

Article

Windthrow Mapping with Sentinel-2 and PlanetScope in Triglav National Park: A Regional Case Study

Matej Zupan, Krištof Oštir  and Ana Potočnik Buhvald * 

Faculty of Civil and Geodetic Engineering, University of Ljubljana, Jamova Cesta 2, 1000 Ljubljana, Slovenia; kristof.ostir@fgg.uni-lj.si (K.O.)

* Correspondence: ana.potocnik-buhvald@fgg.uni-lj.si

Highlights

This regional case study uses multi-temporal Sentinel-2 and PlanetScope imagery to map windthrow and estimate parcel-level timber damage in alpine forests. The findings reflect within-sample performance for a single event in Triglav National Park. The comparison with official sanitary-logging records ($n = 8$, non-probability) is preliminary and not generalisable.

What are the most important results?

- Within-sample overall accuracy for PlanetScope 72.9% (95% CI: 71.2–74.6%) and Sentinel-2 69.2% (95% CI: 67.4–71.2%) in this alpine regional case study.
- Detection was size-dependent: gaps larger than 0.5 ha were consistently detected, while gaps smaller than 0.1 ha were frequently omitted. Omissions were higher for Sentinel-2 and lower for PlanetScope, indicating a modest advantage for smaller fragmented patches in this sample.
- Linking satellite-derived change maps with available forest stand data enabled parcel-level estimates of damaged timber volume. Across $n = 8$ non-probability parcels, compared with official sanitary-logging records, mean absolute deviations were 5–7%; these figures are preliminary and not generalisable.

What are the implication of the main finding?

- The study documents within-sample performance from a regional case study in alpine terrain. Any broader generalisation will require larger, probability-based validation across additional events and forest types, as well as broader access to parcel-level official records.
- In our sample, PlanetScope omitted fewer smaller fragmented gaps than Sentinel-2, while gaps smaller than 0.1 ha often required field verification or VHR/UAV follow-up. The reported bootstrap confidence intervals express within-sample uncertainty and do not constitute operational performance guarantees.

Abstract

Extreme weather increasingly damages forest ecosystems, and affected areas are often remote or inaccessible, limiting field surveys. In such contexts, remote sensing can complement damage assessment. This study presents a regional case study evaluating established multi-temporal optical change detection for windthrow mapping in Triglav National Park (Slovenia) using Sentinel-2 and PlanetScope imagery. Bitemporal index differencing and fixed thresholds were applied, with accuracy quantified via a block bootstrap to account for spatial autocorrelation. Within-sample overall accuracy was 69.2% (95% CI: 67.4–71.2%) for Sentinel-2 and 72.9% (95% CI: 71.2–74.6%) for PlanetScope. Detection was strongly



Academic Editors: Hubert Hasenauer, Nikolay S. Strigul and Antoine Masse

Received: 9 September 2025

Revised: 20 October 2025

Accepted: 26 October 2025

Published: 28 October 2025

Citation: Zupan, M.; Oštir, K.; Potočnik Buhvald, A. Windthrow Mapping with Sentinel-2 and PlanetScope in Triglav National Park: A Regional Case Study. *Remote Sens.* **2025**, *17*, 3568. <https://doi.org/10.3390/rs17213568>

Copyright: © 2025 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 (<https://creativecommons.org/licenses/by/4.0/>).

size-dependent: gaps greater than 0.5 ha were consistently detected, whereas gaps smaller than 0.1 ha were frequently omitted, particularly with Sentinel-2. Linking satellite-derived change maps with forest stand data enabled parcel-level estimates of damaged timber volume; this linkage was examined on a small, non-probability set of parcels and is therefore preliminary. We position the work strictly as a case study documenting within-sample performance in alpine terrain. Broader generalisation will require probability-based validation across additional events and forest types, and wider access to parcel-level official records.

Keywords: multi-temporal; high-resolution satellite imagery; windthrow (windstorm) damage; alpine forests; timber volume estimation; change detection

1. Introduction

Forest biomass is one of nature's most valuable resources, serving both as a renewable source of energy and as a raw material for a wide range of products. However, the resilience of forests is increasingly threatened by climate change and associated extreme weather events, which can cause widespread and severe damage. In Europe, natural disturbances such as windstorms, bark beetle outbreaks, and forest fires have become more frequent and intense over the past decades [1–3]. These disturbances have significant impacts on forest carbon storage, biodiversity, and long-term ecosystem stability. Because timely and accurate assessment of forest disturbances is crucial, remote sensing has become an indispensable tool, especially for capturing storm-induced damage such as windthrow, which is the focus of this study.

Forest damage caused by storms can be monitored using a range of remote sensing platforms, including satellites, aircraft, and unmanned aerial vehicles (UAVs). These platforms provide high-resolution spatial data over large areas, enabling efficient and cost-effective mapping of disturbances. Change detection (CD) techniques are commonly used to monitor damage, with different approaches having different strengths and limitations. The choice of a suitable method depends on the event characteristics, forest structure, and data availability.

Remote sensing based windthrow detection is already a well-established and widely applied topic in scientific research. Numerous studies [4–8] have used different datasets and methods, primarily based on optical or radar satellite imagery, with some also incorporating LiDAR data.

Optical satellite imagery has been used extensively for windthrow detection, utilising datasets such as Sentinel-2, Landsat 5 TM, Landsat 7 ETM+, PlanetScope, QuickBird, and GeoEye. Due to their spatial resolution, these data enable the detection of both large-scale forest disturbances and, in some cases, individual damaged trees [9]. Numerous studies have combined bi- or multi-temporal optical imagery with different CD techniques and report an overall accuracy ranging from 51% to 98% [4–6,8,10,11]. Frequently used CD approaches include principal component analysis (PCA) and change vector analysis (CVA).

Radar sensors, in particular Synthetic Aperture Radar (SAR) systems such as Sentinel-1 and COSMO-SkyMed, also play an important role in forest damage assessment. In contrast to optical sensors, SAR is unaffected by cloud cover or illumination, which is an advantage for time-critical analysis. Schwarz et al. [11] compared optical and radar data and found a comparable accuracy of interpretation, although the choice of sensor depends largely on methodological and environmental factors. Other studies, such as Dalponte et al. [7] and Rüetschi et al. [12], have demonstrated that SAR data can achieve similar results to

optical imagery. However, the accuracy of radar-based assessments can vary considerably depending on seasonal and phenological conditions [7].

LiDAR, particularly data acquired through airborne laser scanning (ALS) via UAVs, offers highly accurate assessments of windthrow damage at finer spatial scales [13–15]. LiDAR enables the calculation of detailed forest structure metrics such as the canopy height model (CHM) and aboveground biomass (AGB). While this method can provide highly accurate estimates of damaged timber volume, it is generally more complex and costly, which limits its wide application.

This study is primarily based on the methods presented by Dalponte et al. [6] and Jonikavičius & Mozgeris [5], which demonstrated the effectiveness of optical satellite data for assessing storm-induced forest damage. Dalponte et al. [6] used Sentinel-2 and PlanetScope imagery in conjunction with CVA to map windthrow in the Trentino–South Tyrol region of Italy following the Vaia storm in October 2018. Their approach yielded an overall accuracy of over 90% and a producer accuracy of over 80%. The use of very high-resolution PlanetScope imagery enabled a better delineation of windthrow boundaries and the detection of smaller damaged areas compared to Sentinel-2.

Jonikavičius & Mozgeris [5], in contrast, used Landsat 5 TM imagery to assess storm damage in the forest in Lithuania in August 2010. While the CD methods and accuracies were comparable to other studies, the main contribution was the estimation of timber volume loss. This was achieved by integrating windthrow mapping with forest inventory data from the Lithuanian State Forest Cadastre. Their results showed an overall accuracy of more than 95%, with the timber volume estimate differing by less than 5% from the official records.

This study assesses the applicability of multitemporal, high-resolution optical satellite imagery for the rapid assessment of windthrow damage in inaccessible mountain forests. We focus on Sentinel-2 and PlanetScope data to test their potential for detecting and delineating storm-induced forest disturbances in Triglav National Park (Slovenia), which was severely affected by local wind vortices on 19 July 2023. Using a change detection approach, we investigate both the spatial extent of the windthrow and the estimation of the damaged timber volume. By comparing the performance of Sentinel-2 and PlanetScope imagery, this study documents the within-sample behaviour of established change detection methods under alpine conditions, highlighting trade-offs between medium- and high-resolution optical data.

2. Materials and Methods

2.1. Study Area

The study area is in the Julian Alps in Slovenia and encompasses the forested Alpine valleys Krma, Kot, Radovna and their surroundings (Figure 1). It covers approximately 5200 ha, with homogeneous forest stands dominated by Norway spruce (*Picea abies*) and European beech (*Fagus sylvatica*). The terrain is very rugged and largely inaccessible; therefore, it presents an excellent opportunity for applying the use of remote sensing data.

In July 2023, the area was struck by several strong local eddies, an uncharacteristic weather phenomenon for this region. The disturbance primarily affected spruce stands, whose relatively shallow and weak root systems make them particularly vulnerable to windthrow, especially on poorly drained or homogenous forest sites [16–18]. Similar patterns were observed by Ogris et al. [16], who analysed windthrow factors on the nearby Pokljuka plateau. Given the geographical proximity and comparable forest structure, their findings provide an important reference for understanding the wind disturbance event examined in our study.

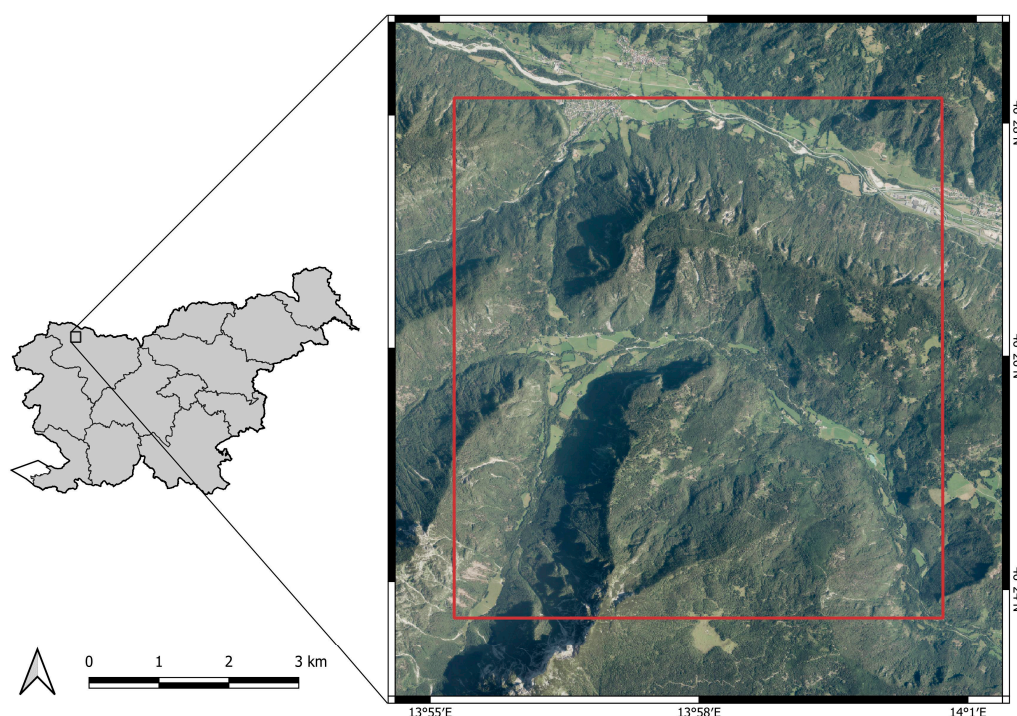


Figure 1. Digital orthophoto (DOF with 0.5 m spatial resolution) of the study area with the boundary (red line) of interest. (DOF was taken in 2020—before the windstorm). (Source: Surveying and Mapping Authority of the Republic of Slovenia).

According to the preliminary estimates of the Slovenia Forest Service [19], more than 60,000 m³ of timber was damaged during the storm (Figure 2). This estimate was based on photointerpretation of DOF 2023 imagery acquired shortly after the event. At the time of our analysis, salvage logging was still ongoing, so the exact amount of damaged timber remained uncertain [20].

2.2. Satellite Images Dataset

To obtain optimal results in the alpine study area and to reduce the frequent effects of cloud cover, we used images from Sentinel-2 (S2) and PlanetScope (PS). These two constellations differ in spatial, spectral, and temporal resolution, as well as in the radiometric quality and accessibility of the data (PS is commercial but free for research purposes, while the S2 data are freely available). Given these differences, we expected the results to vary between the two datasets.

For this study, the red, green, blue, red-edge, near-infrared and swir spectral bands were used. First, true-colour composites were generated to enable visual interpretation and facilitate the presentation of results (see Figure 3). Second, the Normalised Difference Vegetation Index (NDVI), Normalized Difference Red Edge (NDRE) and Normalized Difference Moisture Index (NDMI) were calculated and applied as the primary basis for change detection. Due to the persistent cloud cover in the alpine area of the interest, images from different dates were merged into a mosaic during pre-processing.



Figure 2. Field image of one of the areas in Radovna valley affected by the windstorm. (Source: Matej Zupan).

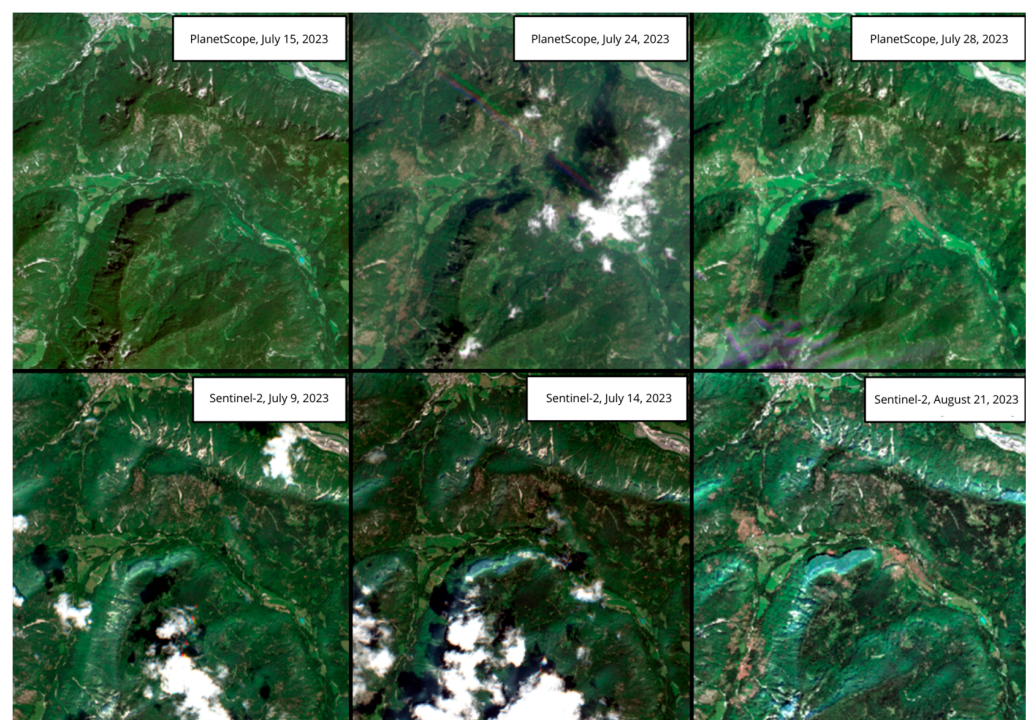


Figure 3. True-colour composites of the acquired PlanetScope and Sentinel-2 datasets [21,22] illustrate the challenges of alpine environments: frequent cloud cover, cast shadows, and terrain-driven illumination differences.

2.2.1. Sentinel-2 Data

The S2 mission, developed by the European Space Agency (ESA) as part of the Copernicus programme, currently consists of three operational satellites—Sentinel-2A, 2B, and

the recently launched 2C. The satellites provide high-resolution multispectral images with a global revisit time of 5 days. The temporal frequency and free access to the data via the Copernicus Browser make S2 a widely used source for land surface monitoring [23].

The images acquired by S2 include 13 spectral bands with three spatial resolutions (10 m, 20 m, and 60 m), as shown in Table 1. For this study, three S2 images were selected that covered the area of interest at different times. Two images were taken shortly before the event, while one image was taken after the windstorm. Due to the alpine mountainous terrain and the resulting frequent cloud cover, the images from 9 July, 14 July, and 21 August 2023, were the most suitable for processing. We were, unfortunately, unable to obtain usable images closer in time to the event (19 July 2023).

Table 1. Spectral band and spatial resolution for Sentinel-2 [23].

Band	Wavelength (nm)	Spectral Band	Spatial Resolution (m)
1	443	Atmospheric Corrections (Aerosols)	60
2	490	Blue	10
3	560	Green	10
4	665	Red	10
5	705	Red Edge 1 (Vegetation)	20
6	740	Red Edge 2	20
7	783	Red Edge 3	20
8	842	Near-Infrared (NIR)	10
8A	865	Red Edge 4 (Narrow NIR)	20
9	940	Atmospheric Corrections (Water Vapour)	60
10	1375	Atmospheric Corrections (Cirrus)	60
11	1610	Short-Wave Infrared 1 (SWIR 1)	20
12	2190	Short-Wave Infrared 2 (SWIR 2)	20

2.2.2. PlanetScope Data

PS is a satellite constellation operated by Planet Labs and currently consists of more than 120 small Dove satellites designed for daily Earth [22]. The constellation provides multispectral images at a spatial resolution of 3 m over eight spectral bands (Table 2). Although PS is primarily used for commercial purposes, limited research access is available free of charge [24].

PS images acquired on 15 July, 24 July and 28 July 2023 were used for this study to complement the S2 dataset and ensure sufficient temporal coverage around the windthrow event in the alpine area.

Table 2. Spectral band and spatial resolution for PlanetScope images [22].

Band	Wavelength (nm)	Spectral Band	Spatial Resolution (m)
1	443	Coastal Blue	3
2	490	Blue	3
3	531	Green I	3
4	565	Green	3
5	610	Yellow	3
6	665	Red	3
7	705	Red Edge	3
8	865	Near Infrared (NIR)	3

2.3. Ancillary Data

In addition to satellite datasets, four ancillary datasets were used for the analysis and validation.

1. A forest stand map of Slovenia, provided by the Slovenian Forestry Institute (SFI) in a vector format, was used to estimate the volume of damaged timber. The database is updated on a 10-year cycle, with approximately 10% of the national forest area revised each year. For the study area analysed in this research, the most recent update corresponds to 2021. As no more recent data were available, this may introduce a potential bias in cases where forest stand conditions have changed since the last update.
2. Digital Orthophoto (DOF) of Slovenia—Acquired in August 2023 by the Surveying and Mapping Authority of the Republic of Slovenia (GURS) [25]. DOF was used as a reference dataset for validating the change detection results. Manually digitised polygons derived from DOF served as reference data.
3. Administrative and cadastral layers—Municipality, cadastral municipality, and land parcel boundaries, provided by GURS, were used to assess storm damage at smaller administrative units and to compare with in situ data.
4. In situ timber data—Field data on damaged timber volume per land parcel, provided by the Slovenia Forest Service (SFS), were used for validation of volume estimates [26].

2.4. Methods

Figure 4 shows the workflow used to map forest windthrow using a change detection approach and to estimate the volume of damaged timber based on ancillary data.

2.4.1. Data Pre-Processing

Both S2 and PS datasets were obtained from their official repositories, with prior geometric and orthorectification corrections already applied. For atmospheric correction, the PS images were provided at Level-3B (surface reflectance), while S2 images were accessible at Level-2A (atmospherically corrected surface reflectance).

For S2, additional pre-processing steps included clipping the images to the study area, masking clouds and cloud shadows using the official Scene Classification Layer (SCL) and retaining only pixels classified as vegetation (SCL value = 4). As no completely cloud-free image was available for the pre-disturbance period, two S2 scenes acquired on 9 July and 14 July were mosaicked into a single composite to ensure full coverage of the study area.

For PS, additional pre-processing steps included clipping the images to the study area and applying the UDM2 (Unusable Data Mask) provided with each scene. Only

pixels flagged as clear in the UDM2 mask were retained, while all others (e.g., clouds, cloud shadows, haze) were set to missing values. The masking was applied across all eight spectral bands, and the resulting clear-sky composites were subsequently used for vegetation index calculations. Because no completely cloud-free image was available for the post-disturbance period, two PS scenes acquired on 24 July and 28 July were mosaicked into a single composite to ensure full coverage of the study area.

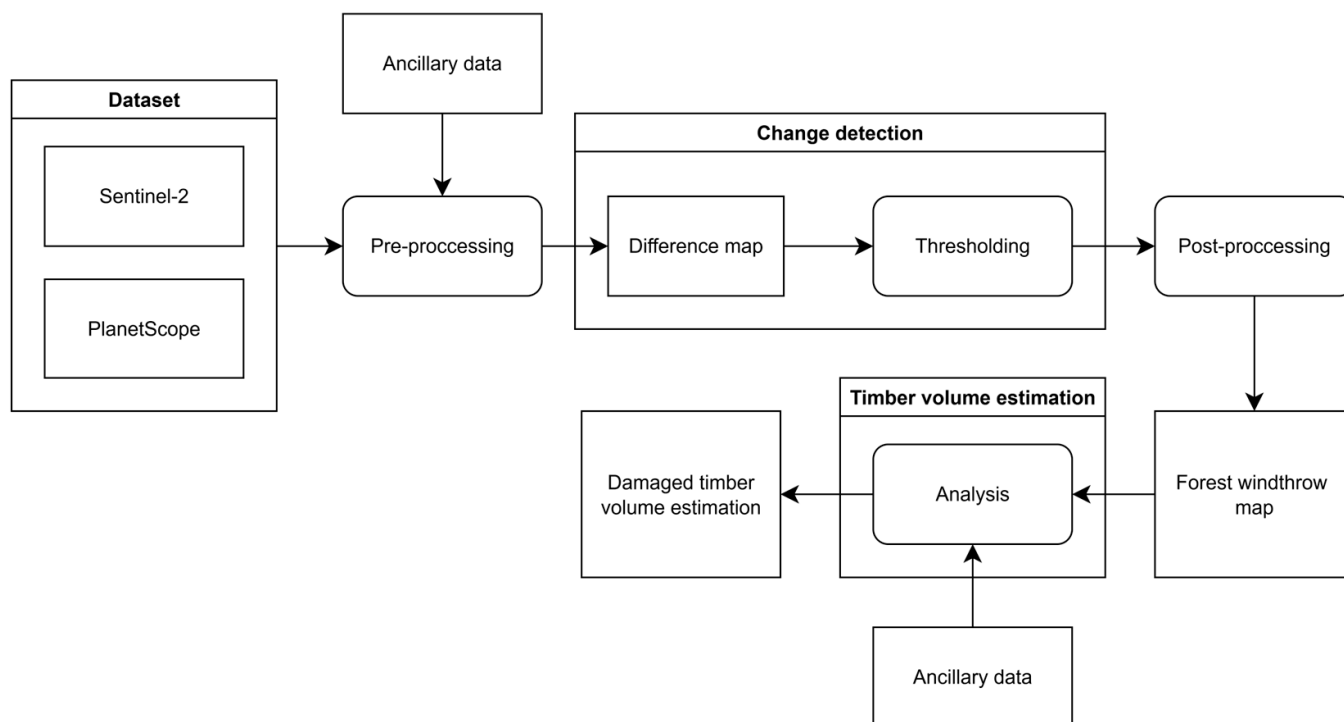


Figure 4. Block diagram of the workflow for windthrow mapping and timber volume estimation.

2.4.2. Windthrow Detection

To assess vegetation changes, we applied a traditional pixel-based change detection approach using different spectral indices, NDVI, NDRE, and NDMI. These indices capture different biophysical properties relevant for windthrow detection. NDVI calculated by Equation (1) is a standard indicator of green biomass and canopy vitality, and a sudden drop is commonly associated with canopy loss. NDRE calculated by Equation (2) is more sensitive to subtle variations in chlorophyll content and canopy structure, which are often affected in partially damaged stands [27]. NDMI calculated by Equation (3) is sensitive to canopy water content and has been shown to respond strongly to structural disturbance and canopy opening following windthrow [28]. Together, these indices provide complementary information on vegetation vigour, canopy structure, and moisture, making them particularly suitable for detecting and quantifying windthrow disturbances.

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \quad (1)$$

$$\text{NDRE} = \frac{\text{NIR} - \text{RedEdge}}{\text{NIR} + \text{RedEdge}} \quad (2)$$

$$\text{NDMI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}} \quad (3)$$

The change detection method was based on pixel-wise differences in spectral index values, producing raster layers that highlight decreases or increases in canopy greenness

and moisture. To illustrate these dynamics, Figure A1 shows the time series of NDVI, NDRE, and NDMI over the windthrow area in the alpine valleys. All three indices display relatively stable values before the disturbance, followed by a sharp decline immediately after the windthrow event in mid-July 2023. This abrupt drop reflects the sudden loss of canopy vitality, chlorophyll content, and moisture associated with windthrow damage.

To delineate areas of significant vegetation loss, we computed histograms of index differences (dNDVI, dNDRE, dNDMI) for both S2 and PS (Figure 5). All histograms show a consistent negative shift, indicating changes associated with windthrow. For each index, we calculated descriptive statistics, including the mean, standard deviation, and 5th percentile (q05) of the distribution. Thresholds were set at q05, corresponding to the lower tail of the distribution where disturbance signals are most strongly expressed. This approach was chosen because windthrow causes an abrupt and substantial reduction in canopy greenness and moisture, resulting in a distinct shift in pixel values into the negative tail of the distribution.

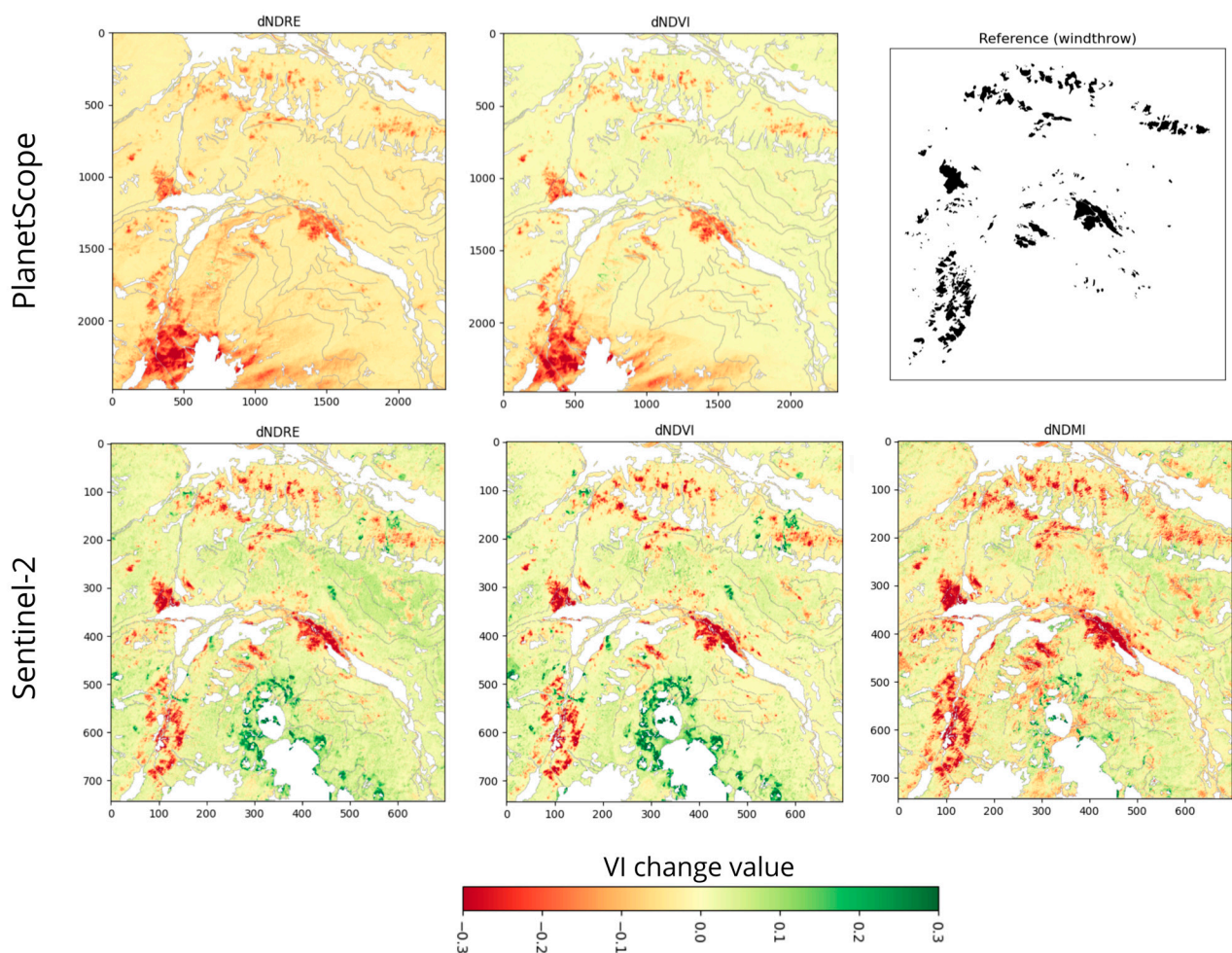


Figure 5. Change detection in vegetation indices (dNDVI, dNDRE, dNDMI) from PlanetScope and Sentinel-2, alongside the mapped windthrow reference.

To further evaluate the robustness of the chosen thresholds, we tested their performance by varying the cutoff values by ± 0.05 . For each adjusted threshold, standard classification metrics (precision, recall, specificity, F1-score, accuracy, and intersection-over-union, IoU) were computed by comparing the detected damage against the reference polygons detected from high-resolution orthophoto (DOF). This allowed us to assess the

sensitivity of the results to small changes in threshold selection and to quantify the stability of the detection approach.

In addition to fixed-threshold analysis, we performed a sensitivity analysis by systematically varying the thresholds of dNDVI, dNDRE, and dNDMI. For each threshold value, the deviation of the estimated damaged area from the reference polygons was calculated and expressed as a percentage. This analysis enabled us to quantify how small changes in threshold selection affect the mapped area and to evaluate the robustness of the final thresholds. In addition, all this analysis helps us to select the best windthrow mask from satellite imagery which was finally used for windthrow estimation in spatial and volume dimension.

Accuracy was assessed on a per-pixel basis by comparing the best windthrow mask outputs with reference data. As per-pixel estimates assume all pixels are independent, they may underestimate uncertainty due to spatial autocorrelation among neighbouring pixels [29]. To address this, we applied a block-bootstrap procedure [30]. The study area was divided into spatially contiguous 1-ha blocks, and 500 bootstrap iterations were performed. In each iteration, blocks were randomly sampled with replacement and accuracy was recalculated as the proportion of correctly classified pixels. The mean of all bootstrap iterations was taken as the corrected accuracy estimate, while the 2.5th and 97.5th percentiles of the bootstrap distribution defined the 95% confidence interval. This approach provides more conservative and realistic measures of accuracy than naïve per-pixel calculations [31].

In addition to per-pixel accuracy, we quantified omission errors to evaluate how frequently reference canopy gaps were missed by S2 and PS. Omission errors were calculated by spatially overlaying reference polygons with satellite-derived gap maps: each reference gap was classified as detected if it intersected a corresponding satellite-derived polygon, and as omitted otherwise [32]. To examine the influence of gap size on omission rates, reference polygons were stratified into three classes (0–0.1 ha, 0.1–0.5 ha, >0.5 ha), following conventions in disturbance mapping [33]. For each size class, omission rate was calculated as the proportion of reference gaps not detected by the satellite-derived maps. In addition, logistic regression was applied to model the relationship between gap size and the probability of detection. This analysis complements per-pixel accuracy by providing scale-dependent insights into detection performance.

This method is consistent with the framework described by [27], who emphasised the effectiveness of NDVI-based image differencing for detecting forest disturbances. It also aligns with refinements proposed by Wang and Xu [3], who highlighted the flexibility and effectiveness of threshold-based techniques for vegetation change analysis, particularly when reference data are available.

2.4.3. Post-Processing for Damaged Timber Volume Estimation

In the post-processing stage, the raster-based change detection outputs were converted into vector format. This conversion enabled seamless integration with supplementary datasets, including cadastral parcels, forest stand boundaries, and official sanitary-logging records. Working in vector format also facilitated the calculation of damage statistics at both the stand- and parcel-level, the application of spatial filtering to reduce noise, and the overlay with official forest management layers. These steps were used to derive estimating the volume of damaged timber and comparing a subset of results against sanitary-logging reports.

The estimation of damaged timber volume was carried out to quantify the impact of the windthrow event at both the stand- and parcel-level (Figure 6). The workflow combined change detection results with forest stand data, which included information on standing volume (m^3/ha). First, the change detection layer was intersected with forest

stand polygons to calculate the volume of damaged timber per stand. These results were then aggregated to the parcel-level ($n = 8$ non-probability parcels) for comparison with official records where post-storm sanitary logging had already been completed.

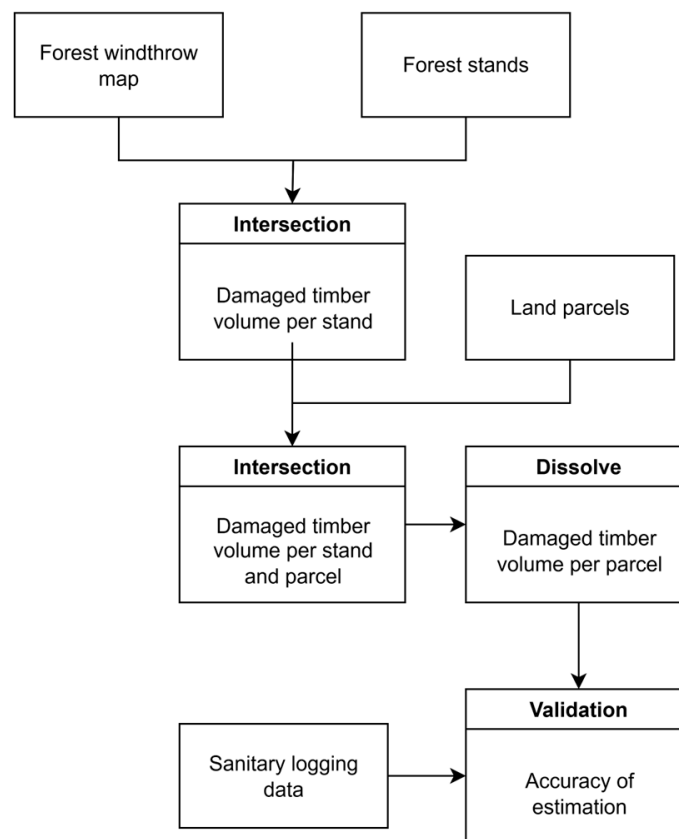


Figure 6. Workflow for the calculation and validation of damaged timber volume.

A partial within-sample comparison indicated the plausibility of the remote-sensing-based estimates while highlighting key limitations: the forest stand dataset used for volume estimation originated from 2021 and therefore did not reflect subsequent sanitary or salvage logging activities [34,35]. To mitigate this bias, we incorporated sanitary-logging reports from 2021–2023, representing the period immediately preceding the storm event. This adjustment was intended to reduce the influence of outdated stand data by accounting for extensive spruce (*Picea abies*) logging associated with bark beetle (*Ips typographus*) outbreaks [28], and it likely reduced bias in the resulting parcel-level estimates.

3. Results

3.1. Threshold Selection

In this regional case study of the Triglav National Park windthrow, thresholds were derived from the lower tails of the spectral-index-difference distributions, where the disturbance signal peaks. Complete histograms and intermediate candidates are provided in Appendix B.1 Figure A2. Candidate cut-offs were selected by visual inspection of the histograms; the final thresholds applied in the analysis were -0.11 (dNDVI, S2), -0.12 (dNDVI, PS), -0.12 (dNDRE, S2), -0.13 (dNDRE, PS), and -0.19 (dNDMI, S2). Because these thresholds were tailored to this specific event and imagery pairing, they should not be generalised without local calibration.

3.2. Sensitivity Analysis of Thresholds

Sensitivity analysis demonstrated that the delineated windthrow area is highly dependent on threshold selection. Figure A3a (S2) and Figure A3b (PS) in Appendix B show F1-score as a function of threshold. For both sensors, NDVI consistently outperformed NDRE and NDMI in terms of F1-score, intersection-over-union (IoU), and Cohen's Kappa, indicating that NDVI is the most robust index for canopy gap detection. The highest F1-scores were obtained with thresholds of -0.10 (S2) and -0.09 (PS).

A detailed summary of classification performance across all tested thresholds is provided in Table A1 in Appendix B. For Sentinel-2, the optimal threshold of -0.10 for dNDVI achieved an F1-score of 0.70, an IoU of 0.54, and a Cohen's Kappa of 0.69, offering the best balance between precision (0.74) and recall (0.67). For PS, the best-performing threshold was -0.09 for dNDVI, resulting in an F1-score of 0.531 and a Kappa of 0.51. Across both sensors, NDVI clearly outperformed NDRE and NDMI, which consistently showed lower F1-scores, IoU, and Kappa values, confirming that NDVI is the most reliable index for windthrow detection. Differences in the "diff area" column further illustrate the tendency of stricter thresholds to underestimate and looser thresholds to overestimate the affected area.

To further assess robustness, we analysed the deviation of mapped area from the reference dataset (DOF, 252.5 ha) as a function of threshold (Appendix B.3, Figure A4). For all indices, looser thresholds (closer to zero) substantially overestimated the affected area, while stricter thresholds progressively underestimated it. For S2, the delineated area ranged from 164% of the reference at -0.07 to 43% at -0.17 . A similar trend was observed for PS, though the deviation from the reference was less pronounced near the optimal thresholds. These results highlight that even small changes in threshold selection can lead to large differences in estimated disturbance extent, emphasising the importance of objective threshold definition.

Based on the sensitivity analysis, NDVI was identified as the most robust index for canopy gap detection across both sensors (Appendix B.4, Table A1). The optimal thresholds were -0.10 for S2 and -0.09 for PS, providing the best balance between precision and recall, as well as the closest match to the reference windthrow extent. These thresholds produced the highest F1-scores and Cohen's Kappa values, while minimising deviations in mapped area relative to the ground-truth dataset. Figure 7 shows the resulting binary masks for the reference, PS, and S2, confirming that both sensors captured the general distribution of canopy gaps, although with differences in spatial detail. These optimised masks were subsequently used as the basis for further accuracy assessment, including the block-bootstrap analysis.

3.3. Accuracy Assessment

3.3.1. Spatial Agreement with Reference Polygons

The total area of windthrow damage detected in this study is 252.5 ha. It was determined by manually delineating reference polygons from a high-resolution digital orthophoto (DOF) acquired in August 2023, immediately after the windthrow event, with a spatial resolution of 0.5 m. This represents reference data for windthrow detection by remote sensed imagery.

The spatial agreement between the satellite-based detections and the reference polygons was assessed through an intersection analysis (see Table 3, Figure 8). Agreement refers to areas where the satellite-detected damage overlaps with reference polygons. Underestimation denotes areas present in the reference data that were not detected in the satellite imagery (i.e., false negatives), while overestimation denotes areas detected by the satellite data that are not present in the reference polygons (i.e., false positives). PS images achieved a higher spatial agreement with reference data (80.9%) than the S2 images (77.2%). The

overestimation was greater for S2 (22.8%) compared to PS (19.1%), mainly due to the coarser spatial resolution of S2 (10 m vs. 3 m), which led to a generalisation of the boundaries and partial inclusion of undamaged areas. Similarly, the underestimation was more pronounced for the S2 results (30.8%) than for PS (27.5%), reflecting the lower sensitivity to small or fragmented areas. Figure A5 in Appendix C.1 shows the numerical (quantitative) results, while Figure 8 illustrates the spatial (qualitative) correspondence between the reference polygons and the satellite-derived detections. The area automatically detected from S2 images overlapped with reference by 174.5 ha, which corresponds to almost the entire part of the S2 coverage that matches the reference dataset (174.7 ha = 99.9%).

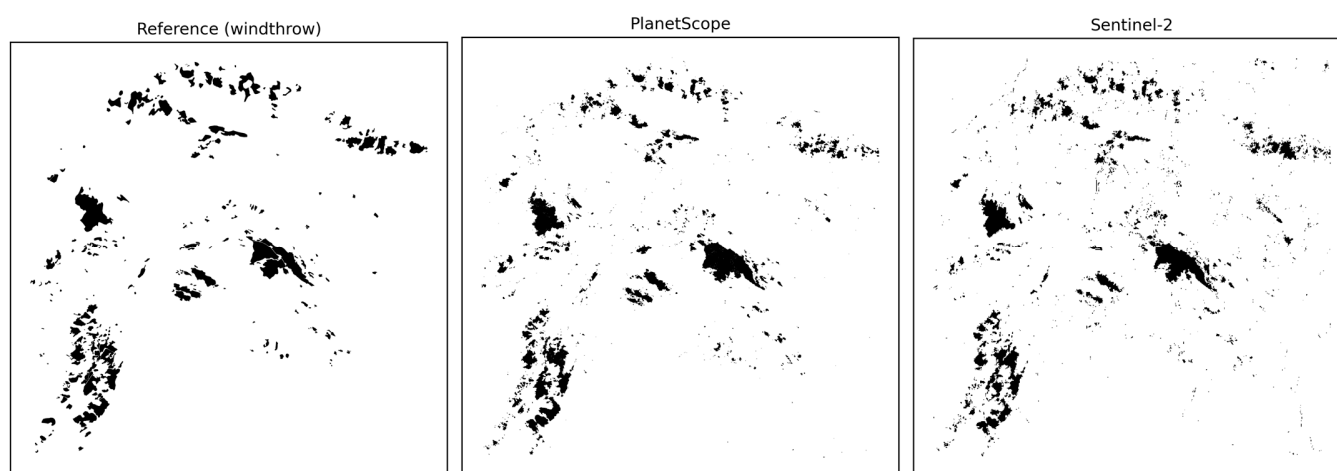


Figure 7. Binary windthrow masks derived from **(left)** reference data (high-resolution orthophoto), **(middle)** PlanetScope NDVI at threshold -0.09 , and **(right)** Sentinel-2 NDVI at threshold -0.10 . Reference polygons represent ground-truth windthrow delineation, while the satellite-based masks correspond to the optimised thresholds identified through sensitivity analysis.

Table 3. Comparison of detected changes with reference data.

	Sentinel-2	PlanetScope
Reference dataset	252.5 ha	
Detected area	226.2 ha	226.3 ha
Matching with reference dataset	174.7 ha (77.2%)	183.1 ha (80.9%)
Overestimation	51.5 ha (22.8%)	43.2 ha (19.1%)
Underestimation	77.8 ha (30.8%)	69.4 ha (27.5%)
S2-PS match	174.5 ha	

While the intersection analysis provides valuable insights into the spatial overlap and the nature of omission and commission errors, it assumes independence of individual pixels. Neighbouring pixels are spatially autocorrelated, meaning that errors tend to cluster rather than occur randomly. As a result, naïve accuracy estimates may overstate the reliability of the results by producing confidence intervals that are too narrow. To overcome this limitation, we applied a block-bootstrap approach, which accounts for spatial autocorrelation and provides more robust uncertainty estimates of classification accuracy.

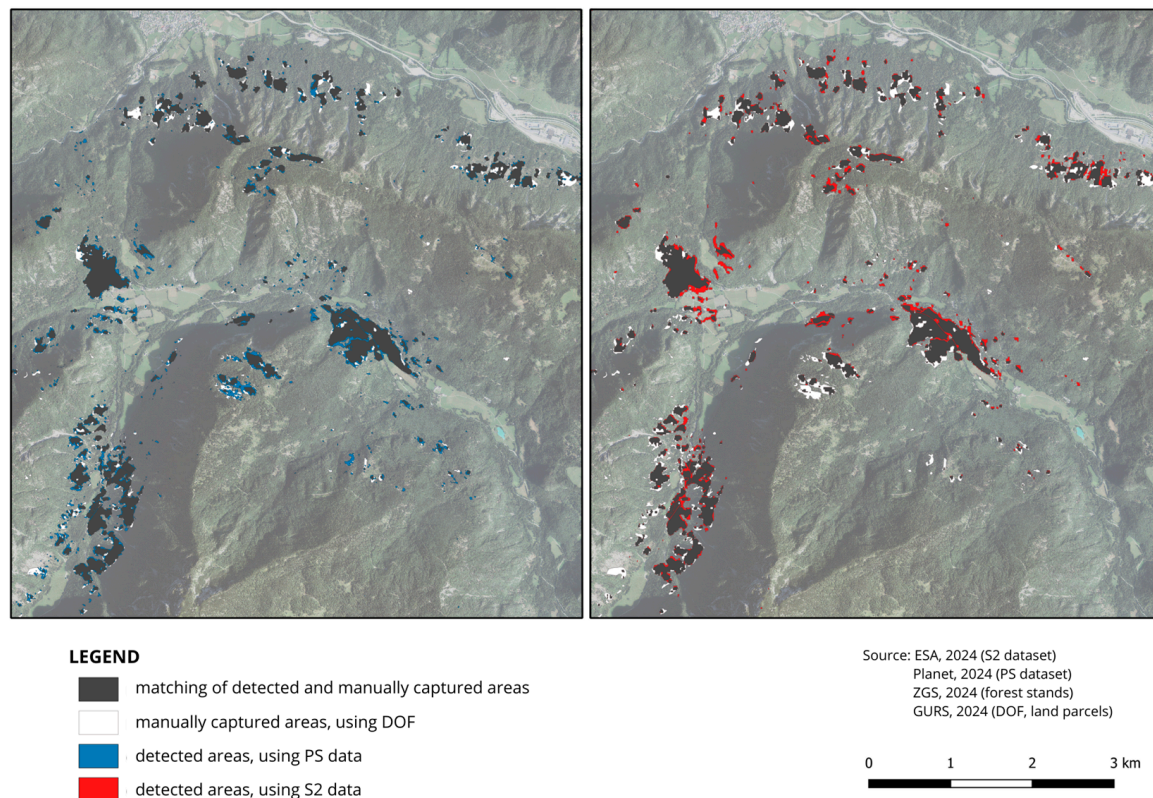


Figure 8. Spatial comparison of satellite-derived change detection areas with manually delineated reference (orthophoto) polygons. The figure shows two separate maps: one for PlanetScope (PS) and one for Sentinel-2 (S2). Grey areas indicate reference polygons that overlap with the satellite detections, white areas show reference polygons not detected by the satellite, blue areas correspond to PlanetScope detections outside the reference polygons, and red areas correspond to Sentinel-2 detections outside the reference polygons. This visualisation complements the numerical evaluation presented in Figure 5.

3.3.2. Bootstrap Accuracy and Confidence Intervals

To account for spatial autocorrelation among neighbouring pixels, classification accuracy was evaluated using a block-bootstrap approach (1-ha blocks, 500 iterations). This method provides more conservative, yet realistic uncertainty estimates compared to naïve per-pixel assessments. In this within-sample evaluation, Sentinel-2 achieved a mean accuracy of 69.2% with a 95% confidence interval of 67.4–71.2%, while PS achieved a slightly higher accuracy of 72.9% with a 95% confidence interval of 71.2–74.6%. Within this sample, the results are consistent with PS's finer spatial resolution improving the detection of small or fragmented windthrow patches. Both sensors' accuracies fall within the reported bootstrap intervals and should be interpreted as within-sample uncertainty rather than as external validity.

To further evaluate the consistency between satellite systems, we compared the S2 mask directly against the PS-derived mask, treating the latter as a high-resolution reference. This cross-sensor comparison yielded an agreement of 76.9% with a 95% confidence interval of 75.3–78.4%. These results highlight that while S2 can reproduce most of the canopy gaps detected by PS, discrepancies remain in smaller and more fragmented areas, which are more accurately captured by the higher spatial resolution of PS imagery.

3.3.3. Omission Analysis

Omission analysis revealed clear differences in detection performance between S2 and PS, particularly for smaller canopy gaps (Figure 9). For S2, omission rates reached 64.2% for the smallest gaps (<0.1 ha) and 23.1% for medium gaps (0.1–0.5 ha), while almost all larger gaps (>0.5 ha) were correctly identified (1.6% omitted). In contrast, PS demonstrated substantially higher sensitivity to small and medium gaps, with omission rates of 34.5% and 8.9%, respectively, and similarly low omission for large gaps (1.3%). These findings are within-sample and consistent with the higher spatial resolution of PS substantially reduces the omission of small and fragmented windthrow patches, while both sensors perform reliably for larger canopy gaps.

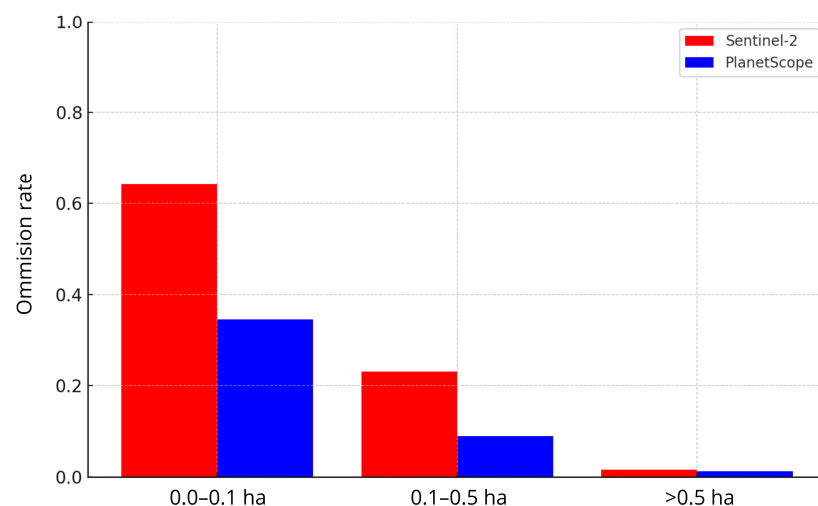


Figure 9. Omission rates of canopy gaps by size class (0–0.1 ha, 0.1–0.5 ha, >0.5 ha) for Sentinel-2 (red) and PlanetScope (blue). Both sensors perform well for large canopy gaps (>0.5 ha), while omission is more pronounced for smaller and fragmented gaps, especially in the Sentinel-2 dataset.

Taken together with the bootstrap intervals in Section 3.3.2, these omission patterns indicate size-dependent behaviour in this sample. PlanetScope detected a larger fraction of small and fragmented gaps than Sentinel-2, whereas both sensors showed high detection rates for gaps greater than 0.5 ha. The confidence intervals remain within-sample and should be considered alongside the size-dependent omissions reported here.

3.4. Validation of Damaged Timber Volume Estimates

To assess the plausibility of the parcel-level estimates of damaged timber volume, we compared them with official sanitary-logging records from SFS for a small set of cadastral parcels where post-storm sanitary-logging had been completed. The eight parcels included in this comparison are shown in Figure 10. This analysis was designed as a within-sample plausibility check, not full statistical validation, because parcel-level records are not publicly available and the validation sample ($n = 8$) was small and non-probability.

Timber volumes estimated from Sentinel-2, PlanetScope, and DOF-based delineations were compared with the official sanitary-logging records (Figure 11). At the parcel-level ($n = 8$, non-probability), all three remote sensing-based estimates followed the same general trend as the official records, with parcel-to-parcel variability. In this sample, PlanetScope produced the lowest mean absolute deviations (approximately 5–7%), whereas Sentinel-2 occasionally under- or overestimated volumes, consistent with its coarser spatial resolution and small-gap omissions. DOF-based delineations were broadly consistent but sometimes included pre-existing canopy openings not caused by the windstorm, which can introduce

positive bias. These comparisons constitute a within-sample plausibility check and should not be interpreted as a general validation.

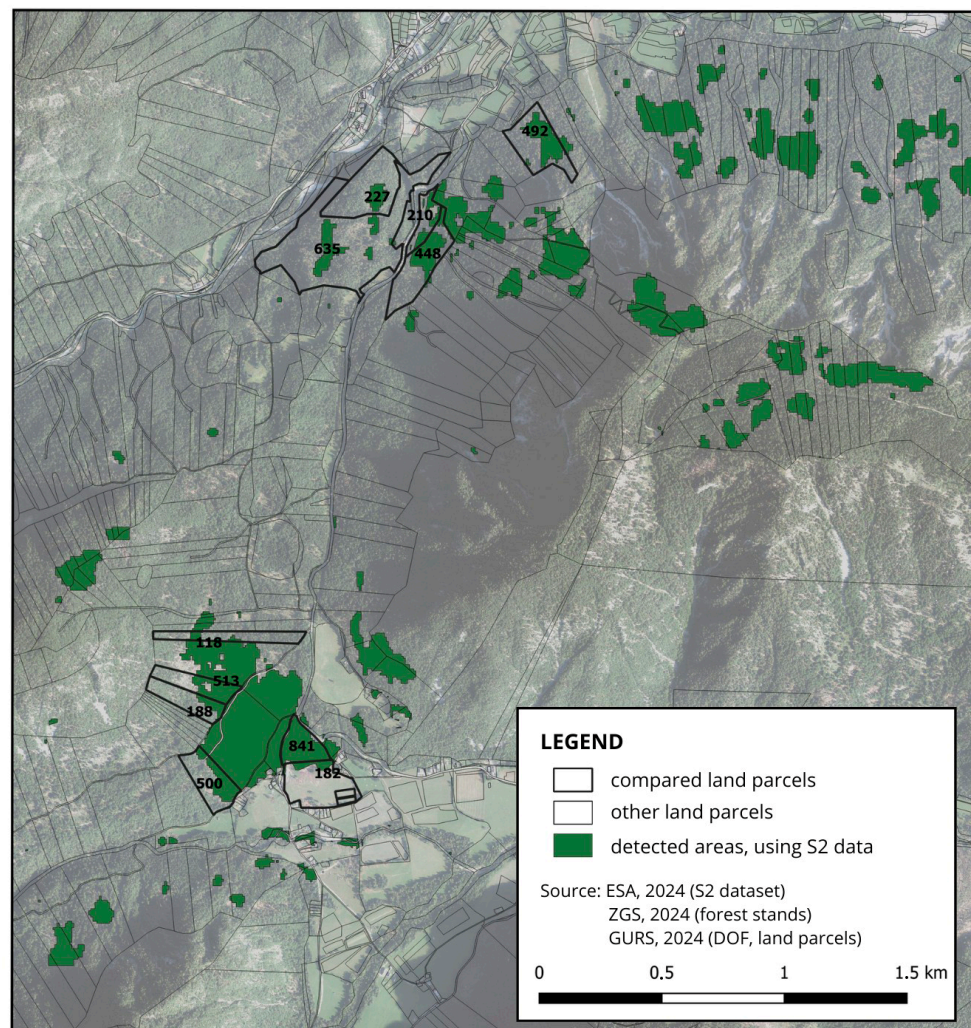


Figure 10. Parcel-level damaged timber volumes derived from Sentinel-2 for the eight cadastral parcels used in the plausibility check (m^3). The subset includes parcels where post-storm sanitary logging had been completed and could be matched to entries in the SFS sanitary-logging register.

Taken together, the comparisons indicate that, for this event and sample ($n = 8$, non-probability), the workflow yields internally consistent parcel-level estimates (Appendix C.2, Figure A6). However, given the small sample size and potential bias in the 2021 stand data—only partially mitigated by incorporating 2021–2023 sanitary-logging records—these results should not be interpreted as a quantitative accuracy assessment or external validation. We therefore treat them as a within-sample plausibility check, showing that the satellite-derived estimates are of the same order of magnitude as the operational forestry records.

At the study-area scale, aggregated timber volume estimates derived from Sentinel-2, PlanetScope, and DOF-based delineations were broadly consistent with totals reported in the SFS sanitary-logging database (Table 4; Figure A7 in Appendix C.3). In this sample, PlanetScope showed the closest agreement with the official totals. These results are within-sample and preliminary and should not be interpreted as evidence of operational performance or general accuracy; rather, they indicate consistency in order of magnitude between the mapping-and-aggregation workflow and the figures reported in official records.

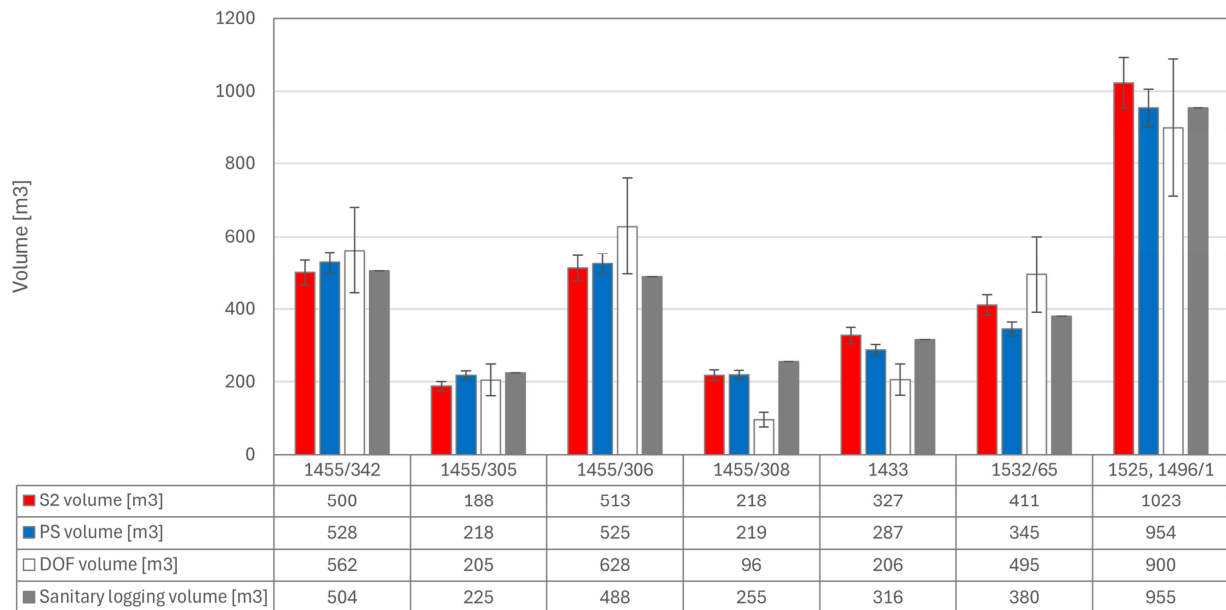


Figure 11. Comparison of parcel-level damaged timber volumes derived from Sentinel-2 (S2), PlanetScope (PS), and orthophoto-based (DOF) delineations with the volumes recorded in the official sanitary-logging database.

Table 4. Aggregated damaged timber volumes at the study-area scale (m³). Values from remote sensing are totals obtained by aggregating parcel-level estimates; SFS refers to totals from the official sanitary-logging database.

Data Source	Timber Volume (m ³)
SFS estimation	>60,000
DOF	80,000
Sentinel-2	69,000
PlanetScope	72,000

4. Discussion

This paper presents a regional case study of established multi-temporal optical change detection applied to the July 2023 windthrow in Triglav National Park. All findings are within-sample and do not claim methodological novelty or operational readiness. In this sample, overall, within-sample accuracy was 69.2% for Sentinel-2 and 72.9% for PlanetScope, with a marked patch-size effect: gaps greater than 0.5 ha were consistently detected, whereas gaps smaller than 0.1 ha were frequently omitted. A preliminary parcel-level comparison with official sanitary-logging records ($n = 8$, non-probability) indicated small, within-sample deviations and is treated as a plausibility check, not a general accuracy estimate.

Previous studies reporting higher accuracies for windthrow mapping typically analyse large, contiguous blowdowns with strong spectral contrast [6,36], use optimally paired image dates and illumination to maximise separability [37], and rely on very high resolution (VHR) optical [4] or SAR data [7,12]. Across the reviewed literature, reported accuracies chiefly concern 2D mapped area rather than timber volume. Here, we complement that perspective by examining 3D damaged-volume estimates at a regional scale for a limited, non-probability sample of parcels, while adhering to a within-sample interpretation of the results. External accuracy assessments should rely on probability-based reference samples

and estimators aligned with the sampling design [38,39]. Most published windthrow studies report overall accuracies above approximately 80% (sometimes approaching 90%) under favourable conditions such as large, contiguous blowdowns and optimally paired imagery. By contrast with those settings, our event features a long tail of small, fragmented openings where 10 m pixels are prone to mixing; imagery acquisition required cloud mosaicking and temporal offsets; and the reference data carry uncertainty (orthophoto delineations may include pre-existing openings, and stand data originate from 2021 and were only partially adjusted with 2021–2023 sanitary-logging records). These factors plausibly account for the lower within-sample performance reported earlier (see Section 3.3).

Rather than proposing new algorithms, this work provides a gap-size-explicit analysis that quantifies omissions of small and fragmented openings in alpine terrain; reports block-bootstrap confidence intervals that account for spatial autocorrelation, offering transparent within-sample uncertainty; and examines a parcel-level linkage to official sanitary-logging records, focusing on management-relevant units while explicitly acknowledging the limitations of a small, non-probability parcel sample. Together, these elements document how established methods perform under the topographic, illumination, and patch-size constraints characteristic of the Alps and clarify where small-gap omissions and reference-data lineage most affect downstream parcel-level estimates.

This study has several important limitations. First, the parcel sample in the timber-volume analysis is small ($n = 8$) and non-probability, so deviations cannot be interpreted as general accuracy; the comparison serves only as a within-sample plausibility check. Second, the reference data are uncertain: orthophoto-based polygons may include pre-existing canopy openings, salvage logging progressed during the observation window, and parcel-level administrative records are incomplete or not fully accessible. Third, temporal pairing was constrained—cloud cover and date offsets likely reduced separability compared to optimally paired pre- and post-event scenes. Fourth, topography and resolution are significant: rugged relief, shadows, and 10 m pixel mixing contribute to boundary smoothing and omissions of gaps smaller than 0.1 ha. Fifth, the fixed thresholds were tailored to this event and are not assumed to transfer without local adjustment. Accordingly, we report within-sample accuracies with block-bootstrap confidence intervals (accounting for spatial autocorrelation), present a gap-size-explicit analysis to reveal small-patch omissions, and treat the parcel-level comparison as preliminary rather than an external validation. Producing generalisable accuracy and area estimates would require a probability-based reference sample and estimators aligned with the sampling design [38,39].

In this sample, PlanetScope omitted fewer small or fragmented gaps than Sentinel-2, consistent with its finer spatial resolution, while both sensors performed similarly on larger openings. Very small openings (<0.1 ha) often require field verification or VHR/UAV follow-up. The reported bootstrap intervals express within-sample uncertainty and should not be interpreted as operational performance bounds. Broader generalisation will require probability-based validation across additional events and forest types, as well as wider access to parcel-level official records.

Priority directions include, first, adopting probability sampling to produce unbiased accuracy and area estimates with standard errors; second, improving reference data using VHR/UAV or LiDAR to reduce small-gap omissions and label noise; third, systematically testing temporal pairing and topographic or illumination corrections; and fourth, evaluating multi-sensor approaches (optical and SAR) together with size-stratified metrics to better characterise detection limits across sensors and event types.

5. Conclusions

This regional case study assessed the applicability of multi-temporal Sentinel-2 and PlanetScope imagery for mapping windthrow in alpine forests and for deriving parcel-level estimates of damaged timber volume by linking satellite-derived change maps with cadastral and forest-management layers. All results are within-sample for the July 2023 windstorm in Triglav National Park.

Both sensors performed consistently on larger openings, with gaps greater than 0.5 ha detected more reliably, while PlanetScope showed a modest advantage on small or fragmented patches and Sentinel-2 exhibited more omissions for gaps smaller than 0.1 ha. A preliminary parcel-level comparison with official sanitary-logging records, based on a small, non-probability set of parcels ($n = 8$), serves as a plausibility check and does not establish general accuracy.

The contribution is not methodological novelty or an operational system, but documentation of how established methods perform in alpine terrain under constraints of patch size, topography and illumination, and available reference data, together with transparent block-bootstrap confidence intervals. Broader generalisation will require probability-based validation across additional events and forest types, improved reference data (e.g., VHR/UAV or LiDAR to address small-gap omissions and label noise), careful attention to temporal pairing, and cautious threshold transfer, as thresholds used here were tailored to this event and are not assumed to transfer without local calibration.

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

Funding: This research was funded by Slovenian Research Agency as part of the research project J2-3055 (ROVI—Innovative radar and optical satellite image time series fusion and processing for monitoring the natural environment) and the research programme P2-0406 (Earth observation and geoinformatics).

Data Availability Statement: Sentinel-2 imagery used in this study is openly available through the Copernicus Data Space Ecosystem (CDSE) portal at <https://dataspace.copernicus.eu> (accessed on 4 August 2025). PlanetScope imagery was provided by Planet Labs and can be accessed through the official website at <https://www.planet.com> (accessed on 2 August 2025); however, redistribution of these data to third parties is not permitted. In situ data on timber volumes and sanitary-logging records were obtained from the Slovenian Forestry Service. These datasets are confidential and therefore cannot be shared publicly.

Acknowledgments: The authors thank the Slovenian Forestry Institute for providing access to the relevant datasets and the Slovenia Forest Service for their support during the study. Special thanks go to Jure Legat, at SFS, for dedicating his time and expertise to assist with the review and analysis of the data. The authors also thank Planet Labs for providing limited free access to PlanetScope satellite imagery for research aim.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CD	Change Detection
DOF	Digital Orthophoto

GURS	Surveying and Mapping Authority of the Republic of Slovenia
NIR	Near Infrared
NDVI	Normalised Difference Vegetation Index
PS	PlanetScope
S2	Sentinel-2
SFI	Slovenian Forest Institute
SFS	Slovenian Forest Service
UAV	unmanned aerial vehicles

Appendix A. Sentinel-2 Time Series (NDVI, NDRE and NDMI)

Figure A1 presents Sentinel-2 time series for the damaged area from 1 June to 1 October 2023. Each panel displays a vegetation or water-content index (NDVI, NDRE, NDMI) for every available cloud-free observation; the vertical dashed red line marks the windthrow date (19 July 2023), and the grey dashed line indicates the pre-event mean. All three indices show a clear level shift after the event, consistent with canopy disturbance: NDVI and NDRE decline from early-summer values and stabilise at a lower plateau, while NDMI decreases and remains depressed, indicating a reduced canopy moisture or biomass signal. These patterns are consistent with the change-detection results reported in the main text.

Sentinel-2 L2A scenes were atmospherically corrected and cloud or shadow masked; indices were calculated from surface reflectance and aggregated as observation-level values over the mapped damaged area (with no temporal smoothing beyond the native revisit). Dates on the x-axis represent scene acquisition dates.

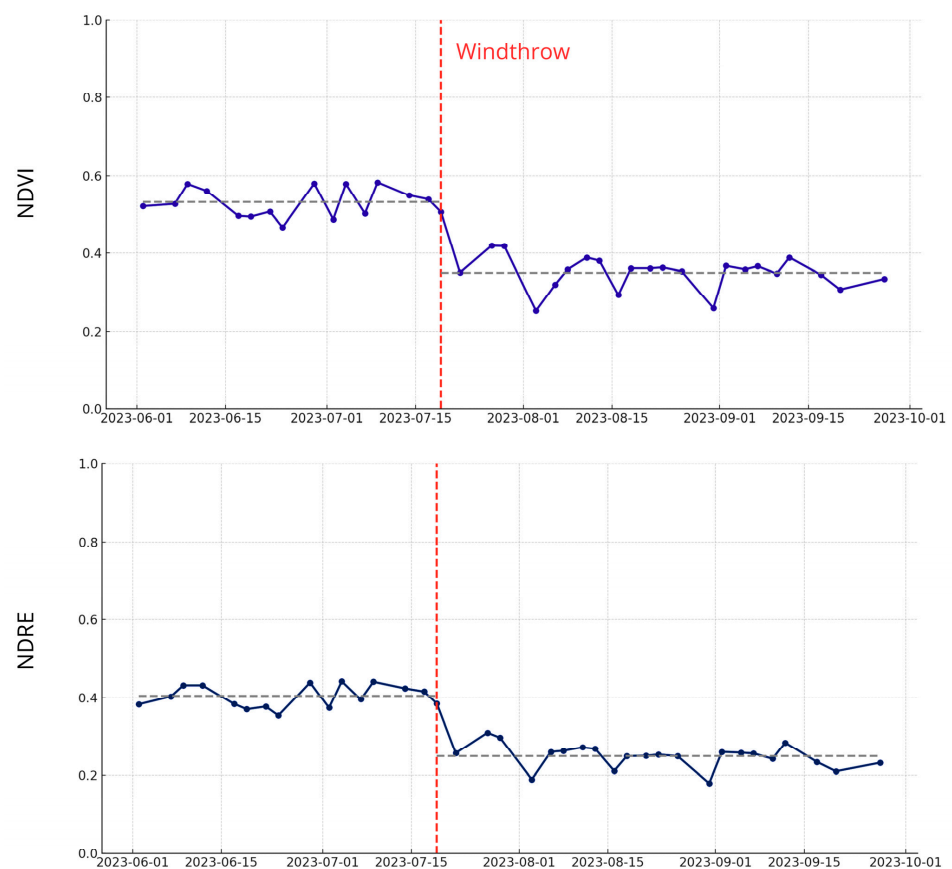


Figure A1. Cont.

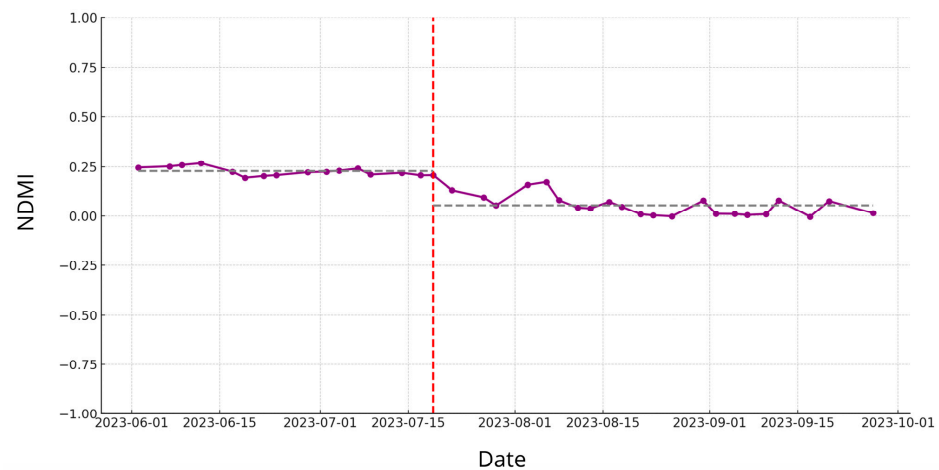


Figure A1. Sentinel-2 time series over the damaged area shows a marked drop in NDVI, NDRE, and NDMI during the windthrow event (red dashed line), followed by persistently lower post-event levels relative to the pre-event mean (grey dashed line).

Appendix B. Threshold Selection Histograms, Sensitivity and Additional Result Evaluation

This appendix provides the full materials underlying the threshold selection and performance evaluation reported in the main text. We include complete threshold-sensitivity curves, extended accuracy metrics, and size-stratified omission/commission summaries. These materials support the operating points and results presented in Section 3.3, while keeping the main text concise.

Appendix B.1

This appendix presents the distributions of index differences used to select fixed thresholds for change detection. For each sensor (Sentinel-2 and PlanetScope), we show histograms of dNDVI, dNDRE, and, for Sentinel-2, dNDMI over the analysis area. The vertical dashed line indicates the mean change value, while the vertical dotted line indicates the 5th percentile, which was used as the lower-tail cutoff for the threshold. These plots support the choices reported in Section 3.1.

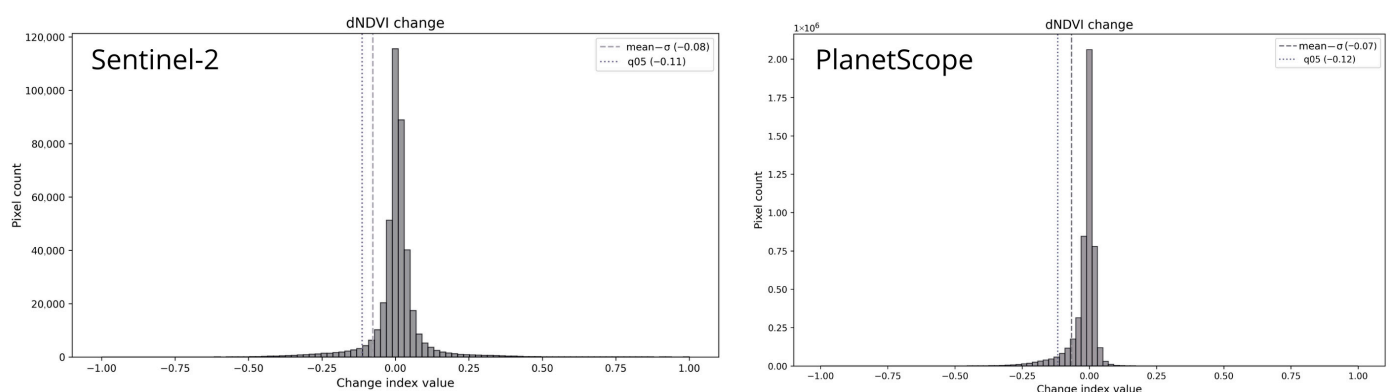


Figure A2. Cont.

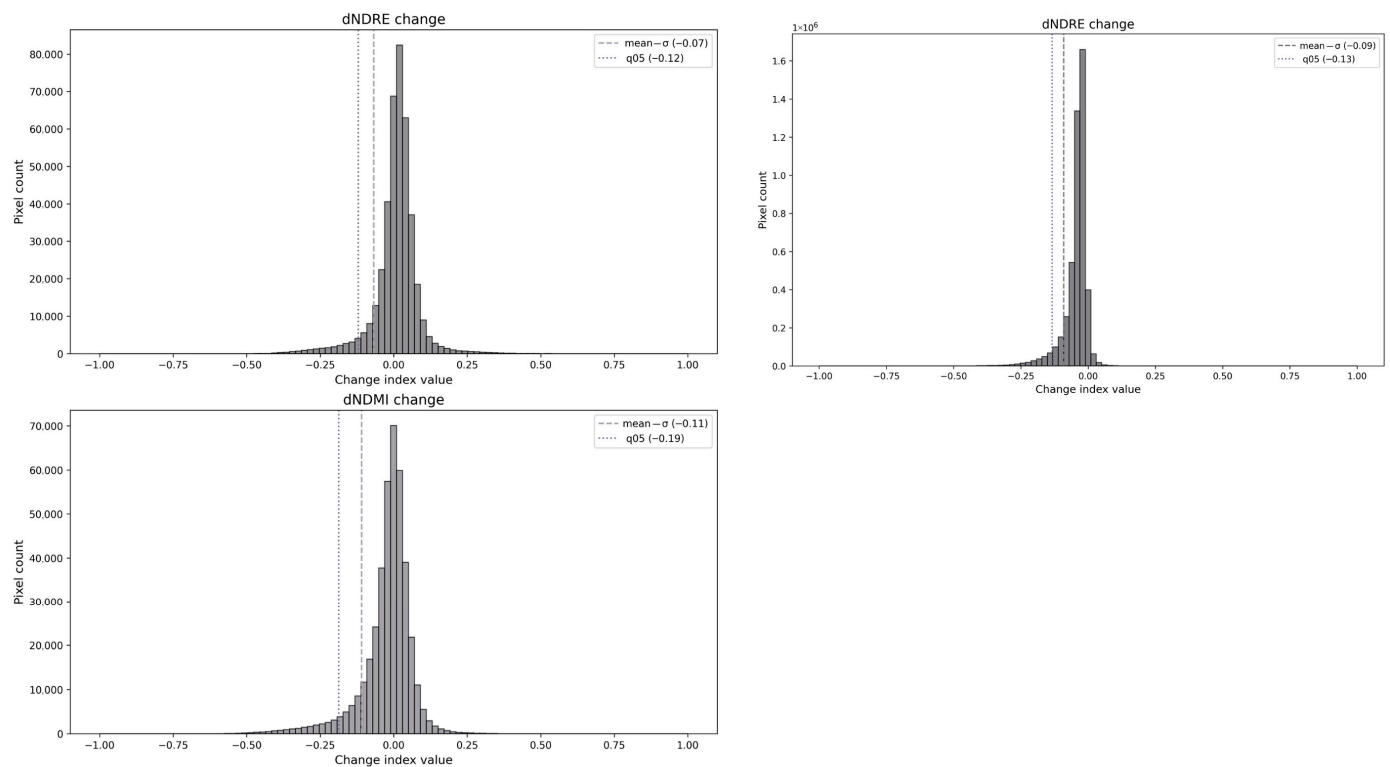


Figure A2. Histograms of dNDVI, dNDRE, dNDMI derived from Sentinel-2 and PlanetScope imagery. Vertical dashed lines indicate the mean change value and the selected threshold based on the 5th percentile (vertical dotted lines).

Appendix B.2. Threshold Sensitivity by Spectral Index (Lead-In)

This section presents the full threshold–sensitivity analysis underpinning the index selection used in the main text. For each sensor we swept negative change thresholds (τ) for multiple indices— Δ NDVI, Δ NDRE, and (for Sentinel-2) Δ NDMI—computing the F1-score at each τ . Histogram-based thresholds were derived to provide an objective operating point. Across sensors, Δ NDVI delivered the strongest performance over most τ , so it was retained for the headline results.

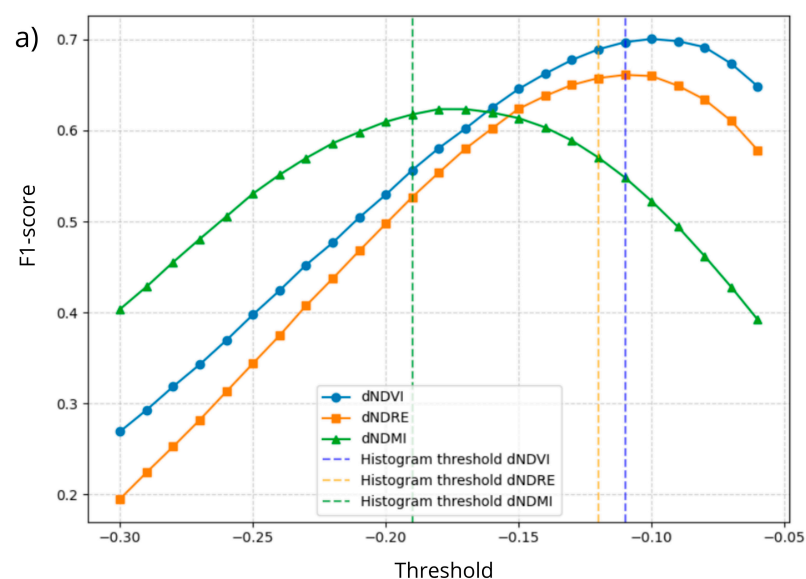


Figure A3. Cont.

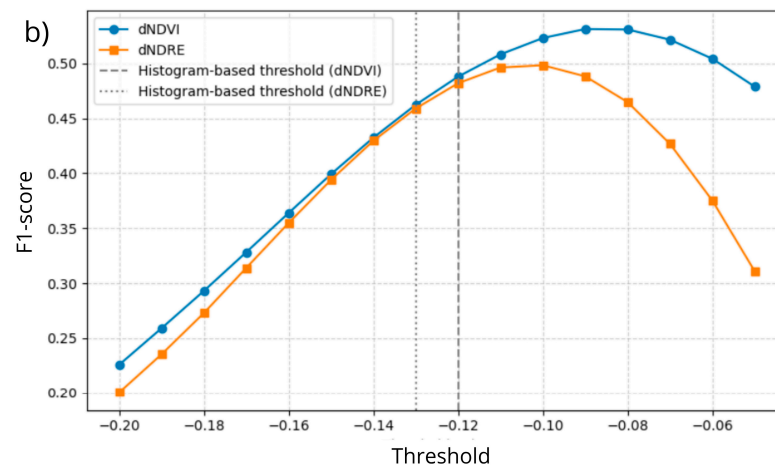


Figure A3. F1-score as a function of threshold values for different spectral indices. (a) Sentinel-2: performance curves for dNDVI, dNDRE, and dNDMI, with vertical dashed lines indicating histogram-based thresholds. (b) PlanetScope: performance curves for dNDVI and dNDRE, with histogram-based thresholds indicated. NDVI consistently outperformed other indices for both sensors.

Appendix B.3. Sensitivity of Mapped Windthrow Area to the Change-Threshold (τ) by Sensor and Index

This appendix quantifies how the mapped windthrow area responds to the choice of change threshold (τ) for each sensor and spectral index. For Sentinel-2, we evaluated Δ NDVI, Δ NDRE, and Δ NDMI; for PlanetScope, Δ NDVI and Δ NDRE. For each index, we swept a grid of negative τ values (from strong to weak change) on the post – pre composites and, at each τ , computed the percentage deviation of the delineated area from the DOF reference (252.5 ha). Negative deviations indicate underestimation, while positive deviations indicate overestimation relative to reference.

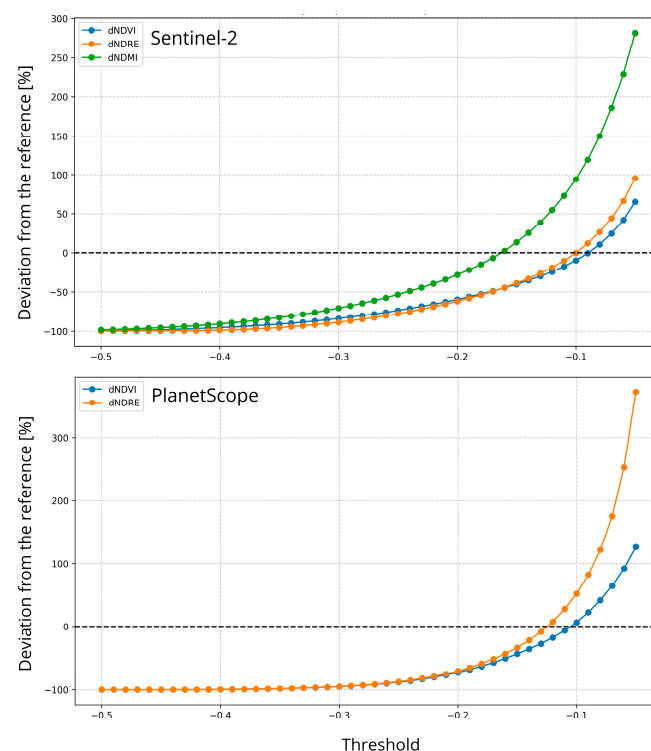


Figure A4. Sensitivity analysis of windthrow detection thresholds for Sentinel-2 (**top**) and PlanetScope (**bottom**). Lines show the deviation (%) of the delineated area from the reference dataset (DOF,

252.5 ha) as a function of the threshold value for different spectral indices (dNDVI, dNDRE, dNDMI). Negative deviations indicate underestimation, while positive deviations indicate overestimation of the affected area.

Appendix B.4

Table A1. Classification performance of windthrow detection using different spectral indices (NDVI, NDRE, NDMI) and thresholds for PlanetScope (PS) and Sentinel-2 (S2). Reported metrics include accuracy, precision, recall, specificity, F1-score, intersection-over-union (IoU), Cohen’s kappa, and difference in mapped area relative to reference 252.5 ha. Bold values indicate the selected optimal thresholds per index.

Data	Index	Threshold	Accuracy	Precision	Recall	Specificity	F1-Score	IoU	Kappa	Diff Area [%]
PS	NDVI	−0.08	0.945	0.453	0.641	0.961	0.531	0.361	0.503	+41.4
		−0.09	0.949	0.482	0.591	0.968	0.531	0.362	0.505	+22.6
		−0.10	0.952	0.507	0.541	0.973	0.523	0.354	0.498	+6.74
		−0.11	0.954	0.526	0.492	0.977	0.508	0.341	0.484	−6.6
		−0.12	0.955	0.540	0.445	0.981	0.488	0.323	0.465	−17.7
	NDRE	−0.09	0.951	0.490	0.382	0.980	0.430	0.274	0.404	+82.2
		−0.10	0.939	0.413	0.627	0.955	0.498	0.332	0.467	+51.6
		−0.11	0.944	0.443	0.564	0.964	0.496	0.330	0.467	+27.5
		−0.12	0.948	0.465	0.501	0.971	0.482	0.318	0.455	+7.71
		−0.13	0.950	0.481	0.439	0.976	0.459	0.298	0.433	−8.6
	NDMI	−0.14	0.951	0.490	0.382	0.980	0.430	0.274	0.404	−22.0
		−0.09	0.971	0.700	0.696	0.985	0.698	0.536	0.683	−0.56
		−0.10	0.972	0.739	0.666	0.988	0.701	0.539	0.686	−9.77
		−0.11	0.973	0.770	0.637	0.990	0.697	0.535	0.683	−17.33
S2	NDVI	−0.10	0.967	0.660	0.659	0.983	0.660	0.492	0.642	−0.15
		−0.11	0.969	0.699	0.627	0.986	0.661	0.494	0.645	−10.21
		−0.12	0.970	0.735	0.595	0.989	0.658	0.490	0.642	−18.97
	NDRE	−0.17	0.965	0.645	0.603	0.983	0.623	0.453	0.605	−6.50
		−0.18	0.966	0.677	0.578	0.986	0.623	0.453	0.606	−14.61
		−0.19	0.967	0.704	0.550	0.988	0.618	0.447	0.601	−21.82
		−0.20	0.967	0.729	0.524	0.990	0.610	0.438	0.593	−28.21

Appendix C. Parcel-Level Damaged-Volume Checks (Within-Sample)

This appendix summarises the within-sample comparisons used to assess the plausibility of parcel-level damaged volume estimates. Satellite-derived change maps (Sentinel-2, PlanetScope) and orthophoto delineations (DOF) were intersected with forest stand data to estimate timber volume per parcel and compared with official sanitary-logging records (SFS). As noted in the main text, the parcel set is small ($n = 8$) and non-probability; the figures below therefore illustrate within-sample behaviour and do not constitute external validation.

Agreement, underestimation, or overestimation in C1 refers to area correspondence between satellite maps and the orthophoto reference; it is not a direct measure of volume accuracy. Figures A6 and A7 show volume estimates aggregated to parcels; these are affected by reference uncertainty, as orthophoto polygons may include pre-existing openings, and stand volumes originate from 2021 and were only partially adjusted using 2021–2023 sanitary logging records.

Appendix C.1

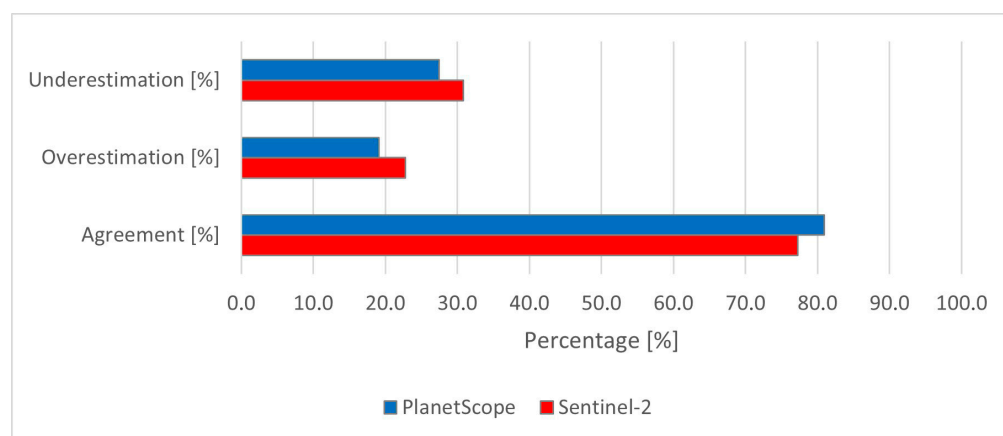


Figure A5. Accuracy evaluation of satellite-based forest damage mapping (PlanetScope, Sentinel-2) using orthophoto reference areas. The bars represent the proportion of detected areas classified as agreement (correctly detected damage overlapping with reference), underestimation (areas present in the reference but missed by satellite), and overestimation (areas detected by satellite but not present in the reference).

Appendix C.2

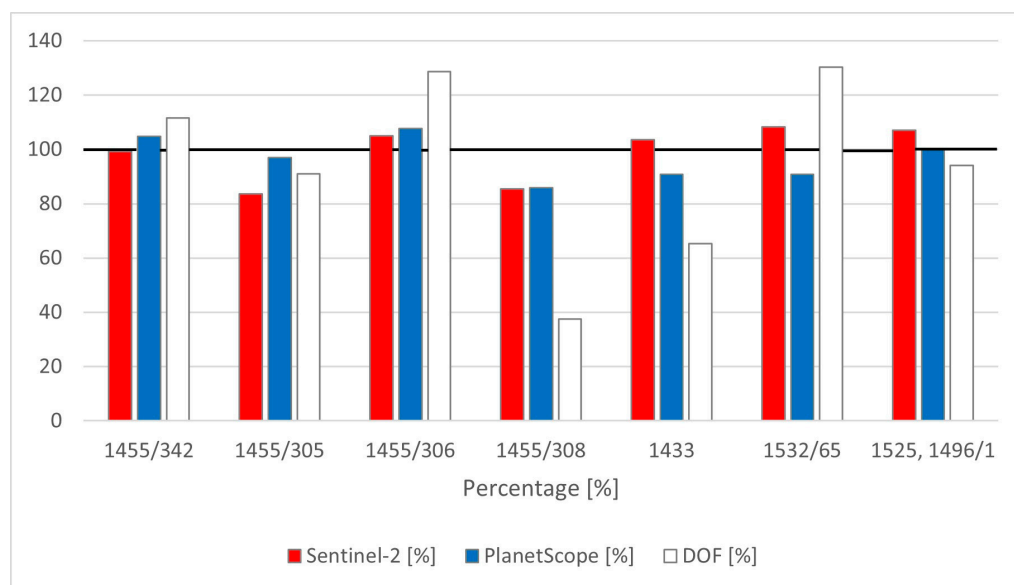


Figure A6. Damaged timber volume per land parcel, estimated from Sentinel-2, PlanetScope, and DOF data, compared with official sanitary-logging records.

Appendix C.3

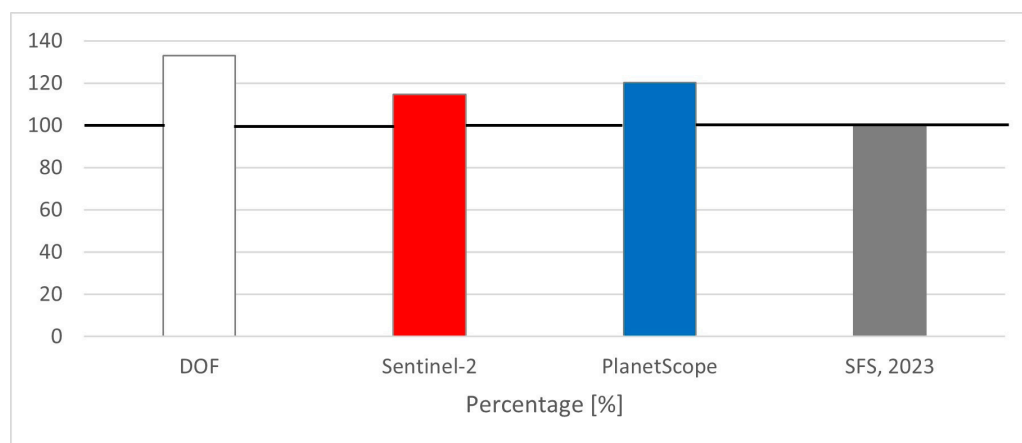


Figure A7. Correspondence of estimated damaged timber volume between different data sources and the reference SFS's timber database.

References

- Seidl, R.; Schelhaas, M.-J.; Rammer, W.; Verkerk, P.J. Increasing Forest Disturbances in Europe and Their Impact on Carbon Storage. *Nat. Clim. Change* **2014**, *4*, 806–810. [\[CrossRef\]](#) [\[PubMed\]](#)
- Usbeck, T.; Wohlgemuth, T.; Dobbertin, M.; Pfister, C.; Bürgi, A.; Rebetez, M. Increasing Storm Damage to Forests in Switzerland from 1858 to 2007. *Agric. For. Meteorol.* **2010**, *150*, 47–55. [\[CrossRef\]](#)
- Hlásny, T.; Krokene, P.; Liebhold, A.; Montagné-Huck, C.; Müller, J.; Qin, H.; Raffa, K.; Schelhaas, M.-J.; Seidl, R.; Svoboda, M.; et al. *Living with Bark Beetles: Impacts, Outlook and Management Options*; From Science to Policy; European Forest Institute: Joensuu, Finland, 2019.
- Einzmann, K.; Immitzer, M.; Böck, S.; Bauer, O.; Schmitt, A.; Atzberger, C. Windthrow Detection in European Forests with Very High-Resolution Optical Data. *Forests* **2017**, *8*, 21. [\[CrossRef\]](#)
- Jonikavičius, D.; Mozgeris, G. Rapid Assessment of Wind Storm-Caused Forest Damage Using Satellite Images and Stand-Wise Forest Inventory Data. *Iforest-Biogeosci. For.* **2013**, *6*, 150–155. [\[CrossRef\]](#)
- Dalponte, M.; Marzini, S.; Solano-Correa, Y.T.; Tonon, G.; Vescovo, L.; Gianelle, D. Mapping Forest Windthrows Using High Spatial Resolution Multispectral Satellite Images. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *93*, 102206. [\[CrossRef\]](#)
- Dalponte, M.; Solano-Correa, Y.T.; Marinelli, D.; Liu, S.; Yokoya, N.; Gianelle, D. Detection of Forest Windthrows with Bitemporal COSMO-SkyMed and Sentinel-1 SAR Data. *Remote Sens. Environ.* **2023**, *297*, 113787. [\[CrossRef\]](#)
- Wang, F.; Xu, Y.J. Comparison of Remote Sensing Change Detection Techniques for Assessing Hurricane Damage to Forests. *Environ. Monit. Assess.* **2010**, *162*, 311–326. [\[CrossRef\]](#)
- Rich, R.L.; Frelich, L.; Reich, P.B.; Bauer, M.E. Detecting Wind Disturbance Severity and Canopy Heterogeneity in Boreal Forest by Coupling High-Spatial Resolution Satellite Imagery and Field Data. *Remote Sens. Environ.* **2010**, *114*, 299–308. [\[CrossRef\]](#)
- Baumann, M.; Ozdogan, M.; Wolter, P.T.; Krylov, A.; Vladimirova, N.; Radeloff, V.C. Landsat Remote Sensing of Forest Windfall Disturbance. *Remote Sens. Environ.* **2014**, *143*, 171–179. [\[CrossRef\]](#)
- Schwarz, M.; Steinmeier, C.; Holec, F.; Stebler, O.; Wagner, H. Detection of Windthrow in Mountainous Regions with Different Remote Sensing Data and Classification Methods. *Scand. J. For. Res.* **2003**, *18*, 525–536. [\[CrossRef\]](#)
- Rüetschi, M.; Small, D.; Waser, L.T. Rapid Detection of Windthrows Using Sentinel-1 C-Band SAR Data. *Remote Sens.* **2019**, *11*, 115. [\[CrossRef\]](#)
- Borsah, A.A.; Nazeer, M.; Wong, M.S. LIDAR-Based Forest Biomass Remote Sensing: A Review of Metrics, Methods, and Assessment Criteria for the Selection of Allometric Equations. *Forests* **2023**, *14*, 2095. [\[CrossRef\]](#)
- Chirici, G.; Bottalico, F.; Giannetti, F.; Del Perugia, B.; Travaglini, D.; Nocentini, S.; Kutchartt, E.; Marchi, E.; Foderi, C.; Fioravanti, M.; et al. Assessing Forest Windthrow Damage Using Single-Date, Post-Event Airborne Laser Scanning Data. *For. Int. J. For. Res.* **2018**, *91*, 27–37. [\[CrossRef\]](#)
- Mokroš, M.; Výboštok, J.; Merganič, J.; Hollaus, M.; Barton, I.; Koreň, M.; Tomašík, J.; Čerňava, J. Early Stage Forest Windthrow Estimation Based on Unmanned Aircraft System Imagery. *Forests* **2017**, *8*, 306. [\[CrossRef\]](#)
- Ogris, N.; Džeroski, S.; Jurc, M. Windthrow Factors—A Case Study on Pokljuka. *Zb. Gozdarstva Lesar.* **2004**, *74*, 59–76.
- Štofko, P.; Kodrík, M. Comparison of the Root System Architecture between Windthrown and Undamaged Spruces Growing in Poorly Drained Sites. *J. For. Sci.* **2008**, *54*, 150–160. [\[CrossRef\]](#)

18. Puhe, J. Growth and Development of the Root System of Norway Spruce (*Picea Abies*) in Forest Stands—A Review. *For. Ecol. Manag.* **2003**, *175*, 253–273. [CrossRef]
19. Slovenia Forest Service (SFS, ZGS). *Neurje Najbolj Prizadelo Gozdove na Gorenjskem*; Slovenia Forest Service: Ljubljana, Slovenia, 2023. Available online: <https://www.zgs.si/informacije/sporocila-za-javnost/neurje-najbolj-prizadelo-gozdove-na-gorenjskem-1214> (accessed on 2 August 2025).
20. Slovenia Forest Service (SFS, ZGS). *Poročilo Zavoda za Gozdove Slovenije o Gozdovih za Leto 2023*; Slovenia Forest Service: Ljubljana, Slovenia, 2024.
21. Copernicus Data Space Ecosystems Copernicus Browser. Available online: <https://browser.dataspace.copernicus.eu/> (accessed on 4 August 2025).
22. Planet Labs Planet Labs: Satellite Imagery & Earth Data Analytics. Available online: <https://www.planet.com/> (accessed on 2 August 2025).
23. SentiWiki. European Space Agency S2 Mission. Available online: <https://sentiwiki.copernicus.eu/web/s2-mission> (accessed on 2 August 2025).
24. Planet Labs Education and Research Program. Available online: <https://www.planet.com/industries/education-and-research/> (accessed on 8 September 2025).
25. Surveying and Mapping Authority of the Republic of Slovenia (SMARS, GURS). Geodetska Uprava Republike Slovenije. Available online: <https://www.e-prostor.gov.si/en/access-to-geodetic-data/> (accessed on 2 August 2025).
26. Slovenia Forest Service (SFS, ZGS). Zavod za gozdove Slovenije. Available online: <https://www.zgs.si/en/> (accessed on 2 August 2025).
27. Çinar, T.; Uslu, A.; Aydin, A. Monitoring the Rehabilitation Process of the Windthrow Area Using UAS Images and Performance Comparison of Sentinel-2A Based Different Vegetation Indexes. *Earth Sci. Inform.* **2025**, *18*, 199. [CrossRef]
28. 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]
29. Legendre, P. Spatial Autocorrelation: Trouble or New Paradigm? *Ecology* **1993**, *74*, 1659–1673. [CrossRef]
30. Marriott, P.; Efron, B.; Tibshirani, R.J. An Introduction to the Bootstrap. *J. R. Stat. Soc. Ser. A Stat. Soc.* **1995**, *158*, 347. [CrossRef]
31. Vaart, A.W.V.D. *Asymptotic Statistics*, 1st ed.; Cambridge University Press: Cambridge, UK, 1998; ISBN 978-0-511-80225-6.
32. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2008; ISBN 978-0-429-14397-7.
33. Senf, C.; Seidl, R. Natural Disturbances Are Spatially Diverse but Temporally Synchronized across Temperate Forest Landscapes in Europe. *Glob. Change Biol.* **2018**, *24*, 1201–1211. [CrossRef]
34. Slovenian Forestry Institute (SFI, GOZDIS). Sanitarni Posek Dreves. Available online: https://www.zdravgozd.si/sanitarni_index.aspx (accessed on 8 September 2025).
35. Legat, J. (Slovenian Forest Services, OE Bled, KE Jesenice, Slovenia). Personal communication, 2024.
36. Elatawneh, A.; Wallner, A.; Manakos, I.; Schneider, T.; Knoke, T. Forest cover database updates using multi-seasonal RapidEye data—Storm event assessment in the Bavarian forest national park. *Forests* **2014**, *5*, 1284–1303. [CrossRef]
37. Vaglio Laurin, G.; Francini, S.; Luti, T.; Chirici, G.; Pirotti, F.; Papale, D. Satellite Open Data to Monitor Forest Damage Caused by Extreme Climate-Induced Events: A Case Study of the Vaia Storm in Northern Italy. *Forestry* **2021**, *94*, 407–416. [CrossRef]
38. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [CrossRef]
39. Stehman, S.V.; Foody, G.M. Key issues in rigorous accuracy assessment of land cover products. *Remote Sens. Environ.* **2019**, *231*, 111199. [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.