# Species Identification/Selection

## Endries Aquatic Occurrence Data

Mark Endries, of the US Fish and Wildlife Service (Asheville Office) provided us with a shapefile of aquatic species occurrences. Each species was represented as a multipoint feature, tagged with the species common name, scientific name, and the number of occurrences in the dataset.

## Disaggregating the multipoint features and tagging them with catchment/HUC data

The Python script “EEP\_CreateHabitatModelInput.py” converts the multipoint features in Endries’ aquatic species occurrence dataset in to single point features, and then intersects these points with the NHD catchment polygons. It then adds new field, “Present”, and sets all values to “1”.

Next, the attribute table of this result is transformed so that each species scientific name becomes a column with the rows set to the NHD catchment FEATUREID. Values of ‘1’ indicate the species was observed in the catchment, values of ‘0’ indicates it was not.

## Identifying species with enough occurrence records to sustain modeling

# Environment Layer Selection

# MaxEnt Analysis

# MaxEnt Results

# GLM Analysis

* **R\_CreateDataFile.py** – Merges records for a species selected from the HabModel table and the EnvStats feature class into a single csv table ready for analysis in R. This table sets the species columns to “1” in the catchments where the species was observed, and “0” in all other catchments. The catchments included in the table is limited to only those in HUC8s where the species was observed.
  + Makes a list of the HUC8s in which the species was recorded
  + Extracts catchments records in those HUC8s from the EnvStats feature class
  + Joins species presences records to those extracted records
  + Writes all occurrence records to an output file, setting species = 1
  + Writes all non-occurrence records to an output file, setting species = 0
* Correlate spp presence/absence with habitat variables.

# GLM Results