Kernel Density Estimation Model (KDEM) Vignette Instructions

J.L. Tracy

28 June, 2018

This vignette provides steps for using example monarch roosting (migratory) occurrence data and a human population grid from Tracy et al. (submitted 2018) to generate a training set ensemble (TSE) of kernel density estimate models (KDEMs) for eastern monarch migration pathways.

1) Open a new empty ArcMap project and add the raster *pop10kmn3* (year 2000 human population density in a 3 km radius generalized to a 10 km resolution) (raster located in *KDMVignetteData* folder). This raster is in the North America Albers Equal Area Conic projection that is used for generating KDEs. Make sure the monarch roost data point shapefile *MonarchRoosts\_2002\_2006East.shp* is in the same folder as the raster.

2) Use Python script editor such as IDLE to open the ArcPython script *MonarchRoostDataProcessingBatch.py* (in the ArcPythonScript folder). Follow the instructions at the top of the script for setting the Geoprocessing Environments in the above ArcMap project so that the output coordinates, processing extent, snap raster and cell size set are set to match the raster *pop10kmn3*. Make sure the specified directories in the script match the above directory for data in step 1 on your computer. Open the ArcPython window in the ArcMap project, and paste in several lines at a time of commands to run from the ArcPython script *MonarchRoostDataProcessingBatch.py*. This script will project the roost data and use the human population density raster to upweight monarch roost data in low population areas, creating the shapefile *rstdenpopall.shp*.

3) Use an R script editor like R Studio or Tinn-R to open the R script *KDEModel\_MonarchRoostTSE.R* (located in the RScript folder). Make sure FunctDirect (line 25) matches the location of *KDEMSubset\_GridTrainTestEvalAIC\_Calib\_Function.R* on your computer. Also, make sure the InDirect directory in the script matches the desired location on your computer (line 56). Run the R script from lines 1 to 128 to establish various parameters and plot the shapefile *rstdenpopall.shp.*

4) Similar to step 2, use the ArcPython script *MonarchRoostBackgroundExtentRaster\_PseudoabsenceRasterGeneration.py* to generate a background raster and pseudoabsence raster for later deriving Background points (not used for KDEM, but used if running MaxEnt) and Pseudoabsence points in R script below.

5) Going back to R script of step 3, start running code from line 136 to 364. This step creates Pseudoabsence points and monarch roost data point training set shapefiles for the KDE training set ensembles in the next step.

6) Similar to step 2, use the ArcPython script *MonarchRoostKfoldDataProjection\_KDEProcessingBatch\_PopDenInd.py* to generate a training set ensemble of three KDE surface rasters.

7) Back in the R script, run lines 374 to 518 for generating the training set ensemble of three KDEM rasters and an output a csv file of evaluation statistics in the InDirect directory. A new output subdirectory will be generated called something like “MonRstKDEMCNAA” (unless you changed parts of name in R script) for producing the three training set KDEM rasters.The function *KDEMSubset\_GridTrainTestEvalAIC\_Calib\_Function.R* is used to normalize the KDE surface from zero to one in creating the KDEM, which is calibrated to maximize the true skill statistic to create a binary presence/absence KDEM projection raster. Evaluation statistics are generated using held out testing data for each KDEM, including AUC, AICc, and AUCDiff.

8) Similar to step 2, use the ArcPython script *MonarchRoostKfoldDataProjectionTSE\_KDEProcessingBatch.py* to combine the three above generated training set KDEMS into a Training Set Ensemble (TSE) that can be used to display a minimum consensus (value 0 is no color, 1-3 is same color) or 100% consensus (values 0-2 is no color, 3 is color) binary calibration model.

After running step 1 above, the subsequent steps can also be applied to create yearly KDEMs from the monarch roost data after running the following provided ArcPython script to make separate shapefiles for each year: *MonarchRoostYearlyDataProcessingBatch.py*. The above R and Arc Python scripts have commented code that can be modified to process the individual years in loops.

The steps below related to analysis of the KDEM centroid locations and widths. Step 2 requires creation of annual KDEMs (see above paragraph).

1) Similar to step 2 above, use the ArcPython script *MonarchRoostYearlyKfoldDataProjectionAllYearTSE\_KDECentroidNorth.py* to generate dbf files with the centroid coordinates for the northwestern section of the KDEM created above.

2) Similar to step 2 above, use the ArcPython script *MonarchRoostYearlyKfoldDataProjectionTSE\_KDECentroidPathAreaNorthProcessingBatch.py* to create a csv file for individual annual KDEM centroid north/south and east/west shifts in km from the all year centroid and the KDEM widths. NOTE: This step requires annual KDEMs (see above).

3) Going back to R script of step 3 above, start running code from line 529 to 559. This step creates a summary output of mean values for annual KDEM centroid shifts and widths.

**Reference**

Tracy JL, Kantola T, Baum KA, Coulson RN (submitted 2018) Modeling fall migration pathways and spatially identifying potential migratory hazards for the eastern monarch butterfly. Submitted to Landscape Ecology