



Open satellite archives

Forest Event Records

with every control **published**

Storm damage ranking, burnt area, forest change, forest health and tree species, each scored against a public reference, with the result published whether it passed or failed.



Contents

■ Introduction	01
■ Why Forest Events Are Hard to Record	03
■ The Axine Method and Its Controls	05
■ Results and Controls	06
■ Burnt Area and Damage Degree	09
■ Forest Cover Change	11
■ Forest Health	13
■ Tree Species	15
■ Limits and What Is Not Evaluated	17
■ How the Record Is Used	18
■ About Axine Laboratories	19
■ References	20

Introduction



This paper describes five forest records that we build from the open Sentinel-1 and Sentinel-2 archive: a ranked stand list after a storm, a burnt-area map with a damage degree, a dated record of forest cover change, a forest health screen and a tree species layer. Each is scored against a public reference before it is offered, and each score is published here whether it passed or failed.

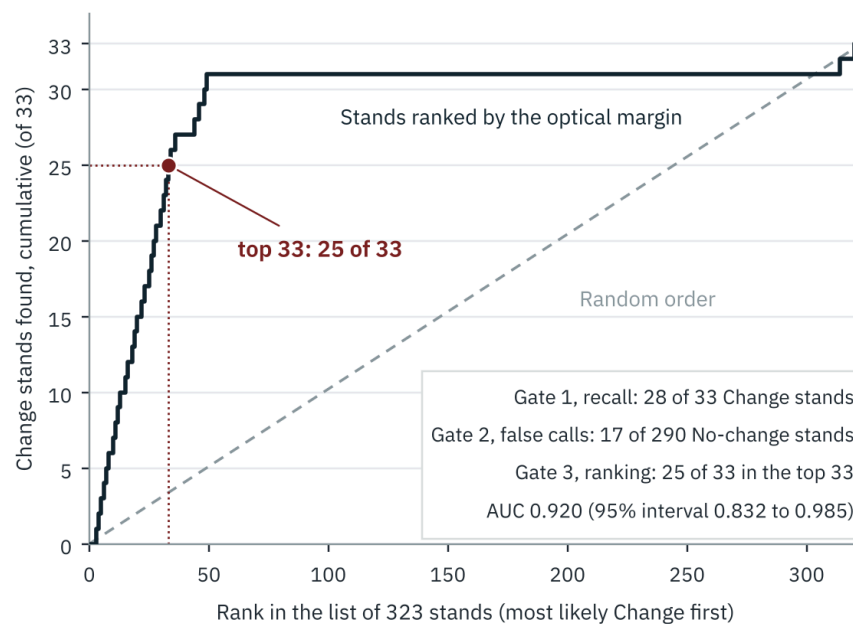


Figure 1. Storm Johannes, central Sweden, 27 and 28 December 2025: 323 stands ranked by the optical margin. The curve counts the 33 stands that an independent reader marked as changed, and 25 of them sit in the top 33 ranks.

Motivation: a forest event needs a dated record

A forest event produces a list of questions before it produces a report. Which stands came down, how much of the burnt ground was severely damaged, where did clearing start, which stands are declining, and what grows here. Foresters, loss adjusters and compliance officers ask these under time pressure, and the open satellite archive holds an answer to many of them. The difficulty is knowing how far an answer can be trusted.

The events are large. In 2025 the European Forest Fire Information System recorded 1,079,538 ha burnt in the EU27, the highest total since its records begin in 2006 and nearly double the 2006 to 2024 average¹. Global tree cover loss in 2025 was 25.5 million ha, with fires responsible for 42 percent of it, and tropical primary forest loss was 4.3 million ha, 36 percent below the 6.7 million ha of 2024².

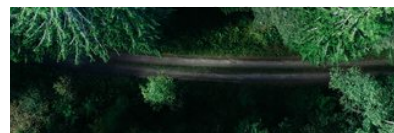
Regulation (EU) 2023/1115 asks operators to show that the commodities they place on the Union market were produced on land not deforested after 31 December 2020³. It applies to large and medium operators from 30 December 2026 and to most micro and small operators from 30 June 2027⁴. The Commission's May 2026 simplification review counts about 275,000 operators in scope, down from 352,000⁵. A dated record of forest cover change is one way an operator can show what a plot was and when it changed.

- + **Loss adjusters.** After a storm or a fire the first need is a ranked list of stands to visit, with the evidence behind each position.
- + **Forest managers.** One estate boundary has to answer a storm, a fire, a clearing, a beetle season and an inventory update within a single year.
- + **Compliance officers.** A filing needs a dated change record that a competent authority can check against a public reference.



Mixed forest crowns beside a dark lake with thin wisps of mist, seen from directly above.

Why Forest Events Are Hard to Record



The agents of forest disturbance differ, several of the available references are models themselves, and what a satellite can read depends on the season. Each of these shapes the controls in this paper.

46%

wind, share of reported forest damage in Europe, 1950 to 2019

24%

fire, share of reported forest damage in Europe, 1950 to 2019

17%

bark beetles, share of reported forest damage in Europe, 1950 to 2019

Many agents, different signals

The shares above are volumes of reported damage⁶, and each agent leaves a different trace in the imagery. Measured as area from satellite patches for 1986 to 2016, storm and fire each made up about 7 percent of all disturbances recorded in Europe, storm area grew over time and fire area showed no trend⁷. Drought adds a further agent: about 1.56 million ha of European forest was disturbed in 2018 and 4.74 million ha over 2018 to 2020⁸.

References that are models themselves

A control is only as good as its reference. The Copernicus Emergency Management Service activation we use for storm Johannes has one tier that a person read from sub-metre imagery and a second tier produced by a radar change-detection model. We score against the first and read the second only as an agreement check⁹. Sweden's open dead-spruce layer is produced with an AI model¹⁰, and the German canopy loss layer from DLR is a classification of Sentinel-2 and Landsat imagery¹¹. Each reference carries its own errors.

Season and sensor

Radar sees through cloud, which makes it attractive after a storm, but a winter event brings freeze, thaw and snow. Ruetschi and colleagues mapped windthrow in Switzerland and northern Germany with Sentinel-1 and reported a producer's accuracy of 0.88 for areal windthrow. They expect winter conditions to bring more obstacles to the method, and their paper contains no measured winter and summer comparison¹². Our own storm control is a winter event.

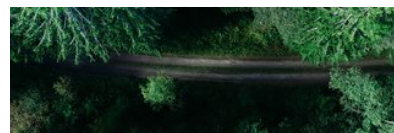
Stress runs ahead of loss

A vegetation index registers stress before a canopy opens. Fernández-Carrillo and colleagues validated Sentinel-2 bark beetle damage maps against field data in the Czech Republic and found accuracy above 0.95 and F1 above 0.88 for high-severity damage, with commission errors of 0.30 to 0.42 in the no-damage class¹³. A screen that flags stress and a reference that records realised loss will disagree for structural reasons, and we report that gap as a limit.



A fixed-wing aircraft releasing a white spray of water or retardant over forested hills under haze.

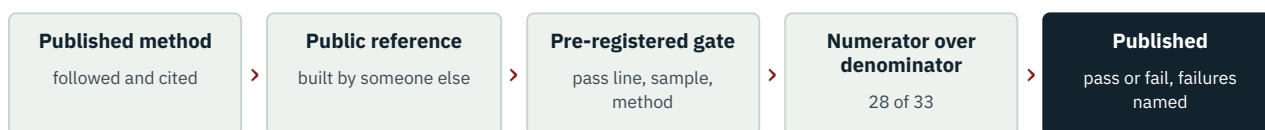
The Axine Method and Its Controls



Every record has the same shape going in and the same shape coming out. An area of interest, or a stand register cut from one, goes in. A ranked list, an area figure or a class map comes out, with before and after imagery that identifies every scene and a validation section that prints a numerator over a denominator from that job's own evidence.

Five rules for every control

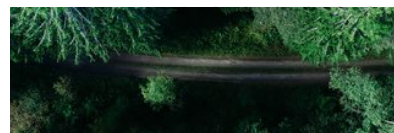
- + **A published method is followed and cited.** We implement it and name it.
- + **A public reference we did not build.** An agency layer, a benchmark archive or an independent reader of the imagery.
- + **The gate is written first.** Pass line, sample and method are fixed before any result is read.
- + **A result is a count.** We state 28 of 33 so the denominator is visible.
- + **A failed gate is published as failed.** Failures stay visible beside the passes.



The five records and their references

Record	Published method followed	Public reference
Storm and beetle stand records	Ruetschi et al. 2019; Dalponte et al. 2020	Copernicus EMS activation EMSN231, layer 4
Burnt area and damage degree	Key and Benson (FIREMON); Miller and Thode 2007; Parks et al. 2014	MTBS fires and Copernicus EMS grading
Forest cover change	Kennedy et al. 2010 (LandTrendr)	JRC Tropical Moist Forest and Hansen loss year
Forest health	Fernández-Carrillo et al. 2020; Baig et al. 2014	DLR Forest Canopy Cover Loss and the Skogsstyrelsen dead-spruce layer
Tree species	Rußwurm and Körner 2018; Blickensdörfer et al. 2024	TreeSatAI benchmark, official held-out split

Results and Controls



Two of the five controls carry a stated pass line, and both pass. The other three are published as measured, with the shortfalls named beside the numbers.

Record	Control	Result	Outcome
Storm and beetle stand records	Three gates on 323 stands	Recall 28 of 33; false calls 17 of 290; 25 of 33 in the top 33; AUC 0.920	Pass on all three gates
Burnt area and damage degree	Kappa at or above 0.45; AC1 at or above 0.60 on graded fires	Kappa 0.462 on 21 held-out fires; AC1 0.684 on the 12 graded fires	Pass on both gates
Forest cover change	F1 against TMF and Hansen, before and after the dense series	Against Hansen F1 rose in all three defined areas; against TMF it rose in one of three	Mixed; Ghana against TMF fell
Forest health	Coverage and commission at each lever, four areas	Timing lever raised coverage in all four; commission stays above 50 percent in three	Commission shortfall named
Tree species	Accuracy on 5,044 held-out patches	48.8 percent overall (2,462 of 5,044)	Below both reference points

The write-ups for the last three records give measured values and named shortfalls and no single pass line, so we do not label them pass or fail in this table. The outcome column reads each one against its own reference.

How to read the numbers

Cohen's kappa and Gwet's AC1 both score the agreement between two classifications after allowing for the agreement chance alone would produce. We report both because they respond differently when one class dominates. F1 combines precision and recall into one value between 0 and 1. AUC is the probability that a stand the reference calls Change is ranked above a stand it calls No change, so 0.5 is chance.

For the forest health records we report coverage, the share of dated reference loss by area that the strong-decline flag overlaps, and commission, the share of flagged area with no dated loss nearby. Where a number comes from a small sample we say so next to it. Each figure in this paper was redrawn from the numbers its own control produced.

Storm and beetle stand records

After a named storm, a snow load or a beetle season, can a dated per-stand record built from the open Sentinel-1 and Sentinel-2 archive put the stands that an independent reader calls changed at the top of the list?

Published method followed

Ruetschi, Small and Waser read windthrow from paired Sentinel-1 composites taken before and after an event, in both VV and VH polarisation¹². Dalponte and colleagues mapped the windthrow of storm Vaia in Italy from Sentinel-2 and PlanetScope and report detection accuracy above 80 percent for data acquired in optimal conditions¹⁴. A companion study of storm Vaia reads a Sentinel-2 NBR time series¹⁵, and further Sentinel-2 work on the same kind of disturbance favours moisture-sensitive indices of the NDWI and NDMI family over NDVI and NBR¹⁶.

Our method follows these findings. It uses phenology-matched optical pairs read by NDMI. A radar windthrow margin is read alongside and does not decide. Each stand is judged against its own quiet-year range, with no fixed cut.

Data

The event is storm Johannes, which crossed central Sweden on 27 and 28 December 2025. The reference is Copernicus EMS activation EMSN231, layer 4 (Confidence Assessment): 323 points that an independent reader of very high resolution imagery classified as Change (33) or No change (290). This is the one reference in the paper that is a person reading a sub-metre picture, with no model between the pixels and the label. A secondary tier, layer 3, holds 1,152 Damaged and Destroyed polygons from a radar change-detection model. We read it as an agreement check only, because it is itself a model output.

Controls

Three gates were pre-registered on the very high resolution tier before any stand was read. Figure 1 in the introduction plots the full ranked list behind the third gate.

Gate	What it counts	Result
Gate 1, recall	Change stands called above their own quiet range	28 of 33 (84.8%)
Gate 2, false calls	No-change stands called above range	17 of 290 (5.9%)
Gate 3, ranking	Change stands among the top 33 ranks (base rate 10.2%)	25 of 33 (75.8%)

All three gates pass. The AUC of the optical margin is 0.920, with a 95% stand-bootstrap interval of 0.832 to 0.985.

What failed, and what changed

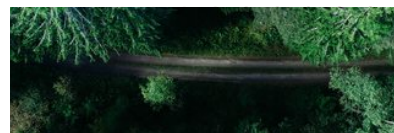
Radar did not decide a stand on this event. Its AUC was 0.390, below chance: on this winter storm the radar margin ranked the stands that the reference confirmed as damaged below the undamaged ones. That fits the expectation in the literature that winter conditions obstruct C-band windthrow detection. The shipped default lets radar decide nothing, and the radar reading is still printed on every record.

We looked at every false call. Seven of the 17 No-change stands called above range carry the EMS producer's own Disagreement tag, about twelve times the base rate of that tag among all No-change points. A share of the method's disagreement with the reference therefore sits on points the reference's two readers could not settle either. The other ten are not explained this way. After the pre-registered result was in hand we tried a variant that de-means each pair against the estate's own median shift before scoring. It recovers two more true positives at the cost of one more false positive. We report it beside the pre-registered result as a variant. It replaces nothing and it is not shipped.



A stack of cut round logs in a misty forest clearing, with tall conifers fading into fog behind.

Burnt Area and Damage Degree



Can burnt extent and a damage-degree class (unburnt, low, moderate, high) be read from Sentinel-2 at parity with a commercial forestry and insurance-claims service, and does the class agree with an independent grading product well enough to use as claims evidence?

Published method followed

Key and Benson's FIREMON methodology from the USDA Forest Service sets dNBR-based severity classes from a trainer reference table¹⁷. Miller and Thode introduced RdNBR and fitted both indices to the field Composite Burn Index in 741 plots on 14 fires, with R^2 of about 0.49 for dNBR and 0.61 for RdNBR¹⁸. Those ceilings are the reason we state damage degree as a spectral severity index and never as a percent-loss figure. Parks, Dillon and Miller introduced RBR to damp the dependence on biomass magnitude, and across 18 western US fires it fitted field measurements better (average R^2 0.786) than dNBR (0.761) or RdNBR (0.766)¹⁹. Our extent rule follows RBR and is unchanged by the second round. The references are MTBS, the US public burn severity record²⁰, and Copernicus EMS grading.

Data

The second round was trained on 19 US MTBS fires from 2020 to 2024, none shared with the original 30-fire extent campaign. The held-out set is 18 further US MTBS fires and 3 Copernicus EMS-graded European fires (EMSR592 Gironde/Landes in France, EMSR602 Castilla y León in Spain and EMSR624 Thassos in Greece), 21 fires in all, scored fire-weighted.

Controls

Fire-weighted agreement on the 21 held-out fires, with every pre-registered variant scored on the same sample, is shown in the left panel of Figure 2. The round one rule replicated fresh scores kappa 0.334 and AC1 0.525. The round two rule, with the RBR extent unchanged and a refit dNBR severity, scores 0.462 and 0.658.

Both pre-registered gates clear on the shipped variant: kappa at or above 0.45 (0.462) and, on the 12 of 21 fires that carry a graded severity reference, AC1 at or above 0.60 (0.684). The right panel ranks the 21 fires by per-fire kappa, so the average can be read beside where agreement is and is not strong. Per-fire kappa runs from 0.10 to 0.79: twelve of the 18 US fires score 0.53 or higher, four score 0.10 to 0.14, and the European fires score 0.32 (Gironde/Landes) and 0.21 (Thassos, Castilla y León).

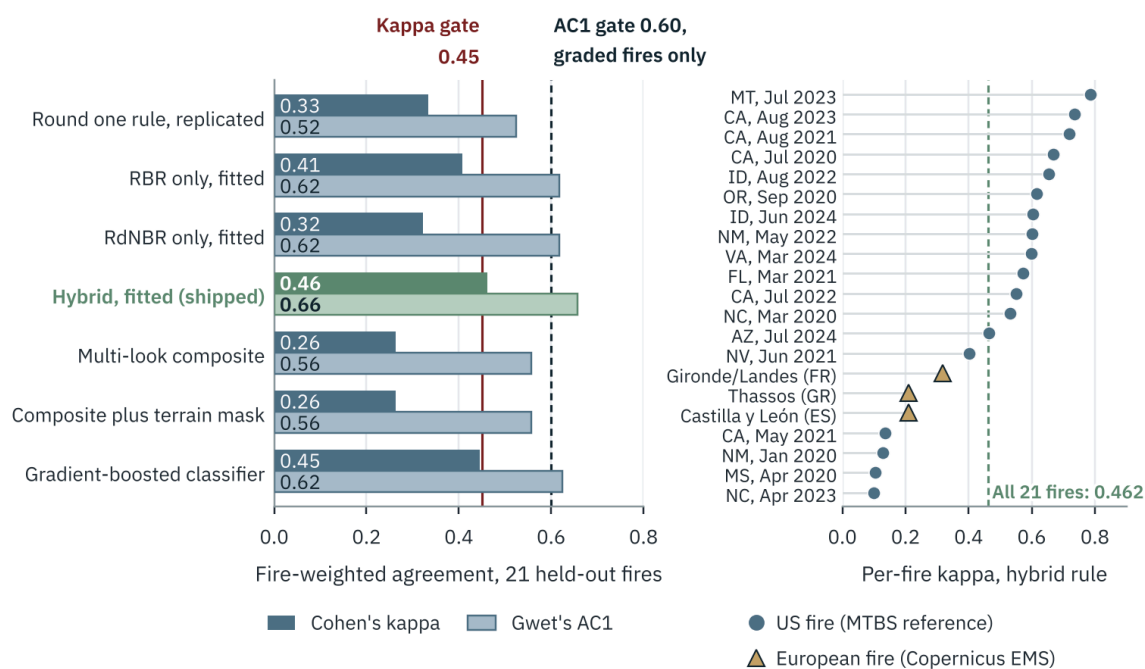
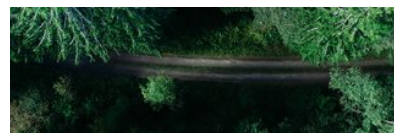


Figure 2. Left: fire-weighted Cohen's kappa and Gwet's AC1 for seven damage-degree variants on the 21 held-out fires, with the shipped hybrid in green. The AC1 gate of 0.60 applies to graded fires only. Right: per-fire kappa of the hybrid rule for the 18 US MTBS fires (circles) and 3 European fires (triangles), ranked, with the all-fire value of 0.462 dashed.

What failed, and what changed

Three tried variants did not beat the shipped hybrid on the held-out sample. A multi-look composite scored kappa 0.26 and AC1 0.56, below a plain nearest-scene pair, because averaging several looks damps the severity signal that a single sharp dNBR pair carries. A DEM-based terrain illumination mask left those numbers unchanged. The training sample held too little steep ground to exercise it, so the result applies to this sample only. A gradient-boosted classifier over several spectral and terrain features reached kappa 0.45 and AC1 0.62, ahead of round one and just short of the hybrid, so the simpler, auditable threshold change shipped in place of a packaged model.

Forest Cover Change



Does reading a dense time series across a monitoring window catch real forest loss that a single before and after pair misses, and does a trained classifier add real signal to the threshold rule it sits beside?

Published method followed

Kennedy, Yang and Cohen's LandTrendr decides whether a pixel changed from its trajectory across a time series and ignores the two endpoints²¹. We implement the idea at product scale as a dense semi-annual composite series with a short rolling-mean smoothing pass before the loss rule is applied. The RADD radar alerts use a confirmation window on the same principle before an alert counts²². The trained classifier follows the lineage of CNNs over multi-temporal composites²³.

Data

Two references that Axine did not build: JRC Tropical Moist Forest (TMF) dated loss²⁴ and Hansen Global Forest Change loss year²⁵. We chose Hansen because it does not inherit TMF's class scheme. Four areas were found by probing Hansen's own forest-cover and loss-year bands for real, substantial, dated clearing: Riau, Sumatra (137 forest pixels, kept as a sanity check and too small to carry a claim); its replacement Sumatra/Jambi (10,092 Hansen dated-loss pixels, fifty times the pre-registered floor); a Ghana cocoa frontier; and a Brazilian Amazon arc. Riau stays in the record because a second area in the same regime answers the question whether Sumatra is special.

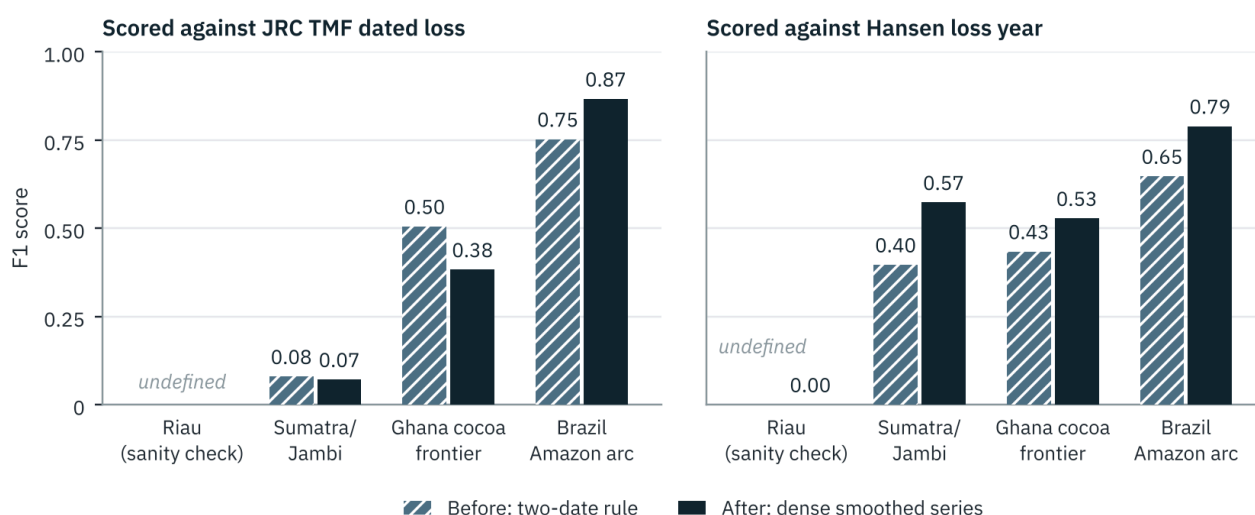


Figure 3. F1 of the two-date rule (hatched) and the dense smoothed series (dark) against JRC TMF dated loss and Hansen loss year in four areas. Riau is undefined against TMF both ways and before the fix against Hansen, and scores 0.00 after it.

Controls

F1 of the two-date rule (before) and the dense smoothed series (after), per area and per reference, is plotted in Figure 3.

Area	TMF, before	TMF, after	Hansen, before	Hansen, after
Riau (sanity check)	undefined	undefined	undefined	0.00
Sumatra/Jambi	0.08	0.07	0.40	0.57
Ghana cocoa frontier	0.50	0.38	0.43	0.53
Brazil Amazon arc	0.75	0.87	0.65	0.79

What failed, and what changed

The dense series raised recall against both references in the areas with a real amount of clearing. In Sumatra/Jambi recall went from 0.416 to 0.891 against TMF and from 0.374 to 0.934 against Hansen. In the Brazil arc it went from 0.618 to 0.874 and from 0.487 to 0.710. Precision fell with it. F1 rose against Hansen in all three defined areas and against TMF in Brazil, and it slipped against TMF in Sumatra/Jambi (0.08 to 0.07). The Brazil result is the clearest confirmation that a single pair misses mid-window felling that a trajectory-aware read catches.

Ghana is the area where the trade is mixed. Against TMF, recall rises from 0.687 to 0.925 and F1 falls from 0.50 to 0.38, while against Hansen F1 rises from 0.43 to 0.53. We tested two explanations. A diagnostic TMF-deforestation-only reference, which drops TMF's degraded-forest class, partly supports a reference-side explanation for the older two-date rule. It does not give the dense method its best precision on this area. Hansen does, and since Hansen does not inherit TMF's class scheme, Hansen is the reference we cite for Ghana's precision.

The trained classifier reads four regions (Riau, Rondônia, the Congo Basin and Tasmania), with Hansen loss unioned with RADD alerts (CC BY 4.0) as labels. Round three drew 100 samples per region, 234,507 labelled pixel rows. Leave-one-region-out AUC runs from 0.68 to 0.83 across the four regions, a real but geographically uneven cross-region signal. One shared probability cut cannot be both precise and high-recall across regions whose base rates differ this much: an earlier uncalibrated 0.5 cut gave recall from 0.2 percent to 94 percent by region. Round three replaced it with a pooled, pre-registered threshold selection (best recall at a minimum precision on the pooled held-out curve, with a documented fallback). The cut is now a measured choice. It does not make one shared cut the right operating point for every region.

Forest Health



Can a vegetation-index deviation from a baseline flag the forest area most likely to carry real stand stress, most pointedly early bark beetle decline, and cover an identifiable share of independently dated canopy loss while bounding how much flagged area has no nearby dated loss?

Published method followed

Fernández-Carrillo and colleagues validated multi-temporal, phenology-matched Sentinel-2 composites as a bark beetle stress signal against field data¹³. That is the basis for reading deviation from a same-season composite. The NDMI finding from the Vaia storm work, already adopted for the storm record, is reused here. Tasseled Cap Wetness coefficients follow Baig, Zhang, Shuai and Tong²⁶, applied to Sentinel-2 bands with matching spectral response. A Sentinel-2-specific derivation could not be retrieved from any open source, and we state the substitution here.

Data

Two open public damage layers in two countries. DLR's Forest Canopy Cover Loss layer for Germany is monthly at 10 m from September 2017 to October 2025, served by WMS, WCS and WFS with no fee under CC-BY-4.0¹¹. Skogsstyrelsen's AI-detected dead-spruce layer for Sweden is read from an open ArcGIS endpoint that needs no account¹⁰. Four areas were used: Bavaria and Harz/Oberharz in Germany, Västra Götaland and Kronoberg in Sweden. Truth samples hold 205, 1,668, 4,000 and 50 points. Kronoberg's 50 points are read with that caution throughout.

Controls

Coverage and commission were measured at every pre-registered lever in each area (Figure 4). The headline result is the first lever alone. Pinning the baseline to a window before the outbreak and adding a later-season read raised coverage in all four areas: from 32 to 74 percent in Bavaria, 55 to 91 percent in Harz/Oberharz, 9 to 43 percent in Västra Götaland and 0 to 50 percent in Kronoberg. Commission at that step went from 93 to 85 percent in Bavaria, 79 to 80 in Harz/Oberharz, 83 to 67 in Västra Götaland and 100 to 99 in Kronoberg.

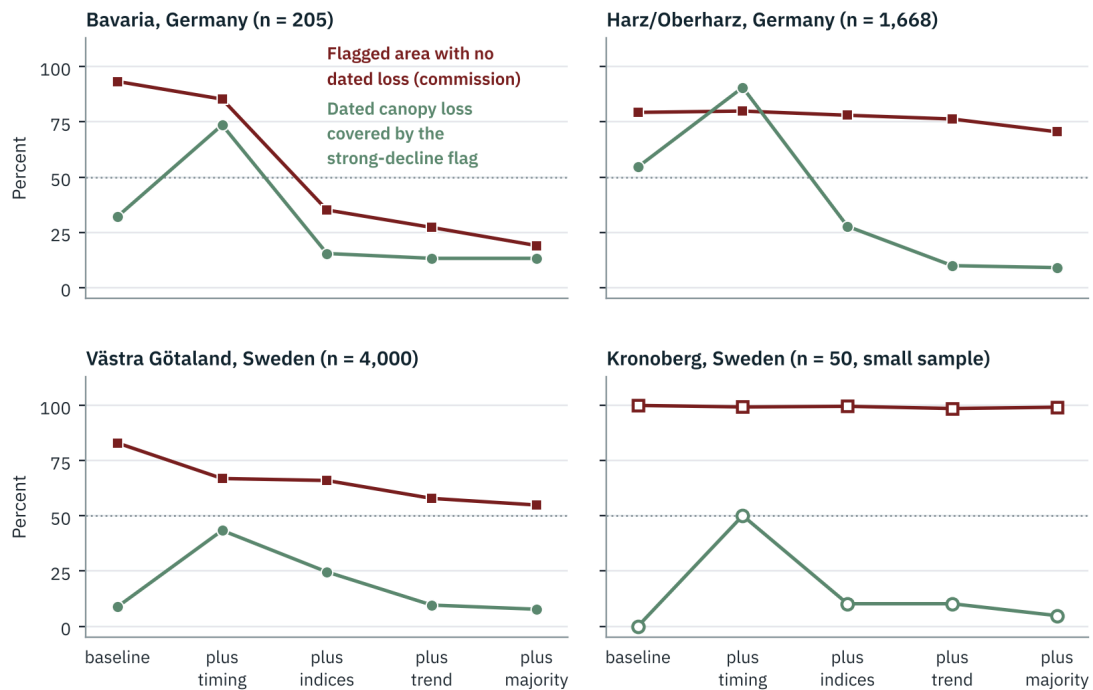


Figure 4. Share of dated canopy loss covered by the strong-decline flag (green circles) and share of flagged area with no dated loss nearby (red squares) at five steps in four areas. The dotted line marks 50 percent. Kronoberg uses hollow markers because its sample is 50 points.

What failed, and what changed

The timing fix is a clear win, replicated in four independent areas. It confirms the round-one finding that a baseline window contaminated by the early years of the outbreak was suppressing real signal. The other three levers are a wider index composite with a stricter agreement rule, a multi-year trend gate and scoring-time smoothing. Each gives back coverage. With all of them applied, coverage is 13.2, 9.0, 7.7 and 4.8 percent against baselines of 32.1, 54.6, 8.8 and 0.0 percent. That is below baseline in Bavaria and Harz/Oberharz and close to it in Västra Götaland. Commission falls to 19.0 percent in Bavaria and stands at 70.5, 54.9 and 99.2 percent in the other three areas. We ship the full set as the better of the two measured options, and we do not describe commission as solved: it stays above half in three of the four areas.

A coordinate-handling error in the Swedish queries was caught on an implausible all-zero first result for both Swedish areas, and fixed before any number was scored.

Tree Species



Can species composition be classified from Sentinel-2 at parity with a named architecture family, a spatial encoder feeding a temporal model, and how does the per-class accuracy compare with what an open benchmark has published for the same problem?

Published method followed

The TreeSatAI Benchmark Archive holds 50,381 patches of 60 by 60 m with aerial, Sentinel-2 and Sentinel-1 data, covers 15 genera, and fixes a 90/10 train and test split²⁷. It is both the proving ground of the architecture family and our validation reference. Its authors report that the sole use of Sentinel-2 still allows weighted precision of up to 74 percent. Our model gives each time step a spatial encoder and passes the sequence through a recurrent layer, classifying from the final hidden state, which follows Rußwurm and Körner²⁸ and the lineage cited in the forest change section. For Germany the national default route follows Blickensdörfer and colleagues, who mapped tree species from Sentinel-1 and Sentinel-2 time series and the German National Forest Inventory²⁹.

Data

We used the full official split: 45,337 training and 5,044 held-out test patches, read at the benchmark's own 15-class scheme, taken from the labelling rule in the benchmark's reference notebook. Fourteen of the classes are genera and one is a cleared-ground class. We separately confirmed that national default layers can be reached live for Germany (the Thünen Institute classified raster), Sweden (SLU Skogskarta per-species cover) and Finland (Luke's MS-NFI through the Paituli WCS).

Controls

The five best-represented classes score between 67 and 77 percent. *Pseudotsuga* and *Larix* score 43 and 21 percent. The eight classes at the right of Figure 5 score 10 percent or less, and five of them (*Alnus*, *Abies*, *Populus*, *Prunus* and *Tilia*) score 1 percent or less, with four at zero. Overall the model gets 2,462 of 5,044 patches right (48.8 percent). Two reference points sit on the same official split: the published ceiling for a tuned, Sentinel-2-only, non-temporal model of up to 74 percent weighted precision, and our own untuned single-label gradient-boosted proxy on the same split and features at 54.47 percent weighted precision. The proxy is a faithful lower bound that shows how much of the published number comes from the paper's own tuning. Weighted precision and overall accuracy are different metrics, and the figure labels them as such.

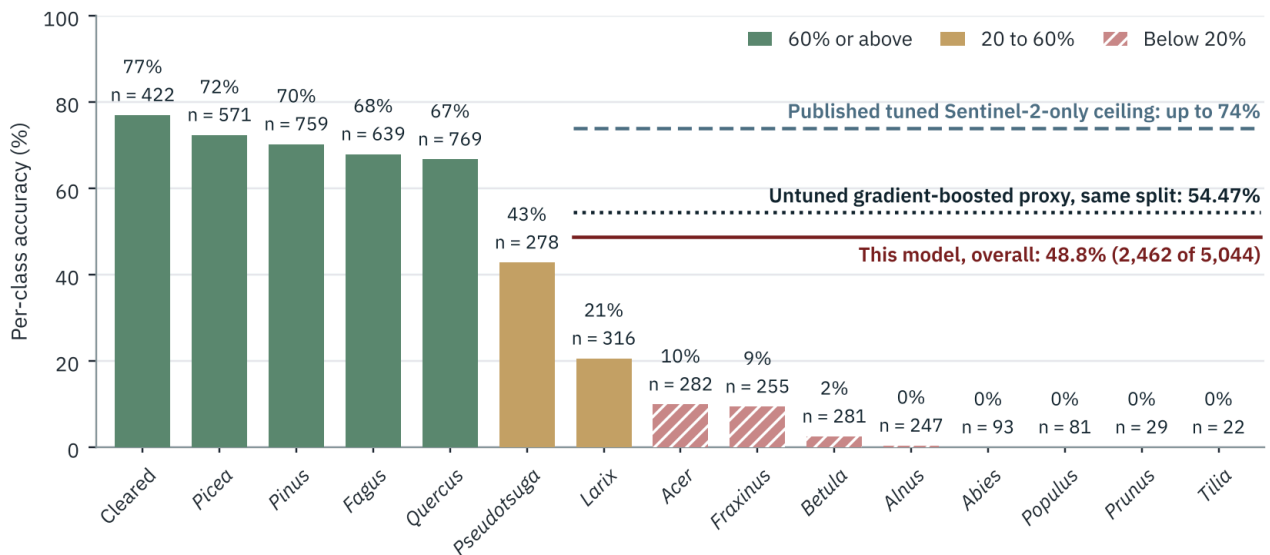


Figure 5. Per-class accuracy of the 15 classes on the 5,044 held-out patches, with the number of patches per class above each bar. The reference lines are the published Sentinel-2-only ceiling of up to 74 percent, an untuned gradient-boosted proxy at 54.47 percent weighted precision, and this model's overall accuracy of 48.8 percent.

What failed, and what changed

The first reproduction scored 30.4 percent against a published ceiling of 74 percent. Part of that gap was a measurement error. The run used a finer binomial species scheme than the benchmark's own evaluation classes, on a small subsample. We re-read the benchmark's reference notebook, found its 15-class genus labelling rule, and moved to the full official split. Under that fair comparison the same small model reaches 48.8 percent, a large gain that still sits below both the published ceiling and our untuned proxy. The reason is capacity. The model has under 10,000 parameters, uses no class weighting, and sees one summer mosaic per patch with no time series, so it cannot separate similar minority classes that it almost never meets in training.

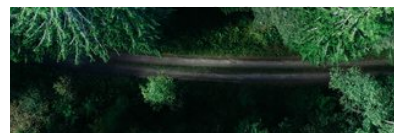
Limits and What Is Not Evaluated



Each record was tested on a stated sample. This section lists what that sample did not cover and what we have not run.

- + **Storm and beetle stand records.** The control covers one storm, one region, one forest type and one season: Swedish conifer forest and a winter event. No control exists yet for a beetle season, a summer storm, broadleaf forest or any region outside central Sweden. The reference is a point-level read of very high resolution imagery, with no pixel-exact outline. Passing this control is a precondition for offering the record more widely and does not settle that question.
- + **Burnt area and damage degree.** Terrain shadow from cloud is masked upstream and hillslope shadow is not. The terrain mask we tried was not tested on steep ground. The second round used Landsat throughout, for comparability with the MTBS and EMS references, so the Sentinel-2 compositing path that runs in production was not re-scored. Whether Sentinel-2's shorter revisit changes the composite result is untested. The boreal and Alaska regime was checked and not reached, because no accessible MTBS mosaic bundle for that region was available within the round. Damage degree remains a spectral severity index and never a percent-loss figure, whatever the agreement number.
- + **Forest cover change.** The trained classifier is verified against its own saved weights and is not yet part of the running service, so a live job today runs on the threshold rule alone. Round three is the latest committed result at the time of writing. A further round was running while this paper was written and is not reflected here.
- + **Forest health.** Both truth layers record realised canopy loss, and this product measures vegetation-index stress, which the cited literature says can run ahead of a visible gap. That is a structural reason for high commission, and two attempts to close it with index and threshold changes did not succeed. Kronoberg's 50-point sample is one noisy data point and is not weighted like the other three areas. The second fix round is the latest committed result, and a further round was running while this paper was written. The next lever on record is a structural or SAR corroboration pass on a different input.
- + **Tree species.** No published per-species Sentinel-2-only table exists to check Figure 5 against directly, because the benchmark's own per-species table is for a late-fusion model that uses aerial, radar and optical data together. Figure 5 therefore reads our per-class result against the published aggregate ceiling. The benchmark number is not what a client report prints: the client-trained route always prints that job's own held-out count from its own uploaded truth. We reason that closing the remaining gap is a capacity question (a deeper encoder, class-balanced training, several GPU hours), because epochs were not the limiting factor in this run. That reasoning has not yet been run on a GPU.

How the Record Is Used



All five records take the same inputs and give the same kind of evidence back. What a reader does with them depends on the desk.

For a claims desk, a storm record, a burnt-area damage-degree map and a forest health screen are three different questions that arrive in the same evidentiary form: which stands or pixels, over what dates, from which public archive, checked against which independent reference, with the scope of that check stated plainly. The scope is one event, one region, one season, named as such. A desk that has learned to read one of these records can read all five without relearning what evidence means.

For a forest manager running one estate through a year, a single area of interest or stand register follows the estate through whatever the calendar brings: a storm ranking after a windthrow event, a burnt-area class map after a fire, a change record for a compliance filing, a health screen ahead of a beetle season and a species layer for an inventory update. The estate boundary stays the same, and the open archive and the published method change with the event.



A narrow white-water river running through a steep gorge of dark conifers, seen from directly above.

For a compliance officer, the change record gives a dated reading from two independent public references, with the areas where they disagree stated. Ghana in this paper is an example: the two references give different F1 results for the same rule, and we report both.

A failure or a scope limit in one record is reported against that record's own control. It does not travel to the other four through the frame they share, and a reader can check each claim in this paper against the reference named beside it.

About Axine Laboratories

Axine Laboratories builds satellite spectral analytics for forestry, agriculture, water monitoring, energy and mining. The company is registered in the United States and run by a technical team of geologists and engineers.

Our approach rests on three commitments. Detection is physics-based and reproducible, derived from measured spectra and published formulations. Every published figure is backed by a pre-registered test against an independent reference, with the validation regime stated alongside it, and a result that fails its gate is published as failed. A plot or field the method cannot assess is reported as not evaluated, because a silent clear is the most expensive output a record can give.

Every record ships with the scenes, dates and parameters behind it, so a client can check any figure against the public archive it came from.

Axine Laboratories

axinelabs.com | contact@axinelabs.com

References



1. Joint Research Centre. *2025 was the EU's most destructive wildfire season on record*. JRC news item, 31 March 2026. https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/2025-was-eus-most-destructive-wildfire-season-record-2026-03-31_en
2. World Resources Institute, Global Forest Watch. Press release on 2025 tree cover loss data and tropical primary forest loss, 29 April 2026. <https://www.wri.org/news/release-tropical-rainforest-loss-drops-36-2025-fires-threaten-global-progress>
3. Regulation (EU) 2023/1115 of the European Parliament and of the Council on the making available on the Union market and the export from the Union of certain commodities and products associated with deforestation and forest degradation. Article 2.
4. European Commission. *Regulation on deforestation-free products*. Application dates. https://environment.ec.europa.eu/topics/forests/deforestation/regulation-deforestation-free-products_en
5. European Commission. *Report on the simplification review of Regulation (EU) 2023/1115*. COM(2026) 191, Brussels, 4 May 2026.
6. Patacca M, et al. Significant increase in natural disturbance impacts on European forests since 1950. *Global Change Biology* 29(5), 2023.
7. Senf C, Seidl R. Storm and fire disturbances in Europe: distribution and trends. *Global Change Biology*, 2021.
8. Senf C, Seidl R. Persistent impacts of the 2018 drought on forest disturbance regimes in Europe. *Biogeosciences* 18:5223, 2021. <https://bg.copernicus.org/articles/18/5223/2021/>
9. Copernicus Emergency Management Service. *EMSN231: Forest damage assessment after storm Johannes in central Sweden*. Risk and Recovery Mapping activation. <https://riskandrecovery.emergency.copernicus.eu/EMSN231>
10. Skogsstyrelsen (Swedish Forest Agency). *Geodata for forest damage*, including a map of clustered areas with dead spruce produced using an AI model. <https://skogsstyrelsen.se/en/digital-open-forest-data/geodata-for-forest-damage/>
11. German Aerospace Center (DLR). *Forest Canopy Cover Loss (FCCL), Germany, monthly, 10 m*. CC-BY-4.0. <https://geoservice.dlr.de/web/datasets/fccl>
12. Ruetschi M, Small D, Waser LT. Rapid detection of windthrows using Sentinel-1 C-band SAR data. *Remote Sensing* 11(2):115, 2019. doi:10.3390/rs11020115
13. Fernández-Carrillo A, Patocka Z, Dobrovolsky L, Franco-Nieto A, Revilla-Romero B. Monitoring bark beetle forest damage in Central Europe: a remote sensing approach validated with field data. *Remote Sensing* 12(21):3634, 2020. doi:10.3390/rs12213634
14. Dalponte M, Marzini S, Solano-Correa YT, Tonon G, Vescovo L, Gianelle D. Mapping forest windthrows using high spatial resolution multispectral satellite images. *International Journal of Applied Earth Observation and Geoinformation* 93:102206, 2020. doi:10.1016/j.jag.2020.102206
15. Sentinel-2 NBR time-series study of storm Vaia. *Forests* 12(6):680, 2021. doi:10.3390/f12060680
16. Sentinel-2 study of disturbance indices (NDWI and NDMI family against NDVI and NBR). *Remote Sensing* 13(8):1530, 2021. doi:10.3390/rs13081530

17. Key CH, Benson NC. Landscape Assessment (LA): sampling and analysis methods. In: FIREMON, USDA Forest Service General Technical Report RMRS-GTR-164-CD, 2006. <https://research.fs.usda.gov/treearch/24066>
18. Miller JD, Thode AE. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment* 109(1):66-80, 2007.
19. Parks SA, Dillon GK, Miller C. A new metric for quantifying burn severity: the Relative Burn Ratio. *Remote Sensing* 6(3):1827-1844, 2014. <https://www.mdpi.com/2072-4292/6/3/1827>
20. U.S. Geological Survey. *Monitoring Trends in Burn Severity (MTBS)*. <https://www.usgs.gov/centers/eros/science/monitoring-trends-burn-severity>
21. Kennedy RE, Yang Z, Cohen WB. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr, temporal segmentation algorithms. *Remote Sensing of Environment* 114(12):2897-2910, 2010. doi:10.1016/j.rse.2010.07.008
22. Reiche J, et al. Forest disturbance alerts for the Congo Basin using Sentinel-1. *Environmental Research Letters* 16:024005, 2021. doi:10.1088/1748-9326/abd0a8
23. Kussul N, Lavreniuk M, Skakun S, Shelestov A. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters* 14(5):778-782, 2017.
24. Vancutsem C, Achard F, Pekel J-F, et al. Long-term (1990-2019) monitoring of forest cover changes in the humid tropics. *Science Advances* 7(10):eabe1603, 2021. JRC Tropical Moist Forest product: <https://forobs.jrc.ec.europa.eu/TMF>
25. Hansen MC, Potapov PV, Moore R, et al. High-resolution global maps of 21st-century forest cover change. *Science* 342:850-853, 2013.
26. Baig MHA, Zhang L, Shuai T, Tong Q. Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. *Remote Sensing Letters* 5(5):423-431, 2014. doi:10.1080/2150704X.2014.915434
27. Ahlswede S, et al. TreeSatAI Benchmark Archive: a multi-sensor, multi-label dataset for tree species classification in remote sensing. *Earth System Science Data* 15:681-695, 2023. doi:10.5194/essd-15-681-2023
28. Rußwurm M, Körner M. Multi-temporal land cover classification with sequential recurrent encoders. *ISPRS International Journal of Geo-Information* 7(4):129, 2018.
29. Blickensdörfer L, et al. National tree species mapping using Sentinel-1/2 time series and German National Forest Inventory data. *Remote Sensing of Environment* 304:114069, 2024.



FOREST EVENT RECORDS