Past Movement and New Models: Reconstructing Past Mobility in the Absaroka Mountains by Applying Bayesian Neural Networks Towards Refining Trace Element Modeling

Daniel Dalmas [1,2]
Lawrence Todd [1]
David Rapson [3]

[1] GRSLE
[2] Dept. Anthropology, University of Utah
[3] Dept. Anthropology, University of Wyoming

To date, the GRSLE Project has recorded >11,000 obsidian artifacts in the Greater Yellowstone ecosystem. Of these, >1,400 have been sourced with Elemental X-Ray Fluorescence (EXRF) technology. One of the tenants of the GRSLE Project has been to practice "catch and release" archaeological recording where >99.9% of recorded artifacts are left in place. Obsidian sourcing is among the few data sources that requires removal of artifacts. By developing and testing sourcing accuracy using portable XRF (pXRF) data we may accurately source obsidian artifacts in the field and continue to minimize our disturbance to the archaeological record which may be preserved for future generations.

Scanning Method

Obsidian artifacts were scanned using a Bruker Tracer Titan III pXRF. Artifacts were scanned for 30 seconds at a kV interval of [1, 40], uA = 12, with a Cu 100um: Ti 25um: Al 300 um filter wheel for 9 cycles. Resulting data were processed using Artax to perform Bayesian Deconvolution.

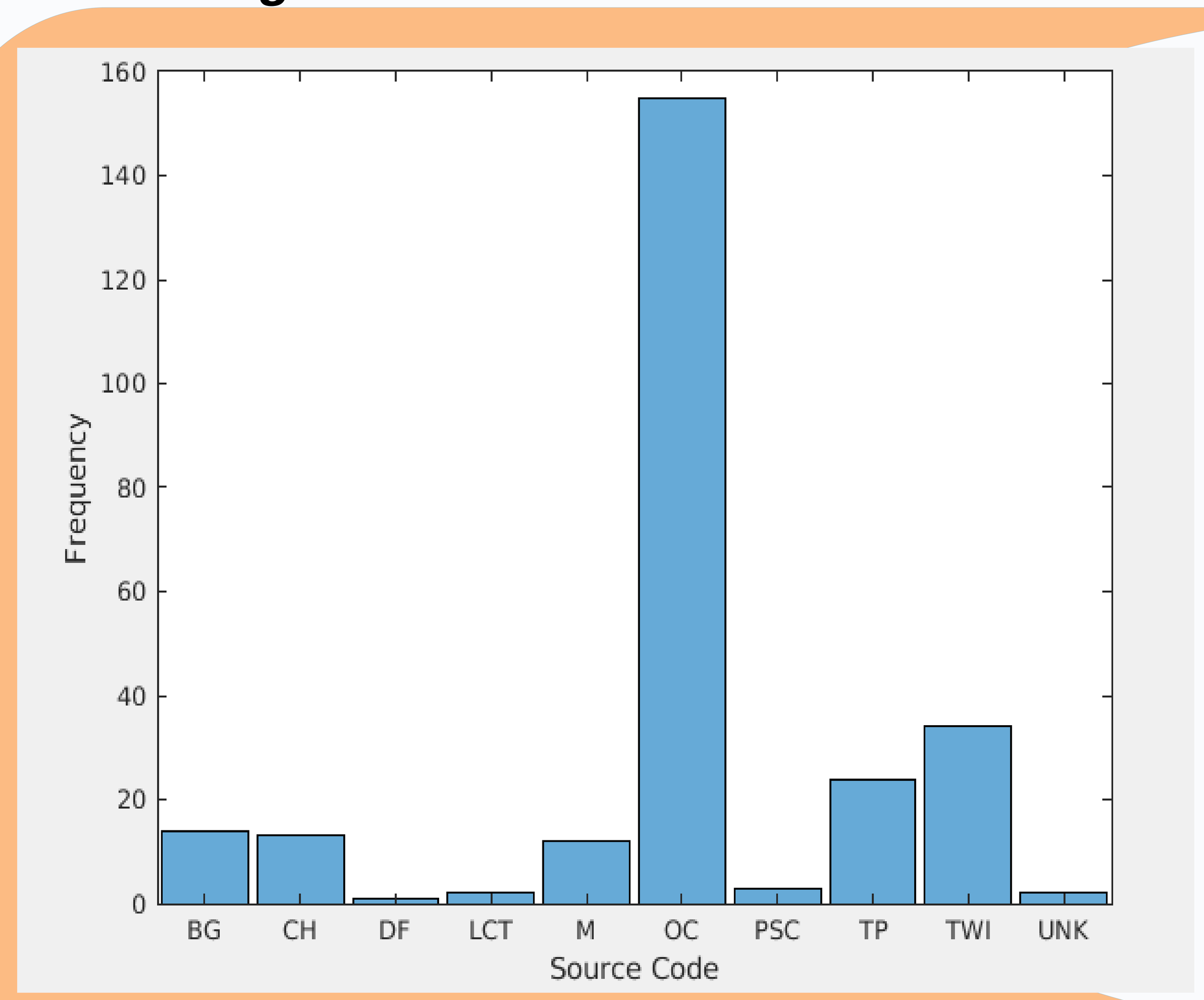
Methodology

Previous research on sourcing Absaroka obsidian artifacts using pXRF data employed Bayesian neural networks (BNNs) to develop sourcing models (Dalmas & Todd 2024). In this study, we compare the performance of BNNs and Random Forests to identify the most effective parameters and models for predicting obsidian sources. Both methods were tested on the same randomly sampled dataset of sourced obsidian from the Absaroka mountains.

Bugas-Holding

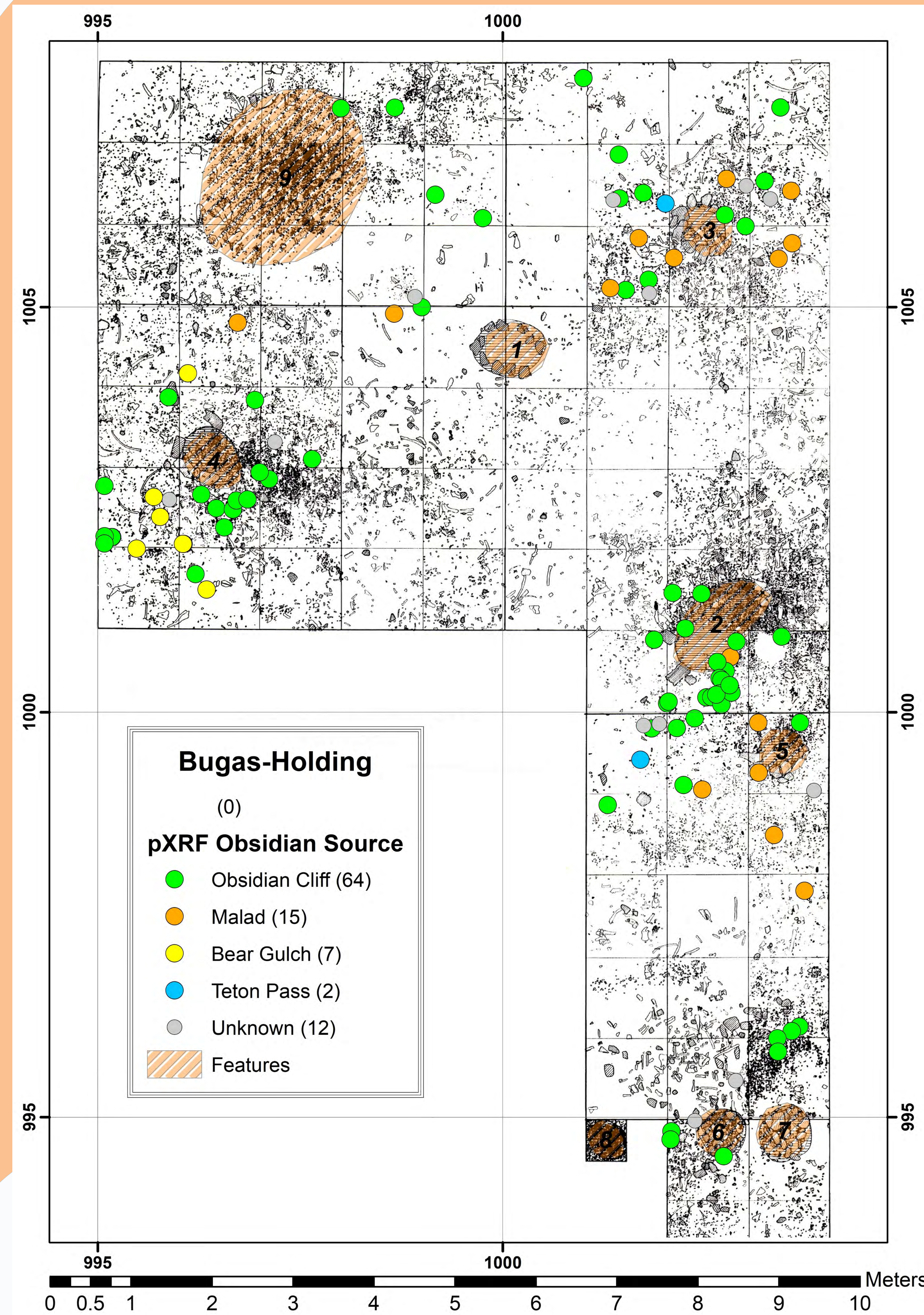
From 1983 to 1986, the Bugas-Holding site was excavated by the University of Wyoming, revealing a Late Prehistoric winter occupation. Findings included evidence of *Bison bison, Cervus canadensis* and *Ovis canadensis* processing, along with a large number of obsidian lithics (Rapson 1990). After identifying the best-performing machine learning model, sources for a sample of 100 randomly selected obsidian artifacts from the Bugas-Holding site were analyzed.

Training Data



Distribution of sourced obsidian from the GRSLE project used for model validation.

Obsidian artifacts were originally sourced by Richard E. Hughes at the Geochemical Research Laboratory in Sacramento, California, and this data served as our control for sourcing accuracy. The same artifacts were subsequently scanned using pXRF to evaluate the potential performance of field-based sourcing methods.



Using the best-performing BNN model, we predicted source probabilities for pXRF-scanned obsidian artifacts from the Bugas-Holding site in Wyoming. The predicted obsidian source results are presented above.

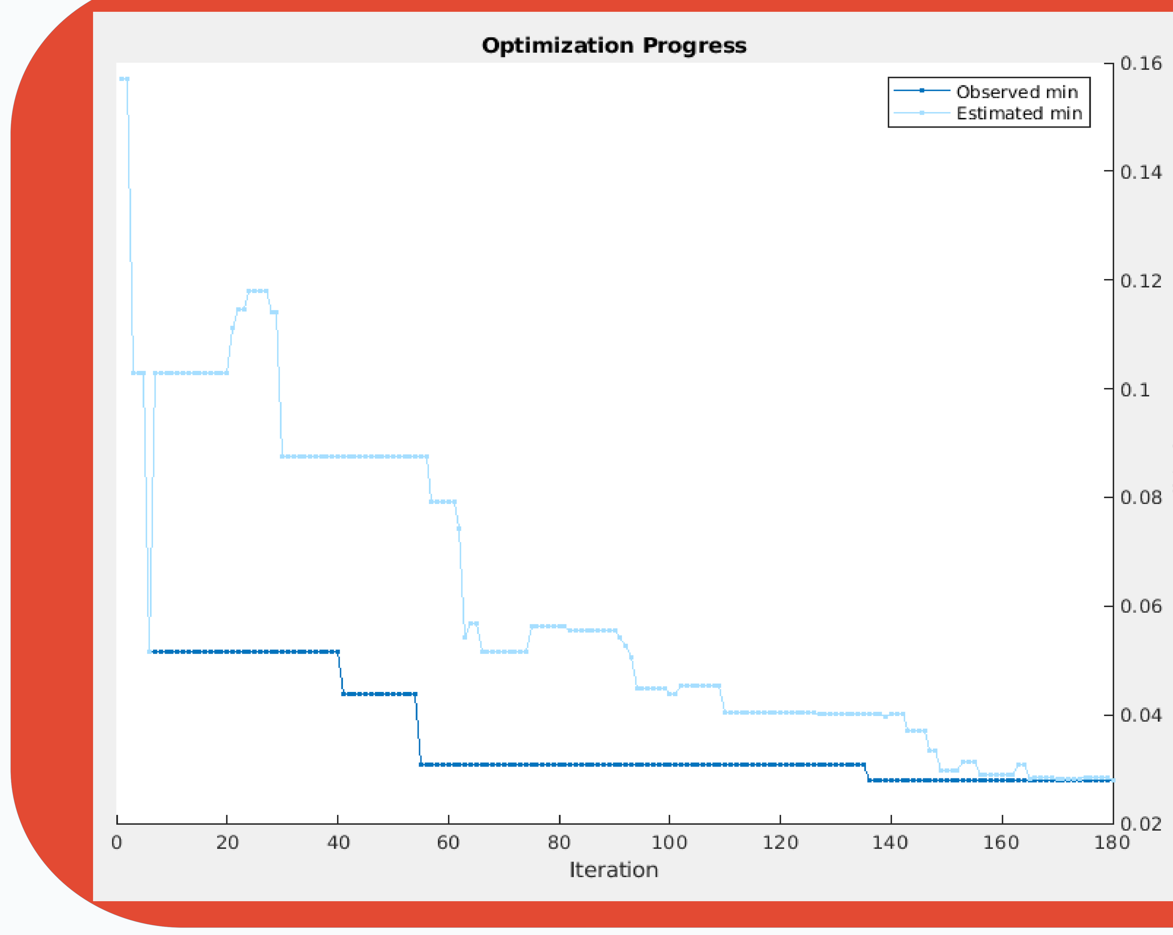
Random Forest vs BNN

Interestingly, the best-performing Random Forest model relies on several trace elements that are not significant predictors in the BNN, including rhodium (Rh), iron (Fe), manganese (Mn), gallium (Ga), and silver (Ag).

On average, the Random Forest models resulted in 6.4% more misclassifications than the BNNs.

The greater performance of the BNNs is likely due to the inclusion of rubidium (Rb) as a hyperparameter. Previous studies have highlighted the importance of Rb in obsidian source identification, supporting its role as a key predictor in the model.

Optimization - BNN

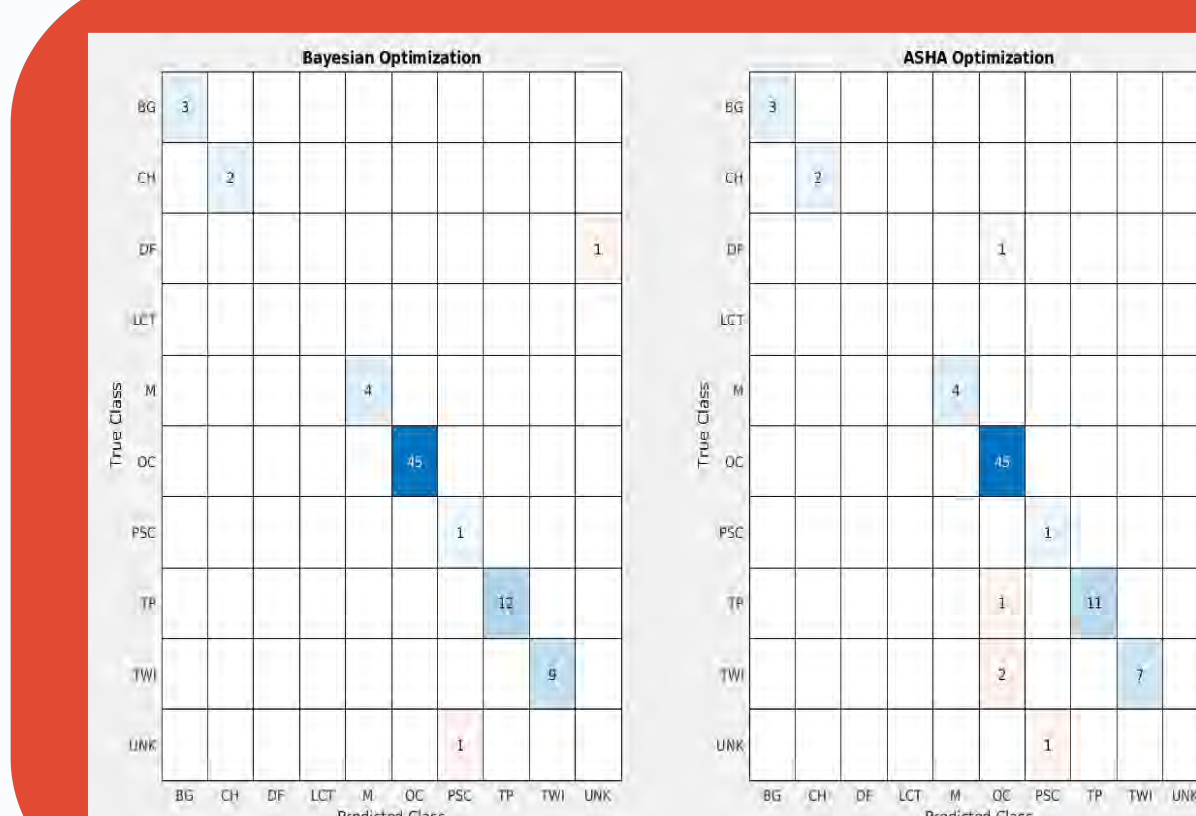


Optimization progress across iterations for both Bayesian optimization and the Asynchronous Successive Halving Algorithm (ASHA) in classification model tuning. While ASHA is less computationally intensive than Bayesian optimization, it resulted in a higher proportion of classification errors on the training dataset. Optimized classifiers were identified when validation loss reached an observed minimum greater than 0.03.

Handling Unknown Sources

Two unknown classifications were identified in the training data using EXRF. To improve the identification of samples from unknown sources, we implemented a posterior confidence threshold of 0.95. Samples that did not meet this threshold were classified as originating from an unknown source, indicating variability not adequately captured in the training dataset.

Classification - BNN



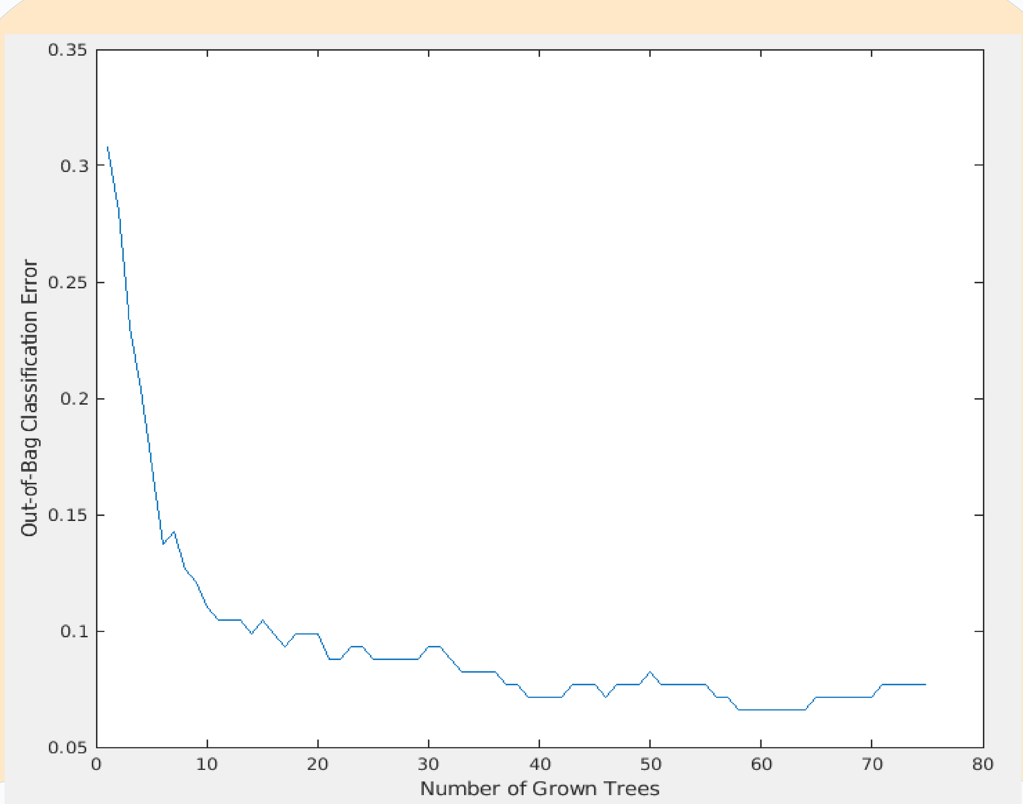
Classification and hyperparameter optimization results for Bayesian optimization and ASHA. Colors scale with the proportion of correct (blue) and incorrect (red) classifications.

Bayesian and ASHA optimization both identified the best performing model as:

$$Source \sim Nb + Rb + Sr + Y + Zn$$

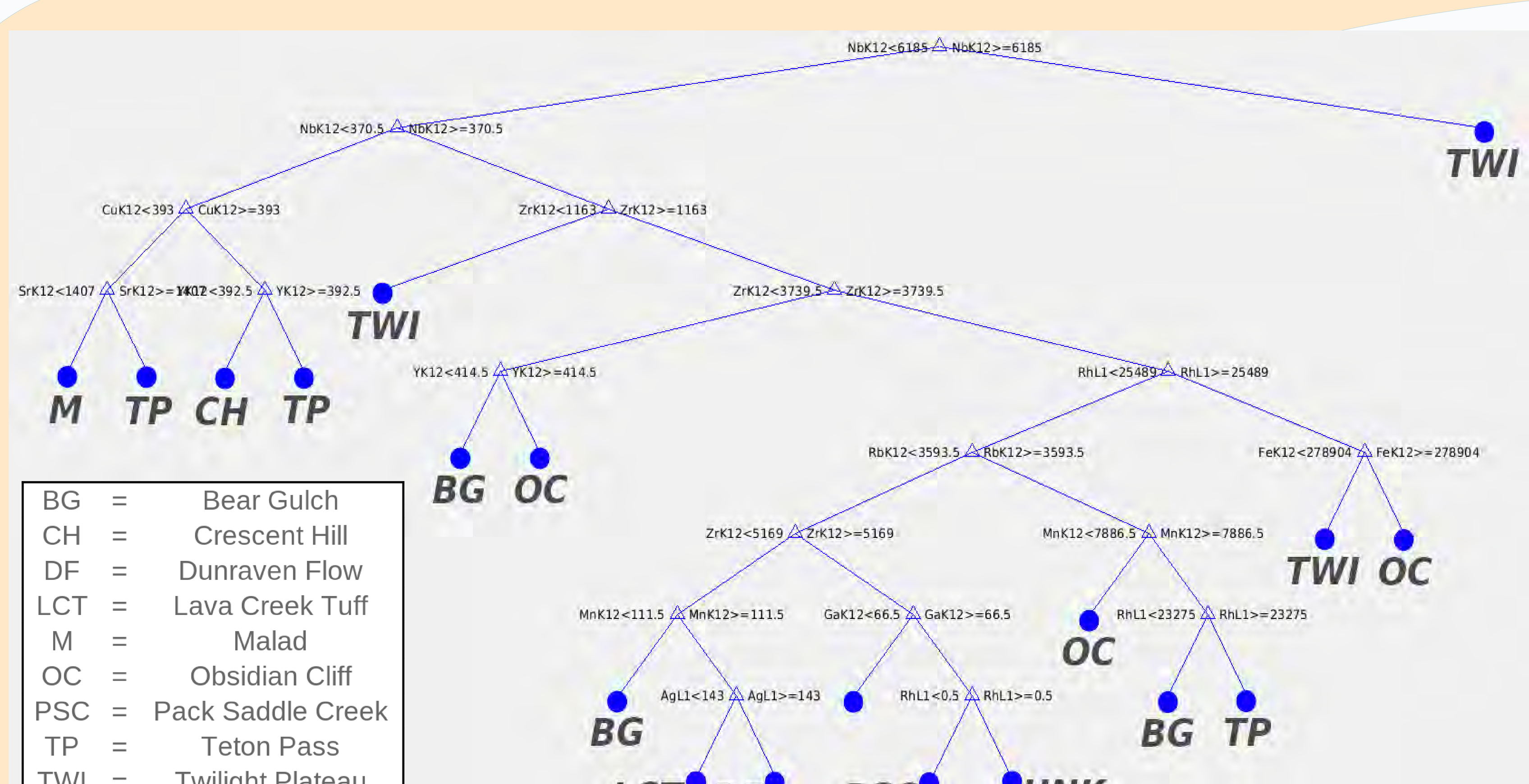
with hyperparameter Rb

Optimization – Random Forest



Out-of-Bag classification error across number of trees. Classification errors reach their minima at ~0.07 demonstrating a considerably low error rate.

Classification – Random Forest



BG	=	Bear Gulch
CH	=	Crescent Hill
DF	=	Dunraven Flow
LCT	=	Lava Creek Tuff
M	=	Malad
OC	=	Obsidian Cliff
PSC	=	Pack Saddle Creek
TP	=	Teton Pass
TWI	=	Twilight Plateau

The classification tree shown represents the best-performing Random Forest model, determined by the lowest classification error. The values to the left and right of each node indicate the range of net reads for each element, while the terminal nodes (tips) display the predicted obsidian source.

Overall, the iterative Random Forest models achieved a median consistency of 0.935, with an approximate classification error rate of 6.5%.

Field sourcing of obsidian artifacts will reduce the disturbance of in situ archaeological materials by eliminating the need to remove them for sourcing analysis. This approach demonstrates the pXRF data provides ample resolution for obsidian sourcing. Our findings bolsters the GRSLE Project's goal of minimizing site disruption through detailed field recording and 'catch and release' archaeology.

Acknowledgments

Special thanks to the University of Utah's Red Lab in the Department of Geography for providing access to and training for their Bruker Tracer Titan III pXRF. We would also like to thank the U.S. Forest Service and Shoshone National Forest for their continued support of the GRSLE project.