

# *Survey Analysis Process*

**Anthony (Tony) R. Olsen**

**USEPA NHEERL Western Ecology Division**

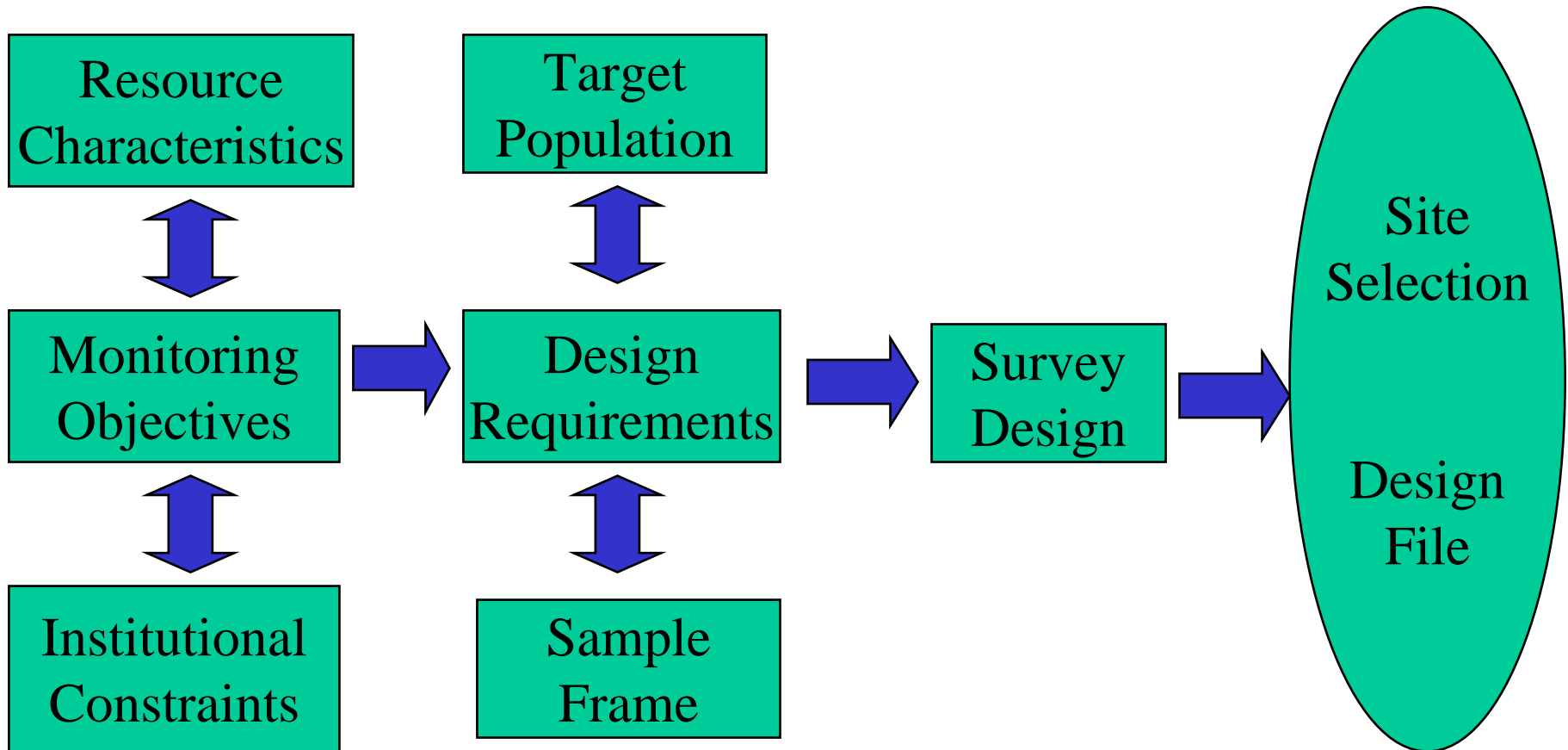
**200 S.W. 35th Street**

**Corvallis, OR 97333**

**Phone: 541 754-4790**

**Email: [olsen.tony@epa.gov](mailto:olsen.tony@epa.gov)**

## *Design Structure*



## *Example Designs*

- Everglades marshes and canals
- Streams and rivers in 12 western states
- Headwater watersheds in coastal plains of Mid-Atlantic
- Prairie pothole wetlands in North Dakota and South Dakota
- 6-th field hydrologic units in Pacific Northwest
- FIA and FHM monitoring of forests
- Amphibians in Olympic National Park and Southeast Oregon
- Riverine wetlands associated with the Great Lakes
- All Lakes >1 ha for fish tissue contaminants

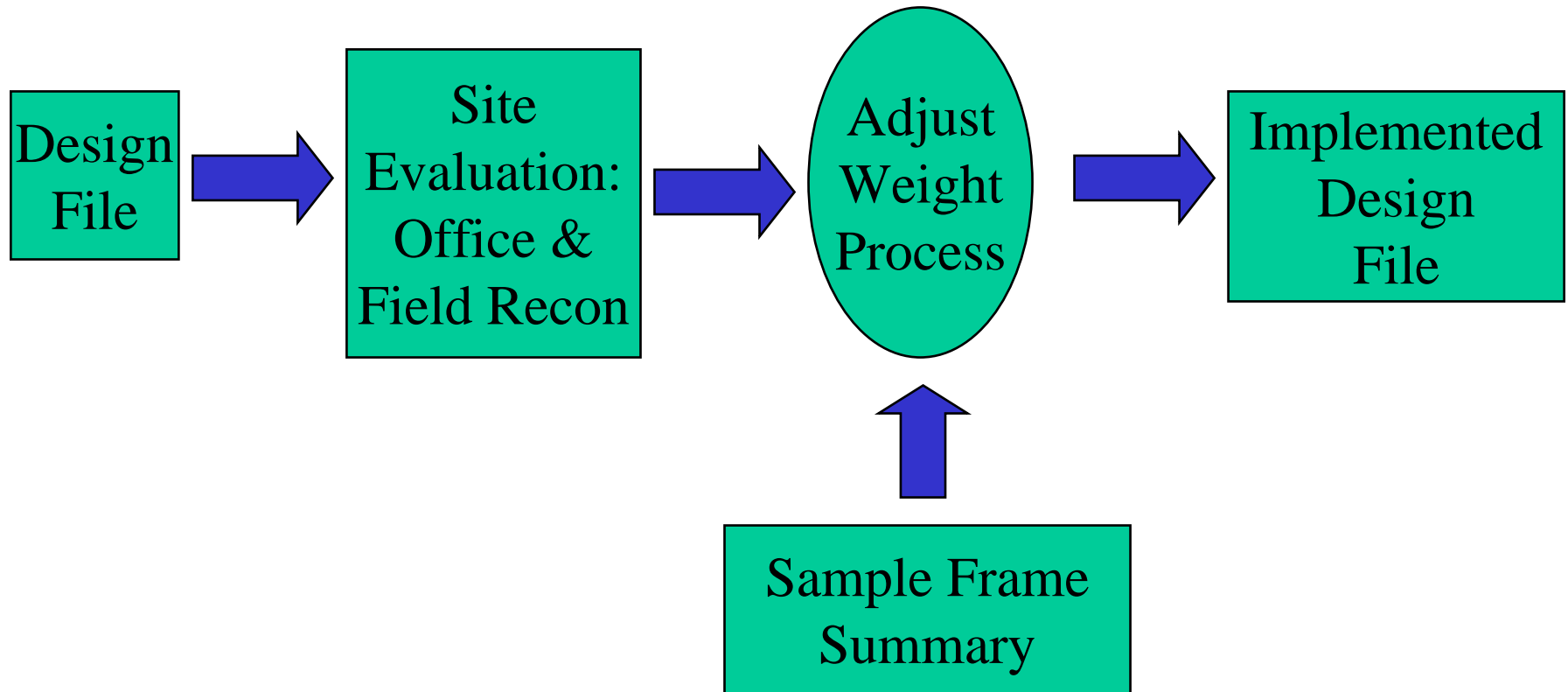


# *Indiana Rotating Basin Stream and River Monitoring Program*

- Target population: all perennial streams and rivers in Indiana
- 5 years to cover the state where each year focus on a collection of basins
- 5 sets of basins are explicit strata
- Unequal probability based on Strahler stream order to achieve approximately equal number of sites: 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup>+
- Biological indicators



# *Incorporating Site Evaluation Information in Design File*



# *Design File Contents:*

## *Has all sites selected by survey design*

- Site Identification
  - Site ID
  - Latitude/Longitude
- Auxiliary Frame Information
  - Site Name
  - Sample Frame ID
  - County
  - Map names
  - State
  - Omernik Ecoregion
  - Other
- Survey Design
  - Stratum
  - Panel
  - Oversample
  - Multi-Density Category
  - Initial Weight
- Site Evaluation
  - Evaluation Status
  - Evaluation Reason



# *Typical Site Evaluation Process*

- Each site evaluated to determine
  - Stream channel existence
  - Perennial or non-perennial
  - Other characteristics
- Three phase evaluation
  - Office assessment based on existing information
  - Phone call to knowledgeable local person
  - Field visit



# *Evaluation Status*

- Critical that know what happened to each site selected by the design
  - Ensures that know how the design was implemented
  - Use information to adjust weights for statistical analysis
  - Use to estimate the extent (size) of the sampled population
- Evaluation Status codes (example)
  - TS: Sampled
  - LD: Landowner Denied Access
  - PB: Physically inaccessible
  - NT: Non-Target
  - NS: Not Sampled
  - NN: Not Needed
- Reason for evaluation status





# *Population Estimation: What can estimate?*

- Population extent estimates for Evaluation Status categories
- Proportions for indicators that are categorical
- For continuous indicators
  - Population mean and variance
  - Cumulative Distribution Function estimates for continuous variables
  - Percentile estimates for continuous variables
  - Testing for difference between two CDFs



# *Statistical Computing Environment*

- Information management for data
  - Monitoring organization responsible
  - STORET for archival
- Data file format will be:
  - ASCII: CSV or Tab delimited
  - Excel spreadsheet
- Monitoring staff has
  - Limited Time for statistical analyses
  - Minimal statistical training
- R Statistical Software
  - Flexible statistical analysis environment
  - Library for survey analyses
  - Software free
- S-Plus Professional
  - Flexible statistical analysis environment
  - Library for survey analyses
  - Software cost
- SAS
  - Flexible analyses
  - SAS macros developed
  - Software cost



# *R (S-Plus) Operation*

## *Initial Step*

- Place all estimation functions in single folder
- Start R
- Install psurvey.analysis library
  - Only have to do this part once or whenever an updated set of functions are provided

## *Each Project*

- Create a folder for each project
- Place an R shortcut icon in folder and change target folder to project folder
- Start R
- Start word or text processor program with example R script file
- Complete analysis by modifying example R script



# *Data Analysis Sequence*

- Set up workspace initially
- Import Design File and Site Evaluation File OR Design Status File
  - Adjust weights if not done
  - Export final Design Status file
  - Do population extent estimation
  - Save extent results
- Import data results file
- Do population status estimation
  - Category estimates
  - CDF estimates
  - Mean, Total, Var estimates
- Construct plots
  - Save as pdf files
- Save status results



# *Monitoring Statistical Analysis Information*

## *<http://WWW.EPA.GOV/NHEERL/ARM>*

- Go to **Design & Analysis > Analysis Information**
- Three levels
  - Overview of the analysis process for information from probability surveys, with examples
  - Detailed information to guide analyses
  - Example results and reporting methods
  - Technical guidance and software to support statistical analyses and estimation procedures
- Bibliography: consult papers by Stevens and Olsen

