**Slide 1:**

Hello, my name is Tobin Haefele and my project is the Russian Olive Watch. Predicting the Spread of Russian Olive in Missoula County.

**Slide 2:**

A quick introduction to this project, I aimed to address two Key Objectives:

First includes assessment of the Current Distribution of Russian Olive across the county . Second is predicting the potential spread using habitat suitability modeling techniques to identify areas of interest.

*Why now?*

Russian olive was recently classified as an invasive species by the Missoula county ecology extension. So when I approached them with my initial idea of analyzing the spread of an invasive species, they pointed me in the direction of Russian Olive. This project hopes to fill any gaps in the knowledge currently out there.

*My Approach*

I used habitat suitability modeling as it is a common in the industry and considers important environmental factors when predicting the presence of Russian olive. This can then allow me to highlight priority areas that stand out for surveying, treatment, and management for the city.

**Slide 3:**

Russian olive was originally introduced in the United States during colonial times and grew to be a popular plant due to its value as an ornamental plant and as a windbreak.

It is currently illegal to distribute or sell in a number of states in the central and western US due to the rapid spread. Montana banned it in 2010, there is a heavy presence especially in Eastern Montana. There have been a number of studies on it’s spread along the rivers.

Ecologically Russian olive can grow in a variety of conditions. This leads to it displacing native plants such as cottonwood and the willow, which in turn disrupts the local ecosystem. There have also been studies that have shown changes in soil and hydrology conditions in some instances.

As I referred to earlier, Missoula county recently listed the Russian Olive in the noxious weeds list. This presented me an opportunity to apply my data science and analytics skills to a local issue that is important to me.

**Slide 4:**

Just a quick project overview:

This project has a fairly straightforward process including:  
data collection from various sources

Data processing and preparation for modeling

Model development

Finally generating the habitat suitability map

**Slide 5:**

A brief description of a random forest model and why I selected it for my project:

Random forest models build multiple decision trees using different data points. Then it combines these trees to make a final prediction either via majority vote or average prediction. This can help reduce overfitting and improve model accuracy.

Some other benefits include it’s ability to handle complex, non-linear data, and it’s ability to handle categorical and continues variables in a single dataset.

I also like that it allows me explore the impact the variables have on the outcome, which can make it more interpretable than other machine learning models. which can be somewhat of a black box as to why it came up with said result.

**Slide 6:**

Moving on to data collection, I had 3 main groups of data including:

Presence points where Russian Olive presence has been confirmed by field surveys from the Missoula county ecology extension.

Pseudo-absence points generating within Missoula county, where there is assumed to be no Russian olive presence.

And finally the environmental variables that will be the “independent variables” in this case.

These are broken down into 3 main categories of Land Cover, Climate, and Soil.

**Slide 7:**

The preprocessing steps I took included:  
Filtering the MTNHP points to retain only those in Missoula county with a spatial precision (accuracy) less than 800 meters.

Creating a presence-absence column to be the dependent variable.

Extracted the environmental variables for each point location

Converting any categorical variables such as land cover to a factor

Finally removing any unnecessary columns and filtering out any NA rows.

Next I had to address the issue of Spatial Autocorrelation before modeling.

Spatial Autocorrelation occurs because nearby points often share environmental traits which can bias the model results.

To address this, I used the BlockCV package in R to create spatially distinct folds. The size of the folds is chosen using a variogram which allows me to see the distance where the variance between points is at it’s peak. This helps prevent data leakage and improve the model accuracy.

**Slide 8:**

Looking at the model results, we achieved an Out of Box error rate around 4% indicating the model accurately classified around 96% of points during testing.

The area under the curve was around .976 which suggests the model did very well discriminating between suitable and unsuitable habitats.

Finally some of the top predictors or variables that contributed to the model includes Relative annual precipitation, frost free days, soil Ph was the highest, and land cover.

**Slide 9:**

Looking at the breakdown of the land cover level 2 variables, which is the highest detail one I could use in the model.

Conifer forest, montane grassland, and deciduous shrublands all had high input into the probability of presence with all other variables being equal. This is interesting as it strays a little from conventional wisdom that Russian Olive only occurs in Riparian and wetland zones. You can also see that developed/mining portion is negative which makes sense as well.

These can be important clues when combined with the other preditors like soil pH and frost-free days.

**Slide 10:**I used an interpolation technique called “kriging” to all the point predictions into a raster or layer surface. This can help visualize the conditions where the model predicts russian olive is most likely to occur.

Several interesting areas in this including a portion in east Missoula, lolo, and central Missoula. As you can see I also added that actual locations to compare with what the model is predicting.

These areas could be potential candidates for priority monitoring and further surveying.

**Slide 11:**

A brief review of management recommendations, many of these are already in place in fashion so much of this is review.

Targeted surveys in those locations of interest can help with confirming presence and prioritize follow up efforts.

Removal and treatment Russian olive in these critical areas. Based on the literature I’ve consumed for this project, these are just a basic recommendation.

Potentially replacing any trees removed with native species, can help restore the native habitat and support ecosystem recovery.

And finally having a long term tracking system of treated and replanted sites can help guide future efforts.

**Slide 12:**

A big item that could help is community engagement.

Raising awareness about russian olive among landowners and other property managers can help prevent new planting and encourage responsible land stewardship.

Promoting citizen science programs such as the montana native plant society as you see on the left. This can help engage the public and might produce early infestation detection in hard to reach areas or private lands.

Along with this, encouraging volunteering where appropriate can help reduce the workload on department staff for routine checks or basic removal. It will also boost the local investment in invasive species management.

Finally these measures can help save time and money for the department. Which in turn frees up resources restoration and removal efforts in highly impacted areas.

**Slide 13:**A Couple limitations of this model include:

Data uncertainty, some of the points may be outdated or misidentified.

When the pseudo-absence points are generated, it is assumed that russian olive is not there, which is not confirmed.

The sample size of around 800 presence points offers predictive value but is still relatively small.

Finally this model assumes stable environmental conditions which may or may not hold under future climate change or land use.

**Slide 14:**This project aims to provide a science-based decision support tool for Missoula County Ecology Extension and any other land management agencies.

Visualizing high risk areas allows for targeted surveying and treatment, helping prioritize efforts and reduce wasted resources. In turn this hopefully contributes to efficient conservation planning by identifying where to act and why.

In the future this model aims to be adaptable, allowing someone to test a new species with minimal changes to the modeling framework.

Ultimately this project aims to save time and money, allowing more focus on tasks like treatment and restoration.