NICHOLAS SCHOOL OF THE ENVIRONMENT AND EARTH SCIENCES
DUKE UNIVERSITY

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
U.S. DEPARTMENT OF COMMERCE

Introduction to GIS, geospatial analyses, and EDC

Elliott Hazen, SWFSC-ERD postdoctoral researcher

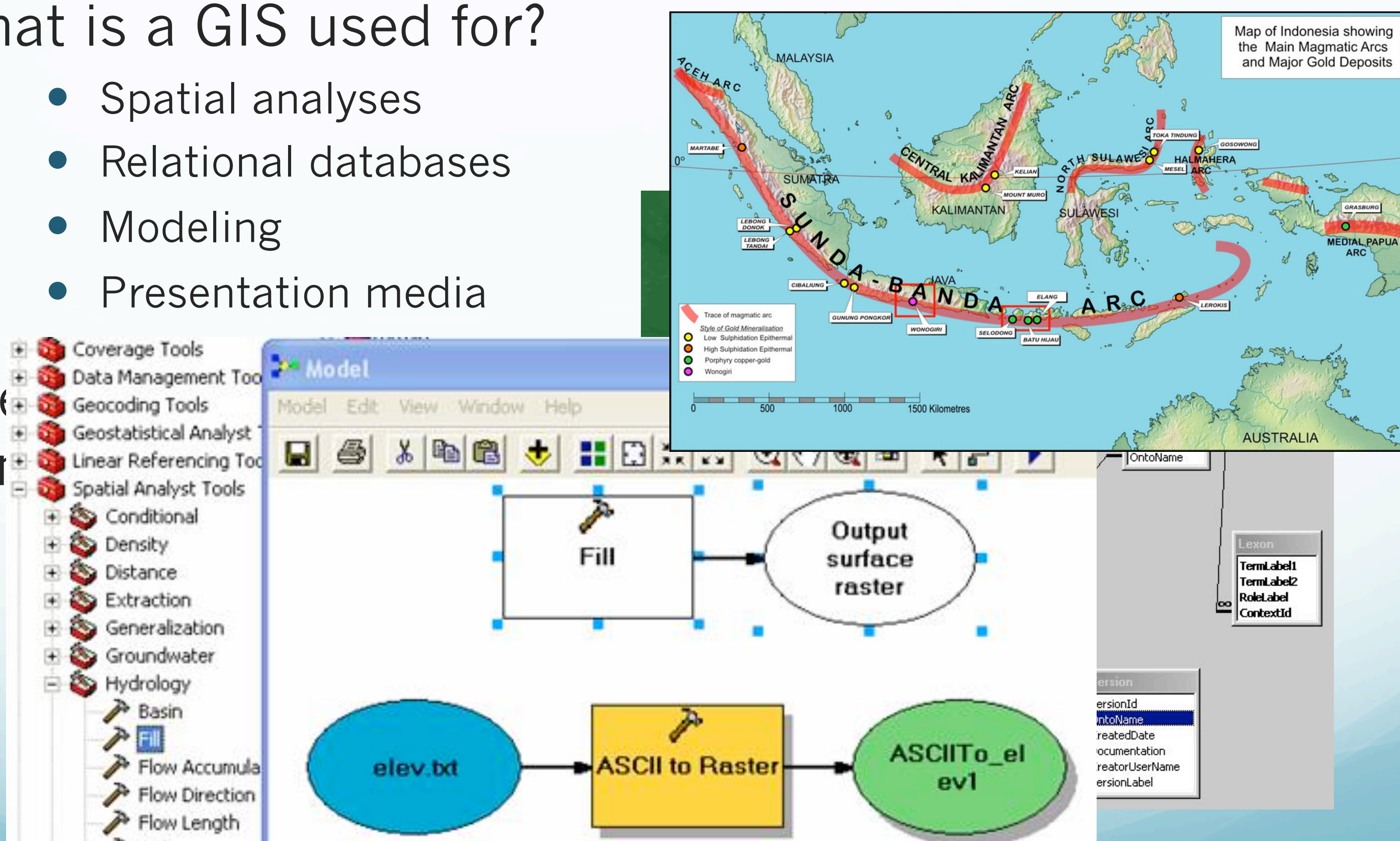
August 2-11, 2010

Introduction to GIS

- What is a GIS? - **Geographic Information System**
- **Geographic** because primarily deals with spatial features. Objects are referenced to a specific location in space and may be physical, biological, cultural, or economic.
- **Information** because data are processed and manipulated within a GIS – both spatial and non-spatial (e.g. attributes, metadata).
- **Systems** because complex analyses are broken down into component parts but ultimately form a unified map or prediction.

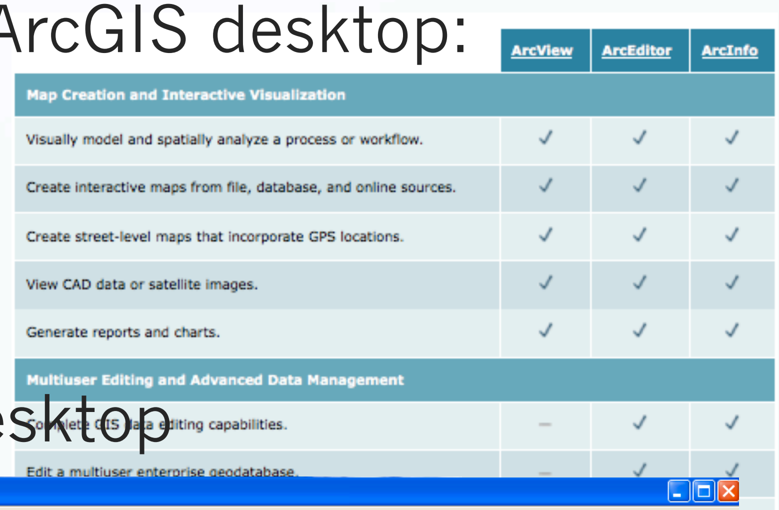
Introduction to GIS

- What is a GIS used for?
  - Spatial analyses
  - Relational databases
  - Modeling
  - Presentation media
- The
  co



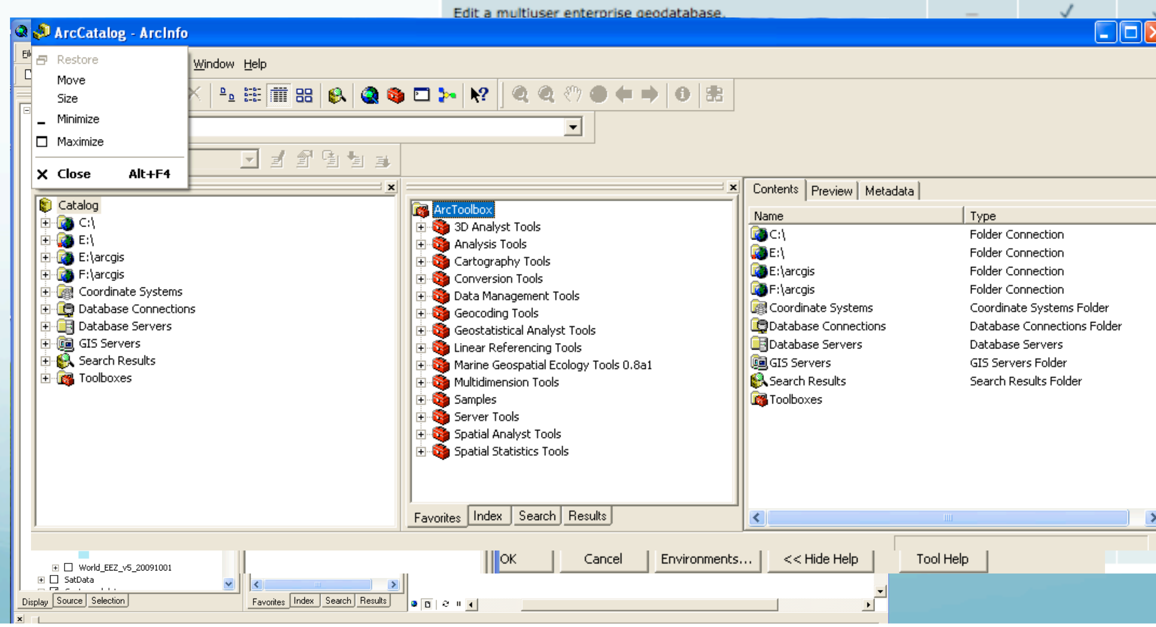
ArcGIS Desktop

- There are three versions of ArcGIS desktop:
  - Arcview
  - Arceditor
  - Arcinfo



	ArcView	ArcEditor	ArcInfo
Map Creation and Interactive Visualization			
Visually model and spatially analyze a process or workflow.	✓	✓	✓
Create interactive maps from file, database, and online sources.	✓	✓	✓
Create street-level maps that incorporate GPS locations.	✓	✓	✓
View CAD data or satellite images.	✓	✓	✓
Generate reports and charts.	✓	✓	✓
Multiuser Editing and Advanced Data Management			
Complete GIS data editing capabilities.	–	✓	✓
Edit a multiuser enterprise geodatabase.	–	✓	✓

- Three components within desktop
  - ArcMap
  - ArcCatalog
  - ArcToolbox



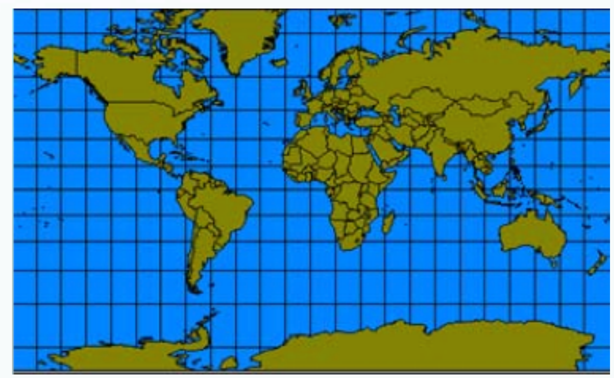
ArcGIS Projections & Coordinate Systems

- Converting a round earth to a flat map

Spherical (3D)

Datum

projection

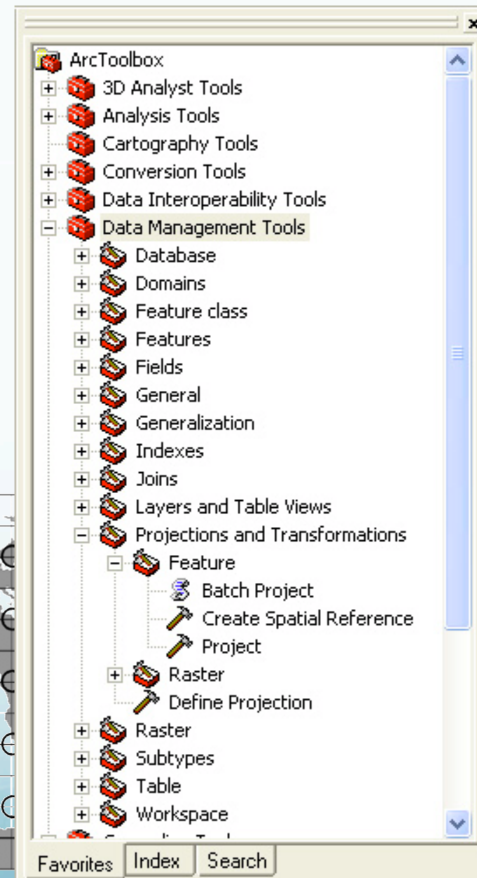


Planar (2D)

- Geographic projections (lat / lon) change in relative distance with position
  - 1° Lon ~ **56km** at 60°N/S
  - 1° Lon ~ **111km** at 0°N

ArcGIS Projections & Coordinate Systems

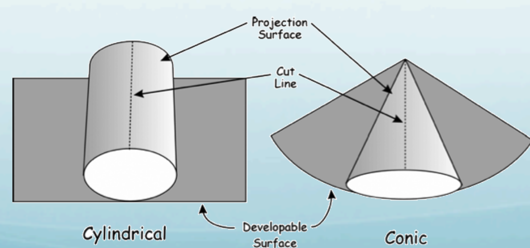
- Planar systems provide a constant distance anywhere on the map allowing calculations of:
  - Location (X,Y intersection)
  - Distance
  - Area
  - Direction
- Requires projections to convert from planar

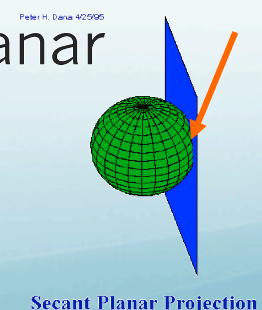


ArcGIS Projections

- Projected maps are often classified by the geographic *properties that they preserve*
  - **Conformal** *maps preserve local shape*
  - **Equal-area** *or equivalent maps retain all areas at the same scale*
  - **Equidistant** *maps maintain certain distances*
  - **True-direction** *maps express certain accurate directions*

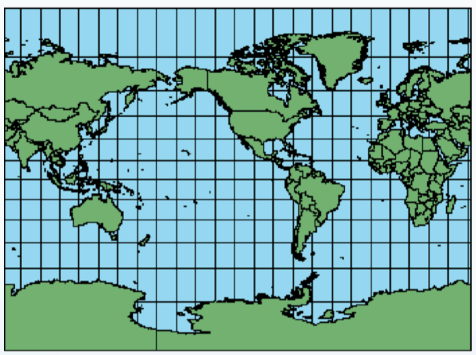
- Surfaces determine how the globe is ‘unraveled’
  - E.g. Cylindrical Conic Planar



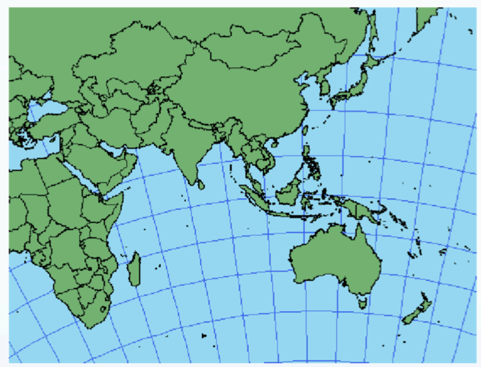


ArcGIS Projections

- Each projection has limitations – e.g. global vs. local
- Example projections



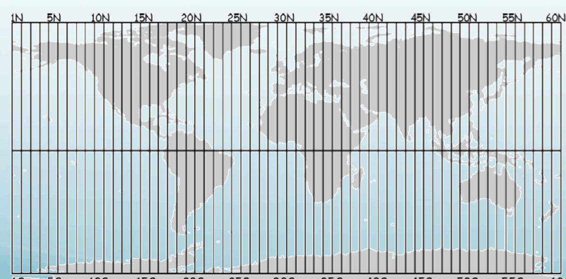
Mercator (secant)



Lambers Conformal Conic

60 N

Universal Transverse Mercator



60 S

State plane

ArcGIS: Extensions

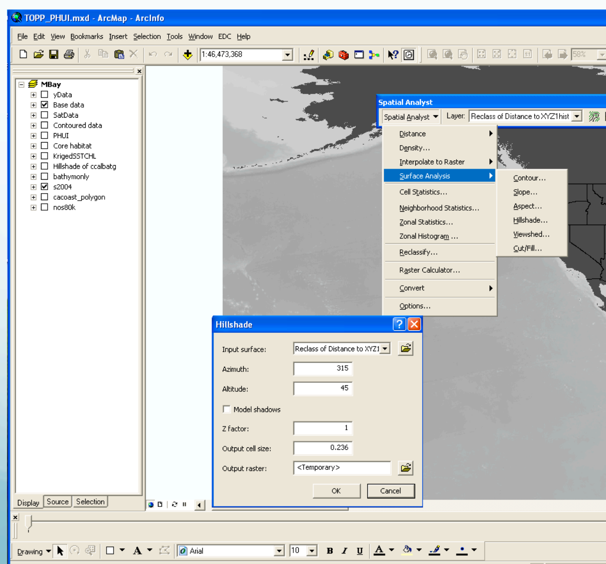
- Spatial Analyst
- Geostatistical Analyst
- Environmental Data Connect.
- Many more both included in Arc and downloadable

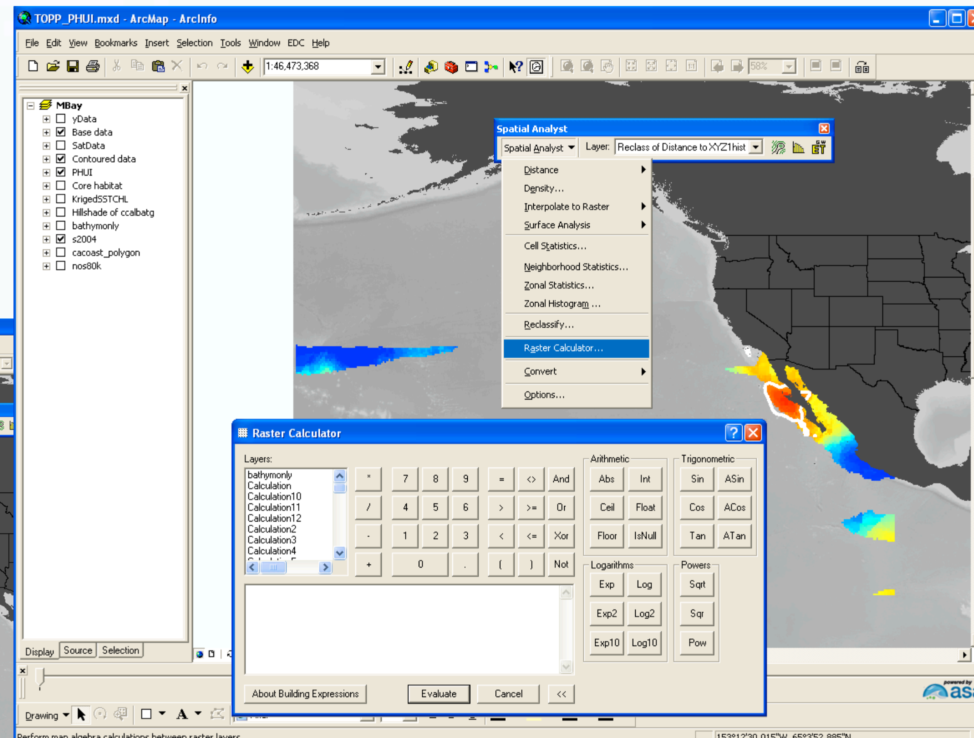
ArcGIS Spatial Analyst	Advanced raster modeling ArcGrid™ calculator with ArcGrid algebra VBA for raster analysis
ArcGIS 3D Analyst	ArcScene™: real-time interactive 3D scenes Scene views in ArcCatalog 3D Modeling tools ArcTIN™ tools
ArcGIS Geostatistical Analyst	Advanced kriging and surface modeling Exploratory spatial data analysis tools Probability, threshold, and error mapping
ArcGIS Survey Analyst	Comprehensive survey information management using the geodatabase Advanced survey computation Improved GIS data accuracy via links to survey locations
ArcGIS StreetMap™ USA	Complete nationwide streets database Nationwide geocoding
ArcGIS StreetMap Europe	European street mapping and geocoding
ArcGIS Tracking Analyst	Time-based map display and rendering Playback tools (Play, Pause, Forward, Rewind) Work with any time-based data (e.g., features that move/change or whose values change through time)

ArcScan™ for ArcGIS	Integrated raster-vector editing Vectorize features from raster Raster snapping (Requires an ArcEditor or ArcInfo license
ArcGIS Schematics	Database-driven schematic rendering and display Schematic views of GIS networks and tabular information Multiple schematic representations
Maplex for ArcGIS	Advanced label placement and conflict detection for high-end cartographic production Greatly simplifies the labor-intensive placement of map text
ArcGIS Publisher	Publish ArcMap documents as PMFs for use with free ArcReader seats Also used with the ArcMap Server extension to ArcIMS
ArcPress for ArcGIS	Advanced map printing
MrSID® Encoder for ArcGIS	Image compression and mosaicking Up to 500 megapixels
TIFF/LZW Compression	Right to use Unisys® patent for writing compressed data using TIFF/LZW

ArcGIS: Extensions

- **Spatial Analyst**
  - Raster (cell) analysis
  - Surface analysis
  - Terrain analysis
  - Hydrologic/ Bathymetric analysis

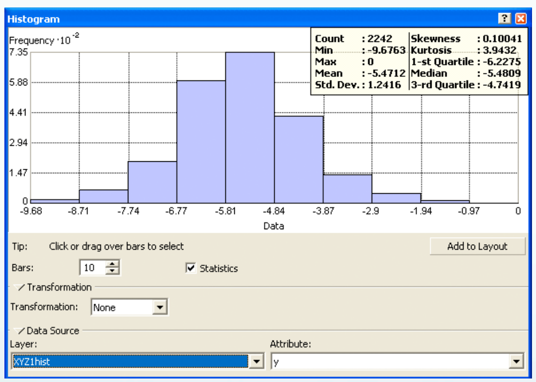


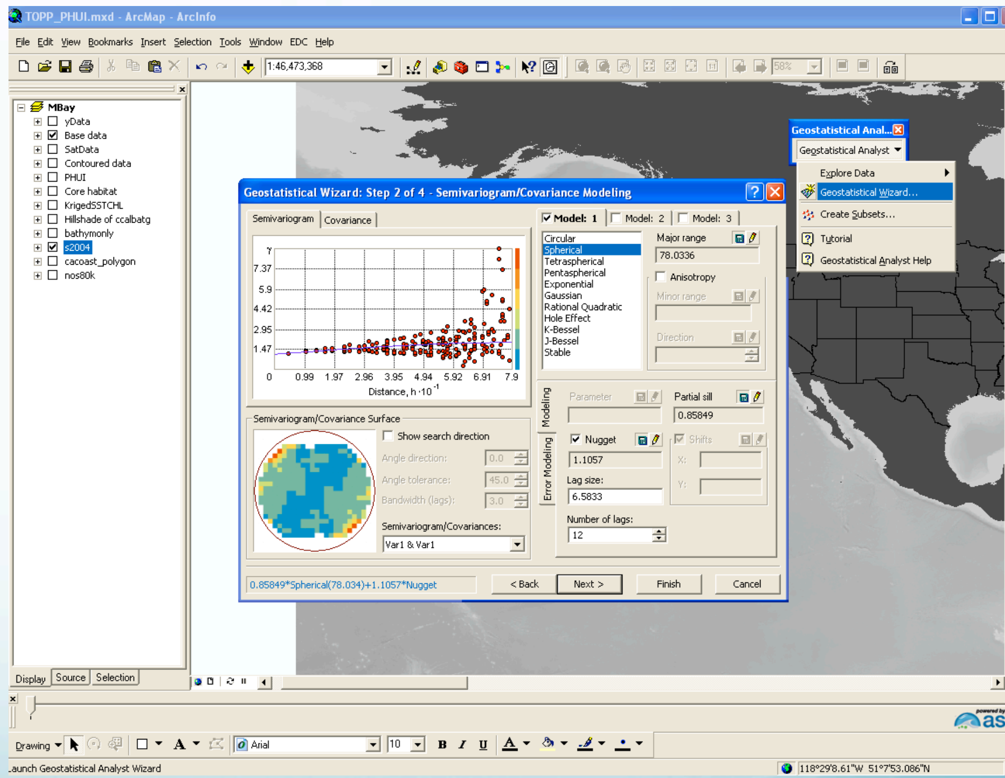


Raster calculator is accessed from ArcGIS spatial analyst toolbar and is used to perform raster modeling on any number of raster layers.

ArcGIS: Extensions

- **Geostatistical Analyst**
  - Data exploration
  - Surface interpolation
  - Kriging

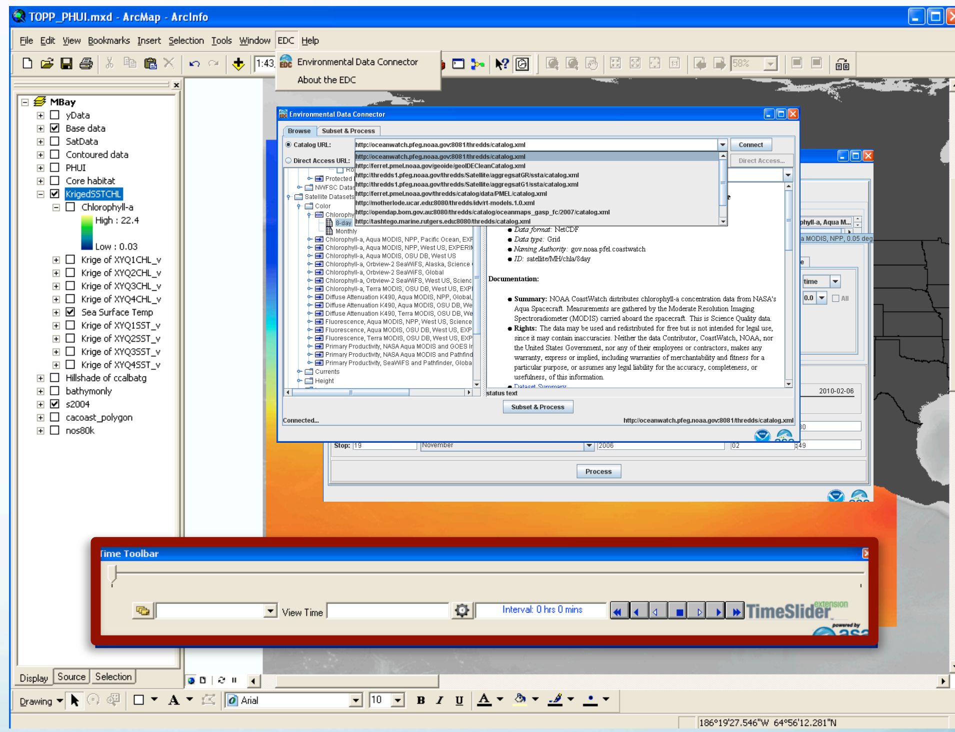




Although also available in Spatial Analyst, Geostatistical Analyst allows more control over interpolations and can also generate summary statistics and analyze trends

ArcGIS: Extensions

- **Environmental Data Connector**
  - Access hosted data from within Arc
  - Subset data based on time and space
  - Adds a temporal component to ArcGIS with the time slider toolbar

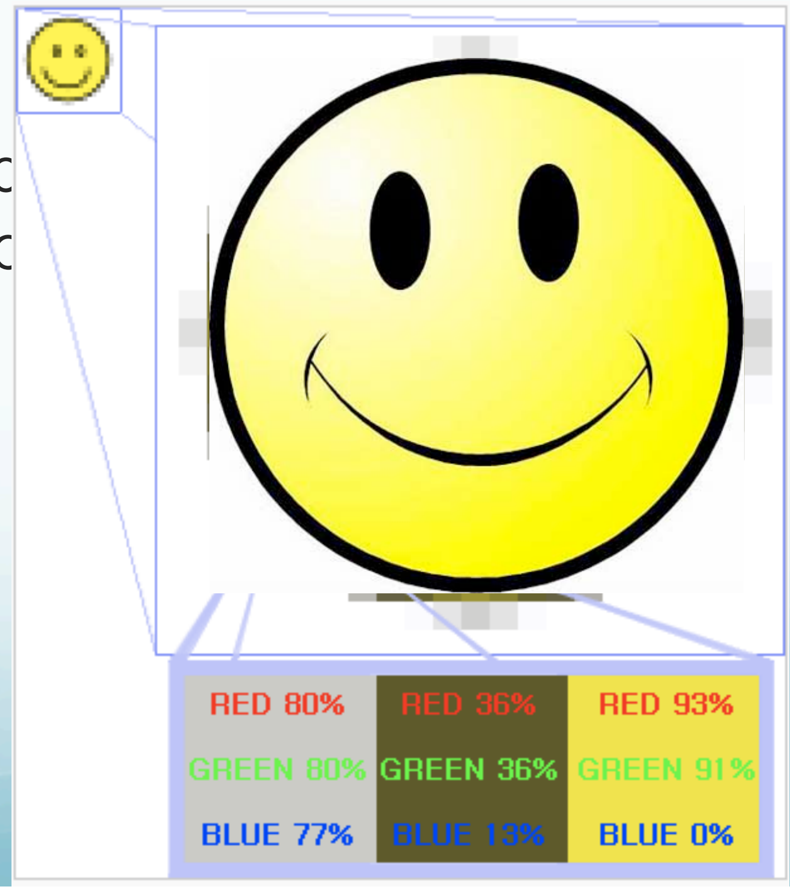


http://www.pfeg.noaa.gov/products/EDC/index.html

ArcGIS data types

- Raster vs. Vector
- Raster files are georeferenced images with explicit cell size (grain) and extent

  Stored as Shapefile,
  Raster, Geodatabase,
  ~~Coverage~~
- Vector files are scaleable and c
  no "cells" (points, lines, polygo
- **Examples:**
  - GHRSST, MODIS-Aqua
  - CA Coastline, I-40
  - Turtle locations, CTD data
  - Global bathymetry, altimetry



NICHOLAS SCHOOL OF THE ENVIRONMENT AND EARTH SCIENCES
DUKE UNIVERSITY

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NOAA
U.S. DEPARTMENT OF COMMERCE

My Focus

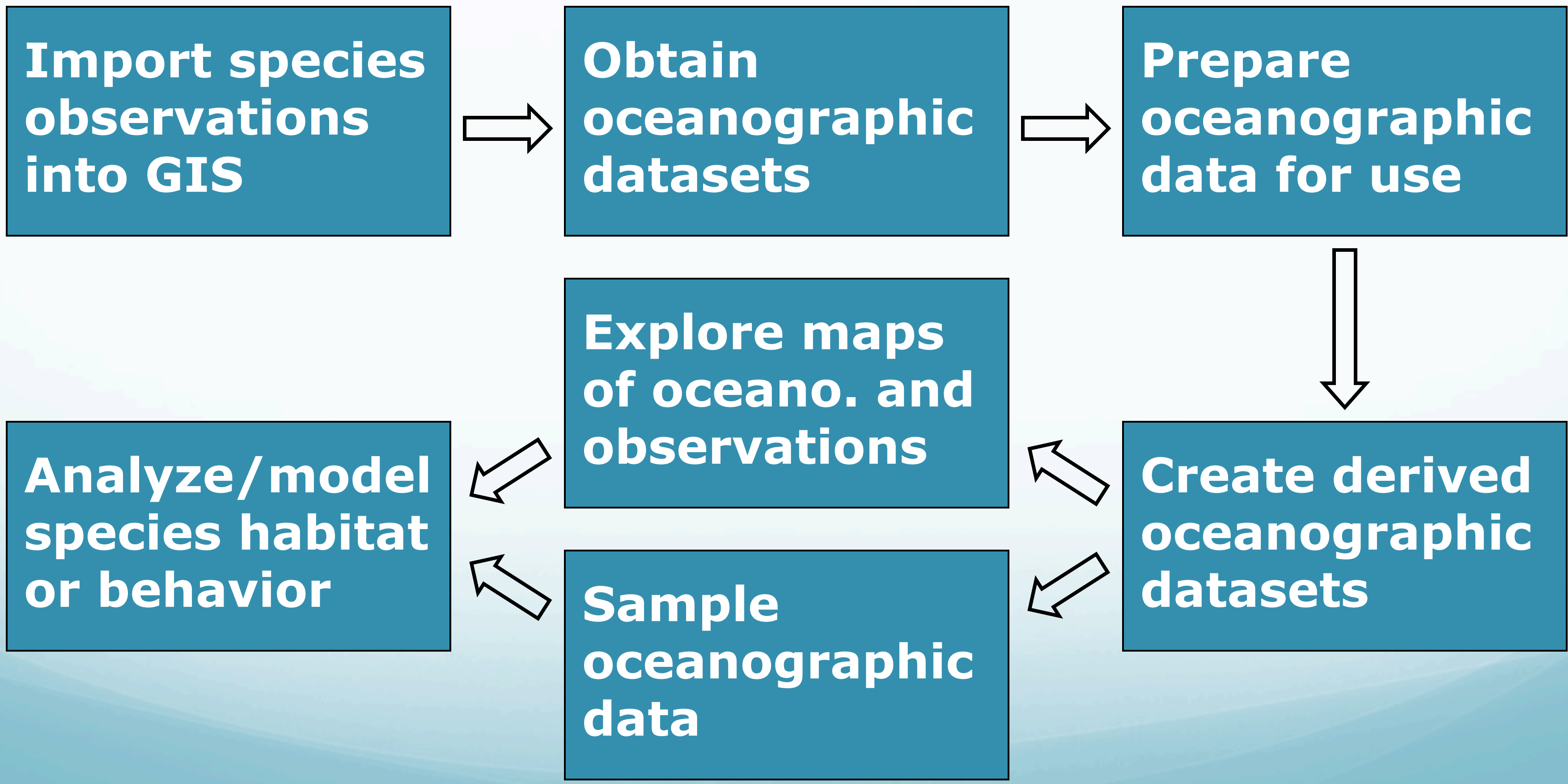
- Disciplinary
  - Space (and Time)
  - Environment + Prey
  - Statistics to prediction
- Flowchart

Data inputs
- Distribution
- Predictors

→ Statistical model fitting

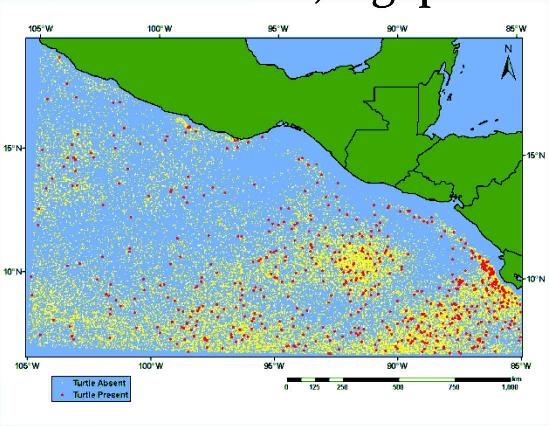
→ Habitat prediction models

Hypothetical workflow



Modeling habitat (overview)

Distribution data, e.g. presence/absence



Presence only data, e.g.
- Vessels of opportunity
- Hydrophones

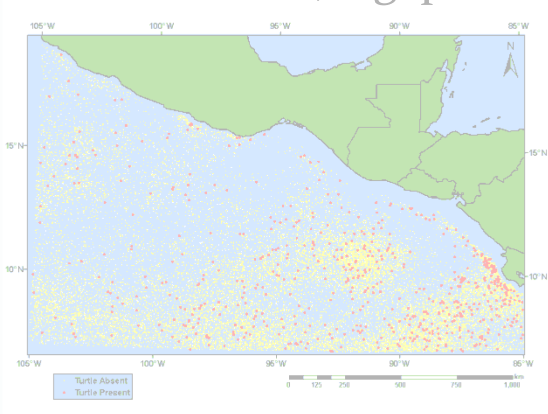
Presence / absence, density data
- Survey sightings w/ effort
- Bycatch

Event based data
- Focal follows
- Short term tag data

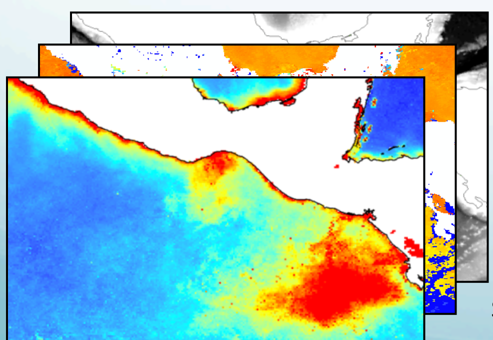
Movement data
- Short and long term tag data

Modeling habitat (overview)

Distribution data, e.g. presence/absence



Sampled predictive data



Bathymetry

SST

Chlorophyll

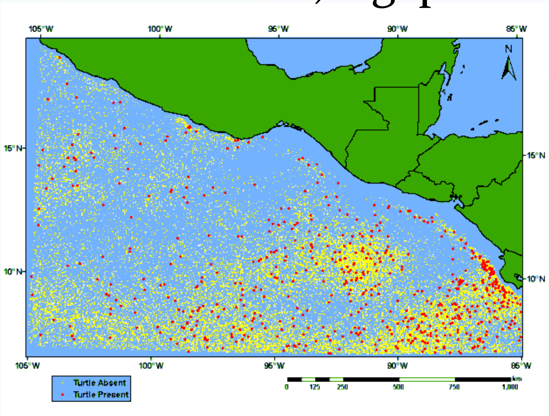
In situ data
•Continuous data – surface sensors, fisheries acoustics, ADCP
•Station data – CTDs, trawls

Remotely sensed data
•SST, SSH, Chl - Data centers

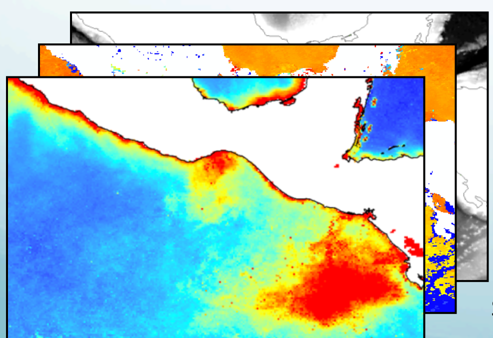
Physical / spatial data
•Bathymetry, distance from feature (e.g. slope, shore, break, front)

Modeling habitat (overview)

Distribution data, e.g. presence/absence



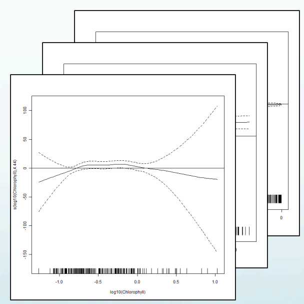
Sampled predictive data



Fit

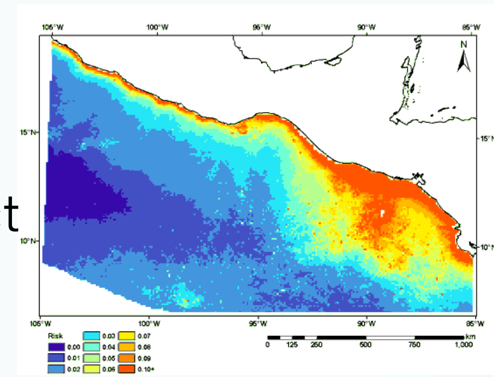
Statistical models

$$g(\mu) = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m$$

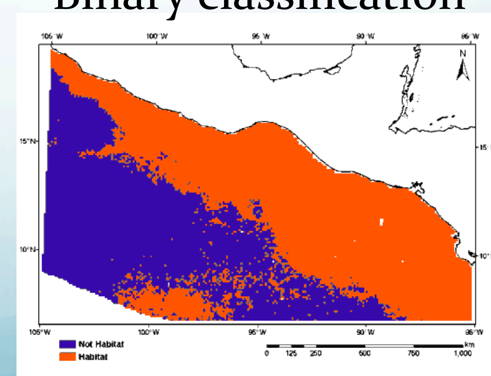


Predict

Probability of occurrence predicted from environmental covariates



Binary classification



Types of statistical models

- There are many, and constantly changing / growing
- Correlation/Regression techniques – GLMs, GAMs (Austin 2002), Mixed models (Wood 2006), regression trees & random forests (Breiman 2001)
- Ordination –Multivariate dimensional scaling, e.g. CCAs (Guisan et al. 1999),
- Maximum Entropy models – species distributions "closest to uniform" (Phillips et al. 2006)
- Recent reviews of modeling approaches (Redfern et al. 2006, Elith et al. 2006, Dormann et al. 2007, Aarts et al. 2008)

Useful software packages

- MATLAB / IDL – multipurpose scientific programming language; PERL
- WinBugs – toolset for Bayesian analysis
- R / S^+ / SAS – statistical programming language
- Python – scripting language used by Arc
- ArcGIS Desktop – Geographic Information System
  - Model builder
  - Hawth's tools, ET Geowizards, Biomapper, MGET
  - **EDC**

What is MGET?

http://code.env.duke.edu/projects/mget

- A collection of geoprocessing tools for marine ecology
  - Oceanographic data management and analysis
  - Habitat modeling, connectivity modeling, statistics
  - Highly modular; designed to be used in many scenarios
  - Emphasis on batch processing and interoperability
- Free, open source software
- Written in Python, R, MATLAB, C#, and C++
- Minimum requirements: Win XP, Python 2.4
- ArcGIS 9.1 or later currently needed for many tools
- ArcGIS and Windows are only non-free requirements

What is EDC?

EDC™ Environmental Data Connector
an extension for connecting ArcGIS to Thredds/OPeNDAP

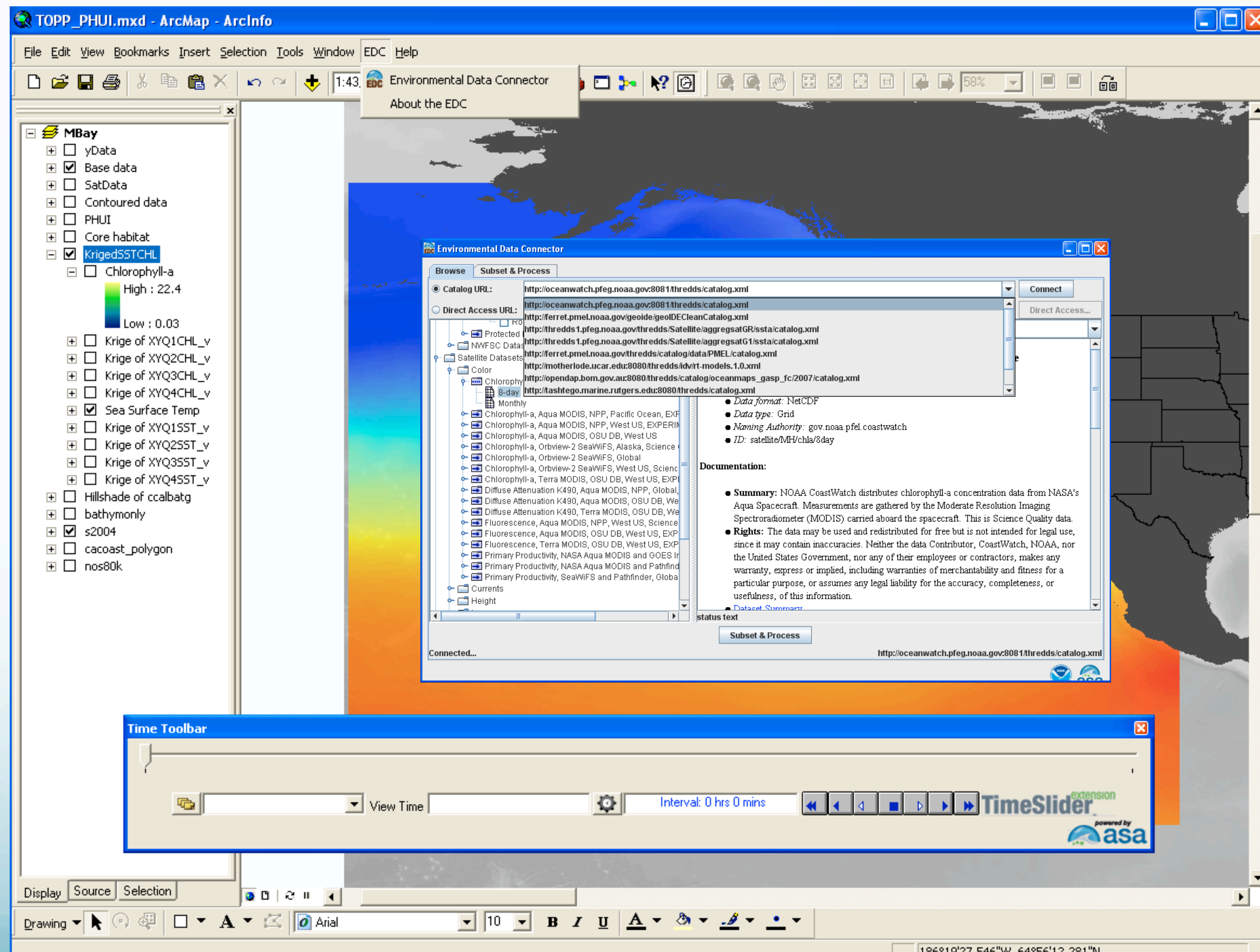
- Allows users to connect to data servers from within ArcGIS and download environmental data
- Works with THREDDS/OPeNDAP servers to provide feature and raster data
- Incorporates time in ArcGIS (great for videos)

EDC installation

- http://www.pfeg.noaa.gov/products/EDC/index.html
- Requires:
  - ArcGIS 9.2 sp 3 or higher (not yet 10)
  - Java Run-Time Environment 6
  - .NET Framework 2.0 or 3.0
  - .NET Support for ArcGIS Desktop

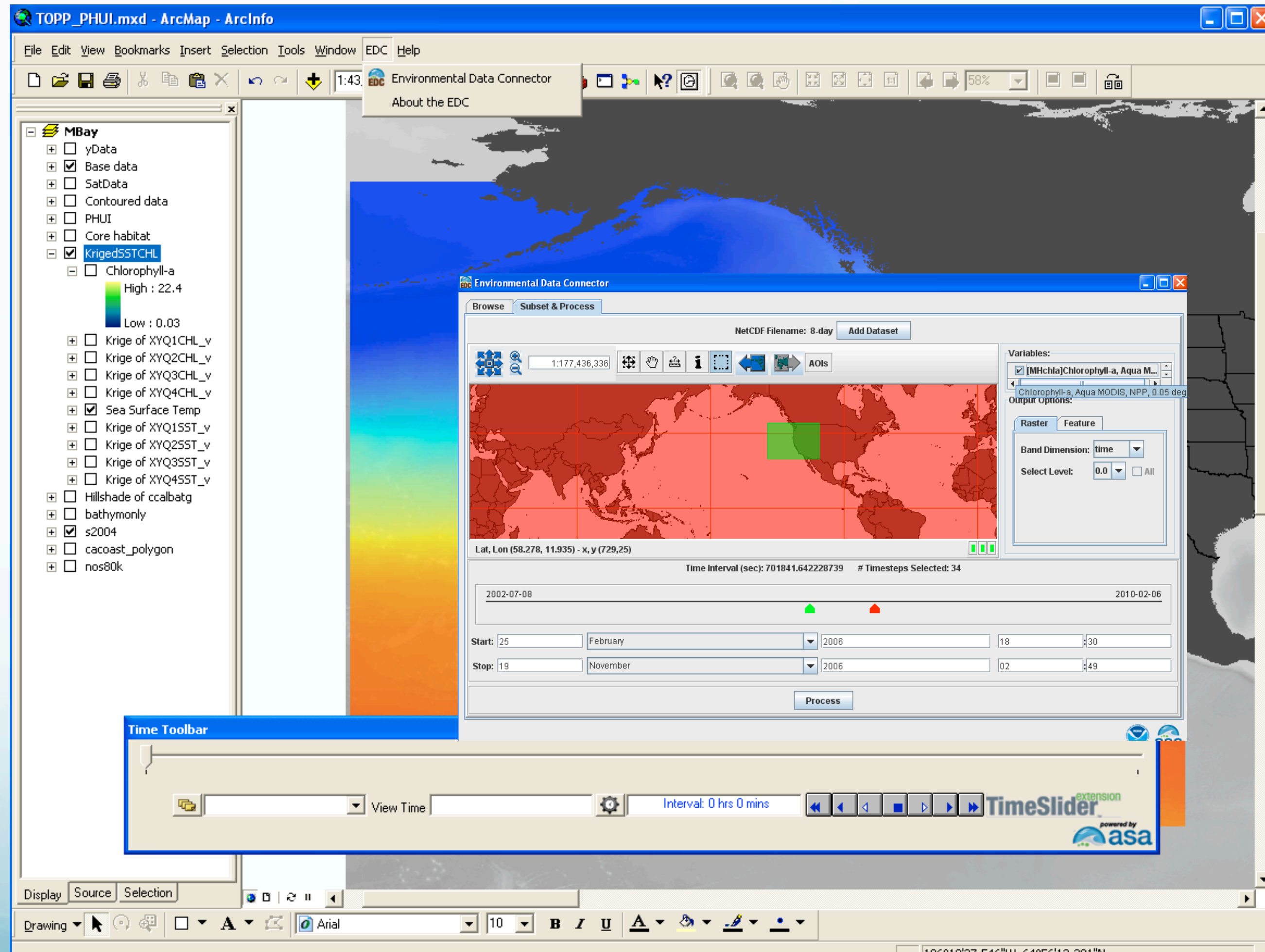
EDC workflow

- Finding data



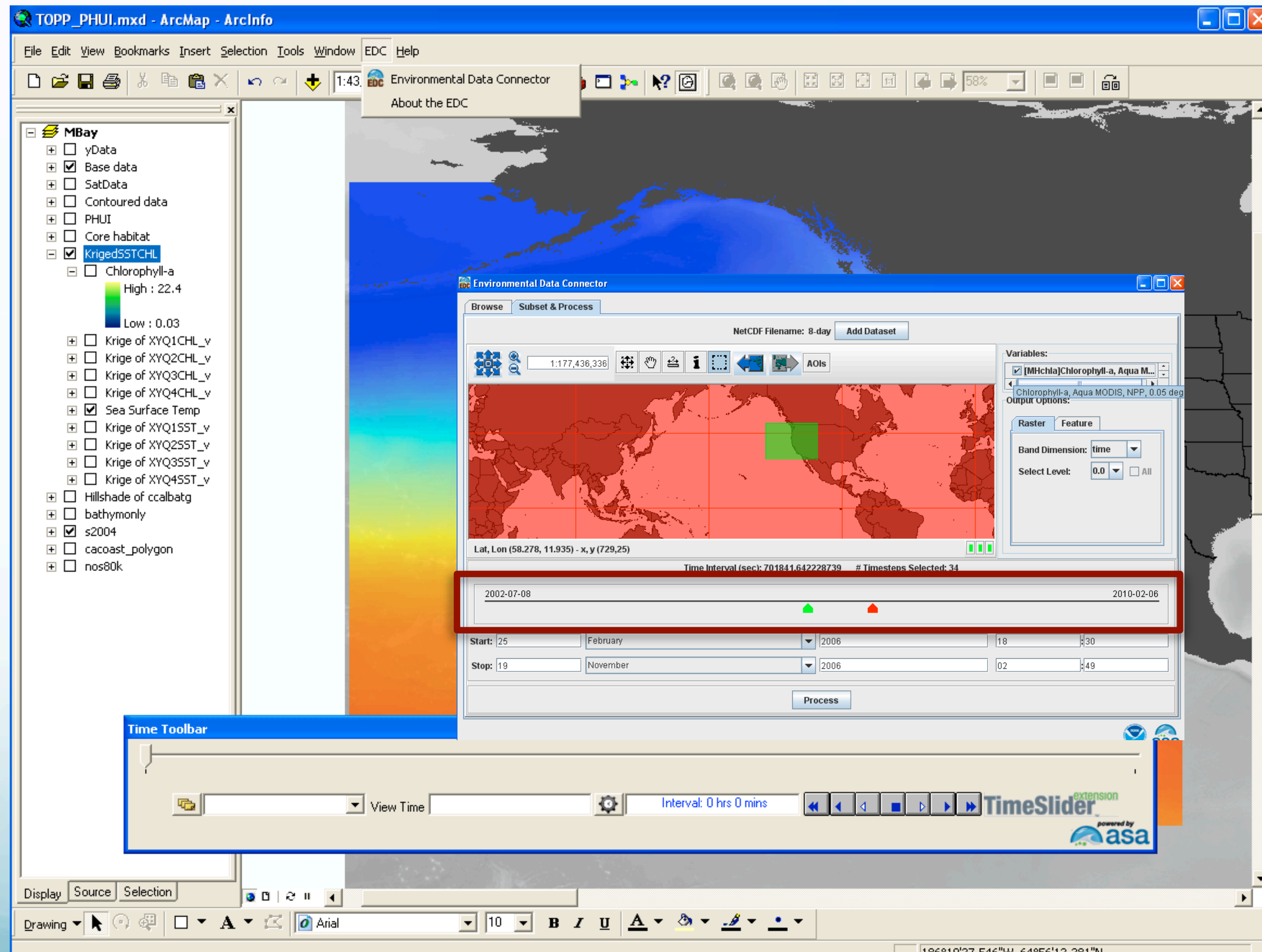
EDC workflow

- Finding data
- Subset data



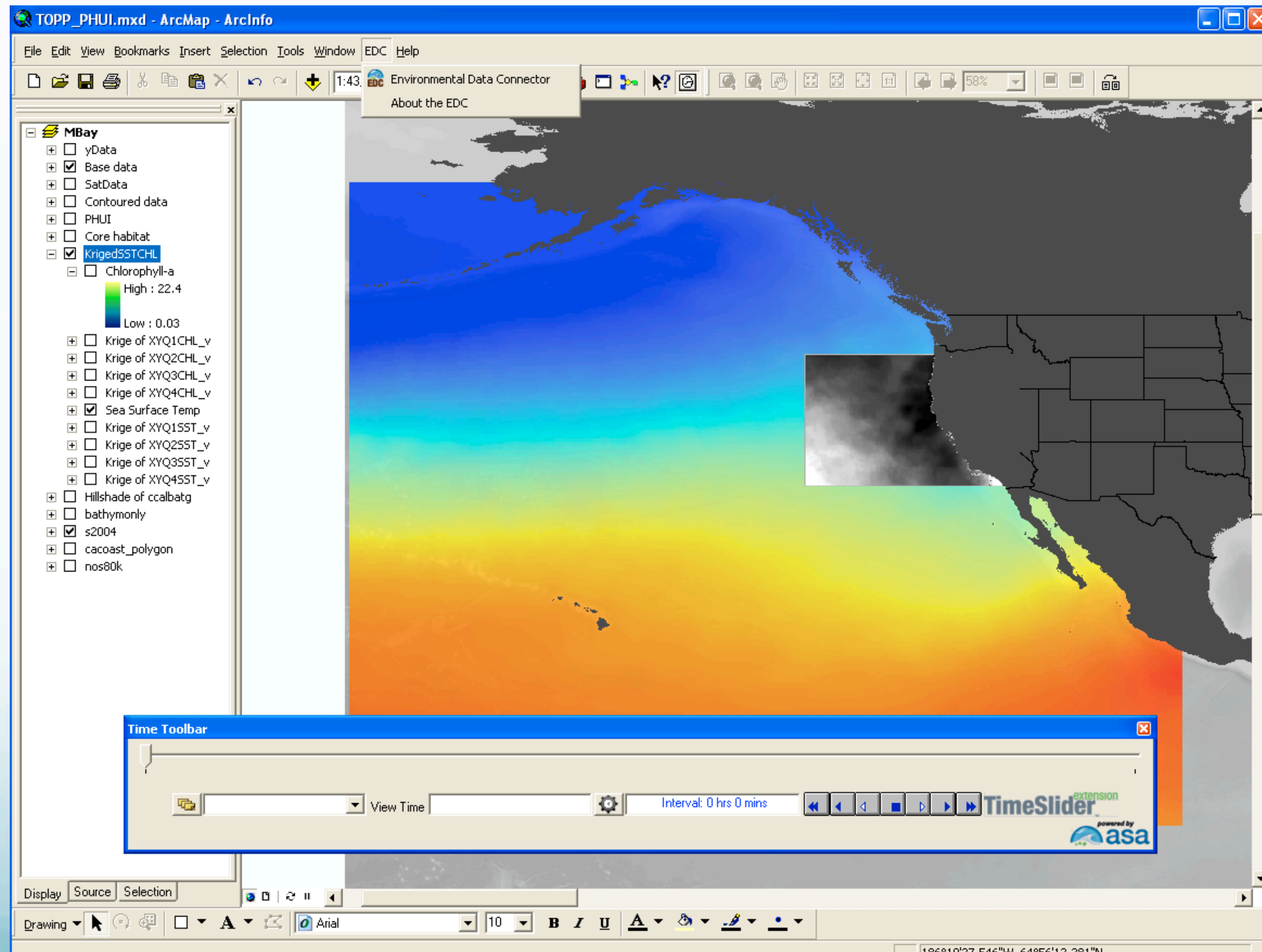
EDC workflow

- Finding data
- Subset data
- Choose time



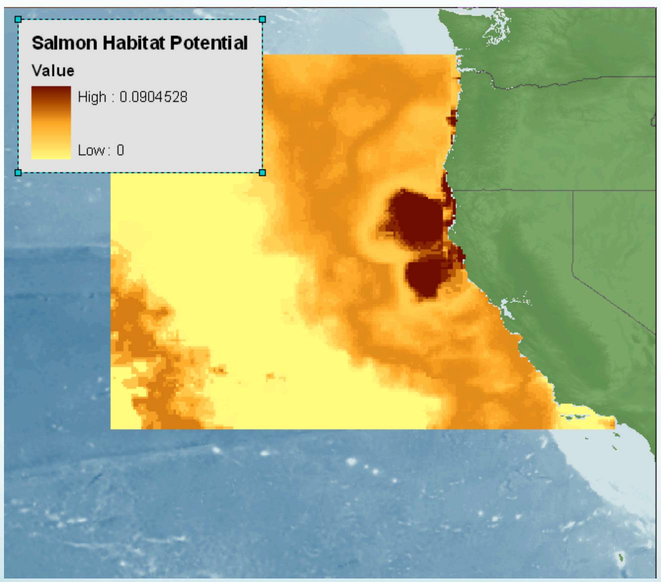
EDC workflow

- Finding data
- Subset data
- Choose time
- Process



Example: Salmon & SST

- Identifying ideal habitat for Chinook Salmon using EDC and ArcGIS Spatial Analyst



Vol. 285: 181–192, 2005 | MARINE ECOLOGY PROGRESS SERIES Mar Ecol Prog Ser | Published January 19

Ocean habitat use in autumn by Chinook salmon in coastal waters of Oregon and California

Jefferson T. Hinke[1,2,*], George M. Watters[2], George W. Boehlert[3], Paul Zedonis[4]

[1]Joint Institute for Marine and Atmospheric Research, University of Hawaii, 1000 Pope Road, Honolulu, Hawaii 96882, USA
[2]NOAA Fisheries, Pacific Fisheries Environmental Laboratory, 1352 Lighthouse Ave, Pacific Grove, California 93950, USA
[3]Hatfield Marine Science Center, Oregon State University, 2030 S.E. Marine Science Dr, Newport, Oregon 97365, USA
[4]US Fish and Wildlife Service, 1655 Heindon Road, Arcata, California 95521, USA

ABSTRACT: Describing the ocean habitats used by Chinook salmon *Oncorhynchus tshawytscha* is an important step towards understanding how environmental conditions influence their population dynamics. We used data from archival tags that recorded time, temperature and pressure (depth) to define the coastal habitats used by Chinook near Oregon and California during the autumns of 2000, 2002 and 2003. We used a clustering algorithm to summarize the data set from each year and identified 4 general habitats that described the set of ocean conditions used by Chinook. The 4 habitats, defined primarily by depth and the time of day that these depths were occupied, were characterized as (1) shallow day, (2) shallow night, (3) deep and (4) deepest. The definitions and use of each habitat were similar across years and the thermal characteristics of all habitats included water temperatures between 9 and 12°C. This temperature range provided the best indicator of Chinook habitat in the coastal ocean. Chinook used 9 to 12°C water at least 52% of the time. Less than 10% of surface waters within the area where Chinook were released and recovered provided these temperatures. Cross sections of subsurface temperatures suggest that between 25 and 37% of the coastal water column was available to Chinook and contained water in the 9 to 12°C range. These results support hypotheses that link salmon-population dynamics to ocean temperatures. Continued monitoring of surface and subsurface thermal habitats may be useful for assessing the extent and quality of conditions most likely to sustain Chinook salmon populations.

KEY WORDS: Chinook salmon · Archival tag · California current · Essential fish habitat

INTRODUCTION

It is increasingly understood that both freshwater and marine habitats are critical to the maintenance of healthy Pacific salmon *Oncorhynchus* spp. populations. Understanding the linkages between environmental conditions, ocean habitats, survival and growth of these fishes is a primary goal of current salmon research initiatives (Boehlert 1997, Gargett 1997, Bisbal & McConnaha 1998, Welch et al. 2003). Correlations of environmental conditions with catches (Mantua et al. 1997) and production (Beamish et al. 1997, Cole 2000, Hobday & Boehlert 2001) have demonstrated a strong coupling of oceanographic variability and salmon population dynamics. In particular, temperatures are correlated with salmon survival throughout much of the Pacific Ocean (Mantua et al. 1997, Cole 2000, Mueter et al. 2002a). The effects of environmental conditions on population dynamics, however, are realized through the continuous interactions of individuals with their environment. Unfortunately, individual patterns of habitat use over time and the environmental conditions experienced in those habitats have rarely been measured.

Recognition of the importance of marine habitats for survival and production has fostered research aimed at identifying the environments that salmon actually experience (Boehlert 1997, Welch et al. 2003). Such

*Email: jefferson.hinke@noaa.gov

Hinke, J.T., Foley, D.G., Wilson, C. & Watters, G.M. (2005) Persistent habitat use by Chinook salmon *Oncorhynchus tshawytscha* in the coastal ocean. Marine Ecology Progress Series, 304, 207–220.