**Chapter 1:** Colony size in relation to land use change in the Great Plains for an expanding aerial insectivore

Question: Are cliff swallows choosing colony location & size based on landscape characteristics such as land use diversity, habitat patchiness, temperature, & drought?

* Have cliff swallows breeding grounds changed at a local level (compared with global and national trends)?
* What variation exists between foraging/colony sites?

 Key terms: land use changes, population expansion, foraging, sociality, habitat choice, landscape ecology, Great Plains

**Abstract**:

Changes in the Great Plains due to population expansion, climate change, and agricultural focus especially in the Western Corn Belt region. My aim was to determine the land usage that drives expansion of cliff swallows within western Nebraska. I used NLCD to construct maps of land use 2001 to 2019 in areas within a 1-km foraging range of each colony location. I determined how land usage in the area has influenced cliff swallow occurrence and abundance in agricultural habitats using random forest models (RFM). Simpsons Diversity index correlated with larger colony sizes. RFMs indicated that presence of open water, XX, and XX, were important features influencing cliff swallow occurrence and abundance within study area of Nebraska. Cliff swallows likely rely on open waters for foraging – with historical usage perhaps directing the present need today.

**Introduction**

Urbanized environments have widespread effects on behavior, reproduction, and other aspects of organisms’ ecology which are normally attributed to land-use change and increasing levels of pollution. Urbanization (=patchiness)– and all that it entails – changes the natural landscape through destruction and/or disruption of the natural environment to better fit humans needs. Urbanization drastically alters the surrounding natural ecosystem.As the world become more connected, understanding how human-mediated dispersal can affect organisms, especially those with health implications becomes more important. However, cities are complicated. None have the same population size, land area, geographic location, or age.

The expansion of urban areas has resulted in land use changes leading to habitat loss and species homogenization (McKinney 2006; Shochat et al. 2006) and behavioral changes (Lowry et al. 2013).

but with expansion usually comes land use change. Indeed, there has been a shift in land use within NE and other western plain states. Although it has been well studied how urban areas impact organisms based on land use changes, increased temperatures, and pollution, there is a lack of long-term studies on the ecological impacts of transportation methods on impact of species who uses them to expand their population in a tim when most aerial insectivores are declining (--) farmland associated birds declining(---).

The loss of native prairies to rangeland and pastureland as the initial shift, and now more recently the shift from rangeland to agricultural land is leading to changes in the ecosystem that we don’t fully understand.

However the change that occurred with the rise of industrial farms and monocultures like corn and wheat should not be underestimated. Water usage for corn, soy and sorghum are enough to shift seasonal climates and precipitation. In addition, the infrastructure needed for transportation of people and goods destroy and fragment existing landscapes while also creating new ecological niches to exploit for invasive species. The shift reduces biodiversity and insect populations.

Cliff swallow colonies can serve as a simplified model for understanding not just parasite dispersal, but other organisms that spread through human-mediated transport. Cliff swallow colonies can serve as a simplified model for understanding not just parasite dispersal, but other organisms that spread through human-mediated transport. If major differences exist between populations, some factors must be driving these shifts.

The Red Queen Hypothesis proposes that organisms are in an “evolutionary arms race” with one another, that is as one species evolves the other species it interacts with, especially in competitive or parasitic relationships, will also evolve to maintain the initial dynamic. The cliff swallow is a great system to explore the red queen hypothesis as cliff swallows have multiple nest and feather parasites, clear morphological changes due to selective pressures, extensive and expanding breeding population in western US, specialized on a declining food source – aerial insects. How has the cliff swallow managed to not only persist, but expand - one of the few bird species doing so in the US? One potential reason for their expansion might be due to their evolutionary potential (ability to change due to genetic lottery/adaptability). Cliff swallows historically nest on cliffs and steep clines in the natural environment. The earliest record of them in NE comes from year 1---. Since then, they have expanded to nest on anthropogenic structures as they become available (cite nests on mission). As railways and roads expanded, new structures became important nesting structures. Today, the largest colonies are on anthropogenic structures. Cliff swallows living on traditional cliffs remain consistent in size across years but must rebuild nests lost to weathering more frequently than on anthropogenic structures.

The frequent rebuilding may seem costly, but it may hold an advantage as new nests have reduced parasites that can be detrimental to the birds, especially juveniles. Like the mud nests on bridges and culverts, swallow bugs can hold on for years.

In a beautiful example of the Red-Queen Hypothesis, as nests became more permanent, swallow bugs survived in higher numbers overwinter which meant that cliff swallows had increased parasite loads in nests that didn’t fall. A comparison of cliff swallows from fumigated and non-fumigated colonies in the 80s demonstrated the clear detriment over-wintering bugs have on cliff swallows. Repeated more recently, and the difference between nestlings was less evident. Cliff swallows now show less of an effect, and with more roads than ever before, it makes sense that we are seeing an expansion in cliff swallow populations.

Although other parasites exist in cliff swallow colonies, the reduction of harm the swallow bug may allow for the expansion into thousands of birds at some colonies. But we still have these differences in size that may be less dependent on parasites in the past. If that’s the case, we can hopefully more easily identify other reasons for cliff swallow colony size.

We can increase our understanding on host-parasite dispersal by looking at

**HISTORY OF LAND USE IN STUDY AREA**

URBANIZATION

NATIVE TRIBES IN AREA -- ACCOUNTS OF CLIFF SWALLOWS AND/OR BUGS?

Early records of cliff swallows date back to 18XX with exploration party-man XX. irst recorded in the state in 1877 by ornithologist Robert Ridgway. In his report, Ridgway described seeing cliff swallows in the Platte River Valley in central NebraskaSaw cliff swallows nesting on cliff banks around XX - within XX from study-area. Quickly following exploration of western US, people began migrating westward. Major trails that went through Nebraska included the Oregon Trial, the Mormon Pioneer Trail, and XX transported on avg XX, XX, XX people respectively (Figure XX - trails w/timelines - lines weight based on number of people and time of year.). Following trails, roads and bridges were built and by 1895 legislation was passed to fund a state highway system. In 1913 - bond for construction of highways, 1920s included Lincolnd Highway. It sits within the study area and was very busy and transcontinental.

In 1956, congress passed funding for the construction of the Interstate Highway System which soon lead to the construction of I-80 and I-29.

(Figure XX - overlay of known colony sites with the trains and roads of different times)

Transportation of people and goods increased with urbanization as well. The history of railroads in Nebraska dates back to 1869 with the construction of of the Union Pacific Railroad from Omaha, Nebraska to Sacramento, California. Other railroads followed and during the peak of train in 19XX’s density of X railroads. Disruption of natural ecosstjmes are amuch larger scale began with the trails and railways of early united states. Although railways continue to be used for transportation of goods, particularly agricultural products, roadways quickly reclaimed status of habitat destruction and patchiness.

--

Effects of railroads and roadways: A brief history of

Habitat fragmentation, urbanization, road mortality, invasive speccies, new dispersal methods and barriers

--

Movement within and between habitats is an important process for finding resources, avoiding competition and predation, and finding mates. When movement is limited, it can reduce gene flow between populations affecting the spatial population dynamics of the species. As such, identifying landscape barriers allows us to understand the movement patterns of the organisms in the area. However, identifying barriers can be difficult as movement patterns can be difficult to track, especially for highly mobile individuals. Additionally, the effects of landscape barriers depend on the species and their dispersal capabilities. For example, one study found that fragmented landscapes can have different effects dependent on species (Bélisle and St. Clair 2002).

For migratory birds that undergo long-distance dispersal, minimizing energy expenditure is paramount. In addition to flight-distance, weather, time of day, and wind patterns have been shown to affect the energy expenditure of birds (Liechti 2006). As such, flight paths based solely on distance may not always be the most cost prohibitive route. For instance, some migratory birds follow Lake Erie’s coastline rather than flying directly over the lake (Gesicki et al. 2019) which may be more energetically favorable than a straight-line path, especially during weather events (Diehl et al.. 2014). Unfortunately, there is little research on the effects of landscape features on bird flight in a non-migratory context (Pigot et al. 2010).

Transient thing about energy here

Previous work has shown that large open bodies of water can be costly to traverse for migrating birds, but few studies have examined their effect on non-migrating birds. It can be difficult to determine for non-migratory birds. Satellite imagery is typically used to follow migration patterns and identify potential barriers, but individual birds moving between populations would not be visible using this method. Instead, landuse data can be used to understand foraginging data and reveal facilitation of habitat choice.

 One way that I am going to try to answer this is by looking at the size of cliff swallow colonies and how it relates to land use. What changes are A in these populations that allow them to adapt to these changes? How are these changes actually impact an organism? Very lucky I have a population with colony site and usage data for over 40 years and one that also has clear adaptations.

As cliff swallows shifted away from nesting on ancestral cliffs to artificial structures in the 1980s, they underwent population expansion that is still occurring today. Throughout this time, Discuss Where has the roadkill gone; brain size morphology; beak hook changes the negative impact of swallow bugs on their hosts’ survival has varied in intensity, with less costs to cliff swallow’s reproduction in more recent years and at larger nesting sites.

Land use changes in the great plains since colonization has resulted in a patchy, highly disturbed, and remnant natural landscapes. The high ranching and eventually crop value of the area in western Nebraska of our study area has changed significantly over the years (--), and fit with larger gobal land use patterns (--) as well.

Directly gathering information on food availability can be difficult. Previous attempt at identifying food species for cliff swallows have proven difficult measuring further down than family, although general patterns were clear. Frequency of foraging as well as amount foraged is known, but how does it interact with the landscape within a foraging area? Studying land use within foraging range scale can allow for indirect knowledge about habitats important for prey abundance and social foraging. Swallows are aerial insectivores, catching insects in flight as food source. Swallows can be social or non-social foragers. Swallows that do forage socially are constrained by morphology. Aerial foraging and reproduction, especially in densely populated areas, are confined by morphological traits. Swallows with larger tails and shorter wings lead to quicker turns and perhaps more insects. Cliff Swallows have shown patterns of shorter wings, bills, and --- within more vs less social cliff swallows. They can range from 2 to over 3000 nests with smaller colonies showing trends similar to those of swallows at large.

What does this mean in terms of current and future morphology? How do birds expand based on these constraints? Finally allow to be fit to social living or is it that they are responding to foraging declines based on general urbanization and climate changes? from maneuverability and

As social foraging is a main benefit of living in larger groups (xx) and as links between foraging success and reproductive success (--), increased colony size (), likelihood of return (), brain size () – foraging efficiency cannot be understated in its importance for morphology (), so why not then colony size?? How are the foraging needs for a large aerial insectivore population, especially one that is expanding, being met? Can colony size be explained by land use variables? And can information from this be used to impy anything about foraging? Foraging is hard to approximate as insect populations are hard to identify and typically ephemeral in nature. Swallows rely on social foraging to identify their grouping insects and are typically located high in the air on wind columns or above trees or in open clearings. Some sampling has been done in the past ---, identifying down to family and species in some cases. These insects are associated with XX landscapes, and fly during XX conditions. However, knowing this information for all insect populations the cliff swallows fed on is difficult and time consuming. For a proxy of information about what are important land use for insects iwhting the great plains that feed thousands of birds and other orgnansims a year (xx). Perhaps colonies with much larger populations, who must find more insects within 1-km foraging range, tend to appear in areas with certain land use attributes.

Western NE from earliest records ~1860s of nesting alongcliffs to nesting under culverts and bridges in 2019. What changes have occurred over that time? Any records of changes in cliff swallows during that time?

With Winkler et al’s etimate that land use change has affected almost a third (32%) of the global land area in just six decades (1960-2019), and that they observed phases of accelerating (~1960–2005) and decelerating (2006–2019) land use change explained by the effects of global trade on agricultural production (2021). I want to examine what land use changes occurred at a local scale in an agricultural region of the US and how they affect an aerial insectivore that breeds in the corn belt.

Our sampling sites are primarily located in an agricultural areas, but were variable in the agricultural usage.. Because of evidence that land usage effects insect populations differently ( main food source for these aerial insectivores), taxonomic richness and diversity (general predictor of insect health), and the potential barriers some of these areas can be (evergreen forests - example of the Capistrono sparrows) will be examined in context of land use and colony size to understand changes in agricultural area.

Will i be able to make a connection between changes in cliff swallow population and a shift in colony size such that at more stable LU sites, the colony size will be more stable? Or will certain land use types be the main correlate with colony size.

My aim in the study was to assess the colony size and abundance of cliff swallows in an area of intense agriculture area of southwestern Nebraska in relation to changing agricultural land use, the structure of nesting structure, and historical nesting locations of cliff swallows. In addition, I examined effect of climatic variables in relation to temperature and precipitation for each year analysis. I hypothesized that cliff swallows would be more abundant in 1-km foraging ranges with higher land use diversity (provides more patches; presence of open water and grasslands will support larger/stable populations, areas with increasing tree presence will show decrease in population. Additionally, substrate will correspond with increased colony size simply because ancestral cliffs could not support such large colonies. Cliff swallows examine colonies before selecting next years (==) and also rely on familiarity (--), personal experience (--), and certain landscape traits (water). Transient birds move between colonies during the breeding season, especially during the early months. They are usually young adult birds or those who migrated late in the season and did not have enough stored energy to breed in a given year (Brown and Brown 2004a). Because they frequently move between colonies, potentially scoping out the quality of these sites, (Brown and Brown 2004a). Transients typically have no active nests. Because of this, they may have more energy available to flight and foraging which would reduce the cost of traveling over certain landscape features. Hence, the presence and number of cliff swallows at a colony may partially reflect habitat choice by individuals within the colony. What can we say about the choice that they are making? Habitats are op Are there other landscape traits we can identify as useful for habitat choice? Does habitat choice even represent foraging vice versa?

Cliff swallows are associated with farmlands (--), (BBS). Reasons for this may be from precipitation from irrigation or livestock water/wells is used for nest building (--); insects increase with cattle and crop attraction of insects, including mosquitoes, . In study areas with barren land, large tree cover, and no water will have lower numbers of cliff swallow nests.

**Materials and methods**

**Study site and sample collection**

The cliff swallow is a migratory, highly colonial passerine found throughout the Great Plains and westward in North America (Brown et al. 2017). Cliff swallows began to arrive in the study area in late April with most arriving by May and early June before departing by late July.They nest closely together in gourd-shaped mud nests attached to cliff faces and rocky outcrops, underneath bridges, and inside culverts (Moore et al. 2007, Brown et al. 2013).

The cliff swallow (Petrochelidon pyrrhonota) is a migratory, highly colonial passerine found throughout the Great Plains and westward in North America (Brown et al. 2017). They nest closely together in gourd-shaped mud nests attached to cliff faces and rocky outcrops, underneath bridges, and inside culverts (Moore et al. 2007, Brown et al. 2013). The colony sites can range in size from 1 to 6000 nests with most of the colony sites located on buildings, bridges, and culverts (Brown et al. 2017). Colonies are typically separated by a distance of 1-10km, but sites have been found up to 20 km away (Brown and Brown 2005). Swallow bugs disperse to other nest sites colony sites by attaching to the legs and feet of cliff swallows (Brown and Brown 2004b), allowing for long distance dispersal. However, they are limited to dispersing during the summer months when the colony sites are active.

Population information

The Breeding Bird Survey (BBS) was used as an index for population of cliff swallows and avian scavengers. The BBS is conducted yearly along designated routes across the United States to evaluate the status and trends of North American bird with each survey route consisting of 50, 3-minute point counts (Pardieck et al. 2019). The Paxton route ran through the study area beginning in 1967 (Pardieck et al. 2019). The BBS recorded any birds within sight or sound of the roads along the route. This data on cliff swallows was used for inferring their use of roadside habitats and for the usage of roadside nesting sites over time. The annual population size of cliff swallows was estimated by summing all the nests in the active colonies each year (Brown et al. 2013).

**STUDY LOCATION**

The study site is a 200 x 50 km area centered at the Cedar Point Biological Station (41°12.591’ N, 101° 38.969’ W) in western Nebraska and included Keith, Garden, Lincoln, and Morill   counties. In the study area, there were more than 200 different colony sites that had been used by birds at least once during the course of the study (1982 to 2012). The colony sites can range in size from 1 to 6000 nests with most of the colony sites located on buildings, bridges, and culverts (Brown et al. 2017). However, not all colony sites are active every year. Colonies are typically separated by a distance of 1-10km, but sites have been found up to 20 km away (Brown and Brown 2005).

The study site was a 200 x 50 km area centered at the Cedar Point Biological Station (41°12.591’ N, 101° 38.969’ W) in western Nebraska and included Keith, Garden, Lincoln, and Morill counties (Figure 1). Located approximately in the middle of the study site is Lake McConaughy (41.24806, -101.82583). It is a man-made lake that is 22 miles long, 4 miles wide, and has a surface area of 30,000 acre. Our N colonies are located across 5 countries and around a xxx sq.km radius (Figure 1). Colony sizes ranged from 1 to 300X individuals per colony (mean +\_ SD = X +\_ x). All colonies were established for (HAVE CHARLES SPEAK TO COLONY SELECTION)

The study site was a 200 x 50 km area centered at the Cedar Point Biological Station (41°12.591’ N, 101° 38.969’ W) in western Nebraska and included Keith, Garden, Lincoln, and Morill counties (Figure 1). Located approximately in the middle of the study site is Lake McConaughy (41.24806, -101.82583). It is a man-made lake that is 22 miles long, 4 miles wide, and has a surface area of 30,000 acres.

Land-use Changes in the Western Great Plains of NE

**Substrate**

The populations that contained ancestral nesting cliff were pulled fromt eh analysis and used separately to examine if differences exist between land usages around ancestral sites, or if they are similar in land use acrpss locations. Cliff swallows

**Foraging Area**

Quantified the total area of different land use in a radius of 1 km around the center GPS coordinate around the colony because cliff swallows forage within a 1-km radius around the nesting location (--). The majority of feeding comes from within this area except for in bad weather conditions which can be further out as they seek active insects. The land use categories were generated using NLCD categories with a raster method.

within a colony location. See --- for more details on size methods.

Table 1: Categories of Land use and relative extent within a 1-km radius, corresponding to the foraging range of the cliff swallows of all colony sites in the study

A picture containing text, screenshot, number, font

Description automatically generated

without long-term samples is difficult to compare diets from larger colonies in areas with higher land-use diversity, does size depend on yearstable land use? re any land use types that explain the size of the colony? Is that even related to foraging?

Elevation was obtained as a Digital Terrain Model (DTM) at X m resolution. Aspect and slope were calculated from the DTM. Average monthly temperature for April-July was obtained at ----, PDSI – Palmer Drought Severity Index was used for drought and soil dryness measures.

Changes between land use categories between year for county comes from the NLCD 2019 change index

**Land-Use Analysis**

Locations for our sampled colony sites were plotted in a GIS using ARCGIS 10.5. These locations were then overlaid with land use data from the Nebraska’s Department of Transportation. These locations were overlaid with 10-m digital elevation models (DEM) from the USGS. Previous studies indicate that colonies with share... GPS used to identify the location of the middle of each colony. A 1-km buffer is around each gps was defined as it represents the foraging range for each colony. (Because we are focusing on insects & land use). The average distance between colony site in this study was 33. km  (ranging from --- to --- km).

Previous research has suggested that the lake may be acting as a barrier between colony sites (Moore 2016). Therefore, we estimated two pairwise measures of distance between colony locations: a straight-line Euclidean distance, and a least-cost path distance. The least-cost path distance was created based on the fact that open waters can be costly for birds to fly over (Figure 2). A cost surface was then created that assigned a high cost to land use cells that contained open water. All other land use types were assigned no cost. Because open water was assigned a high cost, the value of the least-cost path is equivalent to the distance between colony sites when Lake McConaughy is avoided. Patterns of population structure can be revealed through analysis of allele frequencies using Wright’s fixation index values FST (Wright 1978). FST is the measures the level of population genetic differentiation at single loci, which reflects the proportion of allelic variation (Holsinger and Weir 2009). Because I lack the dataset to actually find the FST for each colony site, I used values that would indicate that the colonies on the east side of the lake are not genetically distinct from the colonies on the west side of the lake. Finally, to determine if geographic distances affect genetic distances between swallow bug colonies, a mantel test was conducted in R. For the test, the least-cost path distance between colony sites were correlated with the FST values for the swallow bugs at each colony site.

All map manipulations were carried out in ArcGIS—(ESRI Redlands, CA, USA). Statistical analysis were carried out in R x.xx.1 (R Development Core Team 2011). Packages used ggplot, dyplyr, randomForest (Liaw & Wiener 2002), e121, ---, ----, ---.

**Random Forest Modeling**

A flexible Machine Learning Model – the Random Forest, can approximate a variety of non-linear functions especially with high-dimensional data. Random Forest have been suggested to perform better than traditional regression-based algorithms for variable selection and model predictions and is robust to small number of data points compared with number of predictors (Strobl et al. 2007). Random forest (RF) is a tree-based ensemble classifier that eliminates over-fitting through pruning an ensemble of finite regression trees (Breiman, 2001). At each iteration, new node is created based on lowest tree impurity. They are built through a process to minimize tree impurity. RF samples randomly from dataset and are used to build trees constructed from boostrap samples split at each node by the best predictor from a smaller, stochastic predictor variable pool. Overfitting is avoided by controlling the number of predictors randomly used at each split. Land use regression (LUR) models may fail to capture complex interactions and non-linear relationships between pollutant concentrations and land use variables.

Random forest has many advantages, but one fault is that highly correlated predictors will split the importance levels – so if mixed and evergreen forest are correlate then the low score of evergreen forest does not mean it is not a low importance variable, instead it has similar importance to forest variables. One way to deal with this is to remove or combine highly correlated predictors. For the land use variables within the study area (11), I created a correlation matrix in using the corrPlot package. Significance cut-off of 0.01 and correlation values > 0.6 correlation were combined together. All Developed land use variables were significantly correlatied with each other and were combined together into one category (Figure C).

*PCA loadings plot. Arrows indicate strength and direction of different loadings. This was used to ensure combined variables had similar directional effects. If corr > .6 and similar pulling then variables were combined and model was run and performance increased then combination/removal remained. Additionally, if types were similar such as “forest” or “Developed” then they were combined by simple addition.*

I used a random forest classification model (RFCM) and random forest regression model (RFRM) to evaluate the influence of environmental variables (land use, precipitation, temperature) use on cliff swallow occurrence. NLCD raster data from 2001, .. 2019, were converted to point data using the raster-to-point tool in ArcGIS ---. Atribute tables of the environmental data were joined with the land use points maps.

Random Forest was applied to map colony site distributions for cliff swallows to determine importance of land use variables on cliff swallow colony size. RFCM was performed to evaluate the effects of environmental variables on cliff swallow presence and absence. RFRM was performed to evaluate the effects of environmental variables on cliff swallow population size

For both random forest model constructions, --- trees were specified with tenfold cross-validation across 3 replicates using R package “randomForest”. I trained random forests using hold out cross validation method. The random forest parameter “mtry” was set to (p/3), common for regression tasks where p is equal to the number of features. The RF parameter “ntrees: was varied from 300 to 2000. The RF parameter “nodesize” was set to vary between 3-23.

Mds plot

Importance graphs produced from RFCM and RFRM were used to identify and rank most important land use variables influencing cliff swallow occurrence and abundance. The RFCM performance was evaluated using AUC of ROC. RFRM performance was evaluated with adjusted R^2. The predictor variables with the lowest important values were removed based on drop-column importance method. All the RFMs were implemented in the randomForest package (Liaw and Wiener 2002) in R version 4.2.xx (R Core Team 2021).

I determined the degree to which features in my model were independent or interdependent on one another in their effect on random forest predictions with Friedman’s H-statistic – interactivity between single pair or a single feature with all other input features.

I measured RF performance on validation/test data using RMSE.

Variable Importance of features with Mean Increase in MSE. Feature selection in RF calculates the difference of Out-OfBag (OOB) errors before and after feature permutation, which gives an indication of relative importance of input features (Breiman 2001).

I used two methods for variable importance in random forest is based on drop-column importance method by ----.

RF feature ranking was found to be more informative than graphs of average features for determining the most appropriate feature inputs for land-use classification.

Land use regression (LUR) models may fail to capture complex interactions and non-linear relationships between pollutant concentrations and land use variables.

Random forest has many advantages, but one fault is that highly correlated predictors will split the importance levels – so if mixed and evergreen forest are correlate then the low score of evergreen forest does not mean it is not a low importance variable, instead it has similar importance to forest variables. Pne way to deal with this is to remove or combine highly correlated predictors. For the land use variables within the study area (11), I created a correlation matrix in using the XX packages. Any variables with greater .6 correlation were combined with inputs from PCA loadings plot. Arrows indicate strength and direction of different loadings. This was used to ensure combined variables had similar directional effects. If corr > .6 and similar pulling then variables were combined and model was run and performance increased then combination/removal remained. Additionally, if types were similar such as “forest” or “Developed” then they were combined by simple addition.

**Land Use**

Locations for the colonies were plotted in a GIS with ARCGIS 9.0. These locations were overlaid with XX-m digital elevation models (DEM) from the US Geological Survey (USGS) and specifically the National Land-use Category Variables. Previous studies suggest a 1-km range is an important scale for foraging.

Land use changes in the great plains since colonization has resulted in a patchy, highly disturbed, and remnant natural landscapes. The high ranching and eventually crop value of the area in western Nebraska of our study area has changed significantly over the years (--), and fit with larger gobal land use patterns (--) as well. The more data the better and replicate sites with controlled variation is ideal environment but it can be simulated with multiple locations and variation within those sites of traits of interest. Our study site contains XXX colonies with up to 9 samples per colonies with the year data.

**Results**

The top five land use based on size wer XX, based on total colonies that contained these - maybe pie chart data?

Colony distributions are non-random based on Moran’s I - an indicator of spatial autocorrelation.

Does the land use around a colony affect size?

How much variation in land use has occurred?

How much land change is needed to shift colony site usage?

Is there an identifiable shift towards/based on water?

Random forest modeling

According to ranking of feature importance values based on RFCM with all the variables, ---, ---- had strong relationships with cliff swallow presence and absence (Figure 1a). Feature importance from RF

The relative importance of the features used with the RFRM with all the variables indicates that the abundance of cliff swallows was strongly influenced, in decreasing order of importance, by Substrate, Open\_Water, Mixed\_Forest, Developed, Cultivated\_Crops, Herbaceous, Deciduous\_Forest, Hay\_Pasture, Wetlands, Barren\_Land, Shrub\_Scrub, Evergreen\_Forest, and Year. According to the importance ranking, year and the climatic variables had the lowest importance in both models. A negative value is inidicative of ---. The best RFCM (AUC = 0.851) and RFRM (adj R2 = 0.527) included the variables river incision, distance to river, 1984 LULC classes, flood return intervals and soil drainage as the most important variables for predicting redcedar occurrence and abundance.

Parial dependency plots indicate that cliff swallow populations are greater on sites with increased land use diversity, open-water, and mixed forests. The partial dependency plot indicated that

Substrate refers to the nesting structure of the cliff swallows – the two most common types are bridges and culverts. A 3-D PCA with score loadings?? was colored based on substrate types. This clear grouping of two types – ancestral and bridges indicates the differences between areas surrounding these types. ---

randomForest(formula = Size ~ ., data = testing\_set, metric = "RMSE", importance = TRUE, mtry = 8, nodesize = 5, ntree = 1000, maxnodes = NULL)

Type of random forest: regression

A picture containing text, font, screenshot, diagram

Description automatically generated

Number of trees: 1000

No. of variables tried at each split: 8

Mean of squared residuals: 207775.6

% Var explained: 31.76

> rmse(testing\_set$Size, y\_predict)

[1] 212.8442

> cor(testing\_set$Size, y\_predict) ^ 2

[1] 0.9144253

fit\_rf2<- train(Size ~., training\_set, method = "rf", metric = "RMSE", tuneGrid = tuneGrid,

trControl = trControl, importance = TRUE, nodesize = 5, ntree = 1000, maxnodes = NULL)

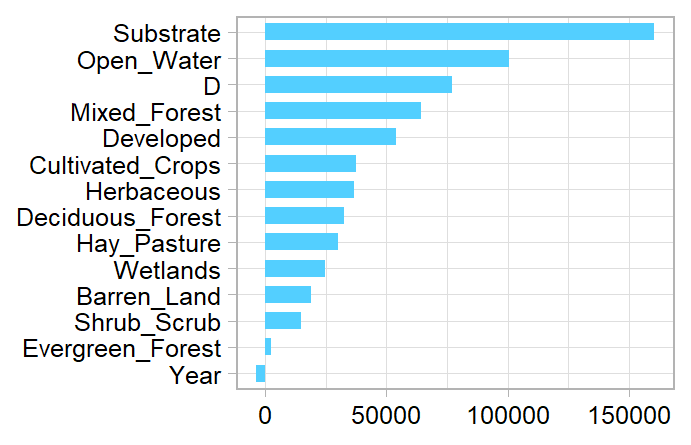


Figure 1: Out-of-bag permutation importance of features from random forest model. Explain land use types

A picture containing text, screenshot, diagram, font

Description automatically generated

Figure 1X: Variable Importance based on percent increase MSE. Year and Substrate

;lk;k

jjj

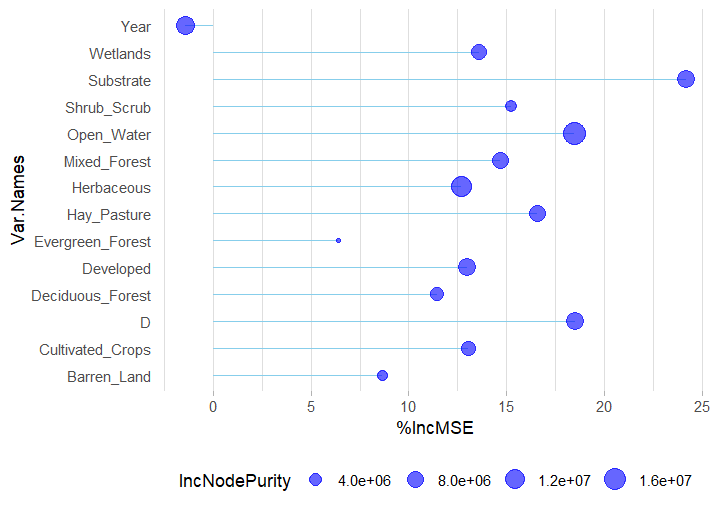


Figure 2: Sub

;lk;k

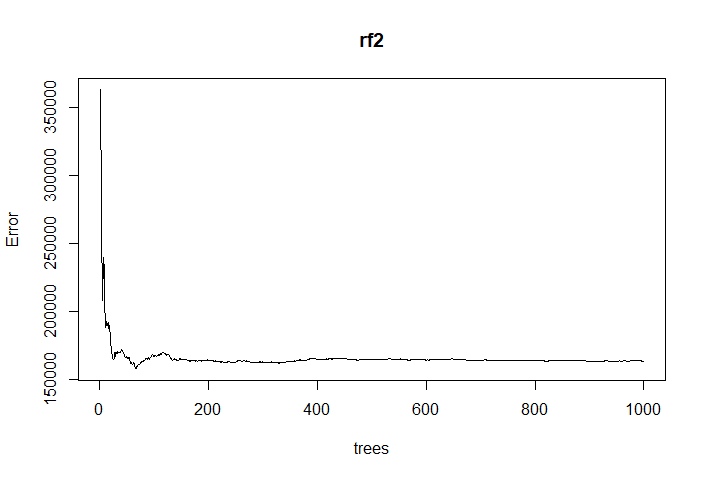


Figure 3: Sub

Results ^^^^^^