Homework 11

This assignment uses data from the paper, Benson et al. 2012, looking at a hybrid zone between wolves and coyotes in Ontario. The authors sampled a number of wolves in these areas:
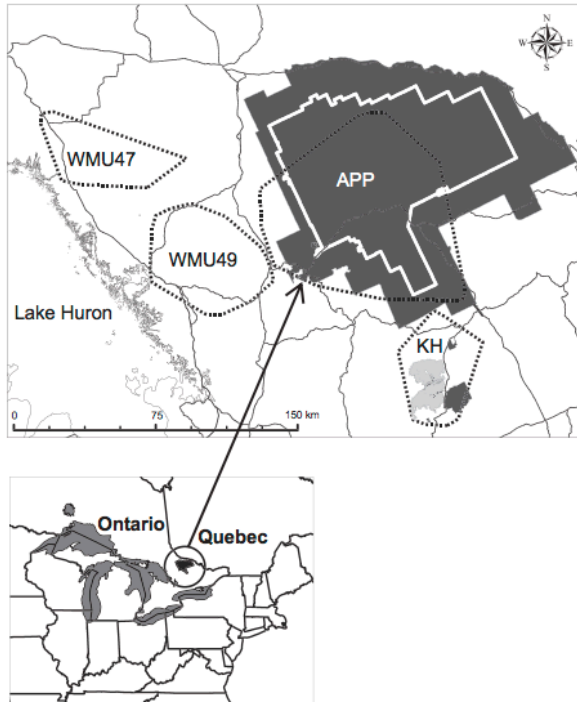

Fig. 1 The four study units (APP, WMU47, WMU49, KH) in central Ontario denoted by minimum convex polygons (dashed outlines) created using telemetry data from study animals. Dark grey shading represents areas where wolves and coyotes were protected from harvest, whereas light grey shading indicates trapping (but no hunting) was allowed. White polygon shows the APP boundary, and black lines represent major roads.

The authors performed genetic analyses to assign to each individual % ancestry from coyotes, eastern wolves, and grey wolves. The results look like this:
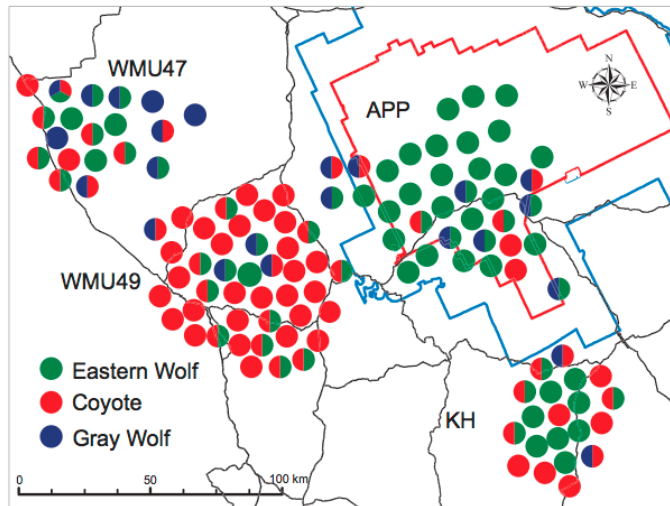
Fig. 5 Study area with resident individuals plotted (approximately) at home range centroids with pie charts showing genotypes based on individual assignment to genetic clusters using Structure and PCA. >1 colour in pie charts indicates admixture. Pie charts are simplified to show 100% ancestry for highly assigned individuals, and 50–50% or 33–33–33% for individuals admixed between 2 or 3 parental clusters, respectively. Also shown are major roads (black lines), APP boundary (red line) and harvest ban buffer area boundary (blue line).

The authors wanted to use this data to address several questions about the structure and causes of this hybrid zone:

1. Does the variation in % coyote ancestry look like a simple cline, or is the spatial structure more complicated?
2. Is variation in % coyote ancestry well-predicted by whether the individuals are mostly inside the Algonquin Provincial Park (APP), where harvesting wolves or coyotes is prohibited?
3. How well can spatial variation in % ancestry be predicted by environmental conditions?

We can address these questions using GAMs. The dataset "wolf_coyote_hybrids.csv" includes the following info:

- % coyote ancestry of each individual, logit-transformed for normality
- coordinates of the centroid of the home range of each individual (in east & north metric coordinates [units of meters])
- density of primary roads in the individual's home range
- density of secondary roads
- density of tertiary roads
- an index of deer abundance in the home range
- moose density in the home range
- study area (inside APP or outside APP)

To get at question 1, fit a GAM where % ancestry is the response and the predictor is a 2D smoother that uses the east and north coordinates. Plot the fitted smooth in a way that is interpretable to you. To what extent does spatial variation in % ancestry look like a smooth cline vs. a more complex pattern? Use gam.check to do some diagnostics. How do they look?

To get at question 2, we need a new variable that measures distance from APP. Let's say the center of APP is at the coordinates 1440000 east, 12160000 north. Make a new column where you calculate the euclidean distance between each individual and the center of APP.

Now make a GAM where % ancestry is a smooth function of distance from APP. Do you think this model supports the idea that the relative contribution of coyotes vs. wolves is primarily affected by this park? How does the % deviance explained compare between this model and the model where you fit a 2D smoother to the data? How does estimated degrees of freedom and AICc compare between the two models? How do you interpret these results?

Question 3 can be addressed by using the environmental predictors in a model. Coyotes are more tolerant of disturbed/human habitats, so predictors like road density (a proxy for human impact) might explain % ancestry. Deer are eaten by both coyotes and wolves, while moose are taken primarily by wolves. Make a GAM where % coyote ancestry is the response, and the predictors are the three road variables, plus deer, moose, and StudyArea. Use a smoother for each predictor (except StudyArea), to see whether it looks like the effect is linear or nonlinear. You may want to transform SecondaryRds, because it has a very skewed distribution. Which of the predictors explain significant variation? Plot the fitted effects, and interpret them in light of what I've already discussed about coyotes and wolves.

The authors hypothesized that the effect of tertiary roads (smaller country roads) might differ if the animals are inside or outside of APP, because those roads are the most important for hunting/trapping access, but the animals are protected inside of APP. Amend the model you just made to include an interaction between the smoother for tertiary roads and StudyArea. Does it look like the interaction is important? How does this model compare based on AICc? What do you think the interaction means? Consider that wolves and coyotes may be differentially susceptible to human effects.

For the models you've used for Question 3, we haven't explicitly accounted for spatial structure. That may lead to spatial autocorrelation of the data, based on the

spatial structure we looked at for Q1. Based on the code from the lecture notes, plot a bubble plot and semivariogram of the residuals of the model that includes the TertiaryRds*StudyArea interaction. Does it look the residuals are spatially autocorrelated? Now make a new model that adds in the 2D spatial smoother that you used previously. Compare the evidence for spatial autocorrelation between this model and the model without the smoother. Does accounting for autocorrelation with the 2D smoother change whether you think the other predictors are important?