[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, canopy cover, 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: Can remotely-sensed spectral diversity be used to identify vegetation types?

1. **Hypotheses**

**Q1 Can Arctic Vegetation types be identified based on the mean and variance of hyperspectral signatures?**

Arctic vegetation types can be identified based on the mean of their spectral signature (H1a) and spectral variance of their spectral signature (H1b). The Spectral variance of spectral signatures will be better than the mean of spectral signatures at discriminating vegetation types (H1c). When ordinated, spectral signatures will discriminate among vegetation types (H1d) and year of measurement (corresponds with change in sensor type) (H1e).

Predictions

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. Spectral variance will be better at discriminating vegetation types, at it captures more internal complexity of the difference between spectral signatures of a vegetation type (Wang et al., 2018). When ordinated, spectral signatures will discriminate by year, due to the potential difference of reflectance measurements made by the different sensor types used in between 2018 & 2019, as well variation in flowering phenologyand community compositions between the two years.

**Q2: How does spectral diversity relate to species richness, evenness, canopy cover, and soil-background?**

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 richness having the stronger relationship(H2d). There will be a positive relationship between canopy cover and spectral diversity (H2e). Soil background cover will have the strongest positive influence on spectral diversity (H2f).

Predictions

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). Canopy cover and structure interact and scatter light, altering the reflective properties that while sometimes contradictory, overall tend to increase optical diversity (Asner and Martin, 2009). Soil background is a significant predictor of spectral diversity, as soil has a distinctly different reflectance from vegetation (Gholizadeh et al., 2018)

3**: What parts of the full spectrum best discriminate vegetation types? What plant traits (structural versus chemical) are most related to those spectral bands?**

4: **Are closer measurements more similar than more distant measurements? What is the distance at which measurements are no longer spatially autocorrelated?**

5: **Do plot and airborne spectra differ?** **Does 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?**

Vegetation type will significantly affect remotely sensed spectral diversity (H3a). Remotely sensed spectral diversity will correspond with the topographic variables of elevation (H3b), slope (H3c), aspect (H3d) and soil wetness (H3e). Soil background cover will have the strongest positive influence on spectral diversity (H3f). When ordinated, spectral diversity will discriminate by vegetation type (H3d). The spatial distribution of vegetation types based on remotely sensed spectral data correspond with existent vegetation maps (H3e).

Understanding

Slope and aspect strongly affect the interception of solar radiation, a key determinate of vegetation patterns (Bennie et al., 2008). Furthermore, topology influences microclimatic factors, such as surface energy balance, surface temperatures, potential evapotranspiration, and soil moister content, which all influence vegetation type distribution (Bennie et al., 2008; Moeslund et al., 2013). Following an ordination of spectral diversity, discriminated points will be used to categorize vegetation type. 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 should correspond to an existent mapping of vegetation types (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) canopy cover and 4) soil background 5) 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, canopy cover, and soil-background influence spectral diversity. Using hyperspectral data obtained at a plane scale, I aim to identify vegetation types using remote sensing data. An existing digital elevation model (ArcticDEM, 2018) will, be used to interpolate other relevant variables for determining microclimatic conditions, such as slope, aspect, and wetness (TWI) (Bennie et al., 2008), as these may be predictors of vegetation type (Moeslund et al., 2013). 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 (“ArcticDEM”, 2018; “JPL | AVIRIS”, 2019; Obu et al., 2016). I have not been granted access to any of the data thus far and therefor 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 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 of previously collected data. From this existing data, point-framing data, canopy cover and percent bare ground 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. Multi-spectral imagery (400-1000 nm), as well as percent ground cover 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 squares, with 2 replicate spectral measurements occurring in each using (device name) . 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 greater region. 3 measurement 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)*

Canopy cover: Percent of canopy hits obtained during point-framing *(continuous variable)*

Exposed bare ground: Percent of bare ground hits obtained during point-framing (*continuous variable)*

Elevation: Elevation of plot *(categorical variable)*

Slope: Percent change in elevation *(continuous variable)*

Aspect: Compass direction that slope faces (*continuous variable)*

Wetness (topographic wetness index) (TWI): Steady state soil moister index (see index 3) *(continuous variable)*

Confidence: The self-defined confidence that a field spectral measurment corresponds to its assigned categorical type *(categorical variable: confidence levels = “definitely plant, probably plant, probably panel, definitely panel”)*

**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. Topographic wetness index (TWI): TWI quantifies the trends of soil water distribution, with respect to local topology.

**(4)**

TWI = ln (α/tan β)

*α is the specific upslope area draining through the local point per unit contour length, and β is the local slope.*

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) .

*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).

**(6)**

*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 (ggfortify).

**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*

*Spectral mean ~ vegetation type + plot + year*

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

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

Both spectral mean and spectral variance will be used as response variables of hyperspectral signatures. Not all spectral signature measurements have plot data available, so it will only be included when available.

**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, canopy cover and soil-background on spectral diversity.

Spectral diversity (CV) ~ vegetation type + richness + evenness + canopy cover + 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 for each factor to show the variability of each explained by each component.

**3: Can remotely sensed spectral diversity be used to identify vegetation types?**

A general linear model would be used to quantify, if remotely sensed spectral diversity differs significantly between vegetation types.

Spectral diversity (CV) ~ vegetation type + elevation + slope + aspect + wetness + soil background

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.

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)

A spectral comparison of normalized plot and plane data will be conducted to check the variance between plot level and remotely sensed data.

Band selection using spectral zone unmixing (SZU) will be conducted to identify which wavelength bands are the most discriminative for differentiating vegetation types/soil cover. This will be done using a inStability index (ISI) (Beamish et al., 2017; Somers et al., 2010).

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.

Data may be checked for spatial autocorrelation using the INLA package for R.

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/>