

Machine Learning to Model Wetland Vegetation Trajectories Using Multispectral Sentinel-2 Imagery

Carl Ayer, B.Sc., EPt.

Meghan Hellman, B.Sc. P.Ag.

May 4, 2022

Agenda

- Program Rationale
- Get “techy”
- Preliminary Model Results
- Next Steps



Introductions

Connacher Oil and Gas Limited:

- Dawn Emes

Matrix Project Team:

- Carl Ayer, B.Sc., EPt | environmental data analyst
- Meghan Hellman, B.Sc., P.Ag | wetland lead and project manager
- Kelly Ostermann, M.Sc., P.Ag. | wetland technical advisor
- Niandry Moreno, Ph.D. | geospatial technical advisor

Program Rationale

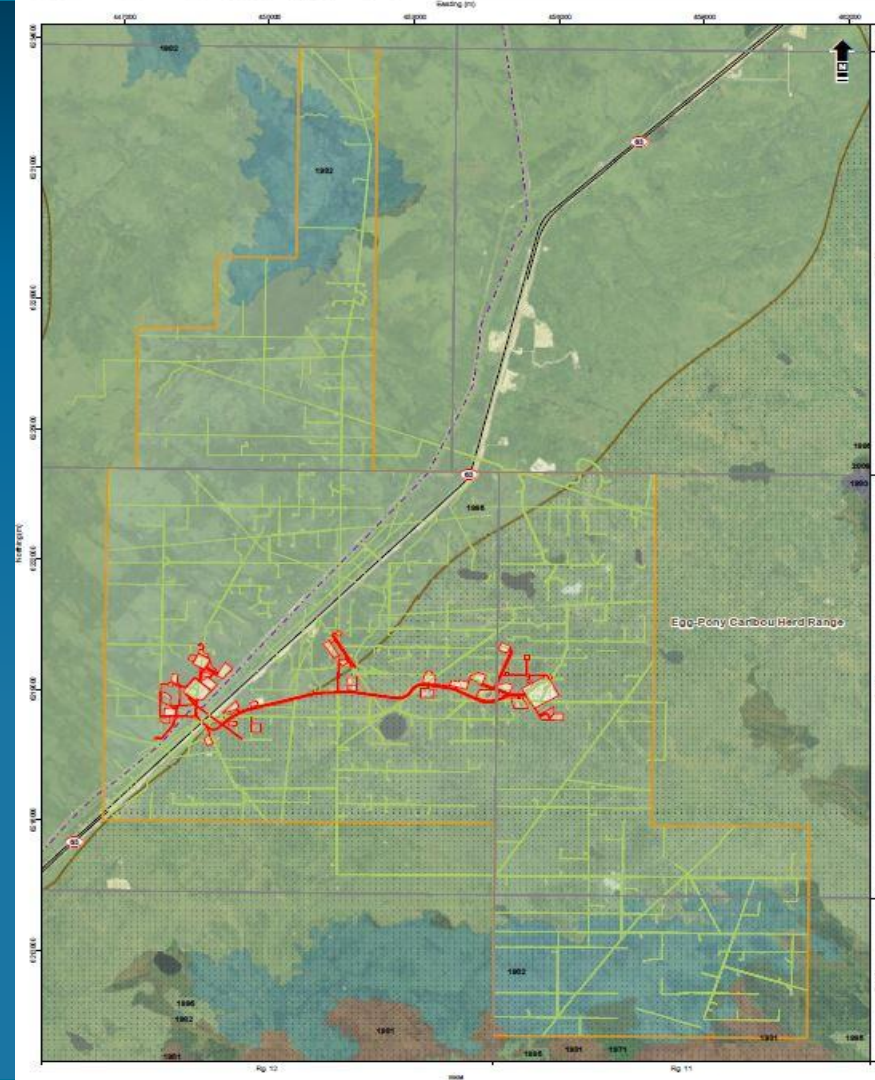
Regulatory Driver: field based or alternative research approach

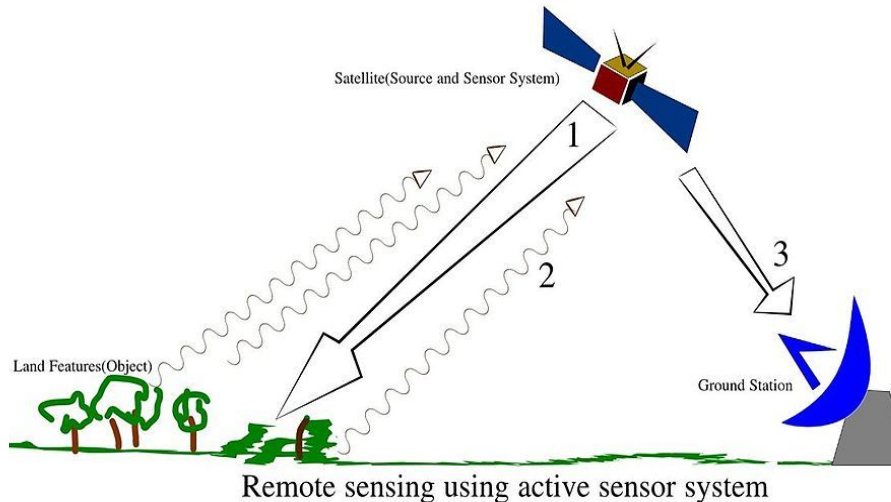
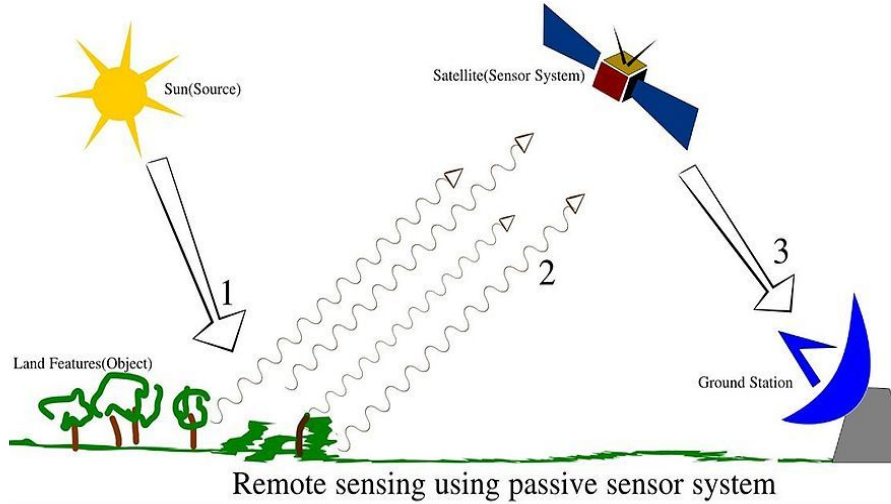
Why did we decide on the alternative research approach?

Hypothesis: metrics of wetland restoration can be quantified from remote sensing imagery

Project Area Overview

- Egg pony caribou herd range
- Multiple fires within the lease boundary





Remote Sensing

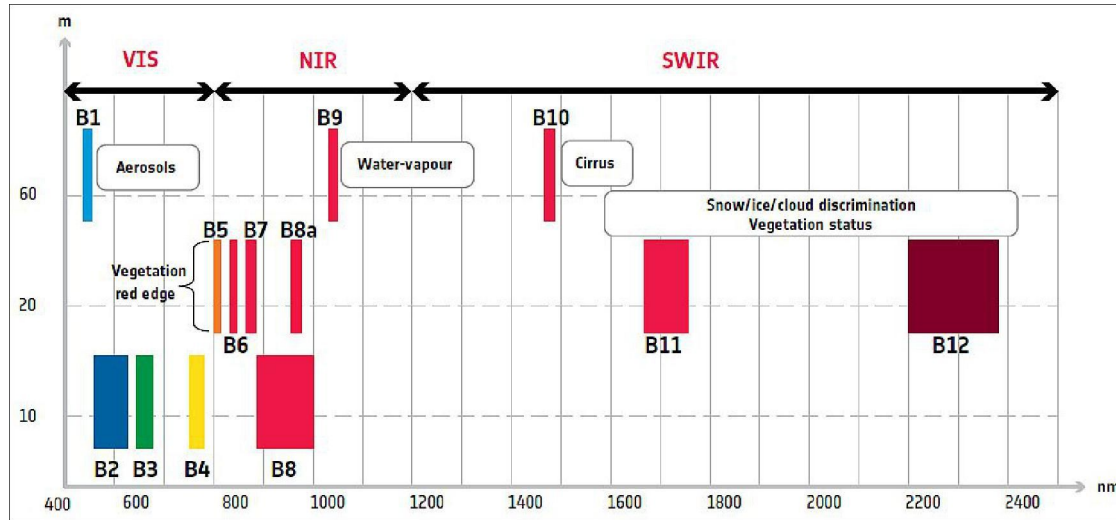
“Remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance.” – USGS

- Passive vs. Active
- Satellite or aircraft platform

Remote Sensing Earth Observation Examples

Sensor Type	Preferred Platform	Sensor Type	Source Examples	Usage Examples
Light Detection and Ranging (LiDAR)	Airplane	Active (infrared)	Commissioned from specialized private vendors (airplane); Public	Digital elevation model (DEM); Vegetation/building heights; Archaeology
High-resolution visible imagery	Airplane	Passive	Many environmental consultants can do this (inexpensive UAV/DSLR)	Visual air photo interpretation; Some ML capabilities, especially object detection
Multispectral	Airplane/Satellite	Passive	Public: LANDSAT (NASA); Sentinel-2 (ESA); Private vendors	All Earth science and geographical disciplines
Synthetic Aperture Radar (SAR)	Satellite	Active (microwave)	Sentinel-1 (ESA); NISAR (NASA/ISRO)	Generally lower resolution

Multispectral Imagery



Source: https://esamultimedia.esa.int/docs/EarthObservation/Sentinel-2_ESA_Bulletin161.pdf

- Divides the EM spectrum into small pieces
- Each band provides different information
- Sentinel-2 has 12 bands
- Highest resolution: 10×10 m

Modelling Approach – Starting Point



10×10 m grid of Sentinel-2 pixels



Field point data from 2016

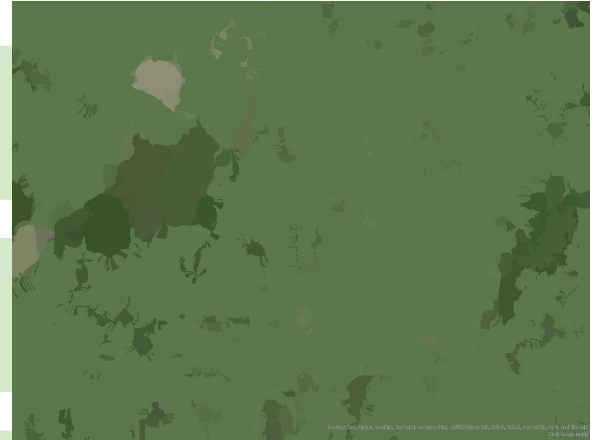


Segmentation: Spatial groupings of similarly-valued pixels



Classification: Assigning spatial groupings of pixels to a set of landcover classes

Wetland classes
Forest
Water
Developed



Don't Reinvent the Wheel: Wetland Detection

- Ancillary data sources can tell us where to look for wetlands
 - ABMI Wetland Inventory
- Can evidence of recovery from a disturbed state be identified within likely wetland areas?

Modelling Approach – Current

- Identify homogeneous wetland vegetation communities on the landscape
- Vegetation assessment
- Predictive models for each of a set of vegetation metrics
 - Predictions generated in annual time steps
- Weighted overlay/”suitability model” approach by vegetation ecologists to discern different community trajectories

Vegetation Metrics

Metric	Purpose	Status Update
Canopy composition and presence of tamarack	<ul style="list-style-type: none">Identify wetland communitiesIndicator species of fen wetlands	<ul style="list-style-type: none">In progress
Presence of jack pine and change in cover over time	<ul style="list-style-type: none">Species is present due to fire disturbanceExpect desirable wetland trees increases; jack pine decreases	<ul style="list-style-type: none">Preliminary model developed
Vegetation height	<ul style="list-style-type: none">Evaluate vegetation health (e.g., stunted trees) and success of the vegetation community	<ul style="list-style-type: none">Method is still TBD
Persistent graminoid species	<ul style="list-style-type: none">Present in graminoid fens and in early stages of regeneration	<ul style="list-style-type: none">In progress

Field Program

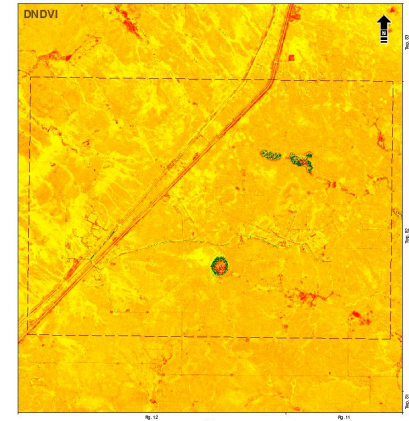
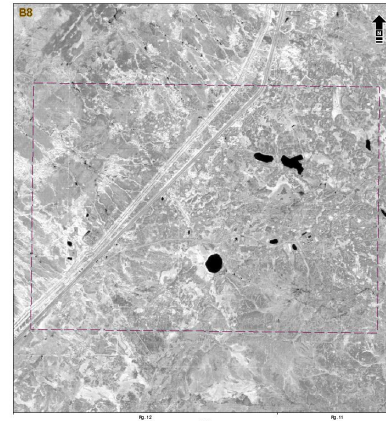
- Aerial survey with opportunistic ground-truthing
- Data were collected as polygons
- Work within the 1995 fire area only



Current Modelling Approach

Using machine learning models to relate multispectral imagery bands and derived metrics to field data

- Peak summer Sentinel 2 imagery
- Summer normalized difference vegetation index (NDVI)
- Spring vs. summer difference NDVI (DNDVI; green up)
- Summer vs. fall DNDVI (brown up)
- Fall normalized difference wetness index (NDWI; persistent late-season wetness)



Why Random Forest Models?

Principle:

- Targeted sampling induces correlated training samples
 - Not random or independent, as would be preferred in statistical models
- Multispectral imagery contains correlated features
- De-correlating training samples and features will yield a model that generalizes better
- Generalization outside the training data = more accurate predictions across the study area

Random Forest Algorithm

For each iteration:

- Take a random subset of the training data and a subset of the imagery bands
- Construct a decision tree:
 - Iterate through the bands and find the one that best separates the training data based on the target value
 - For the next decision, do the same thing again with the remaining bands
 - Continue until the desired tree depth is reached

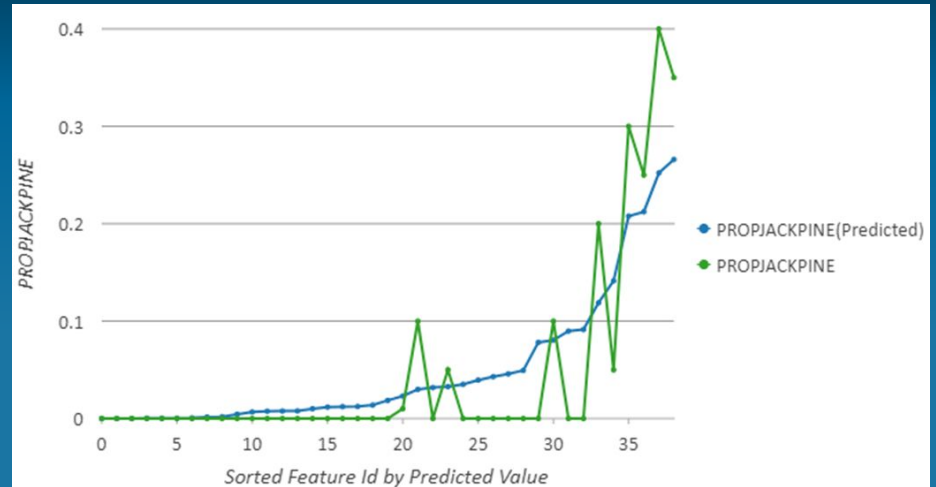
Aggregate:

- Classification: Majority vote
- Regression: Averaging

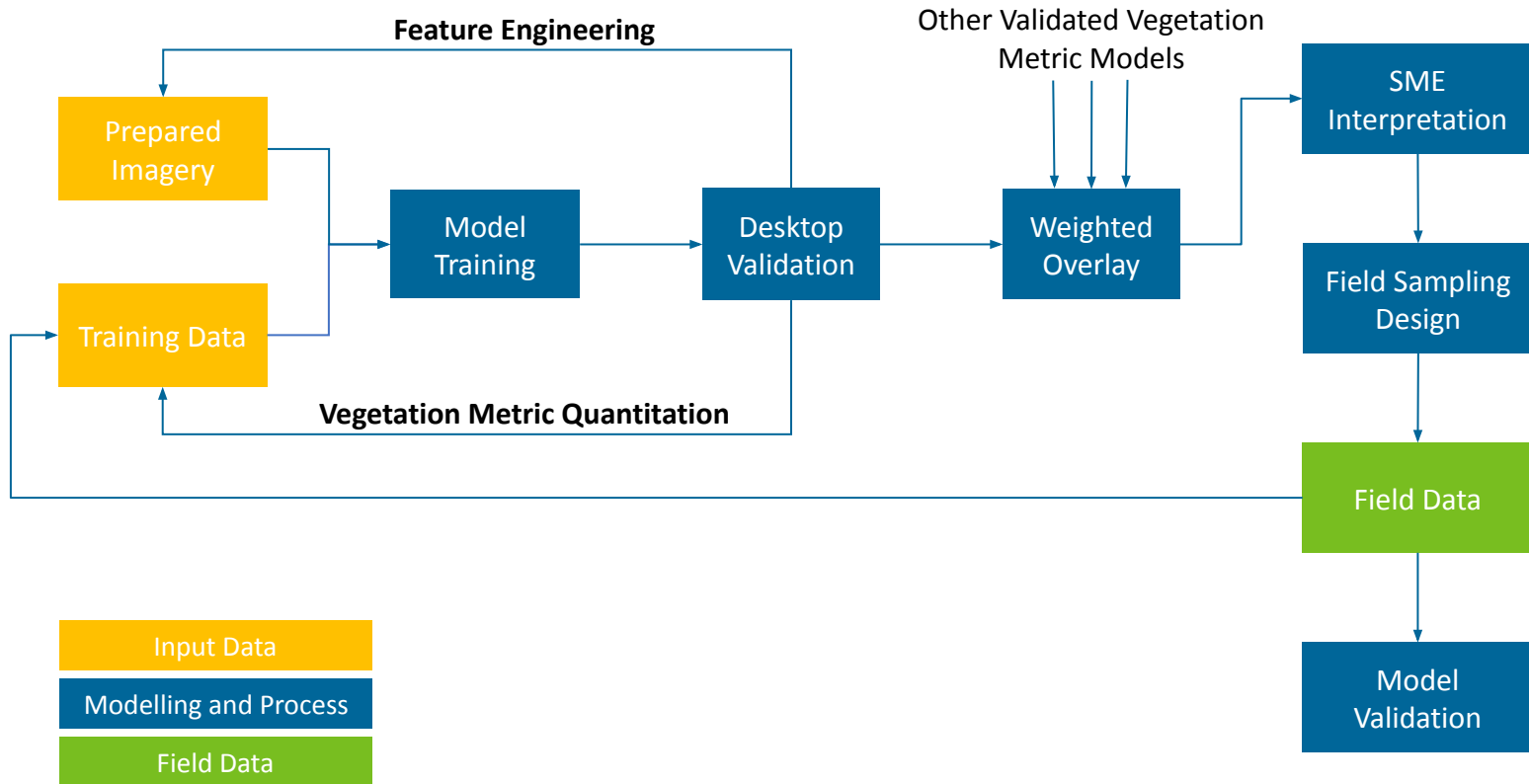
Modelling Outcomes

Proportion of Jack Pine in the Tree Stratum

- Promising results from the initial model run
- Apparent overprediction in some polygons is due to jack pine in the shrub stratum



Review of Modelling Process



Next Steps

- Data model refinement
- Suitability model and dataset review
- Field program to collect additional training data
- Model outcome review



Contact Us

Carl Ayer, B.Sc., Ept.
Environmental Data Analyst
cayer@matrix-solutions.com
403-237-0606

Meghan Hellman, B.Sc., P.Ag.
Senior Ecologist
mhellman@matrix-solutions.com
403-581-6318

matrix-solutions.com

