**Report Outline:**

1. **Data Processing**
   1. Collected (or were sent) photos.
   2. Sorted photos, deleting all blank photos or photos with squirrels, mice, birds, or researchers.
   3. Annotated images within WildID, identifying the species present.
   4. This data is exported as a csv, including all metadata associated with the photo (date, time).
   5. Our data was subsetted into different regions (North, South, Central).
   6. I handled the data for northern Costa Rica.
   7. Camera site and name were combined to form a unique camera station id. There are 116 unique stations within the northern dataset.
   8. There are 31 unique species within this dataset.

[1] "Bassaricyon gabbii" "Bassariscus sumichrasti" "Canis familiaris"

[4] "Canis latrans" "Coendou mexicanus" "Conepatus semistriatus"

[7] "Cuniculus paca" "Dasyprocta punctata" "Dasypus novemcitus"

[10] "Didelphis marsupialis" "Eira barbara" "Felis catus"

[13] "Galictis vittata" "Herpailurus yaguarondi" "Homo sapiens"

[16] "Leopardus pardalis" "Leopardus tigrinus oncilla" "Leopardus wiedii"

[19] "Lutra longicaudis" "Mazama temama" "Mustela frenata"

[22] "Nasua narica" "Odocoileus virginianus" "Panthera onca"

[25] "Pecari tajacu" "Philander opossum" "Procyon lotor"

[28] "Puma concolor" "Sylvilagus dicei" "Tamandua mexicana"

[31] "Tapirus bairdii"

* 1. Data was subset into necessary columns ("Camr\_Nm", "Species", "Date", "Time", "Latitud", "Longitd") and all NAs in the date column were removed.
  2. If there is more than one record at the same camera station on the same date, the first record is always taken, and the subsequent records are compared to determine if they are independent.
  3. If the subsequent records are greater than 24 hrs apart (from the first record and from each other) they are considered independent.
  4. My dataset has 7342 independent records (recs.all variable) → recs.all\_CR\_Dataset03\_North\_07Jun2022.csv
  5. Independent records were plotted with the camera station name on the x axis and date (month and year) on the y axis. The number of independent records for each camera during each month/year was included in the corresponding cell. This table was exported as a csv into an excel spreadsheet.
  6. We manually selected a time period with the most cameras running. Each camera in this window was included once, running <= 4 months, and not necessarily all running at the same time.
  7. To abide by the assumptions of a single season model, the maximum time period we can use for each camera is 4 months.
  8. Our window of selection for the northern dataset was 4 years (2012-2016) and included 65 cameras, 26 species, and 1322 independent records.

[1] "Conepatus semistriatus" "Cuniculus paca" "Dasyprocta punctata"

[4] "Didelphis marsupialis" "Nasua narica" "Pecari tajacu"

[7] "Philander opossum" "Puma concolor" "Dasypus novemcitus"

[10] "Leopardus pardalis" "Leopardus wiedii" "Panthera onca"

[13] "Sylvilagus dicei" "Tapirus bairdii" "Eira barbara"

[16] "Mazama temama" "Tamandua mexicana" "Canis latrans"

[19] "Leopardus tigrinus oncilla" "Canis familiaris" "Coendou mexicanus"

[22] "Procyon lotor" "Herpailurus yaguarondi" "Bassariscus sumichrasti"

[25] "Felis catus" "Odocoileus virginianus"

* 1. Based on the selected dates, a separate table was created to include the start month, start year, end month, and end year for each camera (tw).
  2. The main dataset was then filtered via the selected dates for each chosen camera (creating recs\_tw).
  3. The filtered dataset (recs\_tw) was used to create a table including each camera station, and the first and last date each camera was active for (active). This is different from our other table in that it includes the day, month, and year of the first and last active date during the selected period. Our previous table simply included the start and end month and year.
  4. This data was used to create a detection table (0’s and 1’s) for each camera on every single day from the earliest time a camera was set to the latest time a camera was stopped (ston) → ston\_CR\_Dataset03\_North\_07Jun2022.csv
  5. A function was created to collapse records by seven days (comm\_hist\_maker).
  6. Once records are collapsed (ch\_cr\_north), we can separate the new records by species, and export this subsetted data to a csv. These exported files include 1 if there is a record, 0 if there is no record but the camera was running, and NA if the camera was not running. Exported files of this type include:

“Paca\_Detection.csv”

“Peccary\_Detection.csv”

“Puma\_Detection.csv”

“Ocelot\_Detection.csv”

“Oncilla\_Detection.csv”

1. **Extracting Site Covariates**
   1. A file containing latitude and longitude information for each camera station is created (loc).
   2. Latitude and longitude data is converted in R to spatial coordinates (essentially telling R that these numbers are coordinates).
   3. Covariate .tif files are imported using raster function
   4. Data is then extracted from each covariate raster using coordinate data.
   5. Extracted data for each covariate is combined and exported as a .csv.
2. **Running Occupancy Models**
   * 1. We run our occupancy models utilizing the packaged “unmarked” in R.
     2. In order to utilize this model type we require:
        1. (obs) - detection data for species of choice: “Peccary\_Detection.csv” converted to a matrix
        2. (siteCovs) - site covariates extracted in the previous step: “Site\_Covariates\_North\_Updated5.csv”
        3. (obsCovs) - survey covariates, including effort, in list format: “Effort\_North.csv”
     3. Once the required files are imported and checked for the correct data type, we create and unmarkedFrameOccu (umf), and calculate the naive occupancy (does not account for imperfect detection).
     4. We next run a null model and run all univariate models. Univariate models are compared to the null model, and if better than the null model (even if not significant) are kept. The covariates from the univariate models which are better than the null model are then used to create multivariate models. This step reduces time spent creating unnecessary models.
     5. Once models are created, they are compared based on AIC in order to determine which models are best.
     6. The best models are summarized to determine significance, and if significance is found they are then plotted against occupancy or detection probability (depending on the type of covariate) in order to visualize the relationship.
3. **Code**
   1. Code to make the table of independent records (giant green table)
   2. Data processing code (code to make the green table and create record table)
   3. Extracting site covariate code
   4. Extracting survey covariate code (still need to figure this out)
   5. Running an occupancy model code
4. **Results**
   1. Camera site and name were combined to form a unique camera station id. There are 116 unique stations within the northern dataset.
   2. There are 31 unique species within this dataset.

[1] "Bassaricyon gabbii" "Bassariscus sumichrasti" "Canis familiaris"

[4] "Canis latrans" "Coendou mexicanus" "Conepatus semistriatus"

[7] "Cuniculus paca" "Dasyprocta punctata" "Dasypus novemcitus"

[10] "Didelphis marsupialis" "Eira barbara" "Felis catus"

[13] "Galictis vittata" "Herpailurus yaguarondi" "Homo sapiens"

[16] "Leopardus pardalis" "Leopardus tigrinus oncilla" "Leopardus wiedii"

[19] "Lutra longicaudis" "Mazama temama" "Mustela frenata"

[22] "Nasua narica" "Odocoileus virginianus" "Panthera onca"

[25] "Pecari tajacu" "Philander opossum" "Procyon lotor"

[28] "Puma concolor" "Sylvilagus dicei" "Tamandua mexicana"

[31] "Tapirus bairdii"

* 1. My dataset has 7342 independent records
  2. Our window of selection for the northern dataset was 4 years (2012-2016) and included 65 cameras, 26 species, and 1322 independent records.

[1] "Conepatus semistriatus" "Cuniculus paca" "Dasyprocta punctata"

[4] "Didelphis marsupialis" "Nasua narica" "Pecari tajacu"

[7] "Philander opossum" "Puma concolor" "Dasypus novemcitus"

[10] "Leopardus pardalis" "Leopardus wiedii" "Panthera onca"

[13] "Sylvilagus dicei" "Tapirus bairdii" "Eira barbara"

[16] "Mazama temama" "Tamandua mexicana" "Canis latrans"

[19] "Leopardus tigrinus oncilla" "Canis familiaris" "Coendou mexicanus"

[22] "Procyon lotor" "Herpailurus yaguarondi" "Bassariscus sumichrasti"

[25] "Felis catus" "Odocoileus virginianus"

* 1. Paca
     1. 11 total Paca observations across 5 total sites, with a maximum of 5 observations at one camera station.

SAVEGRE VALLEY 2

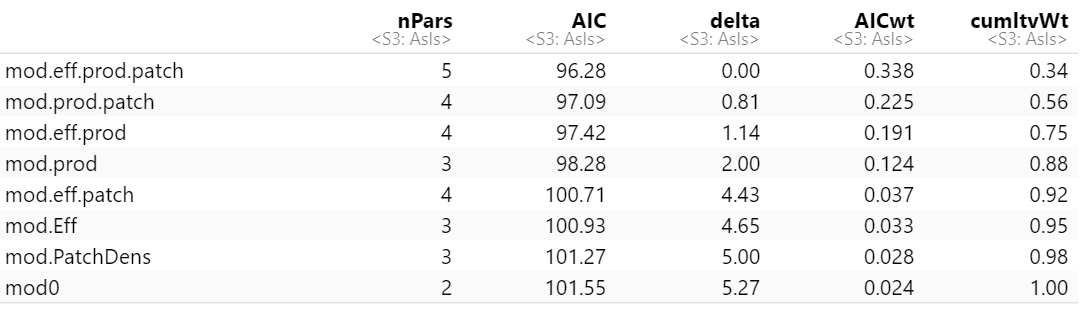
SAVEGRE VALLEY 6

SAVEGRE VALLEY 8

SAVEGRE VALLEY 10

SAVEGRE VALLEY 19

* + 1. Naive Occupancy (5/65): 0.07692308
    2. Best models:



* 1. Peccary
     1. There were 5 total peccary observations across 5 cameras with a maximum of 1 observation per camera.

COPAL EC13-14

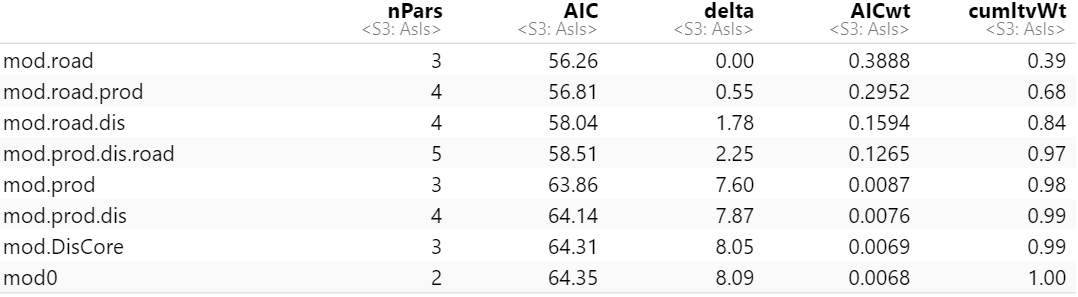
COPAL EC301-302

MARTA LM1-2

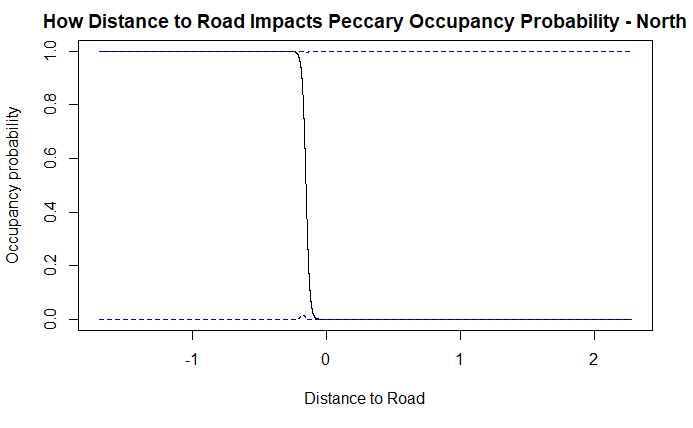
SAVEGRE VALLEY DANTICA29

SAVEGRE VALLEY VUELTAS26

* + 1. Best Models:



* + 1. Naive Occupancy (5/65) = 0.07692308
    2. Plotted Model (not great)



* 1. Puma
     1. 8 total puma observations across 8 camera stations, with a maximum of one puma observation at each camera station.

COPAL EC301-302

SAVEGRE VALLEY 4

SAVEGRE VALLEY 8

TAPANTI T1-2

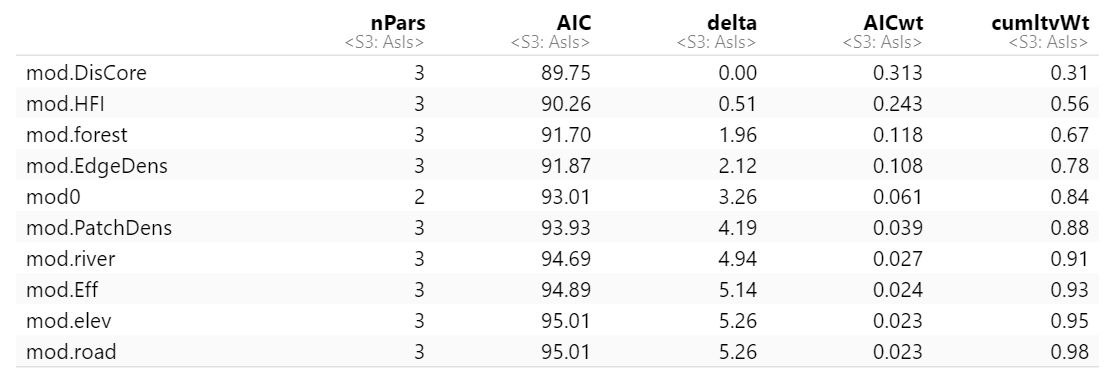
TAPANTI T07-08

TAPANTI T15-16

TAPANTI T21-22

TAPANTI T105

* + 1. Naive occupancy (8/65) = 0.1230769
    2. Best Models:



* 1. Ocelot:
     1. 8 total ocelot observations across 5 camera stations, with a maximum of two observations at one camera station.

COPAL EC301-302

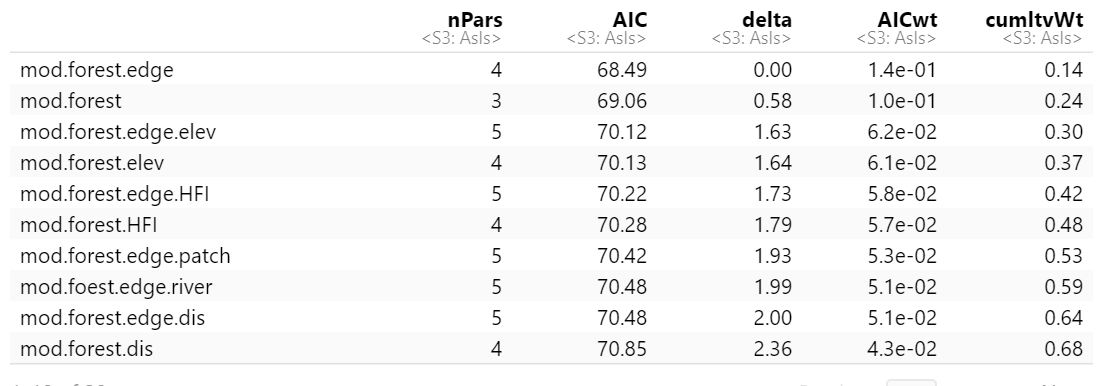
MARTA LM1-2

MARTA LM3

MARTA LM6-7

SAVEGRE VALLEY 1

* + 1. Naive Occupancy (5 sites observed/ 65 total sites): 0.07692308
    2. Best Models:



* + 1. Occupancy probability for top models:

