DEVELOP Technical Report

Northern Rockies Ecological Conservation II

Determining the Distribution of Whitebark Pine in the Intermountain West Through Spectral Signature Classification to Assess Forest Health Within the Region

Spring 2025 | Idaho – Pocatello
April 4th, 2025

Authors: Dustin Corbridge, Nicholas Cramer, Genevieve Harman, Benjamin Shostak, Anna Thario (Analytical Mechanics Associates)

Abstract: Whitebark pine (*Pinus albicaulis; WBP*) plays a critical ecological role in the Northern Rockies and was listed as a Federally Threatened species in 2023. WBP decline due to disease and insect infestations has been further exacerbated by climate change, requiring improved monitoring tools for natural resource managers. Previous models of WBP distribution lacked confidence, limiting their use in conservation planning. This project partnered with and received invaluable in-situ tree location data from the National Park Service, the United States Forest Service, the Bureau of Land Management, the U.S. Fish and Wildlife Service, the Whitebark Pine Ecosystem Foundation, and the Yellowstone Club in order to assess the feasibility of using earth observation data to classify and map WBP with high accuracy. Our team processed Landsat 8 Operational Land Imager, Landsat 9 Operational Land Imager-2, Sentinel-2 Multispectral Instrument, and digital elevation data to extract spectral signatures and classify tree species using a random forest model trained on our in-situ tree location data. The Sentinel-2 model was used to to classify WBP areas compared to limber pine (*Pinus flexilis*), other non-target tree species, and non-forested land. The WPB classification showed a low error of omission (3.93%) and commission of (4.69%), plus a high Kappa Index of Agreement of 0.93 for the WBP class compared with reference data. The final maps provide a tool to assess WBP distribution and identify areas for future data collection across the Intermountain West. Our team also produced an Esri Field Maps template for our partners for aiding their field surveys and WBP health monitoring efforts. Limitations included difficulty with classifying understory vegetation, leading to over-prediction of WBP in some areas. This feasibility study demonstrated that NASA Earth observations can support long-term WBP monitoring and distribution assessment across the Intermountain West, but more reference data is needed to further understand the results.

Key Terms: Whitebark pine, forest ecology, habitat suitability, Northern Rockies, Intermountain West, spectral signatures, Landsat 8 OLI and 9 OLI-2, Sentinel-2 MSI

Advisors: Keith Weber (Idaho State University GIS Training and Research Center), Joe Spruce (Analytical Mechanics Associates), Kait Lemon (NASA DEVELOP)

Lead: Isaac Goldings (Idaho – Pocatello)

Previous Contributors: Hannah Rogers, Dustin Corbridge, A. H. M. Mainul Islam, Joshua Daniel Carrell

1. Introduction

Whitebark pine (WBP; *Pinus albicaulis*; PIAL) is a keystone species in subalpine ecosystems in mountain ranges in the western U.S. and Canada (Burns & Honkala, 1990; Keane et al., 2017). Its unique physiological adaptations allow it to thrive in harsh conditions such as rocky soils, high winds, steep slopes, and prolonged snow cover. WBP provides habitat and key nutrients for wildlife, regulates water availability by retaining snow and slowing snow melt, improves water quality, reduces erosion, enhances biodiversity, and impacts recreational activities (Burns & Honkala, 1990; Keane et al., 2017). However, WBP populations throughout their range have declined by approximately 51% due to multiple factors (Goeking & Izlar, 2018). Widespread WBP population decline prompted its designation as a Federally Threatened species in 2023 (U.S. Fish and Wildlife, 2022). Factors contributing to WBP decline include mountain pine beetle (*Dendroctonus ponderosae*) outbreaks, the spread of non-native white pine blister rust (*Cronartium ribicola*), and historic wildland fire suppression and ongoing wildland fire activity, all of which are increasingly exacerbated by climate change (Tomback & Achuff, 2010; Buotte et al., 2016).

In response to these pressing concerns, DEVELOP and our project partners collaborated to work toward conserving WBP populations. In Summer 2024 and Spring 2025, the terms 1 and 2 teams, respectively, partnered with the Greater Yellowstone Inventory and Monitoring Network of the National Park Service, U.S. Forest Service Region 1, Bureau of Land Management Salmon Field Office, U.S. Fish and Wildlife Service Montana and Wyoming Ecological Services Field Offices, the Whitebark Pine Ecosystem Foundation, and the Yellowstone Club Community Foundation. All collaborators are focused on restoring WBP populations, monitoring their health, and mitigating threats such as pests and wildfire.

The remote, mountainous habitat WBP grows in makes its restoration a more difficult task. Additionally, WBP and limber pine (*Pinus flexilis*) overlap for parts of their range and share similar morphological characteristics, which further complicates restoration efforts. Research suggests synoptic remote sensing is a valuable tool for monitoring and managing WBP populations. For example, spectral signature analysis using Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data was used to effectively identify vegetation types to provide spectral signature data for WBP in Yellowstone National Park (Kokaly et al., 2003). Similarly, Landenburger et al. (2008) employed boosted classification trees and Landsat 7 Enhanced Thematic Mapper Plus (Landsat 7 ETM+) data to map WBP locations in the Greater Yellowstone Ecosystem, providing a validated remote sensing methodology for identifying WBP stands. More recently, Jenkins et al. (2022) developed a GIS-based restoration strategy for WBP management, using spatial data to identify priority conservation areas and guide restoration efforts.

Standardized remote sensing methodologies and sufficiently high-quality satellite data aid partners with identifying, monitoring, and managing WBP populations across land ownership boundaries. Our project includes a satellite-based spectral signature classification framework that enhances WBP mapping and distribution modeling across our study area. This term, we expanded upon the initial distribution model developed by the 2024 team (Rogers et al., 2024), utilizing more training data collected by the 2024 team and partners to improve accuracy. Additionally, we incorporated spectral signatures of non-target tree species into our classification framework. Using our expanded in-situ dataset of target species while refining the spectral separability of non-target species to improve species-level mapping precision in our distribution model. Lastly, we developed a Field Maps template to standardize WBP data collection across all management entities.

The study area encompasses 123,328 km² of area containing parts of the Northern Rockies, Middle Rockies and Idaho Batholith Ecoregions (Omernik, 1987). These ecoregions span parts of Idaho, Montana, and Wyoming (Figure 1). This area was chosen due to its relatively high concentration of suitable WBP habitat. Temporally, the study period spans from 4/1/2024 to 9/1/2024 to match data collected by the 2024 team, reduce possible snow cover, and follow the yearly growing season. Our project used surface reflectance data from Landsat 8 Operational Land Imager (OLI) & Landsat 9 OLI-2 and top of atmosphere (TOA) reflectance from Sentinel-2 Multispectral Instrument (MSI)

Our project used satellite imagery from Landsat 8 Operational Land Imager (OLI), Landsat 9 OLI-2, and Sentinel-2 Multispectral Instrument (MSI).

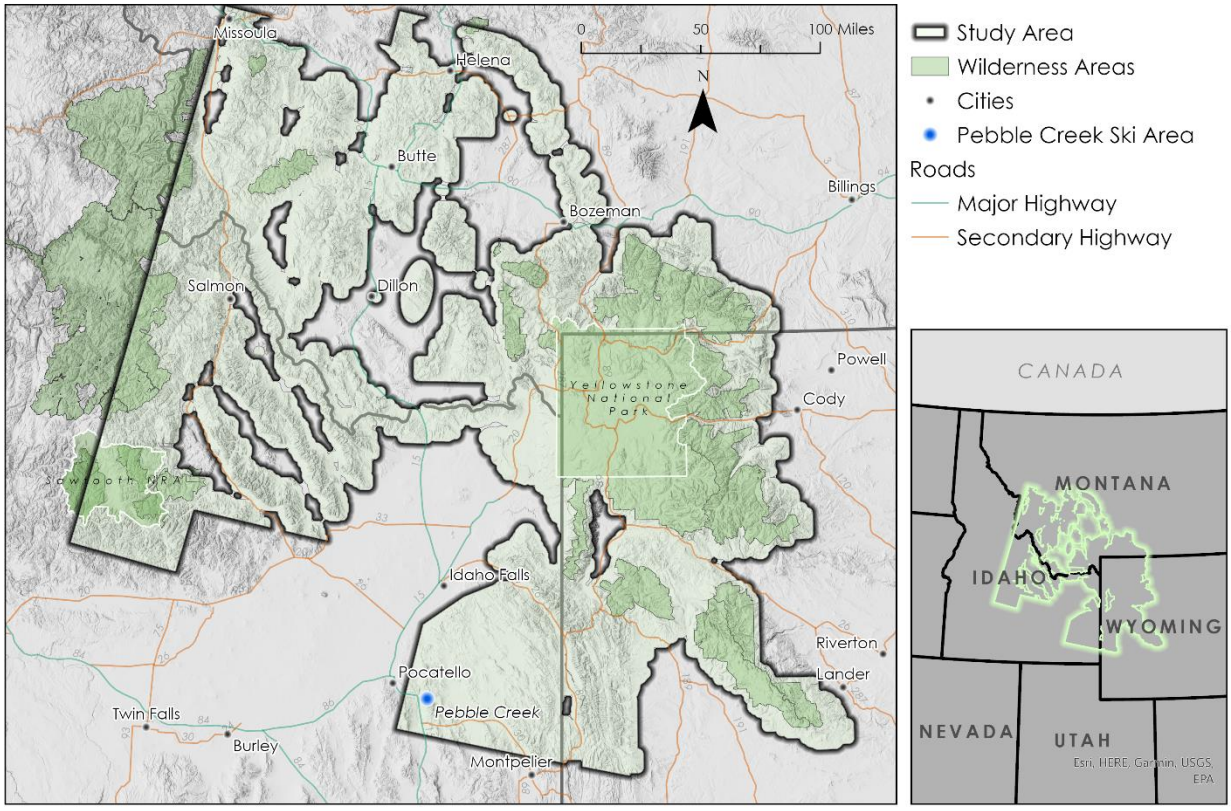

Figure 1. Map depicting spatial extent of the study area (Basemap Credit: Esri, NASA, USFS, NPS).

2. Methodology

2.1 Data Acquisition

We acquired multispectral satellite imagery from three earth observation satellites. All available Landsat 8 OLI and Landsat 9 OLI-2 Collection 2 Level 2 scenes from April 1st to September 1st, 2024, that overlapped our study area and with cloud cover less than 10% were downloaded through United States Geological Survey (USGS) Earth Explorer (2025) (Table A1, Appendix A). All available Sentinel-2 MSI Level 2A scenes were retrieved from the ESA SciHub (Copernicus Open Access Hub), which provided an OData-based REST API, and followed the same query parameters as Landsat. A 10meter digital terrain model (DTM) was acquired from the NASA RECOVER project, derived from USGS, National Elevation Dataset (NED) 3DEP program (Table A1).

The 2024 team, Whitebark Pine Ecosystem Foundation, Yellowstone Club and other researchers collected ground-validated field data throughout the field season of 2024 using similar methodology. Field collected WBP and limber pine samples were identified using dichotomous keys and were double-checked by Chris Earle, Conifer and Wildlife Biologist from The Gymnosperm Database. Trees, along with optional field site notes and site photos, were recorded as geographic point data in Field Maps using WGS 1984 Web Mercator (auxiliary sphere). Geographic coordinates were recorded using either cell phones within Field Maps, or a Trimble R1 GNSS receiver, to achieve 1 – 5-meter positional accuracy. No additional metadata was attached.

Points were collected in Idaho at Pebble Creek Ski Area and the Sawtooth National Forest, and in Montana at the Yellowstone Club (Figure 2).

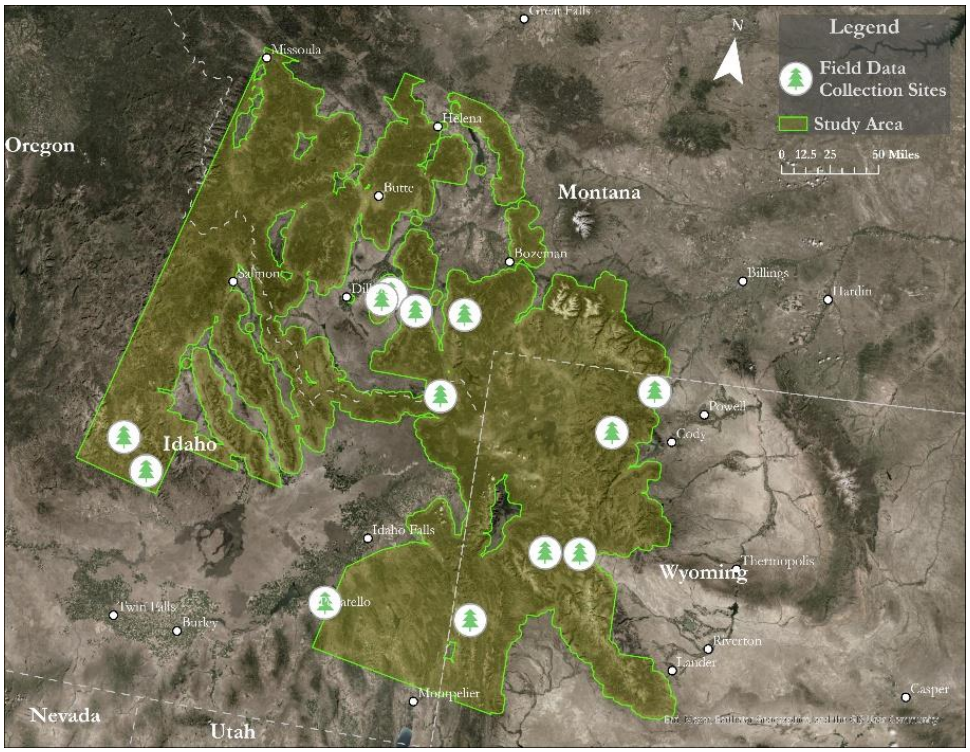

Figure 2. Map of in-situ field data collection locations (Basemap Credit: Esri, Maxar, Earthstar Geographics, and the GIS User Community).

Data from USFS Region One followed Region One adapted National Common Stand Exam methods. Points used were largely mixed stand data, which provided valuable spectral signature data for other timber species like Douglas-fir (*Pseudotsuga menziesii*) and lodgepole pine (*Pinus contorta*). A detailed review of point collection methods is outlined in the *Region 1 Common Stand Exam and Inventory and Monitoring Field Guide* (U.S. Forest Service Region One, 2019).

Data from the Interagency Whitebark Pine Monitoring Program included points for WBP and limber pine from their long-term monitoring program. A detailed review of point collection methods is outlined in the *Interagency Whitebark Pine Monitoring Protocol for the Greater Yellowstone Ecosystem, Version 1.1* (Greater Yellowstone Whitebark Pine Monitoring Working Group, 2011).

2.2 Data Processing

We processed all 78 available Landsat 8 OLI, and Landsat 9 OLI-2 Collection 2 Level 2 scenes in the ArcGIS Pro 3.4.0 ArcPy interface to remove suspected clouds and cloud-shadows using the Landsat QA Band. We applied a scaling factor and additive offset to the data to convert them from 16-bit integer depth to surface reflectance as a floating point fraction varying from zero to one and removed any erroneous pixels with values outside the valid range of the data type. Cloud free scenes were subsequently composited together using median pixel value to create a single cloud-free composite for the entire study area. We extracted each individual band from our cloud-free composite into its respective raster to facilitate importing into TerrSet liberaGIS (version 20.0.0), an open-source remote sensing software.

All Sentinel-2 data was processed on-the-fly using the API library sentinelhub-py in SciHub. All 100+ sentinel scenes used a scene classification layer (SCL) for cloud masking from cloud medium probability, cloud high

probability and cirrus *([8, 9, 10].indexOf(sample.SCL)! == -1)*. Additionally, we used the Normalized Difference Snow Index (NDSI) (ndsi >= 0.4) to mask out pixels where snow would be highly probable during the months when snow was still present in high-altitude biomes. Scenes were subsequently composited using median pixel value and exported as tiles for further processing in ArcGIS Pro. In ArcGIS pro, tiles were first clipped to the study area and masked for no data values (nan). We subsequently mosaicked the scenes together and extracted each band into a single raster for further import into TerrSet. Then, our 10-meter DTM acquired from the NASA RECOVER project was resampled to match the spatial extent and pixel grid of our Sentinel-2 data. Additionally, our 10-meter DTM was resampled to 30 meters, matching the spatial extent, resolution, and pixel grid of Landsat data.

2.2.2 Deriving Ancillary Datasets

We derived the Normalized Difference Vegetation Index (NDVI; Equation 1; Kriegler et al., 1969), Enhanced Vegetation Index 2 (EVI2; Equation 2; Jiang et al., 2007), and Normalized Difference Moisture Index (NDMI; Equation 3; Jin & Sader, 2005) vegetation indices in our Landsat and Sentinel spectral bands using the raster calculator in ArcGIS Pro. In these equations, NIR is the near infrared band's surface reflectance, RED is the red band's surface reflectance, and SWIR1 is the first shortwave infrared band's surface reflectance. We conducted a principal component analysis (PCA) using the PCA module in TerrSet to determine the relative importance of each spectral band by measuring how much it contributed to the overall variation in the data. PCA components were subsequently utilized as inputs to our decision forest analysis. We derived slope and aspect from our 10-meter resampled DTM for Landsat and Sentinel-2 data by inputting the DTM into the "SURFACE" module in TerrSet.

$$\text{NDVI} = \frac{(NIR - RED)}{(NIR + RED)} \tag{1}$$

$$\text{EVI2} = 2.4 \cdot \frac{(NIR - RED)}{(NIR + RED + 1)} \tag{2}$$

$$\text{NDMI} = \frac{(NIR - SWIR1)}{(NIR + SWIR1)} \tag{3}$$

2.2.3 Preprocessing and Data Cleaning Ground Sampled Data

Field data points were overlaid with a Landsat NDVI composite to determine the extent of each pixel in ArcGIS Pro. All points that fell into the same pixel were screened based on percentage of canopy filling each pixel and stand homogeneity to isolate unique spectral signatures for each species of interest. The canopy cover was visually assessed using pictures taken during in-situ data collection and through comparison of very high resolution Esri World Imagery Basemap (Esri, 2025a) imagery with the pixel boundary of the NDVI layer. Pixels with 50% or greater canopy of either WBP (PIAL), limber pine (PIFL2), Douglas-fir (PSME), lodgepole (PICO) or non-forest ("bare ground") were selected as classes used for modeling (See Table B2 in Appendix B: Field Data for a summary of each classification and sample size obtained). Field data points were reassessed for the Sentinel-2 imagery using an NDVI composite in the native Sentinel-2 10m resolution and pixel grid.

WBP and limber pine tend to grow in more open environments, which reduced the number of usable points, especially for the coarser resolution Landsat imagery (Table B3). To increase model accuracy following the Law of Large Numbers, additional points for WBP and limber pine were added to pixels surrounding ground-truthed points. We applied aerial tree identification methods using characteristic tree crown shapes seen from a mix of National Agriculture Imagery Program (NAIP) imagery from Esri's living Atlas and Firefly & Hybrid Imagery Basemap imagery from ArcGIS Pro and operated under the assumption that the location hosts the same species of tree as those identified by ground surveyors (Sayn-Wittgenstein, 1961; Paine & Kiser, 2003).

We assessed manual data entry results for misspellings, missing values, and categorical inconsistencies in R (R Core Team, 2024; Wickham et al. [2019, 2024, 2025]; Hester et al., 20204) before being imported back into ArcGIS Pro. After finalizing quality points, our team used a random generator expression in ArcGIS Notebook, *random()*, to evenly split up our training and validation points for each class to assess the accuracy of our distribution model. Our resulting training and validation point feature classes were converted to rasters using the "Point to Raster" tool in ArcGIS pro for easier import into TerrSet. We used the native spatial resolution, pixel grid, and x/y extent for the Landsat and Sentinel-2 data, resulting in a training raster and a validation raster for the Landsat data and a training raster and a validation raster for the Sentinel-2 data. All data was subsequently reprojected into the USA Contiguous Albers Equal Area Conic map projection.

2.3 Data Analysis

2.3.1 Spectral Separability Assessment

We initially evaluated the spectral separability of tree species in TerrSet. We merged ground sampled training raster and validation raster into a single raster representative of all in-situ data. All spectral bands selected from Landsat and Sentinel, plus our merged training and validation raster, were imported into TerrSet using the GDAL conversion utility. Spectral signatures were computed using the MakeSig module and graphed using the SigComp module. We then used the SepSig module to statistically assess the separability of spectral signatures for each band of each class relative to other mapped classes. We used the "Transformed Divergence" between class separability measure with a scaling factor of 2000.

2.3.2 Decision Forest & Accuracy Assessment

We ran a random forest classifier for Landsat and Sentinel-2 using the Decision Forest module in TerrSet. We provided the training raster as training data for the classifier. We included our spectral bands as input variables to the classifier, along with other raster data such as vegetation indices from satellite data, PCA components from satellite data, and topographic data (Table C1, Appendix C). We specified 3 variables for Landsat and 4 variables for Sentinel-2 to be selected at split and 100 trees. We utilized the same input variable and parameters for the decision forest module for both Landsat and Sentinel-2 in their native spatial resolution and pixel grid. The Decision Forest module outputs a classified raster of our study area showing the extent of each class. Additionally, we relied on the probability output layer which specifies the relative frequency of the tree prediction of each class. We independently assessed the accuracy of the output distribution model for both Landsat and Sentinel using the reserved validation raster, made using in-situ data that was not used for training. We used the ErrMat module in TerrSet to produce error matrices and Kappa Index of Agreement (KIA) values.

2.4 Field Maps Template Design

The Field Maps template was developed in cooperation with the Whitebark Pine Ecosystem Foundation to help enable collection of field data on WBP occurrence so that it can be more effectively used with remote sensing imagery. The survey will be used for the Foundation's Whitebark Pine Friendly Ski Area Certification program to establish monitoring programs at ski areas with WBP habitat. There are two surveys included in the template: one for WBP population surveying and one for Monitoring Tree surveys. Questions were designed for quick, accurate reporting by surveyors who may not have forestry experience and include basic tree location and health information. The Field Maps template was created using Field Maps Designer in ArcGIS Online. All questions are editable so they can be adapted to meet future needs. Application accuracy settings were set for a balance of scientific accuracy and field feasibility. For offline navigation, .mmpk files were created incorporating the Sentinel-2 WBP distribution model and topo maps. See Table D1 in Appendix D: Field Maps for point collection accuracy settings, Table D2 in Appendix D for survey questions and Table D3 in Appendix D for QR codes to each .mmpk file.

3. Results

3.1 Discussion

3.1.1 Spectral Separability Results

Our team analyzed the spectral separability of all 5 classes for Sentinel-2 using a spectral profile chart (Figure 3) and a statistical spectral separability analysis (Table E1). The focused species, WBP and limber pine, showed an average separability over all 11 bands of approximately 103.6. Additionally, the largest separability is in band 11 (SWIR), with the highest meaningful reflectance separability value of 265.91; band 12 (SWIR) showed a separability of 117.18, indicating potential utility in using these spectral bands and vegetation indices derived from these bands in our distribution model (Table E1).

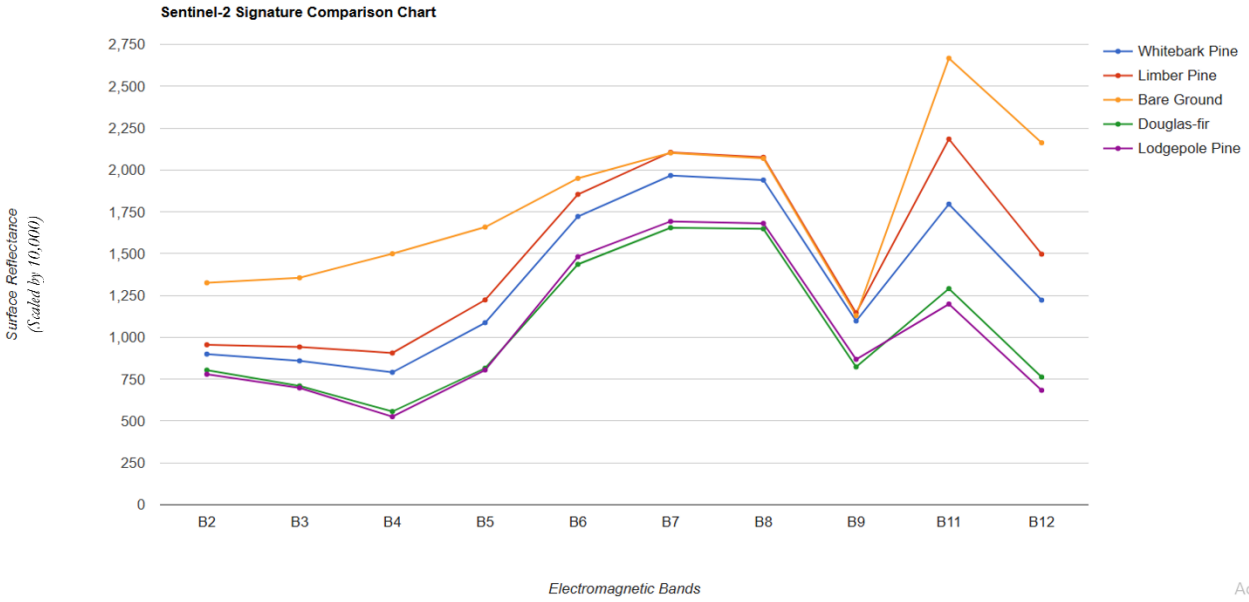

Figure 3. Spectral profile chart showing mean reflectance for each class at each Sentinel-2 band.

Our Landsat spectral bands showed very low separability between spectral bands for the WBP and Limber Pine classes (Figure 4). Average separability for the WBP and Limber Pine class across all bands for Landsat was 26.33 (Table E1, Appendix E).

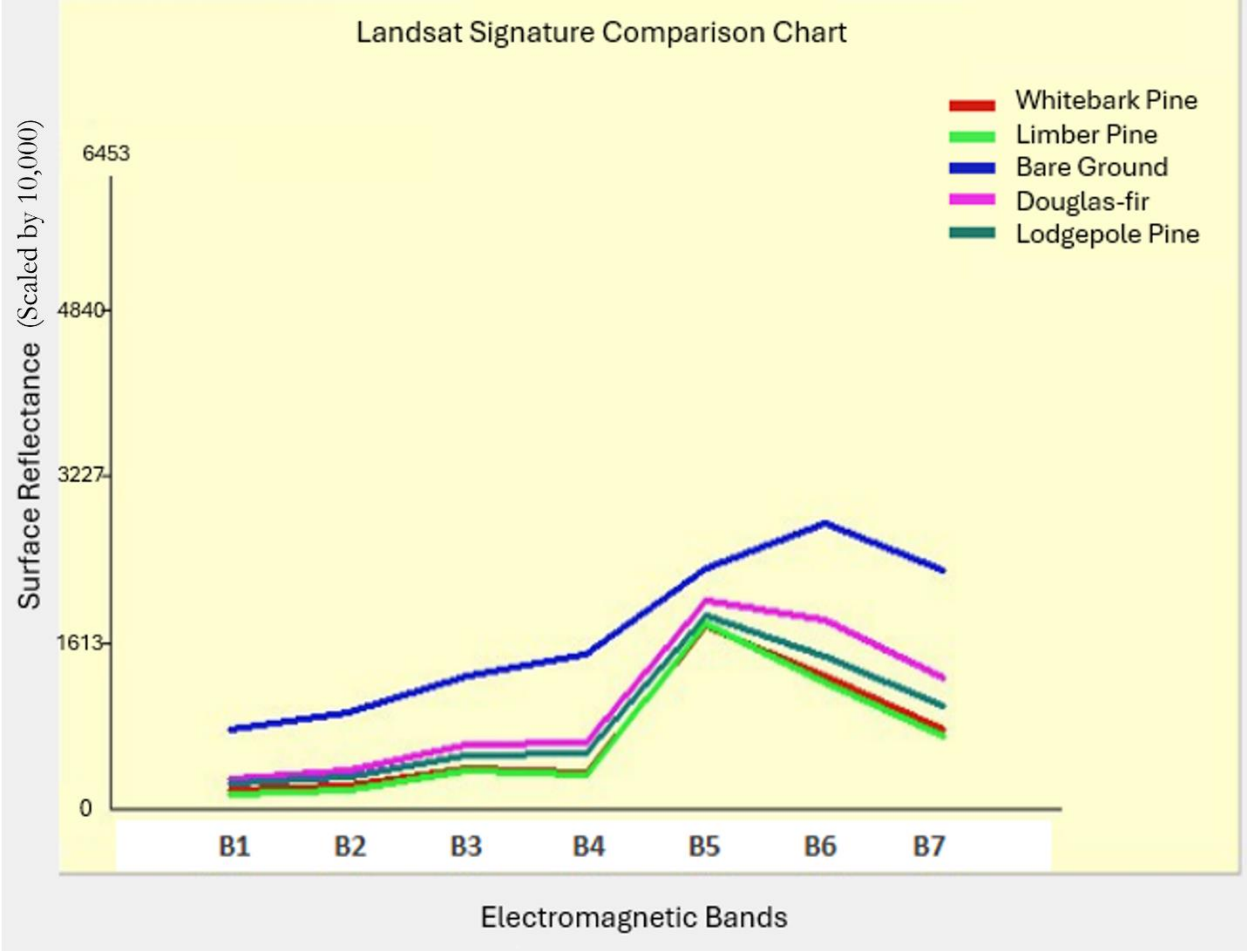

Figure 4. Spectral profile chart showing mean reflectance for each class at each Landsat 8/9 band.

3.1.2 Sentinel-2 Distribution Model

We used the output probability raster from our Sentinel-2 random forest model to map the probability of WBP across the study area. We visualized the WBP probability raster using a probability threshold of 50% for WBP and limber pine to show the highest classification confidence in the model (Figure 5; Figure 6). Setting this threshold allowed our team and partners to have the most reliable predictions for data collection when out in the field.

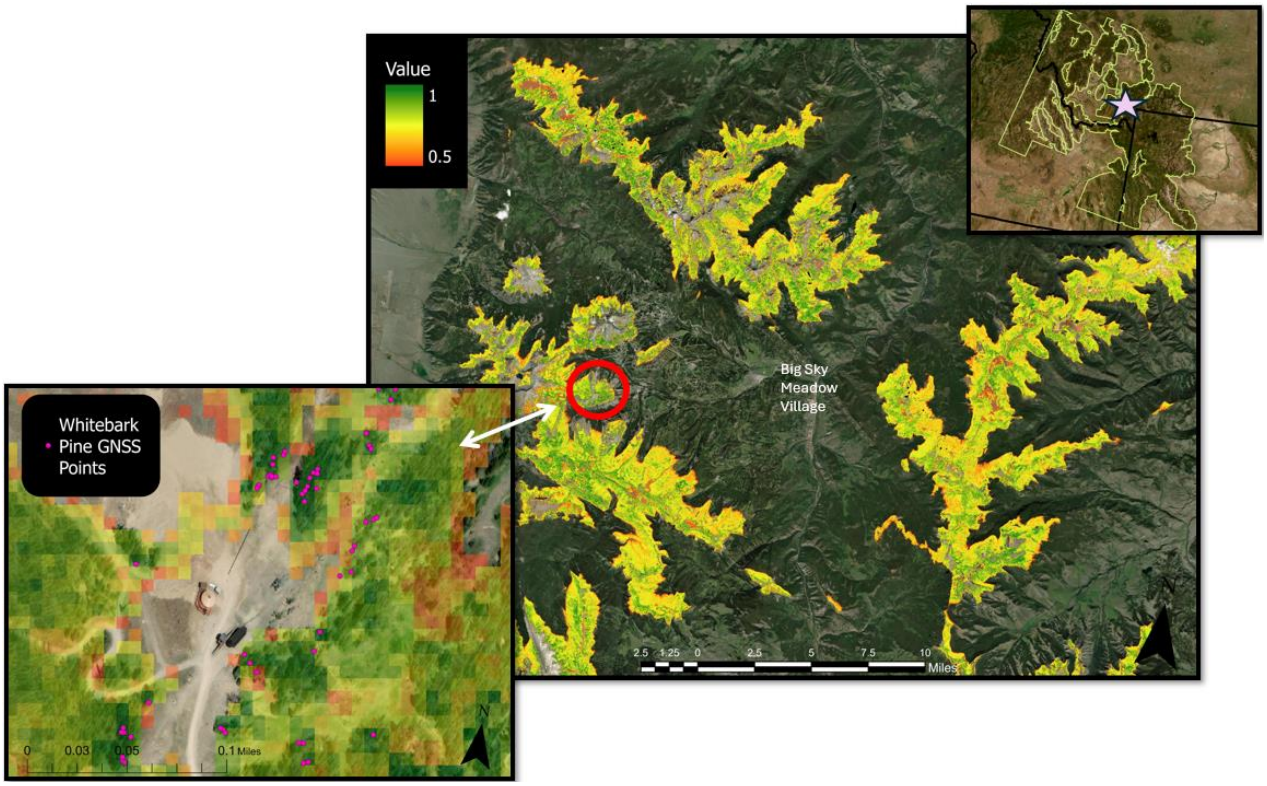

Figure 5. Sentinel-2 distribution model output for the WBP class is shown as moderate (red) to high probability (green).

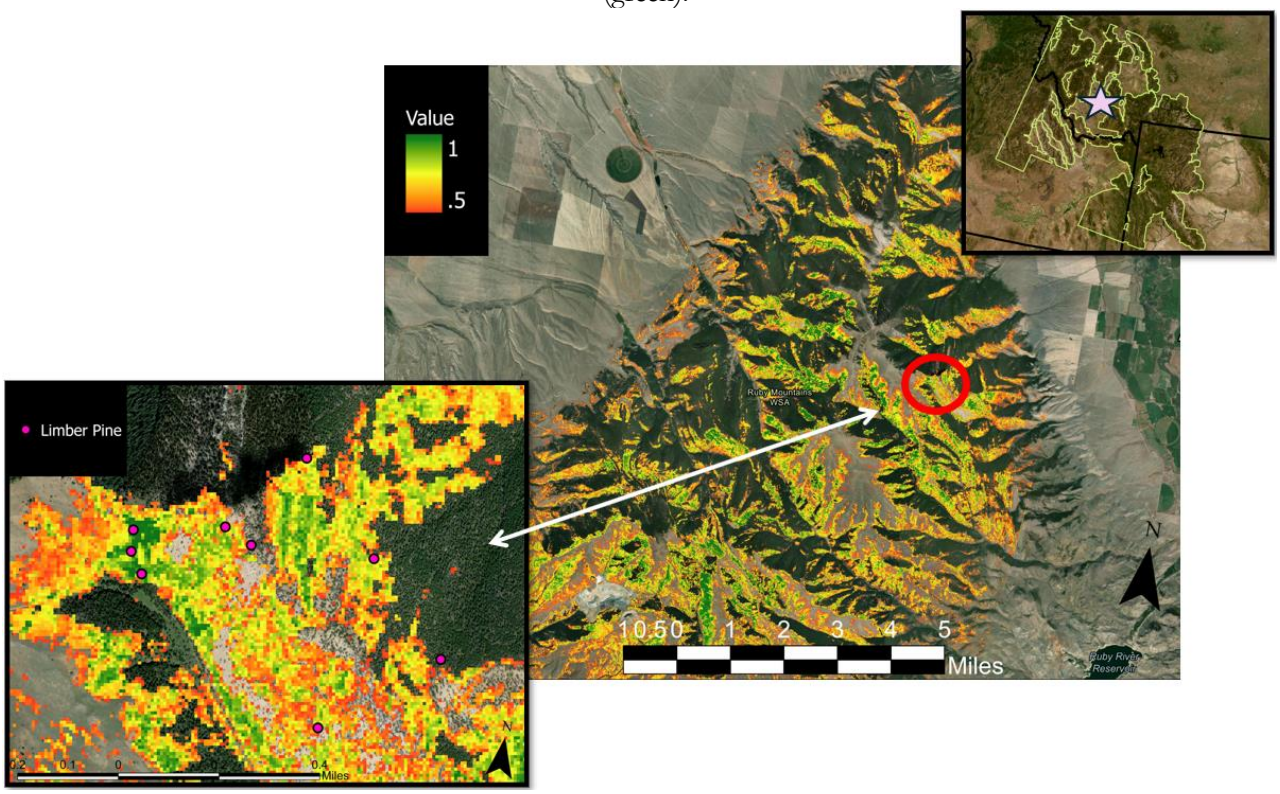

Figure 6. Sentinel-2 distribution model output for the Limber Pine class is shown as moderate (red) to high probability (green).

Comparing our output distribution model to our reserved validation data, we achieved an overall KIA value for all classes of 0.76, which reflects a moderately high agreement between the overall classification and the designated reference data. Furthermore, KIA accounts for random choice errors, resulting in a more conservative and more statistically robust measurement of how the model performs beyond random expectation. Class-specific KIA scores showed the highest agreement for WBP (0.94), with a slightly lower agreement for limber pine (0.72; See Table F1 in Appendix F: Sentinel-2 Model Accuracy Assessments for all KIA values). Additionally, errors of omission and commission illuminated class-specific performance within the model. WBP exhibited exceptionally low errors of omission (3.9%; 96.1% producer accuracy) and commission (4.7%; 95.3% user accuracy). Limber Pine showed a moderate omission error of 23.5% (76.5% producer accuracy) and a commission error of 22.4% (77.6% user accuracy), suggesting that the model failed to capture a notable portion of limber pine's actual presence; this level of confusion is perhaps due to the spectral and ecological similarities of WBP and limber pine. Lastly, the error matrix, see Table F2 in Appendix F, reported a total error of 18.31% and an overall classification accuracy of 81.69%, compared to KIA. The overall accuracy results do not necessarily demonstrate the model's reliability across the study area.

3.1.3 Landsat Distribution Model

We used the output probability raster from our Landsat random forest classification to map the probability of WBP across the study area, setting a threshold at 50% or greater probability (Figure 7). Comparing our output distribution model to our reserved validation data, we achieved an overall KIA value for all classes of 0.68. Among the five classes, WBP achieved an individual KIA score of 0.71, with an omission error of 18.2% (81.8% producer accuracy) and a commission error of 29.4% (70.6%), reflecting moderately strong classification confidence. Limber pine, however, showed more classification uncertainty, with a KIA of 0.51, a high omission error of 39.2% (60.8% producer accuracy), and a commission error of 21.1% (78.9% user accuracy), possibly due to the spectral similarities with WBP and other conifers in question. While performance was slightly lower than the Sentinel model, the Landsat-based approach still produced beneficial classification results. See Appendix G Landsat Model Accuracy Assessments Table G1 for full KIA and Table G2 for the Landsat Distribution Model Error Matrix.

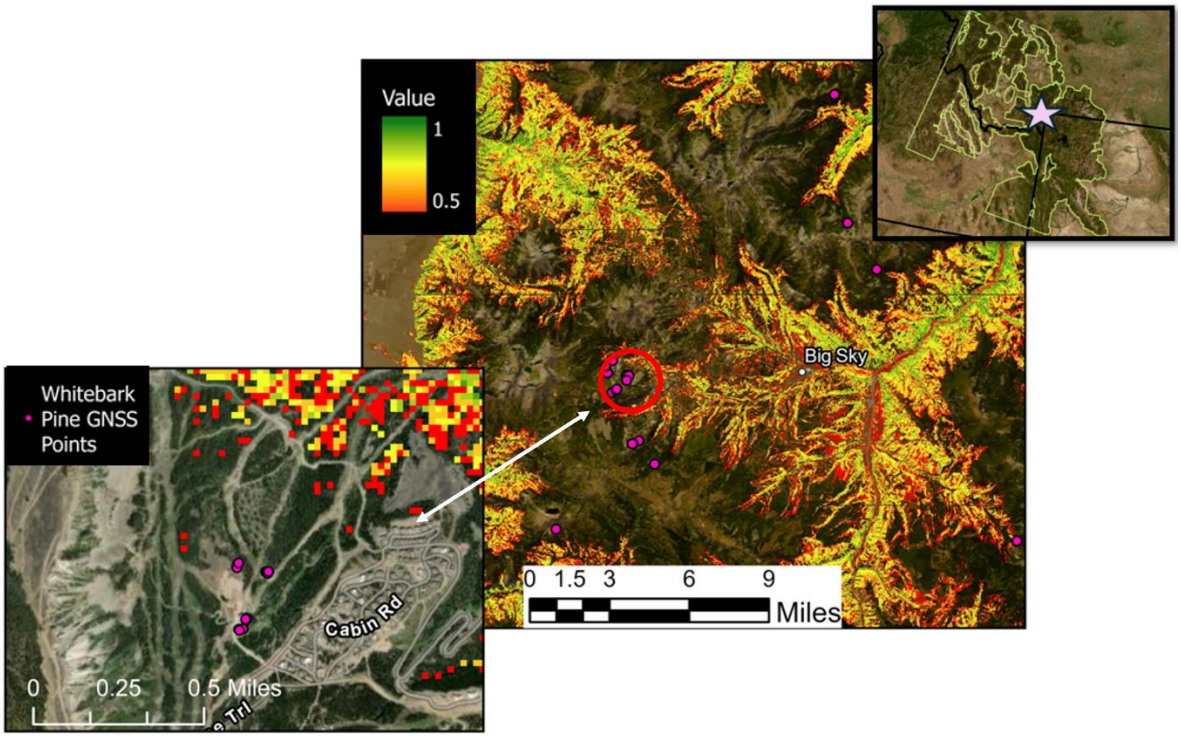

Figure 7. Landsat distribution model probability output for the WBP class. Areas in green with a value closer to 1 are higher probability for WBP according to the random forest classification.

3.2 Errors & Uncertainties

The differing spatial resolution of Landsat and Sentinel-2 somewhat confounded the comparison of these two maps in terms of the classification accuracy for each respective forest type distribution model. While both Landsat and Sentinel-2 data-based models targeted the same output tree classes, the actual training pixels also differed due to the larger Landsat pixel size, possibly leading to more variation in classification outcomes between the Landsat and Sentinel-2 model. Additionally, spectral similarities between WBP and Limber Pine pose a persistent challenge because of their physiological and ecological similarities, especially under mixed canopy conditions. Furthermore, neighboring field-identified points could have influenced the model to over-classify specific regions of high probability. The lack of snow-masking for Landsat data could also have introduced error by leading to snow covered areas earlier in the study period. The lack of dedicated understory vegetation classes likely contributed to the over-prediction of WBP and limber pine in both models, even after implementing a 50% probability threshold, as the model had limited capacity to distinguish sparse vegetation from tree canopy cover. Furthermore, the introduction of an Engelmann Spruce class and a sub-alpine fir class with sufficient in-situ GPS data could also help to reduce uncertainties since these grow in the same elevation zone as WBP. Compositing imagery over a short growing season may have obscured subtle ecological changes such as snow cover and deciduous shrub or understory vegetation growth; therefore, a plausible year-long or multi-year composite could impact the accuracy of our distribution model (Islam & Assal, 2023). Another possible source of error in our models was spatial autocorrelation of our training data. Spatial autocorrelation in the reference data could lead to possible error classifying areas of our study area without sufficient in-situ data that is not spatially autocorrelated (Figure 2). Another potential source of error is that our training data was not randomly created. It was ground truthed points which, while randomly assigned to be training or validation, were not truly random. Lastly, while the fieldwork we provided had a good foundation, additional ground-truthing and cross-validation would further strengthen the confidence in these models and possibly mitigate problems with spatial autocorrelation.

4. Conclusions

4.1 Interpretation of Results

Having scalable and reliable distribution models is crucial for conservation efforts aimed at monitoring WBP health and reducing the impact of WBP population decline. Our results show it is possible to distinguish differences in spectral signatures of WBP and other associated tree species. This spectral separability of Sentinel-2 bands for the WBP and Limber Pine class allowed us to create a scalable distribution model a large study area within the Intermountain West; furthermore, a field maps template for further in-situ data collection was developed for future data collection efforts. Our Sentinel-2 based WBP distribution model achieved a KIA of 0.94, and a very high apparent user and producer accuracy compared to available reference data (Table F2), producing more accurate results compared to the Landsat data (Table G2). Additionally, our Sentinel-2 based Limber Pine distribution model, a subset of our objectives, provided important insight on this crucial non-target species. The spatial distribution of WBP also differed between Sentinel-2 and Landsat distribution models. Landsat mapped large areas of WBP in drainage basins, possibly due to the sparse vegetation there, lack of snow masking, or coarser pixel resolution. On the other hand, Sentinel-2 accurately mapped WBP in sparsely vegetated areas near tree line and in the southeastern portion of our study area, an area widely known to be densely populated with WBP. However, for both models, parts of our study area outside of training areas should be interpreted with caution due to spatial autocorrelation of our training data and will need further ground validation to assess accuracy. The project's WBP distribution models provide general indications of occurrence and do not give information on the characteristics of WBP in areas where WBP is predictively distributed. In some cases, WBP is the dominant forest tree species, though in other cases it may be a minor component of stands dominated by other forest tree species.

4.2 Feasibility & Partner Implementation

The findings from this term demonstrate our project's methodologies and end products are applicable for our partners' decision-making processes for WBP conservation. The distribution model and spectral classification framework developed this term built upon previous work and provide partners with refined tools for identifying and monitoring WBP populations across wide portion of the Intermountain West. Our remote sensing-based approach accurately classified WBP presence, particularly in dense, homogenous stands. Our distribution model could improve through ground validation of areas outside of training and validation sites and by incorporating more sampling for future distribution models. We have streamlined this future work by providing partners with an Esri Field Maps template to standardize data collection and improve long-term monitoring. Partners can immediately include these products in their monitoring workflows, quickly identifying areas of high WBP probability and enabling more efficient allocation of resources for surveying and restoration.

5. Acknowledgements

We extend our gratitude to the following individuals for their invaluable contributions to this project:

- **DEVELOP Personnel:**
  - Isaac Goldings – Node Lead (Idaho – Pocatello)
  - Kait Lemon – Analytical Mechanics and Associates, DEVELOP Collaboration Coordinator
  - Laramie Plott – Senior Fellow (Virginia – Langley)
- **Advisors:**
  - Keith Weber - Idaho State University, GIS Director & NASA DEVELOP Science Advisor
  - Joe Spruce - Analytical Mechanics Associates, NASA Langley Research Center
- **Partners:**
  - Julee Shamhart - Executive Director, Whitebark Pine Ecosystem Foundation
  - Bob Keane - Board Associate Chair & Research Ecologist, Whitebark Pine Ecosystem Foundation
  - Melissa Jenkins - Board Secretary & Forest Silviculturist, Whitebark Pine Ecosystem Foundation
  - Joe Fortier - Remote Sensing Coordinator, USDA, USFS Region 1
  - Erin Shanahan - National Park Service, Yellowstone Inventory and Monitoring
  - Kristin Legg - National Park Service, Yellowstone Inventory and Monitoring
  - Jeff Cadry - Yellowstone Club, Environmental Manager
- **Collaborators:**
  - Chris Earle - Wildlife Biologist, The Gymnosperm Database
  - Jim Lindstrom - Cartographer, U.S. Forest Service
  - Destin Harrell - Biologist, U.S. Fish and Wildlife Service
  - Rachel Arrick - Ecologist, National Park Service
  - Laura Strong - Consulting Biologist, Independent Consultant
  - Hannah Alverson - Field Officer, Bureau of Land Management
  - Betsy Siszell – Sustainability Manager, Sun Valley Company
- **Previous DEVELOPers in Term 1 of project:**
  - Hannah Rogers, Dustin Corbridge, A H M Mainul Islam, Joshua Daniel Carrell

This material contains modified Copernicus Sentinel data (2024), processed by ESA. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration.

This material is based upon work supported by NASA through contract 80LARC23FA024.

6. Glossary

Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) - An optical sensor that delivers calibrated images of the upwelling spectral radiance in 224 contiguous spectral bands. The main objective of AVIRIS is to measure, identify, and monitor constituents of the Earth's surface and atmosphere based on molecular absorption and particle signature.

Airborne Laser Scanning (ALS) - A remote sensing technique that uses a laser scanner attached to an aircraft to create 3 dimensional models of earth's surface. Also commonly known as LiDAR (Light Detection and Ranging)

Band - A wavelength range in the spectrum of reflected or radiated electromagnetic (EM) energy to which a remote sensor is sensitive. Sensors collect data from a band and store the data in a file or a portion of a file devoted to that range. These files are also referred to as bands or spectral bands.

CTA - Classification Tree Analysis. A decision-tree-based classification method which segments data into hierarchical groups based on predictor variables, making it useful for land cover classification.

Earth observations - Satellites and sensors that collect information about the Earth's physical, chemical, and biological systems over space and time.

Ecoregion - A major ecosystem defined by distinct geography and receiving uniform solar radiation and moisture.

EVI - Enhanced Vegetation Index. It optimizes the vegetation signal with improved sensitivity in high biomass regions and improves vegetation monitoring through a de-coupling of the canopy background signal and a reduction in atmospheric influences.

MSI - Multispectral Instrument. We used Sentinel-2 MSI imagery for spectral signature analysis.

NDMI - Normalized Difference Moisture Index. An index that measures water content, calculated using the near-infrared (NIR) and shortwave infrared (SWIR) bands to detect moisture stress in vegetation.

NDVI - Normalized Difference Vegetation Index. It is a widely used index that measures the difference between NIR (which vegetation strongly reflects) and red light (which vegetation absorbs).

OLI - Operational Land Imager. We used Landsat 9 OLI-2 imagery for developing study area and NDVI for habitat suitability model.

PCA - Principal Component Analysis. A statistical technique used to extract significant spectral features from multispectral or hyperspectral imagery by transforming correlated bands into uncorrelated principal components.

RF - Random Forest is a machine learning model which provides predictions based on decision trees.

Spectral Reflectance - Spectral reflectance is the ratio between the energy reflected by the surface and energy incident on the surface. It is measured as a function of the wavelengths.

WBP - Whitebark pine. A high-elevation pine species (*Pinus albicaulis*) found in North America, known for its ecological importance in alpine and subalpine ecosystems.

7. References

Buotte, P.C., Hicke, J.A., Preisler, H.K., Abatzoglou, J.T., Raffa, K.F., & Logan, J.A. (2016). Climate influences on whitebark pine mortality from mountain pine beetle in the Greater Yellowstone Ecosystem. *Ecological Applications*, *26*(8), 2507-2524. https://doi.org/10.1002/eap.1396

Burns, R. M., & Honkala, B. H. (1990). Silvics of North America: Volume 1. Conifers (Agriculture Handbook 654). United States Department of Agriculture, Forest Service. https://www.srs.fs.usda.gov/pubs/misc/ag_654_vol1.pdf

Esri. [basemap]. Scale Not Given. "World Imagery". January 29th, 2025a. https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9. (April 3rd, 2025).

Esri. "Figure 2" [basemap]. Scale Not Given. "World Imagery (Firefly)". January 29th, 2025b. https://www.arcgis.com/home/item.html?id=a66bfb7dd3b14228bf7ba42b138fe2ea. (April 3rd, 2025).

European Space Agency. (2021). Copernicus Sentinel-2 MSI Collection 1 Level-2A Data [Data set]. https://doi.org/10.5270/S2_-znk9xsj

Greater Yellowstone Whitebark Pine Monitoring Working Group. (2011). *Interagency Whitebark Pine Monitoring Protocol for the Greater Yellowstone Ecosystem, Version 1.1*. Greater Yellowstone Coordinating Committee. National Park Service. https://irma.nps.gov/DataStore/DownloadFile/434245

Goeking, S. A., & Izlar, D. K. (2018). Pinus albicaulis Engelm. (Whitebark Pine) in mixed-species stands throughout its US range: Broad-scale indicators of extent and recent decline. *Forests, 9*(3), 131. https://doi.org/10.3390/f9030131

Hester, J., Wickham, H., & Csárdi, G. (2024). _fs: Cross-Platform File System Operations Based on 'libuv'_. R package version 1.6.5. https://CRAN.R-project.org/package=fs.

Islam, A. M., & Assal, T. J. (2023). Rapid, Landscape-Scale Assessment of Cyclonic Impacts on Mangrove Forests Using MODIS Imagery. *Coasts*, *3*(3), 280-293. https://doi.org/10.3390/coasts3030017

Jenkins, M. B., Schoettle, A. W., Wright, J. W., Anderson, K. A., Fortier, J., Hoang, L., Jr, T. I., Keane, R. E., Krakowski, J., LaFleur, D. M., Mellmann-Brown, S., Meyer, E. D., Pete, S., Renwick, K., & Sissons, R. A. (2022). Restoring a forest keystone species: A plan for the restoration of whitebark pine (Pinus albicaulis Engelm.) in the Crown of the Continent Ecosystem. *Forest Ecology and Management*, *522*, 120282. https://doi.org/10.1016/j.foreco.2022.120282

Jiang, Z., Huete, A., Kim, Y., & Didan, K. (2007). 2-Band enhanced vegetation index without a blue band and its application to AVHRR data. *Proceedings of SPIE - the International Society for Optical Engineering, 6679*. https://doi.org/10.1117/12.734933

Jin, S., & Sader, S. A. (2005). Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment, 94*(3), 364–372. https://doi.org/10.1016/j.rse.2004.10.012

Keane, R.E., Holsinger, L.M., Mahalovich, M.F., & Tomback, D.F. (2017). Restoring whitebark pine ecosystems in the face of climate change. Gen. Tech. Rep. RMRS-GTR-361. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 123 p. https://www.climatehubs.usda.gov/sites/default/files/rmrs_gtr361.pdf

Kokaly, R. F., Despain, D. G., Clark, R. N., & Livo, K. E. (2003). Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. *Remote Sensing of Environment, 84*(3), 437-456. https://doi.org/10.1016/S0034-4257(02)00133-5.

Kriegler, F., Malila, W., Nalepka, R., & Richardson, W. (1969). Preprocessing transformations and their effect on multispectral recognition. *Proceedings of the 6th International Symposium on Remote Sensing of Environment*. Ann Arbor, MI: University of Michigan, 97-131.

Landenburger, L., Lawrence, R. L., Podruzny, S., Schwartz, C. C. (2008). Mapping Regional Distribution of a Single Tree Species: Whitebark Pine in the Greater Yellowstone Ecosystem. *Sensors, 8*(8), 4983-4994. https://doi.org/10.3390/s8084983.

Omernik, J.M. (1987). Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers, 77*(1), 118 – 125. https://doi.org/10.1111/j.1467-8306.1987.tb00149.x

Paine, D. P., & Kiser, J. D. (2003). Forestry, *Aerial Photography and Image Interpretation* (2nd ed., pp. 413 - 434). John Wiley & Sons, Inc.

R Core Team (2024). _R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rogers, H., Corbridge, D., Islam, A. M., & Carrell, J. D. (2024). *Northern Rockies Ecological Conservation: Leveraging Earth Observations to Monitor and Predict Populations of Federally Threatened Whitebark Pine (Pinus albicaulis) across the Intermountain West.* NASA DEVELOP National Program, Idaho – Pocatello. https://ntrs.nasa.gov/citations/20240010394

Sayn-Wittgenstein, L. (1961). "Recognition of Tree Species on Air Photographs by Crown Characteristics." *Photogrammetric Engineering*, *27*(5), 792 – 809.

Tomback, D.F., & Achuff, P. (2010). Blister rust and western forest biodiversity: ecology, values and outlook for white pines. *Forest Pathology, 40,* 186-225. https://doi.org/10.1111/j.1439-0329.2010.00655.x

U.S. Fish and Wildlife Service. (2022). Endangered and Threatened Wildlife and Plants; Threatened Species Status with Section 4(d) Rule for Whitebark Pine (Pinus albicaulis). *Federal Register*, *87*, 76882 – 76917 https://www.govinfo.gov/content/pkg/FR-2022-12-15/pdf/2022-27087.pdf

U.S. Forest Service Region One. (2019 April 11). Region 1 Common Stand Exam and Inventory and Monitoring Field Guide. Retrieved from https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fseprd699576.pdf.

United States Geological Survey. (2025). EarthExplorer. https://earthexplorer.usgs.gov/

USGS. (2020). Landsat 8-9 OLI/TIRS Collection 2 Level-2 Science Products. [Data set]. USGS EROS Archive. https://doi.org/10.5066/P9OGBGM6

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). "Welcome to the tidyverse." *Journal of Open-Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686.

Wickham, H., Hester, J., & Bryan, J. (2024). _readr: Read Rectangular Text Data_. R package version 2.1.5. https://CRAN.R-project.org/package=readr.

Wickham, H., Vaughan, D., & Girlich, M. (2024). _tidyr: Tidy Messy Data_. R package version 1.3.1. https://CRAN.R-project.org/package=tidyr.

Wickham, H., Henry, L. (2025). _purrr: Functional Programming Tools_. R package version 1.0.4. https://CRAN.R-project.org/package=purrr.

7. References

Appendix A: Data Specifications for Satellite Data

Table A1.
Data Specifications for Satellite Data.

Data Product	**Data Source**	**Spatial Resolution**	**Temporal Resolution**	**Temporal Range**	**Acquisition Method**
Landsat 8 OLI Collection 2 Level 2 (Surface Reflectance)	USGS and NASA Earth Observing Systems	30 m	8 days	04-01-2024 - 09-01-2024	Earth Explorer
Landsat 9 OLI-2 Collection 2 Level 2 (Surface Reflectance)	USGS and NASA Earth Observing Systems	30 m	8 days	04-01-2024 - 09-01-2024	Earth Explorer
Sentinel-2 MSI Level 2A (TOA Reflectance)	European Space Agency	10 m, 20 m, 60 m	5 days	04-01-2024 - 09-01-2024	Sentinel Hub
3DEP elevation data	National Elevation Data 3DEP Program	10 m			NASA RECOVER program

Appendix B: Field Data

Table B1.
Data Specifications for Ground Sampled Data.

Data Source (Organization and/or researcher)	Number of GNSS points provided	Sampling Area	Acquisition Date	Data collection method	Average horizontal accuracy (meters)	Species recorded
Summer, 2024 DEVELOP Team	12	Bonneville Peak, Idaho	June 1, 2024	Field Maps, or a Trimble R1 GNSS receiver	1 - 5	PIAL, PIFL2, ABLA, PIEN
Whitebark Pine Ecosystem Foundation and Yellowstone Club	480	Yellowstone Club, Sawtooth National Forest, and Whitefish Range	September 2024-January 2025	Field Maps, or a Trimble R1 GNSS receiver		WBP
Joe Fortier, USFS Region One Remote Sensing Coordinator	965	Northern Rockies	Unknown	Adapted National Common Stand Exam methods	1 - 5	PSME, PICO, mixed stand
Erin Shannahan, National Park Service Greater Yellowstone Inventory and Monitoring Network	214	Northern Rockies	Unknown	Interagency Whitebark Pine Monitoring Protocol for the Greater Yellowstone Ecosystem		PIAL, PIFL2

Table B2.
Total number of points and pixels used after data cleaning for Sentinel-2 and Landsat.

Class ID	Class name	Total GNSS Points	Total Sentinel Pixels	Total Landsat Pixels
1	Whitebark Pine	196	255	190
2	Limber Pine	37	135	113
3	Lodgepole Pine	148	142	131
4	Douglas-fir	175	175	30
5	Bare Ground	109	109	92

Appendix C: Decision Forest Input Variables

Table C1.
Decision Forest Input Variables.

Decision Forest Input Variables Landsat	**Decision Forest Input Variables Sentinel-2**
Band 1 - Coastal Blue	Band 2 - Blue
Band 2 - Blue	Band 3 - Green
Band 3 -Green	Band 4 - Red
Band 4 - Red	Band 5 - Vegetation Red Edge
Band 5 - Near Infrared (NIR)	Band 6 - Vegetation Red Edge
Band 6 - Short Wave Infrared (SWIR) I	Band 7 - Vegetation Red Edge
Band 7 - Short Wave Infrared (SWIR) 2	Band 8 - Near Infrared (NIR)
PCA Component 1 (59% of the variability)	Band 8a- Narrow NIR
PCA Component 2 (35% of the variability)	Band 10 - Water vapour
NDVI	Band 11 - Short Wave Infrared (SWIR) I
EVI2	Band 12 - Short Wave Infrared (SWIR) 2
NDMI	PCA Component 1 (96% of the variability)
Elevation	NDVI
Slope	EVI2
Aspect	NDMI
	Elevation
	Slope
	Aspect

Appendix D: Field Maps

Table D1.
Field Maps template location accuracy settings.

Location accuracy	30 feet
Horizontal and vertical accuracy confidence level	95 percent
GPS averaging	100 points
Fix type threshold	Any
Manual location	allowed

Table D2.
Field Maps template survey questions. The form opens with a set of introductory instructions and links to training materials. Survey questions are the same for both templates, with the exception of the Monitoring Tree question.

WBP Monitoring Instructions

Only mark live, 5-needle pine trees. After you are familiar with the form, each tree record should take no more than 2 - 5 minutes to complete.

- Study more on 5-needle pine identification here.
- Study more on blister rust signs and symptoms of infection here.
- Study more on mountain pine beetle here.

**We recommend downloading or printing identification material to your device prior to conducting field surveys, as there may not be adequate service to view information.*

Surveyor name	Manual entry
Date	Autofill
Tree species?	Whitebark pine Limber pine Unknown 5-needle pine
Cones present?	Yes No Unknown
Stand type	Mixed stand Stand-alone tree WBP dominant stand
Blister rust observed? (Select all that apply)	Canker (take a photo) Flagging (take a photo) Not observed Unknown
Mountain pine beetle activity observed? (Select all that apply)	Pitch tubes (take a photo) Frass (take a photo) No signs observed Unknown
Is this a monitoring tree? Characteristics for potential include trees: representative of average size and condition of stand, large/unique, special to ski area, "Plus"/"Elite" tree designated by land manager	Yes, record data in monitoring tree form No (take a photo) Potential (take photos from four sides of the tree)
Notes	Open space for additional field observations.

Table D3.
Links to .mmpk files for use with Field Maps templates during offline surveys. WBP distribution probability is included as an overlay, with National Forest System Boundaries, National Park Boundaries and Wilderness Areas.

Idaho Whitebark Population Mapping Project Northernmost Idaho.		**Wyoming** NW Wyoming	
Idaho Whitebark Population Mapping Project N Idaho 2, Lake Pend Oreille area.		**Montana** Western Montana 1: northwestern-most region	
Idaho Whitebark Population Mapping Project N Idaho3.		**Montana** Western Montana 2: westernmost central region	
Idaho Central Idaho 1		**Montana** Western Montana 3: Westernmost southern region.	
Idaho Central Idaho 2.		**Montana** Western Montana 4: southeastern-most bit of western Montana	
Idaho Central Idaho 3.		**Montana** Western Montana 5: central-easternmost bit of western Montana	
Wyoming Southeastern Idaho and Western Wyoming			

Appendix E: Spectral Separability for Landsat and Sentinel-2

Table E1.
Spectral separability for Landsat and Sentinel-2 bands for the WBP and Limber Pine class.

Landsat Bands	Separability of WBP and Limber Pine	Sentinel-2 Bands	Separability of WBP and Limber Pine
Band 1 - Coastal Blue	76.60	Band 2 - Blue	72.57
Band 2 - Blue	45.99	Band 3 - Green	73.04
Band 3 -Green	9.08	Band 4 - Red	68.96
Band 4 - Red	15.74	Band 5 - Vegetation Red Edge	78.11
Band 5 - Near Infrared (NIR)	14.94	Band 6 - Vegetation Red Edge	93.34
Band 6 - Short Wave Infrared (SWIR) I	13.24	Band 7 - Vegetation Red Edge	87.59
Band 7 - Short Wave Infrared (SWIR) 2	8.73	Band 8 - Near Infrared (NIR)	95.00
		Band 9 - Water Vapor	117.81
		Band 10 - Cirrus	9.99
		Band 11 - Short Wave Infrared (SWIR) I	265.91
		Band 12 - Short Wave Infrared (SWIR) 2	117.18

Appendix F: Sentinel-2 Model Accuracy Assessments

Gauging our model's performance: Our WBP KAPPA Index of Agreement (Table F1) showed a high accuracy of 0.93. Additionally, our distribution model error matrix for our WBP class (Table F2) identified only six commission errors, where the model missed a WBP pixel. In contrast, the error of omission falsely predicted three limber pine and two bare ground pixels.

Table F1.
Kappa Index of Agreement (KIA) for Sentinel-2 Model.

Class	**KIA**
WBP	0.943
Limber Pine	0.718
Lodgepole Pine	0.578
Douglas-fir	0.639
Bare Ground	0.931

Table F2.
Sentinel-2 distribution model error matrix.

	WBP	**Limber Pine**	**Lodgepole Pine**	**Douglas-Fir**	**Bare Ground**	**Total**	**Error of Commission**
WBP	122	6	0	0	0	**128**	**4.69%**
Limber Pine	3	52	3	6	3	**67**	**22.39%**
Lodgepole Pine	0	1	48	19	0	**68**	**29.41%**
Douglas-Fir	0	2	22	63	0	**87**	**27.59%**
Bare Ground	2	7	1	0	48	**58**	**17.24%**
Total	**127**	**68**	**74**	**88**	**51**	**408**	
Error of Omission	**3.93%**	**23.53%**	**35.13%**	**28.41%**	**5.88%**		**Total Error 18.38%**

ErrorO = Errors of Omissions (expressed as proportions).
ErrorC = Error of Commission (Expressed as proportions).
Total Error = all incorrect points/all data points

90% Confidence Interval = +/- 0.0315 (0.1523 – 0.2154)
95% Confidence Interval = +/- 0.0376 (0.1462 – 0.2214)
99% Confidence Interval + +/- 0.0495 (0.1343– 0.2333).

Appendix G: Landsat Model Accuracy Assessments

Gauging our model's performance: Our WBP KAPPA index of agreement (Table G1) showed a substantial accuracy of 0.71, or 71% agreement. Additionally, our distribution model error matrix for our WBP class (Table G2) identified only 30 total commission errors and 26 total omission errors.

Table G1.
Kappa Index of Agreement (KIA) values for Landsat Model.

Category	KIA
WBP	0.713
Limber Pine	0.507
Lodgepole Pine	0.950
Douglas-Fir	0.288
Bare Ground	0.818

Table G2.
Landsat Distribution Model Error Matrix.

	WBP	Limber Pine	Lodgepole Pine	Douglas-Fir	Bare Ground	**Total**	**Error of Commission**
WBP	72	26	0	4	0	**102**	**29.41%**
Limber Pine	12	45	0	0	0	**57**	**21.05%**
Lodgepole Pine	2	3	49	2	6	**62**	**20.97%**
Douglas-Fir	2	0	2	6	1	**11**	**45.45%**
Bare Ground	0	0	0	7	39	**46**	**15.22%**
Total	**88**	**74**	**51**	**19**	**46**	**278**	
Error of Omission	**18.18%**	**39.19%**	**3.92%**	**68.42%**	**15.22%**		**24.10%**

ErrorO = Errors of Omission (expressed as proportions)

ErrorC = Errors of Commission (expressed as proportions)

Total Error = all incorrect points/all data points

90% Confidence Interval = +/- 0.0422 (0.1988 - 0.2832)
95% Confidence Interval = +/- 0.0503 (0.1907 - 0.2913)
99% Confidence Interval = +/- 0.0662 (0.1748 - 0.3072)