**Assignment:**

**How to run Multivariate Ecological Analysis for Unconstrained Non-metric Multidimensional Scaling (NMS) Ordination, perMANOVA, and Indicator Species Analysis using R (4.0.2) and RStudio Desktop to Explore Ecological Recovery of Plant Communities on Oil and Gas Well Pads in Alberta’s Boreal Forests After Reclamation**

**Learning outcomes:**

By the end of this activity, you should be able to:

* Write a high quality ecological hypothesis to be tested by multivariate analysis.
* Explain what types of data are appropriate for multivariate analysis
* Briefly explain what an ordination shows the reader.
* Run a Non-metric multidimensional scaling (NMS) ordination in R that visualizes plant community composition patterns between well pads and adjacent reference forests.
* Apply permutational multivariate analysis of variance (perMANOVA) to distinguish between plant community groups on the well pad and in the adjacent forest reference sites.
* Use indicator species analysis to identify which species are contributing to separation among *a priori* groups (well pad vs forest).
* Describe and interpret the output of their multivariate statistical analyses, including their plant community ordination plot, perMANOVA, and indicator species analyses

**Background Information and Metadata:**

Ecological succession has been a long-standing area of focus in ecology (Connell and Slatyer 1977, Pickett et al. 1987). Boreal forest ecosystems have a somewhat predictable successional pathway, albeit with variation among individual stands. The general pathway of Canada’s boreal forest after stand-replacing disturbance (fire is the dominant natural stand-replacing disturbance agent and forest harvest is the dominant anthropogenic stand-replacing disturbance agent) is initial development of a shade-intolerant broadleaf forest (e.g., aspen and balsam poplar) that then transitions to conifers as the shorter-lived broadleaf forest species die off (see Bergeron et al. 2014). An alternative pathway occurs in shade-intolerant fire adapted lodgepole pine forests that have serotinous cones, who self replace after fire. So while there is variation in the relative abundance of the species, the species composition is deterministic after succession. So generally we would expect the forest to return (eventually) to its original community structure after disturbance.

However, less studied are the ways in which anthropogenic disturbance agents, such as natural resource extraction, may alter the successional trajectories of ecosystems. For example, crude oil and natural gas well pads established for development of energy resources are estimated to have a footprint that covers over 500 000 ha of Alberta, Canada’s landscape. Reclamation of these sites is a critical step in alleviating impacts of this industrial human footprint, with reclamation criteria evolving over time (Powter et al. 2012). Approximately 130,000 well pads have been certified as reclaimed since 1963 in Alberta. However, given the significant additional disturbance to the soils that may result during oil and gas development and the subsequent reclamation practices, there is less certainty that these reclaimed sites will return to their original community structure – which is what you are exploring this in this dataset. Instead, there is a concern that resource development will have long-lasting legacy effects on soil, biological, and spatial characteristics of ecosystems and that their succession may be slowed or arrested.

Site preparation for oil and gas extraction is a high intensity disturbance agent that removes the vegetation and surface soil on the well pad. After the well pad is decommissioned, subsequent reclamation, which includes requirements to meet criteria to receive certification (e.g., in Alberta - ESRD 2013), then attempts to restore vegetation and soil properties on the well pad (e.g., recontouring the ground that had been leveled so the heavy duty equipment could be on it, sowing of non-native (historically) or native (newer criteria) seed to conserve top soil). The surface stripping of the forest floor often removes the native seedbank, thus resetting the successional pathway for the well pad. Therefore the potential to shift (or even arrest) the well pad’s successional trajectory is high. In this research activity you will use the accompanying datasets to investigate patterns in the understory plant communities (the percent cover of shrubs, herbs, grasses, clubmoss, fern, lichens, including non-native species) in clustered study units that include both certified reclaimed well pads and adjacent undisturbed forested reference sites. You will explore whether there is recovery of the plant communities in the reclaimed well pads compared with the adjacent undisturbed reference forest sites. This dataset investigates patterns in the understory plant communities (the percent cover of shrubs, herbs, grasses, clubmoss, fern, lichens, including non-native species) as a function of measuring vegetation recovery in clustered study units that include both certified reclaimed well pads and adjacent undisturbed forested reference sites. The study units were located in both the Central Mixedwood (conifer & deciduous; n=15) and Lower Foothills (n=15) Natural Subregions of Alberta (The Natural Regions Committee, 2006).These study units contained reclaimed well pads, ranging from 7-48 years post-certification. Other ecological variables measured included: age post-reclamation, the LFH depth (organic soil layer), soil bulk density, soil pH, organic soil carbon, soil nitrogen, the soil C:N ratio, number of live and dead trees, and live and dead basal area. Detailed sampling methods are described in McIntosh et al. (2019). [For more detailed background information about the study and data files, review WellvsRefMetadata.docx]

**Primarydataset.csv** file contains a unique identifier column (ID) that indicates the sampling unit that was measured – it is a combination of the Site\_ID (which includes BOR for Boreal or FOOT for Foothills – which represent the two natural regions from which they were sampled) and the WellorRef fields from Secondataset.csv so there is one row of plant community data for each of the 60 sampling units (30 sites \* 2 locations within each site – wellsite (well) or adjacent reference (ref)). For example, the first line ID is Bor1\_Ref is the ID for the Reference well pad that was sampled in the Boreal Region Study Site #1. The columns after ID are the species codes for each of the 106 understory plant species that were sampled within the 0.25 m2 quadrats. For example Abiebal is *Abies balsamea*. The values reported here are the mean percent cover values for all of the quadrats that were sampled within a wellsite or reference forest within a given study unit that was sampled. See Appendix A for species code.

**Seconddataset.csv Entity-level Metadata** file contains both categorical and quantitative information for each sampling unit that was measured. The file starts with the same unique identifier column (ID) as Primarydataset.csv so that you can match up the data from both datasets. The data are organized in two separate .csv files because this is how the R statistical environment will need to import and use them. The columns after ID are the various data there were collected for each sampling unit in addition to the plant cover species that are captured in the primary dataset.csv. See Table 1 below for a description of all of these variables.

Table 1. Seconddataset.csv Column Descriptors

|  |  |  |  |
| --- | --- | --- | --- |
| Column identifier | Type | Units | Description |
| ID | Unique identifier (Primary Key) | n/a | Unique identifier that indicates the sampling unit that was measured – it is a combination of the Site\_ID and the WellorRef fields so there is one row and one unique identifier for each of the 60 sampling units (30 sites \* 2 locations within each site) |
| Site\_ID | Categorical | n/a | Identification of wellsite (There are 30 total, with 15 in the Boreal Natural region (Bor1-15) and 15 in the Foothills Natural region (Foot1-15). |
| WellorRef | Categorical | n/a | Whether it is the wellsite (Well) or adjacent reference (Ref) site that was sampled in that row |
| Bor1Foot2 | Categorical | n/a | Natural subregion within the Boreal or Foothills Natural Regions: 1 = Central Mixedwood Boreal Natural Subregion; 2 = Lower Foothills Natural Subregion |
| Age\_postcert | Quantitative | yrs | The number of years that has passed between when the reclamation certificate was issued after the wellsite was decommissioned and when the wellsite was sampled. |
| LFHmean\_mm | Quantitative | mm | LFH soil horizon depth (organic layer) |
| BD\_0-15cmdepth\_cm | Quantitative | g/cm3 | Bulk density of the soil - 0-15 cm depth |
| pH\_0 | Quantitative | n/a | measured pH of the soil - 0-15 cm depth |
| TOC\_0 | Quantitative | % | Total organic carbon in the soil - 0-15 cm depth |
| TN\_0 | Quantitative | % | Total nitrogen in the soil - 0-15 cm depth |
| CNratio\_0 | Quantitative | n/a | Carbon to nitrogen ratio of the soil - 0-15 cm depth |
| tph\_total | Quantitative | #/ha | Number of live and dead trees/ha |
| LiveBA\_m2/ha | Quantitative | m2/ha | Live basal area (BA; m2/ha) for all trees combined. |
| DeadBA\_m2/ha | Quantitative | m2/ha | Dead basal area (BA; m2/ha) for all trees combined. |
| herb\_cover | Quantitative | % | Percent herb cover |
| shrub\_cover | Quantitative | % | Percent shrub cover |
| graminoid\_cover | Quantitative | % | Percent graminoid cover |
| lichen\_cover | Quantitative | % | Percent lichen cover |
| clubmoss\_cover | Quantitative | % | Percent clubmoss cover |
| fern\_cover | Quantitative | % | Percent fern cover |
| non\_native\_cover | Quantitative | % | Percent non-native vegetation cover |

**Assignment:**

You will perform a non-metric multidimensional scaling (NMS) ordination, permutational multivariate analysis of variance (perMANOVA), Indicator Species Analysis, and Summary Statistics to explore the effect of the grouping factor in this dataset (i.e., well pad or reference forest – the variable “WellorRef” in your seconddataset.csv) on the plant communities of these research sites. Use the WellvsRef\_RPrimer.docx to guide you through providing answers for the following numbered inquiries.

*Research Hypothesis*

Write out an ecological hypothesis for how you think the plant communities will recover over time after well pad reclamation. This will inform the multivariate analyses that you complete below.

*NMS Ordination (Section 4 of the Primer)*

1. Export and paste your scree plot here (see section 4.1):
2. Export and paste your ordmain output (see section 4.2). This information will be useful for writing your results section (See #10 below).
3. Run a i) 3-dimensional and ii) a 4-dimensional NMS ordination. Export and paste your output including your stress and number of tries on the 3-dimensional and 4-dimensional NMS ordinations.
4. How much of an improvement in stress was there from the 2-D to the 4-D solution?
5. Export and paste an NMS plot of your choice (section 4.3): While going through the primer, you will notice that you were given various methods on how to produce your ordination plot (i.e., different colors, symbols etc.) For your NMS plot, include only a polygon border and legend, and use a different symbol and color code for your datapoints (cannot be the same ones as in the primer!). No need to include vectors on this ordination plot (but you can if you like – if you do, provide two plots one with the vectors and one without). Be sure to include a detailed figure caption describing your plot.
6. Vectors. i) Copy and paste the vector output from your analyses (section 4.4). Hint: Using the snip tool may make the output more user friendly than just copying and pasting from RStudio). ii) Based on the vector output - is age post-certification an important attribute in contributing to the separation of the groups? Why or why not? What about non-native cover?
7. Describe clearly and completely the purpose of NMS. Be sure to include in your description information about the nature of the data analyzed.

*perMANOVA (Section 5 of the Primer)*

1. Run the perMANOVA and copy and paste your perMANOVA output (reminder: using the snipping tool sometimes gives a better output).
2. Briefly explain what the perMANOVA output is telling you in the context of this study dataset and the results of your perMANOVA (include your P-value here (alpha=0.05).

*Indicator species analysis (Section 6 of the Primer)*

10. i) Run Indicator Species Analysis and export and paste your ISA output, first for just the significant species and then for all species. ii) Is green alder (*Alnus crispa* – see Appendix for species code) a good indicator for the reference forests? What are you basing your decision on? iii) is *Cornus canadensis* a good indicator for either group – if so – which one – what are you basing your decision on? iv) Create a table for the species that are significant indicators (alpha=0.001) and be sure to include the P, A, and B values. v) How many indicator species are there for the reference forests and the well pads? vi) Briefly describe the purpose of Indicator Species Analysis - what does the indicator analysis species tell you in the context of this study?

*Summarizing your results*

11. Summary Statistics (Section 7): i) Run the summary/descriptive statistics that reports the mean and standard deviation (SD) for all of the plant species, grouped by well pad vs reference and copy and paste your results.

1. write out a results section as if you were going to submit these results as part of a scientific report/manuscript. Describe and interpret the outcome of the multivariate statistical analysis incorporating the information from each of the sections above. Be sure to include your customized plant community ordination plot (vector portion not required), your perMANOVA results and indicator species analysis results (HINT: you only need the relevant information in your tables- what is significant? Use alpha = 0.001 and don’t forget to include your A and B values). In addition – include the mean and SD for each of your indicator species in a Table. Refer to the “what to report” (Section 8) section of your primer on how to report results.

**References:**

Bergeron Y, Chen HYH, Kenkel NC, Leduc AL, Macdonald SE. 2014. Boreal mixedwood stand dynamics: Ecological processes underlying multiple pathways. The Forestry Chronicle, 90(2):202-213.

Connell JH, Slatyer RO. 1977. Mechanisms of succession in natural communities and their role in community stability and organization. The American Naturalist, 111: 1119–1144.

Environment and Sustainable Resource Development (ESRD). 2013. 2010 Reclamation criteria for wellsites and associated facilities for forested lands (Updated July 2013). Edmonton, Alberta. https://open.alberta.ca/publications/9780778589846.

McIntosh ACS, Drozdowski B, Degenhardt D, Powter CB, Small CC, Begg J, Farr D, Janz A, Lupardus RC, Ryerson D, Schieck J. 2019. Monitoring ecological recovery of reclaimed wellsites: protocols for quantifying recovery on forested lands. MethodsX. https://doi.org/10.1016/j.mex.2019.03.031

Natural Regions Committee. 2006. Natural regions and subregions of Alberta. Publication Number T/852. Government of Alberta, Canada.

Pickett STA, Collins SL, Armesto JJ. 1987. Models, mechanisms and pathways of succession. Botanical Review, 53: 335–371.

Powter C, Chymko N, Dinwoodie G, Howat D, Janz A, Puhlmann R, Richens T, Watson D, Sinton H, Ball K, Etmanski A, Patterson B, Brocke L, Dyer R. 2012. Regulatory history of Alberta's industrial land conservation and reclamation program. Canadian Journal of Soil Science, 92(1):39-51.

**Appendix A:**

Table A1. Species list scientific names and common names. Visit plants.usda.gov to get more information on individual plant species.

| Species Code | Genus | Species | Common |
| --- | --- | --- | --- |
| Abiebal | *Abies* | *balsamea* | Balsam Fir |
| Achimill | *Achillea* | *millefolium* | Common Yarrow |
| Agrosca | *Agropyron* | *scabra* | Tickle Grass |
| Agrotra | *Agropyron* | *trachycaulum* | Slender Wheatgrass |
| Alnucri | *Alnus* | *crispa* | Green Alder |
| Alnurug | *Alnus* | *rugosa* |  |
| Amelaln | *Amelanchier* | *alnifolia* | Saskatoon |
| Apocand | *Apocynum* | *androsaemifolium* | Spreading Dogbane |
| Aquican | *Aquilegia* | *canadensis* | Canada Columbine |
| Aralnud | *Aralia* | *nudicaulis* | Wild Sarsparilla |
| Arnicor | *Arnica* | *cordifolia* | Heart-Leafed Arnica |
| Astecil | *Aster* | *ciliolatus* | Lindley's Aster |
| Astecon | *Aster* | *conspicuus* | Showy Aster |
| Betupap | *Betula* | *papyrifera* | Paper Birch |
| Botrvir | *Botrychium* | *virginianum* | Virginia Grape Fern |
| Bracsal | *Brachythecium* | *salebrosum* | Golden Ragged Moss |
| Bromcil | *Bromus* | *ciliatus* | Fringed Brome |
| Bromine | *Bromus* | *inermis* | Smooth Brome |
| Brompum | *Bromus* | *pumpellianus* | Pumpelly brome |
| Calacan | *Calamagrostis* | *canadensis* | Bluejoint |
| Carespp | *Carex* | *Spp* | Upland Carex Spp |
| Castmin | *Castilleja* | *miniata* | Red Indian Paintbrush |
| Chamang | *Chamerion* | *angustifolium* | Fireweed |
| Circalp | *Circaea* | *alpine* | Small Enchanter's-Nightshade |
| Cirsarv | *Cirsium* | *arvense* | Canada Thistle |
| Cladspp | *Cladonia* | *spp* |  |
| Corncan | *Cornus* | *canadensis* | Bunchberry |
| Cornsto | *Cornus* | *stolonifera* | Dogwood |
| Desccae | *Deschampsia* | *caespitosa* | Tufted Hairgrass |
| Dicrsco | *Dicranum* | *scoparium* | Broom Moss |
| Dryoaus | *Dryopteris* | *austriaca* | Spinulose Shield Fern |
| Elymspp | *Elymus* |  |  |
| Equiarv | *Equisetum* | *arvense* | Common Horsetail |
| Equipra | *Equisetum* | *pratense* | Meadow Horsetail |
| Equisyl | *Equisetum* | *sylvaticum* | Woodland Horsetail |
| Eurhpul | *Eurhynchium* | *pulchellum* | Common Beaked Moss |
| Evermes | *Evernia* | *mesomorpha* | Spruce Moss |
| Fragves | *Fragaria* | *vesca* | Woodland Strawberry |
| Fragvir | *Fragaria* | *virginiana* | Wild Strawberry |
| Galetet | *Galeopsis* | *tetrahit* | Hemp-Nettle |
| Galibor | *Galium* | *boreale* | Northern Bedstraw |
| Galitri | *Galium* | *triflorum* | Sweet-scented Bedstraw |
| Geumale | *Geum* | *aleppicum* | Yellow Avens |
| Gymndry | *Gymnocarpium* | *dryopteris* | Common Oak Fern |
| Habehyp | *Habenaria* | *hyperborea* | Northern Green Orchid |
| Haledef | *Halenia* | *deflexa* | Spurred Gentian |
| Heralan | *Heracleum* | *lanatum* | Cow-Parsnip |
| Hierumb | *Hieracium* | *umbellatum* | Narrow-leaved Hawkweed |
| Hylospl | *Hylocomium* | *splendens* | Stairstep Moss |
| Impacap | *Impatiens* | *capensis* | Spotted Touch-Me-Not |
| Ledugro | *Ledum* | *groenlandicum* | Labrador Tea |
| Leyminn | *Leymus* | *innovatus* | Hairy Wild Rye Grass |
| Linnbor | *Linnaeus* | *borealis* | Twinflower |
| Loniinv | *Lonicera* | *involucrata* | Bracted Honeysuckle |
| Lycoann | *Lycopodium* | *annotinum* | Stiff Club-moss |
| Maiacan | *Maianthemum* | *canadense* | Wild Lily-of-the-valley |
| Melioff | *Melilotus* | *officinalis* | Yellow Sweet Clover |
| Mertpan | *Mertensia* | *paniculata* | Tall Lungwort |
| Mitenud | *Mitella* | *nuda* | Bishop's Cap |
| Orthsec | *Orthilia* | *secunda* | One-sided Wintergreen |
| Oryzasp | *Oryzopsis* | *asperifolia* | Rough-Leaved Ricegrass |
| Peltcan | *Peltigera* | *canina* | Dog Lichen |
| Petapal | *Petasites* | *palmatus* | Palmate-Leaved Coltsfoot |
| Petasag | *Petasites* | *sagitatus* | Arrow Leaved Coltsfoot |
| Phalaru | *Phalaris* | *arundinacea* | Reed Canary Grass |
| Phlepra | *Phleum* | *pratense* | Timothy |
| Picegla | *Picea* | *glauca* | White Spruce |
| Picemar | *Picea* | *mariana* | Black Spruce |
| Plagcus | *Plagiomnium* | *cuspidatum* | Woodsy Leafy Moss |
| Platyrep | *Platygyrium* | *repens* | Common flat-brocade moss |
| Pleusch | *Pleurozium* | *schreberei* | Big Red Stem moss |
| Poapal | *Poa* | *palustris* | Fowl Bluegrass |
| Pohlnut | *Pohlia* | *nutans* | Copper Wire Moss |
| Polyjun | *Polytrichum* | *juniperinum* | Juniper Hair-Cap |
| Popubal | *Populus* | *balsamifera* | Balsam Poplar |
| Poputre | *Populus* | *tremuloides* | Trembling Aspen |
| Ptilcri | *Ptilium* | *crista-castrensis* | Knight's Plume moss |
| Pyroasa | *Pyrola* | *asarifolia* | Common Pink wintergreen |
| Rhinbor | *Rhinanthus* | *borealis* | Yellow Rattle |
| Ribelac | *Ribes* | *lacustre* | Black Gooseberry |
| Ribeoxy | *Ribes* | *oxycanthoides* | Canadian gooseberry |
| Ribetri | *Ribes* | *triste* | Wild Red Currant |
| Rosaaci | *Rosa* | *acicularis* | Prickly Rose |
| Rubucha | *Rubus* | *chamaemorus* | Cloudberry |
| Rubuida | *Rubus* | *idaeus* | Raspberry |
| Rubupub | *Rubus* | *pubescens* | Dewberry |
| salix | *Salix* | *Spp* | Willow spp. |
| Scirmic | *Scirpus* | *microcarpus* | Small-Fruited Bulrush |
| Shepcan | *Shepherdia* | *canadensis* | Canada buffaloberry |
| Smileste | *Smilacina* | *stellata* | false Solomon's seal |
| Solican | *Solidago* | *canadensis* | Canada Goldenrod |
| Soncarv | *Sonchus* | *arvensis* | Perennial Sow-thistle |
| Soncasp | *Sonchus* | *asper* | Spiny Annual Sow-thistle |
| Stelspp | *Stellaria* | *Spp* | Chickweed Spp |
| Sympalb | *Symphoricarpos* | *albus* | Common Snowberry |
| Sympcil | *Symphyotrichum* | *ciliolatum* | Lindley’s Aster |
| Sympocc | *Symphoricarpos* | *occidentalis* | Buckbrush |
| Taraoff | *Taraxacum* | *officinale* | Common Dandelion |
| Thuiabi | *Thuidium* | *abietinum* | Wiry Fern Moss |
| Triebor | *Trientalis* | *borealis* | Northern Starflower |
| Trifhyb | *Trifolium* | *hybridum* | Alsike Clover |
| Trifpra | *Trifolium* | *pratense* | Red Clover |
| Urtidio | *Urtica* | *dioica* | Stinging Nettle |
| Vacccae | *Vaccinium* | *caespitosum* | Dwarf Bilberry |
| Vaccmyr | *Vaccinium* | *myrtilloides* | Common Blueberry |
| Vaccvit | *Vaccinium* | *vitis-idaea* | Lingonberry |
| Vibuedu | *Viburnum* | *edule* | Mooseberry |
| Viciame | *Vicia* | *americana* | American Vetch |
| Violcan | *Viola* | *canadensis* | Canadian white violet |
| Violren | *Viola* | *renifolia* | Kidneyleaf Violet |