SciAps
ApNotes
May 2023

Endangered Species
& Region of Origin ID

Multi-element provenance analysis of endangered tropical hardwoods

Caelin Celani & Karl Booksh University of Delaware

Introduction

Handheld LIBS offers an intriguing option for classification of Dalbergia species (rosewoods) because there is a strong correlation between Dalbergia species and geographic origin. The metal signature of the soil, as transferred into the wood, may form a pathway to determine both species and growing region of Dalbergia lumber and products.

SciAps Z-300 and Z-200 are now the Z-903 and Z-902.

As rosewoods are the world's most trafficked wild product, accounting for nearly 35 percent of all seized illegal flora and fauna[1], one potential application for handheld spectroscopic instrumentation is to help combat the international illegal timber trade by rapidly and reliably identifying timber from endangered species in the field or at ports of entry. By applying chemometric and machine learning classification models, the species and provenance may be rapidly discerned. LIBS combined with chemometric analysis offers an alternative route to policing this trade via field detection, with handheld instrumentation being used by customs agencies at the port of entry.

Method

Two different sets of Dalbergia timber exemplars were analyzed by LIBS. The first study consisted of 90 samples taken from 9 tropical hardwoods —seven Dalbergia and two other tropical hardwoods from similar geographic regions. These samples were measured in a three-by-two grid with 10 cleaning shots and 16 data shots per spot (96 averaged spectra) with a SciAps Z-200 LIBS spectrometer, with a spectral range of 190 to 620nm. In the second study, samples consist of 159 wood chips across 12 Dalbergia species/classes (Table 1) with class assignment based on vendor-reported species and verified by Coplen et al.[2] Each wood chip (Scheme 1) was analyzed as is—i.e., no sample preparation—with a SciAps Z-300 LIBS spectrometer with the following parameters: 12 locations in a single point (one-by-one grid) with one cleaning shot and four averaged collection shots. The main difference between the Z-200 and Z-300 LIBS is that the Z-300 has an extended spectrometer range from 190 nm out to 950 nm.

In Study 1[3], spectra were preprocessed with a Whittaker filter and a Savitzky-Golay first derivative smoothing filter. The absolute value and square root of each value was taken, followed by decluttering with external parameter orthoganalization, and naïve variable selection using an arbitrary intensity cutoff.

For Study 2, collected spectra were normalized to the mean intensity of each spectrum to account for shot-to-shot variance, then separated into training and validation sets via bootstrapped Latin partitioning by taking all spectra associated with a fixed percentage of wood chips from each species (stratified by sample) before further preprocessing.

Preprocessing included baseline correction and smoothing with a Savitzky-Golay first derivative filter using a second order polynomial and 15-point window. An absolute value and square root of each variable was then taken. Lastly, variable selection was naïvely completed by removing all variables less than the mean value of each spectrum in the training set. This data were then mean centered and variance scaled (auto scaled) prior to use in various classification models.

Scheme 1: Example wood chip analyzed in Study 2.

Class	Number of Wood Chips	Location	Genus Species	Common Name
1a	12	Honduras	Dalbergia Stevensonii	Rosewood
2a	11	Madagascar	Dalbergia Spp.	Rosewood
3b	6	Para, Brazil	Dalbergia Spruceana	Rosewood
4a	11	Surinam	Dicorynia Paraensis	Angelique
5a	16	Brazil	Dalbergia Cearensis	Kingwood
6a	13	Malaysia	Dalbergia Spp.	Indonesian Rosewood
7a	15	Brazil	Dalbergia Futascans	Tulipwood
9b	6	Madagascar	Dalbergia Maritima	Rosewood
10a	17	Tanzania	Dalbergia Melanoxylon	African Blackwood
11a	8	Mexico	Caesalpinia Platyobe	Chakte Viga
12a	13	India	Dalbergia Latifolia	Indian Rosewood
14a	17	Brazil	Dalbergia Nigra	Brazilian Rosewood

Table 1: Class Information of Analyzed Samples in Study 2

Results & Discussion

In Study 1, data were analyzed with k-nearest neighbors (k-NN) and partial least squares discriminant analysis (PLS1-DA). K-NN yielded 90 percent classification success. PLS1-DA was used to build a series of one-versus-all decisions in a tree-style classification (Figure 1), which resulted in prediction sensitivity of 0.97 and specificity of 0.99.

In Study 2, data were analyzed with five classification algorithms: PLS2-DA, k-nearest neighbors (k-NN), classification and regression trees (CART), random forests (RF), and support vector machines (SVM), then compared using Cohen's kappa statistic for overall classification performance. Kappa is an adjusted accuracy measure, designed to account for random chance correct classifications and can be defined as $\frac{p_0 - p_e}{1 - p_e}$ where p_0 is accuracy and p_e is the no information rate. It can be seen from Figure 2 that even with minimal preprocessing and naïve variable selection, SVM & RF classify with K> 0.80 across all species, and k-NN and CART classify with K> 0.50 across all species, and PLS2-DA can classify with only >0.25 percent adjusted accuracy across all species.

The benefit of chemometrics in LIBS analysis is that targeted elemental line identities are not needed for classification problems such as this. Instead, the combination of all included wavelengths contributes to model performance, giving the analyst the ability to retroactively identify wavelengths (and therefore chemistry) of importance.

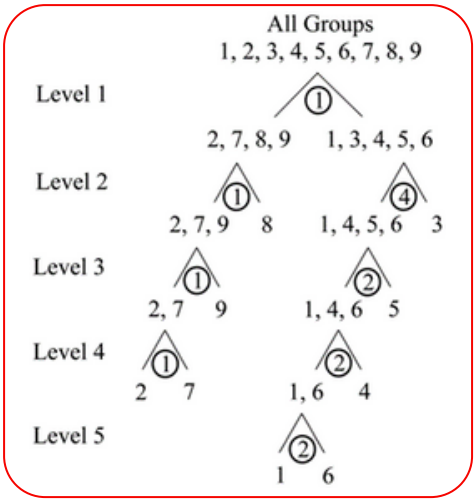

Figure 1: Decision tree model for calibration, training, and validation sets. Circled number indicates number of latent variables used at each branch for the PLS1-DA model.

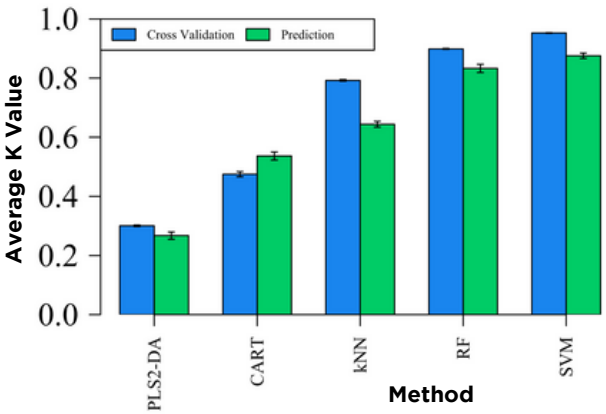

Figure 2: Bar plot showing the average k coefficient across classification algorithms for the validation and prediction sets. Error bars indicate the standard error for ± 1σ.

Summary

This case study shows the power of the SciAps LIBS Z-200 and Z-300 for research level applications. Through simple data collection of many samples, macro- and micronutrients are detected by the spectrometer and used for multivariate classification of Dalbergia. After minor preprocessing and unguided variable selection, highly accurate prediction models can be generated for potential field use.

References

[1] Zhu, A. L. China's Rosewood Boom: A Cultural Fix to Capital Overaccumulation. Ann. Am. Assoc. Geogr. 2020, 110 (1), 277–296.

[2] Coplen, T. B.; Qi, H.; Jordan, J. A.; Miller, K.; Espinoza, E.; Tarbox, L. V.; Wieser, M. E. Chemical and isotopic compositions of tropical wood samples: U.S. Geological Survey data release https://www.sciencebase.gov/catalog/item/5e345d42e4b0a79317daf1fb (accessed Sep 14, 2022).

[3] Celani, C. P.; Lancaster, C. A.; Jordan, J. A.; Espinoza, E. O.; Booksh, K. S. Assessing Utility of Handheld Laser Induced Breakdown Spectroscopy as a Means of Dalbergia Speciation. Analyst 2019, 144 (17), 5117–5126. https://doi.org/10.1039/C9AN00984A.

Multi-element Provenance Analysis of Endangered Tropical Hardwoods was compiled by Caelin Celani, a fifth-year graduate student (2022-2023 academic year) in Analytical Chemistry at the University of Delaware in Dr. Karl Booksh's lab studying applications of chemometric analysis methods using handheld instrumentation across various applications.

SciAps
7 Constitution Way
Woburn, MA 01801
www.sciaps.com | +1 339.927.9455