

# OSM\_2022-2023

This repository contains data, R scripts and associated outputs, and other materials necessary for the Applied Conservation and Macro Ecology (ACME) lab's Oil Sands Monitoring project for the 2022-2023 season.

## GENERAL INFORMATION

### Project Information

Details for the Oil Sands Monitoring Program study design can be found [here](#)

Also visit the ACME website [more information](#) about the OSM project and the ACME lab.

### Author Information (data):

Principal Investigator Contact Information

Name: Jason T. Fisher, PhD

Institution: University of Victoria

Address: 3800 Finnerty Rd, Victoria, BC V8P 5C2

Email: [fisherj@uvic.ca](mailto:fisherj@uvic.ca)

### Author Information (code):

Data Analysis Contact Information

Name: Marissa A. Dyck

Institution: University of Victoria

Address: 3800 Finnerty Rd, Victoria, BC V8P 5C2

Email: [marissadyck@uvic.ca](mailto:marissadyck@uvic.ca)

**Date of data collection:** 2022-2023

**Geographic location of data collection:** Alberta, Canada

## DATA & FILE OVERVIEW

### File List:

*Files in main folder*

- **OSM\_2022-2023.Rproj**; R project to run code for data cleaning and analyses.
- **README**; this README file with various extensions for viewing (.html, .pdf) and editing (.md)

*Files in data folder*

*/processed*

This folder includes data for the current year (2022-2023) that has been cleaned and reformatted using scripts within this repository. It also includes some data from the prior year (2021-2022) that was cleaned and formatted using similar scripts in a repository for that year.

- **OSM\_covariates\_2021.csv**; contains cleaned HFI and land cover data from LUs sampled in 2021-2022
- **OSM\_covariates\_2022.csv**; contains cleaned HFI and land cover data from LUs sampled in 2022-2022
- **OSM\_covariates\_grouped\_2021\_2022.csv**; contains cleaned HFI and land cover data that have also been grouped based on feature similarity from LUs sampled in 2021-2022 and 2022-2023
- **OSM\_covariates\_merged\_2021\_2022.csv**; contains cleaned HFI and land cover data from LUs sampled in 2021-2022 and 2022-2023 (*Note these have not been grouped yet*)

- **OSM\_covariates\_wide\_2022.csv**; contains cleaned data pertaining to the deployment location, data, time, and functionality of the cameras used in the project
- **OSM\_deployment\_2022.csv**; contains cleaned HFI and land cover data that have also been grouped based on feature similarity from LUs sampled in 2021-2022 and 2022-2023
- **OSM\_ind\_det\_2021.csv**; contains independent detections for species detected on cameras from LUs sampled in 2021-2022
- **OSM\_ind\_det\_2022.csv**; contains independent detections for species detected on cameras from LUs sampled in 2022-2023
- **OSM\_proportional\_detections\_2021.csv**; contains proportional monthly presence/absences data for a subset of mammals detected on cameras from LUs sampled in 2021-2022
- **OSM\_proportional\_detections\_2022.csv**; contains proportional monthly presence/absences data for a subset of mammals detected on cameras from LUs sampled in 2022-2023
- **OSM\_proportional\_detections\_merged\_2021\_2022.csv**; contains proportional monthly presence/absences data for a subset of mammals detected on cameras from LUs sampled in 2021-2022 and 2022-2023
- **OSM\_timelapse\_2021.csv**; contains cleaned image data from program Timelapse for all species and LUs sampled in 2021-2022 (not formatted into independent detections)
- **OSM\_timelapse\_2022.csv**; contains cleaned image data from program Timelapse for all species and LUs sampled in 2022-2023 (not formatted into independent detections)
- **OSM\_total\_detections\_2022.csv**; contains total independent detections for all species and LUs sampled in 2022-2023
- **OSM\_total\_presence\_absence\_2022.csv**; contains binary presence/absences (0s, 1s) for all species and LUs sampled in 2022-2023

*/raw*

This folder includes raw data for the current year (2022-2023), and some for the previous year (2021-2022).

- **OSM\_Deployment\_Data\_2022.csv**; contains raw data pertaining to the deployment location, data, time, and functionality of the cameras used in the project
- **OSM\_Deployment\_Site\_Data\_2021.csv**; contains raw data pertaining to the location, gps coordinates, date, time, and description of the sites where cameras were deployed for 2021-2022
- **OSM\_Deployment\_Site\_Data\_2022.csv**; contains raw data pertaining to the location, gps coordinates, date, time, and description of the sites where cameras were deployed for 2022-2023
- **OSM\_LU01\_LU13\_LU15\_HFI\_2022\_2024-04-19.csv**; contains raw data on Human Footprint Inventory (HFI) for each site sampled in 2022-2023 at varying buffer distances
- **OSM\_LU01\_LU13\_LU15\_VEG\_2022\_2024-04-19.csv**; contains raw data on land cover for each site sampled in 2022-2023 at varying buffer distances

*Files in data deposit folder*

This folder includes data and relevant files needed for an internal (OSM TBM) data deposit requested by Tobi Oke in 2024. It is maintained in this folder in case future data deposits are required, ACME has a template to work from of what the files structure should look like. This entire folder was zipped and sent to Tobi Oke.

- **00\_README**; readme file for viewing by data requisitor (.html) or editing by ACME (.md)
- **01\_ACME\_Camera\_Image\_processing\_Protocol.pdf**; document that describes how meta-data are defined and entered for raw image data in Timelapse
- **OSM\_ABMI\_covariates\_grouping\_table**; document defining how HFI covariates from ABMI wall-to-wall data were grouped in the the data submitted with this data deposit
- **OSM\_camera\_data\_2021-2022.csv**; contains cleaned image and camera deployment data for LUs sampled in 2021-2022 and 2022-2023
- **OSM\_landscape\_data\_2021-2022.csv**; contains cleaned, formatted, and group HFI and land cover data from ABMI wall-to-wall data sets for sites sampled between 2021-2022 and 2022-2023

#### *Files in figures folder*

This folder contains various plots generated in the scripts of this repository for the purposes of data visualization.

- **OSM\_body\_size\_buffers\_2022.jpg**; plot depicting the top buffer size for each species from the global models; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_detections LU01.jpg**; plot depicting the number of independent detections per species for LU 01; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_detections LU02.jpg**; plot depicting the number of independent detections per species for LU 02; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_detections LU03.jpg**; plot depicting the number of independent detections per species for LU 03; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_detections LU13.jpg**; plot depicting the number of independent detections per species for LU 13; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_detections LU15.jpg**; plot depicting the number of independent detections per species for LU 15; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_detections LU21.jpg**; plot depicting the number of independent detections per species for LU 21; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_ind\_det\_per\_LU\_2022.jpg**; plot depicting the number of independent detections per species for the 4 LUs sampled in 2022-2023; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_LU01\_det\_occ\_plots.jpg**; plot depicting the number of independent detections and naive occupancy per species for LU 01; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_LU02\_det\_occ\_plots.jpg**; plot depicting the number of independent detections and naive occupancy per species for LU 02; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_LU03\_det\_occ\_plots.jpg**; plot depicting the number of independent detections and naive occupancy per species for LU 03; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_LU13\_det\_occ\_plots.jpg**; plot depicting the number of independent detections and naive occupancy per species for LU 13; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_LU15\_det\_occ\_plots.jpg**; plot depicting the number of independent detections and naive occupancy per species for LU 15; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_LU21\_det\_occ\_plots.jpg**; plot depicting the number of independent detections and naive occupancy per species for LU 21; generated in the `3_OSM_glm_analysis_2022.rmd` script

- **OSM\_Naive Occupancy LU01.jpg**; plot depicting the naive occupancy per species for LU 01; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_Naive Occupancy LU02.jpg**; plot depicting the naive occupancy per species for LU 02; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_Naive Occupancy LU03.jpg**; plot depicting the naive occupancy per species for LU 03; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_Naive Occupancy LU13.jpg**; plot depicting the naive occupancy per species for LU 13; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_Naive Occupancy LU15.jpg**; plot depicting the naive occupancy per species for LU 15; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_Naive Occupancy LU21.jpg**; plot depicting the naive occupancy per species for LU 21; generated in the `3_OSM_glm_analysis_2022.rmd` script
- **OSM\_proportional\_detections\_2022.jpg**; plot depicting the response metric (months present) for each species for each site in the 4 LUs sampled in 2022-2023; generated in the `1_ACME_camera_script_2024-05-06..rmd` script
- **OSM\_species\_anthro\_buffers\_2022.jpg**; plot depicting the top buffer for each species in the anthropogenic analysis; generated in the `4_OSM_glm_figures_2022.rmd` script
- **OSM\_species\_global\_buffers\_2022.jpg**; plot depicting the top buffer for each species in the global analysis; generated in the `4_OSM_glm_figures_2022.rmd` script
- **OSM\_species\_land\_buffers\_2022.jpg**; plot depicting the top buffer for each species in the landscape analysis; generated in the `4_OSM_glm_figures_2022.rmd` script
- **OSM\_total\_detections\_site\_2022.jpg**; plot depicting the response metric (total independent detections) for each species for each site in the 4 LUs sampled in 2022-2023; generated in the `1_ACME_camera_script_2024-05-06.rmd` script
- **OSM\_total\_presence\_absence\_2022.jpg**; plot depicting the response metric (binary presence/absence) for each species for each site in the 4 LUs sampled in 2022-2023; generated in the `1_ACME_camera_script_2024-05-06.rmd` script

/publication\_figures

This sub folder contains final publication quality figures for the scale analysis conducted in this repository.

- **figure\_2\_a.jpg**; plot with three panels, each depicting the top buffer sizes for each species for one of the models run in the `3_OSM_glm_analysis.rmd` script; generated in the `4_OSM_glm_figures_2022.rmd` script
- **figure\_2\_b.jpg**; plot depicting the top buffer sizes for each species for the 3 analysis run `3_OSM_glm_analysis.rmd` script, with points color coded by model type; generated in the `4_OSM_glm_figures_2022.rmd` script
- **figure\_3.jpg**; plot depicting the top buffer sizes for each species based on log body size for the 3 analysis run `3_OSM_glm_analysis.rmd` script, with points color coded by model type; generated in the `4_OSM_glm_figures_2022.rmd` script
- **figure\_S1.jpg**; plot depicting the model weights for all models (buffer widths) for each species for the global analysis; generated in the `4_OSM_glm_figures_2022.rmd` script
- **figure\_S2.jpg**; plot depicting the model weights for all models (buffer widths) for each species for the anthropogenic analysis; generated in the `4_OSM_glm_figures_2022.rmd` script

- **figure\_S3.jpg**; plot depicting the model weights for all models (buffer widths) for each species for the landscape analysis; generated in the `4_OSM_glm_figures_2022.rmd` script

#### *Files in scripts folder*

This file contains the various scripts needed for data formatting, visualization, and analysis.

- **0\_ACME\_read\_timelapse\_script-2024-04-04**; .rmd file and knitted .html file that will take all of the individual imagery folders and move them to a common location in the ACME Netdrive
- **1\_ACME\_camera\_script-2024-05-06**; .rmd file and knitted .html/.pdf files for data cleaning and formatting
- **2\_ACME\_landscape\_covariate\_exploration\_script**; .rmd file and knitted .html/.pdf files for further data cleaning, formatting, and exploration of the covariate (HFI and land cover) data
- **3\_OSM\_glm\_analysis\_2022**; .rmd file and knitted .html/.pdf files for running th GLMM anaylsis to look at top buffer scales for nine species of mammals
- **4\_OSM\_glm\_figures\_2022**; .rmd file and knitted .html/.pdf files for making the publication quality figures

## RAW DATA

### DATA-SPECIFIC INFORMATION FOR: [OSM\_2022\_Deployment\_Data.csv]

- **Number of variables/columns:** 6
- **Number of observations/rows:** 155 (one per camera site)

#### Variable List:

- **Project.ID**, factor where the first element abbreviation describes the project (e.g. OSM for Oil Sands Monitoring) and the second describes the landscape unit (i.e. LU13, LU15, LU21, or LU01).
- **Deployment.Location.ID**, factor where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **Camera.Deployment.Begin.Date.**, date indicating when the camera was deployed.
- **Camera.Deployment.End.Date**, date indicating when the camera was retrieved.
- **Deployment**, factor indicating which the deployment for the associated camera (all are deployment 1 or D1).
- **Camera.Failure.Details**, character describing any issues with the camera during deployment.

### DATA-SPECIFIC INFORMATION FOR: [OSM\_Deployment\_Site\_Data\_2021.csv]

- **Number of variables/columns:** 21
- **Number of observations/rows:** 79 (one per camera site)

#### Variable List:

- **Array**, a factor with 2 levels where the first element abbreviation describes the project (e.g. OSM for Oil Sands Monitoring) and the second describes the landscape unit.

- **DeployDate**, date indicating when the camera was deployed.
- **Deploy Time**, time (in 24hrs) indicating when the camera was deployed.
- **Crew**, factor with the initials of the crew who deployed the camera.
- **Grid Cell #**, a factor with 79 levels where the first element abbreviation describes the landscape unit (i.e. LU02 or LU03) and the second element describes the camera site.
- **Camera Site #**, [same as Grid Cell #] - factor where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **Camera Unit #**, a unique character identifier for the camera unit that was deployed.
- **SD Card #**, a unique character identifier for the SD card that was deployed with the camera unit.
- **Lat**, numeric latitudinal geographic coordinates for the camera location.
- **Long**, numeric longitudinal geographic coordinates for the camera location.
- **GPS Label**, factor with 79 levels where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **Forest Type**, factor with 2 levels describing the type of forest at the camera site (i.e. Broadleaf, Mixed, Conifer, or Coniferous).
- **Camera Site Description**, brief description characterizing the habitat of the camera site, completed by the technicians who deployed the camera at the camera site
- **Topography**, a factor with 3 levels describing the topography of the camera site (i.e. Flat, Slope, or Ridge).
- **Grade (%)**, a numeric measurement of the slop (incline/decline) of the camera site.
- **Elevation (m)**, a numeric measurement of the elevation in meters of the camera site.
- **Distance to Trail (m)**, a numeric measurement of the distance to the nearest trail in meters.
- **Camera Direction**, a factor with 8 levels indication the cardinal direction the camera was pointed (i.e. NE, N, NW, W, E, SE, S, SW).
- **Trail Use Rating**, a factor with 3 levels indicating the trail use (1 low, 3 high) of the animal trail the camera was deployed near
- **Distance to Lure**, a numeric measurement of the distance from the camera to the lure in meters.
- **Comments and Access Notes**, a character with notes for the field crew.

#### DATA-SPECIFIC INFORMATION FOR: [OSM\_Deployment\_Site\_Data\_2022.csv]

- **Number of variables/columns:** 21
- **Number of observations/rows:** 155 (one per camera site)

#### Variable List:

- **Array**, a factor with 4 levels where the first element abbreviation describes the project (e.g. OSM for Oil Sands Monitoring) and the second describes the landscape unit (i.e. LU13, LU15, LU21, or LU01).
- **DeployDate**, date indicating when the camera was deployed.

- **Deploy Time**, time (in 24hrs) indicating when the camera was deployed.
- **Crew**, factor with the initials of the crew who deployed the camera.
- **Grid Cell #**, a factor with 155 levels where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **Camera Site #**, [same as Grid Cell #] - factor where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **Camera Unit #**, a unique character identifier for the camera unit that was deployed.
- **SD Card #**, a unique character identifier for the SD card that was deployed with the camera unit.
- **Lat**, numeric latitudinal geographic coordinates for the camera location.
- **Long**, numeric longitudinal geographic coordinates for the camera location.
- **GPS Label**, factor with 155 levels where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **Forest Type**, factor with 4 levels describing the type of forest at the camera site (i.e. Broadleaf, Mixed, Conifer, or Coniferous).
- **Camera Site Description**, brief description characterizing the habitat of the camera site, completed by the technicians who deployed the camera at the camera site
- **Topography**, a factor with 3 levels describing the topography of the camera site (i.e. Flat, Slope, or Ridge).
- **Grade (%)**, a numeric measurement of the slop (incline/decline) of the camera site.
- **Elevation (m)**, a numeric measurement of the elevation in meters of the camera site.
- **Distance to Trail (m)**, a numeric measurement of the distance to the nearest trail in meters.
- **Camera Direction**, a factor with 8 levels indication the cardinal direction the camera was pointed (i.e. NE, N, NW, W, E, SE, S, SW).
- **Trail Use Rating**, a factor with 3 levels indicating the trail use (1 low, 3 high) of the animal trail the camera was deployed near
- **Distance to Lure**, a numeric measurement of the distance from the camera to the lure in meters.
- **Comments and Access Notes**, a character with notes for the field crew.

#### DATA-SPECIFIC INFORMATION FOR: [OSM\_LU01\_LU13\_LU15\_LU21\_HFI\_2022.csv]

*Information on exact methods for data extraction and more specific variable descriptions can be found on the ABMI human footprints wall to wall data download website for Year 2021 OR in the relevant\_literature folder of this repository (HFI\_2021\_v1\_0\_Metadata\_Final.pdf).*

- **Number of variables/columns:** 61
- **Number of observations/rows:** 3100

#### Variable List:

- **Site**, a factor with 155 levels where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.

- **BUFF\_DIST**, a numeric measurement in meters ranging from 250 - 5000, of the buffer radius around the camera for which the proportion of associated human factors variables were calculated.
- **FEATURE\_TYHARVEST.AREA**, a numeric variable indicating the proportion of harvest area within the buffer. Harvest area is defined as, areas where forestry operations have occurred (clear-cut, selective harvest, salvage logging, etc.), specifically areas in Alberta's forested Green Zone where forestry operations have occurred (clear-cut, selective harvest, salvage logging, etc.).
- **FEATURE\_TYCROP**, a numeric variable indicating the proportion of crop within the buffer. Crop is defined as, cultivated cropland or cropland planted with annual crop species, including farmlands that are in cultivation rotation.
- **FEATURE\_TYWELL.ABAND**, a numeric variable indicating the proportion of abandoned wells within the buffer. Abandoned wells are defined as, ground cleared for an oil/gas well pad where the well is currently abandoned.
- **FEATURE\_TYWELL.OIL**, a numeric variable indicating the proportion of oil wells within the buffer. Oil wells are defined as, ground cleared for an oil well pad where at least one well is currently active.
- **FEATURE\_TYTRAIL**, a numeric variable indicating the proportion of trails within the buffer. Trails are defined as, a polygon feature class derived from a 2-meter buffer (4 meter total width) of a pre-low-impact-seismic centerline.
- **FEATURE\_TYHARVEST.AREA.WHITE.ZONE**, a numeric variable indicating the proportion of harvest area white zone within the buffer. Harvest area white zone is defined as, areas where forestry operations have occurred (clear-cut, selective harvest, salvage logging, etc.). specifically, areas in Alberta's unforested White Zone where woody vegetation (i.e. shrub, trees, etc..) have been removed and the purpose of the clearing has not yet been determined.
- **FEATURE\_TYCONVENTIONAL.SEISMIC**, a numeric variable indicating the proportion of conventional seismic lines within the buffer. Conventional seismic lines are defined as, a polygon feature class derived from a 3-meter buffer (6 meter total width) of a pre-low-impact-seismic centerline.
- **FEATURE\_TYPIPELINE**, a numeric variable indicating the proportion of pipelines within the buffer. Pipelines are defined as, a line of underground and overground pipes, of substantial length and capacity, used for the conveyance of petrochemicals.
- **FEATURE\_TYTAME\_PASTURE**, a numeric variable indicating the proportion of tame pasture within the buffer area. Tame pasture is defined as, lands where the soil has been disturbed and planted to perennial grass species used primarily for grazing livestock. Specifically, tame pasture represents areas of grasses, legumes or grass-legume mixtures planted for livestock grazing or hay collection.
- **FEATURE\_TYROUGH\_PASTURE**, a numeric variable indicating the proportion of rough pasture within the buffer area. Rough pasture is defined as, lands where the forest and/or shrubs have been removed so that native or introduced grasses can flourish for the grazing of livestock. Specifically, this pastureland has not been irrigated or fertilized and the soil has not been disturbed to improve productivity.
- **FEATURE\_TYRURAL.RESIDENCE**, a numeric variable indicating the proportion of rural residence within the buffer area. Rural residence is defined as, country-residential developments with density of 10 - 100 buildings per quarter section.
- **FEATURE\_TYTRANSMISSION.LINE**, a numeric variable indicating the proportion of transmission lines within the buffer area. Transmission lines are defined as, cleared corridors designated for the location of power transmission line infrastructure.



- **FEATURE\_TYWELL.GAS**, a numeric variable indicating the proportion of gas wells within the buffer area. Gas wells are defined as, ground cleared for an gas well pad where at least one well is currently active.
- **FEATURE\_TYMISC.OIL.GAS.FACILITY**, a numeric variable indicating the proportion of misc. oil/gas facilities within the buffer area. Misc. oil/gas facilities are defined as, human footprint features related to various industrial activities. Specifically, industrial facility used for the purpose of oil and gas. BATTERY SITE, COMPRESSOR SITE, FLARE STACK, METER STATION SITE, and VALVE SITE.
- **FEATURE\_TYCLEARING.UNKNOWN**, a numeric variable indicating the proportion of unknown clearings within the buffer area. Unknown clearings are defined as, a human-made clearing with unknown purposes and contains no visible buildings, fences or equipment.
- **FEATURE\_TYVEGETATED.EDGE.ROADS**, a numeric variable indicating the proportion of vegetated edge roads within the buffer area. Vegetated edge roads are defined as, disturbed vegetation alongside road edges and railway edges including ditches.
- **FEATURE\_TYROAD.UNIMPROVED**, a numeric variable indicating the proportion of unimproved roads within the buffer area. Unimproved roads are defined as, a roadway surfaced with dirt, which is constituted as a minor access route. The road surface is up to 7 meters in width, and the road clearing is up to 20 meters in width. The surface and ditches are poorly maintained, and the bridges are narrow.
- **FEATURE\_TYROAD.GRAVEL.1L**, a numeric variable indicating the proportion of 1L gravel roads within the buffer area. 1L gravel roads are defined as, a roadway surfaced with gravel constituting a main access route. The road surface is about 6 meters in width, and the road clearing is about 20 meters or greater in width. The surface, ditches, bridges and intersections are in good condition.
- **FEATURE\_TYROAD.GRAVEL.2L**, a numeric variable indicating the proportion of 2L gravel roads within the buffer area. 2L gravel roads are defined as, a roadway surfaced with gravel constituting as a main access route. The road surface is 7 meters or greater in width, and the road clearing is 30 meters or greater in width. The surface, ditches, bridges and intersections are in good condition.
- **FEATURE\_TYTRUCK.TRAIL**, a numeric variable indicating the proportion of truck trails within the buffer area. Truck trails are defined as, a roadway surfaced with dirt or low vegetation constituting a minor access route.
- **FEATURE\_TYBORROWPITS**, a numeric variable indicating the proportion of borrow pits within the buffer area. Borrow pits are defined as, excavation outside of the road right-of-way, made solely for the purpose of removing or providing borrowed material for the construction of the sub-base for a specific roadway project. It includes any other associated infrastructure such as access roads. (ALBERTA TRANSPORTATION; GUIDE TO RECLAIMING BORROW EXCAVATIONS – 2013 Edition). Specifically, includes pits dug to build forestry and well-site roads. They are usually associated with a road or another structure.
- **FEATURE\_TYSUMP**, a numeric variable indicating the proportion of sumps within the buffer area. Sumps are defined as, An artificial holding or treatment pond for industrial wastewater. Drilling waste storage system – holding of drilling waste on well sites or remotely. Either earthen excavation (in clayey soils) or sumps lined with a synthetic liner.
- **FEATURE\_TYBORROWPIT.WET**, a numeric variable indicating the proportion of wet borrow pits within the buffer area. Wet borrow pits are defined as, excavation outside of the road right-of-way, made solely for the purpose of removing or providing borrowed material for the construction of the sub-base for a specific roadway project. It includes any other associated infrastructure such as access roads. (ALBERTA TRANSPORTATION; GUIDE TO RECLAIMING BORROW EXCAVATIONS – 2013 Edition). Specifically, includes pits dug to build forestry and well-site roads. They are usually associated with a road or another structure. Presence of water confirmed by visual interpretation.

- **FEATURE\_TYCULTIVATION\_ABANDONED**, a numeric variable indicating the proportion of abandoned cultivation within the buffer area. Abandoned cultivation is defined as, agricultural land that has been formally seeded and tilled, but no evidence of present day production use. Landscape appears to have a heterogeneous mix of vegetation and closely resembles natural cover.
- **FEATURE\_TYURBAN.RESIDENCE**, a numeric variable indicating the proportion of urban residence within the buffer area. Urban residence is defined as, residential areas in cities, towns, villages, hamlets and ribbon developments. Areas that are dominated by dwellings.
- **FEATURE\_TYCOUNTRY.RESIDENCE**, a numeric variable indicating the proportion of country residence within the buffer area. Country residence is defined as, country-residential developments with density of 10 - 100 buildings per quarter section.
- **FEATURE\_TYRECREATION**, a numeric variable indicating the proportion of recreational area within the buffer area. Recreation is defined as, urban/rural greenspace and recreation that does not fit into other categories (e.g. graveyards, baseball diamonds, parks, shelterbelts, ski hills, clearings from old industrial activity that is now vegetated). This layer was also used to identify green-space features that do not fit into other categories such as storage areas and parking lots.
- **FEATURE\_TYWELL.OTHER**, a numeric variable indicating the proportion of other wells within the buffer area. Other wells are defined as, ground cleared for an oil/gas well pad where at least one well is currently active.
- **FEATURE\_TYWELL.BITUMEN**, a numeric variable indicating the proportion of bitumen wells within the buffer area. Bitumen wells are defined as, ground cleared for an oil/gas well pad where at least one well is currently active.
- **FEATURE\_TYWELL.CASED**, a numeric variable indicating the proportion of cased wells within the buffer area. Cased wells are defined as, ground cleared for an oil/gas well pad where at least one well is currently active.
- **FEATURE\_TYROAD.PAVED.UNDIV.2L**, a numeric variable indicating the proportion of 2L paved roads within the buffer area. 2L paved roads are defined as, non-vegetated, impermeable surfaces used for motorized vehicle or aircraft transportation or access. Specifically, a major roadway, which is paved with asphalt or concrete, and consists of two (2) roadbeds separated by a median. Each road bed usually consists of two (2) or more lanes.
- **FEATURE\_TYROAD.UNCLASSIFIED**, a numeric variable indicating the proportion of unclassified roads within the buffer area. Unclassified roads are defined as, non-vegetated, impermeable surfaces used for motorized vehicle or aircraft transportation or access. Specifically, a temporary coding for an unknown class of road, which will be updated after a field check or verification.
- **FEATURE\_TYRUNWAY**, a numeric variable indicating the proportion of unclassified runways within the buffer area. Runways are defined as, human footprint related to vegetated facilities and recreation, specifically a vegetated runway.
- **FEATURE\_TYCLEARING.WELLPAD.UNCONFIRMED**, a numeric variable indicating the proportion of unconfirmed clearing well pads within the buffer area. Unconfirmed clearing well pads are defined as, human footprint features related to various industrial activities. Specifically, roughly square in shape clearing, roughly 90-120 meters wide (approximately 1 ha). Not confirmed as a well pad by available reference sources.
- **FEATURE\_TYFACILITY.UNKNOWN**, a numeric variable indicating the proportion of unknown facilities within the buffer area. Unknown facilities are defined as, human footprint features related to various industrial activities. Specifically, identifies areas where the reclamation liability associated for the disturbance is currently held by another industry operator.

- **FEATURE\_TYBORROWPIT.DRY**, a numeric variable indicating the proportion of dry borrow pits within the buffer area. Dry borrow pits are defined as, excavation outside of the road right-of-way, made solely for the purpose of removing or providing borrowed material for the construction of the sub-base for a specific roadway project. It includes any other associated infrastructure such as access roads. (ALBERTA TRANSPORTATION; GUIDE TO RECLAIMING BORROW EXCAVATIONS – 2013 Edition). Specifically, Includes pits dug to build forestry and well-site roads. They are usually associated with a road or another structure. No presence of water.
- **FEATURE\_TYGRVL.SAND.PIT**, a numeric variable indicating the proportion of gravel and sand pits within the buffer area. Gravel and sand pits are defined as, a human footprint features directly related to mining activities. Specifically, an area of surface disturbance for the purpose of extracting sand and/or gravel consistently open and/or expanding over multiple years, usually close to lakes or rivers.
- **FEATURE\_TYDUGOUT**, a numeric variable indicating the proportion of dugouts within the buffer area. Dugouts are defined as, small water storage excavations that collect water from runoff from summer rains, a surplus of surface water that occurs during snowmelt in the spring or from groundwater. (Alberta Agriculture and Rural Development, QUALITY FARM DUGOUTS).
- **FEATURE\_TYLAGOON**, a numeric variable indicating the proportion of lagoons within the buffer area. Lagoons are defined as, an artificial holding or treatment ponds for agricultural or municipal wastewater. Human made water and sewage lagoons used for municipal purposes.
- **FEATURE\_TYOPEN.PIT.MINE**, a numeric variable indicating the proportion of open pit mines within the buffer area. Open pit mines are defined as, human footprint features directly related to mining activities. Specifically, an area of surface disturbance for the purpose of mining (with the exception of sand and/or gravel), consistently open and/or expanding over multiple years, usually close to lakes or rivers.
- **FEATURE\_TYLOW.IMPACT.SEISMIC**, a numeric variable indicating the proportion of low impact seismic lines within the buffer area. Low impact seismic lines are defined as, a polygon feature class derived from a 1.5-meter buffer (3 meter total width) of a pre-low-impact-seismic centerline.
- **FEATURE\_TYSURROUNDING.VEG**, a numeric variable indicating the proportion of surrounding vegetation within the buffer area. Surrounding vegetation is defined as, human footprint related to vegetated facilities and recreation. Specifically, disturbed vegetation surrounding airport runways, highway ramps and other industrial features.
- **FEATURE\_TYTRANSFER\_STATION**, a numeric variable indicating the proportion of transfer stations within the buffer area. Transfer stations are defined as, human footprint used for the transportation, storage, and disposal of waste/garbage. Specifically, smaller area of land, less than one hectare, usually fenced with a U-shaped road and two entry ways. Used primarily for garbage drop-off and located close to municipalities or present in rural areas.
- **FEATURE\_TYFACILITY.OTHER**, a numeric variable indicating the proportion of other facilities within the buffer area. Other facilities are defined as, human footprint features related to various industrial activities. Specifically, industrial facilities characterized by large non-residential buildings most often surrounded by concrete for parking purposes. The purpose of the facility is not disclosed.
- **FEATURE\_TYVEGETATED.EDGE.RAILWAYS**, a numeric variable indicating the proportion of vegetated edge railways within the buffer area. Vegetated edge railways are defined as, disturbed vegetation alongside road edges and railway edges including ditches.
- **FEATURE\_TYFRUIT.VEGETABLES**, a numeric variable indicating the proportion of fruit and vegetables within the buffer area. Fruit and vegetables are defined as, crops based on Agriculture and Agri-Food Canada (AAFC) 2014 Crop Types.

- **FEATURE\_TYRESIDENCE\_CLEARING**, a numeric variable indicating the proportion of residence clearing within the buffer area. Residence clearing is defined as, areas cleared for building developments that do not yet have any buildings.
- **FEATURE\_TYCF0**, a numeric variable indicating the proportion of confined feeding operations within the buffer area. Confined feeding operations are defined as, confined feeding operations and other high density livestock features.
- **FEATURE\_TYLANDFILL**, a numeric variable indicating the proportion of confined landfill within the buffer area. Landfills are defined as, human footprint used for the transportation, storage, and disposal of waste/garbage. Specifically, larger area of raised land, indicating buried garbage. Some landfills have evidence of surface revegetation and garbage dispersed throughout designated extent. They may also have large perimeter berms or fences.
- **FEATURE\_TYWELL\_CLEARED\_NOT\_CONFIRMED**, a numeric variable indicating the proportion of unconfirmed wellpad clearings within the buffer area. Unconfirmed wellpad clearings are defined as, human footprint features related to various industrial activities. Specifically, Roughly square in shape clearing, roughly 90-120 meters wide (approximately 1 ha). Not confirmed as a well pad by available reference sources.
- **FEATURE\_TYOIL.GAS.PLANT**, a numeric variable indicating the proportion of oil/gas plants within the buffer area. Oil/gas plants are defined as, human footprint features related to various industrial activities. Specifically, Industrial facility used for oil production. REFINERIES, PLANTS, FACTORIES.
- **FEATURE\_TYURBAN.INDUSTRIAL**, a numeric variable indicating the proportion of urban-industrial area within the buffer area. Urban-industrial area is defined as, human footprint features related to various industrial activities. Specifically, an industrial facility within the boundary of an urban residence.
- **FEATURE\_TYROAD.PAVED.1L**, a numeric variable indicating the proportion of 1L paved roads within the buffer area. 1L paved roads are defined as, non-vegetated, impermeable surfaces used for motorized vehicle or aircraft transportation or access. Specifically, a roadway, paved with asphalt or concrete, consisting of one (1) lane.
- **FEATURE\_TYROAD.PAVED.UNDIV.1L**, a numeric variable indicating the proportion of UNDIV 1L paved roads within the buffer area. UNDIV 1L paved roads are defined as, non-vegetated, impermeable surfaces used for motorized vehicle or aircraft transportation or access. Specifically, a roadway, paved with asphalt or concrete, consisting of one (1) lane, and usually found servicing rural acreages that are close to large urban centers.
- **FEATURE\_TYROAD.WINTER**, a numeric variable indicating the proportion of winter roads within the buffer area. Winter roads are defined as, non-vegetated, impermeable surfaces used for motorized vehicle or aircraft transportation or access. Specifically, a clearing that is vehicular accessible in winter only.
- **FEATURE\_TYWELL\_CLEARED\_NOT\_DRILLED**, a numeric variable indicating the proportion of not drilled wells within the buffer area. Not drilled wells are defined as, ground cleared for an oil/gas well pad where at least one well is currently active. Specifically, a well site - confirmation of the boundary outline is provided by reference sources.
- **FEATURE\_TYWELL.UNKNOWN**, a numeric variable indicating the proportion of unknown wells within the buffer area. Unknown wells are defined as, ground cleared for an oil/gas well pad where at least one well is currently active. Specifically, a well site - ground cleared, well status unknown or license location.
- **FEATURE\_TYAIRP.RUNWAY**, a numeric variable indicating the proportion of airplane runways within the buffer area. Airplane runways are defined as, non-vegetated, impermeable surfaces used for

motorized vehicle or aircraft transportation or access. Specifically, an active landing facility for aircraft, usually associated with paved and lighted runways, an operating control tower, and services for aircraft and passengers.

- **FEATURE\_TYRESERVOIR**, a numeric variable indicating the proportion of reservoirs within the buffer area. Reservoirs are defined as, an artificial lake or storage pond resulting from human-made dam. A body of water created by excavation or the man-made damming of a river or stream.
- **FEATURE\_TYCAMPGROUND**, a numeric variable indicating the proportion of campgrounds within the buffer area. Campgrounds are defined as, human footprint related to vegetated facilities and recreation. Specifically, disturbed vegetation with frequently changing facilities of RVs and tents used for overnight stay. Most often consists of several individual clearings surrounded by vegetation and gravel or asphalt roads connecting clearings.
- **FEATURE\_TYCANAL**, a numeric variable indicating the proportion of canals within the buffer area. Canals are defined as, a human-made watercourse built to convey water for irrigation. An irrigation canal is larger than a ditch, with reinforced banks that are usually well maintained. A human-made drainage network channels built to prepare wetland areas for anthropogenic land use.
- **FEATURE\_TYCAMP.INDUSTRIAL**, a numeric variable indicating the proportion of industrial camps within the buffer area. Industrial camps are defined as, human footprint features related to various industrial activities. Specifically, buildings used for temporary residence by employees on or in close proximity to an industrial activity such as mining, forestry, or oil and gas activities.
- **FEATURE\_TYRLWY.SGL.TRACK**, a numeric variable indicating the proportion of single track railways within the buffer area. Single track railways are defined as, hard, steel rail lines designed for train use. Specifically, a road or track for trains, consisting of parallel steel rails, supported on wooden crossbeams. The single track consists of one parallel sets of tracks.

**DATA-SPECIFIC INFORMATION FOR:** [/data/raw/OSM\_LU01\_LU13\_LU15\_LU21\_VEG\_2022.csv]

*Information on exact methods for data extraction and more specific variable descriptions can be found on the ABMI data archive wall to wall landcover inventory data download website for Year 2010 OR in the relevant\_literature folder of this repository (LandcoverMapABMIGuide2010v1.0.pdf.pdf).*

- **Number of variables/columns:** 11
- **Number of observations/rows:** 3100

#### Variable List:

- **Site**, a factor with 155 levels where the first element abbreviation describes the landscape unit (i.e. LU13, LU15, LU21, or LU01) and the second element describes the camera site.
- **BUFF\_DIST**, a numeric measurement in meters ranging from 250 - 5000, of the buffer radius around the camera for which the proportion of associated human factors variables were calculated.
- **LC\_class20**, a numeric variable indicating the proportion of landcover class 20 within the buffer area. Landcover class 20 is defined as water.
- **LC\_class32**, a numeric variable indicating the proportion of landcover class 32 within the buffer area. Landcover class 33 is defined as rock and rubble.
- **LC\_class33**, a numeric variable indicating the proportion of landcover class 33 within the buffer area. Landcover class 33 is defined as exposed land.
- **LC\_class34**, a numeric variable indicating the proportion of landcover class 34 within the buffer area. Landcover class 34 is defined as developed land.

- **LC\_class50**, a numeric variable indicating the proportion of landcover class 50 within the buffer area. Landcover class 50 is defined as shrubland.
- **LC\_class110**, a numeric variable indicating the proportion of landcover class 110 within the buffer area. Landcover class 110 is defined as grassland.
- **LC\_class120**, a numeric variable indicating the proportion of landcover class 120 within the buffer area. Landcover class 120 is defined as agriculture.
- **LC\_class210**, a numeric variable indicating the proportion of landcover class 210 within the buffer area. Landcover class 210 is defined as coniferous forest.
- **LC\_class220**, a numeric variable indicating the proportion of landcover class 220 within the buffer area. Landcover class 220 is defined as broadleaf forest.
- **LC\_class230**, a numeric variable indicating the proportion of landcover class 230 within the buffer area. Landcover class 230 is defined as mixed forest.

## Processed Data

### DATA-SPECIFIC INFORMATION FOR: [OSM\_covariates\_2022.csv]

*Information on exact methods for data extraction and more specific variable descriptions can be found on the ABMI human footprints wall to wall data download website for Year 2021 OR in the relevant\_literature folder of this repository (HFI\_2021\_v1\_0\_Metadata\_Final.pdf).*

- **Number of variables/columns:** 119
- **Number of observations/rows:** 3100

#### Variable List:

This csv contains combined variables from the HFI and VEG raw data files ( ‘OSM\_LU01\_LU13\_LU15\_LU21\_HFI\_2022.csv’ & ‘OSM\_LU01\_LU13\_LU15\_LU21\_VEG\_2022.csv’). Variable descriptions are the same as above except the variables names are in lowercase type and the HFI variables no longer contain the prefix ‘FEATURE\_TY’.

### DATA-SPECIFIC INFORMATION FOR: [OSM\_covariates\_2021.csv]

*Information on exact methods for data extraction and more specific variable descriptions can be found on the ABMI human footprints wall to wall data download website for Year 2021 OR in the relevant\_literature folder of this repository (HFI\_2021\_v1\_0\_Metadata\_Final.pdf).*

- **Number of variables/columns:** 80
- **Number of observations/rows:** 1560

#### Variable List:

This csv contains the same data as OSM\_covariates\_2022.csv (variables from the HFI and VEG raw data files) but for the 2021 data which was formatted in a separate github repository.

## DATA-SPECIFIC INFORMATION FOR: [OSM\_covariates\_merged\_2021\_2022.csv]

- **Number of variables/columns:** 118
- **Number of observations/rows:** 4660

### Variable List:

This csv contains the combined data from OSM\_covariates\_2022.csv and OSM\_covariates\_2021.csv

## DATA-SPECIFIC INFORMATION FOR: [OSM\_covariates\_grouped\_2021\_2022.csv]

This csv contains data from OSM\_covariates\_merged\_2021\_2022.csv, where individual features have been grouped with similar features to simplify the number of potential variables and ensure enough data to use them in a modeling framework. Additionally, features not needed for the analysis or with too few data that could not be grouped, were removed. Grouped features are defined below, and the individual features that were included within each group are provided.

*Information on specific variable descriptions can be found on the ABMI human footprints wall to wall data download website for Year 2021 OR in the relevant\_literature folder of this repository (HFI\_2021\_v1\_0\_Metadata\_Final.pdf).*

- **Number of variables/columns:** 19
- **Number of observations/rows:** 4660

### Variable List:

- **array**, a factor with 6 levels where the first element abbreviation describes the project (e.g. OSM for Oil Sands Monitoring) and the second describes the landscape unit.
- **site**, a factor with where the first element abbreviation describes the landscape unit and the second element describes the camera site.
- **buff\_dist**, a numeric measurement in meters ranging from 250 - 5000, of the buffer radius around the camera for which the proportion of associated human factors variables were calculated.
- **harvest**, a numeric measurement representing the proportion of timber harvest within the buffer. Harvest is defined as, Areas where forestry operations have occurred (clear-cut, selective harvest, salvage logging, etc.) and includes, Harvest-area and Harvest-area-white-zone.
- **pipeline**, a numeric measurement representing the proportion of pipelines within the buffer. Pipelines are defined as, a line of underground and overground pipes, of substantial length and capacity, used for the conveyance of petrochemicals. The physical clearing that contains underground and above-ground high pressure pipelines and were not grouped with any other variables.
- **roads**, a numeric measurement representing the proportion of roads within the buffer. Roads are defined as, non-vegetated, impermeable surfaces used for motorized vehicle or aircraft transportation or access and includes, Airp-runway, Interchange-ramp, Ris-airp-runway, Ris-road, Road-gravel-1L, Road-gravel-2L, Road-paved-1L, Road-paved-2L, Road-paved-3L, Road-paved-4L, Road-paved-5L, Road-paved-6L, Road-paved-7L, Road-paved-div, Road-paved-undiv-1L, Road-paved-undiv-2L, Road-unclassified, Road-unimproved, Road-unpaved-1L, Road-unpaved-2L, Road-winter, and Transfer station.
- **seismic\_lines**, a numeric measurement representing the proportion of seismic lines within the buffer. Seismic lines are defined as, cleared corridors created during hydrocarbon exploration with a 3-meter buffer (6-meter total width), and were not grouped with any other variables.

- **seismic\_lines\_3D**, a numeric measurement representing the proportion of 3D seismic lines (also called low-impact seismic lines) within the buffer. 3D seismic lines are defined as, cleared corridors created during hydrocarbon exploration with a 1.5-meter buffer (3-meter total width), and were not grouped with any other variables.
- **trails**, a numeric measurement representing the proportion of trails within the buffer. Trails are defined as, cleared corridors surfaced with dirt or low vegetation for human/vehicle access, and include Trail, and Truck-trail.
- **transmission\_lines**, a numeric measurement representing the proportion of transmission lines within the buffer. Transmission lines are defined as, cleared corridors designated for the location of power transmission line infrastructure, and include Transmission-line and Ris-transmission-line.
- **veg\_edges**, a numeric measurement representing the proportion of vegetated edges within the buffer. Vegetated edges are defined as, disturbed vegetation alongside road edges, railway edges including ditches, and other industrial features, and include Vegetated-edge-railways, Vegetated-edge-roads, and Surrounding-veg.
- **wells**, a numeric measurement representing the proportion of wells within the buffer. Wells are defined as, ground cleared for an oil/gas well pad, and include Well-abandoned, Well-bitumen, Well-cased, Well-cleared-not-confirmed, Well-cleared-not-drilled, Well-gas, Well-oil, Well-other, and Well-unknown.
- **lc\_grassland**, a numeric measurement representing the proportion of grassland within the buffer. This was renamed from LC\_class110 to be more informative and was not grouped with any other variables.
- **lc\_coniferous**, a numeric measurement representing the proportion of coniferous forest within the buffer. This was renamed from LC\_class210 to be more informative and was not grouped with any other variables.
- **lc\_broadleaf**, a numeric measurement representing the proportion of broadleaf forest within the buffer. This was renamed from LC\_class220 to be more informative and was not grouped with any other variables.
- **lc\_mixed**, a numeric measurement representing the proportion of mixed forest within the buffer. This was renamed from LC\_class230 to be more informative and was not grouped with any other variables.
- **lc\_developed**, a numeric measurement representing the proportion of developed land within the buffer. This was renamed from LC\_class34 to be more informative and was not grouped with any other variables.
- **lc\_shrub**, a numeric measurement representing the proportion of shrubland within the buffer. This was renamed from LC\_class50 to be more informative and was not grouped with any other variables.
- **osm\_industrial**, a numeric measurement representing the proportion of various industrial features and clearings within the buffer. This includes, borrowpits (i.e., Borrowpit-dry, Borrowpit-wet, Borrowpits, Ris-borrowpits, Dugout, Lagoon, and Sump), clearings (i.e., Clearing-unknown, Clearing-wellpad-unconfirmed, Ris-clearing, Ris-clearing-unknown, and Runway), facilities (i.e., Camp-industrial, Facility-other, Facility-unknown, Mill, Misc-oil-gas-facility, Oil-gas-plant, Ris-camp-industrial, Ris-facility-operations, Ris-facility-unknown, Ris-plant, Ris-tank-farm, Ris-utilities, and Urban-industrial), and mines (i.e., Grvl-sand-pit, Mines-oilsands, Mines-pitlake, Open-pit-mine, Peat, Ris-drainage, Ris-mines-oilsands, Ris-oilsands-rms, Ris-overburden-dump, Ris-reclaim-ready, Ris-soil-salvaged, Ris-tailing-pond, Ris-waste, and Tailing-pond).

#### DATA-SPECIFIC INFORMATION FOR: [OSM\_covariates\_wide\_2022.csv]

- **Number of variables/columns:**



- Number of observations/rows:

Variable List:

DATA-SPECIFIC INFORMATION FOR: [OSM\_deployment\_2022.csv]

- Number of variables/columns:
- Number of observations/rows:

Variable List:

DATA-SPECIFIC INFORMATION FOR: [OSM\_glm\_anthro\_model\_sel\_data.csv]

- Number of variables/columns:
- Number of observations/rows:

Variable List:

DATA-SPECIFIC INFORMATION FOR: [OSM\_glm\_global\_model\_sel\_data.csv]

- Number of variables/columns:
- Number of observations/rows:

Variable List:

DATA-SPECIFIC INFORMATION FOR: [OSM\_glm\_land\_model\_sel\_data.csv]

- Number of variables/columns:
- Number of observations/rows:

Variable List:

DATA-SPECIFIC INFORMATION FOR: [OSM\_ind\_det\_2021.csv]

This csv contains independent detections (images taken at least 30-min apart) for all species and sites during 2021.

- Number of variables/columns: 8
- Number of observations/rows: 6696

Variable List:

DATA-SPECIFIC INFORMATION FOR: [OSM\_ind\_det\_2022.csv]

This csv contains independent detections (images taken at least 30-min apart) for all species and sites during 2022.

- Number of variables/columns: 8
- Number of observations/rows:

Variable List:

#### DATA-SPECIFIC INFORMATION FOR: [OSM\_proportional\_detections\_2021.csv]

This csv contains proportional monthly detection data for **selected species of interest** from all sites during 2021. There are columns for each species that total the number of months the species was detected, and columns that total the number of months the species was not detected - when these two columns are combined for each species that is used as our proportional binomial response variable.

- **Number of variables/columns:** 25
- **Number of observations/rows:** 78

##### Variable List:

- **site**, a factor with where the first element abbreviation describes the landscape unit and the second element describes the camera site.
- **black\_bear - caribou**, each of these columns is a numeric integer representing the number of months (when a camera was active for  $\geq 0.5$  month) that a species was detected (controlling for hibernation months for black bears by removing Dec-March).
- **absent\_black\_bear - absent\_caribou**, each of these columns is a numeric integer representing the number of months (when a camera was active for  $\geq 0.5$  month) that a species was **not** detected (controlling for hibernation months for black bears by removing Dec-March).

#### DATA-SPECIFIC INFORMATION FOR: [OSM\_proportional\_detections\_2022.csv]

This csv contains proportional monthly detection data for **selected species of interest** from all sites during 2022. There are columns for each species that total the number of months the species was detected, and columns that total the number of months the species was not detected - when these two columns are combined for each species that is used as our proportional binomial response variable.

- **Number of variables/columns:** 27
- **Number of observations/rows:** 154

##### Variable List:

- **site**, a factor with where the first element abbreviation describes the landscape unit and the second element describes the camera site.
- **black\_bear - caribou**, each of these columns is a numeric integer representing the number of months (when a camera was active for  $\geq 0.5$  month) that a species was detected (controlling for hibernation months for black bears by removing Dec-March).
- **absent\_black\_bear - absent\_caribou**, each of these columns is a numeric integer representing the number of months (when a camera was active for  $\geq 0.5$  month) that a species was **not** detected (controlling for hibernation months for black bears by removing Dec-March).

#### DATA-SPECIFIC INFORMATION FOR: [OSM\_timelapse\_2021.csv]

- **Number of variables/columns:**
- **Number of observations/rows:**

##### Variable List:

**DATA-SPECIFIC INFORMATION FOR: [OSM\_timelapse\_2021.csv]**

- **Number of variables/columns:**
- **Number of observations/rows:**

**Variable List:**

**DATA-SPECIFIC INFORMATION FOR: [OSM\_total\_detections\_2022.csv]**

This csv contains the total number of independent detections for all species and sites during 2022.

- **Number of variables/columns:** 40
- **Number of observations/rows:** 155

**Variable List:**

- **site**, a factor with where the first element abbreviation describes the landscape unit and the second element describes the camera site.
- **\*\*Blackbear - Canada goose\***, each of these columns is a numeric integer representing the number of independent detections (images taken at least 30-min apart) for each species at each site during the entire study period.