srfn_workingdirectory

output: pdf_document: default html_document: default — # OSM_2022-2023 This repository contains data, R scripts and associated outputs, and other materials necessary for the Swan River First Nation camera project and manuscript - TBD

GENERAL INFORMATION

Project Information
This research was conceived by members of Swan River First Nation (SRFN) as part of their priorities for understanding wildlife and landscape change within their traditional territory. Data collection and image tagging were carried out by SRFN staff, with non-Indigenous collaborators contributing expertise in data analysis and statistical interpretation. Elders and knowledge holders were consulted to integrate Indigenous perspectives on mammal ecology and habitat change throughout the study. The project reflects an Indigenous-led, co-produced approach grounded in respect and reciprocity. Details on the study design of this project can be found in the Materials and Methods section of the associated manuscript - TBD

Author Information:
Contact Information
Name: Marissa A. Dyck
Institution: University of Victoria
Address: 3800 Finnerty Rd, Victoria, BC V8P 5C2
Email: marissadyck@uvic.ca

Date of data collection:
2022-2023

Geographic location of data collection: The Swan River First Nation's traditional territory (Wildlife Management Unit 350), including Grizzly Ridge Traditional Use Preserve (TUP)

DATA & FILE OVERVIEW

File List:

Files in main folder

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

Files in data folder

/processed
This folder includes all processed data (raw data that has been cleaned and reformatting using the scripts in this repo) to run the anaylsis for this project and associated manuscript

- **prop_det_data.rds**; an r object that contains outputs from the top-performing GLMs for each species; this object is used to generate publication figures in script #4
- **srfn_covariates_grouped.csv**; contains the final covariate (Human Features aka HFI and land cover from ABMI data sources) data that were grouped into ecologically meaningful categories for analysis
- **srfn_covariates.csv**; contains the covariate (Human Features aka HFI and land cover from ABMI data sources) data that have been renamed but not yet grouped into ecologically relevant categories
- **srfn_deploy_fixed.csv**; contains the start and end date for each camera location in the project
- **srfn_ind_det.csv**; contains list of independent detections (detections of s pecies >30min apart) for all species and cameras in the project

- **srfn_presence_absence.csv**; contains binary 0s and 1s indicating presence or absences for each species and each camera site. **This file was not use for analysis but was generating for some internal reports*
- **srfn_proportional_detections.csv**; contains response variable data for analysis - proportional detections - the number of months a species was detected and the number of months a species was absent from each site.
- **srfn_total_detections.csv**; contains the summed number of independent detections for each species at each camera site. **This file was not use for analysis but was generating for some internal reports*

/raw
Raw data files with sensitive information and raw images have been omitted. These data are the property of Swan River First Nation (see the CARE Principles for Indigenous Data Governance: (https://www.gida-global.org/care). Please contact Swan River First Nation coauthors with further questions regarding access to these datasets.

Files in figures folder

This folder contains various plots generated in the scripts of this repository for the purposes of data visualization. For figures in the manuscript (TBD) see the figure legends for further description of figures.

- **bbear_plot.jpg**; figure S3
- **camera_operability.jpg**; shows duration of time each camera was deployed and functioning
- **corr_plot_250.png**; correlation plot to assess multicolinearity of predictor variables at the 250 m buffer radius
- **corr_plot_500.png**; correlation plot to assess multicolinearity of predictor variables at the 500 m buffer radius
- **corr_plot_4000.png**; correlation plot to assess multicolinearity of predictor variables at the 4000 m buffer radius
- **corr_plot_4250.png**; correlation plot to assess multicolinearity of predictor variables at the 4250 m buffer radius
- **corr_plot_5000.png**; correlation plot to assess multicolinearity of predictor variables at the 5000 m buffer radius
- **coyote_plot.jpg**; figure S5
- **deer_plot.jpg**; figure S9
- **figure_2_odds.jpg**; figure 2
- **lynx_plot.jpg**; figure S4
- **moose_plot.jpg**; figure S7
- **snowshoe_hare_plot.jpg**; figure S8
- **srfn_all_images.jpg**; figure S1
- **bbear_plot.jpg**; similar to figure S1, but for just the seven focal species used in the analysis
- **wolf_plot.jpg**; figure S6

Files in main folder

This folder contains relevant literature that may be useful with interpreting data etc.

- **ABMI_HFI_metadata_final_2021.pdf**; metadata file from Alberta Biodiversity Monitoring Institute (ABMI) with details and definitions for the human features data (see: https://abmi.ca/data-portal/46.html)

- **ABMI_landcover_metadata_final_2010.pdf**; metadata file from Alberta Biodiversity Monitoring Institute (ABMI) with details and definitions for the land cover data (see: https://abmi.ca/data-portal/63.html)

Files in scripts folder

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

- **1_ACME_SRFN_camera_script-2024-12-10**; .rmd file and knitted .html files for data cleaning and formatting
- **2_ACME_SRFN_landscape_covariate_exploration_script**; .rmd file and knitted .html files for further data cleaning, formatting, and exploration of the covariate (ABMI HFI and land cover) data
- **3_ACME_SRFN_glm_finalanalaysis_2022**; .rmd file and knitted .html/.pdf files for running th GLM analysis
- **4_ACME_SRFN_figures**; .rmd file and knitted .html/.pdf files for making the publication quality figures

PROCESSED DATA

DATA-SPECIFIC INFORMATION FOR: [srfn_covariates_grouped.csv]

- **Number of variables/columns:** 18
- **Number of observations/rows:** 1201

Variable List:

- **site_number**, numeric factor indicating the unique camera site number
- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number.
- **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 and land cover variables were calculated.
- **harvest**, a numeric variable indicating the proportion total area harvested for timber within the designated buffer around each camera
- **harvest_2000**, a numeric variable indicating the proportion of area harvested for timber from year 2000 onward within the designated buffer around each camera
- **harvest_pre2000**, a numeric variable indicating the proportion of area harvested for timber from prior to the year 2000 within the designated buffer around each camera
- **pipeline**, a numeric variable indicating the proportion of area defined as pipelines within the designated buffer around each camera
- **roads**, a numeric variable indicating the proportion of area defined as roads within the designated buffer around each camera
- **seismic_lines**, a numeric variable indicating the proportion of area defined as seismic lines within the designated buffer around each camera
- **veg_edges**, a numeric variable indicating the proportion of area defined as vegetated edges within the designated buffer around each camera
- **wells**, a numeric variable indicating the proportion of area defined as a well pad within the designated buffer around each camera
- **lc_agriculture**, a numeric variable indicating the proportion of area defined as agriculture land within the designated buffer around each camera

- **lc_broadleaf**, a numeric variable indicating the proportion of area defined as broadleaf forest within the designated buffer around each camera
- **lc_coniferous**, a numeric variable indicating the proportion of area defined as coniferous forest within the designated buffer around each camera
- **lc_developed**, a numeric variable indicating the proportion of area defined as developed land within the designated buffer around each camera
- **lc_grassland**, a numeric variable indicating the proportion of area defined as grassland within the designated buffer around each camera
- **lc_mixed**, a numeric variable indicating the proportion of area defined as mixed forest within the designated buffer around each camera
- **lc_shrub**, a numeric variable indicating the proportion of area defined as shrub land within the designated buffer around each camera

DATA-SPECIFIC INFORMATION FOR: [srfn_covariates.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:** 144
- **Number of observations/rows:** 1201

Variable List:

- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number.
- **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 and land cover variables were calculated.
- **airp_runway - well_unknown** all variables from 'airp_runway' to 'well_unknown' are a numeric variable indicating the proportion of area defined as the listed feature (e.g. airp_runway) within the designated buffer around each camera. Full definitions/descriptions of each HFI feature can be found in the metadata for this dataset from ABMI listed at the top of this data section
- **110**, a numeric variable indicating the proportion of area defined as grassland within the designated buffer around each camera
- **120**, a numeric variable indicating the proportion of area defined as agriculture land within the designated buffer around each camera
- **20**, a numeric variable indicating the proportion of area defined as water within the designated buffer around each camera
- **210**, a numeric variable indicating the proportion of area defined as coniferous forest within the designated buffer around each camera
- **220**, a numeric variable indicating the proportion of area defined as broadleaf forest within the designated buffer around each camera
- **230**, a numeric variable indicating the proportion of area defined as mixed forest within the designated buffer around each camera
- **33**, a numeric variable indicating the proportion of area defined as exposed land within the designated buffer around each camera
- **34**, a numeric variable indicating the proportion of area defined as developed land within the designated buffer around each camera

- **50**, a numeric variable indicating the proportion of area defined as shrub land within the designated buffer around each camera
- **1940 - 2021** all variables from '1940' to '2021' are a numeric variable indicating the proportion of area harvested for timber in the defined year within the designated buffer around each camera.

DATA-SPECIFIC INFORMATION FOR: [srfn_deploy_fixed.csv]

- **Number of variables/columns:** 6
- **Number of observations/rows:** 65

Variable List:

- **project_id**, factor indicating the project the data are associated with. All are SRFN as that is the only project data included in this repository
- **site_number**, numeric factor indicating the unique camera site number
- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number
- **array**, factor that describes the primary landcover type at the camera (LUD = developed, LUC = coniferous forest, LUS = shrub, LUBF = broadleaf forest, LUUK = unknown, LUG = grassland, etc.)
- **start_date**, date (m/d/yy) indicating the initial day in which the camera was deployed
- **end_date**, date (m/d/yy) indicating the last day in which the camera was deployed and subsequently collected

DATA-SPECIFIC INFORMATION FOR: [srfn_ind_det.csv]

- **Number of variables/columns:** 9
- **Number of observations/rows:** 8429

Variable List:

- **array**, factor that describes the primary landcover type at the camera (LUD = developed, LUC = coniferous forest, LUS = shrub, LUBF = broadleaf forest, LUUK = unknown, LUG = grassland, etc.)
- **site_number**, numeric factor indicating the unique camera site number
- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number
- **species**, factor indicating the common name of the species in the camera image
- **datetime**, date (yyyy/mm/dd) and time (24hr format) indicating the date and time the image was captured
- **month**, numeric variable indicating the month the image was captured (1 = January, 12 = December)
- **year**, numeric variable indicating the year the image was captured
- **timediff**, numeric variable used to calculate the difference between the previous detection of the same species on the same camera and the current detection, used to calculate independent detections (only those >30 min apart are retained in this file and used for analysis to ensure independence)
- **event_id**, factor indicating the detection event (only unique (independent) detections are kept [those >30 min apart]from the full capture history)

DATA-SPECIFIC INFORMATION FOR: [srfn_presence_absence.csv]

This file was not use for analysis but was generating for some internal reports

- **Number of variables/columns:** 33
- **Number of observations/rows:** 64

Variable List:

- **array**, factor that describes the primary landcover type at the camera (LUD = developed, LUC = coniferous forest, LUS = shrub, LUBF = broadleaf forest, LUUK = unknown, LUG = grassland, etc.)
- **site_number**, numeric factor indicating the unique camera site number
- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number
- **black_bear - owl** all variables remaining variables are binary variables (0s & 1s) indicating whether a species was detected (1s) or not detected (0s) at a given camera site across the course of the study

DATA-SPECIFIC INFORMATION FOR: [srfn_proportional_detections.csv]

- **Number of variables/columns:** 29
- **Number of observations/rows:** 64

Variable List:

- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number
- **black_bear - cougar** variables from 'black_bear' to 'cougar' indicate the number of months a species was detected while a camera was active for at least 15 days of the given month
- **absent_black_bear - absent_cougar** variables from 'black_bear' to 'cougar' indicate the number of months a species was NOT detected while a camera was active for at least 15 days of the given month

DATA-SPECIFIC INFORMATION FOR: [srfn_total_detections.csv]

This file was not use for analysis but was generating for some internal reports

- **Number of variables/columns:** 33
- **Number of observations/rows:** 64

Variable List:

- **array**, factor that describes the primary landcover type at the camera (LUD = developed, LUC = coniferous forest, LUS = shrub, LUBF = broadleaf forest, LUUK = unknown, LUG = grassland, etc.)
- **site_number**, numeric factor indicating the unique camera site number
- **site**, factor where the first element abbreviation describes the primary landcover type and the second is the site number
- **black_bear - owl** all variables remaining variables are numeric variables indicating the total number of independent detections (images taken >30 min apart) for a given species across the duration of the study at a given site