Supplement of Biogeosciences, 23, 1089–1101, 2026
https://doi.org/10.5194/bg-23-1089-2026-supplement

Biogeosciences
Open Access

Supplement of

Northward shift of boreal tree cover confirmed by satellite record

Min Feng et al.

Correspondence to: Min Feng (mfeng@terrapulse.com) and Joseph O. Sexton (sexton@terrapulse.com)

Contents

S1. Study Area

The study region (Fig. S1) was defined by all boreal forest or taiga ecoregions, as well as a selection of immediately adjacent temperate conifer forests and tundra regions predominantly surrounded by boreal forest or taiga (Dinerstein et al., 2017).

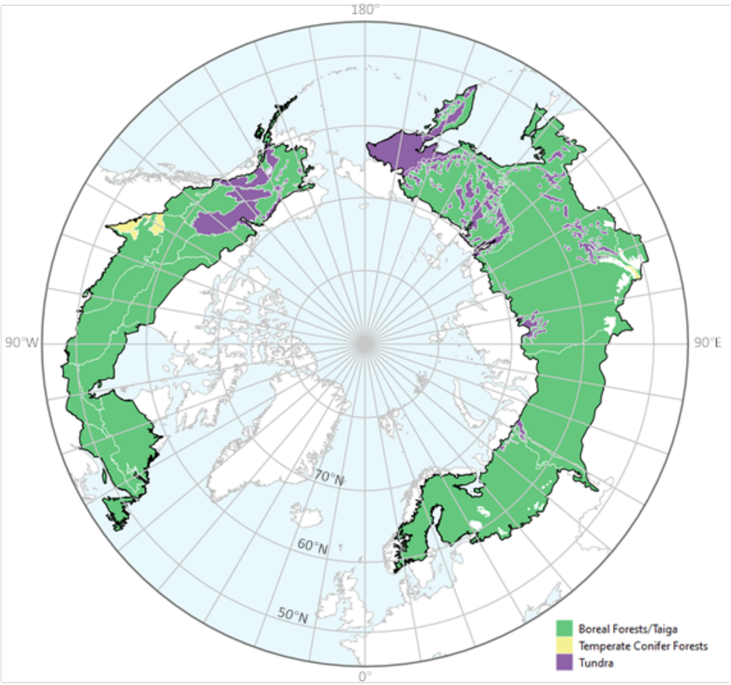


Fig. S1. Study area extent and inclusion of boreal forests/taiga with a selection of immediately adjacent temperate conifer forests and tundra from Dinerstein et al. (2017).

S2. Boreal calibration of MODIS VCF tree-cover input data

Downscaling and extension of the high-resolution, 36-year historical time series of tree cover estimates was based on the MODIS Vegetation Continuous Fields (VCF) Collection 6 Tree Canopy product (MOD44B; Carroll et al., 2011), following calibration to the boreal region (Fig. S2). The 250-m MODIS VCF data were acquired from NASA's Land Processes Distributed Active Archive Center (LP DAAC) (https://lpdaac.usgs.gov/tools/data-pool/) and masked using the MODIS Cropland Probability layer (Pittman et al., 2010) (https://glad.umd.edu/dataset/gce/modis-global-crop-extent-discrete-croplandnot-cropland-data) to exclude agricultural lands.

To improve characterization of boreal tree cover, the original MODIS tree cover estimates were calibrated to a region-wide sample of airborne LiDAR-based reference measurements (Montesano et al., 2016). Calibration stratified by topographic and climatic gradients, including elevation, slope, and aspect derived from ASTER GDEM v3 (Abrams et al., 2020) and bioclimatic variables from WorldClim v2 (Fick and Hijmans, 2017). Calibration models were trained using the Cubist regression tree algorithm, with reference data partitioned into independent training and testing sets via stratified random sampling across the range of observed tree cover values. Multiple Cubist models were fit to the training sample and subjected to iterative variable selection to reduce overfitting. At each iteration, the lowest-ranked predictor was removed and the model rebuilt. After 23 iterations, the model achieving the highest coefficient of determination (R^2) on the test sample was selected for application.

The calibrated MODIS VCF estimates were then rescaled to 30-m resolution and extended to the full Landsat archive (1984-2020), producing an annual, high-resolution record of tree cover suitable for biome-wide change detection.

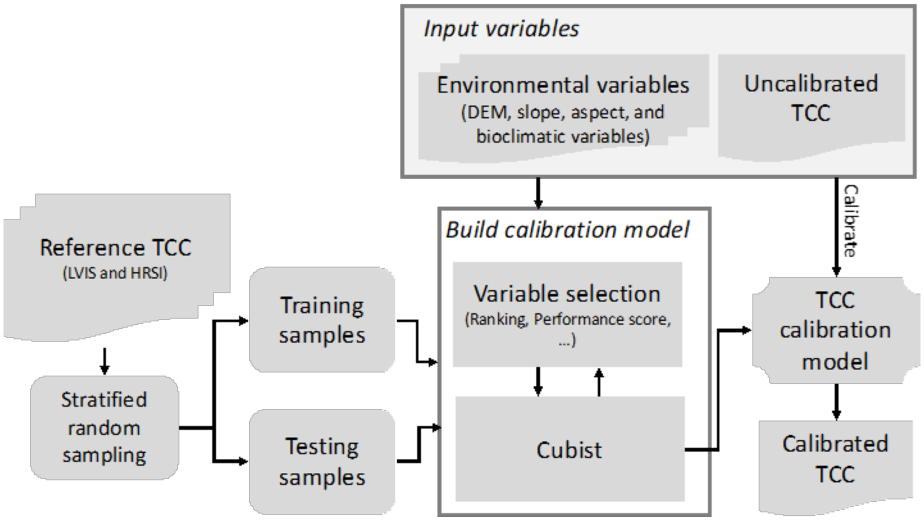


Fig. S2. Regression-tree calibration of tree cover (TCC) for the boreal ecoregion.

S3. Reference measurements of tree cover

Calibration and posterior validation of tree-cover estimates were based on two sources of independent reference data (Fig. S3). The first source comprised waveform LiDAR measurements collected in 2017 across Alaska and western Canada by NASA's Land, Vegetation, and Ice Sensor Facility (LVISF; Blair et al., 1999). For each 10-m LVISF footprint, tree cover was defined as the fraction of returned energy originating above a 1.37-m height threshold—a convention commonly used to distinguish trees from shrubs and ground vegetation. These footprint-level estimates were aggregated to 30-m resolution by averaging all returns whose centroids fell within each 30-m grid cell. The resulting gridded dataset (Montesano et al., 2021) provided a continuous, structure-based reference for calibrating MODIS VCF estimates across the full gradient of boreal vegetation density. The second reference source consisted of 425 visually interpreted observations of tree cover derived from very high-resolution spaceborne imagery, primarily from QuickBird (~0.6 m resolution), in Google Earth. Interpretation focused on identifying individual tree crowns from pan-sharpened color-infrared images, acquired circa 2008, and distributed across the northern boreal zones of North America and Eurasia (Montesano et al., 2009; 2016; 2020). To reduce spatial misregistration errors, interpretation was restricted to homogeneous 500 × 500 m scenes where tree cover could be confidently assessed.

The LVIS canopy cover reference was evaluated relative to NASA G-LiHT airborne LiDAR, which was assumed as reference. Montesano et al. (2023) reported agreement between LVIS and G-LiHT canopy heights with

R^2 values up to 0.87 and RMSE values in the 1–2 m range. No comparable independent validation exists for the high-resolution optical imagery interpretations; these are based on expert identification of crowns in QuickBird scenes, which have been used extensively in boreal validation but without published quantitative their own error estimates.

S4. Validation metrics

Accuracy of tree-cover and other continuous estimates was quantified by Mean Bias Error (MBE), and their precision was quantified by Root-Mean-Squared Error (RMSE) (Willmott, 1982):

$$MBE = \sum_{i=1}^{n} \frac{L_i - M_i}{n} \qquad \text{(S1)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (L_i - M_i)^2}{n}} \qquad \text{(S2)}$$

where L_i and M_i are values from the Landsat-based model and the reference data, respectively, at a sample location *i,* and *n* is the count of joint observations in the sample. After modeling the relationship between *L* and *M*, the (squared) difference between *L* and *M* were disaggregated into systematic error (MSE_S) and unsystematic error (MSE_U):

$$MSE_S = \sum_{i=1}^{n} \frac{(\widehat{L_i} - M_i)^2}{n} \qquad \text{(S3)}$$

$$MSE_U = \sum_{i=1}^{n} \frac{(L_i - \widehat{L_i})^2}{n} \qquad \text{(S4)}$$

where $\widehat{L_i}$ is a cover, age, or year value predicted by the modeled relationship ($Y = a + b\,X$) between *L* and *M*. Ordinary least squares (OLS) regression was applied to fit the parameters *a* (*intercept*), *b* (*slope*), and R^2 (Sokal and Rohlf, 1994).

Reference datasets provide spatial coverage but not temporal continuity; therefore, stability (bias change through time) cannot be quantified here. We note this explicitly and highlight the need for sustained reference time series in future validation efforts. Calibration and validation were conducted using stratified random partitions of reference data drawn across ecological and topographic gradients, with independent test samples withheld at each stratum to guard against overfitting. This design reduces—but does not eliminate—the possibility of unincorporated variance at ecotonal boundaries. A full “leave-tile-out” validation would require exclusion of entire LVIS flightlines and/or high-resolution imagery tiles during calibration and subsequent reprocessing of the Landsat time series across the boreal biome. Such an analysis was beyond the scope of the present study and not feasible within the short revision period, but we identify it as an important avenue for future refinement.

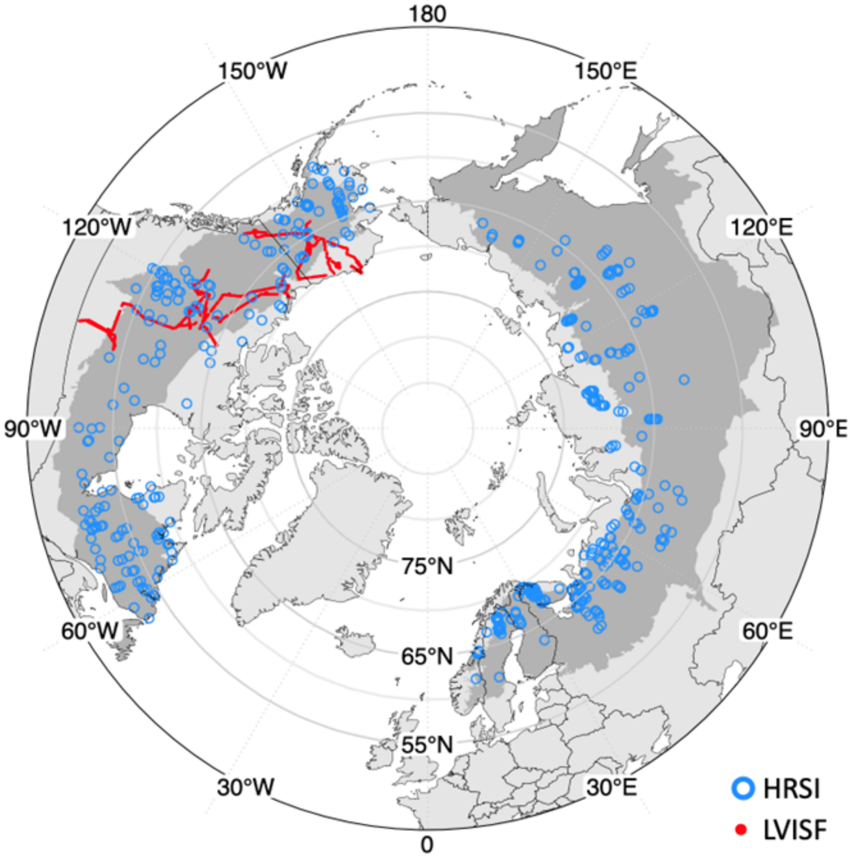


Fig. S3. Distribution of high-resolution satellite images (HRSI) and 2017 Land and Vegetation Imaging Sensor "Facility" version (LVISF) derived measurements for calibration and validation of boreal tree cover estimates.

S5. Historical retrieval of tree cover

Calibrated MODIS VCF tree cover estimates were downscaled and extended to Landsat spatial resolution and temporal extent through a nonparametric, machine learning regression-tree model f (Sexton et al., 2013):

$$\hat{c_i} = f(X;\hat{\beta}) + \varepsilon \qquad \text{(S5)}$$

where $\hat{c_i}$ is the percentage of a pixel (i)'s area covered by woody vegetation taller than 3 to 5 meters; $\hat{\beta}$ is a set of empirically estimated parameters; ε is residual error or uncertainty; X is a set of Landsat measurements of surface reflectance, derived indices, image acquisition date, and sensor identification (Fig. S4). The model was fit to spatiotemporally coincident values of calibrated MODIS VCF as response and Landsat images as covariates and then applied to each complete Landsat image to produce the map of estimates.

Model parameters were fit in 3×3 moving windows of WRS-2 tiles by a gradient-boosted regression tree (Dorogush et al., 2018) and applied to the center tile of the window to map tree-cover estimates across the Landsat images in the center tile. Each model's training sample was pooled from 2000 to 2019 to minimize overfitting to inter-annual noise, thus conservatively incorporating phenological and atmospheric variability into ε. Each WRS-2 tile's fitted model was applied to all Landsat images from 1984 to 2020 within the tile to retrieve a time series of tree-cover estimates at 30-m spatial resolution. The median estimate of cover and its uncertainty within each year was reported as the estimate (c, ε) for that pixel in that year. In addition to minimizing inter-annual noise, this compositing filled gaps due to clouds, snow, and cloud shadows.

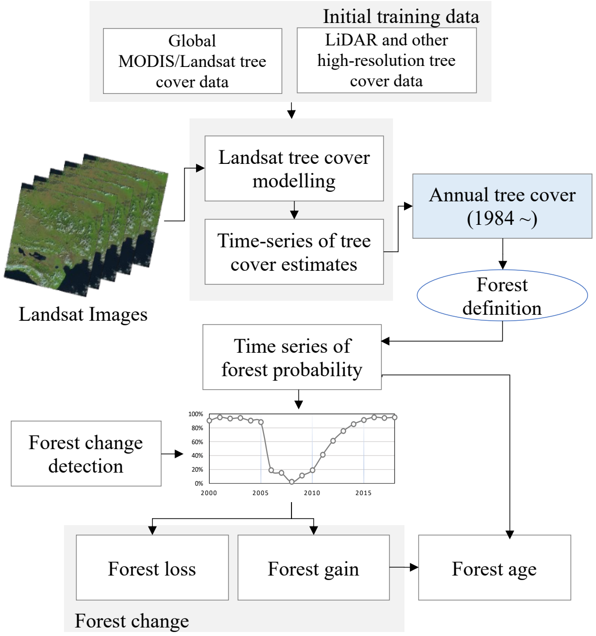


Fig. S4. Process for estimating tree cover (TCC) and forest probability, change, and age.

S6. Landsat time series

Tree cover (TC) estimation was based on the Landsat Collection 1, Level-1 Terrain Corrected (L1T) archive covering the period from 1984 to 2020, including images from the Landsat 4 and 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI) sensors. All images were downloaded from the USGS Earth Resources Observation and Science (EROS) Center (http://landsat.usgs.gov). Each image was converted to units of surface reflectance; the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Schmidt et al., 2013) was used for TM and ETM+ images, and the Landsat Surface Reflectance (LaSRC) (Vermote et al., 2018) was used for OLI images. Clouds and their shadows were removed following Zhu and Woodcock (2012).

Leveraging the high degree of image overlap in the high latitudes, a total of 2,189 World Reference System 2 (WRS-2) tiles was selected to cover the region. A maximum of four Landsat images within the growing season of each year and WRS-2 tile were chosen to avoid errors from clouds, snow, and phenological variation. All images were scored by cloud coverage, seasonality, and image quality flags (e.g., SLC-off, Landsat collection 1 processing levels), and images with the highest scores in each year were selected for analysis:

$$\text{score} = ((1 - c) * (1 - ws) + (s * ws)) * wq, \qquad \text{(S6)}$$

where c is the ratio of cloudiness in an image (0 = clear, 1 = fully cloudy); s represents the seasonality of the image calculated as the number of days of an image acquisition to the mid-summer day:

$$s = \cos(\mathrm{abs}(d - ds) * 2 / 366), \qquad (S7)$$

where d is the Julian day of the acquisition and ds is the value of Julian day of mid-summer; ws is a seasonality weight (higher in high latitudes and lower in low latitudes):

$$ws = \sin(\mathrm{latitude}), \qquad (S8)$$

and wq is an image quality weight, which is 0.1 for Landsat 7 ETM+ SLC-off images collected after May 31, 2003, when the Scan Line Corrector (SLC) failed, and 1.0 for all other images.

Images with the highest composite scores were retained, allowing a maximum of 148 images per tile across the study period. In total, 224,026 images were selected, including 110,407 TM images, 59,791 ETM+ images, and 53,828 OLI images (Fig. S5). Image density varied spatially due to historical archive limitations (Wulder et al., 2016). Of the 2,189 WRS-2 tiles, 72.6% had at least 90 images available. Persistent data gaps—particularly in central and eastern Russia—reflect limited ground-based reception capabilities during the 1980s and 1990s for Landsat 4 and 5 (Fig. S6).

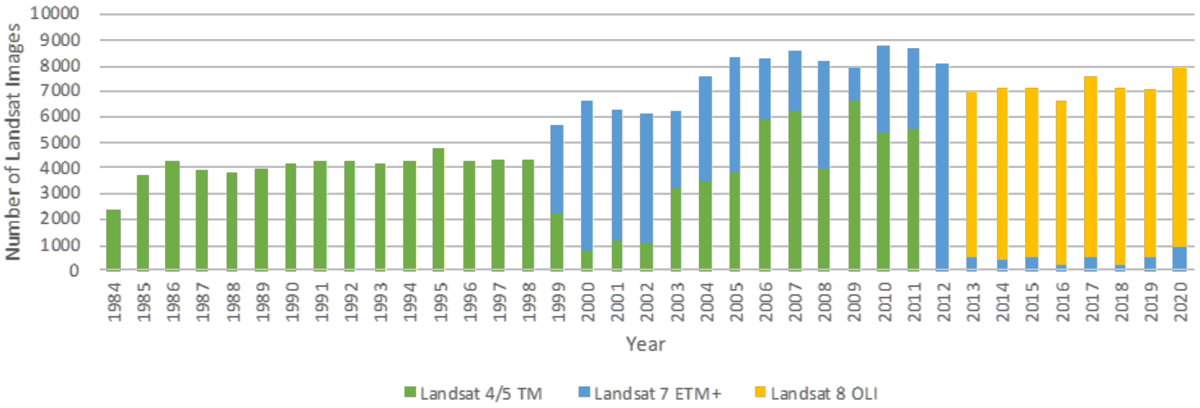


Fig. S5. Number of Landsat TM, ETM+, and OLI images selected in each year.

This biome-specific calibration and rescaling improved characterization of the complex gradient of boreal tree-cover across the region (Fig. S7). Calibration increased accuracy, decreased uncertainty, and improved the linear correlation of tree-cover estimates to reference measurements (Fig. S8). MAE decreased to 11.13%, RMSE decreased to 16.44%, and the coefficient of determination (R^2) of the linear model between estimated and measured data increased to 0.60. The residual bias of the rescaled Landsat-based estimates relative to the LiDAR reference was slight (~2%). All subsequent analyses, including region-wide summaries, change detection, and forest-age estimation, were based on the calibrated, rescaled dataset.

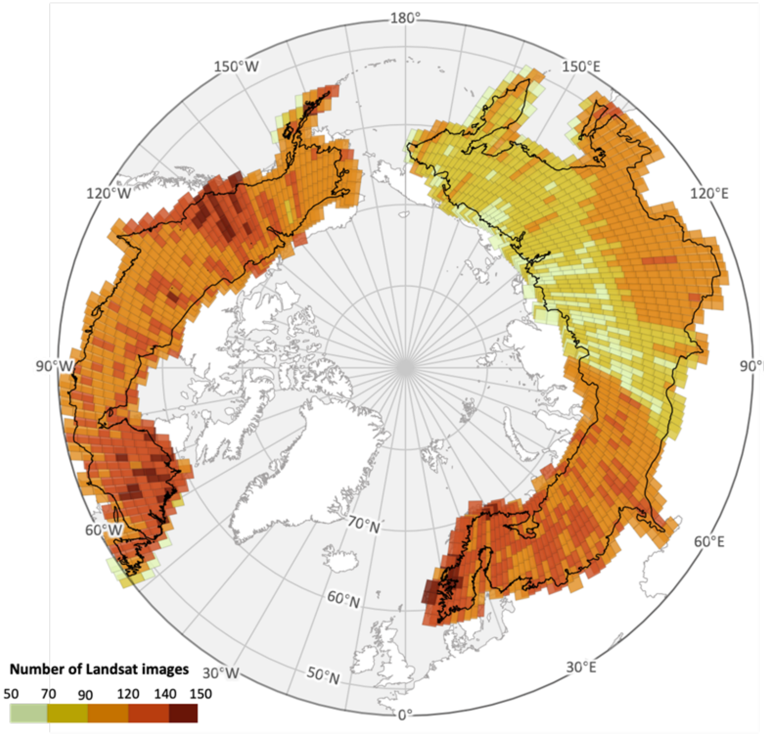
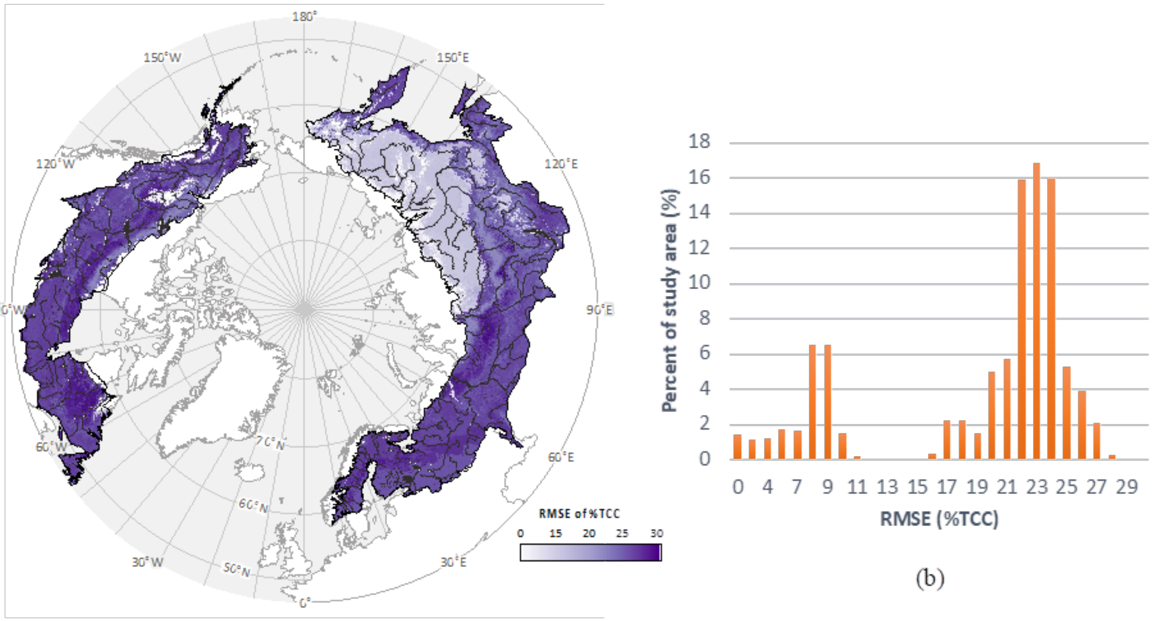


Fig. S6. Sampling density of Landsat images across the study region from 1984 to 2020.



(a)

(b)

Fig. S7. Spatial distribution (a) and histogram (b) of RMSE across the boreal region in the Landsat-derived tree cover for 2020.

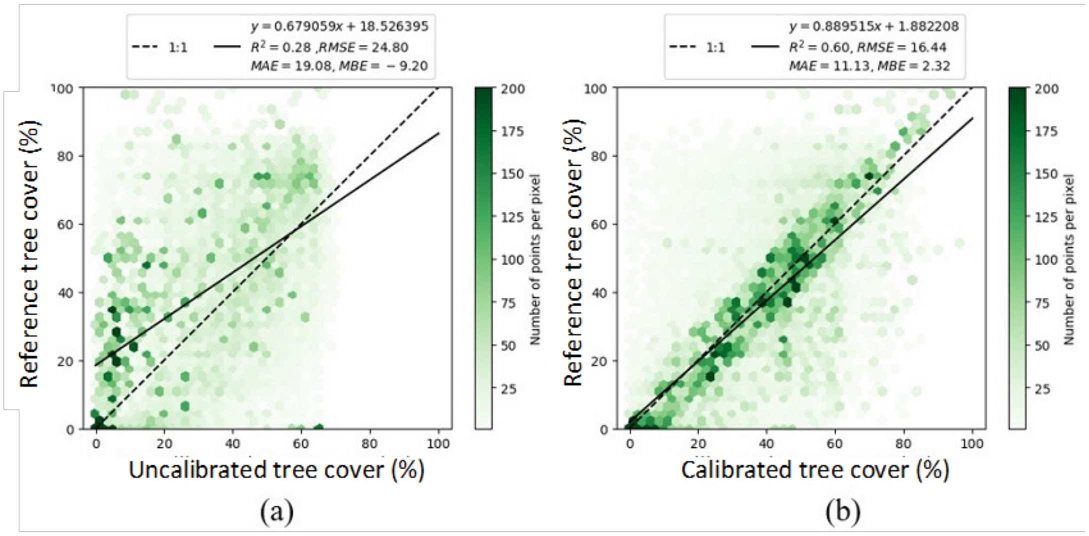


Fig. S8. Effect of calibration on estimates of tree cover. Scatterplots show the joint distribution of uncalibrated (a) versus calibrated (b) tree-cover estimates relative to LiDAR-based reference measurements.

S7. Example retrievals of tree-cover disturbance and regrowth

Three examples of forest disturbance and regrowth illustrate the wide variability in post-disturbance recovery trajectories across the boreal biome (Fig. S9). In Two Creeks, Alberta, Canada (54°21′20.2″N, 116°20′03.8″W), extensive timber harvesting in 1988 removed mature forest stands. Tree cover declined abruptly from dense pre-disturbance values to below 20% in 1988, followed by steady regrowth to approximately 80% by 2020. In Gorod Ivdel', Sverdlovsk Oblast, Russia (60°34′47.6″N, 61°56′54.5″E), clear-cut logging occurred in 1986 and again in 2011. Tree cover declined sharply to near zero in 1987, recovered to over 60% by 2010, and then declined again to around 20% following the second harvest. By 2020, partial regrowth had restored tree cover to over 40%. In northern Saskatchewan, Canada (57°44′21.7″N, 104°34′05.4″W), a wildfire in 2009 interrupted a trend of gradual tree-canopy growth. Tree cover had increased from the mid-1980s to approximately 40% by 2008 before dropping below 20% in 2010. A slow trajectory of recovery followed in the ensuing decade.

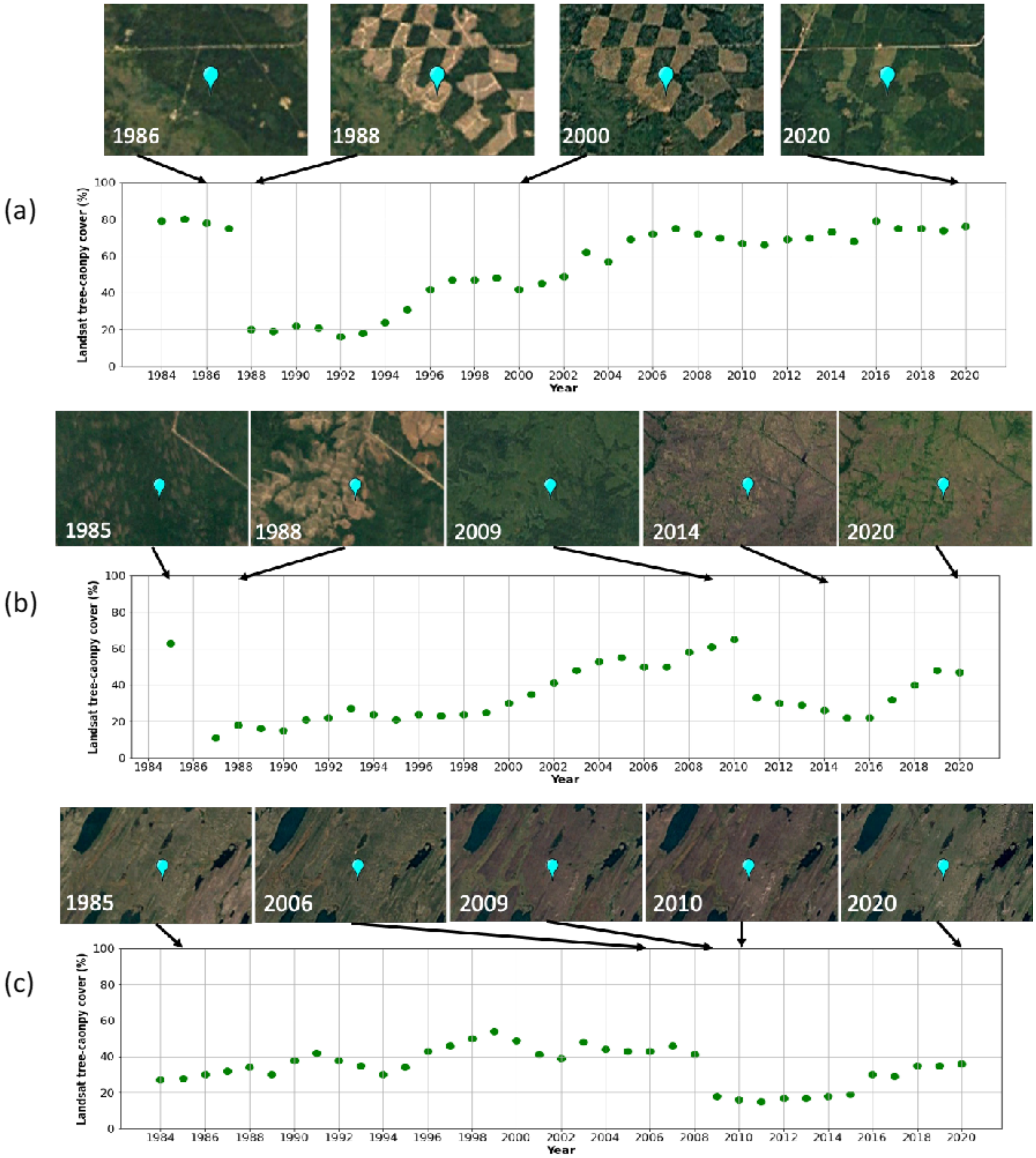


Fig. S9. Historical retrieval of tree cover collected at three locations across the boreal biome. (a): the forests of Two Creeks, Alberta, Canada (54°21'20.2"N, 116°20'03.8"W) recorded an anthropogenic disturbance for timber harvest in 1988 and subsequent recovery over 30 years; (b) the forests in Gorod Ivdel', Sverdlovsk Oblast, Russia (60°34'47.6"N 61°56'54.5"E) records two timber harvests in 1987 and 2010 and subsequent recoveries. (c): the sparse forests in northern Saskatchewan, Canada (57°44'21.7"N 104°34'05.4"W) recorded a slow increase of tree cover until a fire disturbance in 2009 and subsequent recovery afterward. The annual tree cover records are presented in the bottom panel for each location. The true-color Landsat images are presented in the top panel to show the location (in blue balloon symbol) and historical landscapes in representative years around the location.

S8. Trend analysis of tree cover

The rate (slope), coefficient of determination (R^2), and significance (p-value) of tree-cover change over time were estimated for each pixel using ordinary least squares (OLS) regression. Annual estimates of tree cover were summarized as pan-boreal means and medians to assess overall trends across the entire study area over the 36-year period (Fig. 2). To examine spatial variability in these trends, annual mean and median tree cover values were also aggregated by degree of latitude for the circumpolar boreal domain (47°N to 70°N), and separately for North America and Eurasia (Fig. S10). This enabled quantification of latitudinal patterns in tree-cover changes at both global and continental scales. Pixels with 30 or fewer valid annual observations were excluded from the trend analysis to avoid geographic bias due to gaps in Landsat image availability, particularly in central and northeastern Siberia. In addition, data from 1984 were excluded because of incomplete spatial coverage during the first operational year of Landsat 5 (Wulder et al., 2016).

Most of the boreal region exhibited minimal change in tree cover, with approximately 70% of the area showing trends smaller than ±0.5% per year. Tree-cover increases exceeding 0.5% per year occurred in 19.75% of the region, while decreases greater than 0.5% per year were observed in 9.74% of the region (Fig. S11).

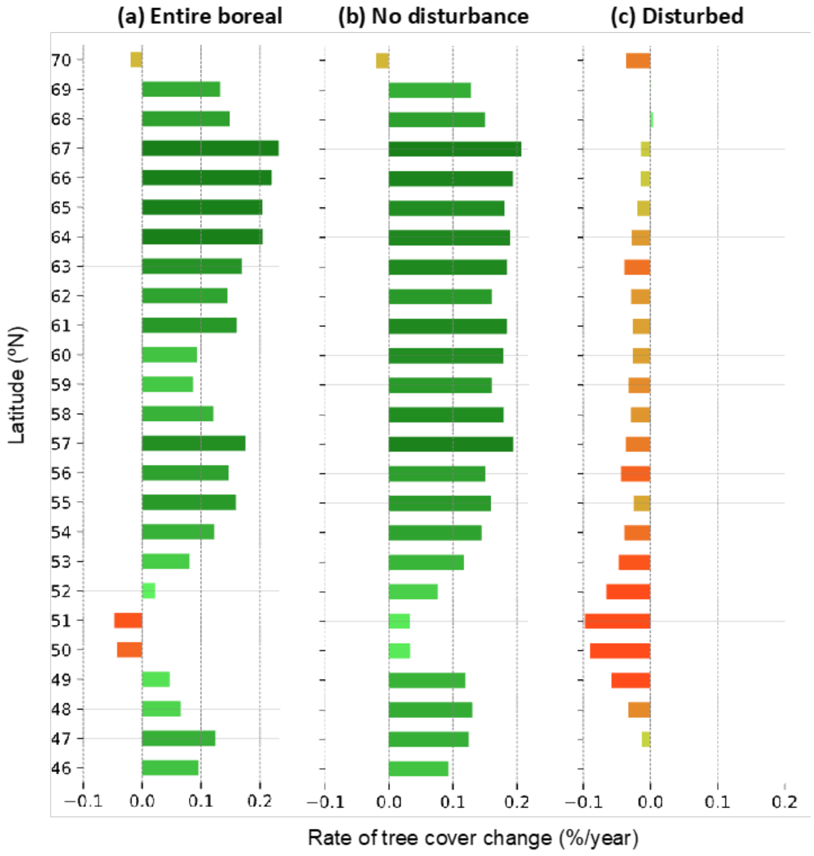


Fig. S10. Linear regression slope of latitude-stratified tree cover change over time for the entire boreal region (a) as well as for the region where no disturbance (b) or disturbance (c) was detected during the study period (1984-2020) separately.

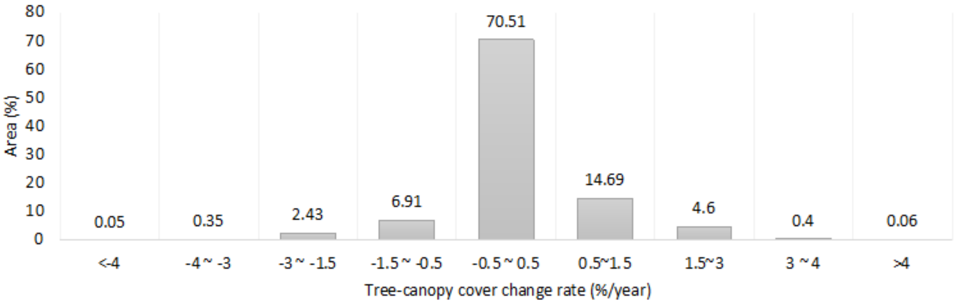


Fig. S11. Frequency distribution of local rates of tree cover change from 1984 to 2020.

Beginning in 1999 with the addition of Landsat 7 ETM+ to the Landsat fleet, the USGS adopted a comprehensive global acquisition strategy that greatly increased sampling intensity across the boreal region (Wulder et al., 2016). To test the possible effect of the increased data availability, trend analyses were conducted on pre- and post-1999 subsets of the time series, as well as the entire series from 1985 to 2020. Although significance decreased due to shortening of the temporal span, trends of all pan-boreal regressions for both mean and median tree cover in both pre- and post-1999 periods remained positive and significant at $p < 0.05$ (Fig. S12).

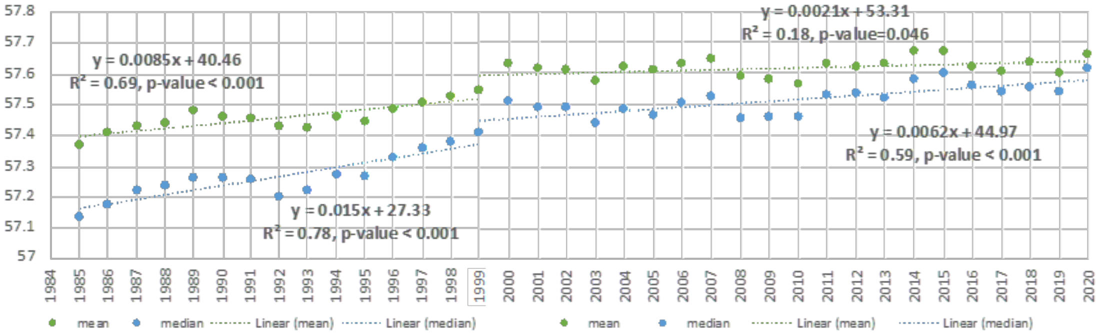


Fig. S12. Northward shift of mean and median latitudes of boreal tree cover in periods pre- and post-1999.

To evaluate the potential influence of forest disturbance on long-term tree-cover trends, we repeated the trend analysis (Fig. 2) using subsets of disturbed and undisturbed forest. The total study area spanned 17,694,070.5 km², comprising 575,461,722,221 valid pixel-year observations from 1984 to 2020. Within this subset, 1,746,167,768 pixel-years (0.303%) were classified as disturbed, corresponding to 1,576,707 km²—approximately 10% of the study area. Using the forest disturbance map (Fig. 3) as a spatial mask, the full pixel population (N = 575,461,722,221) was partitioned into disturbed (N = 57,819,280,985) and undisturbed (N = 517,642,441,236) subsets.

Trend analysis revealed a slight but confirmatory effect of disturbance: tree cover in undisturbed areas exhibited lower interannual variability than in disturbed areas but closely tracked pan-boreal trends across nearly all latitudes. The exception occurred between 58° and 61°N, where total and undisturbed tree-cover trends diverged. This difference was attributable to extensive wildfires in Siberia during the observation period and is corroborated by regional studies (Kukavskaya et al., 2016; Ponomarev et al., 2021).

S9. Forest categorization, change detection, and estimation of forest age

To estimate the timing of forest disturbance and establishment, we defined "forest" as a pixel-level condition where tree cover c exceeds a predefined threshold $c^* = 30\%$, following Sexton et al. (2015). Accordingly, the probability of a pixel being forested, p(F), is the probability that $c > c^*$, given the estimated tree cover distribution:

$$p(F) \stackrel{\text{def}}{=} p(c > c^*) = \int_{c^*}^{100} p(c)dc. \tag{S9}$$

Tree cover c is modeled as a normally distributed variable:

$$p(c) \stackrel{\text{def}}{=} N(\hat{c}, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(c-\hat{c})^2}{2\sigma^2}} \tag{S10}$$

where $\hat{c}$ is the estimated tree cover and σ is the root-mean-square error (RMSE) of the estimate, treated as its uncertainty $\sigma = \varepsilon$.

Using the resulting 37-year time series of 30-m, annual-resolution forest probabilities p(F), forest disturbance and establishment events were identified as statistically significant transitions across the 50% forest-probability threshold. A two-sample z-test was applied in a moving temporal kernel to detect such changes:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} - \frac{\sigma_2^2}{n_2}}} \tag{S11}$$

where $\bar{x}_1$ and $\bar{x}_2$ are antecedent and trailing means, respectively, σ_1 and σ_2 are their standard deviations, and n_1 and n_2 are the number of forest-probability estimates contributing to the values in all years.

The test was applied with the kernel centered on each forest-probability value of 50% in the series that was also increasing over time—i.e., $p(F_{t_1}) = \bar{x}_1 < 50\%$ and $p(F_{t_2}) = \bar{x}_2 >= 50\%$. If a statistically significant ($p <= 0.05$) difference was identified between the two ascending groups, the focal year was labeled as a forest gain or loss. If multiple significant losses or gains were detected in a pixel over the 36 years (1985-2020), up to three events were recorded. The detected forest disturbance was categorized as "incomplete" if its input annual tree cover had records missing at more than 7 years over the observable period (1985-2020); otherwise, "complete". The "incomplete" disturbance mainly occurred after 1999 due to the limited coverage of Landsat data before Landsat ETM+ (Fig. 3).

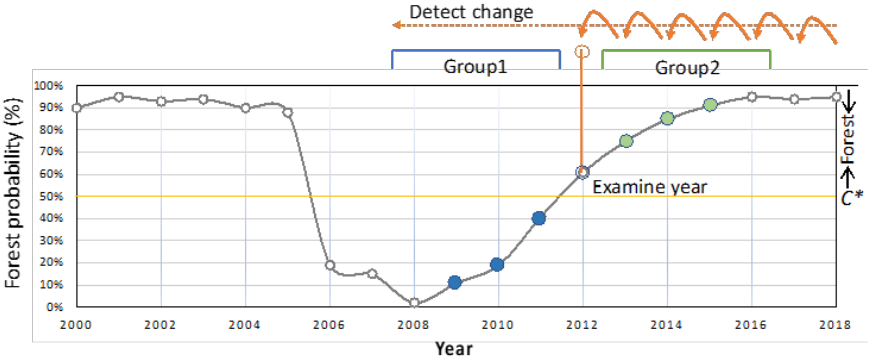


Fig. S13. Time-series of derived forest probability and division of groups for forest gain identification. C* defines the minimum tree cover for a forest.

Forest age, in years, at any year t^* and location *(x, y)* was calculated by subtracting the year of the most recent significant forest gain t^+ from the focal year:

$$Forest\ Age = t^* - t^+. \tag{S12}$$

While insensitive to the early stages of seedling establishment and growth, this formalization is logically consistent with the definition of forest and enabled consistent measurement across the entire boreal domain. Forested pixels were further categorized into "new" and "recovering" forests: "new" forests were identified as pixels with forest cover following a gain but no prior forest cover or loss earlier in the time series within a 150-m radius (5 pixels) over the observable period (1985 – 2020); "recovering" forests were identified as pixels with forest cover following a gain where a forest loss had been observed previously in the series.

S10. Validation of forest changes

Detected forest changes were validated using a two-tier, stratified sampling design. In the first tier, 41 WRS-2 tiles in North America and 43 in Eurasia were selected to represent the diversity of bioclimatic conditions, topographic gradients, and forest characteristics across the region. Within each selected tile, a second-tier sample was drawn consisting of 60 points in the posterior "change" stratum and 30 points in the "no change" stratum, which included both persistent forest and persistent non-forest. After excluding pixels with invalid observations and removing duplicates from overlapping WRS-2 tiles, a total of 4,320 unique sample points were retained (Fig. S14). These were randomly divided into 12 equally sized groups and assigned to 12 expert interpreters. For each point, interpreters assessed annual forest status (forest or non-forest) and identified the timing of any disturbance events using time series of Landsat NDVI and, where available, time-serial high-resolution imagery from Google Earth. Each interpretation was assigned a confidence score ranging from 0 (uninformative) to 3 (high confidence): 0 = uninterpretable, 1 = low, 2 = medium, and 3 = high. A total of 2,404 points received medium or high confidence scores and were retained for validation analyses.

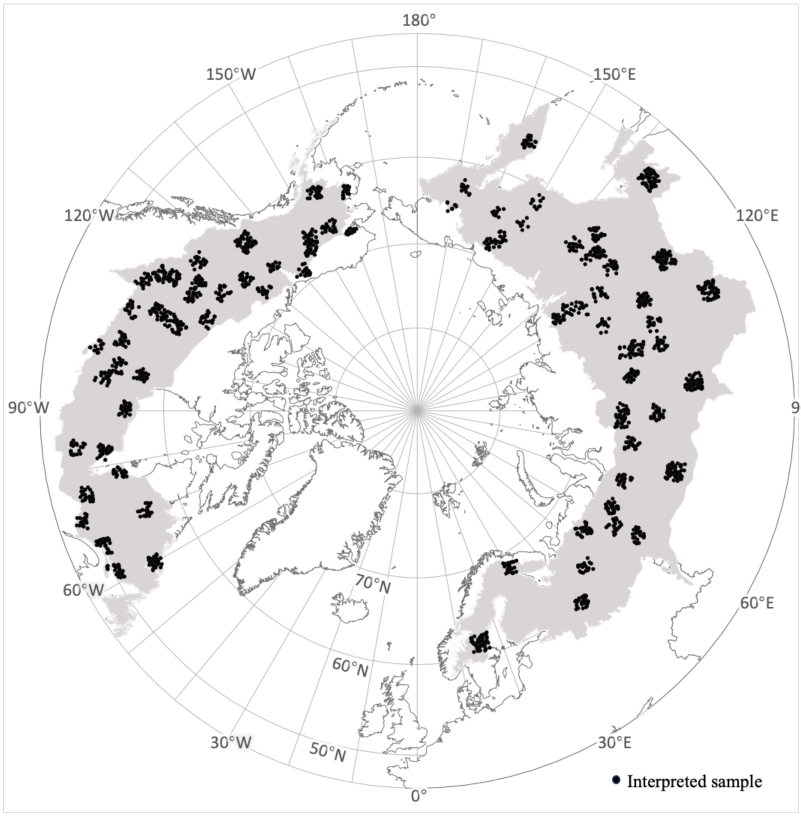


Fig. S14. Spatial distribution of the visually interpreted reference sample of forest change.

Estimates of the year of most recent forest loss showed strong agreement with visually interpreted reference data, with $R^2 = 0.79$, mean bias of 1.45 years, and dispersion (*RMSE*) of 4.66 years (Fig. S15~~Fig. S15~~a). The densest concentration of paired values occurred along the 1:1 line, and errors were skewed toward later detection of change by the algorithm, likely due to the preference of expert interpreters to earlier changes. Estimated year of the most recent forest gain showed a weaker linear relationship ($R^2 = 0.20$) to visually identified reference data (Fig. S15~~Fig. S15~~b); *RMSE* was 8.9 years, and *MBE* was 2.49 years. As with forest-loss year, estimated forest-gain years were higher than reference observations, likely due to the same cause. The error was dominated by unsystematic noise ($MSE_U > MSE_S$) in all three variables, i.e., forest-loss year, forest-gain year, and forest age (Table S1~~Table S1~~).

The accuracy assessment of forest age yielded an RMSE of 17.96 years and bias of –3.27 years. These values indicate high uncertainty in stand age retrieval. Accordingly, the forest age dataset should be interpreted as a broad-scale indicator of age distribution rather than an exact estimator at individual pixels. We recommend caution in map interpretation and highlight the need for further work to better characterize regional and class-specific errors.

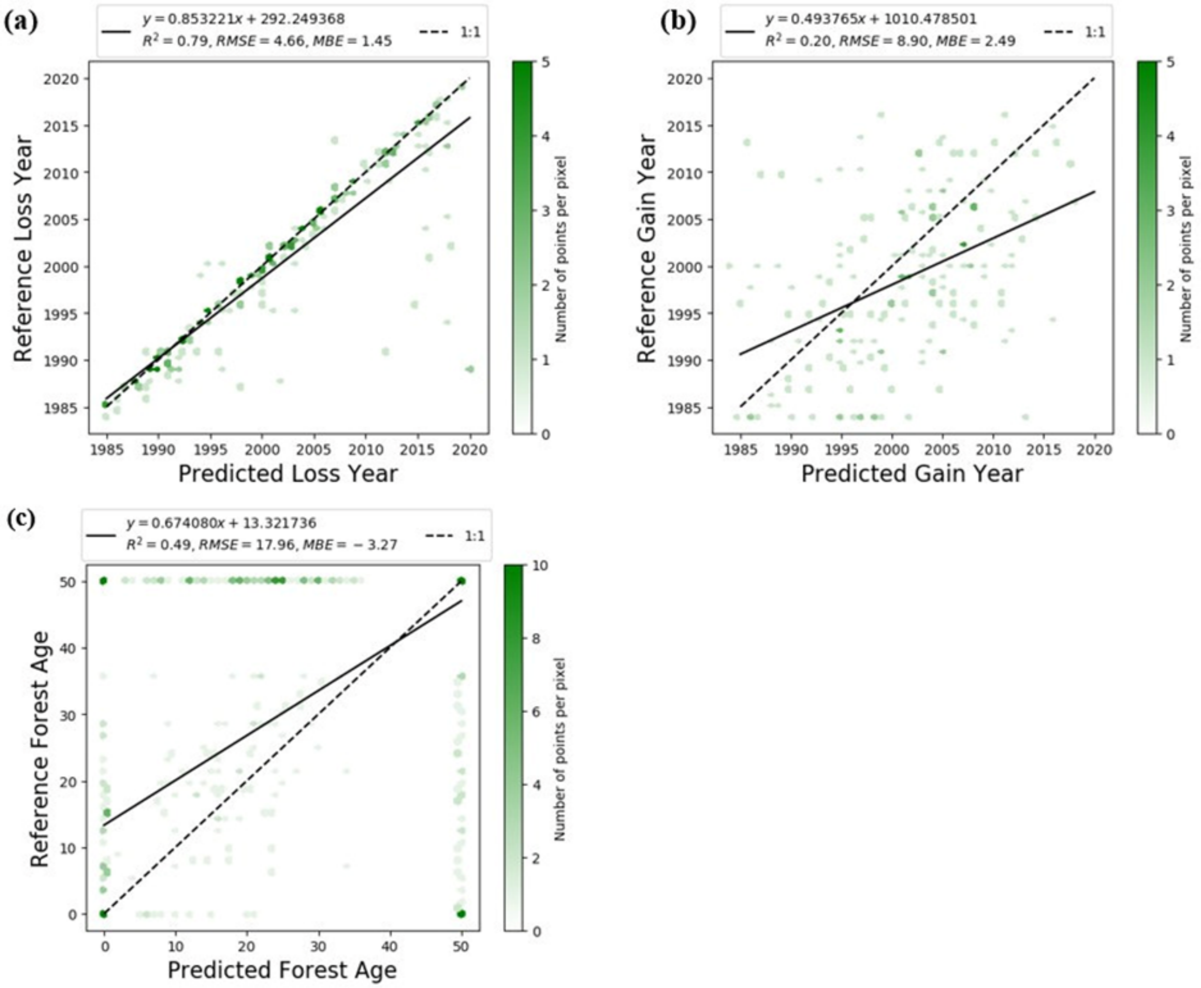


Fig. S15. Scatterplots and linear regression of estimated forest loss (a), gain (b), and age (c) relative to visually identified reference observations.

Table S1. Validation of forest loss, gain, and age estimates against visually identified reference observations.

Estimate	n	MBE	MSEs	MSEu	MSE	RMSE
Forest Loss (year)	231	1.45	4.08	17.67	21.76	4.66
Forest Gain (year)	178	2.49	21.23	58.05	79.28	8.90
Forest Age (years)	1,648	-3.27	67.41	255.29	322.70	17.96

S11. Estimation of carbon in aboveground biomass by stand age

Aboveground biomass carbon was modelled as a function of forest age (Cook-Patton et al., 2020) (Fig. S16~~Fig. S16~~), estimating for each age a range of possible carbon estimates by applying ±1 standard deviation from the intercept (μ = -35.7, σ = 12.6) and slope coefficients (μ = 23.2, σ = 3.2):

$$\mathrm{Mg\,C\,ha^{-1}} = (-35.7 \pm 12.6) + (23.2 \pm 3.2) \times \ln(stand\ age)\,. \qquad \text{(S13)}$$

Because the average age of forests older than 36 years could not be directly determined from the satellite record, total aboveground carbon stocks for these undated stands were bracketed using three hypothetical scenarios. These assumed mean stand ages of 36, 100, and 300 years, corresponding to estimated AGB carbon stocks of 19.1–58.4 Pg C, 35.8–80.5 Pg C, and 42.4–89.2 Pg C, respectively. The range of estimates accounts for parametric uncertainty in the growth model as well as variation in the assumed age structure of undated forest. However, these estimates do not incorporate potential variability related to changes in soil moisture or other edaphic factors.

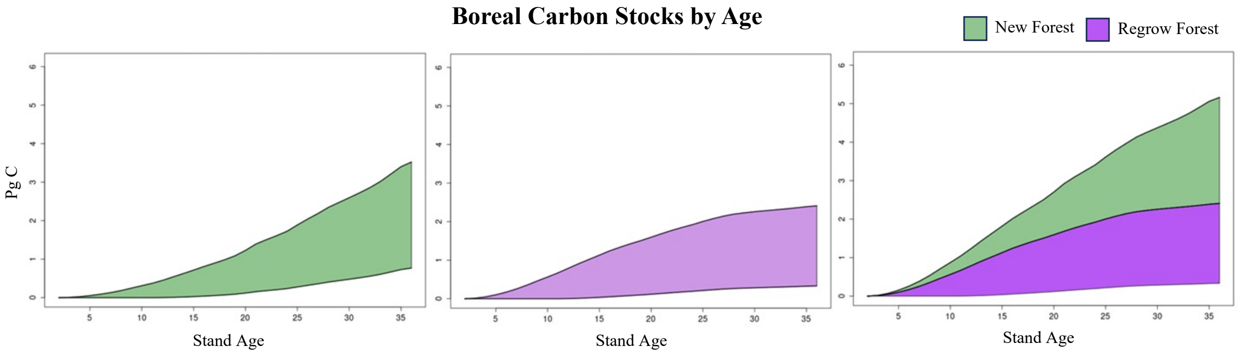


Fig. S16. Cumulative forest aboveground biomass (AGB) as a function of forest age. The area under the curves is calculated as a cumulative sum of AGB $_{\text{stand age}}$ × Area $_{\text{stand age}}$ for new forest, regrow forest, and both.

S12. Estimation of a potential range in ecosystem respiration from realistic temperatures

To assess the role of young and recovering forests in offsetting temperature-driven increases in boreal respiration, we compared their carbon sink potential against both empirical and modeled estimates of terrestrial ecosystem respiration (TER). Upscaled flux estimates from Jung et al. (2011), based on eddy covariance data for the late 1990s and early 2000s and applied to the 1982–2008 period, yield a mean annual boreal TER of 7.37 Pg C $\mathrm{yr^{-1}}$. During this interval, TER increased at a rate of 0.021 Pg C $\mathrm{yr^{-2}}$ ($r = 0.52$, $p < 0.005$). Extrapolated across the 1984-2020 period under a linear trend assumption, this corresponds to a cumulative efflux of approximately 9.87 Pg C (Wei et al., 2014; Dee et al., 2011; Forkel et al., 2016). In contrast, a higher-end model estimate using Q_{10}-based respiration dynamics projects 28.36 Pg C over the same interval (Mahecha et al., 2010).

To independently estimate the cumulative effect of warming on respiration, we applied a first-order Q_{10} model assuming constant baseline respiration and a linear 1.5°C increase in temperature over the 36-year period:

$$TER_{cum} = \sum_{i=1}^{36}[Q_{10}{}^{(T_i - T_{ref})/10} - Q_{10}{}^{(T_{avg} - T_{ref})/10}] * TER_{avg} \qquad \text{(S14)}$$

where T_i is the temperature departure from the average (T_{avg}) based on a 1.5°C/36-yr trend, Q_{10} is the temperature sensitivity of *TER,* and T_{ref} is the reference temperature.

Depending on the values of Q_{10} (1.4 to 2.2) and T_{ref} (5°C to 15°C), from more conservative (lower Q_{10}, higher T_{ref}) to more aggressive (large Q_{10}, low T_{ref}), the cumulative respiration fluxes from the temperature trend can range between 5 Pg C and 25 Pg C for the 36 years. The trend in tree cover could dampen 18% (percentile 10), or even exceed by twofold (percentile 90) the increase in TER only driven by temperature. Considering a T_{ref} of 15°C and a Q_{10} of 1.4 ± 0.1, as derived from eddy covariance data (Mahecha et al., 2010), the expected (median) buffering effect would represent 65% of the temperature-driven increase in TER, ranging between 15% and 125%.

Temperature trends were derived from two independent sources: the Climate Research Unit (CRU) dataset (1979–2016; Wei et al., 2014) and the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (1979–2016; Dee et al., 2011). Both records show significant warming trends over the boreal region during the study period: 0.038°C yr^{-1} ($r = 0.69$, $p < 1 \times 10^{-5}$) in the CRU dataset and 0.035°C yr^{-1} ($r = 0.73$, $p < 1 \times 10^{-6}$) in ERA (Fig. S17~~Fig. S17~~).

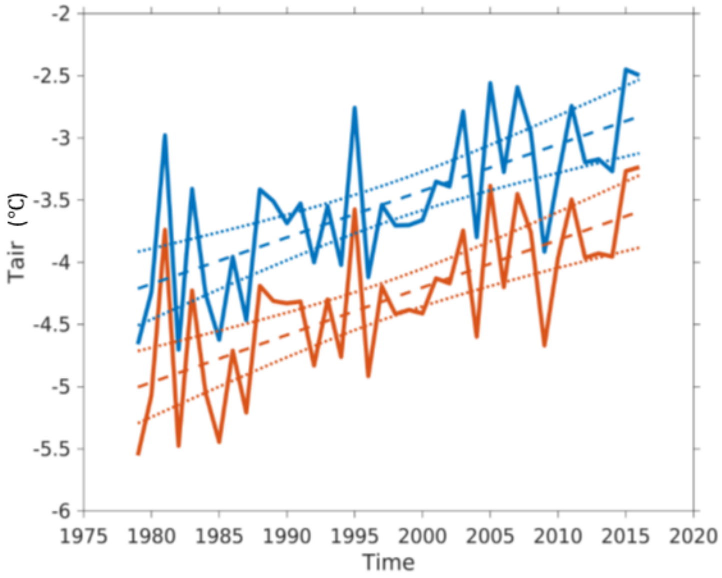


Fig. S17. Trends in air temperature in the boreal zone (orange: ERA, blue: CRU). We checked the consistency of the trend for several other temperature products and found a trend of 0.032 ± 0.006 (N=7) across them. Data sources: CERES/GPCP, CRU-JRA, CRUNCEP v6 and v8, GSWP3, WFDEI, and ERA5.

S13. References

Abrams, M., Crippen, R., and Fujisada, H.: ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD), Remote Sens., 12, 1156, https://doi.org/10.3390/rs12071156, 2020.

Blair, J. B., Rabine, D. L., and Hofton, M. A.: The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography, ISPRS J. Photogramm. Remote Sens., 54, 115–122, https://doi.org/10.1016/S0924-2716(99)00002-7, 1999.

Carroll, M., DiMiceli, C., Sohlberg, R., Huang, C., and Hansen, M. C.: MODIS Vegetative Cover Conversion and Vegetation Continuous Fields, in: Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS, edited by: Ramachandran, B., Justice, C. O., and Abrams, M. J., 725–745, Springer, New York, NY, https://doi.org/10.1007/978-1-4419-6749-7_32, 2011.

Cook-Patton, S. C., Leavitt, S. M., Gibbs, D., Harris, N. L., Lister, K., Anderson-Teixeira, K. J., Briggs, R. D., Chazdon, R. L., Crowther, T. W., Ellis, P. W., and Griscom, B. W.: Mapping carbon accumulation potential from global natural forest regrowth, Nature, 585, 545–550, https://doi.org/10.1038/s41586-020-2686-x, 2020.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., and Andrae, U.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N. D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., and Noss, R.: An ecoregion-based approach to protecting half the terrestrial realm, BioScience, 67, 534–545, https://doi.org/10.1093/biosci/bix014, 2017.

Dorogush, A. V., Ershov, V., and Gulin, A.: CatBoost: gradient boosting with categorical features support, arXiv [preprint], arXiv:1810.11363, 2018.

Feng, M., Sexton, J. O., Channan, S., and Townshend, J. R. G.: Quality assessment of Landsat surface reflectance products using MODIS data, Comput. Geosci., 38, 9–22, https://doi.org/10.1016/j.cageo.2011.04.011, 2012.

Fick, S. E., and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, Int. J. Climatol., 37, 4302–4315, https://doi.org/10.1002/joc.5086, 2017.

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Reichstein, M., and High-Latitude Ecosystem Modeling Group: Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems, Science, 351, 696–699, https://doi.org/10.1126/science.aac4971, 2016.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., and Chen, J.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res.-Biogeosci., 116, G00J07, https://doi.org/10.1029/2010JG001566, 2011.

Kukavskaya, E. A., Buryak, L. V., Shvetsov, E. G., Conard, S. G., and Kalenskaya, O. P.: The impact of increasing fire frequency on forest transformations in southern Siberia, For. Ecol. Manage., 382, 225–235, https://doi.org/10.1016/j.foreco.2016.10.016, 2016.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Sottocornola, M., and Cescatti, A.: Global convergence in the temperature sensitivity of respiration at the ecosystem level, Science, 329, 838–840, https://doi.org/10.1126/science.1189587, 2010.

Montesano, P. M., Macander, M. J., and Hoy, E. E.: ABoVE: Gridded LVIS Vegetation Structure across Alaska and Canada, 2017 and 2019, ORNL DAAC [data set], https://doi.org/10.3334/ORNLDAAC/1846, 2021.

Montesano, P. M., Neigh, C. S. R., Macander, M., Feng, M., and Noojipady, P.: The bioclimatic extent and pattern of the cold edge of the boreal forest: the circumpolar taiga-tundra ecotone, Environ. Res. Lett., 15, 105019, https://doi.org/10.1088/1748-9326/aba945, 2020.

Pittman, K., Hansen, M. C., Becker-Reshef, I., Potapov, P. V., and Justice, C. O.: Estimating global cropland extent with multi-year MODIS data, Remote Sens., 2, 1844–1863, https://doi.org/10.3390/rs2081844, 2010.

Sexton, J. O., Noojipady, P., Song, X.-P., Feng, M., Song, D. X., Kim, D. H., and Hansen, M. C.: Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error, Int. J. Digit. Earth, 6, 427–448, https://doi.org/10.1080/17538947.2013.786146, 2013.

Vermote, E., Roger, J. C., Franch, B., and Skakun, S.: LaSRC (Land Surface Reflectance Code): overview, application and validation using MODIS, VIIRS, LANDSAT and Sentinel 2 data, in: IGARSS 2018 – 2018 IEEE International Geoscience and Remote Sensing Symposium, 8173–8176, https://doi.org/10.1109/IGARSS.2018.8517622, 2018.

Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., and Wang, W.: The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data, Geosci. Model Dev., 7, 2875–2893, https://doi.org/10.5194/gmd-7-2875-2014, 2014.

Wulder, M. A., White, J. C., Masek, J. G., Dwyer, J., Roy, D. P., and Hermosilla, T.: The global Landsat archive: status, consolidation, and direction, Remote Sens. Environ., 185, 271–283, https://doi.org/10.1016/j.rse.2015.11.032, 2016.

Zhu, Z., and Woodcock, C. E.: Object-based cloud and cloud shadow detection in Landsat imagery, Remote Sens. Environ., 118, 83–94, https://doi.org/10.1016/j.rse.2011.10.028, 2012.