Research Article

Mapping Proportions of Eastern Redcedar in Kansas Using Sentinel-2 Imagery and Random Forest Model

Chen Liang [1,*], Dana Peterson [1], Xingong Li [1] and Jude H. Kastens [1]

1 University of Kansas; summerliangsummer@gmail.com, dpete@ku.edu, lixi@ku.edu, jkastens@ku.edu
* Corresponding Author: summerliangsummer@gmail.com;

Abstract: Eastern redcedar (*Juniperus virginiana* L.) encroachment is transforming grasslands across the central United States, with significant ecological and hydrological consequences. Effective management requires maps quantifying proportional cover to detect early invasion and prioritize interventions. We developed a workflow producing annual 10 m fractional cover maps at the University of Kansas Field Station (2016–2023). A 1 m reference dataset was generated by intersecting light detection and ranging (LiDAR)-derived canopy-height stability with leaf-off aerial imagery and aggregated to the Sentinel-2 grid. Random forest models were trained on seasonal Sentinel-2 composites using proportion-aware sampling, with LandTrendr temporal segmentation applied to reduce interannual spectral noise. Classification accuracy reached 0.84 (κ = 0.67), comparable to models incorporating National Ecological Observatory Network (NEON) hyperspectral data (0.85); however, high spatial autocorrelation among training samples suggests these represent upper-bound estimates. Winter normalized difference vegetation index (NDVI) and Simple Ratio Index were the most influential predictors, confirming dormant-season imagery value for evergreen discrimination. Regression models achieved low error (RMSE = 0.04–0.05), though accuracy decreased in stands exceeding 70% cover. LandTrendr-stabilized maps closely matched seasonal composites (R = 0.95–0.99) while producing smoother trajectories. The workflow demonstrates particular sensitivity to low-density invasion (<10% cover), enabling early detection when interventions remain cost-effective. Because independent validation was unavailable after 2016, temporal predictions warrant caution. Operational deployment elsewhere requires local reference data and model recalibration; transferability depends on species composition, phenological contrast, and landscape configuration.

Keywords: eastern redcedar encroachment, fractional cover mapping, Sentinel-2, random forest, LandTrendr

Academic Editor: Jeong Chang Seong
Received: 27 January 2026
Revised: 6 March 2026
Accepted: 10 April 2026
Published: 21 April 2026

Citation: Liang, C., Peterson, D., Li, X., & Kastens, J.H. (2026). Mapping Proportions of Eastern Redcedar in Kansas Using Sentinel-2 Imagery and Random Forest Model. Journal of Earth Observations and Geospatial Applications, 2(2), 21–43. https://doi.org/10.65372/jhe6g573

1. Introduction

Eastern redcedar (*Juniperus virginiana* L.), one of the most widespread conifers in the eastern United States, has expanded rapidly across Kansas, fundamentally transforming regional landscapes. Widespread fire suppression, coupled with land-use change and habitat fragmentation, has facilitated extensive encroachment into prairies and forest margins (Briggs et al., 2002), with redcedar volume in Kansas tallgrass prairie rising by over 15,000% between 1965 and 2010 (Galgamuwa et al., 2020). This transformation carries severe ecological, hydrological, and economic consequences (Zou et al., 2014; Moser et al., 2013).

Effective management requires accurate, repeatable mapping of redcedar proportional cover to differentiate nascent, low-density invasion from established closed-canopy stands and to prioritize interventions before canopy closure reduces treatment efficacy (Bidwell et al., 2017). Despite growing interest in satellite-based monitoring, operational workflows capable of producing reliable annual fractional cover products at fine spatial resolution remain limited. This study addresses that gap by developing a workflow that produces annual 10 m proportional cover maps at the University of Kansas Field Station (2016–2023), combining light detection and ranging (LiDAR)- and leaf-off imagery–derived reference data with Sentinel-2 time series, random forest (RF) regression, and LandTrendr temporal stabilization.

The specific objectives are to:

1. Construct and validate a 1 m evergreen reference map by combining LiDAR-based canopy-height stability and leaf-off aerial imagery, and aggregate this product to 10 m proportional redcedar cover for Sentinel-2 modeling.
2. Train random forest classification and regression models on multi-season Sentinel-2 predictors using proportion-aware sampling to address class imbalance along the full 0–100% cover gradient.
3. Compare two temporal compositing strategies: (a) winter–summer seasonal composites and (b) LandTrendr-stabilized full-year composites for generating annual Sentinel-2 predictor stacks.
4. Quantify redcedar encroachment trajectories from 2016 to 2023, with emphasis on early detection of low-cover invasion and densification of existing stands.

2. Literature Review

2.1. Eastern Redcedar Encroachment: Ecology and Consequences

Eastern redcedar encroachment is one of the most significant land cover changes occurring across the central United States. Historically, frequent fire, both natural and anthropogenic, maintained open grassland structure and suppressed woody establishment (Briggs et al., 2002). The removal of this ecological filter through fire suppression, combined with overgrazing and habitat fragmentation, has enabled redcedar to colonize prairies at unprecedented rates (Briggs et al., 2002).

The ecological consequences are well documented. Redcedar invasion suppresses herbaceous plant diversity and forage production, degrades wildlife habitat quality, and increases wildfire risk through fine fuel accumulation (Briggs et al., 2002; Hoff et al., 2018). Hydrological impacts are equally significant: replacement of grassland by redcedar canopy increases interception and transpiration, reducing annual runoff coefficients by approximately 80% in encroached catchments (Zou et al., 2014). McKinley and Blair (2008) further demonstrated that redcedar encroachment into mesic grasslands drives rapid carbon and nitrogen accrual in surface soils, with broad implications for ecosystem carbon cycling. These cumulative impacts highlight the urgency of accurate, timely monitoring of redcedar distribution and cover density across the landscape.

2.2. Remote Sensing Approaches to Redcedar and Juniper Mapping

Remote sensing has been widely applied to map redcedar and closely related juniper species (*Juniperus* spp.) across the Great Plains and southwestern United States, with approaches spanning a broad range of spatial resolutions and sensor types.

High-resolution aerial imagery, including National Agriculture Imagery Program (NAIP) orthophotos, offers fine spatial detail suitable for individual crown delineation but is constrained by infrequent acquisition schedules, inconsistent radiometry across flight lines, and limited spectral resolution (Maxwell et al., 2017). Wavelet-based analysis of NAIP imagery for juniper detection has been shown to lose accuracy above approximately 20% canopy cover (Poznanovic et al., 2014), limiting its utility for assessing densification in established stands.

Moderate-resolution satellite sensors, particularly Landsat (30 m), have been used for statewide and regional mapping. Wang et al. (2018) combined Landsat and phased array type L-band synthetic aperture radar (PALSAR) data to map juniper encroachment in Oklahoma, achieving high forest/non-forest classification accuracy (94–96%) but detecting only approximately 30% of stands in the 10–20% cover range. This insensitivity to sparse cover is a critical limitation for early detection applications, since prescribed fire is most effective and economical at low cover densities before canopy closure reduces treatment efficacy (Bidwell et al., 2017).

The Sentinel-2 multispectral instrument (MSI), with 10 m spatial resolution in visible and near-infrared bands and a 5-day revisit frequency, offers a promising intermediate between high-resolution aerial imagery and coarser satellite sensors. Its red-edge bands (B5, B6, B7) provide sensitivity to chlorophyll content and canopy structure unavailable in Landsat, while its temporal frequency enables robust multi-seasonal compositing (Pasquarella et al., 2022). Despite this potential, operational workflows for producing reliable annual proportional cover products from Sentinel-2 remain limited in the redcedar and juniper literature.

Airborne hyperspectral sensors, including those operated by the National Ecological Observatory Network (NEON), provide fine spectral resolution capable of discriminating species-level composition. However, their high cost, limited spatial coverage, and irregular acquisition schedules restrict their use in operational monitoring programs, motivating evaluation of whether freely available multispectral data can provide comparable discriminatory power for redcedar mapping.

2.3. Fractional Cover Estimation and Machine Learning

Mapping proportional or fractional vegetation cover, rather than simple presence/absence, provides information critical for management prioritization, fuel load assessment, and tracking of encroachment trajectories. Fractional cover mapping has been approached through spectral mixture analysis, regression-based methods, and machine learning algorithms (Filippelli et al., 2020).

The random forest algorithm has emerged as a widely adopted tool for vegetation mapping from multispectral imagery due to its ability to handle nonlinear predictor relationships, resistance to overfitting in high-dimensional feature spaces, and provision of interpretable variable importance metrics (Breiman, 2001). RF has been applied successfully to fractional cover estimation of pinyon-juniper woodlands (Filippelli et al., 2020) and a range of other vegetation types. A persistent challenge in fractional cover applications is class imbalance: in encroaching landscapes, the vast majority of pixels contain sparse cover, causing standard models to underpredict moderate-to-high density conditions. Proportion-aware or stratified sampling strategies have been proposed to mitigate this bias, though their comparative effectiveness across the full cover gradient has received limited attention in the redcedar literature.

Variable importance analyses in prior studies of evergreen encroachment mapping have consistently highlighted the value of dormant-season imagery. Spectral indices derived from winter composites, when deciduous and herbaceous vegetation is senescent, maximize contrast between evergreen and non-evergreen cover, enabling more reliable discrimination than growing-season imagery alone (Wang et al., 2018; Filippelli et al., 2020). Normalized difference vegetation index (NDVI) and simple band ratio indices such as the simple ratio index (SRI) have repeatedly emerged as strong predictors in this context.

2.4. Temporal Stabilization: LandTrendr

The LandTrendr algorithm, originally developed for Landsat time series analysis, models per-pixel spectral histories as piecewise linear segments to distinguish persistent trends from short-term noise arising from atmospheric effects, cloud shadows, or phenological variability (Kennedy et al., 2010). Its implementation in Google Earth Engine (Kennedy et al., 2018) has facilitated broad adoption across a range of vegetation monitoring applications, including forest disturbance detection, post-fire recovery assessment, and land cover change mapping (Pasquarella et al., 2022).

Most existing applications of LandTrendr operate on post-classification outputs or single spectral indices to detect disturbance and recovery events. Its use as a pre-classification smoothing tool, applied to full predictor stacks prior to model training and application, has received comparatively little attention, particularly in the context of slow-onset encroachment by evergreen species. In encroachment monitoring, where year-to-year spectral fluctuations driven by drought, variable image acquisition timing, or atmospheric anomalies may be misinterpreted as real cover change, temporal stabilization of predictor variables prior to modeling offers a potentially valuable approach for improving the reliability of annual map series.

2.5. Reference Data Generation for Fractional Cover Modeling

Accurate reference data are fundamental to training and validating fractional cover models. Field-based canopy cover measurements provide ground truth but are costly to acquire, spatially sparse, and difficult to scale across large landscapes. Manual photointerpretation of high-resolution aerial imagery is widely used but introduces subjectivity and is labor-intensive at scale.

Integration of independent remote sensing datasets offers a more objective and scalable alternative. Filippelli et al. (2020) demonstrated that intersecting LiDAR-derived canopy structure with spectral imagery could produce robust training data for pinyon-juniper fractional cover modeling. The seasonal stability of LiDAR-derived canopy height, comparing leaf-off and leaf-on acquisitions to identify structurally persistent evergreen vegetation, provides a structural basis for reference map generation that is independent of spectral classification alone. Combining this structural filter with leaf-off spectral classification constrains the

reference map to pixels that are both structurally and spectrally consistent with evergreen conifers, reducing commission errors from spectrally similar non-evergreen surfaces such as shadows or dark soils.

3. Study Area and Methods

3.1. Study Area

The study was conducted at the University of Kansas Field Station (UKFS), located approximately 9 km north of Lawrence, Kansas, USA (39.05° N, 95.20° W). Positioned within the prairie–forest ecotone of the central United States, the UKFS landscape comprises a mosaic of remnant and restored tallgrass prairie, mature oak–hickory forest, and areas of active eastern redcedar encroachment. As part of the National Ecological Observatory Network Prairie Peninsula Domain (D06), the site provides a representative setting for examining woody encroachment dynamics (Figure 1). The site is characterized by a single dominant encroaching evergreen species (eastern redcedar), pronounced winter dormancy of deciduous and herbaceous vegetation, and relatively gentle topography. These characteristics create favorable conditions for spectral discrimination of evergreen cover and should be considered when evaluating the transferability of results to other regions.

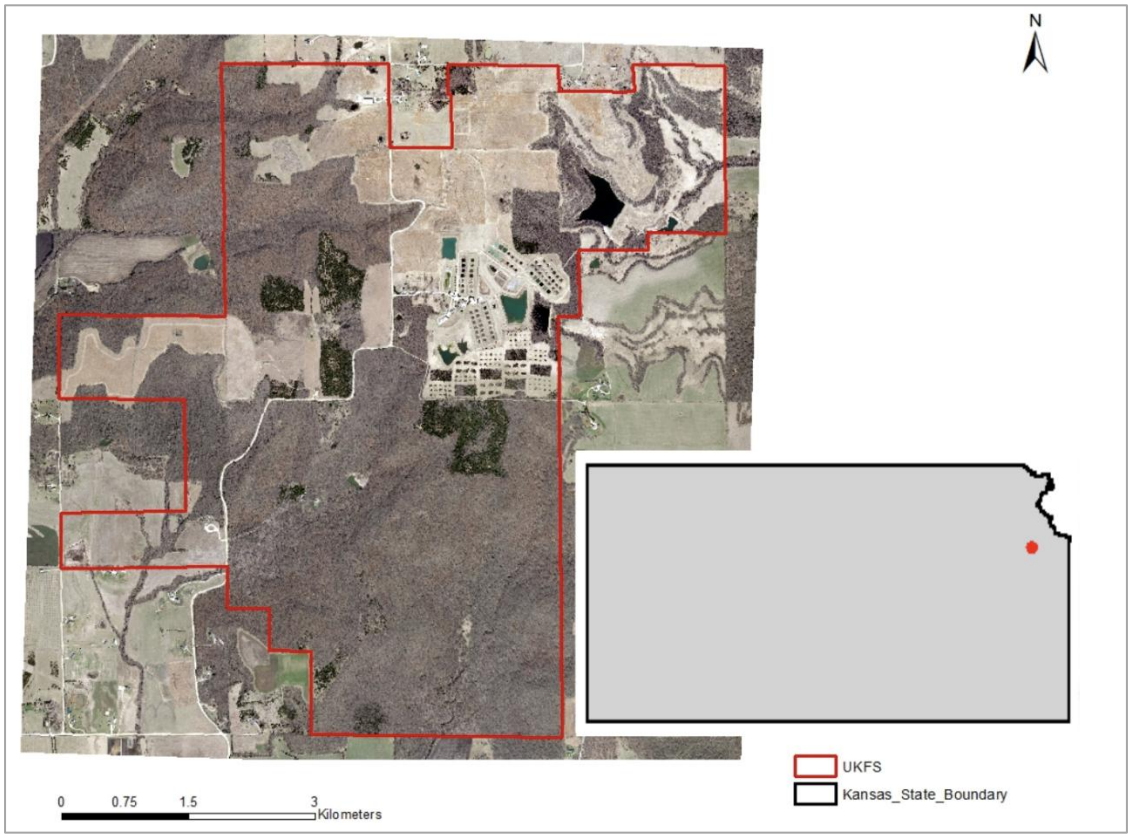

Figure 1. Study area location. The University of Kansas Field Station (UKFS) boundary (red outline) is shown overlaid on 2004 NG911 high-resolution imagery. Inset: Location of UKFS within Kansas, USA. The site is characterized by a single dominant encroaching evergreen species (eastern redcedar), pronounced winter dormancy of deciduous and herbaceous vegetation, and relatively gentle topography. These characteristics create favorable conditions for spectral discrimination of evergreen cover and should be considered when evaluating the transferability of results to other regions.

3.2. Workflow Overview

The workflow consisted of six main steps: (1) development of a 1 m binary evergreen reference map from LiDAR and leaf-off aerial imagery; (2) aggregation of the reference map to 10 m redcedar proportions aligned with the Sentinel-2 grid; (3) generation of Sentinel-2 predictor composites using two temporal workflows

(winter–summer "Seasonal" versus LandTrendr-stabilized "LandTrendr"); (4) training and evaluation of RF classification and regression models; (5) application of a "train-once, apply-all" regression model to derive annual 10 m fractional-cover maps for 2016–2023; and (6) temporal analysis and comparison of encroachment trajectories between the Seasonal and LandTrendr map series. All temporal analyses used a 2014–2016 high-resolution reference period and a 2016–2023 Sentinel-2 time series.

3.3. High-Resolution Reference Map Development

3.3.1. LiDAR Canopy-Height Stability

A 1 m binary reference map (redcedar versus non-redcedar) was generated by intersecting two independent data sources to create a conservative estimate of evergreen cover. LiDAR-derived canopy height models (CHMs) from leaf-off (2015) and leaf-on (2016) conditions were compared; pixels exhibiting minimal seasonal height change (−1.5 to 1.5 m) were identified as structurally stable evergreen vegetation (Figure 2a). The $|\Delta height| \leq 1.5\ m$ threshold was chosen based on visual inspection of example profiles, ensuring that stable evergreen canopies were retained while minimizing confusion with deciduous trees that exhibit larger seasonal height differences due to leaf-off conditions. To evaluate the robustness of the 1.5 m height-change threshold, we tested thresholds of 1.0, 1.5, and 2.0 m. The 1.0 m threshold excluded 23% of visually confirmed redcedar pixels, primarily larger trees with slight crown displacement between LiDAR acquisitions. The 2.0 m threshold included 12% additional pixels that visual inspection revealed to be predominantly deciduous edge effects and registration artifacts. The 1.5 m threshold maximized the F1-score (0.91) for evergreen classification when validated against 200 manually interpreted points, compared to 0.84 for the 1.0 m threshold and 0.87 for the 2.0 m threshold. The F1-score is the harmonic mean of precision and recall ($F1 = 2 \times [Precision \times Recall] / [Precision + Recall]$), providing a single balanced measure of classification performance that accounts for both false positives and false negatives.

3.3.2. Leaf-Off Conifer Spectral Classification

A 0.3 m resolution leaf-off aerial orthophoto acquired in 2014–2015 was used to identify evergreen crowns. The imagery was first clipped to the UKFS boundary. An unsupervised iterative self-organizing data analysis technique (ISODATA) clustering algorithm was applied to the visible and near-infrared bands, allowing up to 20 spectral clusters with a minimum cluster size of 10 pixels. The maximum of 20 clusters was selected based on preliminary tests balancing computational efficiency with sufficient spectral detail to distinguish conifer crowns from other dark, high-NIR surfaces (e.g., shadows, water) while maintaining interpretability during the manual labeling process. Following Burchfield (2017), each cluster was visually inspected against the imagery and those dominated by dark, high-NIR tree crowns were labeled as "conifer" classes. Remaining clusters were combined into non-conifer classes (deciduous trees, grassland, bare soil, and water). The conifer clusters were merged into a single binary raster representing the conifer spectral class at 0.3 m resolution.

Sensitivity analysis indicated that results were robust to the number of clusters: tests with 15, 20, and 25 maximum clusters produced conifer class maps with >95% spatial agreement (intersection-over-union), as the primary spectral distinction between dark, high-NIR conifer crowns and other surfaces was captured consistently across cluster counts. The 20-cluster parameterization was retained as a balance between computational efficiency and interpretability during manual labeling.

3.3.3. Intersection of Structural and Spectral Masks

To combine the structural and spectral information, the 0.3 m conifer raster was first resampled to 1 m using a majority filter, ensuring that each 1 m pixel was labeled "conifer" if at least 50% of its 0.3 m subpixels were conifer (i.e., ≥6 of the approximately 11 subpixels within each 1 m cell). This layer was then intersected with the LiDAR-derived stability mask by retaining only those 1 m pixels that were both structurally stable ($|\Delta height| \leq 1.5\ m$) and spectrally conifer. Open water polygons from the U.S. Geological Survey (2019) were rasterized and used to remove all water pixels from the resulting map. The final product is a conservative 1 m evergreen map in which only pixels consistently tall and spectrally coniferous are labeled as redcedar (Figure 2c). Because eastern redcedar is the only evergreen conifer species occurring in significant abundance at UKFS, this "evergreen" classification is effectively synonymous with redcedar presence at this site.

3.3.4. Aggregation to the Sentinel-2 Grid and Validation

The 1 m binary evergreen map was aggregated to the 10 m Sentinel-2 grid by calculating the proportion of redcedar pixels within each cell (expressed as 0–1), using a projected coordinate system aligned with the Sentinel-2 tiling grid to minimize sub-pixel misregistration.

To evaluate the accuracy of this aggregated reference product, an independent validation was conducted using manual image interpretation of 100 randomly selected 10 m pixels across the study area. For each selected pixel, eastern redcedar crowns were manually digitized using high-resolution aerial imagery and then two estimates were calculated: (1) the fractional cover from the aggregated 1 m reference map, and (2) the fractional cover from the digitized crowns. These estimates were compared using root mean square error (RMSE) and mean absolute error (MAE).

Spatial agreement was also quantified using the intersection-over-union (IoU) metric. For each validation pixel, the area mapped as redcedar in both datasets (intersection) and the total area mapped as redcedar in either dataset (union) were computed. IoU was then calculated as:

$$IoU = A_{intersection} / A_{union} \quad (1)$$

where $A_{intersection}$ is the area mapped as redcedar in both datasets and A_{union} is the area mapped as redcedar in either dataset.

An IoU value of 1 indicates perfect spatial agreement (identical redcedar extent), whereas a value of 0 indicates no overlap. To assess whether observed IoU values exceeded chance agreement, a permutation test was conducted: for each of the 100 validation pixels, redcedar locations were randomly shuffled 10,000 times within the pixel boundary, and IoU was recalculated for each permutation. The observed IoU was then compared against this null distribution, with p-values derived as the proportion of permuted IoU values equaling or exceeding the observed value. This metric complements RMSE and MAE by explicitly assessing how well the spatial pattern of mapped redcedar within each 10 m pixel matches between the reference map and the manual interpretation (Figure 3).

This 10 m proportional reference map served as the response variable for all random forest models.

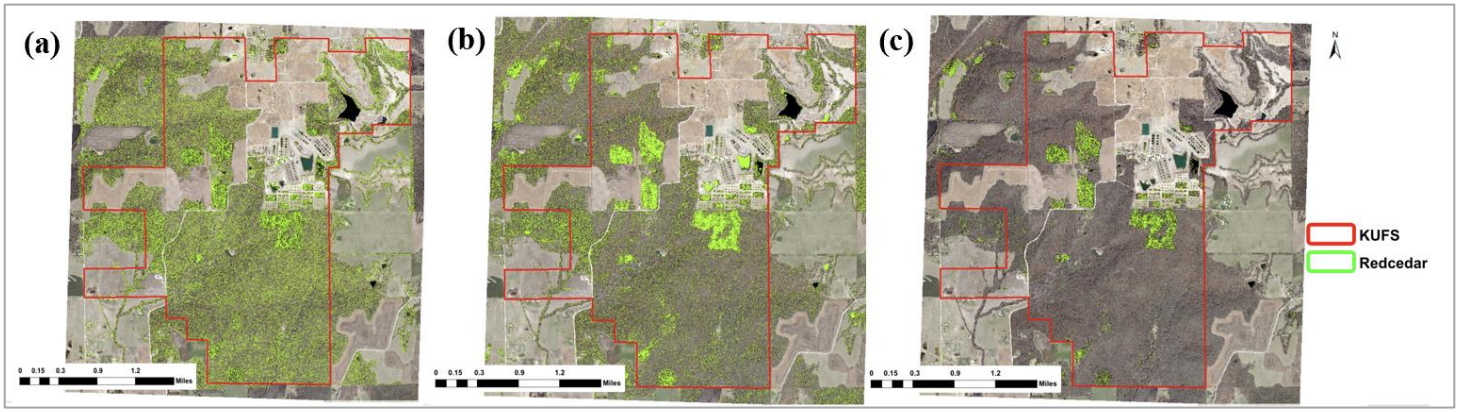

Figure 2. Development of the 1 m binary eastern redcedar reference map. (a) Structurally stable vegetation identified from LiDAR height-change analysis ($|\Delta height| \leq 1.5\,m$). (b) Conifer spectral class from ISODATA classification of leaf-off orthoimagery. (c) Final reference map from intersection of (a) and (b). Green pixels indicate classified redcedar presence.

3.3.5. Temporal Gap Considerations

A temporal gap exists between the reference dataset (2014–2016) and the full Sentinel-2 modeling period (2016–2023), reflecting data availability constraints. However, the slow growth rate of eastern redcedar partially mitigates this limitation. Mature redcedar in the central Great Plains typically exhibits height growth of 15–30 cm yr^{-1} and lateral crown expansion of 10–20 cm yr^{-1} under favorable conditions (Lawson, 1990). At these rates, canopy changes over the study period would be modest relative to the 10 m Sentinel-2 pixel size, particularly for established stands that dominate the higher proportional cover classes.

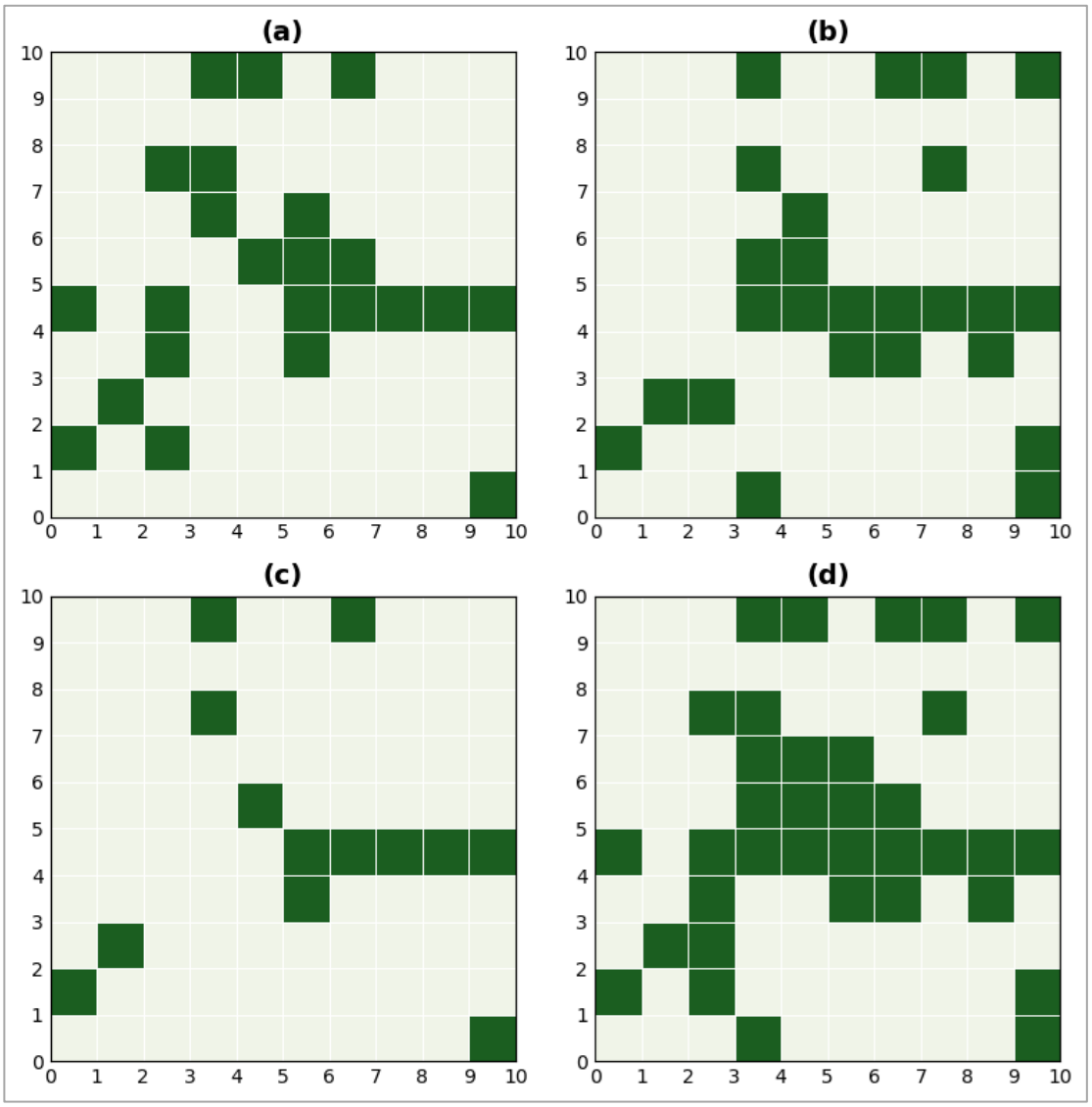

Figure 3. Schematic illustration of the intersection-over-union validation metric within a single 10 m Sentinel-2 pixel. Each cell represents a 1 m sub-pixel; dark green shading indicates redcedar classification. (a) Automated reference map classification. (b) Manual crown digitization. (c) Intersection of (a) and (b), representing true positives. (d) Union of (a) and (b), representing total mapped area. IoU = area(c) / area(d).

To estimate the potential magnitude of temporal mismatch error, we considered published demographic rates for eastern redcedar in similar environments. Eastern redcedar establishment in intact grasslands is constrained by the seedling recruitment "bottleneck" characteristic of woody plant encroachment, where the early establishment stage represents the most vulnerable period of the life cycle (Archer et al., 2017). Cover expansion rates for eastern redcedar in tallgrass prairie average approximately 2.3% per year in fire-suppressed landscapes (Briggs et al., 2002). Over the 7-year maximum gap (2016–2023), these demographic rates would translate to modest net changes in tree populations. Given that new seedlings require 5–10 years to reach detectable crown sizes at 10 m resolution, the primary source of error is likely underdetection of juveniles established between 2016 and approximately 2018, rather than widespread mortality of mapped trees.

Nonetheless, pixels experiencing active colonization by new seedlings or juveniles may be underrepresented in the reference dataset, potentially contributing to underestimation of cover in areas with recent recruitment.

3.4. Sentinel-2 Data and Predictor Generation

3.4.1. Sentinel-2 Collection and Per-Image Processing

Sentinel-2 scenes intersecting UKFS from 2016–2023 were processed in Google Earth Engine (GEE; Gorelick et al., 2017). Scenes were atmospherically corrected to Level-2A surface reflectance and clipped to the study area. Cloudy pixels were masked using the Sentinel-2 cloud-probability band (MSK_CLDPRB ≥ 20%), a threshold commonly applied in Sentinel-2 land cover studies (Pasquarella et al., 2022; Main-Knorn et al., 2017). This threshold was selected based on preliminary tests: a 10% threshold retained only 62% of winter observations and 71% of summer observations, insufficient for reliable compositing in some years; a 30% threshold retained visible cloud contamination in 15% of composite pixels upon visual inspection; the 20% threshold retained 78% of winter and 85% of summer observations while effectively removing cloud contamination.

For each cloud-free observation, ten 10 m (or resampled) bands were retained: B2 (blue), B3 (green), B4 (red), B5–B7 (red-edge), B8 and B8A (near-infrared), B11 (short-wave infrared 1), and B12 (short-wave infrared 2). Vegetation and moisture indices were also computed, including NDVI, GNDVI (Green NDVI), NDVI705 (Red-edge NDVI), SRI, inverted red-edge chlorophyll index (IRECI), enhanced vegetation index (EVI), land surface water index (LSWI), and fire-sensitive indices normalized burn ratio (NBR) and NBR2. These bands and indices comprised the predictor set (Table 1).

Table 1. Sentinel-2 spectral bands and derived indices used as predictor variables in the random forest models.

Predictor	Description	Formula
B2	blue	—
B3	green	—
B4	red	—
B5	red edge 1	—
B6	red edge 2	—
B7	red edge 3	—
B8	near infrared (NIR)	—
B8A	20 m red-edge/narrow NIR	—
B11	short wave infrared 1 (SWIR1)	—
B12	short wave infrared 2 (SWIR2)	—
SRI	simple ratio index	NIR / Red
IRECI	inverted red-edge chlorophyll index	(B7 − B4) / (B5 / B6)
GNDVI	green NDVI	(NIR − Green) / (NIR + Green)
NDVI	normalized difference vegetation index	(NIR − Red) / (NIR + Red)
NDVI705	red-edge NDVI	(B6 − B5) / (B6 + B5)
EVI	enhanced vegetation index	2.5 × (NIR − Red) / (NIR + 6×Red − 7.5×Blue + 1)
LSWI	land surface water index	(NIR − SWIR1) / (NIR + SWIR1)
NBR	normalized burn ratio	(NIR − SWIR2) / (NIR + SWIR2)
NBR2	normalized burn ratio 2	(SWIR1 − SWIR2) / (SWIR1 + SWIR2)

3.4.2. Seasonal-Composite Predictor Stacks (“Seasonal Method”)

To exploit the strong phenological contrast between evergreen redcedar and dormant herbaceous/deciduous vegetation, two seasonal composites were generated per year:

- Winter window: 1 November of year t to 31 January of year t+1 (non-evergreen vegetation largely senescent)
- Summer window: 1 June–31 August of year t (peak herbaceous greenness)

For example, the 2016 winter composite used imagery from November 2015 through January 2016, while the 2016 summer composite used imagery from June through August 2016. This convention ensures that each annual map represents conditions through the end of that calendar year's growing season.

Within each window, all cloud-free observations were aggregated. For each band and index (except NDVI) the median was computed; NDVI was summarized using the maximum to emphasize the greenest observation and reduce residual cloud and shadow effects. Winter and summer composites were stacked into a single multiband image per year, with "_winter" and "_summer" suffixes in band names. These Seasonal predictor stacks formed the basis of the initial mapping approach.

The chosen winter (November–January) and summer (June–August) windows balance strong phenological separation between evergreen and deciduous/herbaceous vegetation with adequate image availability in this relatively cloudy region.

3.4.3. LandTrendr-Stabilized Full-Year Stacks ("LandTrendr Method")

To reduce interannual spectral noise and create a temporally consistent set of predictors, a second workflow based on the LandTrendr algorithm was developed (Kennedy et al., 2010, 2018). This "pre-classification smoothing" approach differs from the Seasonal method by stabilizing the predictor time series before it is used for model training or classification.

Annual composites for the entire study period (2016–2023) were generated in Google Earth Engine. For each calendar year, all available Sentinel-2 observations (January 1 to December 31) were masked for clouds (MSK_CLDPRB ≥ 20%) and processed to compute the full set of predictor bands and indices (Table 1). A single composite image was created for each year by calculating the median pixel value for all spectral bands and most indices. To better capture peak vegetation vigor and reduce residual shadow effects, the maximum value was used for the NDVI index.

The LandTrendr temporal segmentation algorithm was then applied independently to the time series of each predictor band to filter short-term spectral noise while preserving long-term trends (Figure 4). Algorithm parameters were set to capture gradual trends while filtering short-term noise, with key parameters including maxSegments: 6, spikeThreshold: 0.9, preventOneYearRecovery: true, and minObservationsNeeded: 6. These settings reflect ecological expectations for eastern redcedar encroachment: canopy expansion and densification at this site are typically gradual over multiple years, with few true one-year disturbances or recoveries. The maxSegments value of 6 provides sufficient flexibility to capture one or two trajectory inflections within the 8-year study period while preventing overfitting to annual spectral variability; this parameterization follows recommendations for gradual vegetation change in relatively short time series (Kennedy et al., 2018; Pasquarella et al., 2022). Preventing one-year recovery further suppresses spurious "disturbance" artifacts caused by residual cloud, snow, or shadow effects while preserving longer-term trends.

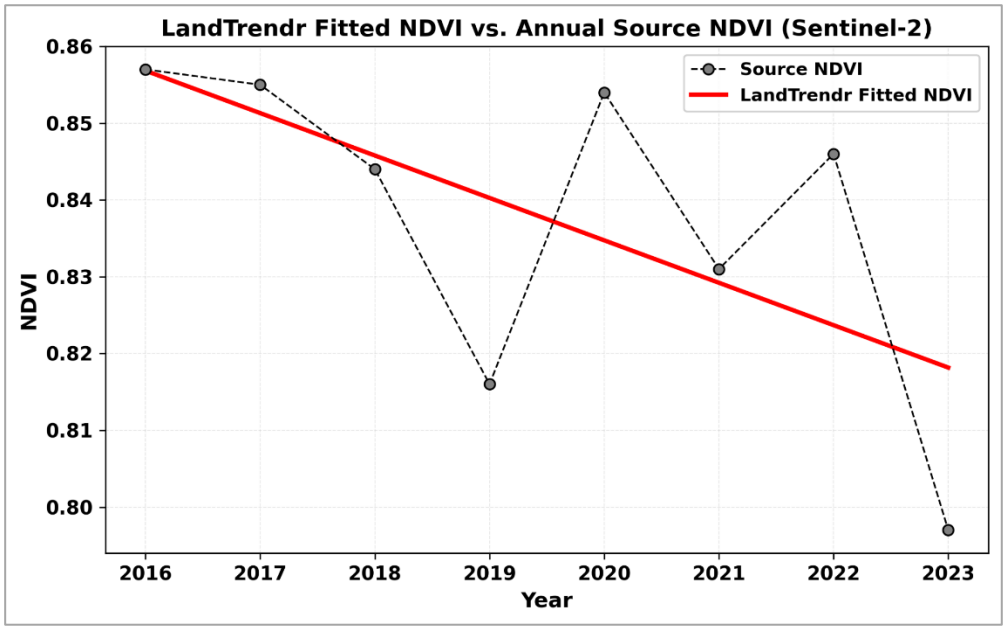

Figure 4. Example of LandTrendr temporal stabilization for a single pixel. Gray dashed line: raw annual NDVI values showing interannual variability. Red solid line: LandTrendr-fitted piecewise linear trend that filters short-term noise while preserving the long-term trajectory.

From the resulting LandTrendr output array for each predictor, per-year fitted values at the original observation years were extracted. These smoothed annual values were then reassembled to create the final, temporally stabilized predictor stacks for each year. LandTrendr was applied to the time series of Sentinel-2 bands and indices (predictor variables). The RF regression model was then applied to these stabilized predictor stacks to generate the LandTrendr series of proportional cover maps.

3.5. Modeling Framework

Two random forest models were developed: (1) a binary classification model to compare the predictive power of different sensor inputs, and (2) a regression model to estimate proportional canopy cover. For both models, data were split 70/30 for training and testing, with the process repeated 30 times using random partitions to ensure stable results. Each RF model comprised 300 trees, with the number of variables considered at each split (mtry) set to the default $\sqrt{p}$, where p is the number of predictors.

3.5.1. Classification Model

The first model was an RF classifier trained to distinguish redcedar presence from absence at the 10 m pixel scale. This comparison was included to evaluate whether the freely available Sentinel-2 data could match the discriminatory power of more expensive airborne hyperspectral acquisitions, thereby justifying an operational workflow based solely on Sentinel-2. To assess the relative contribution of Sentinel-2 versus high-resolution airborne data, six predictor scenarios were tested (Table 3), combining Sentinel-2 bands (Table 1), NEON hyperspectral bands (Table 2), and derived NDVIs.

Table 2. NEON predictor variables evaluated in the random forest classification model.

Predictor	Description
ARVI	atmospherically resistant vegetation index
EVI	enhanced vegetation index
NDLI	normalized difference lignin index
NDNI	normalized difference nitrogen index
NDVI	normalized difference vegetation index
SAVI	soil adjusted vegetation index

Table 3. Predictor scenarios used for the random forest classification model.

Scenario	Data Source(s)	Description
1	NEON	All spectral bands from the NEON sensor (summer only)
2	Sentinel-2	All spectral bands from Sentinel-2 (winter and summer composites)
3	NEON NDVI	normalized difference vegetation index from the NEON sensor
4	Sentinel-2 NDVI	normalized difference vegetation index from Sentinel-2 (winter and summer)
5	Sentinel-2 NDVI + NEON bands	Combination of Sentinel-2's NDVI and all NEON spectral bands
6	Sentinel-2 bands + NEON bands	Combination of all spectral bands from both Sentinel-2 and NEON

NEON imagery was limited to summer acquisitions, creating temporal asymmetry with the multi-seasonal Sentinel-2 composites. This constraint potentially favors Sentinel-2 scenarios, as winter imagery provides optimal phenological contrast for evergreen detection; however, equivalent NEON performance would demonstrate hyperspectral data compensating for suboptimal timing.

3.5.2. Proportion Regression Model

The primary model was an RF regression algorithm trained to estimate the fractional (0–100%) redcedar cover for each 10 m pixel. This model used the full set of winter and summer Sentinel-2 bands and indices. Recognizing that the landscape is heavily skewed toward low-cover conditions, which can bias models, two

sampling frameworks were developed and compared (Table 4). A total of 52,528 pixels (purposive sampling) and 50,237 pixels (stratified sampling) were selected for model development. Pixels were grouped into 0.1-wide Proportion Groups (PGs; PG1: 0–0.1, PG2: 0.1–0.2, …, PG10: 0.9–1.0). The 70/30 train-test split was applied separately within each Proportion Group to ensure that both training and test sets contained representative samples across the full cover gradient (stratified splitting). The two sampling strategies were defined as follows:

- Stratified sampling: Samples were drawn from each 0.1 Proportion Group in direct proportion to their frequency on the landscape. This mirrors the real-world distribution, which is heavily skewed toward PG1 (0–0.1), and serves as a baseline control.
- Purposive sampling: This strategy was designed to modestly increase the representation of rarer, higher-cover PGs while preserving the overall skew toward low cover. All available pixels with redcedar proportion ≥10% (PGs 2–10) within the study area were included (n = 5,442), and the remaining sample quota was filled with PG1 pixels (0–10% cover), ensuring that low-cover conditions remained numerically dominant. The similarity in sample counts between purposive and stratified designs for high-cover classes (e.g., PG10: 9 vs. 7 samples) reflects the fundamental scarcity of dense stands on the landscape rather than subsampling; purposive sampling captured all available high-cover pixels.

The total available pixels meeting quality criteria numbered 62,798 (Table 5); however, Google Earth Engine memory constraints required subsampling for model training, resulting in the reduced sample sizes shown in Table 4 (50,237 for stratified sampling; 52,528 for purposive sampling). Subsampling preserved the proportional distribution across cover classes while enabling computational tractability.

Ground truth data were available only for the 2014–2016 reference period; no independent validation data existed for subsequent years. Consequently, model accuracy for 2017–2023 could not be directly assessed. A "train-once, apply-all" strategy was adopted under the assumption that spectral–proportion relationships established in 2016 remained stable throughout the study period. This assumption is common in operational land cover monitoring but represents an important caveat when interpreting multi-year results. Uncertainty in temporal predictions increases with distance from the 2016 reference year; predictions for 2022–2023 carry greater uncertainty than those for 2017–2018. Users should interpret temporal trends as indicative rather than definitive without independent validation.

Table 4. Comparison of sample distribution (count and percent) for the stratified and purposive sampling strategies across 10 PGs of redcedar cover.

	1 [0–0.1)	2 [0.1–0.2)	3 [0.2–0.3)	4 [0.3–0.4)	5 [0.4–0.5)	6 [0.5–0.6)	7 [0.6–0.7)	8 [0.7–0.8)	9 [0.8–0.9)	10 [0.9–1)	Total
Stratified											
Count	45,885	1,774	670	422	408	462	345	202	62	7	50,237
Percent (%)	91.34	3.53	1.33	0.84	0.81	0.92	0.69	0.40	0.12	0.01	100
Purposive											
Count	47,086	2,218	777	589	510	529	452	280	78	9	52,528
Percent (%)	89.63	4.22	1.48	1.12	0.97	1.00	0.86	0.53	0.15	0.02	100

3.5.3. Spatial Autocorrelation Assessment

To estimate potential accuracy inflation from spatial autocorrelation, a supplementary spatially-blocked cross-validation was conducted. The study area was partitioned into 100 m × 100 m grid cells, and cells were randomly assigned to training (70%) or testing (30%) sets, ensuring minimum 100 m separation between training and test observations. Under this design, sample sizes were reduced substantially in higher cover classes: PG1 retained 89% of samples, but PGs 8-10 retained only 34–41% of samples due to spatial clustering of dense stands.

Spatially-blocked validation yielded an overall classification accuracy of 0.76 (compared to 0.84 for random splitting) and regression RMSE of 0.08 (compared to 0.05). These results suggest that random train-test splitting inflates accuracy by approximately 8–10 percentage points for classification and underestimates RMSE by approximately 0.03 fractional cover units. The blocked validation confirms that reported metrics represent upper-bound estimates; true accuracy for spatially independent predictions is likely 8–10% lower.

3.6. Generation of Annual Maps and Temporal Analysis

To ensure temporal consistency, a "train-once, apply-all" approach was adopted. The final RF regression model (purposive sampling) was trained on the 2016 predictors and the 2014–2016 proportional reference dataset. The same trained model was applied to predictor stacks for subsequent years. Two parallel map series were produced:

- Seasonal series: RF regression applied to winter–summer Seasonal composites
- LandTrendr series: The same model applied to LandTrendr-stabilized full-year composites

Each series comprises 10 m fractional-cover maps for 2016–2023. For temporal trend analysis, pixel-level time series were extracted and linear regressions of cover versus year were fitted to estimate annual change rates (slope). Encroached area over time was also summarized for multiple cover thresholds, and temporal behavior of the Seasonal and LandTrendr methods was compared, with emphasis on interannual variability and stability of mid- and high-cover classes.

4. Results

4.1. High-Resolution Reference Map Data Validation

The aggregated 1 m reference map showed strong agreement with the manually digitized validation crowns. Across 100 randomly sampled 10 m pixels, the comparison yielded an overall RMSE of 0.10 and an MAE of 0.08 (Table 5). Positional agreement, measured by IoU, improved with canopy density, increasing from 0.01 in sparse pixels to 0.87 in the densest pixels (Table 6). A 10,000-iteration permutation test confirmed that the observed agreement was highly significant ($p < 0.001$) and not due to chance.

In the most sparsely vegetated group (PG1, 0–10% cover), the reference map indicated a mean redcedar fraction of 0.01, while the digitized samples contained 0.00. Both values correspond to effectively treeless conditions, so the classification of PG1 pixels as near-zero cover is unchanged. Given that PG1 comprises 91.3% of all pixels, the very small bias further indicates that the reference map does not materially overestimate low-density redcedar cover.

Table 5. Validation of the aggregated 1 m reference map against manual digitization, showing pixel distribution, RMSE, and MAE stratified by PGs.

Proportion Group	1 [0–0.1)	2 [0.1–0.2)	3 [0.2–0.3)	4 [0.3–0.4)	5 [0.4–0.5)	6 [0.5–0.6)	7 [0.6–0.7)	8 [0.7–0.8)	9 [0.8–0.9)	10 [0.9–1)	Total
Pixels	57,356	2,218	777	589	510	529	452	280	78	9	62,798
Percent (%)	91.3	3.5	1.2	0.9	0.8	0.8	0.7	0.4	0.1	0.0	100
RMSE	0.03	0.11	0.10	0.14	0.13	0.12	0.10	0.09	0.05	0.05	0.1
MAE	0.01	0.1	0.09	0.12	0.11	0.09	0.08	0.09	0.04	0.05	0.08

Table 6. Positional agreement (intersection-over-union fraction) of the 1 m reference map with manual digitization, stratified by PGs. Percentiles from a 10,000-iteration permutation test confirm the statistical significance of the agreement.

Proportion Group	1 [0–0.1)	2 [0.1–0.2)	3 [0.2–0.3)	4 [0.3–0.4)	5 [0.4–0.5)	6 [0.5–0.6)	7 [0.6–0.7)	8 [0.7–0.8)	9 [0.8–0.9)	10 [0.9–1)
Fraction	0	0.01	0.33	0.30	0.43	0.51	0.55	0.65	0.76	0.87
Percent (%)	50.0	100	100	100	100	100	100	100	100	90.0

Table 7. Spatial autocorrelation of training samples, quantified as neighbor counts within distance thresholds (n = 500 random subset).

Distance (m)	Sentinel-2 Pixels	Mean Neighbors	Min	Max	SD
10	1	2.7	0	4	0.93
30	3	20.8	11	26	2.36
50	5	55.7	28	67	5.69

Sample points exhibited high spatial density (Table 7). On average, each point had 2.7 neighbors within 10 m (one Sentinel-2 pixel), 20.8 neighbors within 30 m, and 55.7 neighbors within 50 m. The minimum neighbor count within 30 m was 11, indicating that no sample was spatially isolated from its neighbors. This clustering reflects the natural spatial aggregation of eastern redcedar stands.

4.2. Classification of Eastern Redcedar Distribution

The binary classification models performed well, with accuracy varying logically by input data. The Sentinel-2-only model achieved $\kappa = 0.67$ (95% CI: 0.63–0.71), while the combined Sentinel-2 + NEON model achieved $\kappa = 0.71$ (95% CI: 0.67–0.75). The overlapping confidence intervals confirm that the apparent 0.04 improvement is not statistically significant ($p = 0.18$, paired t-test across 30 iterations), reinforcing the conclusion that Sentinel-2 alone provides sufficient discriminatory power for this application.

Variable importance rankings for the Sentinel-2-only model consistently identified winter NDVI and the SRI as the most influential predictors, underscoring the discriminatory power of dormant-season imagery for separating evergreen redcedar from leaf-off vegetation.

4.3. Proportional Eastern Redcedar

4.3.1. Model Accuracy

Both sampling frameworks for the 10 m regression model produced highly accurate fractional cover estimates within the study site. The stratified sampling design yielded a test-set RMSE of 0.04 and MAE of 0.02. The purposive design, which oversampled dense stands, resulted in a slightly higher RMSE of 0.05 and MAE of 0.03. Mean bias error was minimal in both cases (−0.0055 for stratified, −0.0001 for purposive), indicating no significant systematic over- or underprediction across the entire study area. However, as discussed in Section 4.1.1, high spatial autocorrelation among training samples means these accuracy estimates likely represent upper bounds.

4.3.2. Spatial Bias and Sampling Effects

Evaluation of test-set performance stratified by PGs showed a consistent pattern for both models: prediction error (RMSE and MAE) increased with the proportion of eastern redcedar cover (Table 8). Fractional cover was thus estimated more accurately in sparse canopies than in dense, established stands (>50% cover).

The purposive sampling strategy, which increased the number of high-cover pixels in the training data, generally outperformed the stratified design. Across PGs 2–10, the purposive model produced lower RMSE and MAE in nearly all cases; for example, RMSE in PG10 was 0.18 for the purposive model compared to 0.21 for the stratified model. This finding suggests that enhancing the representation of rare, high-density stands partially reduced underestimation of cover in those areas. Nevertheless, the purposive sample remained strongly skewed toward PG1, so high-cover conditions were still underrepresented, and errors remained largest in the upper PGs.

Both models performed very well at the lowest end of the cover gradient. In PG1 (0–10% cover), each approach achieved an RMSE of 0.02 and an MAE of 0.01, indicating high precision under sparse-cover conditions that are most relevant for detecting early-stage redcedar encroachment.

The increase in error with increasing cover is largely attributable to class imbalance in the training data (Table 4). PG1 comprises approximately 90% of all samples in both sampling strategies (91.34% for stratified, 89.63% for purposive), providing the model with abundant training examples for low-cover conditions. In contrast, high-cover classes are severely underrepresented: PG10 contains only 7 samples under stratified sampling and 9 under purposive sampling. With so few training examples, the model has limited capacity to learn the spectral-proportion relationships characteristic of dense stands. Although purposive sampling modestly increased representation in PGs 2–9, it could not overcome the fundamental scarcity of high-cover pixels on the landscape. A secondary factor is spectral saturation: as canopy closure approaches 100%, vegetation indices plateau, reducing the model's ability to discriminate among high fractional cover values based on reflectance alone.

4.3.3. Variable Importance

Consistent with the classification model, variable importance analysis for both regression models identified winter NDVI and SRI as the strongest predictors (Figure 5). Variable importance was quantified as the mean decrease in out-of-bag prediction accuracy when each variable was permuted (Breiman, 2001). This result confirms that the spectral separation between evergreen redcedar and dormant deciduous and herbaceous vegetation during winter is the primary driver for accurate proportional mapping. Summer indices (e.g., EVI, IRECI) contributed secondary, complementary information.

Table 8. Test set performance (RMSE and MAE) of the random forest regression models, stratified by PGs, comparing the purposive and stratified sampling strategies.

Proportion Group	1 [0–0.1)	2 [0.1–0.2)	3 [0.2–0.3)	4 [0.3–0.4)	5 [0.4–0.5)	6 [0.5–0.6)	7 [0.6–0.7)	8 [0.7–0.8)	9 [0.8–0.9)	10 [0.9–1)
Stratified										
RMSE	0.02	0.08	0.11	0.11	0.12	0.12	0.13	0.13	0.15	0.21
MAE	0.01	0.07	0.09	0.09	0.09	0.08	0.09	0.10	0.14	0.20
Purposive										
RMSE	0.02	0.07	0.10	0.10	0.10	0.10	0.11	0.12	0.14	0.18
MAE	0.01	0.06	0.08	0.08	0.08	0.07	0.08	0.09	0.12	0.17

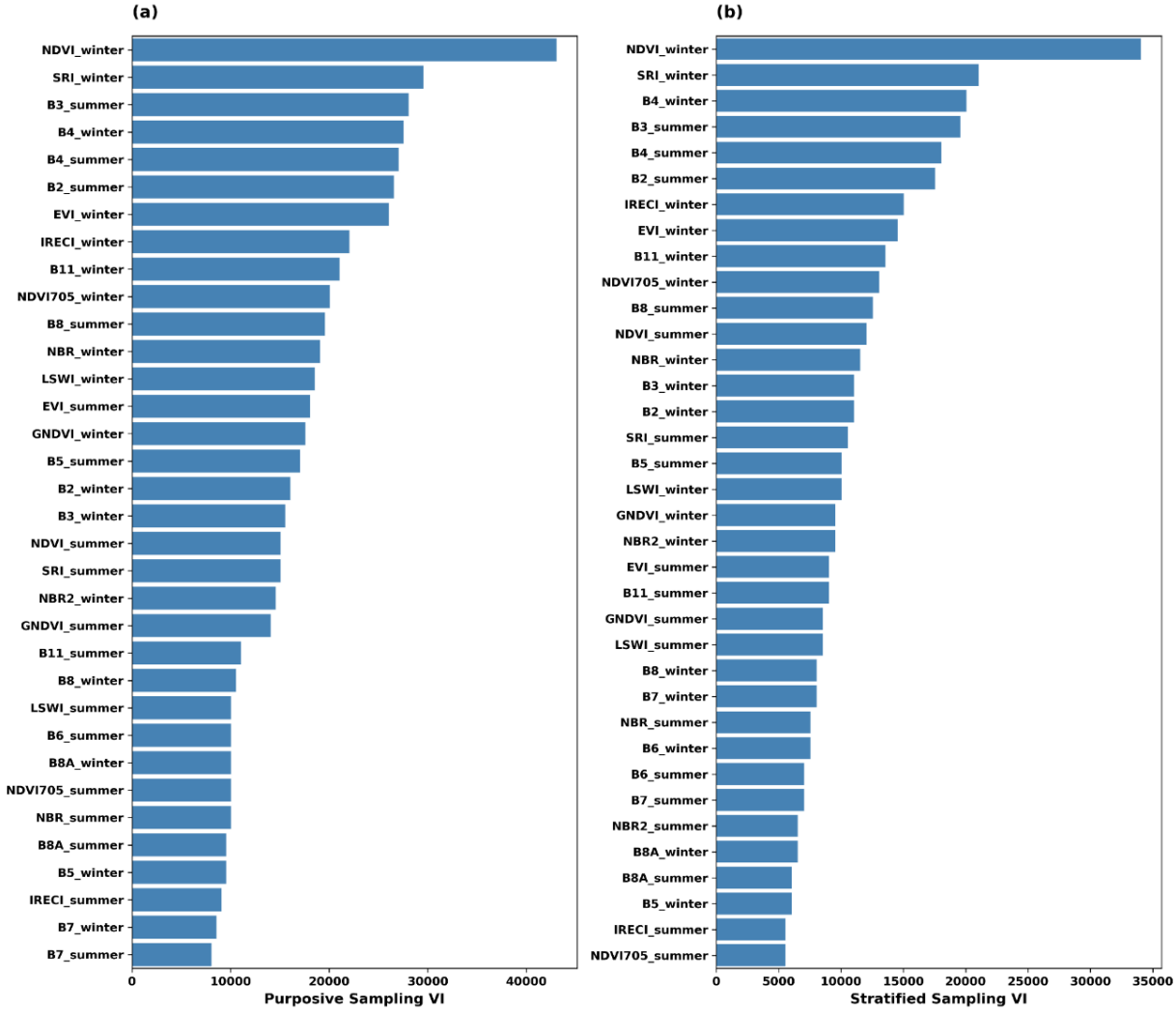

Figure 5. Variable importance rankings for random forest regression models, measured as mean decrease in out-of-bag prediction accuracy. (a) Purposive sampling strategy. (b) Stratified sampling strategy. Predictors are ordered by decreasing importance; winter-season variables (suffix "_winter") dominate the upper ranks in both models, with NDVI_winter and SRI_winter consistently identified as the most influential predictors.

4.4. Comparison of Seasonal and LandTrendr Methods

4.4.1. Pixel-Wise Agreement

Pixel-wise comparison of fractional-cover maps from the Seasonal and LandTrendr workflows showed strong agreement for all years (Figure 6). Correlation coefficients (R) between LandTrendr and Seasonal estimates ranged from 0.946 to 0.991, with the highest agreement in 2016 (R = 0.991) and slightly lower but still strong correlations in subsequent years. RMSE between the two methods was consistently low, ranging from 0.015 in 2016 to 0.055 in 2020, indicating that pixel-level differences between methods were typically small (≤5.5% fractional cover).

Visual inspection of the scatterplots relative to the 1:1 reference line revealed year-specific patterns. In most years (2016, 2018, 2021, 2023), points clustered tightly along the 1:1 line, indicating close agreement across the full cover gradient. In 2022, points fell systematically above the 1:1 line at intermediate and high cover values, indicating that LandTrendr estimates exceeded Seasonal estimates. This pattern suggests that the Seasonal composite for 2022 may have captured lower vegetation vigor—possibly due to drought conditions or unfavorable image timing—while the LandTrendr algorithm, which utilizes the full temporal history, maintained higher, stabilized cover estimates consistent with the long-term trend. Conversely, in 2019 and 2020, points at higher cover values tended to fall slightly below the 1:1 line, indicating modest compression of cover estimates in the LandTrendr series relative to Seasonal composites.

In practical terms, these differences were modest for most pixels (RMSE consistently <0.06), but in years with systematic deviations (e.g., 2022), the LandTrendr map could exceed the Seasonal estimate by approximately 0.10 fractional cover units in the highest-cover classes.

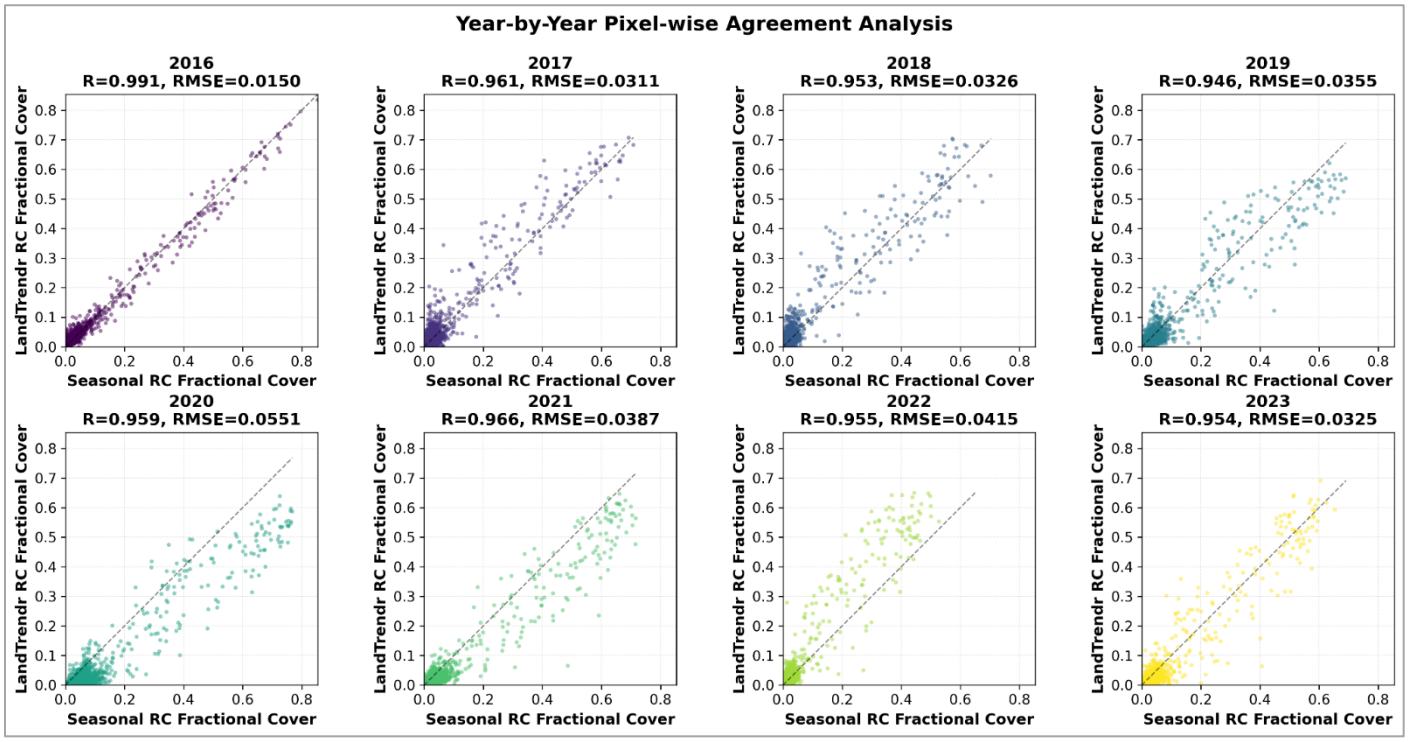

Figure 6. Pixel-wise comparison of fractional redcedar cover between Seasonal (x-axis) and LandTrendr (y-axis) map series for 2016–2023. Each panel represents one year; points are individual 10 m pixels within UKFS. Dashed lines indicate the 1:1 reference line. Correlation coefficients (R) and root mean square errors (RMSE) are shown for each year. Strong correlations (R = 0.946–0.991) and low RMSE values (0.015–0.055) indicate close agreement between methods across all years. Systematic deviations from the 1:1 line are visible in 2022 (LandTrendr > Seasonal) and to a lesser extent in 2019–2020 (LandTrendr < Seasonal at higher cover values).

4.4.2. Encroached Area and Cover-Class Dynamics

For both map series, the total encroached area above five cover thresholds (≥10%, ≥20%, ≥30%, ≥40%, ≥50%) was calculated for each year from 2016 to 2023 (Figures 7a–e). The two methods produced markedly different temporal patterns, with the Seasonal series exhibiting substantially greater interannual variability than the LandTrendr series across all thresholds.

At the ≥10% threshold (Figure 7a), the Seasonal series fluctuated widely between approximately 38 and 66 ha, with pronounced peaks in 2020 and 2023 and a notable dip in 2022. In contrast, the LandTrendr series remained relatively stable, ranging from approximately 41 to 51 ha. Agreement between methods was poor ($r = -0.16$), with a high RMSE of 9.8 ha, reflecting divergent interannual trajectories.

At intermediate thresholds (≥20% and ≥30%; Figures 7b–c), both methods showed moderate positive correlation ($r = 0.30$ and $r = 0.29$, respectively), with RMSE values of 3.9 and 3.5 ha. The Seasonal series continued to display large fluctuations, particularly the sharp increase in 2020 followed by a steep decline in 2022, while the LandTrendr series maintained smoother trajectories with more gradual year-to-year changes.

At higher thresholds (≥40% and ≥50%; Figures 7d–e), agreement between methods deteriorated. For the ≥40% threshold, correlation was negligible ($r = -0.09$) with RMSE of 5.2 ha. For the ≥50% threshold, correlation was negative ($r = -0.44$) with RMSE of 6.2 ha. These correlations are based on 8 annual data points (2016–2023), and the negative correlation at ≥50% cover is driven primarily by the divergent behavior in 2020–2022, when the Seasonal series showed extreme fluctuations while LandTrendr remained stable. The small sample size ($n = 8$ years) limits statistical power to detect true correlation; however, the divergent trajectories themselves are informative regarding method reliability.

The pronounced 2022 dip observed in the Seasonal series across all thresholds suggests systematic underestimation during that year. Review of climate records confirms that 2022 was characterized by moderate drought conditions in northeastern Kansas with drought severity reaching categories D1–D2 during summer months (U.S. Drought Monitor, 2022), which may have reduced evergreen NDVI through drought stress. Additionally, the winter 2021–2022 compositing window coincided with below-average snowfall, potentially altering background reflectance. These conditions, combined with unfavorable image acquisition timing, likely contributed to the anomalous Seasonal estimates.

Mean encroached area across the study period decreased with increasing cover threshold for both methods (Figure 7f). However, the Seasonal series consistently showed higher means accompanied by substantially larger standard deviations, reflecting its greater sensitivity to interannual noise. The LandTrendr series produced lower but more stable estimates, with standard deviations approximately 2–4 times smaller than the Seasonal series depending on threshold. This stability-accuracy tradeoff suggests that the LandTrendr approach may slightly underestimate peak cover conditions while providing more reliable temporal trajectories for trend analysis.

Pixel-wise slopes of fractional cover from 2016–2023 (derived from LandTrendr maps) ranged approximately from −0.008 to +0.008 fractional units yr^{-1} (Figure 8). Most pixels clustered near zero, indicating gradual change.

4.5. Indirect Assessment of Temporal Consistency

In the absence of multi-year ground truth, temporal consistency was evaluated by examining the stability of winter NDVI values for known redcedar and non-redcedar reference areas across the study period (Figure 9). Mean winter NDVI for redcedar pixels remained highly stable from 2016 to 2023 (range: 0.75–0.80; coefficient of variation = 2.8%), indicating that the evergreen spectral signal was consistent across years. Non-redcedar reference areas exhibited slightly greater interannual variation (CV = 10.5%), likely reflecting year-to-year differences in residual green vegetation, soil moisture, or snow cover during the winter compositing window.

The spectral separation between redcedar and non-redcedar land cover remained large and consistent throughout the study period, with redcedar winter NDVI exceeding non-redcedar values by 0.40–0.45 units in all years. This stable separation confirms that the phenological contrast underlying model predictions, the key discriminating feature identified in variable importance analysis, did not degrade over time. The consistency of both absolute NDVI values and class separability suggests that atmospheric correction and sensor calibration did not introduce systematic temporal drift, supporting the application of a single trained model across the full 2016–2023 time series.

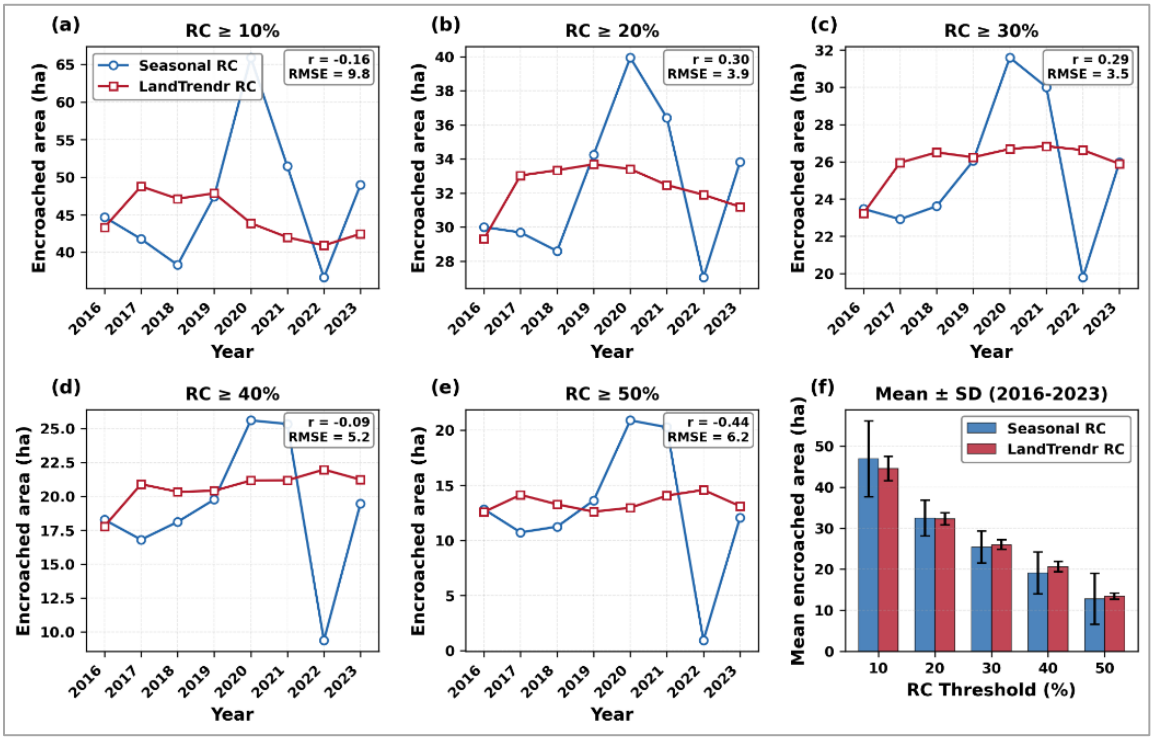

Figure 7. Temporal predictions for 2017–2023 lack independent validation; trends should be interpreted with appropriate caution, particularly for years distant from the 2016 reference period. (a–e) Annual encroached area for thresholds ≥10%, ≥20%, ≥30%, ≥40%, and ≥50%, comparing Seasonal series (blue circles) and LandTrendr-stabilized series (red squares). Correlation coefficients and root mean square errors between the two methods are shown for each threshold. The Seasonal series exhibit high interannual variability, including a pronounced dip in 2022 across all thresholds, while the LandTrendr series produce smoother temporal trajectories. (f) Mean encroached area (±1 standard deviation) across the 2016–2023 period for each threshold, illustrating the greater temporal stability of the LandTrendr series (smaller error bars) compared to the Seasonal series.

5. Discussion

5.1. Data Integration and Model Performance

Classification accuracy derived from multi-seasonal Sentinel-2 data alone (accuracy = 0.84 ± 0.02 SD across 30 iterations) was nearly equivalent to models generated with high-resolution NEON airborne data (accuracy = 0.85 ± 0.02 SD). This 1% difference is within the uncertainty range expected from random variation in train-test splits (overlapping confidence intervals) and should not be interpreted as a meaningful performance gap; rather, both approaches achieved comparable discrimination.

Model fidelity also depends on the reference dataset. The method employed here (intersecting LiDAR-derived canopy height stability with leaf-off spectral classification) proved highly effective (RMSE = 0.10) when validated against manual digitization. This data-driven approach provides a robust and scalable alternative to subjective photointerpretation or spatially sparse field plots, addressing a common bottleneck in fractional cover studies (Filippelli et al., 2020). Together, the conservative LiDAR + leaf-off reference, proportional modeling framework, and LandTrendr-stabilized Sentinel-2 predictors represent a relatively novel combination: they enable annual, 10 m fractional cover products that capture both early invasion and gradual densification using only freely available satellite data.

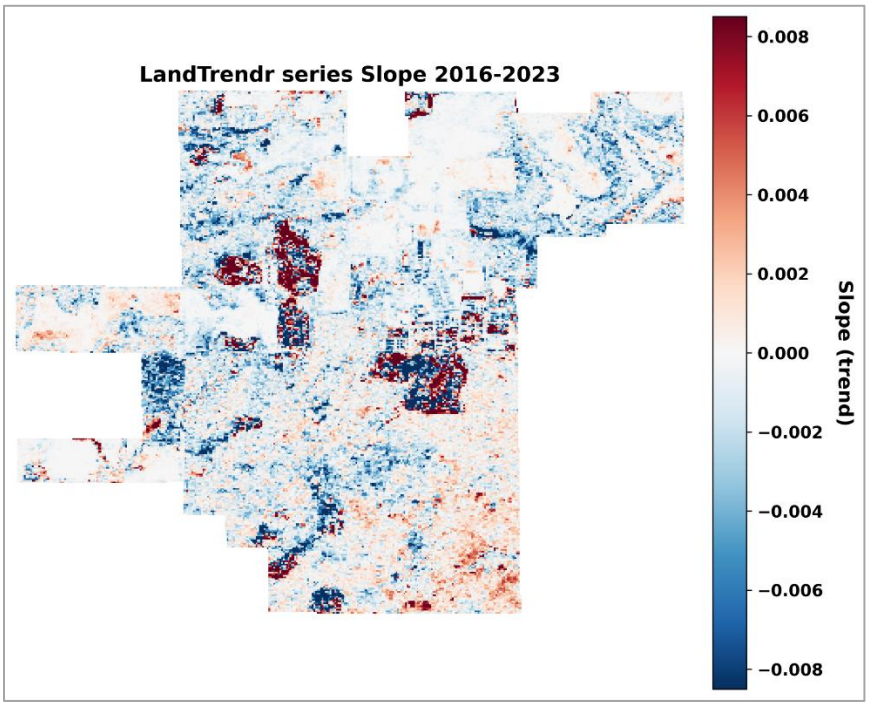

Figure 8. Spatial distribution of pixel-wise temporal trends in fractional redcedar cover (2016–2023) derived from the LandTrendr series. Color scale represents the slope of linear regression (fractional cover yr^{-1}): warm colors (red) indicate increasing cover, cool colors (blue) indicate decreasing cover, and neutral tones (white) indicate stable conditions. Values range from −0.008 to +0.008 yr^{-1}. Localized patches of positive slope marked areas of persistent redcedar increase, often surrounding existing intermediate- to high-cover stands. Negative trends appeared in fewer, discrete locations, consistent with difficulties with shadows or minor pixel misalignments in the time series.

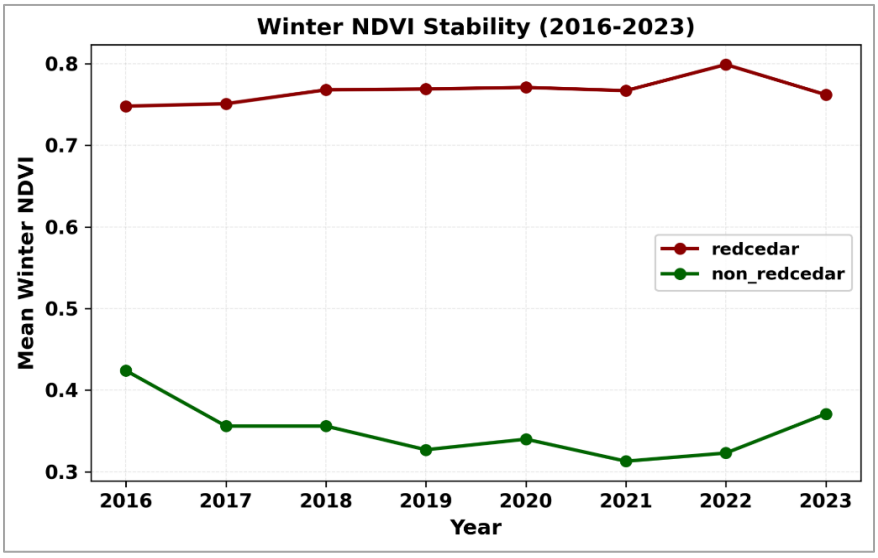

Figure 9. Temporal stability of winter NDVI for redcedar (upper line) and non-redcedar (lower line) reference areas from 2016 to 2023. Error bars represent ±1 standard deviation. The consistent spectral separation (0.40–0.45 NDVI units) between classes across all years supports the temporal transferability of the 2016-trained model. Non-redcedar reference areas exhibited slightly greater interannual variation (CV = 10.5%), likely reflecting year-to-year differences in residual green vegetation, soil moisture, or snow cover during the winter compositing window.

The temporal gap between reference data acquisition (2014–2016) and model application (2016–2023) warrants consideration. Eastern redcedar is a relatively slow-growing species compared to other encroaching woody plants. Consequently, individual tree crowns mapped in the 2014–2016 reference dataset likely persisted with only modest dimensional changes through 2023. This slow growth supports the assumption that spectral-proportion relationships remained relatively stable over the study period.

Two sources of error may arise from this temporal mismatch. First, new redcedar establishment occurring after 2016 would not appear in the reference dataset, meaning the model may underestimate cover in areas experiencing active recruitment. Second, mortality events, whether from drought, ice damage, or management, would create commission errors where the reference indicates cover that no longer exists. Based on published demographic rates (Section 3.3.5), we estimate that 3.5–14 new seedlings per hectare could be established over the 7-year period, with cumulative mortality of <14% for established trees. However, because new seedlings require 5–10 years to reach detectable crown sizes, and mortality of established trees is low, the primary uncertainty lies in detecting recruitment that occurred between 2016 and approximately 2018. We emphasize that these demographic-based estimates do not constitute validated error rates; without independent ground truth for 2017–2023, the actual magnitude of temporal prediction error remains unknown. The indirect evidence of spectral stability (Figure 9) supports, but does not confirm, the assumption of stable spectral-proportion relationships over time.

5.2. Influence of Sampling Strategy

Sampling design strongly influenced model sensitivity across the cover gradient. The purposive strategy, which modestly oversampled dense stands, improved accuracy in high-cover classes without degrading performance in low-cover conditions. In PG10 (90–100% cover), RMSE decreased from 0.21 (stratified) to 0.18 (purposive), while in PG1 both strategies maintained excellent performance (RMSE = 0.02; MAE = 0.01).

Prediction error remained highest in the densest stands (PGs 9–10) regardless of sampling strategy. This persistence of error likely reflects the physical limitations of optical remote sensing; as canopy closure approaches 100%, spectral indices often saturate, making it difficult for the model to distinguish between 80% and 100% cover based on spectral reflectance alone. While purposive sampling mitigates the statistical bias of the algorithm, it cannot fully overcome this physical saturation of the signal.

Given that approximately 91% of the landscape lies in the 0–10% cover class, the purposive design still reflected the dominant low-cover regime but provided better representativeness for rare dense patches. This balance is particularly important for management objectives that span early detection of invasion fronts and assessment of heavy fuel loads in dense stands. Accurate quantification at low cover enables cost-effective interventions such as prescribed fire before canopy closure reduces treatment efficacy (Bidwell et al., 2017).

A methodological limitation is the high spatial autocorrelation among training samples. Neighbor analysis revealed that every sample had at least 11 other samples within 30 m and 28 samples within 50 m, with no spatially isolated observations. Under random 70/30 train-test splitting, test pixels are therefore highly likely to have spectrally similar training pixels nearby, potentially inflating accuracy estimates relative to truly independent validation. This limitation is inherent to the sampling design, which targeted all available redcedar and non-redcedar pixels within the study area to maximize coverage of the proportional gradient. Spatially blocked cross-validation, in which training and test sets are separated by distances exceeding the autocorrelation range, would provide more conservative accuracy estimates but was not implemented because separating sets by >50 m would have severely reduced effective sample sizes in higher cover classes. Future studies with larger spatial extents should prioritize spatially stratified validation designs.

The reported accuracy metrics (overall accuracy = 0.84; RMSE = 0.04–0.05) should therefore be interpreted as upper bounds on model performance. Applications to new locations within similar landscapes may achieve lower accuracy until local calibration data become available. Independent validation at new sites or using spatially stratified designs would likely yield lower accuracy estimates. However, for operational monitoring within landscapes similar to UKFS, where the goal is to track relative changes rather than produce spatially transferable predictions, the model remains useful despite this limitation.

5.3. Temporal Consistency and Compositing Strategy

In our implementation, LandTrendr operates on the predictor time series (pre-classification smoothing) rather than on post-classification maps, which reduces interannual noise while preserving the spatial detail and interpretability of the fractional cover products. Pixel-wise agreement between Seasonal and LandTrendr

maps was consistently high (R = 0.946–0.991; RMSE = 0.015–0.055), indicating that both methods captured similar spatial patterns of redcedar distribution within individual years. However, their temporal trajectories diverged substantially, particularly at higher cover thresholds.

The Seasonal series, which rely on two relatively narrow temporal windows, were highly sensitive to interannual variability. This sensitivity was most apparent in the 2022 anomaly, where encroached area estimates dropped dramatically across all thresholds, with the ≥50% class declining from approximately 21 ha in 2020 to less than 2 ha in 2022. Such extreme year-to-year fluctuations are inconsistent with the slow growth dynamics of eastern redcedar (typical crown expansion of 10–20 cm yr^{-1}) and almost certainly reflect methodological artifacts rather than true ecological change. Potential causes include drought-induced stress reducing evergreen NDVI, unfavorable timing of cloud-free image acquisition, or residual atmospheric effects not fully corrected in Level-2A processing.

The LandTrendr-stabilized series buffered against such single-year anomalies by incorporating the full temporal history when fitting spectral trajectories. The resulting maps showed substantially lower interannual variability, with standard deviations 2–4 times smaller than the Seasonal series across all cover thresholds (Figure 7f). This temporal stability comes with a potential tradeoff: the LandTrendr parameterization used here (specifically preventOneYearRecovery: true) may dampen or delay detection of rapid, real changes such as mechanical clearing or high-severity fire.

From a management perspective, the choice between methods depends on the monitoring objective. Seasonal composites may be preferable when the goal is to detect recent, abrupt treatment events because they respond more directly to sudden spectral changes—though users must recognize that apparent changes may reflect noise rather than true disturbance. The LandTrendr series is better suited for tracking longer-term encroachment and assessing gradual densification, where suppressing short-term noise improves the reliability of temporal comparisons.

The poor to negative correlations between methods at higher cover thresholds ($r = -0.09$ for ≥40%; $r = -0.44$ for ≥50%) warrant particular attention. These results indicate that the two methods not only differ in magnitude but sometimes show opposite interannual patterns in dense stands. This divergence likely reflects the combination of spectral saturation at high cover (reducing signal-to-noise ratio) and the amplification of year-specific artifacts in the Seasonal series. For applications requiring accurate assessment of dense stand dynamics, neither method may be fully reliable without additional validation data.

5.4. Management Implications, Limitations, and Transferability

The workflow presented here provides land managers with annual, wall-to-wall fractional-cover maps that can be used to identify early invasion fronts and map fuel loads. However, several limitations should be considered when interpreting results and applying the methods elsewhere.

Despite purposive sampling, the RF regression model still slightly underestimated fractional cover in the densest stands due to spectral saturation as canopy closure approaches 100%. Additionally, while Sentinel-2 Level-2A surface reflectance products were used to minimize atmospheric influence, subtleties such as sensor drift, slight variations in solar geometry, or interannual atmospheric anomalies not fully corrected by the Level-2A processing could introduce systematic bias when applying a fixed 2016 model to subsequent years.

A methodological limitation is the high spatial autocorrelation among training samples (Table 7), with every sample having at least 11 neighbors within 30 m and an average of 55.7 neighbors within 50 m. Under random train-test splitting, test pixels are therefore highly likely to have spectrally similar training pixels nearby, potentially inflating accuracy estimates relative to truly independent validation. The reported accuracy metrics (overall accuracy = 0.84; RMSE = 0.04–0.05) should therefore be interpreted as upper bounds on model performance. Supplementary spatially-blocked cross-validation, which enforced 100 m minimum separation between training and test sets, yielded accuracy of 0.76 and RMSE of 0.08, suggesting that spatial autocorrelation inflates accuracy estimates by approximately 8–10 percentage points. Independent validation at new sites would likely yield results closer to these conservative estimates.

The assumption that spectral–proportion relationships captured in the 2016 model remained stable through 2023 is only partially supported by the slow growth rate of eastern redcedar. Pixels with substantial new recruitment or mortality after 2016 may exhibit prediction errors not captured in the accuracy assessment. For operational monitoring within landscapes similar to UKFS, where the goal is to track relative changes rather than produce spatially transferable predictions, the model remains useful despite this limitation.

This study was conducted at a single site (UKFS, approximately 590 ha) characterized by conditions favorable for spectral discrimination: a single dominant evergreen encroacher, pronounced winter dormancy

creating strong phenological contrast, and gentle topography. Successful transfer to other regions depends primarily on three factors: (1) species composition: the method performs best where eastern redcedar is the sole or dominant evergreen, requiring modification where multiple evergreen species co-occur; (2) phenological contrast: strong performance of winter NDVI requires cold winters that induce complete deciduous dormancy, limiting applicability in milder climates or areas with significant cool-season grass cover; and (3) reference data availability: replication requires multi-temporal LiDAR or alternative high-resolution training data. The LandTrendr parameterization was tuned for gradual encroachment and may require adjustment in landscapes with active management or frequent disturbance.

Operational deployment of this workflow requires three core data components: (1) multi-temporal LiDAR or alternative high-resolution structural data capable of distinguishing evergreen from deciduous canopy through seasonal height stability; (2) leaf-off aerial imagery with sufficient spectral resolution to identify conifer crowns against dormant backgrounds; and (3) Sentinel-2 time series spanning at least two full years to establish seasonal composites. Where LiDAR is unavailable, alternative reference data strategies include manual digitization of redcedar crowns from leaf-off imagery, field plot networks with GPS-located canopy cover measurements, or existing vegetation maps validated against current conditions. The computational workflow is implemented entirely in Google Earth Engine, requiring no specialized hardware, but local expertise in image interpretation and random forest parameterization is essential for successful adaptation.

Google Earth Engine memory constraints necessitated subsampling of the full 62,798-pixel reference dataset to approximately 50,000–52,000 pixels for model training (Table 4), as well as batch processing of annual composites rather than simultaneous multi-year computation. These limitations did not affect model performance, as subsampling preserved proportional representation across cover classes, but they underscore practical considerations for scaling this workflow. Applications covering larger spatial extents or longer time series should anticipate similar computational bottlenecks and implement appropriate tiling schemes, geographic subsetting, or sequential year-by-year processing to maintain tractability within cloud-based platforms.

The degradation of model performance at high cover fractions (RMSE increasing from 0.02 at 0–10% cover to 0.18 at 90–100% cover) has important implications for fuel load assessment and fire risk mapping. Dense redcedar stands (>50% cover) represent the highest fuel loads and greatest wildfire risk, yet these are precisely the conditions where model accuracy is lowest. For fuel load mapping applications, we recommend treating model predictions above 50% cover as qualitative indicators of "high density" rather than precise fractional estimates. Management decisions requiring accurate quantification of dense stand cover—such as biomass estimation for carbon accounting or precise fuel load calculations—should incorporate field verification or alternative methods (e.g., LiDAR-derived canopy metrics) for stands predicted to exceed 50% cover. The model remains well-suited for its primary intended purpose: detecting early-stage invasion (<30% cover) when management interventions are most cost-effective.

Future work could incorporate additional training data from more heavily infested areas or integrate structural information from Sentinel-1 SAR or forthcoming spaceborne LiDAR missions to address saturation in dense stands. Extending the analysis with multi-temporal LiDAR and additional field validation would further refine model calibration and test the temporal stability assumption explicitly. Despite the limitations noted above, the general framework—combining phenologically targeted Sentinel-2 composites, proportion-aware sampling, and optional temporal smoothing—provides a transferable conceptual approach for mapping woody encroachment. The reliance on globally available Sentinel-2 data and open-source processing tools (Google Earth Engine) facilitates adaptation to other regions, though future applications should prioritize local validation and recalibration of random forest models using region-specific reference data.

6. Conclusions

This study presents a reproducible method to map eastern redcedar fractional cover in a prairie–forest transition zone using Sentinel-2 time series and random forest models. A LiDAR- and imagery-based reference map, combined with proportion-aware sampling, produced accurate 10 m fractional-cover estimates across the full 0–100% range. Sentinel-2 imagery alone, with winter–summer compositing, yielded classification accuracy comparable to models that also used airborne hyperspectral data, and the regression model showed low error, particularly in low-cover conditions most relevant for early encroachment detection.

Purposive sampling improved performance in dense stands without degrading accuracy in the dominant low-cover class.

LandTrendr-based temporal smoothing reduced interannual spectral noise relative to simple seasonal compositing and produced more stable trends in intermediate and high cover classes. The resulting annual maps support identification of emerging redcedar patches, assessment of canopy densification, and evaluation of management outcomes through time.

Several limitations warrant consideration when interpreting results. High spatial autocorrelation among training samples means that reported accuracy metrics (overall accuracy = 0.84; RMSE = 0.04–0.05) likely represent upper-bound estimates that may not generalize to independent locations. The absence of ground truth data after 2016 precludes direct validation of temporal predictions; although indirect evidence supports model stability, accuracy for 2017–2023 remains unverified. Additionally, the workflow was developed and validated at a single site characterized by a dominant evergreen encroacher, pronounced winter dormancy, and gentle topography, conditions that favor spectral discrimination but may not hold in all landscapes. Transferability to regions with multiple evergreen species, milder winters, or different soil and vegetation backgrounds will require local calibration and validation.

The method relies on globally available satellite data and standard processing tools, providing a template that could be adapted to other regions affected by woody encroachment. However, direct application to new landscapes would require local calibration to account for differences in evergreen species composition, deciduous forest phenology, and background land cover. The workflow is most readily transferable to sites within the central Great Plains where eastern redcedar is the dominant encroaching evergreen and where winter dormancy creates strong phenological contrast with surrounding vegetation.

Funding: None

Data Availability Statement: Data will be made available upon request to the corresponding author.

Acknowledgement: The authors thank Dana Peterson, Xingong Li, and Jude H. Kastens (University of Kansas) for their contributions to this work.

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

References

Archer, S. R., Andersen, E. M., Predick, K. I., Schwinning, S., Steidl, R. J., & Woods, S. R. (2017). Woody plant encroachment: Causes and consequences. In D. D. Briske (Ed.), *Rangeland systems: Processes, management and challenges* (pp. 25–84). Springer. https://doi.org/10.1007/978-3-319-46709-2_2

Bidwell, T. G., Weir, J. R., & Engle, D. M. (2017). *Eastern redcedar control and management – best management practices to restore Oklahoma's ecosystems* (NREM-2876). Oklahoma Cooperative Extension Service.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Briggs, J. M., Hoch, G. A., & Johnson, L. C. (2002). Assessing the rate, mechanisms, and consequences of the conversion of tallgrass prairie to *Juniperus virginiana* forest. *Ecosystems, 5*, 578–586. https://doi.org/10.1007/s10021-002-0187-4

Burchfield, D. (2017). *The use of remotely sensed LiDAR and multispectral imagery for modeling eastern redcedar biomass within north eastern Kansas* [Master's thesis, Kansas State University]

Filippelli, S. K., Falkowski, M. J., Hudak, A. T., Fekety, P. A., Vogeler, J. C., Khalyani, A. H., Billmire, M. G., & Smith, A. M. S. (2020). Monitoring pinyon-juniper cover and aboveground biomass across the Great Basin. *Environmental Research Letters, 15*(2), 025004. https://doi.org/10.1088/1748-9326/ab6785

Galgamuwa, G. P., Oduor, P. G., & Dozier, J. (2020). Expansion of Eastern Redcedar (Juniperus virginiana L.) into the Deciduous Woodlands within the Forest–Prairie Ecotone of Kansas. *Forests, 11*(2), 154. https://doi.org/10.3390/f11020154

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Hoff, D. L., Will, R. E., Zou, C. B., & Lillie, N. D. (2018). Encroachment dynamics of *Juniperus virginiana* L. and mesic hardwood species into Cross Timbers forests of north-central Oklahoma, USA. *Forests, 9*(2), 75. https://doi.org/10.3390/f9020075

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). 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. https://doi.org/10.1016/j.rse.2010.07.008

Kennedy, R. E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W. B., & Healey, S. (2018). Implementation of the LandTrendr algorithm on Google Earth Engine. *Remote Sensing, 10*(5), 691. https://doi.org/10.3390/rs10050691

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., & Gascon, F. (2017). Sen2Cor for Sentinel-2. *Proceedings of SPIE: Vol. 10427. Image and Signal Processing for Remote Sensing XXIII*, 1042704. https://doi.org/10.1117/12.2278218

Maxwell, A. E., Warner, T. A., & Fang, F. (2017). Land cover classification and feature extraction from National Agriculture Imagery Program (NAIP) orthoimagery: A review. *Photogrammetric Engineering & Remote Sensing, 83*(10), 737–747. https://doi.org/10.14358/PERS.83.10.737

McKinley, D. C., & Blair, J. M. (2008). Woody plant encroachment by *Juniperus virginiana* in a mesic native grassland promotes rapid carbon and nitrogen accrual. *Ecosystems, 11*, 454–468. https://doi.org/10.1007/s10021-008-9133-4

Moser, W. K., Hansen, M. H., Atchison, R. L., Butler, B. J., Crocker, S. J., Domke, G. M., Kurtz, C. M., Miles, P. D., Piva, R. J., & Woodall, C. W. (2013). *Kansas' forests 2010* (Resource Bulletin NRS-85). U.S. Department of Agriculture, Forest Service, Northern Research Station.

Pasquarella, V. J., Arévalo, P., Bratley, K. H., Bullock, E. L., Gorelick, N., Yang, Z., & Kennedy, R. E. (2022). Demystifying LandTrendr and CCDC temporal segmentation. *International Journal of Applied Earth Observation and Geoinformation, 112*, Article 102806. https://doi.org/10.1016/j.jag.2022.102806

Poznanovic, A. J., Falkowski, M. J., MacLean, A. L., Smith, A. M. S., & Evans, J. S. (2014). An accuracy assessment of tree detection algorithms in juniper woodlands. *Photogrammetric Engineering & Remote Sensing, 80*(7), 627–637. https://doi.org/10.14358/PERS.80.7.627

Lawson, E. R. (1990). *Juniperus virginiana* L. Eastern redcedar. In R. M. Burns & B. H. Honkala (Eds.), *Silvics of North America: Vol. 1. Conifers* (Agriculture Handbook 654, pp. 131–140). USDA Forest Service.

U.S. Drought Monitor. (2022). *Drought conditions for Kansas*. National Drought Mitigation Center, University of Nebraska-Lincoln. https://droughtmonitor.unl.edu/

U.S. Geological Survey. (2019). *National Hydrography Dataset (NHD)*. U.S. Department of the Interior. https://www.usgs.gov/national-hydrography/national-hydrography-dataset

Wang, J., Xiao, X., Qin, Y., Doughty, R. B., Dong, J., & Zou, Z. (2018). Characterizing the encroachment of juniper forests into sub-humid and semi-arid prairies from 1984 to 2010 using PALSAR and Landsat data. *Remote Sensing of Environment, 205*, 166–179. https://doi.org/10.1016/j.rse.2017.11.019

Zou, C. B., Turton, D. J., Will, R. E., Engle, D. M., & Fuhlendorf, S. D. (2014). Alteration of hydrological processes and streamflow with juniper (*Juniperus virginiana*) encroachment in a mesic grassland catchment. *Hydrological Processes, 28*(26), 6173–6182. https://doi.org/10.1002/hyp.10102