Mapping salmonberry concentration through multispectral imaging

Gyabaah Kyere Gyeabour ‘27 and Sean P. Gleason

Department of Biology, Hampden-Sydney College, Hampden-Sydney, VA 23943

Abstract

This summer, I collaborated with the United States Department of Agriculture (USDA) under the mentorship of Dr. Gleason at Hampden-Sydney College. The overarching goal was to investigate the feasibility of using remote sensing and data science methods to locate salmonberry patches in Quinhagak, Alaska, and estimate their potential yield. This research lays the groundwork for integrating machine learning models into ecological monitoring efforts that can ultimately serve Indigenous food systems.

Background Information

Salmonberries (Rubus spectabilis), native to coastal Alaska, represent an important subsistence food source for Indigenous and rural communities. Rich in vitamins and cultural significance, salmonberries are traditionally gathered and play a vital role in maintaining food security, health, and cultural identity in Alaskan villages. However, climate change, shifting ecological patterns, and increased environmental pressures have created challenges in both the growth and distribution of salmonberry populations. Locating viable berry patches has become more difficult, while unpredictable yields present further concerns for communities that rely on these plants.

Advances in geospatial technology, particularly remote sensing combined with computer vision and machine learning, provide a promising approach to addressing this challenge. By leveraging UAV-based and satellite-based multispectral imagery, we can map and identify salmonberry patches, estimate their productivity, and provide Alaskan communities with data-driven tools to enhance food security. Multispectral imaging involves the collection of image data across multiple discrete spectral bands, including those beyond the visible range of human perception. Unlike traditional RGB imagery, which is limited to red, green, and blue wavelengths, multispectral sensors capture additional bands—most commonly the near-infrared (NIR)—that provide critical information about vegetation and soil characteristics. Such approaches not only reduce travel costs and time spent searching for berries but also support the preservation of cultural practices tied to traditional food gathering.

Materials and Methods

Data Acquisition and Processing.

One of the fundamental challenges in applying computer vision and machine learning to ecological problems is obtaining high-quality, labeled data. For this project, we explored multiple data acquisition strategies:

- ***UAV-mounted sensors:** UAVs (drones) equipped with multispectral cameras were deployed to capture high-resolution imagery of berry-growing areas. UAV imagery provides centimeter-level spatial resolution, which is invaluable for detecting fine-grained vegetation patterns. However, UAV data collection faced limitations such as:*
  - *o **Environmental challenges:** Strong coastal winds in Quinhagak caused UAV instability, often resulting in blurred or incomplete image captures.*
  - *o **Logical Constraints:** UAVs require trained operators, have limited battery life, and cover relatively small areas per flight, necessitating repeated missions for larger-scale data collection.*
- ***Satellite imagery:** To address UAV limitations, we considered satellite-based imagery with multispectral bands as a scalable alternative. While spatial resolution is coarser than UAV imagery, satellite data provides broader spatial coverage, enabling real-time monitoring of entire communities. By incorporating freely available or commercial multispectral satellite data, the approach becomes more accessible to rural Alaskan communities without requiring on-site UAV operations. Also, other research studies involving the use of satellite data for remote sensing has been done using image data with resolution of about 30m, and the results proved to be acceptable for technical applications. Using satellite imagery, which is both accessible and offers comparable spatial resolution, can greatly improve the scalability of our approach. Unlike UAVs, satellites provide consistent coverage without the need for on-site operation. This makes the software solution more practical, reliable, and easier for local communities to use.*

Collected UAV imagery was processed using Agisoft Metashape 2.2.1 to generate high-resolution orthomosaics. These orthomosaics were then imported into ArcGIS Pro 3.5.2, where geospatial analyses and vegetation index calculations were conducted.

Vegetation Indices Calculation.

Vegetation indices provide quantitative measures derived from spectral reflectance properties of vegetation. These indices enhance spectral differences and allow researchers to infer vegetation health, water content, and chlorophyll activity. During this project, we implemented the following indices:

- **Normalized Difference Vegetation Index (NDVI):**
  - Formula: (NIR – Red) / (NIR + Red)
  - The Normalized Difference Vegetation Index (NDVI) method is a standardized index allowing you to generate an image displaying greenness (relative biomass). This index takes advantage of the contrast of the characteristics of two bands from a multispectral raster dataset—the chlorophyll pigment absorptions in the red band and the high reflectivity of plant materials in the NIR band.
- **Green Chlorophyll Index (CIg):**
  - Formula: (NIR / Green) – 1
  - The Chlorophyll Index - Green (CIg) method is a vegetation index for estimating the chlorophyll content in leaves using the ratio of reflectivity in the NIR and green bands, which is directly related to photosynthetic activity and plant productivity. It can specify the health status of the vegetation or warn of the start of temporary seasons. This index is particularly useful in distinguishing high-yielding berry patches from surrounding vegetation.
- **Normalized Difference Water Index (NDWI):**
  - Formula: (NIR – Green) / (NIR + Green)
  - The Normalized Difference Water Index (NDWI) method is an index for delineating and monitoring content changes in surface water. It is computed with NIR and green bands. NDWI highlights water content in vegetation, allowing assessment of plant stress and soil moisture conditions. Since salmonberries thrive in moist soils, NDWI may serve as an important marker for suitable berry habitats.

Future work will incorporate additional indices such as the Enhanced Vegetation Index (EVI), Bare Soil Index (BSI), and Visible Atmospherically Resistant Index (VARI) to provide a richer dataset for machine learning classification tasks.

Ground truthing.

Remote sensing models require validation against field data to ensure accuracy. Ground truthing was conducted using GIS methods:

• Differential GPS measurements were taken to identify precise coordinates of salmonberry patches. These locations were digitized into polygon shapefiles, serving as labeled training data for supervised classification models.

• Indicator species were also noted, as certain plants co-located with salmonberries can serve as ecological markers. Incorporating these species as secondary features may enhance classification performance.

Ground truth data thus forms a critical link between spectral imagery and ecological reality, enabling effective training and evaluation of predictive models.

Discussion

Findings.

Despite technical challenges, the project achieved several key milestones:

• **Orthomosaic generation:** High-resolution orthomosaics of the study area were successfully produced from UAV imagery. These orthomosaics serve as foundational datasets for geospatial analyses.

• **Vegetation index mapping:** Multiple vegetation indices (NDVI, GCI, NDWI) were calculated and visualized in ArcGIS Pro. These indices improved feature separability and highlighted vegetation patterns likely corresponding to salmonberry patches.

• **Preliminary patterns:** Early analysis revealed distinct spectral and spatial patterns in areas identified by locals as traditional berry-picking sites, supporting the feasibility of using these indices for patch identification.

Future Work.

The next phase of this project will focus on:

1. **Expanding vegetation indices:** Incorporating additional indices and biophysical variables, such as DEM-derived elevation, slope, and aspect, to refine predictive capability.
2. **Machine learning implementation:** Training models such as XGBoost, Random Forests, and Multiple Linear Regression (MLR) to classify salmonberry patches from spectral and spatial features. These models will be validated using field-labeled training data and evaluated with accuracy, precision, recall, and F1-score metrics.
3. **Predictive testing:** Applying trained models to previously unseen imagery datasets to evaluate real-world predictive power.

4. **Yield estimation:** Developing statistical relationships between spectral signatures and estimated berry yield, potentially integrating regression-based approaches.

5. **Community integration:** Designing user-friendly workflows or platforms (e.g., web dashboards or mobile apps) to allow local communities to access berry patch predictions and yield maps in real time.

Conclusion.

This summer's research successfully demonstrated the potential of remote sensing for locating and assessing salmonberry patches in Alaska. By combining UAV imagery, spectral vegetation indices, and geospatial analysis, we created a preliminary workflow that lays the groundwork for machine learning applications in ecological monitoring

The integration of machine learning algorithms with expanded remote sensing data will significantly improve predictive accuracy and scalability. Beyond its scientific contributions, this project holds strong cultural and social relevance, providing Indigenous communities with practical tools for sustaining food security and preserving traditional berry-harvesting practices.

Ultimately, this research exemplifies how interdisciplinary approaches—merging ecology, geospatial science, and data science—can address pressing challenges at the intersection of climate change, culture, and community resilience.

REFERENCES

S. Gleason, "2025 Alaska Food Festival and Conference - Sendable.pptx," Mar. 24, 2025. Accessed: Aug. 25, 2025. [Online]. Available: https://docs.google.com/presentation/d/1jtqabl2mN4HFHXOYtW4lAiFUQRe4WWG3

Parkinson, L., Mulder, C., Putman, M., & Spellman, K. (2024). "(PDF) Salmonberry in a Changing Climate: Threats and Opportunities," ResearchGate. Accessed: Aug. 25, 2025. [Online]. Available: https://www.researchgate.net/publication/383526590_Salmonberry_in_a_Changing_Climate_Threats_and_Opportunities

S. N. Subhashree et al., "Evaluating Remote Sensing Resolutions and Machine Learning Methods for Biomass Yield Prediction in Northern Great Plains Pastures," Agriculture, vol. 15, no. 5, p. 505, Jan. 2025, doi: 10.3390/agriculture15050505.

Y. Pazmiño, J. J. de Felipe, M. Vallbé, F. Cargua, and L. Quevedo, "Identification of a Set of Variables for the Classification of Páramo Soils Using a Nonparametric Model, Remote Sensing, and Organic Carbon," Sustainability, vol. 13, no. 16, p. 9462, Jan. 2021, doi: 10.3390/su13169462.

B. Zagajewski et al., "Intraspecific Differences in Spectral Reflectance Curves as Indicators of Reduced Vitality in High-Arctic Plants," Remote Sens., vol. 9, no. 12, p. 1289, Dec. 2017, doi: 10.3390/rs9121289.

"NDVI—ArcGIS Pro | Documentation." Accessed: Aug. 25, 2025. [Online]. Available: https://pro.arcgis.com/en/pro-app/3.4/arcpy/image-analyst/ndvi.htm

"CIg—ArcGIS Pro | Documentation." Accessed: Aug. 25, 2025. [Online]. Available: https://pro.arcgis.com/en/pro-app/3.4/arcpy/image-analyst/cig.htm

"NDWI—ArcGIS Pro | Documentation." Accessed: Aug. 25, 2025. [Online]. Available: https://pro.arcgis.com/en/pro-app/3.4/arcpy/image-analyst/ndwi.htm