# Introduction

* Objectives
  + What is current distribution of Oklahoma grassland songbirds?
    - Density estimates use PC and transect distance sampling
    - Compare estimates from transect and PC sampling (PC along roads, transects cross-country)
    - Species distribution model maps
  + What landcover (including crops, conservation easements?), vegetation (from 2014 transects only), and climatic variables predict the distributions of the study species?
    - Use presence/absence (from our surveys and from ebird) in species distribution models
  + How will distributions of selected species move with climate change and land use changes?
    - Distribution changes with predicted climate change
    - IF CAN FIND DATA: predicted landuse/crop cover changes

# Methods

## Study area

Brief description of Oklahoma vegetation and climate, number and types of species found in areas surveyed.

## Response data

* Survey methods
  + Point counts
  + Transects
    - lengths are not even. Some transects longer than others.
* eBird data
  + Find out how to download, what is format
  + Beware duplicates because some of survey data has already been entered
  + Maybe use as test data set to cross validate predictions of species distribution models from our surveys? Or incorporate and use k-fold after?

## Predictors

* Survey vegetation data from 2014 transects (none from 2013 or point counts in 2014?)
* Bioclim (get through R or from website)
* Data I need to find if exists:
  + Forecast changes in landuse in OK
    - <http://tethys.dges.ou.edu/main/?cat=12>
* Data I have downloaded
  + NRCS Conservation Easement Areas by State  
       Size: 0.40 megabytes (46 files).  Download compressed size: 0.19 megabytes (1 map).  
       <http://gws.ftw.nrcs.usda.gov/GWDL/3276698/easements_EASEAREA_ok_3276698_01.zip>  
     National Land Cover Dataset  by State  
       Size: 35.18 megabytes (7 files).  Download compressed size: 29.96 megabytes (1 map).  
       <http://gws.ftw.nrcs.usda.gov/GWDL/3276698/land_use_land_cover_NLCD_ok_3276698_02.zip>  
     Cropland Data Layer by State  
       Size: 235.53 megabytes (3 files).  Download compressed size: 235.57 megabytes (1 map).  
       <http://gws.ftw.nrcs.usda.gov/GWDL/3276698/land_use_land_cover_NASS_CDL_ok_3276698_03.zip>  
     Gridded Soil Survey Geographic (gSSURGO) by State  
       Size: 952.32 megabytes (4 files).  Download compressed size: 952.46 megabytes (1 map).  
       <http://gws.ftw.nrcs.usda.gov/GWDL/3273245/soils_GSSURGO_ok_3273245_01.zip>  
     Major Land Resource Areas by State  
       Size: 1.35 megabytes (46 files).  Download compressed size: 1.00 megabytes (1 map).  
       <http://gws.ftw.nrcs.usda.gov/GWDL/3276698/soils_MLRA_ok_3276698_05.zip>  
     Common Resource Areas by State  
       Size: 1.28 megabytes (45 files).  Download compressed size: 1.03 megabytes (1 map).  
       <http://gws.ftw.nrcs.usda.gov/GWDL/3276698/soils_CRA_ok_3276698_06.zip>

## Analyses

### Density estimations

* Using distance sampling, possibly including detectability from repeated surveys
  + Comparison of point count vs transect effectiveness if sample size large enough for each and geographical overlap sufficient. However, point counts go along road and transects usually walking off-road. Alternative: comparison of estimations from road pcs vs “off road” transects?

### Species distribution models

* Ensemble models
  + Compares species distribution models by weighting averages of each single model prediction “with weights assigned to each modelling technique based on its discriminatory power as measured by the area under the receiver-operated characteristic curve” (Oppel et al. 2012, seabird paper). Will combine niche and other species distribution models.
  + adaSTEM/STEM models
    - Ensemble models of decision trees, used with “bagged decision trees” (a type of classification tree) as base models trees in Fink et al paper
    - Unsure if can do with multiple types of models, like in Oppel paper? They had two regular models (linear and additive) and three machine learning but did not do spatiotemporal adaptive aspect. STEM is type of ensemble model with different bases, unsure if can incorporate multiple model types as bases.

# Results

# Discussion

# Notes to self

* Things I need to do meanwhile:
  + Talk to Todd about getting more detailed landcover if needed beyond what I have downloaded, and crop or land use predictions
  + Continue reading on machine learning <http://cs229.stanford.edu/materials.html>, review classification tree papers from multivariate class
  + Continue reading about distance sampling techniques (Buckland et al book)
  + Continue polishing up dataset
  + Acquire ebird data