[Study Information](#_3o9ssh2swz09)

[Design Plan](#_4mzf79vx2q6j)

[Sampling Plan](#_hu8o0vkz41nk)

[Variables](#_pec3rgxfolor)

[Analysis Plan](#_3mtn7m44krsg)

[Other](#_6wujw18ggcuz)

### Study Information

1. **Title (required)**

**Can Remotely-Sensed Spectral Data Capture Arctic Plant Biodiversity?**

1. **Authors (required)**

Shawn Schneidereit

1. **Description**

Changes in plant communities are one of the most distinct responses to global climate change, yet we lack quantification of plant diversity and composition in the biome experiencing the highest rate of warming – the Arctic. Traditional methods of measuring biodiversity involve field studies and visual surveys, which are both resource intensive and limited in their spatial and temporal resolution. The synthesis of remotely sensed earth observation data with local climatic and topographical conditions, presents itself as a potential cost-effective and standardized technique to monitor biodiversity on an ecosystem wide scale. While high-resolution spectral data are becoming increasingly available at multiple scales, there is little known about the drivers of spectral diversity when transitioning from species to community-specific spectral data, particularly in tundra biomes.

My goal is to evaluate the capabilities of using plot scale and remotely sensed spectral data to assess ecological traits. Understanding how environmental factors correspond with spectral diversity have implications for the feasibility of using hyperspectral data for remote biodiversity estimation.

**Purpose of study:**

This will encompass using spectral signatures for the identification of Arctic vegetation type, as well as quantifying influencers of spectral diversity. Specifically, species richness, species evenness, and exposed soil will be investigated, as these factors contribute to spectral diversity in complex ways.

**Research questions**

1: Can Arctic Vegetation types be identified based on the mean and variance of hyperspectral signatures?

2: How does spectral diversity relate to species richness, evenness, canopy cover, and soil-background?

3: What parts of the full spectrum best discriminate vegetation types?

4: Are closer measurements more similar than more distant measurements?

5: Do plot and airborne spectra differ? Do plot or airborne spectra show more variation among vegetation types?

6: Can airborne remotely-sensed spectral diversity be used to identify vegetation types at the island scale?

1. **Hypotheses**

**1: How do Arctic Vegetation types be discriminate based on the mean and variance of hyperspectral signatures?**

**H1**: Arctic vegetation types can be identified based on the mean of their spectral signature (H1a) and spectral variance of their spectral signature (H1b). Spectral variance will have a larger correspondence with vegetation type (H1c). When ordinated, spectral signatures will discriminate among vegetation types (H1d) and year of measurement (corresponds with change in sensor type) (H1e).

**H1o:** Arctic vegetation types do not differ based on the mean or variance of their spectral signatures, and do not discriminate when ordinated.

Predicted results

I predict that both the mean and variance of spectral signatures can be used to identify Arctic vegetation types. The diversity of a spectral signature correlates with the chemical, anatomical, and morphological traits of communities (Schweiger et al., 2017). Given the compositional difference between communities, this should translate into observable differences in spectral signatures.

I predict that spectral variance will have higher correspondence with vegetation types, as it captures the spectral complexity within each vegetation type (Wang et al., 2018). If spectral variance has a higher correspondence with vegetation types, this means that the vegetation types discriminate based on overall spectral diversity, which is variable at all spectral regions and is influenced by compositional differences, such as biological, functional, and phylogenetic diversity.

If the mean of spectral signatures has greater correspondence with vegetation type, this indicates that the spectral complexity of a vegetation type is less important than between-type differences in spectral signatures. These differences would relate to course optical differences at specific spectral regions, resultant from a limited number of structural or chemical differences. The correspondence between spectral mean and vegetation type is likely to be limited by spectral differences not being concentrated a one region of the spectrum. Spectral differences occurring at both low and high wavelengths, overall result in an in-between spectral mean that is similar and does not account for internal spectral complexity.

When ordinated, spectral signatures will discriminate by year, due to the potential difference of reflectance measurements made by the different sensor types used between 2018 & 2019, as well variation in flowering phenologyand community compositions between the two years.

**2: How does spectral diversity relate to species richness, evenness, and soil-background?**

**H2:** Vegetation type will significantly affect spectral diversity (H2a). Higher spectral diversity will correspond with both increases in species richness (H2b) and species evenness (H2c), with evenness having the stronger relationship (H2d). Soil-background cover will have the strongest positive influence on spectral diversity (H2e).

**H2o:** Species richness, evenness, and soil-background have no effect on spectral diversity, with no visible relationship observed.

Predicted results

Both increases in richness and evenness correlate with higher observed spectral diversity, across an array of different grassland/prairie habitats (Cavender‐Bares et al., 2017; Schweiger et al., 2017, 2018; Wang, Gamon, Cavender-Bares, et al., 2018a; Wang and Gamon, 2019). I predict evenness will have a larger correspondence with spectral diversity, as more even representations of species are likely to increase the detection of the variable chemical and physical structures that ultimately influence spectral diversity. If richness has larger correspondence to spectral diversity, this indicates that 1) sensor resolution is sensitive enough for the detect additional species, (potentially even when at low relative abundances) and 2) this presence of additional species sufficient to significantly increase spectral diversity.

I predict that soil-background is the strongest predictor of spectral diversity, as soil has distinctly different reflectance than vegetation (Gholizadeh et al., 2018). Furthermore, soils reflectance is variable depending on local microclimatic conditions such as moisture and roughness. This potential for plot level variability can therefore alter spectral signatures in more complex ways, resulting in soil having the largest correspondence with spectral diversity.

**3: Does band selection influence correspondence between spectral diversity and vegetation types**

**H3.** Both band selection via using a subset spectral region, as well as spectral zone unmixing will result in greater correspondence between spectral diversity and vegetation type (H3a). spectral zone unmixing will result in the selection of a small but spectrally diversity subset of bands (H3a), with overall regions in the visible part of the spectrum (400-700nm) having higher correspondence with vegetation type than bands in the near infrared range (700-100nm) (H3b).

**H3o:** Band selection does not visible influence the correspondence between spectral diversity and vegetation type. There are no specific spectral regions that are best suited to discriminate vegetation types.

Predicted results

Selecting a subset of bands will result in better discrimination of vegetation types, as reducing spectral dimensionality excludes variable and noisy regions of the spectrum. Spectral mixture analysis using spectral zone unmixing will result in finding what zones of the spectrum are best suited to differentia vegetation types. This is achieved by isolating bands that explain maximum between community variability, while exhibiting the minimum within community variability.

If spectral zone unmixing (SZU) results in the selection of a small subsets of bands, this indicates that specific structural or chemical differences explain most of the spectral variation between the two vegetation types. Due to shared life histories, the two vegetation types would have traits that are similar, aside from key divergent adaptations. These specific divergent traits would be most relevant for spectral discrimination. Furthermore, when collecting and disseminating spectral data, as well as running spectral analysis, high resolution measurement spectral bands/regions are more important for vegetation identification than coarsely capturing reflectance across the full spectrum.

If SZU results in the selection of a large sub-set of bands, this indicates that the differences between the two vegetation types are complex, with many regions of the spectrum (aka structural and chemical traits) being relevant and important for vegetation type discrimination. Furthermore, if SZU results in the selection of bands that correspond with and/or are not visually better than an informal literature derived selection of spectral bands, this indicates that structural and chemical differences may be relatively convergent between different Arctic vegetation types. This implies that trends between biological traits and spectral diversity may be generalizable, and that future research can aim to work at greater scales of observation, rather than synthesizing the findings at small sites. Furthermore, this is promising for the development of future satellites, as potentially selection of spectral regions may be sufficient to discriminate vegetation types.

If spectral regions in the visible spectrum (~400-700 nm) are found to best discriminate vegetation types, then chemical traits are primarily responsible for the key spectral differences between vegetation types. This spectral region corresponds with wavelengths carotenoid absorption (502-511 & 533 nm) green light reflectance (~550 nm), and chlorophyll absorption (644, 653, 672 nm). If spectral regions in the red-edge transition and near infra-red plateau (700-1000 nm) are found to best discriminate vegetation types, then physical structures are primarily responsible for the key spectral differences between vegetation types.

**4: Are closer measurements more similar than more distant measurements?**

**H4:** When tested for spatial autocorrelation, closer measurements will be more similar than more distant measurements (H4a), with no visible autocorrelation occurring across vegetation type areas (H4a).

**H4o:** There is no visible level of autocorrelation between measurements at any distance.

Predicted results

Spatial autocorrelation of measurements will be present but weak between proximal plots and negligible at a regional scale, due to the between plot variability of both species composition and environmental factors such wetness and roughness.

**5: Do plot and airborne spectra differ from each other? Does plot or airborne spectra show more variation among vegetation types?**

**H5:** Plot and airborne spectra do visually differ from each other (H5a), with plot spectra showing more variation both within vegetation types (H5b), and between vegetation types (H5c).

**H5o:** There is no visible difference between plot and airborne spectra, with neither accounting more variation among vegetation types.

Predicted results

I predict that plot spectra show more variance within vegetation types, as they measure a smaller spatial area, resulting in only subsets of the full variation present in a vegetation type to be observed. This higher sensor resolution captures more small-scale compositional variation in each individual measurement. If plot spectra show more variance between vegetation types, then these captured small-scale compositional differences are important factors for discriminating vegetation types. It also indicates that future work that aims to use remotely-sensed spectral data for biodiversity assessment should aim to have a pixel size smaller than individual plants and structures, ensuring that each measurement is of a single or small subset of species.

Airborne spectra how lower spatial resolution, with each pixel potentially including multiple species resulting in the combined spectral signatures to be measured by the sensor. I predict that this will homogenize spectral diversity, with less variation being observed between individual measurements within a vegetation type. Furthermore, reduced spatial resolution will increase the likelihood of capturing external environmental factors, such as exposed soil or canopy structure which are likely to add noise to measurements. This will result in less vegetation specific spectral information being captured, reducing the differentiation between types.

If airborne spectra do show more variance between vegetation types, then course optical differences and capturing the full potential compositional variation in each measurement are more important for discriminating vegetation types. Alternatively, there might be a correlation between vegetation type and environmental factors that alters the airborne spectral signatures, resulting in greater spectral variation among the vegetation types.

**6: Can airborne remotely-sensed spectral diversity be used to identify vegetation types at the island scale?**

**H6:** Airborne remotely-sensed spectral diversity can be used to identify vegetation types at an island scale (H6a), and when visually compared, will correspond with an existent mapping vegetation type (H6b).

**H6o:** Airborne remotely-sensed spectral diversity cannot be used to identify vegetation types at an island scale, and when visually compared does not correspond with an existent mapping of vegetation types.

Predictions

Airborne remotely-sensed spectra diversity should be able to identify vegetation types, based on categorical spectral classifications. Through isolating the most discriminative areas of the spectrum, when mapped new spatial classified vegetation types should visually correspond with an existent vegetation classification on Qikiqtaruk (Obu et al., 2017).

### Design Plan

In this section, you will be asked to describe the overall design of your study. Remember that this research plan is designed to register a single study, so if you have multiple experimental designs, please complete a separate preregistration.

1. Study type
   1. Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.
2. Blinding (required)

6.1. No blinding is involved in this study.

1. Study design (required)

My study aims to discriminate, and map vegetation types based on their hyperspectral signatures, as well as test the relationship between spectral diversity 1) species richness 2) species evenness 3) soil background and 4) scale of observation.

The analysis conducted will rely on plot level, as well as remotely sensed data collected Team Shrub in previous years of field work on Qikiqtaruk-Herschel Island. The hyperspectral signature of an individual or community is the unique expression of electromagnetic radiation interacting with physical plant structures (Schweiger et al., 2018). Different vegetation types have dissimilar chemical, anatomical, and morphological traits, which alter the absorbance and scattering of light, resulting in the distinct patterns reflectance that compose a spectral signature (Cavender‐Bares et al., 2017). It will be aimed to identify Herschel and Komakuk vegetation types based on the mean and variance in spectral signatures. This will use spectral data gathered from a regional field survey on the two vegetation types. Detailed data obtained from long-term plot monitoring, will be used to build a linear model that predicts how species richness, evenness, and soil-background influence spectral diversity. What spectral regions are most relevant at discriminating vegetation types will be investigated, by sub-setting the full spectrum into regions of bands. Band selection will occur through 1) informal literature derived selection of spectral bands and 2) performing spectral band unmixing for optimal band selection. Spectral measurements will be tested for spatial autocorrelation using a variogram. Plot spectral data will be compared to airborne hyperspectral data measured at a plane scale. These airborne spectra will also be used to identify vegetation types. Categorical spectral classifications will be created and applied to remotely sensed spectral data to create a predictive map of the spatial distributions of vegetation types across Qikiqtaruk- Herschel Island.

### Sampling Plan

In this section we’ll ask you to describe how you plan to collect samples, as well as the number of samples you plan to collect and your rationale for this decision. Please keep in mind that the data described in this section should be the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.

1. Existing data (required)
   1. Registration prior to accessing the data: As of the date of submission, the data exist, but have not been accessed by you or your collaborators. Commonly, this includes data that has been collected by another researcher or institution.

My study relies on pre-existing data, that was compiled by other researches in previous years of field work and open access data (“JPL | AVIRIS”, 2019; Obu et al., 2016). I have not been granted access to any of the data thus far, and therefore have not conducted any summary statistics or exploratory analysis. Analysis will commence upon completion of the preregistration process. In supplementary information given by one of my supervisors, I have been made aware that the variability within spectral signatures tend to be better at discriminating vegetation types than difference in means.

1. Explanation of existing data (optional)
   1. If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study.
2. Data collection procedures (required)

The analysis conducted will rely on plot level, as well as remotely sensed data collected by Team Shrub in previous years of field work (2018-19) on Qikiqtaruk-Herschel Island. 6 (1x1 m) long-term monitory plots have been established - each of the two vegetation communities (Komakuk and Herschel type). No blinding or randomization took place during the data collation process, as my analysis exclusively relied on previously collected data. From this existing data, point-framing data, and percent soil-background can be obtained. During point-framing sampling, present species are also recorded, which can be used to calculate the species biodiversity metrics of species richness and evenness. Regional plot scale hyperspectral data exists for both vegetation types. Vegetation type A is partitioned into plots and has data from both 2018 & 2019, while vegetation type B only has data available for 2019, obtained from a greater region without plot subdivision. Percent soil-background data across Qikiqtaruk exists at a drone scale of observation. Plane scale hyperspectral imagery is available from a mission conducted last summer by NASA, which passed over Qikiqtaruk on two occasions (early and late July)

1. Sample size (required)

For each of the two vegetation types, 6 long term monitoring plots exist, which will all be utilized in this analysis. Each plot was partitioned into 9 measurement points, with 2 replicate spectral measurements occurring at each point using a GER 1500 field spectrometer (350–1050 nm; 512 bands, spectral resolution 3 nm, spectral sampling 1.5 nm). Regional plot scale hyperspectral data exists for both vegetation types. Vegetation type A is partitioned into 30 plots and has data from both 2018 & 2019, while vegetation type B only has data available for 2019, obtained from 30 non-georeferenced plots across a greater region. 3 measurements were taken at different areas within the plot. Remotely sensed hyperspectral data span across Qikiqtaruk-Herschel Island at two different dates (2.7.2019 & 27.7.2019) a spatial resolution of 5m per pixel.

1. Sample size rationale (optional)

Sample size for this study was predetermined by the existing data and the amount of long-term sampling plots, which will all be utilized in analysis. The sample size of the landscape-wide spectral measurements was determined by data availability, which was constrained by the time available to conduct field measurements. The rational for including both 2018 and 2019 plot scale spectral data is evaluating the variance of repeat measurements as different sensor types were used between the years.

### Variables

In this section you can describe all variables (both manipulated and measured variables) that will later be used in your confirmatory analysis plan. In your analysis plan, you will have the opportunity to describe how each variable will be used. If you have variables which you are measuring for exploratory analyses, you are not required to list them, though you are permitted to do so.

1. Measured variables (required)

**Response variables:**

Spectral mean: Mean value of spectral band at given wavelength (*continuous variable)*

Spectral reflectance (CV): The average between image variance for each spectral band (*continuous variable)*

**Explanatory variables:**

Vegetation type: If the vegetation corresponds with Herschel or Komakuk type *(categorical variable)*

Species richness: Number of species present in plot *(continuous variable)*

Species evenness: Relative abundance of species in a plot *(continuous variable)*

Soil-background: Percent of bare soil hits obtained during point-framing (*continuous variable)*

**Meta data:**

Year: Year of measurement

Plot: measurement plot

1. Indices (optional)
2. Mean of spectral signature: Mean value of spectral band at given wavelength

**(1)**

*λ is reflectance at wavelength=n*

1. Spectral variance (band coefficient of variance): CVb of spectral reflectance (Equation 1), will be used as the spectral variance metric for this study (Wang, et al., 2018). Here the average between image variance for each spectral band is calculated.

**(2)**

*ρλ is the reflectance at wavelength λ=n.* σ*(ρλ) and μ(ρλ) are the standard deviation and mean value of reflectance at wavelength λ=n across measurands in one band, respectively. (Wang, Gamon, Cavender-Bares, et al., 2018b)*

1. Spectral diversity (coefficient of variance(CV)): CV of spectral reflectance (equation 2), will be used as the spectral diversity metric for this study (Wang, et al., 2018). Here the average variation between all spectral bands is calculated for the pixels within an image.

**(3)**

*ρλ is the reflectance at wavelength λ.* σ*(ρλ) and μ(ρλ) are the standard deviation and mean value of reflectance at wavelength λ across all the pixels in one plot, respectively. (Wang, Gamon, Cavender-Bares, et al., 2018b).*

1. Instability index (ISI): Used to identify discriminative wavelengths used for the stable zone unmixing method (SZU) of optimal band selection (Beamish et al., 2017; Somers et al., 2010) .

**(4)**

*Rmean,1,i and Rmean,2,i are the mean reflectance values of each vegetation type and σmean,1,i and σmean,2,i are the standard deviations of the reflectance values.*

1. Spectral normalization (Rnorm): Used to identify the difference between field and plane spectral measurements (Beamish et al., 2018).

**(5)**

*Rj is the reflectance of remotely sensed plane data at wavelength λ and Rj is the reflectance field measured data at wavelength λ. (Rj may be subject to resolution rescaling to match remotely sensed plane data and then normalisation)*

### Analysis Plan

You may describe one or more confirmatory analysis in this preregistration. Please remember that all analyses specified below must be reported in the final article, and any additional analyses must be noted as exploratory or hypothesis generating.

A confirmatory analysis plan must state up front which variables are predictors (independent) and which are the outcomes (dependent), otherwise it is an exploratory analysis. You are allowed to describe any exploratory work here, but a clear confirmatory analysis is required.

1. Statistical models (required)

I will use R (v 3.5.2) to conduct general linear modeling and ordinations.

**1: Can Arctic Vegetation types be identified based on the mean and variance of hyperspectral signatures?**

Multiple linear regressions will be run to test if hyperspectral signatures of different vegetation types differ significantly from each other.

*Spectral mean ~ vegetation type + plot + year*

*Spectral variance (CVb) ~ vegetation type + plot + year*

Both spectral mean and spectral variance will be used as response variables of hyperspectral signatures.

**2: How does spectral diversity relate to species richness, evenness, canopy cover, and soil-background?**

The following model will be run to quantify relative contribution species richness, evenness, and soil-background on spectral diversity.

*Spectral diversity (CV) ~ vegetation type + richness + evenness + Soil-background*

Spectral diversity will be ordinated using PCA, to identify which components best explain the variation in spectral diversity.

One ordination plot will then be produced, and points will be discriminated by each variable.

**3: What parts of the full spectrum best discriminate vegetation types?**

The band selection methods of using a defined subset of spectral bands and spectral band unmixing will be used to reduce dimensionality of spectral data. Spectral variance will be re-calculated, based on these new subsets and spectral data, the same model as question 1 will be used to assess how band selection affects correspondence between spectral variance and vegetation type.

*Spectral variance (CVb) ~ vegetation type + plot + year*

1. Transformations (optional)

During ordination, variables will be standardized to zero mean and unit variance.

1. Inference criteria (optional)

Variables in linear models will be considered to be significant, if the upper and lower bounds of the 95 percent credible interval do not cause the estimated effect size to cross zero. The results of all statistical tests will be reported, even if these contradict hypothesized direction or are non-significant.

1. Data exclusion (optional)

All relevant available field data will we used in data analysis, with additional meta-data such as year being used when appropriate. Extreme outliers in measurement data, that are resultant of spectrometer malfunctioning will be excluded. Field reflectance spectra will be subset from 287-1094 nm to 400–985 nm due to the radiometer sensor noise at the edges of the detection range.

1. Missing data (optional)
   1. How will you deal with incomplete or missing data?
   2. **Example**: If a subject does not complete any of the three indices of tastiness, that subject will not be included in the analysis.
   3. **More information**: Any relevant explanation is acceptable. As a final reminder, remember that the final analysis must follow the specified plan, and deviations must be either strongly justified or included as a separate, exploratory analysis.
2. Exploratory analysis (optional)

**4: Are closer measurements more similar than more distant measurements?**

This will be assessed using spatial autocorrelation analysis conducted in R and a variogram.

**5: Do plot and airborne spectra differ? Do plot or airborne spectra show more variation among vegetation types?**

This will be assessed by visual comparison of the spectrally normalized differences between plot and airborne spectra, using the previously described equation for spectral normalization (equation 5)

**6: Can airborne remotely-sensed spectral diversity be used to identify vegetation types at the island scale?**

Categorical spectral classifications will be created and applied to remotely-sensed spectral data to create a predictive map of the spatial distributions of vegetation types across Qikiqtaruk- Herschel Island. This created map will be visually compaired to an existent mapping of vegetation types across Qikiqtaruk ((Obu et al., 2016).

Other

1. Other (Optional)

References:

“ArcticDEM”. (2018), , available at:

https://www.arcgis.com/apps/webappviewer/index.html?id=aff5fa8f5d5548c6bff44cc8be385f61 (accessed 22 January 2020).

Bennie, J., Huntley, B., Wiltshire, A., Hill, M.O. and Baxter, R. (2008), “Slope, aspect

and climate: Spatially explicit and implicit models of topographic microclimate in chalk grassland”, *Ecological Modelling*, Vol. 216 No. 1, pp. 47–59.

Cavender‐Bares, J., Gamon, J.A., Hobbie, S.E., Madritch, M.D., Meireles, J.E.,

Schweiger, A.K.

and Townsend, P.A. (2017), “Harnessing plant spectra to integrate the biodiversity sciences across biological and spatial scales”, *American Journal of Botany*, Vol. 104 No. 7, pp. 966–969.

Gholizadeh, H., Gamon, J.A., Zygielbaum, A.I., Wang, R., Schweiger, A.K. and

Cavender-

Bares, J. (2018), “Remote sensing of biodiversity: Soil correction and data dimension reduction methods improve assessment of α-diversity (species richness) in prairie ecosystems”, *Remote Sensing of Environment*, Vol. 206, pp. 240–253.

Moeslund, J.E., Arge, L., Bøcher, P.K., Dalgaard, T., Ejrnæs, R., Odgaard, M.V. and

Svenning,

J.-C. (2013), “Topographically controlled soil moisture drives plant diversity patterns within grasslands”, *Biodiversity and Conservation*, Vol. 22 No. 10, pp. 2151–2166.

Obu, J., Lantuit, H., Myers-Smith, I., Heim, B., Wolter, J. and Fritz, M. (2017), “Effect of

Terrain Characteristics on Soil Organic Carbon and Total Nitrogen Stocks in Soils of Herschel Island, Western Canadian Arctic: Geomorphic Disturbance, SOC and TN”, *Permafrost and Periglacial Processes*, Vol. 28 No. 1, pp. 92–107.

Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Hobbie, S.E., Madritch, M.D.,

Wang, R.,

Tilman, D., et al. (2018), “Plant spectral diversity integrates functional and phylogenetic components of biodiversity and predicts ecosystem function”, *Nature Ecology & Evolution*, Vol. 2 No. 6, pp. 976–982.

Schweiger, A.K., Schütz, M., Risch, A.C., Kneubühler, M., Haller, R. and Schaepman,

M.E.

(2017), “How to predict plant functional types using imaging spectroscopy: linking vegetation community traits, plant functional types and spectral response”, edited by Chisholm, R.*Methods in Ecology and Evolution*, Vol. 8 No. 1, pp. 86–95.

Wang, R. and Gamon, J.A. (2019), “Remote sensing of terrestrial plant biodiversity”,

*Remote Sensing of Environment*, Vol. 231, p. 111218.

Wang, R., Gamon, J.A., Cavender-Bares, J., Townsend, P.A. and Zygielbaum, A.I.

(2018a),

“The spatial sensitivity of the spectral diversity-biodiversity relationship: an experimental test in a prairie grassland”, *Ecological Applications*, Vol. 28 No. 2, pp. 541–556.

Wang, R., Gamon, J.A., Schweiger, A.K., Cavender-Bares, J., Townsend, P.A.,

Zygielbaum, A.I.

and Kothari, S. (2018), “Influence of species richness, evenness, and composition on optical diversity: A simulation study”, *Remote Sensing of Environment*, Vol. 211, pp. 218–228.

* 1. If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other context that will be helpful for future readers would be appropriate here.

**Instructions**:

* Under the “File” menu, select “Make a copy…” or “Download As” to make your own preregistration document.
* Preregister your study by either 1) attaching the document to an OSF project and registering with the “OSF Standard pre-data collection” form or 2) use the “OSF Prereg” form available here: (<https://osf.io/prereg>) (option 2 provides a better formated, final preregistration)
* Information on registering OSF projects and the different forms is available [on the OSF help docs](https://help.osf.io/hc/en-us/categories/360001550953-Registrations).
* General information about preregistration is available at <https://cos.io/prereg> and you can reach out to [prereg@cos.io](mailto:prereg@cos.io) or [@OSFprereg](https://twitter.com/osfprereg?lang=en). A preprint of this template is available at <https://osf.io/preprints/metaarxiv/epgjd/>

Panic notes

NOTES TO DELETE:

Methods

a GER 1500 field spectrometer, which covers a spectral range of 350–1050 nm, partitioned into 512 bands at a spectral resolution 3 nm, (spectral sampling 1.5 nm).

**3.3.3 ordinaiton of addiontal spectral factors**

Vegetation types discriminated from each other when ordinated. HE plots here heavily affected by diversity metrics (and total cover which is strongly correlated), as well as spectral diversity (corresponds with models). KO alined with environmental (visual) variables and mean reflectance (corresponds with models). (add mention relevance of direction)

Spectal had relatively close correspondence with total cover and could suggest that greater cover=greater canopy complexity=greater spectral diversity. For mean reflectance visable dead plant structure seemed to have a great effect and graminoid cover aswell.

Unlike previously when only ordinated according to visual properties, when environmental variables are added, vegetation types and year have slightly better visual discrimination. overlap of ordinated groups exists for at there interfaces (edges) of both measurements vegetation type and year. Overall measurments form 2018 where seen to be the most similar, with overlap to their respective vegetation type 2019 measurements. The largest discrimination in ordinal space was observed between 2019s vegetation type groups.

~~Very important, the trend and importance of variables only minimally changed when measurements where re-ordinated with the addition of 2018 data~~

Note how specific variables affect spec\_mean and CV

Roughly 2019 measuremtns aligned and discriminated along PC1, which strongly corresponds to/with mean reflectance. 2018 additonally varied with PC2, which corresponds to/with spectral diversity. (write about implication of this).

When ordinated with all available environmental variable percent dead cover, spectral diversity and bare ground were the three variables with the best representation in the first two principle components (cos2) (implies most important). Each on contributed equally to pc1 and pc2. Visuable flowering tissue had the weakest contribution in ordinating measurments filter

A high cos2 indicates a good representation of the variable on the principal component. In this case the variable is positioned close to the circumference of the correlation circle.

**4.2.1 discriminating vegetation based on mean reflectance and spectral signatures**

Herschel plot dominated by shrub cover would be expected

Komakuk vegetation This may be a result of the relative dominance of shrubs Herschel and graminoids Komakuk vegetation types (Myers-Smith, Hik, et al., 2011). Graminoid vegetation having greater canopy complexity, would result in higher reflectance in the IR region (Cavender‐Bares et al., 2017). Higher IR reflectance in Komakuk vegetation resulted in vegetation showering greater differentiation by mean reflectance.

Phenological phase may influence how vegetation type discriminate, based on mean reflectance and spectral diversity. Measurements were taken during the maximum canopy, when vegetation tissue density and structure is most pronounced (Bratsch et al., 2016). Increased canopy and vegetation density predominantly amplify reflectance in the NIR and IR regions of the spectrum (Asner and Martin, 2009; Ollinger, 2011). Differences in reflectance would be expressed in mean reflectance values but are not necessarily captured by spectral diversity.

Mean reflectance being a stronger predictor of vegetation type may be explained by between types differences in reflectance being concentrated in only one region of the spectrum. Measurements were taken during the maximum canopy phenological phase, when vegetation tissue density and structure is most pronounced (Bratsch et al., 2016). Increased canopy and vegetation density predominantly amplify reflectance in the NIR and IR regions of the spectrum (Cavender‐Bares et al., 2017). These differences would be expressed in mean reflectance values but are not necessarily captured by spectral diversity.

Phenological phase may influence

Variation in spectral signatures across the growing season may influence the applicability of using spectral diversity.

The senescence phase is when differences in chemical properties, such as photosynthetic pigmentation are most prominent (Beamish et al., 2017). Pigmentation influences reflectance across the multiple regions spectrum resulting in vegetation types showing less differentiation based on their mean reflectance. Spectral diversity accounts for variation across spectral regions and would be the better metric for distinguishing vegetation types. Furthermore, at senescence, environmental factors such as visible reproductive tissue would I found negatively correlated with spectral diversity and have been shown to confound spectral diversity (Heumann et al., 2015), would be less pronounced. Quantifying how spectral diversity varies across the growing season, could provide insight into when vegetation types are most distinct in their spectral diversity, and could influence the success of predicting biodiversity. And the strength of the spectral diversity hypothesis

Previous findings indicate that information content-based metrics such as spectral diversity are better at discriminating vegetation types, as they account for complexity of reflectance patterns across the spectrum (cite).

Mean reflectance being a better predictor of vegetation type is possibly due to measurements being taken during the maximum canopy phenological phase, where vegetation tissue density and structure are most pronounced (Bratsch et al., 2016). As canopy structure and vegetation density predominantly affect (or amplify) reflectance in the NIR and IR regions of the spectrum (Cavender‐Bares et al., 2017), spectral differences would be concentrated at only one end of the spectrum. These differences would be expressed in mean reflectance values but are not necessarily captured by spectral diversity.

*Information content-based metrics extract information from the spectral space in a number of ways, for example, by calculating the variance of vegetation reflectance indices (e.g., NDVI) (Carlson et al., 2007; Gould, 2000****), the coefficient of variation derived from spectral reflectance (Wang et al., 2016a****), or the distance from the spectral centroid (Palmer et al., 2002).* ***Alternatively, information content-based metrics can be obtained from patterns in principal component spac****e, such as the distance from the centroid (Rocchini, 2007), which com- pacts spectral information and removes noise and band collinearity (Thompson et al., 2017).*

While vegetation type identification though spectral signatures was possible, my results highlight the complex

I predict that spectral diversity will have higher correspondence with vegetation types, as it captures the spectral complexity within each vegetation type (Wang et al., 2018). If spectral diversity has a higher correspondence with vegetation types, this means that the vegetation types discriminate based on overall spectral diversity, which is variable at all spectral regions and is influenced by compositional differences, such as biological, functional, and phylogenetic diversity.

Note that mean reflectance is not a great indicator, espically as you begin to consider more veg types, as differences as only dectectle if varation occurs at one end of the specure. Not high and low or somewhere more central.

If the mean of spectral signatures has greater correspondence with vegetation type, this indicates that the spectral complexity of a vegetation type is less important than between-type differences in spectral signatures. These differences would relate to course optical differences at specific spectral regions, resultant from a limited number of structural or chemical differences. The correspondence between mean reflectance and vegetation type is likely to be limited by spectral differences not being concentrated a one region of the spectrum. Spectral differences occurring at both low and high wavelengths, overall result in an in-between mean reflectance that is similar and does not account for internal spectral complexity.

The observed discrimi- nability of the senescence phenological phase could be explained by a combination of over- all lower variability in reflectance and/or comparable or greater differences in mean reflec- tance boosting the differentiability of these three communities from the other communities in comparison to leaf-out and maximum canopy (Table 2–4, Figure 2–3**). Previous charac- terizations of Arctic vegetation community reflectance spectra have suggested that distinct communities are spectrally more similar outside of the peak growing season but have not explicitly explored the influence or trade-off of accompanying spectral variability observed in the late season (Bratsch et al., 2016; Buchhorn et al., 2013).** To date, the late season represents a relatively unexplored phenological phase spectrally, and this paper shows that it has potential given the relatively low ISI values in the visible and red-edge spectrum, providing independence from the litter and water-sensitive NIR plateau for characteriza- tion. beamish

**My measurments were taken at peak growing season, which may have dimished spectral seperation**

MAT and ST are comprised of approximately 70% vascular plant and 30% non-vascular. MNT has the inverse composition, and MT composition is almost equal between vascular and non-vascular vegetation, and was the only vegetation community never misclassified at Ivotuk. During peak growing season, MAT was exclusively misclassified as ST, highlighting the importance of vegetation structure for tundra community discrimination during peak growing season.

**4.3.2 Does band selection improve differentiation vegetation types based on spectral signature?**

Band selections did not result a greater visual discrimination of vegetation types based on spectral diversity. I expected that band selection would improve MLRM predicted differences in mean reflectance and spectral diversity between vegetation types. Only modest visual improvements where observed in discriminating based on mean reflectance and spectral diversity was not affected by band selection.

**4.2.1 improtant spectral regions**

, due to the signicant differences in the graminoid vs shrub covers between the two vegetation types. KO tends to have greater graminoid cover, which tend to have complex leaf and stalk orientation and faceted canopy structure refract and interact with light in ways distinctly different the HE plots, explaining the distinct differences in NIR platou spectral region.

The NIR plateau 750-800 nm provided consistently low ISI bands and is a region that reflectcts “leaf cellular and canopy structure” (Beamish, 2019). (but also (Cavender‐Bares et al., 2017)

~~This region may have had low and stable ISI values, possibly due to differences in graminoid cover, between vegetation types. KO plots tended to have greater cover, which Increases canopy structure/ complexity and has been shown to have alter light refraction and interaction patterns and been shown to alter (Schweiger et al., 2017)(complex leaf and stalk orientation and faceted canopy structure),~~

Relative peak at 718-720 nm (NIR ridge) has the lowest discrimination potential (confirm by (Asner and Heidebrecht, 2002)

920-985 water sensitive NIR region

Partically xxx wavelengths for xxx pigments were relavent. This corresponds with xxx species, or higher shub cover, or high flower visability

i) entering vegetation types enter senescence different dates and ii) that spectral signatures are influenced by temporal variations in specific pigmentation concentrations at time of measurements, a trend observed in (Chavana-Bryant et al., 2017).

spectral signatures are influenced by temporal variations in specific pigmentation concentrations at time of measurements, a trend observed in (Chavana-Bryant et al., 2017). Furthermore this supports discussion xxx about senescence

Bs link to H1 peak vs sensence somehow

4.2.2 band selcetion model predicitons

Write about 3 point vs 5 point minima, and the inclusion of highly variable parts of the spectrum). (could look into 5 point window). Discussion this could significantly improve model accuracy. Especially as SZU implies less bands (ideally 3) are best.

But While other work has demonstrated that having distinct endmember classes are essential to providing accurate and discriminative bands selections, my results indicated that i) more band selection based on disntinct endmemebers (2019) did not improve the differentiation of vegetationand, ii) even when spectrally signicantly less distinct (2018 data), selected bands had close correspondence. When synethising this with the manual selection this may suggest that i) ISI is robust in its selection, even with suboptimal endmembers ii) arctic vegetation types are paticullly well discriminated at distinct spectral regions and that this trend holdes true across landscapes and quality of data, iii) or lack of applicability as it does not do much…

Discussion while theoretically have more distinct endmembers should result in a better selection of discriminative bands, it may be possible that the selection of only 2019 performed worse than the 2018+2019 selection, as measurements were so distinctly different that the spectral trends per vegetation type are not shared between years.

Selecting a subset of bands will result in better discrimination of vegetation types, as reducing spectral dimensionality excludes variable and noisy regions of the spectrum. Spectral mixture analysis using spectral zone unmixing will result in finding what zones of the spectrum are best suited to differentia vegetation types. This is achieved by isolating bands that explain maximum between community variability, while exhibiting the minimum within community variability.

4.4.1

Multi-year data may be helpful to separate the confounding effects of short term drought and anthesis on NDVI–biodiversity relationship because the seasonal meteorology can vary year to year. The exact impact of these multiple factors on the timing of the NDVI–biodiversity relationship, while beyond the scope of this study, might yield additional insights into the mechanisms driving the productivity–biodiversity relationship.

Remarkably, the NDVI-biodiversity correlation peaked at mid-season, a period of warm, dry conditions and anthesis, when NDVI reached a local minimum. These findings confirm a positive, but dynamic, productivity–diversity relationship and highlight the benefit of optical remote sensing as an objective and non-invasive tool for assessing diversity–productivity relationships.

Compared to the spatial patterns of biodiversity, less attention has been paid to the seasonal patterns of biodiversity [[**40**](https://www.mdpi.com/2072-4292/8/2/128/htm#B40-remotesensing-08-00128)] (maggurn) or the effect of phenology on the ability to assess biodiversity with remote sensing. In our study, the NDVI-richness relationship was dynamic and the best regression between NDVI and species richness occurred near peak season, although the exact reasons for this deserve further study. This dynamic relationship was most likely affected by canopy development, as well as by prevailing conditions (mid-season warm, dry conditions) and flowering phenology (timing of anthesis). While both short-term drought and mid-season anthesis clearly reduced NDVI, their effect on the NDVIbiodiversity patterns was less clear, and could have even enhanced this relationship, as illustrated by the enhanced NDVI-biodiversity correlations at mid-season ([**Figure 6**](https://www.mdpi.com/2072-4292/8/2/128/htm#fig_body_display_remotesensing-08-00128-f006), [**Figure 7**](https://www.mdpi.com/2072-4292/8/2/128/htm#fig_body_display_remotesensing-08-00128-f007) and [**Figure 8**](https://www.mdpi.com/2072-4292/8/2/128/htm#fig_body_display_remotesensing-08-00128-f008), [**Table 2**](https://www.mdpi.com/2072-4292/8/2/128/htm#table_body_display_remotesensing-08-00128-t002)), or at least not interfered with it. Multi-year data may be helpful to separate the confounding effects of short term drought and anthesis on NDVI–biodiversity relationship because the seasonal meteorology can vary year to year. The exact impact of these multiple factors on the timing of the NDVI–biodiversity relationship, while beyond the scope of this study, might yield additional insights into the mechanisms driving the productivity–biodiversity relationship.

~~Features like graminoid cover that have significant influence on canopy structure seemed to be important influence of spectral properties~~ Another aspect of spectral diversity not fully considered in this paper is the impact of canopy structure on the species richness-spectral diversity relationship. The information provided by optical passive re- mote sensing data is limited to the uppermost layer of the canopy seen by the sensor. Therefore, understory vegetation is underrepresented, which is a limiting factor in estimating species richness even for low- statured ecosystems. Active remote sensing sensors such as Light impact of vertical and horizontal vegetation structure on spectral di- versity was beyond the scope of this study, but remains a promising direction for future research. (Gholizadeh et al., 2018)

In addition to the biochemical properties and vegetation structure of plants, another factor contributing to the species richness-spectral di- versity relationship is the impact of phenology (the variation of bio- chemical and structural traits of plants over time). Previous research (e.g. Shurin, 2007; Wang et al., 2016b; White et al., 2006) has shown that the species richness is dynamic and varies with time. One source of uncertainty not considered in our analysis is the varying phenology of different species within the same sampling unit (e.g. while some species have senesced, others remain green). Although the time of collecting proximal data in our experiment was coincident with the peak growing season and before appreciable senescence occurred, applying the same approach in other phenological stages or other ecosystems with dif- ferent successional stages requires considering the temporally dynamic characteristics of the species. (Gholizadeh et al., 2018)

Species evenness adds additional information on stand composition, which affects spectral variation. However, it is not clear how or to what extent species rich- ness, evenness, and composition affect the overall optical signal, in part because experimental approaches are difficult to apply in remote sen- sing studies due to the large spatial scales involved. Furthermore, soil is known to confound optical diversity estimation (Gholizadeh et al., 2018) and these effects (species richness, evenness, composition and soil background) can be scale-dependent (Wang et al., 2018) requiring studies to be explicit about the spatial, temporal and spectral scales involved. (Wang, Gamon, Schweiger, et al., 2018)

This explains the widespread and well‐known importance of the NIR region in vegetation remote sensing, but presents an interesting paradox that has yet to be fully explored: that we can often gain more insight about the functioning of plants by examining wavelengths that are not used in photosynthesis than by examining those that are.(Ollinger, 2011)

In reviewing knowledge about the factors affecting canopy reflectance, several interesting themes have emerged. In attempts to identify specific drivers of reflectance, uncer- tainties related to scattering presently outweigh uncertain- ties related to absorbers (pigments, water, etc.). This is perhaps because absorbers influence specific spectral regions and can be measured more easily than structural properties. Factors such as leaf anatomy and leaf angle distribution affect scattering over all wavelengths in ways that are diffi- cult to quantify. This presents a challenge for understanding reflectance in the NIR region, where multiple combinations of interdependent properties can yield similar patterns of reflectance. This does not necessarily restrict our ability to

In addition to the biochemical properties and vegetation structure of plants, another factor contributing to the species richness-spectral di- versity relationship is the impact of phenology (the variation of bio- chemical and structural traits of plants over time). Previous research (e.g. Shurin, 2007; Wang et al., 2016b; White et al., 2006) has shown that the species richness is dynamic and varies with time. One source of uncertainty not considered in our analysis is the varying phenology of different species within the same sampling unit (e.g. while some species have senesced, others remain green). Although the time of collecting proximal data in our experiment was coincident with the peak growing season and before appreciable senescence occurred, applying the same approach in other phenological stages or other ecosystems with dif- ferent successional stages requires considering the temporally dynamic characteristics of the species. (Gholizadeh et al., 2018)

ADD mixed veg and PCA? and the inability to discriminate (mixed plots) based on their spectral properties,

4.2.3 discrimination in ordination

High heterogeneity in community-scale vegetation composition, soil moisture, and an abundance of litter and non-vascular components complicates interpreta- tion of NDVI and other broadband VIs in sparsely vegetated areas and outside of the peak- growing season (Liu et al., 2017). Buchhorn et al (2013) found that surface moisture re- duced reflectance in the NIR, in turn underestimating the biomass signal of low-Arctic tundra. Further, van Leeuwen and Huete, (1996) have demonstrated the importance of standing litter and soil in the interpretation of biophysical parameters. These results suggest that the small-scale heterogeneity of Arctic tundra vegetation and other internal ecosystem components are not well characterized at non-peak times, limiting our ability to gain a complete and detailed picture of vegetation change in the Arctic.

Communities are often broadly divided into moisture regime such as xeric (dry), mesic (moist), or hydric (wet), and soil acidity (acidic, non-acidic) and then further divided depending on the species complexes present (Walker et al., 2005, 1994). Often the same species can be observed in multiple distinct complexes as different phenotypes. Large homogenous patches of one or two species rarely exist outside water tracks, disturbed areas, or dry uplands where erect and prostrate dwarf shrub species dom- inate. This combination of small-scale heterogeneity in vegetation composition and soil moisture as well as the prostrate nature of tundra species leads to highly mixed, variable, and often similar spectral signatures between distinct vegetation communities. The some- times-dominant presence of non-vascular components (mosses and lichens) and barren ar- eas also contribute to the unique spectral signatures of tundra landscapes (Hope et al., 1993). This high spectral similarity can be observed with ground-based Visible-Near Infra- red (VNIR) remote sensing data making spectral separation challenging (Bratsch et al., 2016; Buchhorn et al., 2013; Riedel et al., 2005).

Soil reflectance is highly sensitive to moisture and roughness (Jacquemoud et al., 1992; Pinty et al., 1998) and its re- flectance can vary within and across study sites and through time. Therefore, the spectra extracted from one site (or image) may not be applicable to another site and another time

**Xxx Comparison between manual and automatic selection**

I found there was a general correspondence between bands included in manual and automatic selection. My manual selection was based on bands identified In other work to be associated with key biological functions in Arctic vegetation (Beamish, 2019; Bratsch et al., 2016; Buchhorn et al., 2013) (cite). Automatic band selection yielding similar subset of may suggest that spectral regions important for discriminating vegetation are similar across different sites with Arctic tundra vegetations. This implies observed trends of bands that are best suited to spectrally discriminate vegetation generalizable across sites.

These results from a prairie ecosystem near or during the peak growing season might differ from those from other ecosystems that differ substantially in species traits or from results in similar ecosystems at different times. Our results imply that, for biodiversity mapping, we can focus on spectral regions that cause heterogeneity in reflected radiation and are sensitive to factors such as leaf biochemistry, structural parameters, and phenology of plants, rather than using all bands. However, the exact influence of these factors is likely to vary with species type, environmental conditions, and time (Chavana-Bryant et al., 2017; Feilhauer et al., 2017; Filella and Penuelas, 1994; Gitelson and Merzlyak, 1997; Roberts et al., 1998). Therefore, identifying key wavelengths and plant traits for biodiversity applications is a complex, multidisciplinary effort that requires a deep understanding of remote sensing, ecology, and plant physiology. The effort is constrained by issues such as temporal resolution of the data, lack of data on species traits over time and space, and the challenge of data fusion at different scales (Jetz et al., 2016; Ustin and Gamon, 2010). Considering oppor- tunities afforded by forthcoming spaceborne hyperspectral sensors, such interdisciplinary efforts are needed to significantly improve our ability to use these sensors to obtain biodiversity information. (Gholizadeh et al., 2018)

The resulting wavelength selection from the ISI minima technique in the most discrimina- tive phenological phase of each community predominantly identifies vegetation colour, driven by vegetation pigment concentration, as discriminative spectral regions for identifi- cation of specific communities. The identification of senescence as a discriminative phe- nological phase is particularly interesting, as the communities are visually most distinct at senescence when carotenoid (yellow to orange) and anthocyanin (red) pigments begin to dominate and chlorophyll (green) pigments are degrading. With the exception of DT, all communities had equal or greater wavelength selection in the combined visible and red- edge spectrum than the NIR. **Slope and position of the red edge has been directly related to chlorophyll content in vascular plants (Filella and Penuelas, 1994)** **(maybe link to phase shift between 2018 and 2019) (difference between years & time of measurement)** and this paper is in agreement with previous and well-established research linking absorption in the visible spectrum to vegetation colour and pigment concentration (Curran, 1989; Gitelson and Merzlyak, 1997; Gitelson et al., 2002; Stylinski et al., 2002; Ustin and Curtiss, 1990) but brings new insights to the potential of the senescent phase for optical discrimination of low- Artic vegetation.

Although my results showed no difference between in the improvements to differentiation vegetation types provided by dimensionally reduced manual and automatically bands, persistent issues such as noise reduction would are addressed in manual selection (Jia et al., 2012)

Furthermore, if SZU results in the selection of bands that correspond with and/or are not visually better than an informal literature derived selection of spectral bands, this indicates that structural and chemical differences may be relatively convergent between different Arctic vegetation types. Furthermore, this is promising for the development of future satellites, as potentially selection of spectral regions may be sufficient to discriminate vegetation types.