**Introduction**

Hello all, my name is John Kennedy. I have 25 plus years of experience using GIS technology in support of scientific endeavors related to hydrogeology, environmental science, and now marine biology. My BS, MS, and PhD are in geology, so my last ten years working for Fisheries Science & Technology has been one new adventure after another. My role on the DisMAP project is data processing and metadata management using ArcGIS Pro and Python.

**First Slide**

The current software environment I’m working in is ArcGIS Pro version 2.9, which uses Python version 3.7. I use the Pyscripter IDE version 4.1 for a majority of the scripting, but also occasionally use the Python window, Jupyter Notebooks, and ArcTools within ArcGIS Pro.

The primary Python module I use is arcpy, but others like pandas, numpy, and scipy play a critical part in data wrangling to create the final products.

For example, although there are functions in arcpy that can load Excel and CSV data into ArcGIS Pro, pandas allows for fine-tuning the data importing and formatting before loading the data into an ArcGIS table. These code snippets read the CSV file into a pandas dataframe, set the column names and field types, control for Null values, the insertion of a new column based on the calculation of an existing column (the Cube Root of WTCPUE is calculated), the dataframe is converted to an numpy array, and then arcpy is used to convert the numpy array to an ArcGIS table.

ArcPy is then used to convert the ArcGIS table to a set of geographic points representing the Survey Locations contained in the original CSV file for a region.

The Survey Locations for each region are then queried (using an arcpy tool) to create a sub-set of data for a selected species and a year range containing the selected year (in this example **1980**) and the previous and following two years. **<NEXT SLIDE>**

**Second Slide**

This selected set of points are then used to create a biomass distribution raster using the arcpy IDW or inverse distance weighted tool. The WTCPUECubeRoot field is used as the input values, the cell size is set at 2000m, power of two is set, a search neighborhood of 200 km radius, with maximum and minimum of nearest locations is set at 15 and 10, and the weight field is set to YearWeight. The output from that tool is then cubed. Cube-rooting the data before using the IDW spatial distribution tool and the cubing the data back is a common modeling technique.

The biomass distribution rasters are then combined into mosaics, which are then used to calculate species richness rasters. The final raster mosaics are then converted to Cloud Raster Format or CRF files that are then published online. The biomass rasters are then used as one of the inputs in calculating the distribution metrics that are then stored in the Indicators table. In addition to the biomass rasters, there are three other inputs used in calculating the distribution metrics. The first is bathymetry. For the Alaska region we used bathymetry provided by the Alaska Fisheries Science Center and for the other regions we used the data provided by GEBCO at gebco.net. The bathymetry for each region was sampled using a fishnet grid set at 2000m cell size and tied to the region extent. The fishnet grid was used also to create the other inputs, which are the latitude and longitude value rasters. The centroid of each grid cell provides the latitude and longitude value for the cell. **<NEXT SLIDE>**

**Third Slide**

The survey locations **<CLICK>** for each region, the regions themselves, and the distribution metrics table are then published **<CLICK>** to the ArcGIS Online / NOAA Geoplatform as Feature Layers **<CLICK>** using ArcGIS Pro. **<CLICK>** The biomass and species richness rasters, contained in a CRF, are published to the ArcGIS Enterprise / Fisheries Portal as Image Layers using ArcGIS Pro. As part of the above described processes metadata is created for each dataset in ArcGIS Pro and becomes a part of the dataset when published. **<CLICK>** The metadata is exported as ArcGIS formatted XML files, which are then loaded into the NOAA Fisheries InPort Metadata Catalog. The basic metadata items available in InPort are the dataset title, an abstract, a list of points of contact, and **<CLICK>** distribution links to the data that is published online. That’s my time, thank you. **<NEXT SLIDE>**