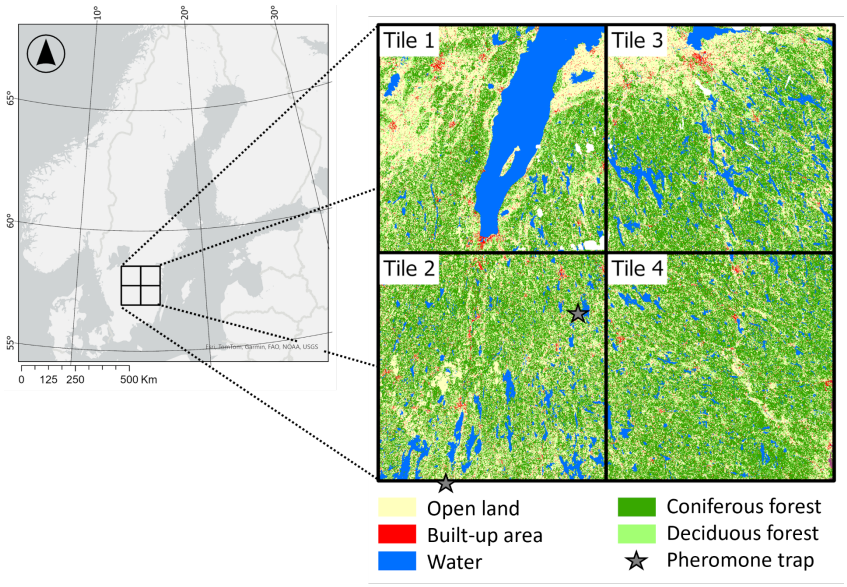


Figure 1. Location of the study area in Sweden divided into four tiles (**left**) and land cover [31] in the four tiles (**right**).

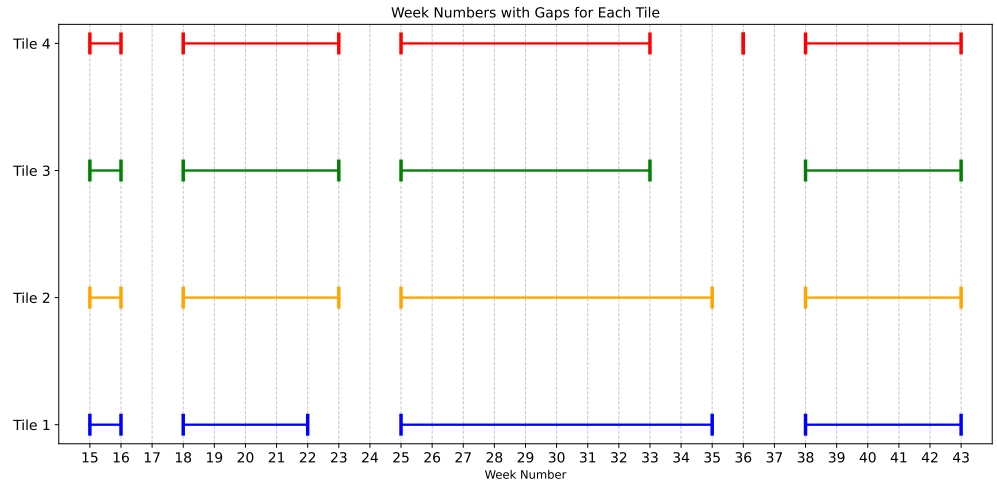


Figure 2. Available weeks of Sentinel-2 data with sufficient cloud-free observations for each tile before preprocessing, where each line and colour represents a different tile, segments show continuous weeks, gaps indicate missing weeks, and vertical bars mark the first and last week of each sequence.

Table 1. Band number, name, spatial resolution, and abbreviations for the Sentinel-2 bands used in this study.

Sentinel-2 Band	Name	Original Spatial Resolution (m)	Abbreviation
2	Blue	10	Blue
3	Green	10	Green
4	Red	10	Red
5	Red-edge 1	20	RE1
6	Red-edge 2	20	RE2
7	Red-edge 3	20	RE3
8	Near Infrared	10	BNR
8a	Narrow NIR	20	NIR
11	Short Wave Infrared 1	20	SWIR1
12	Short Wave Infrared 2	20	SWIR2

2.2. Static Data

We also used geo-spatial, topography, and forestry features that are known to influence the risk of bark beetle infestations [22,24] in this article. Since topography is static and the

forestry parameters change slowly (data is typically updated every few years), we refer to those variables as static, in contrast to the dynamic satellite data that follow the within-season changes in phenology. The static features used here are listed below; they were all pre-processed to the same extent and spatial resolution as the Sentinel-2 data (see [22] for further details). Correlation between the features is illustrated in Figure A2 of Appendix A. These features are not direct spectral indicators of attack symptoms; rather, they provide contextual information on site conditions, stand structure, and forest composition that may show infestation risk and susceptibility of trees.

aspect: The direction that a slope faces.

basalArea: A measure of the cross-sectional area of tree trunks in a forest.

biomass: Total mass of living vegetation in an area.

CHM (Canopy Height Model): Height of vegetation canopies.

DEM (Digital Elevation Model): Elevation data of the terrain, indicating the height above sea level.

Edge: Proximity of a location to the edge of a forest or habitat.

LandC (Land Cover): Type of land cover.

LandForm: Characterizes the shape of the terrain.

slope: The steepness of the terrain.

soilType: Classification of soil.

soilWetness: Moisture content of the soil. Static variable partly derived from a topographic wetness index.

spruceVolume: Volume of spruce trees.

2.3. Response Data

Bark beetle damage data were obtained from Sveaskog, the largest forest owner in Sweden. The location and date of harvest for trees killed by bark beetles were collected by harvester machines; the data do not include the date of attack. We included only trees harvested from April to June 2019, as they were most likely attacked in 2018. The harvester dataset was rasterized to the same 10 × 10 m spatial resolution and extent as the FORCE-processed Sentinel-2 data, with pixels containing at least one attacked tree labeled as attacked, giving a binary response variable (see [24] for further details).

Limited data on forest stands without bark beetle-infested trees, here labeled healthy data, were available as inventory data obtained from the Swedish forest agency. To increase the amount of data on healthy forest stands, we utilized national land cover data with a spatial resolution of 10 × 10 m [31] and a property map [32] in combination with harvester data on attacked trees. The harvester machines were brought to the forest by the management to remove bark beetle-killed trees; hence, we made the assumption that forest stands without harvested trees were healthy if they were located inside a forest estate where trees had been attacked and harvested in other stands. To focus on spruce trees, we only included the land cover classes spruce forest, mixed coniferous forest, and mixed forest when deriving the healthy data. The identified pixels were added as healthy data to the binary raster. The resulting dataset is highly imbalanced; the total number of samples and the percentage of attacked samples across different tiles are shown in Table 2.

Bark beetle swarming data, based on pheromone traps, from the Swedish Forest Agency [33] were used to estimate the timing of attacks (see [24]). Two stations with pheromone traps were located in the study area, one on the southern boundary (Lagan) and one in the central part (Nässjö) of the area (Figure 1). The traps were emptied on a weekly

basis, and in 2018 there was a distinct peak in the swarming data in week 19, indicating that the trees were most likely attacked at that time (the swarming data are included in Figure 3 in the results). In addition, we checked the two pheromone traps closest to the study area; one just west of the area and another near the city of Örebro around 90 km north of the study area, and both followed the same pattern with a very distinct peak during week 19, indicating that week 19 was the main swarming week over the entire study area.

Table 2. Total number of samples and proportion of attacked samples in each tile.

Tile Id	Total Number of Samples	Percentage of Attacked Samples
1	19,882	3.8%
2	28,204	5.2%
3	43,911	0%
4	71,474	6.7%
	163,471	4.3%

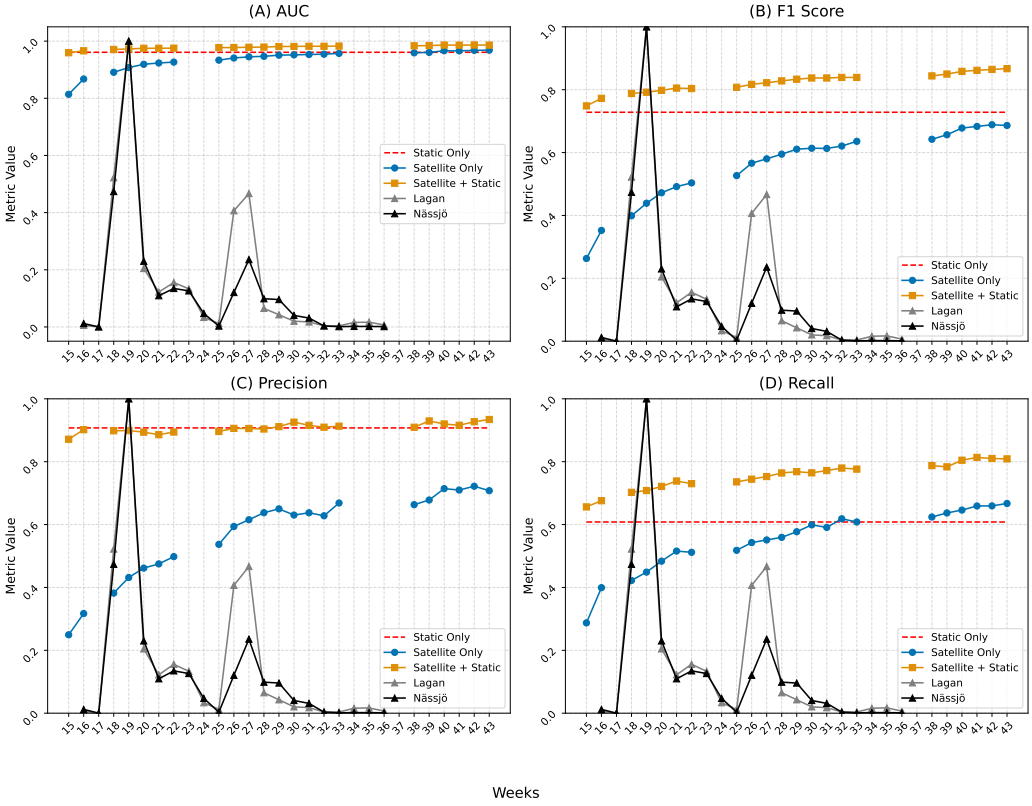


Figure 3. Evaluation metrics over time using: only static data (dashed Red line), only satellite data (blue line/circles), static and satellite data (yellow line/squares). The swarming data for the pheromone traps at Lagan and Nässjö are shown with gray/black lines and triangles; values are scaled so that 1 corresponds to the largest number of captured bark beetles (6000 at Lagan and 11,885 at Nässjö).

3. Methodology

3.1. Sentinel-2 Data Preparation

To obtain a consistent temporal resolution and reduce the effects of missing data and noise—such as cloud cover, cloud shadows, or snow—weekly maxima were computed for each pixel. Temporal aggregation of satellite data using maxima is commonly used in environmental sciences, since maxima have been shown to better indicate ecological processes compared to averages or minima [34]. Additionally, clouds (and cloud shadows)

tend to bias vegetation indices towards lower values [35]. The resulting consistent temporal resolution, consisting of weekly data, facilitates our trend analysis and anomaly detection.

After aggregation, weeks with more than 10% missing values across the study area were dropped. For weeks with less than 10% missing values, the remaining missing values were imputed using the average of the previous three weeks. For the first three weeks in the time series, imputation was performed using forward and backward filling techniques. First, forward filling propagated the first available observation forward, while backward filling assigned the earliest valid observation backward to fill any initial missing values.

3.2. Balancing Response Data

The response dataset is highly imbalanced (see Section 2.3), potentially biasing any classifier towards healthy pixels. We therefore used the Synthetic Minority Oversampling Technique (SMOTE, [36]) to balance the training dataset by creating synthetic examples for the minority class and dropping a proportion of the majority class. It is important to note that the validation data were left untouched, so that the data imbalance is correctly represented when validating the models.

Given a matrix of features, X_{train}, and a vector of labels, y_{train}, SMOTE takes a random sample, x_i, from the minority class and finds the k-nearest neighbors in the minority class. Here, we used $k = 5$ neighbors. Synthetic samples for the minority class were then generated by linearly interpolating between the sample, x_i, and each of the k-nearest neighbors, x_{nn}:

$$x_{new} = x_i + \lambda \cdot (x_{nn} - x_i), \tag{1}$$

where λ is a random number between 0 and 1. This process was repeated, e.g., new random samples x_i were drawn, until the target number of synthetic minority samples was reached. Pseudo-code for the method is given in Algorithm 1.

Algorithm 1 SMOTE Resampling Algorithm

Require:
X_{train}: Feature matrix of training data
Y_{train}: Labels of training data
while more synthetic samples are needed **do**
    Randomly pick a minority class sample x_i from X_{train}
    Find the k nearest neighbors of x_i within the minority class
    for each x_{nn} from the k neighbors **do**
        Generate a random number $\lambda \sim U(0, 1)$
        Compute synthetic sample:
            $x_{\text{new}} = x_i + \lambda \cdot (x_{nn} - x_i)$
        Add x_{new} to X_{new}
        Assign label of minority class to Y_{new}
    end for
end while
Combine the original data and synthetic samples $\hat{X} = \{X_{\text{train}}, X_{\text{new}}\}$ and $\hat{Y} = \{Y_{\text{train}}, Y_{\text{new}}\}$.
Return the resampled $\hat{X}$ and $\hat{Y}$

By interpolating between points in this way, SMOTE creates a more general representation of the minority class, helping to enhance the training of the classifier by providing more examples to learn from.

3.3. XGBoost

XGBoost (Extreme Gradient Boosting), introduced by [25], is a scalable and efficient implementation of gradient boosting decision trees, widely used for classification and regression tasks. Unlike Random Forest, which builds an ensemble of trees in parallel, XGBoost constructs trees sequentially, where each new tree is trained to correct the errors made by the previously built trees. The final prediction is obtained by aggregating the outputs of all trees, typically using a weighted sum for regression or logistic transformation for binary classification. This boosting framework helps minimize both bias and variance, often achieving superior predictive performance compared to Random Forest and other bagging-based methods.

In XGBoost, each tree is built by greedily selecting feature splits that minimize an objective function that combines the training loss with a regularization term that penalizes model complexity. The objective function at iteration t is generally expressed as:

$$\mathcal{L}^{(t)}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t)$$

where $l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i))$ is the training loss for training sample i given the previous prediction $\hat{y}_i^{(t-1)}$ and the new tree $f_t(x_i)$, and $\Omega(f_t)$ is the regularization term for the new tree f_t. A common loss function for binary classification is the logistic, also known as binary cross-entropy, loss:

$$l(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

The algorithm uses a second-order (Newton) approximation of the loss function. By applying a Taylor expansion to the loss function around the current prediction $\hat{y}_i^{(t-1)}$, the objective function can be simplified and optimized in a greedy manner.

To determine the best splits, XGBoost computes the gain, or improvement, in the objective function at each potential split. For a split that divides a node into left (I_L) and right (I_R) child nodes, the gain is calculated as:

$$\text{Gain} = \frac{\left(\sum_{i \in I_L} g_i\right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_R} g_i\right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I_L \cup I_R} g_i\right)^2}{\sum_{i \in I_L \cup I_R} h_i + \lambda} - \gamma$$

where g_i is the first derivative (gradient) of the loss function with respect to the model's prediction for the ith training example and h_i is the second derivative (Hessian) of the loss function with respect to the model's prediction for the ith training example. The **feature importance** is then quantified by the total gain contributed by each feature across all trees in the model. Features that frequently appear in high-gain splits are considered more important, as they contribute more significantly to reducing the model's error.

We used the XGBClassifier from the **XGBoost** (ver. 3.0.2) Python library [37] and Python ver. 3.10. The original satellite imagery was provided in raster format, where each pixel represents a spatial observation, and each raster layer corresponds to a spectral band or derived variable. Data were extracted from these spatial rasters and transformed into a tabular dataset, with one observation per row, one column for each explanatory variable, and one indicator column containing the outcome for each observation (attacked/health). The model was trained with 450 boosting rounds, a logistic loss, and the *average precision* evaluation metric. A fixed random seed was used to ensure reproducibility. Hyperparameters were selected to balance model complexity and performance while preventing overfitting.

3.4. Evaluation

To evaluate model performance on our imbalanced dataset, we utilized metrics that focus on the minority class (attacked pixels). While traditional accuracy can be misleading in imbalanced contexts, we include it here to allow for direct comparison with existing literature. Table 3 summarizes the evaluation metrics used, their mathematical definitions, and their specific relevance to bark beetle detection.

Table 3. Summary of evaluation metrics used to assess model performance for beetle attack detection. When comparing the prediction of an attack with the actual outcome, we divide the results into the classical cases: TP—true positive; TN—true negative; FP—false positive; and FN—false negative.

Metric	Formula	Ecological Significance
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	Included for comparison with other studies.
Precision	$\frac{TP}{TP+FP}$	Minimizes false alarms to ensure resources are not wasted on healthy trees.
Recall	$\frac{TP}{TP+FN}$	Ensures that the maximum number of actual infestations are detected.
F1-Score	$2 \times \frac{Prec \cdot Rec}{Prec+Rec}$	Provides a single measure that penalizes poor performance in either metric.
AUC-ROC	Area under Curve	AUC-ROC (Area Under the Receiver Operating Characteristic curve) measures the model's ability to distinguish between healthy and attacked pixels across all possible classification thresholds.

A calibration curve (or reliability diagram, [38]) is used to assess how well the predicted probabilities from a model reflect the true likelihood of an event. In other words, it measures the agreement between predicted probabilities and observed outcomes. If a model is perfectly calibrated, predictions of probability p should correspond to the observed frequency of the event p. To create a calibration curve, predicted probabilities are typically binned (in this paper we use 10 bins), and for each bin, the mean predicted probability is plotted against the observed fraction of positive outcomes. Deviations from the diagonal line $y = x$ indicate miscalibration.

The Brier score [39] is a numerical measure of calibration and accuracy for probabilistic predictions. It quantifies the mean squared difference between predicted probabilities and actual binary outcomes. Mathematically, for N predictions:

$$\text{Brier Score} = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2$$

where f_i is the predicted probability for observation i and $o_i \in \{0, 1\}$ is the actual outcome. Brier Score ranges from 0 (perfect prediction) to 1 (worst prediction for binary outcomes). An advantage of using the Brier score is that it evaluates both the calibration and refinement of predicted probabilities.

AUC-ROC (Area Under the Receiver Operating Characteristic curve) quantifies the model's ability to discriminate between the two classes across all possible classification thresholds. The ROC curve is constructed by varying the decision threshold and plotting the True Positive Rate against the False Positive Rate. The area under the curve (AUC) can be interpreted as the probability that the model assigns a higher score to a randomly selected attacked pixel than to a randomly selected healthy pixel. An AUC value of 0.5 indicates random performance, while a value of 1.0 corresponds to perfect separability between classes.

To evaluate the contribution of the satellite data and identify the most informative time periods, we began with the initial week and progressively added data from subse-

quent weeks, retraining the XGBoost model at each step to incorporate the new temporal information. This incremental approach allowed us to determine whether certain weeks are particularly important for detecting attacks. Important features of each model were also extracted. For the model with only static features, feature importance was easy to extract. For the models trained on only Sentinel-2 data or a combination of static and Sentinel-2 data, feature importance was found by averaging the importance of each spectral band across all available weeks used in the final model.

4. Results

Following data pre-processing (Section 2) and augmentation (Section 3.2), we trained the XGBoost model using three combinations of the datasets: (1) only static features, (2) only Sentinel-2 data, and (3) a combination of the two. For each Sentinel-2 tile, we standardized the spectral information to mean zero and variance one. Standardization was performed separately for each band and each week.

For all models, we applied stratified five-fold cross-validation to divide the dataset into training and test folds (20% of the data were used in each test fold). The stratification maintained consistent class distributions in all folds, ensuring that the class imbalance ratio is preserved in each split, reduced sampling bias, and provided a more reliable estimate of model generalization under imbalanced conditions. Within each fold, SMOTE was applied only to the training partition, while the corresponding test partition remained untouched and preserved its original distribution. This procedure prevents data leakage from synthetic samples into the evaluation phase and allows the reported metrics to reflect performance on truly unseen, imbalanced data. Cross-validation further ensured that the evaluation was robust to the choice of train–test split and provided stable estimates of model performance.

Our best model, which combined static features and satellite data, achieved an overall accuracy of 98.74%, representing a significant improvement compared to previous studies. However, the main contribution of this paper is the following analysis of feature importance and model performance. A comparison with other classifiers as well as an ablation study comparing different missing-value imputation strategies are provided in Appendices B and C.

4.1. Feature Importance

Feature importance for the different models is illustrated in Figure 4. Among the static features, land cover was by far the most influential predictor, followed by canopy height (CHM) and soil type. When combining satellite data and static features, land cover is still the most important static feature, but it is much less dominant, with the combined model giving a markedly higher weight to landforms, soil type, and spruce volume.

For the model using only Sentinel-2 data, the Red, Blue, and Red-edge bands are the most important. However, the difference between the most and least important bands is small. For the combined model, the Red-edge 3, Blue, and Near Infrared bands are the most important. The combined model also shows a larger difference in feature importance between bands, with the Red-edge 3 band being the most important by a wide margin.

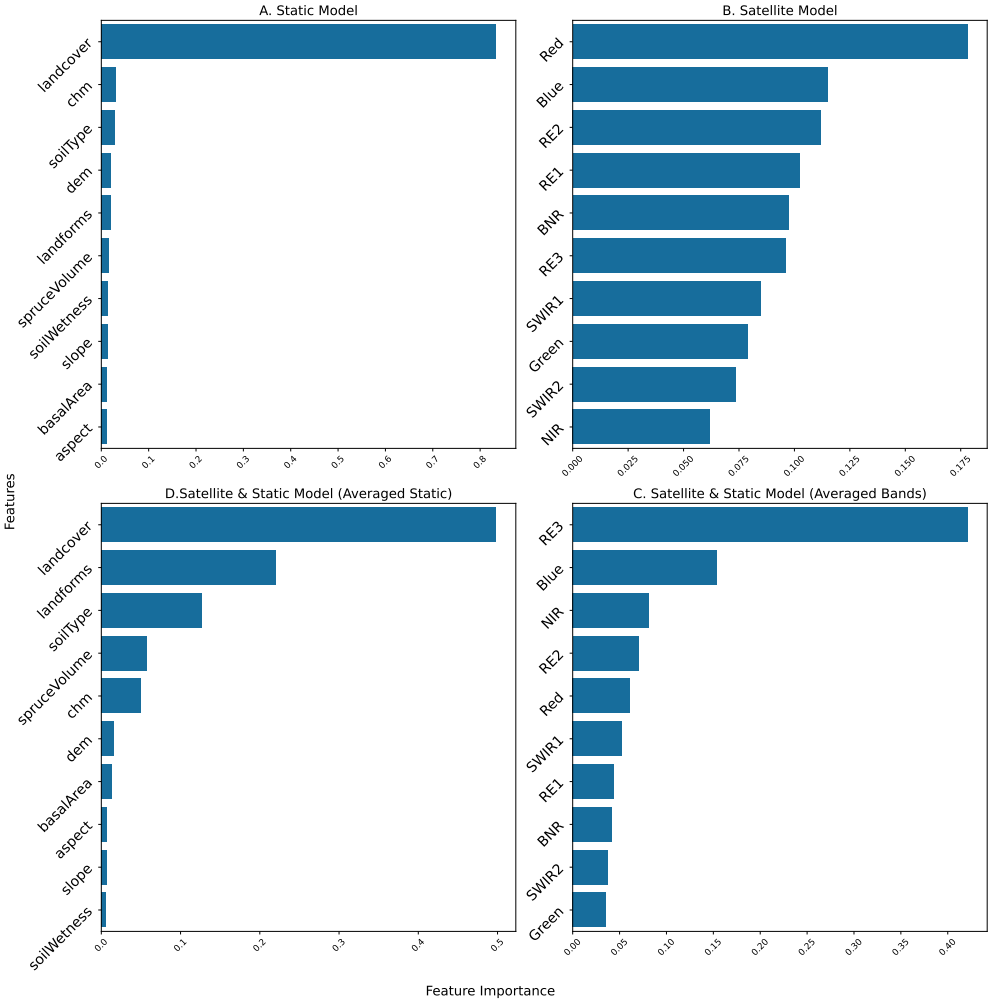


Figure 4. The 10 most important features: (**A**) using static data, (**B**) using satellite data, and (**C**,**D**) using both satellite and static data, where the mean importance is calculated by averaging the importance of each spectral band across all available weeks used in the final model. Note that feature importance is relative, and it is difficult to compare absolute values between models.

4.2. Model Performance over Time

Following the pre-processing in Section 3.1, 22 weeks of satellite data were retained, covering calendar weeks 15 to 43 (mid-April to late October), with some weeks missing due to cloud cover (see Figure 2). The improvement in model evaluation metrics and ROC curves as more weeks are incrementally added is shown in Figures 3 and 5. As expected, model performance increases as additional satellite data become available.

When satellite data are added to the static features, we generally expect the model's performance to improve or at least remain stable, since it has access to more information. This expectation is supported by our results: the combined model either outperforms (for AUC, F1 score, and recall) or performs as well as (for precision) the static-only model throughout the study period.

In contrast to this, the satellite-only model struggles to match the other models. It reaches parity with the static-only model either late, around week 33 (for AUC and recall) or not at all (for F1 score and precision).

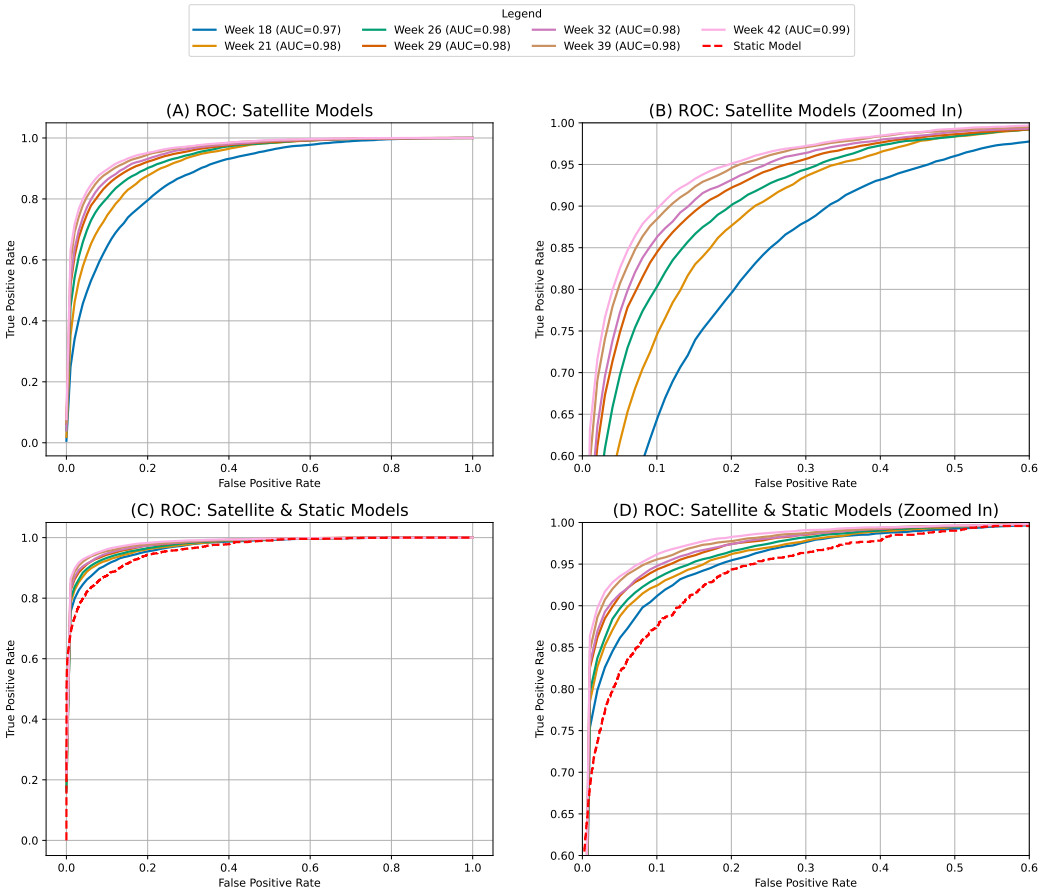


Figure 5. ROC curves comparing the performance of models trained with different input data types across multiple weeks. To improve legibility, only every 3rd week is shown. ROC curves for models trained using only satellite data (**A**) and zoomed-in view of these ROC curves (**B**). ROC curves for models trained using both static and satellite data (**C**), with the static-only model shown as a dashed red line, and zoomed-in view of these ROC curves (**D**).

Comparison of the Satellite Only and Combined Models

To assess whether the differences between the combined and satellite-only model are statistically significant, we used bootstrap resampling, calibration curves, and Brier scores.

Bootstrap was carried out on the unseen test data by repeatedly drawing samples with replacement and computing the difference in evaluation metrics between the two models. The average paired differences and 95% confidence intervals show that the combined model is significantly better across all metrics; AUC values are shown in Figure 6, and the other metrics can be found in Figure A6 of Appendix D.

Calibration curves were used to examine the relationship between predicted and observed outcomes. A well-calibrated model should produce probabilities that closely match the true frequencies. The calibration curves and Brier scores (see legend) in Figure 7 clearly show that the combined models consistently outperform the satellite-only models. The Brier score directly evaluates how close the predicted probabilities are to the truth, making it particularly useful when working with imbalanced datasets and risk estimation.

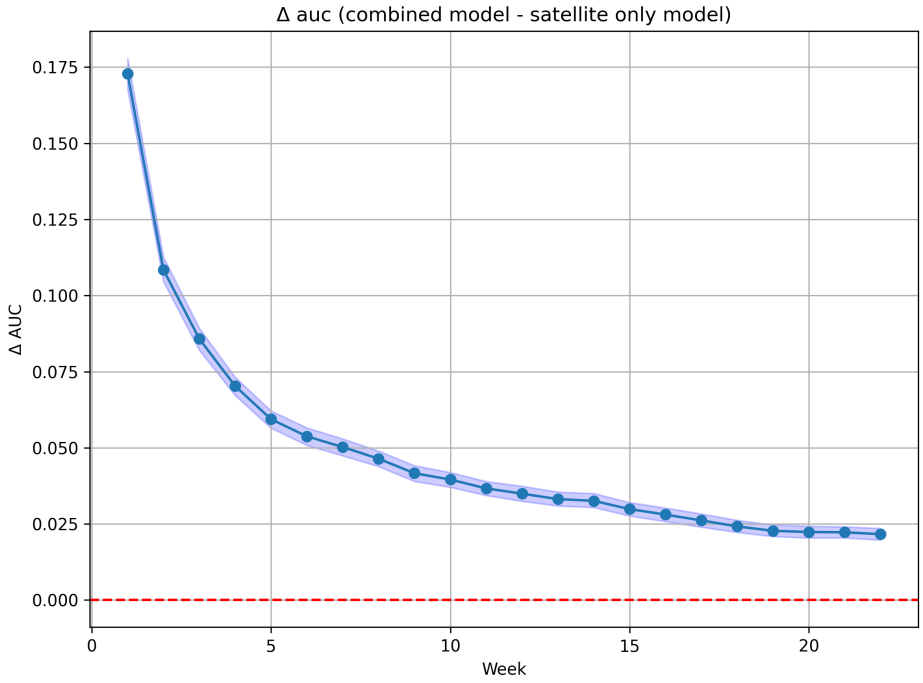


Figure 6. Difference between bootstrapped AUC estimates and paired 95% confidence intervals on the unseen test data. Values above zero (the red dashed line) indicate that the combined model significantly outperforms the satellite-only model.

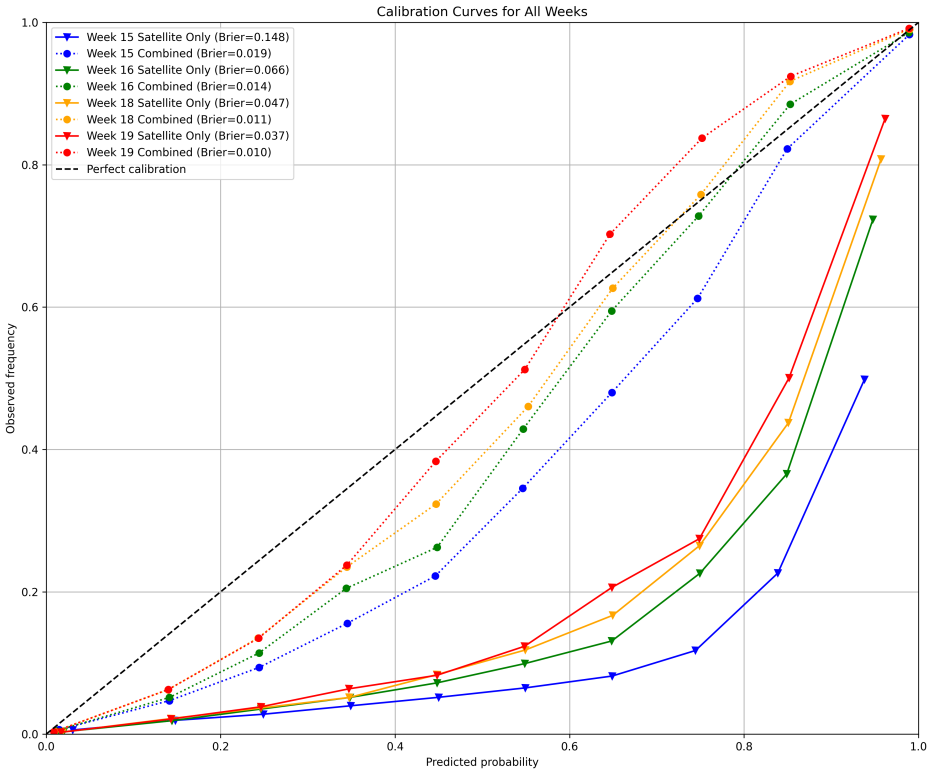


Figure 7. Calibration curves of some representative weeks for satellite-only and combined models; values closer to the 1-to-1 line are better. Brier scores for each model are presented in the legend; lower values are better.

4.3. Most Important Weeks

We used feature importance to systematically evaluate how temporal information contributes to detecting bark beetle attacks and to identify the most important weeks.

For the combined model, using both static features and satellite data, we averaged the importance scores for features belonging to the same week within each model, following the same principle used for band-level aggregation in Section 4.2.

The five most important weeks per model are illustrated using a heatmap; see Figure 8, where the models are represented by columns and the swarming data shown to the right. Several temporal patterns are obvious in the heatmap. Week 19 (swarming week) consistently emerges as highly important, suggesting that it captures early signals of bark beetle activity. Week 25 (six weeks after swarming), and to a lesser extent week 27, initially show high importance. Week 27 is eight weeks after the main swarming, and the second peak in the swarming data suggests that this is when the second generation was swarming. Once data from the highly important weeks 31 and 33 are included in the modeling, the importance of weeks 25 and 27 declines, in contrast to week 19, which remains relevant throughout the period.

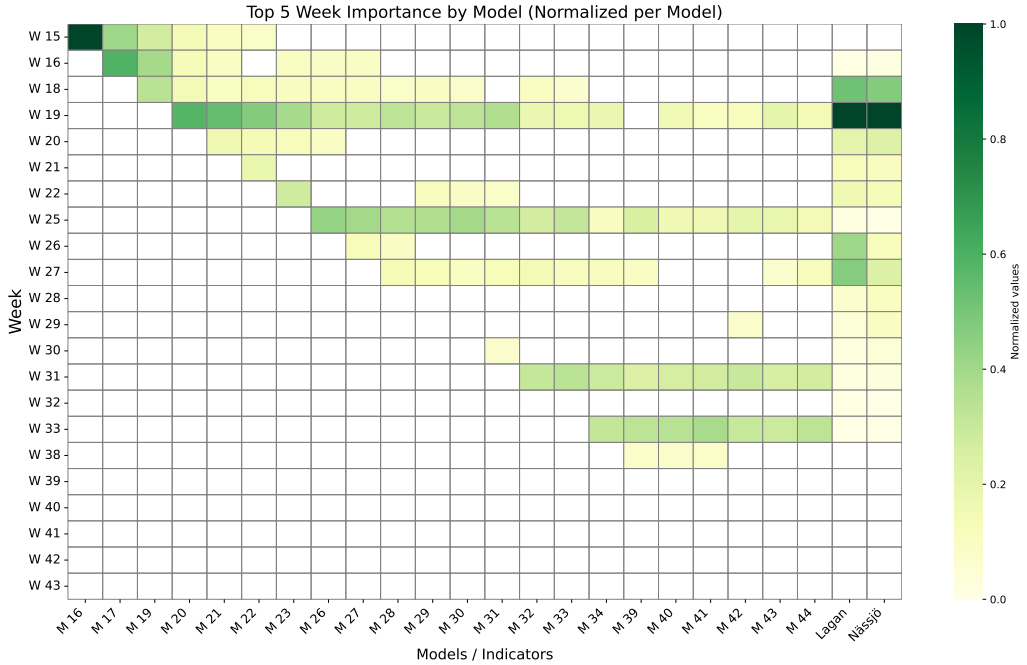


Figure 8. Heatmap showing the five most important weeks per model. The color intensity represents the normalized importance of each week's features (rows) within the respective model (columns). Note that the lower left part of the heatmap is empty since models are not allowed to use future weeks, i.e., Model 32 only uses data from week 15 to 32. Normalized swarming data are also shown.

4.4. Random Forest as an Alternative to XGBoost

Random Forest is a popular method for classification of satellite data, and for completeness, the analysis was also run using Random Forests. Since the Random Forest achieved similar or worse results at substantially higher computational cost, those results are only briefly presented in Appendix B.

5. Discussion

This study investigates the early detection of bark beetle attacks using a combination of Sentinel-2 satellite imagery and static environmental features within an XGBoost-based framework. The results demonstrate that integrating multi-source data consistently improves predictive performance compared to models relying solely on either static or satellite-derived features. This highlights the complementary nature of structural landscape information and temporally evolving spectral signals. The accuracy of 98.74% obtained by our best model compares favorably to other similar studies [12–15,19,20,24] that report

accuracies in the range of 36% to 96%. However, comparisons of accuracies between studies should be taken with a grain of salt due to differences in datasets and experimental settings.

An important characteristic of our study area is that the bark beetle infestation was less intense and more spatially dispersed compared to the large, spatial contiguous outbreaks common in Central Europe [11]. Thus, the attacked trees in this study often occur as isolated individuals or in small groups, likely resulting weak spectral signals within each Sentinel-2 pixel.

A key finding is that static features provide strong predictive power for the susceptibility to attack in the early stages of the infestation, when spectral changes in vegetation are still weak or not yet observable. Variables such as land cover, landforms, and canopy structure likely capture predisposing environmental conditions that influence tree susceptibility to attack. In particular, land cover remains the most influential feature across all models, although its dominance decreases when satellite data are included. This suggests that while static conditions define baseline vulnerability, temporal spectral observations progressively refine detection by capturing physiological stress responses.

As the season progresses and more Sentinel-2 observations become available, spectral information increasingly contributes to model performance. The Red-edge and Red bands emerge as particularly important, which is consistent with previous studies linking these wavelengths to chlorophyll content and vegetation stress [16,18,40,41]. The strong contribution of Red-edge 3 in the combined model further suggests that subtle physiological changes are better captured when multi-source information is jointly exploited. The Blue band shows moderate importance, aligning with mixed findings in the literature, where its contribution appears to be context-dependent [15,20,24,40].

Temporal analysis of feature importance reveals that specific weeks (19, 25, 31, and 33) contain disproportionately high predictive value. Week 19 corresponds to the peak swarming period, where no visible attack symptoms are expected; its importance likely reflects pre-dispositional stress conditions that increase susceptibility. Week 25, just before the new generation leaves the attacked tree, shows a strong contribution to detection performance, which aligns with reported delays between infestation and observable spectral response in previous studies [17,21]. However, this period is also affected by missing observations in weeks 23–24, which may amplify the apparent importance of week 25.

The combined model shows improved stability in recall during early weeks compared to the satellite-only model, indicating that static features help mitigate missing or weak temporal signals. This suggests that static data provide a regularizing effect, improving robustness under incomplete or noisy temporal observations. Later weeks (31 and beyond) contribute less to performance improvement, likely because most discriminative information has already been captured earlier in the temporal sequence.

Despite these promising results, several limitations should be noted. First, the reliance on SMOTE for balancing may introduce synthetic patterns that do not fully reflect ecological variability. Second, missing satellite observations in certain weeks may bias temporal importance interpretation. Third, while XGBoost provides interpretable feature importance estimates, these should not be interpreted as causal relationships. Finally, Sentinel-2's spatial and spectral resolution limits the ability to detect very early physiological stress signals at the individual tree level.

Overall, the findings demonstrate that combining static and temporal remote sensing data improves both predictive performance and temporal robustness, supporting the value of multi-source data fusion for operational forest monitoring systems.

6. Conclusions

This study shows that combining Sentinel-2 satellite imagery with static environmental features and an XGBoost-classifier results in a model that excels at detecting early bark beetle attacks. The results demonstrate that the model trained on Sentinel-2 satellite imagery and static environmental features consistently outperforms models based on either data source alone, particularly in early detection scenarios when the spectral signals are weak.

Static features provide strong baseline predictive capability by capturing site-specific susceptibility, while Sentinel-2 data become increasingly informative as vegetation stress develops over time. The analysis further indicates that only a subset of the temporal observations carries the majority of the predictive information, highlighting the importance of key phenological and infestation-related periods. Temporal feature importance analysis identified weeks 19, 25, 31, and 33 as particularly informative for prediction. The high importance of week 19, which corresponds to the peak swarming period, suggests that pre-dispositional stress conditions that influence attack susceptibility may already be detectable before visible infestation symptoms appear. Week 25 also contributed substantially to model performance, which is consistent with the expected delay between infestation and observable spectral responses reported in previous studies. In addition, the increased importance of later-season observations reflects the progression of infestation-related canopy stress and discoloration.

Overall, the proposed multi-source framework offers a robust and interpretable approach for monitoring bark beetle infestations. It has potential for operational use in forest management systems, where early detection is critical for reducing ecological and economic damage. Future work should focus on improving robustness to missing temporal observations and exploring more advanced spatio-temporal learning methods for further performance gains.

Author Contributions: Conceptualization, P.-O.O., T.K., M.W. and J.L.; methodology, S.Z., P.-O.O., T.K., M.W. and J.L.; software, S.Z.; validation, S.Z.; formal analysis, S.Z., P.-O.O. and T.K.; investigation, S.Z.; resources, P.-O.O.; data curation, P.-O.O.; writing—original draft preparation, S.Z., P.-O.O. and J.L.; writing—review and editing, T.K. and M.W.; visualization, S.Z.; supervision, T.K., M.W. and J.L.; project administration, J.L.; funding acquisition, P.-O.O. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: SZ and JL have received financial support from the Swedish Strategic Research Area (SRA) MERGE (ModElling the Regional and Global Earth System). The study is also a contribution to the SRA BECC (Biodiversity and Ecosystem Services in a Changing Climate) and to the Lund University profile area Nature-based future solutions. PO was funded by the Swedish National Space Agency, grant Dnr. 2020-00173.

Data Availability Statement: Due to the data use agreement signed with the data provider, the bark beetle attack data cannot be shared publicly. Access to these data is restricted. Sentinel-2 data are freely available; however, without the restricted bark beetle dataset, they cannot be used to fully reproduce the results of this study.

Acknowledgments: Sveaskog, the largest forest owner in Sweden, and the forest owner association Södra are acknowledged for providing harvester data, and the Swedish Forest Agency (Skogsstyrelsen) and the National Forest Inventory (NFI) for providing bark beetle presence data.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results

Appendix A. Correlation Between Features

This section presents the correlation matrices of spectral bands and static features in Figure A1 and Figure A2, respectively. To quantify relationships between spectral bands, temporal observations were first aggregated by computing the mean correlation values over time.

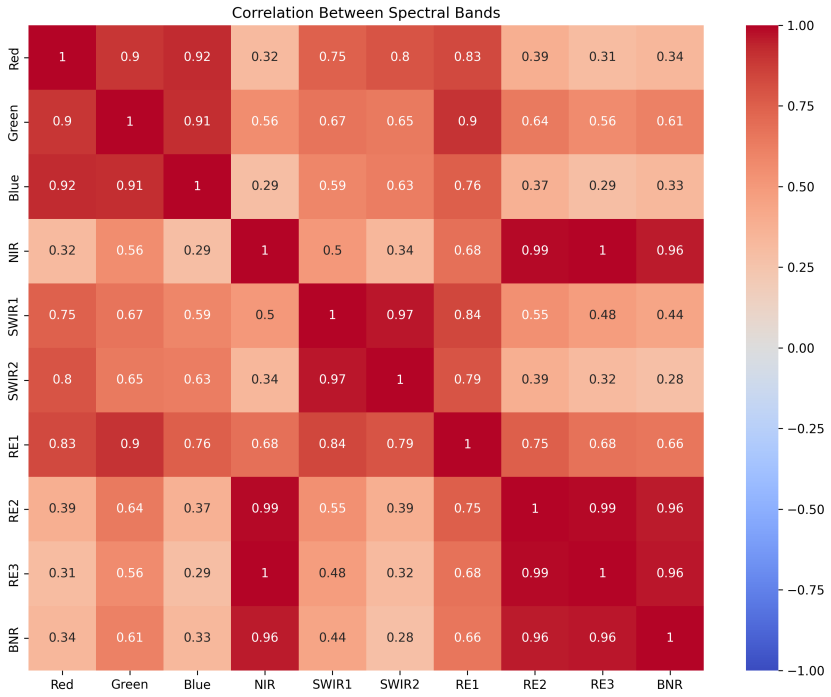


Figure A1. Correlation between bands averaged over the temporal dimension.

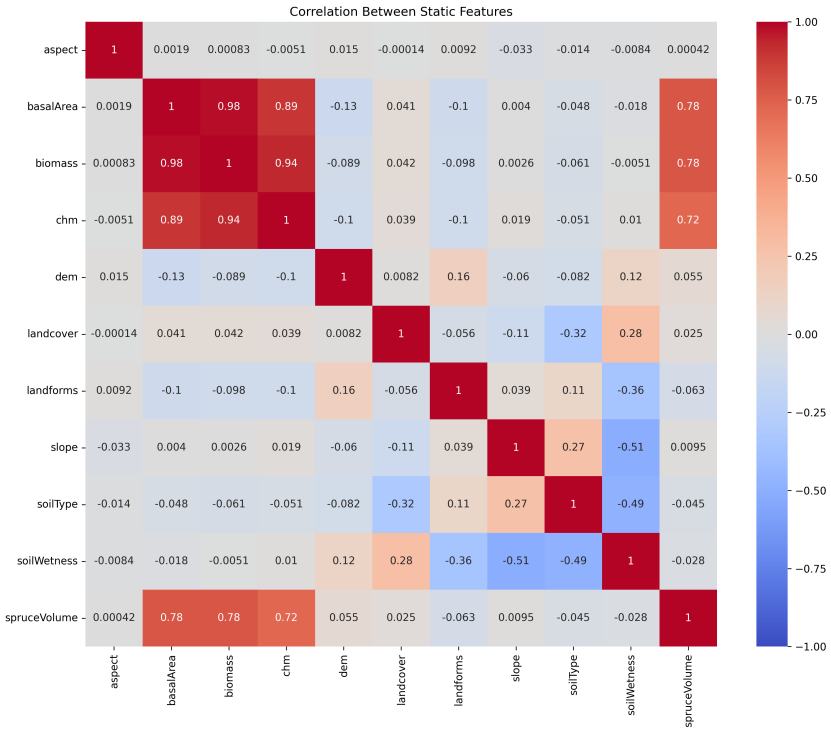


Figure A2. Correlation between static features.

Appendix B. Comparison with Other Classifiers

Appendix B.1. Random Forest

To assess the use of the XGBoost method, the same analysis was performed using Random Forests (RFs); see evaluation metrics in Figure A3. Both methods achieve similar AUC values, suggesting comparable ability to distinguish between pixels with attacked and healthy trees. However, the computational cost of RFs is substantially higher, and the evolution of precision, recall, and F1 score across weeks highlights clear differences in how the methods balance false positives and false negatives.

For both RF and XGBoost, precision increases with more data, indicating improved reliability of positive predictions. However, RF increasingly favors labeling pixels as healthy, reducing false positives at the cost of missing a growing share of attacked trees. Consequently, recall and F1 score decrease or stagnate, further emphasizing the differences between the XGBoost- and RF-based models.

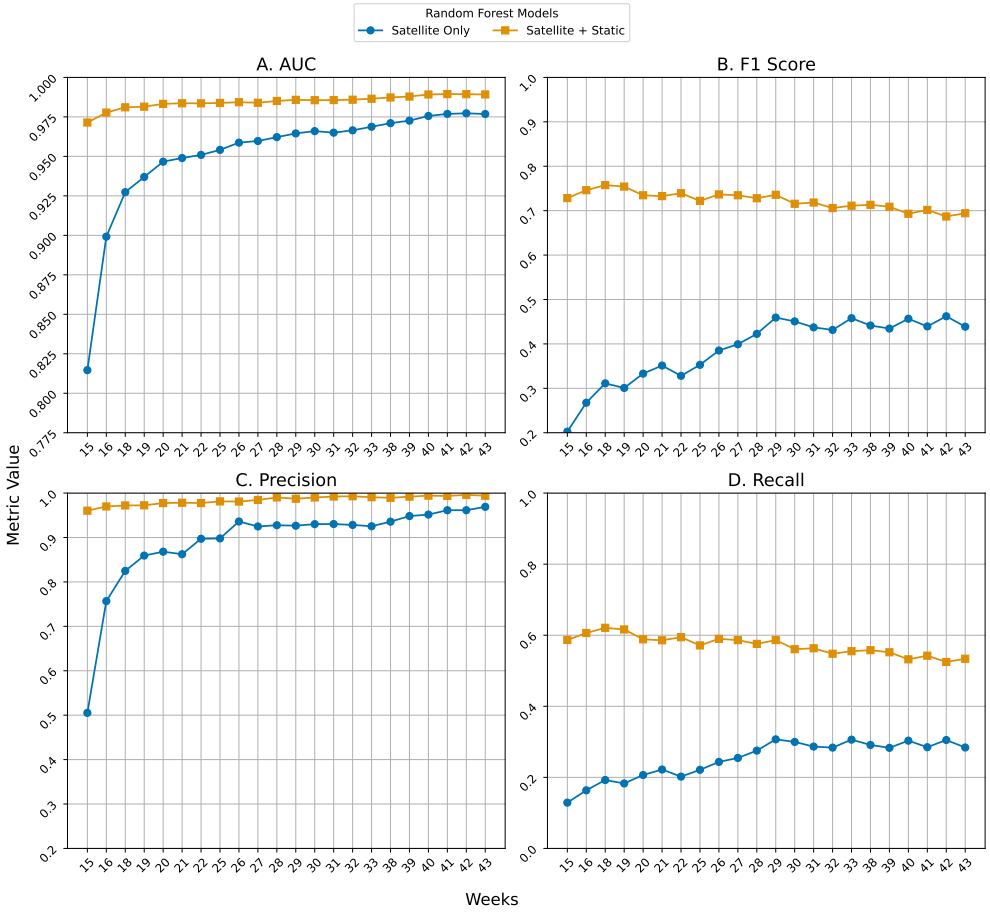


Figure A3. Evaluation metrics of Random Forest models over time using satellite data (blue) and combined satellite and static data (yellow). Overall performance is lower than that of XGBoost (see Figure 3).

Appendix B.2. Deep Learning

The performance of the proposed framework was also compared against a recurrent neural network (RNN)-based [42] baseline. The baseline model consists of a bidirectional Gated Recurrent Unit (GRU) network designed to capture temporal dependencies within the multispectral satellite time series.

The GRU architecture has two stacked recurrent layers with 128 hidden units per direction. By operating bidirectionally, the model processes the temporal sequence in both forward and backward directions, enabling it to learn relationships from both past and

future observations within the input sequence. Dropout regularization is applied between recurrent layers to reduce overfitting.

The output hidden states from all time steps are aggregated using temporal mean pooling. This representation is then passed through a fully connected classification head consisting of two linear layers with ReLU activation and additional dropout regularization. The final layer uses sigmoid activation function to output a single probability.

The RNN-model captured the same temporal patterns as the XGBoost model—see Figure A4—confirming the importance of temporal information for detecting bark beetle-attacked forest stands. However, XGBoost consistently outperformed the RNN across all evaluation metrics.

Several factors may explain these differences. Firstly, the available dataset contains a relatively limited number of attacked samples despite the use of SMOTE. Tree-based ensemble methods, such as XGBoost, are well known for their robustness on limited training samples, whereas deep learning models typically require larger datasets to fully exploit their representational capacity. Secondly, XGBoost can effectively capture non-linear relationships and feature interactions without having to learn any long-term temporal representations. Thirdly, XGBoost is resistant to overfitting through the use of regularization, tree pruning, and ensemble averaging, making it particularly suitable for noisy data with missing-value imputation and class imbalances.

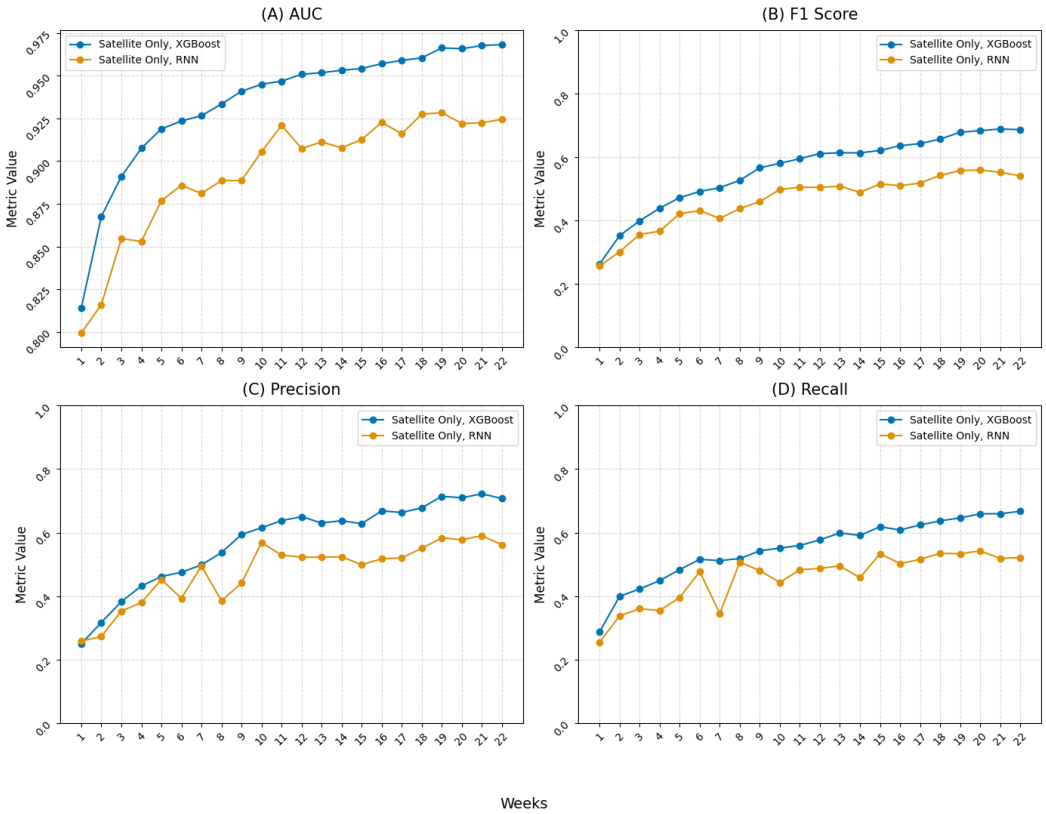


Figure A4. Comparing evaluation metrics of the RNN-deep learning approach to the XGBoost model.

Appendix C. Ablation Study

As discussed earlier, the dataset contains a substantial proportion of missing values. The total number of samples available in each tile is reported in Table 2. When all pixels containing at least one missing observation along the temporal dimension are removed prior to the weekly maximum aggregation step, fewer than ten samples remain in each tile.

This drastic reduction in sample size highlights the necessity of missing-value imputation before any further analysis.

To assess the impact of different strategies for handling missing values and class imbalances, we conducted an ablation study using several different pre-processing approaches; see Figure A5. Specifically, we compared: (1) the original pipeline employing SMOTE for class balancing and temporal forward filling of missing values, (2) spatial, instead of temporal, imputation of the missing-values, and (3) random down-sampling of the majority class (healthy pixels) to match the minority class (attacked pixels). The results indicate that spatial imputation performs poorly due to the sparsity of the available observations, which limits the effectiveness of neighboring pixels for reconstructing missing values. Similarly, down-sampling substantially reduces the amount of training data, degrading model performance. Overall, the original SMOTE-based approach with temporal forward filling provides the most favorable balance between preserving information and addressing class imbalance.

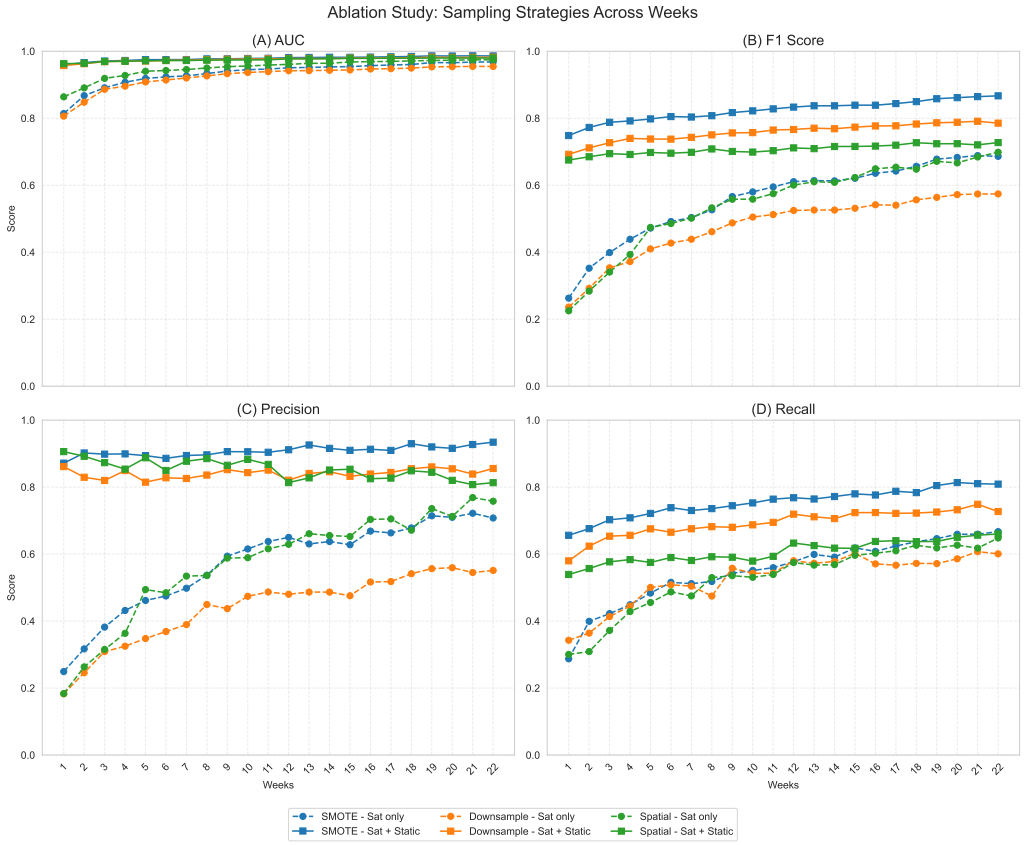


Figure A5. Ablation study evaluating alternative strategies for handling missing values and class imbalance. The baseline approach combines temporal imputation with SMOTE-based oversampling of the minority class. Alternative configurations include spatial imputation and random downsampling of the majority (healthy) class. Spatial imputation yields inferior performance due to the sparsity of observations, while downsampling leads to substantial information loss by discarding a large fraction of healthy samples. The SMOTE-based approach achieves the best overall performance.

Appendix D. Comparison of the Temporal Model

To assess the differences in F1 score, recall, and precision between the satellite-only and combined models, the bootstrap approach described in Section Comparison of the Satellite Only and Combined Models was applied. The mean differences in F1 score, recall, and precision, along with the corresponding 95% confidence intervals, are presented in

Figure A6. This shows that the combined models consistently outperform the satellite-only models across all weeks and metrics.

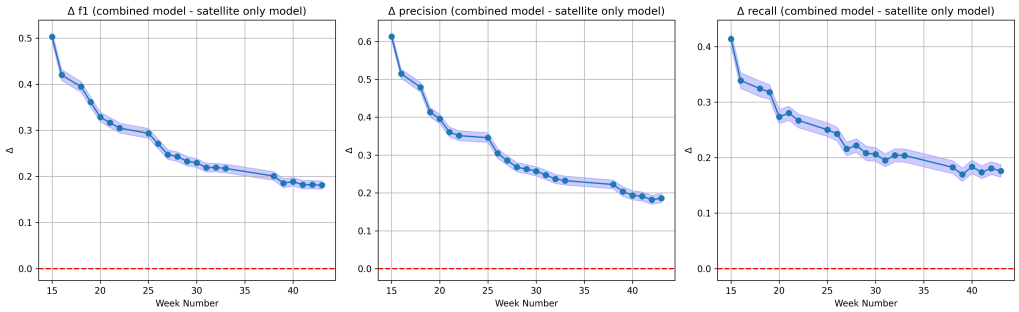


Figure A6. Difference between bootstrapped AUC estimates and paired 95% confidence intervals on the unseen test data. Values above zero (the red dashed line) indicate that the combined model significantly outperforms the satellite-only model.

References

1. Hỳsek, Š.; Löwe, R.; Turčáni, M. What happens to wood after a tree is attacked by a bark beetle? *Forests* **2021**, *12*, 1163. [CrossRef]
2. Robbins, Z.J.; Xu, C.; Jonko, A.; Chitra-Tarak, R.; Fettig, C.J.; Costanza, J.; Mortenson, L.A.; Aukema, B.H.; Kueppers, L.M.; Scheller, R.M. Carbon stored in live ponderosa pines in the Sierra Nevada will not return to pre-drought (2012) levels during the 21st century due to bark beetle outbreaks. *Front. Environ. Sci.* **2023**, *11*, 1112756. [CrossRef]
3. Kurz, W.A.; Dymond, C.C.; Stinson, G.; Rampley, G.J.; Neilson, E.T.; Carroll, A.L.; Ebata, T.; Safranyik, L. Mountain pine beetle and forest carbon feedback to climate change. *Nature* **2008**, *452*, 987–990. [CrossRef] [PubMed]
4. Hlásny, T.; König, L.; Krokene, P.; Lindner, M.; Montagné-Huck, C.; Müller, J.; Qin, H.; Raffa, K.F.; Schelhaas, M.J.; Svoboda, M.; et al. Bark Beetle Outbreaks in Europe: State of Knowledge and Ways Forward for Management. *Curr. For. Rep.* **2021**, *7*, 138–165. [CrossRef]
5. Raffa, K.F.; Grégoire, J.C.; Staffan Lindgren, B. Chapter 1—Natural History and Ecology of Bark Beetles. In *Bark Beetles*; Vega, F.E., Hofstetter, R.W., Eds.; Academic Press: San Diego, CA, USA, 2015; pp. 1–40. [CrossRef]
6. Öhrn, P.; Långström, B.; Lindelöw, Å.; Björklund, N. Seasonal flight patterns of Ips typographus in southern Sweden and thermal sums required for emergence. *Agric. For. Entomol.* **2014**, *16*, 147–157. [CrossRef]
7. Schelhaas, M.J.; Nabuurs, G.J.; Schuck, A. Natural disturbances in the European forests in the 19th and 20th centuries. *Glob. Change Biol.* **2003**, *9*, 1620–1633. [CrossRef]
8. Senf, C.; Pflugmacher, D.; Zhiqiang, Y.; Sebald, J.; Knorn, J.; Neumann, M.; Hostert, P.; Seidl, R. Canopy mortality has doubled in Europe's temperate forests over the last three decades. *Nat. Commun.* **2018**, *9*, 4978. [CrossRef] [PubMed]
9. Fernandez-Carrillo, A.; Patočka, Z.; Dobrovolnỳ, L.; Franco-Nieto, A.; Revilla-Romero, B. Monitoring bark beetle forest damage in Central Europe. A remote sensing approach validated with field data. *Remote Sens.* **2020**, *12*, 3634. [CrossRef]
10. Schroeder, M. *Granbarkborrens förökningsframgång i dödade träd under sommaren 2022 i sydöstra Småland, Värmland och Uppland/Västmanland*; Sveriges lantbruksuniversitet (SLU): Uppsala, Sweden, 2023.
11. Washaya, P.; Modlinger, R.; Tyšer, D.; Hlásny, T. Patterns and impacts of an unprecedented outbreak of bark beetles in Central Europe: A glimpse into the future? *For. Ecosyst.* **2024**, *11*, 100243. [CrossRef]
12. Abdullah, H.; Skidmore, A.K.; Darvishzadeh, R.; Heurich, M. Sentinel-2 accurately maps green-attack stage of European spruce bark beetle (*Ips typographus*, L.) compared with Landsat-8. *Remote Sens. Ecol. Conserv.* **2019**, *5*, 87–106.
13. König, S.; Thonfeld, F.; Förster, M.; Dubovyk, O.; Heurich, M. Assessing combinations of Landsat, Sentinel-2 and Sentinel-1 time series for detecting bark beetle infestations. *Gisci. Remote Sens.* **2023**, *60*, 2226515. [CrossRef]
14. Huo, L.; Persson, H.J.; Lindberg, E. Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). *Remote Sens. Environ.* **2021**, *255*, 112240. [CrossRef]
15. Bárta, V.; Lukeš, P.; Homolová, L. Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *100*, 102335. [CrossRef]
16. Fassnacht, F.E.; Latifi, H.; Ghosh, A.; Joshi, P.K.; Koch, B. Assessing the potential of hyperspectral imagery to map bark beetle-induced tree mortality. *Remote Sens. Environ.* **2014**, *140*, 533–548. [CrossRef]
17. Huo, L.; Koivumäki, N.; Oliveira, R.A.; Hakala, T.; Markelin, L.; Näsi, R.; Suomalainen, J.; Polvivaara, A.; Junttila, S.; Honkavaara, E. Bark beetle pre-emergence detection using multi-temporal hyperspectral drone images: Green shoulder indices can indicate subtle tree vitality decline. *ISPRS J. Photogramm. Remote Sens.* **2024**, *216*, 200–216. [CrossRef]

18. Huo, L.; Lindberg, E.; Bohlin, J.; Persson, H.J. Assessing the detectability of European spruce bark beetle green attack in multispectral drone images with high spatial-and temporal resolutions. *Remote Sens. Environ.* **2023**, *287*, 113484. [CrossRef]
19. Dalponte, M.; Solano-Correa, Y.T.; Frizzera, L.; Gianelle, D. Mapping a European Spruce Bark Beetle Outbreak Using Sentinel-2 Remote Sensing Data. *Remote Sens.* **2022**, *14*, 3135. [CrossRef]
20. Candotti, A.; De Giglio, M.; Dubbini, M.; Tomelleri, E. A Sentinel-2 Based Multi-Temporal Monitoring Framework for Wind and Bark Beetle Detection and Damage Mapping. *Remote Sens.* **2022**, *14*, 6105. [CrossRef]
21. Jamali, S.; Olsson, P.O.; Müller, M.; Ghorbanian, A. Early Detection of Forest Bark Beetle Attack Using Time Series Spatial Variability of Spectral Indexes from Sentinel-2. In Proceedings of the IGARSS 2024—2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece, 7–12 July 2024; pp. 10157–10160. [CrossRef]
22. Müller, M.; Olsson, P.O.; Eklundh, L.; Jamali, S.; Ardö, J. Features predisposing forest to bark beetle outbreaks and their dynamics during drought. *For. Ecol. Manag.* **2022**, *523*, 120480. [CrossRef]
23. Zhao, P.; Olsson, P.O.; Øhrman Wellendorf, A.; Müller, M.; Mansourian, A. Exploring the spatial heterogeneity of bark beetle infestation risk factors with geographically weighted regression and random forest. *Scand. J. For. Res.* **2025**, *40*, 188–204. [CrossRef]
24. Olsson, P.O.; Zhao, P.; Müller, M.; Mansourian, A.; Ardö, J. Combining Sentinel-2 Data and Risk Maps to Detect Trees Predisposed to and Attacked by European Spruce Bark Beetle. *Remote Sens.* **2024**, *16*, 4166. [CrossRef]
25. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: New York, NY, USA, 2016; KDD' 16; pp. 785–794. [CrossRef]
26. Swedish Meteorological and Hydrological Institute. *Climate Normals*; Swedish Meteorological and Hydrological Institute: Norrköping, Sweden, 2025.
27. Nilsson, P.; Roberge, C.; Fridman, J. *Skogsdata 2021: Aktuella uppgifter om de svenska skogarna från SLU Riksskogstaxeringen*; Sveriges lantbruksuniversitet (SLU): Uppsala, Sweden, 2021.
28. European Space Agency. *Copernicus Open Access Hub*; European Space Agency: Paris, France, 2015.
29. Frantz, D. FORCE—Landsat + Sentinel-2 Analysis Ready Data and Beyond. *Remote Sens.* **2019**, *11*, 1124. [CrossRef]
30. Rufin, P.; Frantz, D.; Yan, L.; Hostert, P. Operational Coregistration of the Sentinel-2A/B Image Archive Using Multitemporal Landsat Spectral Averages. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 712–716. [CrossRef]
31. Naturvårdsverket. *National Land Cover Database (NMD)*; Naturvårdsverket: Stockholm, Sweden, 2021.
32. Lantmäteriet. *GSD-Terrängkartan, Vektor*, version 5.10; Lantmäteriet: Gävle, Sweden, 2021.
33. Skogsstyrelsen. *Svärmningsövervakning*. Available online: https://www.skogsstyrelsen.se/svarmningsovervakning (accessed on 23 June 2026).
34. Yang, Y.; Xu, J.; Hong, Y.; Lv, G. The dynamic of vegetation coverage and its response to climate factors in Inner Mongolia, China. *Stoch. Environ. Res. Risk Assess.* **2012**, *26*, 357–373. [CrossRef]
35. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]
36. Nitesh, V.C. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321. [CrossRef]
37. Cho, H.; XGBoost Contributors. XGBoost: Extreme Gradient Boosting (ver. 3.0.2). 2025. Available online: https://pypi.org/project/xgboost/3.0.2/ (accessed on 23 June 2025).
38. Niculescu-Mizil, A.; Caruana, R. Predicting good probabilities with supervised learning. In Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany, 7–11 August 2005; pp. 625–632.
39. Glenn, W.B. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **1950**, *78*, 1–3. [CrossRef]
40. Kautz, M.; Feurer, J.; Adler, P. Early detection of bark beetle (Ips typographus) infestations by remote sensing—A critical review of recent research. *For. Ecol. Manag.* **2024**, *556*, 121595. [CrossRef]
41. Huo, L.; Lindberg, E.; Fransson, J.E.; Persson, H.J. Comparing spectral differences between healthy and early infested spruce forests caused by bark beetle attacks using satellite images. In *Proceedings of the IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*; IEEE: New York, NY, USA, 2022; pp. 7709–7712.
42. Elman, J.L. Finding structure in time. *Cogn. Sci.* **1990**, *14*, 179–211. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.