Near-Infrared Spectroscopy Aids Ecological Restoration by Classifying Variation of Taxonomy and Phenology of a Native Shrub

Brecken C. Robb  
*Boise State University*

Peter J. Olsoy  
*Boise State University*

T. Trevor Caughlin  
*Boise State University*

Stephanie J. Galla  
*Boise State University*

Marcella R. Fremgen-Tarantino  
*Boise State University*

*See next page for additional authors*

**Publication Information**  

For a complete list of authors see article.
Authors
Brecken C. Robb, Peter J. Olsoy, T. Trevor Caughlin, Stephanie J. Galla, Marcella R. Fremgen-Tarantino, Jordan D. Nobler, and Jennifer S. Forbey

This article is available at ScholarWorks: https://scholarworks.boisestate.edu/bio_facpubs/728
Near-infrared spectroscopy aids ecological restoration by classifying variation of taxonomy and phenology of a native shrub

Brecken C. Robb1,2, Peter J. Olsoy1, Jessica J. Mitchell3, T. Trevor Caughlin1, Donna M. Delparte4, Stephanie J. Galla1, Marcella R. Fremgen-Tarantino1, Jordan D. Nobler1, Janet L. Rachlow5, Lisa A. Shipley6, Jennifer S. Forbey1

Plant communities are composed of complex phenotypes that not only differ among taxonomic groups and habitats but also change over time within a species. Restoration projects (e.g. translocations and reseeding) can introduce new functional variation in plants, which further diversifies phenotypes and complicates our ability to identify locally adaptive phenotypes for future restoration. Near-infrared spectroscopy (NIRS) offers one approach to detect the chemical phenotypes that differentiate plant species, populations, and phenological states of individual plants over time. We use sagebrush (Artemisia spp.) as a case study to test the accuracy by which NIRS can classify variation within taxonomy and phenology of a plant that is extensively managed and restored. Our results demonstrated that NIRS can accurately classify species of sagebrush within a study site (75–96%), populations of sagebrush within a subspecies (99%), annual phenology within a population (>99%), and seasonal phenology within individual plants (>97%). Low classification accuracy by NIRS in some sites may reflect heterogeneity associated with natural hybridization, translocation of nonlocal seed sources from past restoration, or complex gene-by-environment interactions. Advances in our ability to detect and interpret spectral signals from plants may improve both the selection of seed sources for targeted conservation and the capacity to monitor long-term changes in vegetation.

Key words: ecological restoration, functional restoration, near-infrared spectroscopy, phenotypes, sagebrush

Implications for Practice
- Near-infrared spectroscopy (NIRS) uses spectral traits of plants to differentiate vegetation based on taxonomy and phenology.
- NIRS offers an emerging restoration tool that can detect changes in plant composition, distribution, and functional traits that are not obvious from morphology.
- NIRS has the potential to detect taxonomic diversity associated with natural hybridization and reseeding or translocation of local and nonlocal plant sources.
- NIRS can help increase the spatial and temporal scale at which plant traits can be used to monitor restoration outcomes.

Introduction
Pervasive landscape alterations from climate change and human disturbance substantially affect vegetation and associated wildlife communities (Coop et al. 2020). These changes are often met with ecological restoration practices meant to prevent further loss of native species and restore vegetative communities.

Past and current restoration efforts include reseeding with locally sourced native seed (Brabec et al. 2015), admixing individuals from different sources to maximize adaptive potential (Bucharova et al. 2019), and introducing species that may shift community types but avoid extirpation (Guerrant Jr &
Kay 2007). These restoration practices can alter taxonomic composition of plants (Carlucci et al. 2020), associated communities (e.g. herbivores, Pyke et al. 2020; fungi, Hovland et al. 2019), land use and cover (Cox & Anderson 2004), and ecosystem health and services (Reynolds et al. 2012). The high economic costs of restoration (Boyd & Davies 2012) demand better tools to properly classify and monitor plants and their functional traits for restoration.

Monitoring restoration outcomes requires an understanding of past restoration practices and optimizing present ones; this generally starts with classifying taxa (e.g. species) and functional traits (e.g. morphology, physiology, and phenology). Classifying plants within a community after translocation or reseeding can identify which plants were resilient to restoration (Davidson et al. 2019). Differentiating species can also reveal the extent of local versus foreign recruits from translocation efforts (Keller et al. 2000) and invasive species (Somers & Asner 2013). Landscapes characterized by morphologically distinct plant species may be relatively simple to measure and monitor (Pasquarella et al. 2016) compared to landscapes dominated by plant species that are difficult to classify due to complex or inconsistent morphological traits.

Sagebrush steppe ecosystems are characterized by complex landscapes that contain individual plants from the genus Artemisia that can vary morphologically within a species and others that are morphologically indistinguishable among species and subspecies (Jaeger et al. 2016). Sagebrush exhibits high intraspecific genetic diversity and has the ability to hybridize between species and subspecies (Bajgain et al. 2011; Richardson et al. 2012; Davidson et al. 2019), which further complicates classification. The diverse intraspecific phenotypes in sagebrush, including flowering phenology and phytochemistry (e.g. secondary metabolites), are underpinned by a combination of genetics, environmental conditions (i.e. phenotypic plasticity), and complex gene-by-environment (G×E) interactions (i.e. heritable plasticity; Nicotra et al. 2010; Karban et al. 2014; Richardson et al. 2017). Given the genetic basis of these morphological traits, it is possible that sagebrush is adapted to local conditions, which should influence where seeds are collected for restoration efforts (Davidson & Germino 2020). Alongside the natural complexities, sagebrush-dominated ecosystems have a convoluted history of management. In the 1940s, herbicide, controlled burning, and manual techniques (e.g. plowing) cleared the land and increased grass production for livestock (Pechanez et al. 1954). After the Federal Land Policy and Management Act was passed in 1976, wildlife-oriented management, restoration, and conservation became a priority for federal land managers (Knick & Connelly 2011). Despite substantial efforts to restore sagebrush ecosystems, sagebrush habitat continues to decline (Arkle et al. 2014; Requena-Mullor et al. 2019) and threaten the survival of species that rely on it (e.g. sage-grouse [Centrocercus spp., Conover & Roberts 2016]; pygmy rabbits [Brachylagus idahoensis, Thines et al. 2004]). To protect and restore sagebrush steppe habitats, we need to identify which species and subspecies of sagebrush exist across the landscape.

Unfortunately, classifying sagebrush taxa based on morphology alone is challenging (Rosentreter et al. 2021) and misclassification may lead to errors in land cover maps (Fremgen-Tarantino et al. 2021). Although genetic tools for classification of sagebrush are emerging (Richardson et al. 2012), there is a need for inexpensive and rapid approaches to identify and monitor specific sagebrush taxa in the field. Near-infrared spectroscopy (NIRS) can meet this need by detecting distinctive phytochemical traits (Vance et al. 2016). NIRS has been shown to be as accurate, if not more, in classifying vegetative phenotypes than manual approaches (Espinoza et al. 2012). Furthermore, spectral information collected by NIRS that differentiates taxa and phenotypes of interest can be scaled up to landscape and global extents with air- and satellite-borne hyperspectral sensors (Ustin & Middleton 2021). We used NIRS to differentiate sagebrush across four biological scales to demonstrate how spectral signals can detect taxonomic and temporal variation that can benefit ecological restoration, including: (1) sagebrush species within a community, (2) geographically distinct populations within a single subspecies, (3) annual variation within a single population, and (4) seasonal variation within individual plants.

Methods

Study Sites

Sagebrush leaves were collected from four field sites in Idaho, U.S.A: the Magic site (43°14′N, 114°19′W, elevation 1,465–1,480 m), Cedar Gulch site (44°41′N, 113°17′W, elevation 1,890–1,940 m), Craters site (42°57′N, 113°23′W, elevation 1,300–1,650 m), and Raft River site (42°9′N, 113°24′W, elevation 1,380–2,140 m). Previous research on Greater Sage-grouse (Centrocercus urophasianus; Ulappa 2011; Fremgen 2015) and pygmy rabbits (Ulappa 2011; Nobler 2016) at these sites demonstrated high phenotypic diversity among sagebrush species. All study sites were predominantly sagebrush (Artemisia spp.) but varied in taxonomic (e.g. species) and functional (e.g. morphology, phytochemistry) traits within the sagebrush community and environmental conditions (i.e. terrain, elevation, climate, year, and season collected) (Fig. 1).

Taxonomic and Temporal Phenotypes

Taxonomic phenotypes included sagebrush species and subspecies and geographically distinct populations within the Wyoming big sagebrush subspecies (Artemisia tridentata ssp. wyomingensis). Sagebrush species and subspecies were identified prior to NIRS analysis using plant morphology, environment type (e.g. elevation, soil type, and depth; Rosentreter 2005), and phytochemical profiles identified in previous research (Fremgen-Tarantino et al. 2020; Olsoy et al. 2020). Sagebrush taxonomic groups included Wyoming big, three-tip (Artemisia tripartita), black (Artemisia nova), low (Artemisia arbuscula), and “dwarf” sagebrush (Fig. 1). Although there are morphological and phenotypic differences between black and low sagebrush (e.g. stem color, stem stature, and soil type; Rosentreter 2005), samples were identified as...
dwarf when they had smaller morphological size relative to big sagebrush species at the site or occurred within microtopographically unique patches but did not have clear morphological features to distinguish them as black or low.

Temporal phenotypes represent phenological states and included year (2013 vs. 2015) within a single Wyoming big sagebrush population at one study site (Magic) and season (summer vs. winter) within individual plants of Wyoming big sagebrush at two different study sites (Magic and Cedar Gulch). Within-species analysis (classifying populations, year, and season) focused on Wyoming big sagebrush because it was the only subspecies present at all of our study sites.

Sample Collection
Sampling collection methods were standardized across sites and seasons. To reduce potential noise from user differences, individuals collecting sagebrush tissues within a site collected all morphotypes across all seasons and a subset of collectors were consistent across sites. Approximately 2.0 g (wet weight) of biomass was clipped per plant and stored in individual bags on ice until transfer to −20°C freezers. Sagebrush samples from Magic (n = 1,089) and Cedar Gulch (n = 625) were representative of plant parts consumed by pygmy rabbits (for the broader purpose of that study, see Olsoy et al. 2020) and included both stems and leaves. Sagebrush samples from Craters (n = 94) and Raft River (n = 263) were representative of plant parts consumed by
Greater Sage-grouse (for the broader purposes of that study, see Fremgen-Tarantino et al. 2020) and included only leaves.

**NIRS Analysis**

The ASD FieldSpec 4 spectroradiometer (Malvern Panalytical, Westborough, MA, U.S.A.) was used to measure continuous near-infrared wavelength reflectance (350–2,500 nm) on samples. Each sagebrush sample was ground (~2 mm) with liquid nitrogen and dried at 60°C for 48 hours and then spread homogeneously on a black surface within a sealed clear plastic bag. After calibrating and optimizing the spectroradiometer, NIRS reflectance was collected with a contact probe (Fig. 2). Thirty replicate scans were collected per sample. The instrument was

![Near-infrared spectroscopy classifies sagebrush](image-url)
recalibrated and optimized every 15 samples using a white reference material. The replicate scans were checked for visual outliers using Camo Analytics (now AspenTech) Unscrambler spectral plotting interface. Any spectral signatures that were the result of mis-scans (e.g. scan signatures at or close to 0 or 100% reflectance) were removed. Replicate spectral scans were then averaged to one spectral profile. Each spectrum was converted to absorbance values using a log(1/R) transformation, where R is reflectance. After visual inspection of reflectance spectra, spectral absorbance values were truncated to 450–2,350 nm to eliminate systematic noise near spectrometer detection limits. To minimize error, all NIRS steps were collected in one lab by the same researcher using this standardized approach.

**Statistical Analysis**

All response variables (i.e. species, population, year, and season) were joined with their associated near-infrared (NIR) spectra and imported into R 3.6.2 (R Core Team 2020). The e1071 package (Meyer et al. 2020) in R was used to perform support vector machine (SVM) analyses to classify taxonomic and temporal phenotypes of sagebrush using NIR spectra. The SVM type used was C-classification, with a linear kernel and C value of 1.0 (i.e. large margin). Each model was independently calibrated and validated using 75:25 train-to-test datasets. Samples were assigned to training and testing datasets in a stratified manner to ensure representative selection among source populations and phenotypes. The resulting confusion matrices from the SVM analyses were used to evaluate the overall accuracy of NIRS to classify phenotypes. Kappa statistics, which reflect the difference between actual agreement and agreement expected by chance, were obtained for each model. Further accuracies were calculated to differentiate the proportion of phenotypes on the ground that were accurately classified by NIRS (i.e. producer’s accuracy) versus the proportion of phenotypes classified by NIRS that were actually present on the ground (i.e. user’s accuracy).

**Results**

Our results demonstrated that NIRS can accurately classify taxonomy in sagebrush. The accuracy of NIRS to classify sagebrush species was high, but varied among study sites (Table 1). At the Magic site, the overall accuracy of the training dataset to predict the sagebrush species of the test dataset was 95.7%, with a Kappa statistic of 0.79. At Cedar Gulch, the overall accuracy was 76.4%, with a Kappa statistic of 0.25. At Craters, the overall accuracy was 96.2%, with a Kappa statistic of 0.91. At Raft River, the overall accuracy was 74.6%, with a Kappa statistic of 0.00. The accuracy of NIRS in classifying geographically distinct populations within Wyoming big sagebrush was 98.7%, with a Kappa statistic of 0.98 (Table 2).

NIRS can also accurately classify temporal phenotypes within a population and within individual plants (Table 3). NIRS classified years of collection (2013 vs. 2015) for Wyoming big sagebrush at one study site with an overall accuracy of 99.5% and Kappa statistic of 0.99. NIRS classified seasons (summer vs. winter) within the same individual Wyoming big sagebrush plants at two study sites with an overall accuracy of 99.2% and Kappa statistic of 0.97.

**Discussion**

Overall, we demonstrated that NIRS can classify species within a genus within and across communities, differentiate geographically distinct populations within a subspecies, and detect annual changes within a population and seasonal changes within...
individual plants. Specifically, NIRS provides a tool for detecting taxonomic phenotypes and phenological states of plants that may not be obvious from morphology alone. Given the standardized collection methods used between sites, the consistency of individuals collecting data at each site, and the ability of NIRS to differentiate between species and locations over time, we are confident NIRS is detecting biologically relevant signals. Although we focused on a widespread plant genus in western North America, this case study showcases how spectral traits could aid in monitoring changes in vegetation within and across landscapes that could benefit restoration ecology.

Proper classification of sagebrush is critical for locating desired species or populations for reseeding of locally adapted (e.g. drought tolerant; Barron et al. 2020) or functional (e.g. palatable; Olsoy et al. 2020) phenotypes and identifying which seed sources of sagebrush survive and reproduce after restoration projects. When distinct morphological characteristics of sagebrush allow for proper identification in the field, the accuracy of NIRS to classify species increases. For example, three-tip, which is easy to identify based on deeply lobed leaf morphology (Rosentreter 2005), was still classified accurately with a very small sample size ($n = 8$), compared to the large sample size of Wyoming big sagebrush plants ($n = 807$) at the Magic field site. We propose that when morphotypes are distinct among species (e.g. Wyoming vs. three-tip [Fremgen-Tarantino et al. 2020]; Wyoming vs. black [Frye et al. 2013]), large sample sizes may not be required for NIRS to correctly classify species.

However, distinct morphology may not always indicate different species. Despite distinct morphology between dwarf and Wyoming big sagebrush at Cedar Gulch, the majority of the dwarf ($n = 100$) samples were classified as Wyoming (i.e. 92% of the time). NIRS data indicate that the dwarf taxa at Cedar Gulch are Wyoming big sagebrush with a small growth form that may reflect young age, shallow soil, or disturbance (e.g. browsing or mowing). In support of our NIRS classification, the dwarf and Wyoming big sagebrush morphotypes had similar levels of crude protein, coumarins, and monoterpenes (Olsoy et al. 2020), further indicating they are the same subspecies. On the other hand, 75% of the dwarf plants at Magic were classified discretely from Wyoming big sagebrush, suggesting that at least some of the dwarf morphotypes at this site represent a distinct taxonomic group, which was supported by distinct phytochemistry (Olsoy et al. 2020). Misclassification of dwarf species as Wyoming big sagebrush occurs both on-ground (this study) and in commonly used land cover datasets (Fremgen-Tarantino et al. 2021). Our results demonstrate that NIRS is capable of differentiating sagebrush species that are morphologically complex or ambiguous and could be used to rapidly validate species classifications that are remotely collected at higher spatial scales (i.e. unoccupied aerial systems and satellites) for conservation and restoration purposes.

To further pursue the idea that spectra can differentiate plants within a subspecies with indistinguishable morphology, we assessed if NIRS could classify populations within Wyoming big sagebrush. NIRS was able to predict source populations of samples with 99% overall accuracy (100% at Magic, 100% at Cedar Gulch, 94% at Raft River, and 83% at Craters), with the worst classification accuracy associated with the smallest sample size (i.e. Craters, $n = 94$). Classifying sagebrush species and differentiating unique phenotypes in the field is challenging, yet imperative to the restoration, monitoring, and future protection of sagebrush landscapes. Errors in classification in the field and by NIRS, such as those at Raft River where all of the plants (low, $n = 197$; Wyoming, $n = 65$) were classified as low sagebrush, may be due to unique environmental factors that elicit plastic responses in plants (Gratani 2014), high phenotypic and genetic variation (Richardson et al. 2012; Davidson et al. 2019), or hybridization (McArthur et al. 1988). Although NIRS can detect known genetic variation in model plant species, including wheat (*Triticum aestivum*; Rincent et al. 2018; Krause et al. 2019), poplar (*Populus nigra*; Rincent et al. 2018), and Eucalyptus (*Eucalyptus globulus*; Raymond & Schimleck 2002), it is unknown to what degree trait variation seen below the species level in sagebrush is determined by genetic variation, phenotypic plasticity, or heritable phenotypic plasticity. To elucidate these mechanisms in sagebrush, we recommend that NIRS be included in G×E studies (Richardson et al. 2017) as a high-throughput measure of phenotyping that could account for trait variation among and within plants (i.e. “modular” concept of plasticity; de Kroon et al. 2005; Forster 2015), especially studies with shared environmental conditions like common gardens. Such G×E studies, coupled with measures of stress (e.g. herbivory) and phytochemicals (e.g. auxins, crude proteins), can further tease apart the phenotypic underpinnings of NIRS signatures.

We suggest that NIRS can aid in restoration by correctly classifying sagebrush species in the field, reveal sites with high phenotypic variation of sagebrush, and improve our ability to select or avoid taxonomic groups for seed collection and planting (Erickson & Halford 2020). For example, further experimentation with NIRS could identify whether sagebrush that appears morphologically homogeneous are spectrally heterogeneous (i.e. divided among more classification classes) or homogeneous (i.e. shared within a classification class). Sites containing diverse taxonomic variation, as indicated by NIRS, may contain plants most appropriate for reseeding at sites with diverse microclimates (Broadhurst et al. 2008) and may benefit from restoration practices that use seed mixes of sagebrush species. The potential for NIRS to detect hybrids (Wan-hong et al. 2019) is important for selecting seeds for restoration, as hybrids may have higher fitness within ecotones but lower fitness than parents within parental habitats (Wang et al. 1997). To better characterize hybrids in the sagebrush steppe, NIRS analyses will need to be coupled with high resolution genetic data (Richardson et al. 2012) to classify and detect hybrid individuals and their parents across ecotones. To confirm that NIRS can detect hybridization and assess fitness consequences of hybrids, NIRS should be coupled to common garden studies that include both parents and hybrids. We anticipate that the application of NIRS to classify species and detect hybrids is most useful in taxonomically homogeneous landscapes, where preservation of the locally adapted genotype for future seed collection may improve successful reseeding within that site.
The ability to detect temporal variation in phenotypic responses of plants to environmental conditions may also benefit restoration efforts. We offer evidence that the annual and seasonal phenology of plants can be spectrally distinguished. NIRS accurately classified the year in which samples of Wyoming big sagebrush were collected (99.5%), as well as the season (>97%). The ability of NIRS to distinguish sagebrush between years may reflect plastic responses to biotic and abiotic stressors that changed leaf growth, senescence, or survival of plants (Wilt & Miller 1992). The ability of NIRS to distinguish sagebrush between seasons within individual plants may be associated with emergence and senescence of photosynthetic (ephemeral leaves; Rosentreter 2005) and reproductive phenology (Rosentreter 2005; Richardson et al. 2017). We propose NIRS is detecting changes in phytochemistry associated with these annual and seasonal changes (Wilt & Miller 1992). The next step is to link NIRS to these phytochemical changes as a spectral biomarker that could predict developmental stage and productivity traits of plants (see Rincent et al. 2018). Relatively dynamic NIRS spectra within a plant without an obvious change in environmental conditions or growth form could offer an early signal of biotic (e.g. pathogen) or abiotic (e.g. drought) stress. Temporal variation observed in NIRS within a season and species at previously restored landscapes may suggest variation in the sources of seed used for restoration, which could create a mismatch in reproductive phenology.

We used sagebrush as a case study to demonstrate that NIRS offers a valuable tool to discriminate among plant phenotypes, taxonomy, and phenology, which is important to the restoration of sagebrush-dominated landscapes. Broader use of NIRS could help monitor changes in plant composition, identify previously translocated plants that are distinct from local populations and are resilient to restoration, and identify phenotypically homogeneous or heterogeneous sites and potential hybrid zones. Using NIRS to quantify dynamics in leaf phenology could be used to detect biotic or abiotic stressors or identify mismatches in timing of reproductive stage among plants. However, realizing the benefits of NIRS for the conservation and restoration of plant communities across landscapes requires scaling up from benchtop measurements to remote sensing platforms. Scaling up is increasingly possible with greater availability of spectral sensors with increased electromagnetic ranges and resolutions and uncoupled aerial systems and satellite platforms (Ustin & Middleton 2021). NIRS data can be integrated with similar spectral traits across sensors to minimize the loss of resolution associated with scaling up spatially and temporally (Xiao et al. 2019). Finally, NIRS can help isolate and identify functional links between plant traits of interest (e.g. crude protein) and key spectral bands to develop more targeted remote sensors that align with restoration needs. Until scaling up is realized, NIRS of collected plants provides a restoration tool to classify phenotypes, identify locations where hybridization or past restoration influenced taxonomic diversity, select specific taxonomic phenotypes to optimize ongoing restoration practices, and monitor vegetative outcomes of restoration into the future.

Acknowledgments

We thank Boise State University, University of Idaho, Washington State University, Idaho State University, University of Florida, Bureau of Land Management, Idaho Department of Fish and Game, United States Geological Survey, United States Forest Service, The Nature Conservancy, and Dairy One Forage Lab for their support in this project. We would also like to thank C. L. Merriman, K. Gehlken, C. R. Milling, M. J. Camp, M. M. Crowell, L. A. McMahon, L. Cross, D. D. Musil, K. Graski, J. Peña, G. Lockwood, J. W. Connelly, R. Rosentreter, N. Glenn, the BCAL lab, and the sage-scape team for field collections, sample processing, and data analysis. Funding was provided by the National Science Foundation (DEB-1146194, DEB-1146368, DEB-1146166, IOS-1258217, OIA-1826801, OIA-1757324), BLM (#LO9AC16253, #LO9AC15385, #L16AC00137, #L17AC00306), USDA National Institute of Food and Agriculture (Hatch Project 1005876), M.J. Murdock Charitable Trust Partners in Science Program, Pittman-Robertson funds from Idaho Department of Fish and Game (IDFG), and Sigma Xi Grants-In-Aid. The data that support the findings of this study are openly available at http://doi.org/10.7923/30ca-6g41.

LITERATURE CITED


