

State-Scale
Probability
Surveys:
An Overview

Garrett Stillings-EPA/OWOW/MAB
Stillings.Garrett@epa.gov

**2022 NATIONAL TRAINING WORKSHOP ON
WATER QUALITY DATA, ASSESSMENT, AND PLANS**

May 31, 2022

Section 305(b): Water Quality Inventory

Each State shall prepare and submit, a report which shall include--

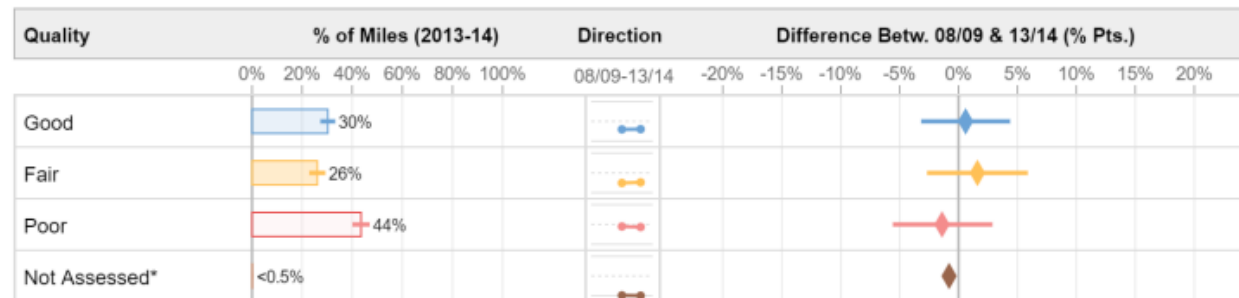
An analysis of the extent to which ALL navigable waters of such State provide for the protection and propagation of a balanced population of shellfish, fish, and wildlife, and allow recreational activities in and on the water.



How can you offer
100% reporting of the
condition of your
waters?

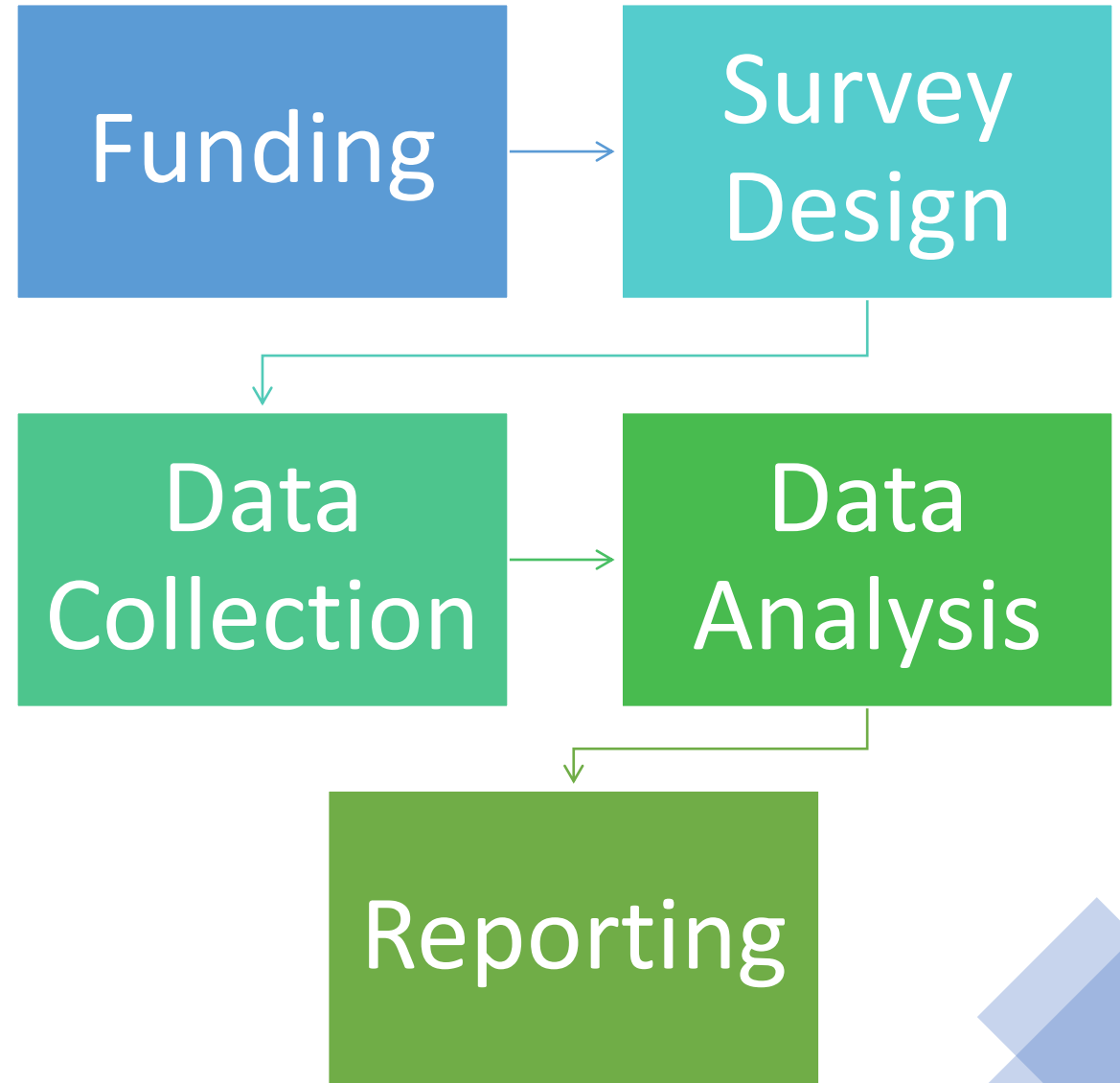
Probability Surveys

- Allows for the extrapolation of information from a subset of the population (sample).
- Provide comprehensive statements about statewide water quality conditions (in line with §305(b) reporting).
- Less time consuming and more cost effective than a Census.
- Provide additions and context to 303(d) list and reference waterbodies.
- Help inform monitoring priorities.



*Reflects a statistically significant change between 2008-09 and 2013-14 (95% confidence).

Components of a Probability Survey



Funding

- Section 106 Allocations
 - Use monitoring program allocation from base 106 grant and/or formula increase
 - Apply monitoring program enhancement allocation (monitoring initiative funds)
- Leverage NARS site allocation with state intensification



What are Intensifications?

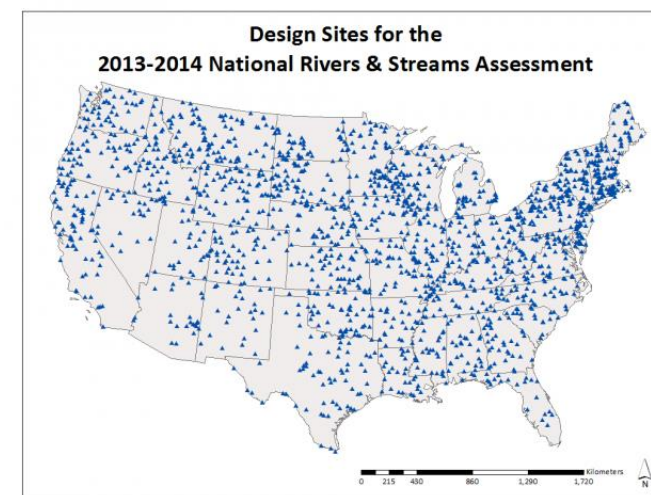
Additional sites or special indicators of interest which an entity wants to sample in conjunction with a NARS survey to complete a desired statistical survey.

Ecoregion/ Multi-State Level

State Scale Surveys

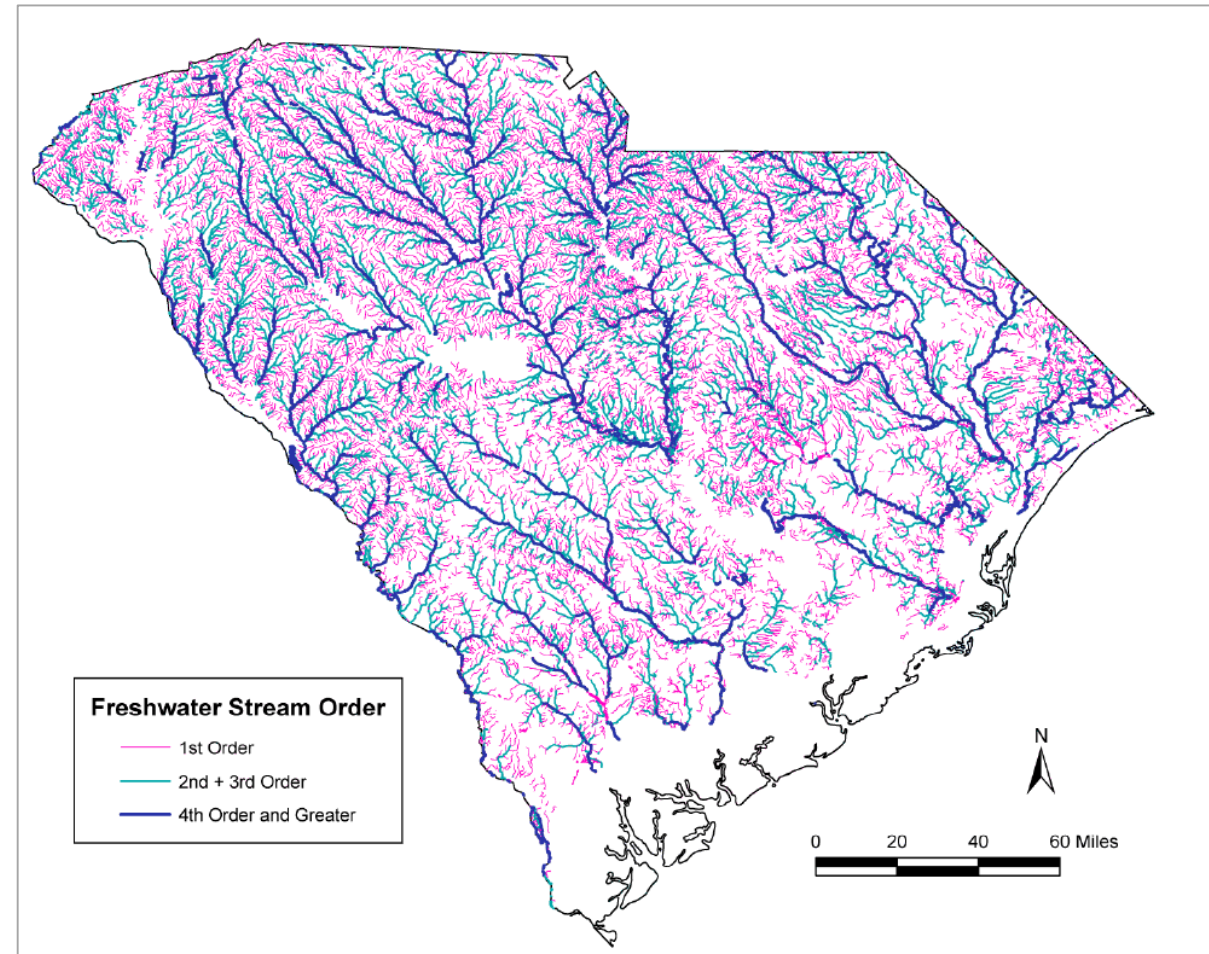
Waterbody Specific

Indicators of Interest




Survey Design

- Developing a Design
 - Define Target Population
 - Sample Frame (GIS Based framework)
- Incorporating state compatible sample frames for an Intensification.
- Develop own survey outside of NARS.



Survey Design Tool

https://survey_design_tool.app.cloud.gov

 Environmental Protection Agency

Environmental Topics Laws & Regulations About EPA

Related Topics: Environmental Topics [Contact Us](#)

Survey Design Tool (v. 1.0.2)

Step 1: Instructions for Use **Step 2: Prepare Survey Design** Step 3: Survey Design Results Step 4: Adjust Survey Weights

Select the Survey Sample Frame

Choose all files of the Sample Frame Required: (.shp, .dbf, .prj, .shx)

Browse... 4 files

Upload complete

Transform CRS to NAD83 / Conus Albers

Design Attributes

Choose Design Type ?

GRTS IRS

Select Attribute Which Contains Strata ?

PerInt

Select Attribute Which Contains Categories ?

Strah_Cat

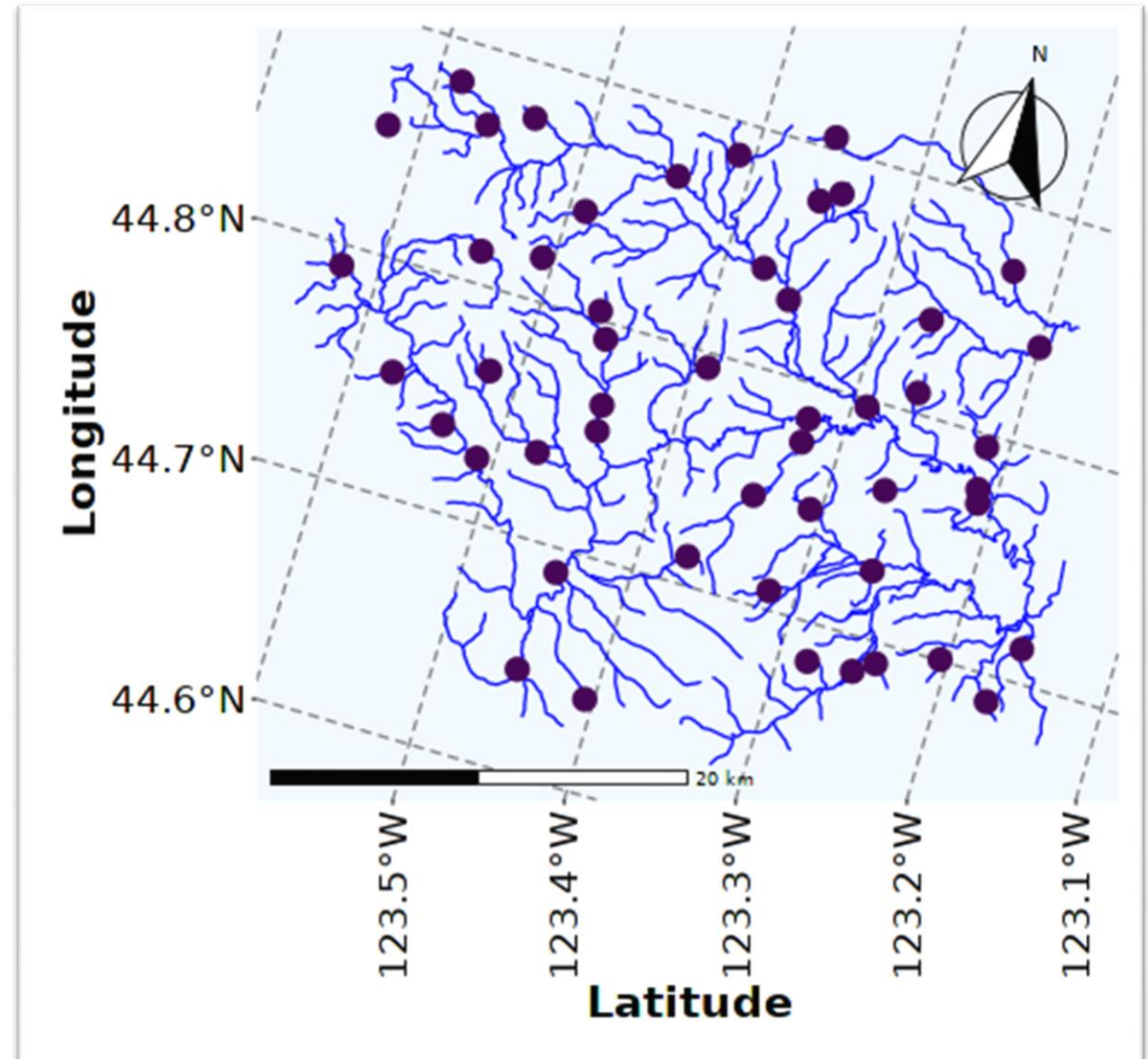
Optional Design Attributes

Calculate Survey Design

Stratum: Intermittent **Stratum: Perennial**

Stratum 1 ?	Base Sites ?	Category 1 ?	Category 1 Sites
Intermittent	0	1st	0
	Replacement Sites ?	Category 2	Category 2 Sites
	0	2nd	0
		Category 3	Category 3 Sites
		3rd	0

Sample Frame Summary ?



Data Collection

- NARS Field collection applications
- Data quality control checks



Data Analysis

- Probability Survey Analyses

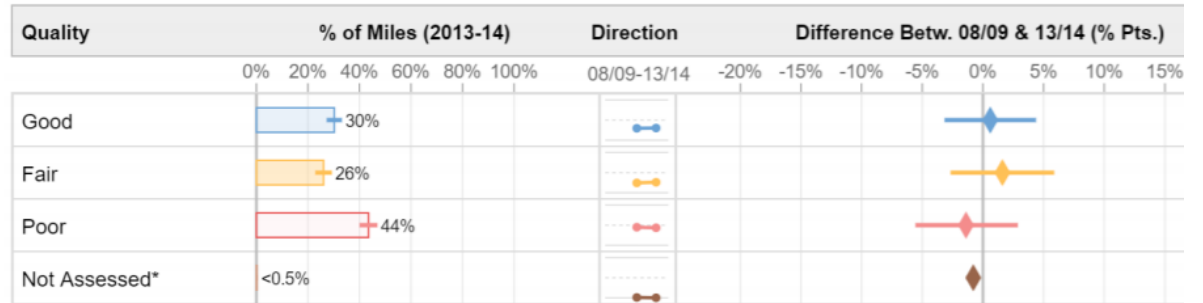
- Population Estimates
- Change and Trend Analyses
- Stressor Extent
- Relative Risks

USEPA/**spsurvey**

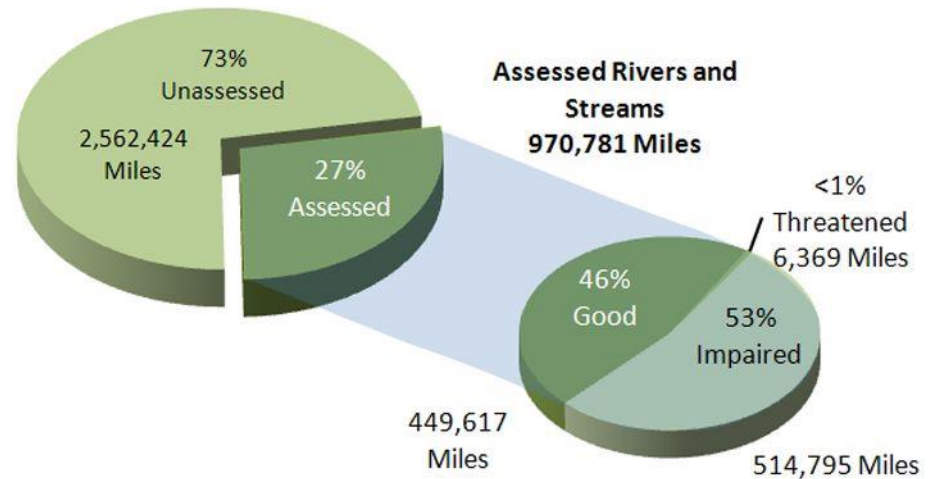
Implements a design-based approach to statistical inference



Macroinvertebrates: NRSA 2013-14 National Results



Total U.S. Rivers and Streams
3,533,205 Miles*



*Total U.S. river and stream miles based on state 2010 Integrated Report.

Population Estimate Tool

<https://github.com/USEPA/NARS-Population-Estimate-Shiny-Tool>

NARS Population Estimate Calculation Tool (v. 2.0)

[Instructions for Use](#)

[Prepare Data for Analysis](#)

[Run Population Estimates](#)

[Run Change Analysis](#)

[Plot Categorical Estimates](#)

[Plot Continuous Estimates](#)

Input file from URL instead of local directory

Select a delimited file for analysis

Browse...

NLA2017_Oregon_NLA_data_

Upload complete

Header

Separator

Comma

Semicolon

Tab

Display

Head

All

Subset data using a single categorical variable

Select site variable

SITE_ID

Select weight variable

WGT_TP

Select up to 10 response variables - All must be either categorical or numeric

CHLA_COND

Check box if performing change analysis or need to subset data by year or cycle for population estimates

Calculate overall (all sites) estimates?

Calculate estimates for subpopulations?

If ANY changes have been made to your choices, you MUST click the button to prepare data for analysis again!

! Type of variance estimate to use (select one)

Local neighborhood variance (recommended, used for NARS, requires site coordinates)

Simple Random Sample (requires stratum but not site coordinates)

Select the X coordinate variable (or longitude) (required only for local neighborhood variance)

XCOORD

Select the Y coordinate variable (or latitude) (required only for local neighborhood variance)

YCOORD

Select the stratum variable in order to calculate variance based on a simple random sample

None

Click HERE to prepare data for analysis

NARS Po

Click to revert back to full dataset or change data display

Type of Analysis

Categorical (for character variables)

Continuous (for numeric variables)

If the Run/Refresh Estimates button is grayed out, return to the Prepare Data for Analysis tab and click the button that says Click HERE to prepare data for analysis

Note that if all values are very small, the results may appear as zeroes. Save and view output file to see the results with full digits.

Run/Refresh estimates

Save Results as .csv file

If output is not as expected, be sure you chose the correct type of analysis (categorical or continuous) for your data.

Warnings

warnings

none

Analysis Output

Type	Subpopulation	Indicator	Category	nResp	Estimate.P	StdError.P	MarginofError.P	LCB95Pct.P	UCB95Pct.P	Estimate.U	StdError.U	MarginofError.U	LCB95Pct.U	UCB95Pct.U
All_Sites	All Sites	CHLA_COND	Good	39.00	86.39	5.82	11.41	74.98	97.80	4163.04	660.74	1295.02	2868.02	5458.06
All_Sites	All Sites	CHLA_COND	Fair	8.00	12.43	5.81	11.39	1.04	23.82	598.95	300.37	588.71	10.23	1187.66
All_Sites	All Sites	CHLA_COND	Poor	2.00	1.18	0.68	1.32	0.00	2.50	56.92	31.57	61.87	0.00	118.79
All_Sites	All Sites	CHLA_COND	Total	49.00	100.00	0.00	0.00	100.00	100.00	4818.91	709.61	1390.80	3428.11	6209.71

Visualize Population Estimates

NARS Population Estimate Calculation Tool (v. 2.0)

[Instructions for Use](#)

[Prepare Data for Analysis](#)

[Run Population Estimates](#)

[Run Change Analysis](#)

[Plot Categorical Estimates](#)

[Plot Continuous Estimates](#)

Choose Categorical Estimate Dataset to Use:

- Upload Estimate Data File
 Current Estimate Data

Select Estimate Type

Proportion Estimates

Select 'Good' Condition Classes

Good

Select 'Fair' Condition Classes

Fair

Select 'Poor' Condition Classes

Poor

Select 'Not Assessed' Condition Classes

Select 'Other' Condition Classes

Add a Plot Title

Optional

Select Indicator

CHLA_COND

Select Subpopulation Group

All_Sites

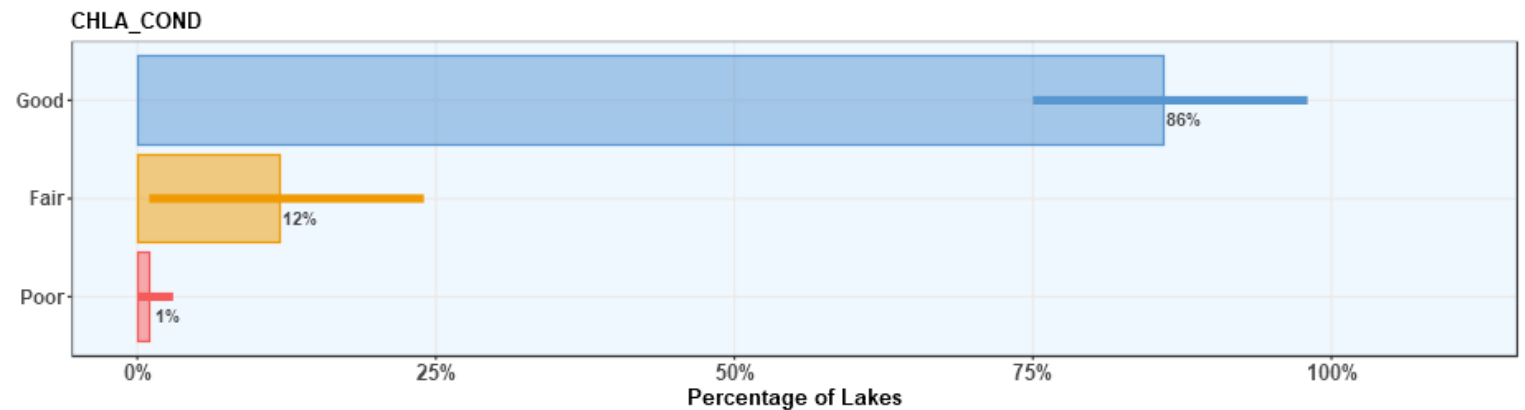
Select Subpopulation

All Sites

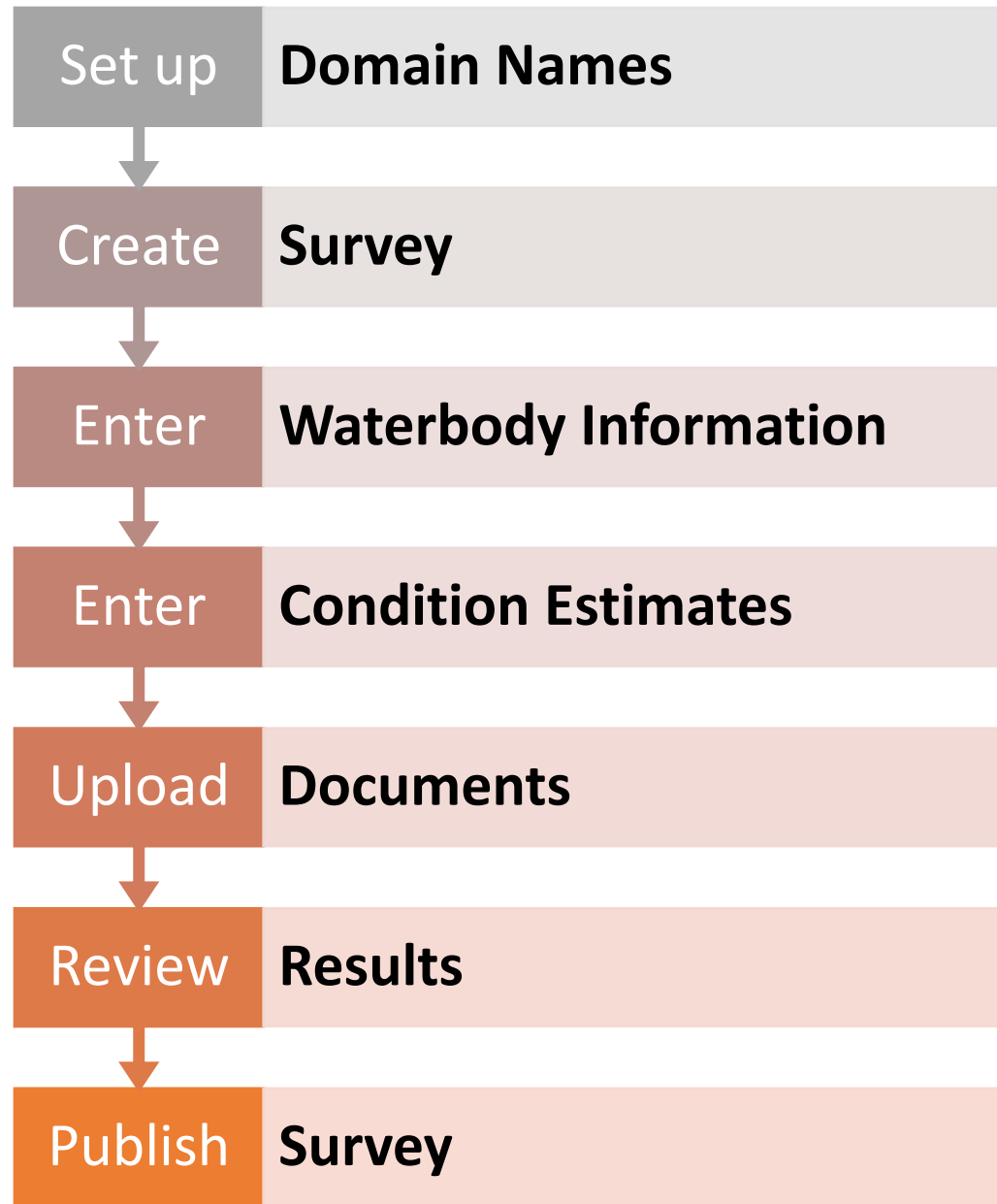
Download Estimate Plot

Categorical Estimates by Population

Add Confidence Limit Values



Reporting Survey Results in ATTAINS



Enter Results

General

Survey Water Type Groups

Documents

Summary

Survey Year 2015 **Organization** TEST_ORG_C
Survey Status Draft
Survey Comment Survey conducted from 2008-2013

Survey Water Type Groups

Survey Water Group ^	Subpopulation	Size	Units	Number of Sites	...
BAY/ESTUARY	Statewide	300	Square Miles	50	
LAKE/RESERVOIR/POND	Statewide	3000	Number of Lakes	50	
STREAM/CREEK/RIVER	Statewide	11000	Miles	50	
WATERSHED	Statewide	6000	Square Miles	600	

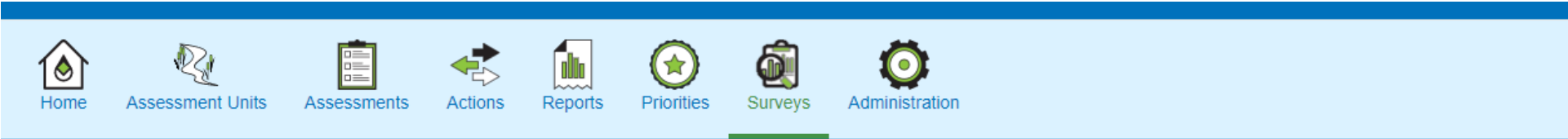
Survey Water Group - STREAM/CREEK/RIVER

Waterbody Type Group STREAM/CREEK/RIVER
Survey Subpopulation Statewide
Size 11000 Miles
Number of Sites 50
Comment Target population is streams less than 4th order.

Survey Water Group - Use Parameters

Type	Use or Condition ^	Stressor ^	Category ^	Statistic	Metric Value	Margin of Error	Confidence Level	...
Use / Stressor	Aquatic Life Use	ALGAE	Fair	Condition Estimate	25	± 15	95	
Use / Stressor	Aquatic Life Use	ALGAE	Good	Condition Estimate	50	± 15	95	
Use / Stressor	Aquatic Life Use	ALGAE	Poor	Condition Estimate	25	± 15	95	

Publish Survey



State Statistical Survey - 2015

Menu Return to list



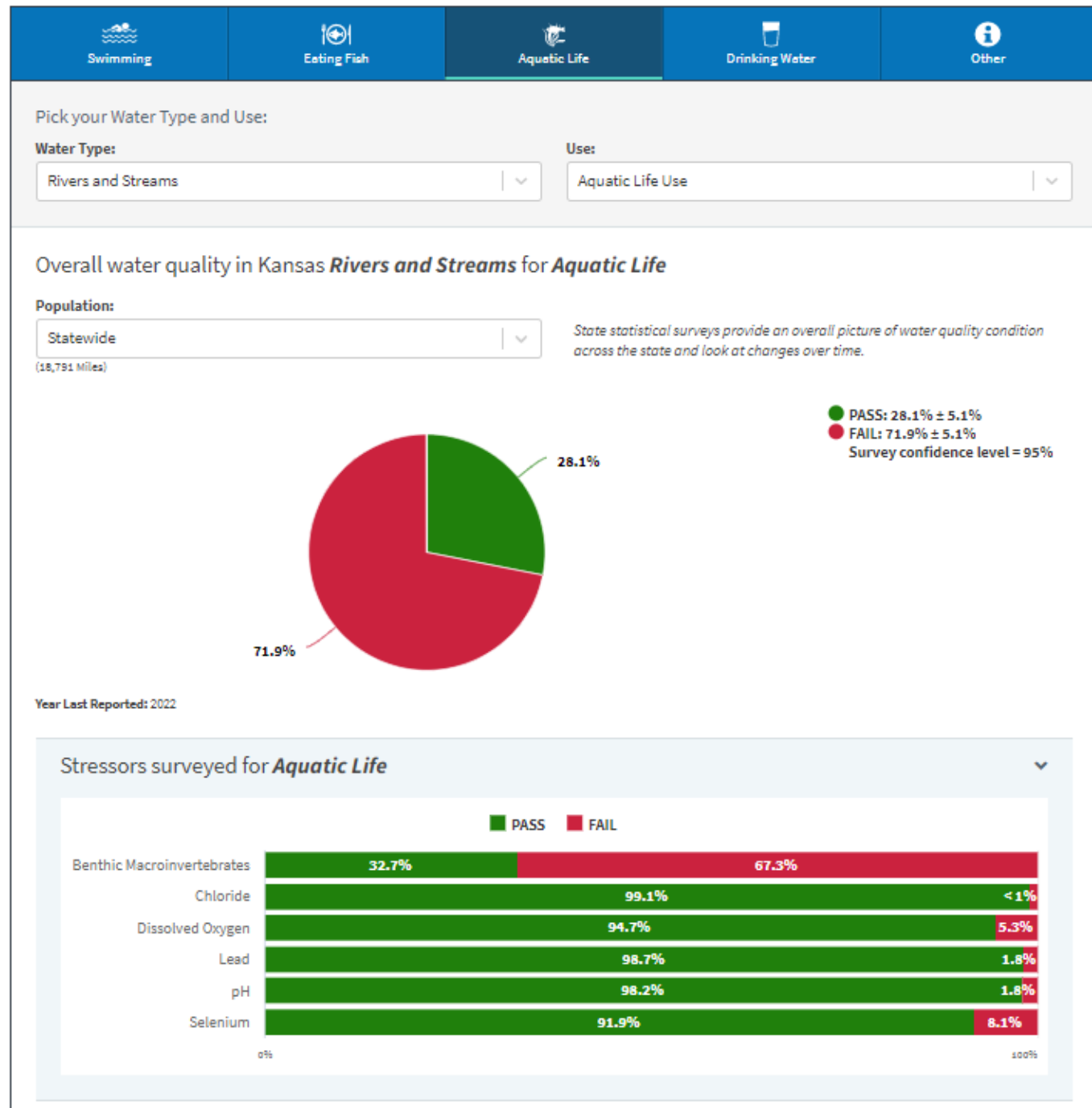
Survey Year 2015 **Organization** TEST_ORG_C
Survey Status Draft
Survey Comment Survey conducted from 2008-2013

Survey Water Type Groups

Survey Water Group ^	Subpopulation	Size	Units	Number of Sites	...
BAY/ESTUARY	Statewide	300	Square Miles	50	
LAKE/RESERVOIR/POND	Statewide	3000	Number of Lakes	50	
STREAM/CREEK/RIVER	Statewide	11000	Miles	50	
WATERSHED	Statewide	6000	Square Miles	600	

Survey Water Group - STREAM/CREEK/RIVER	
Waterbody Type Group	STREAM/CREEK/RIVER
Survey Subpopulation	Statewide
Size	11000 Miles
Number of Sites	50
Comment	Target population is streams less than 4th order.

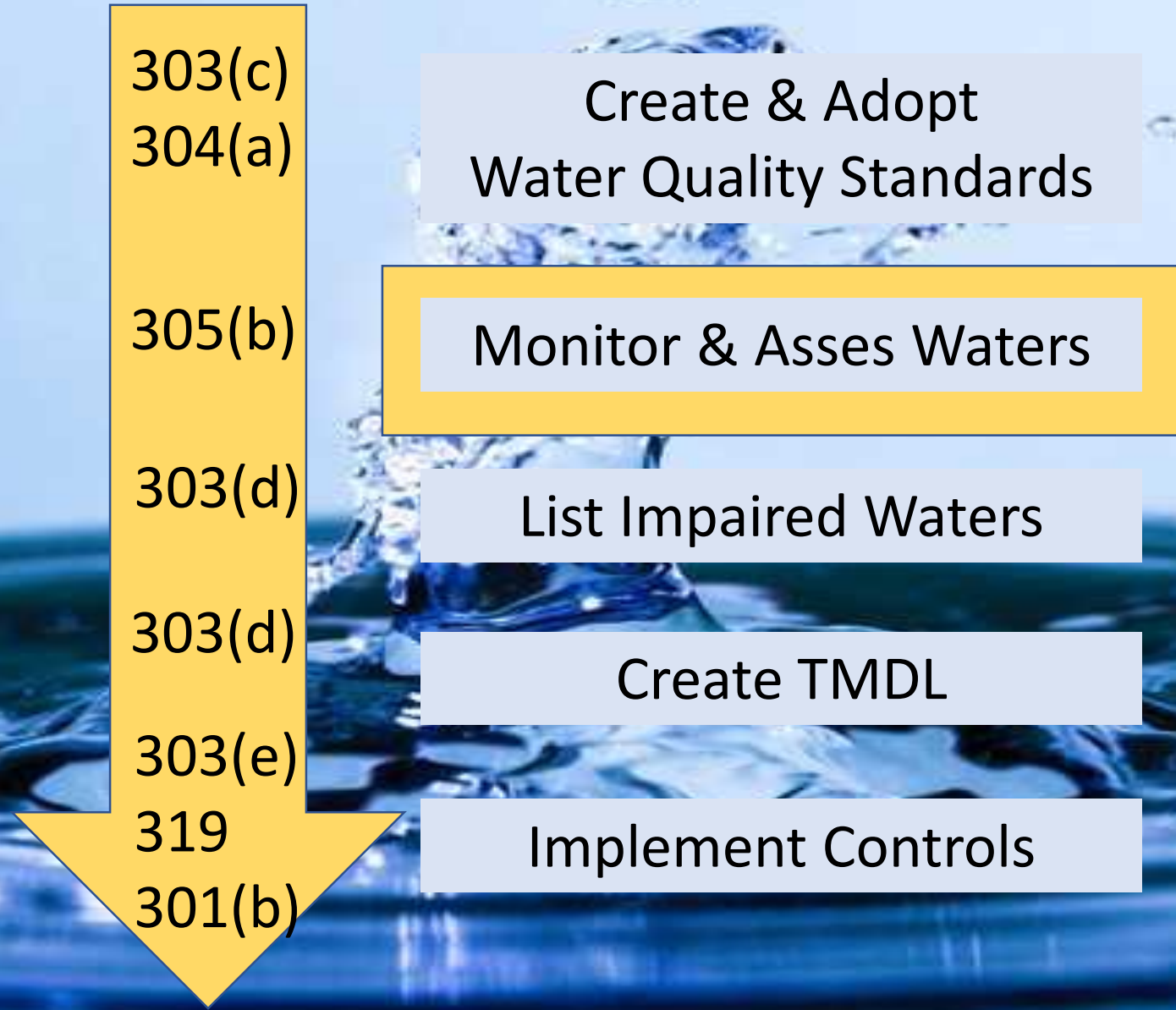
Publish State Scale Survey Results to Hows My Waterway





Thank You
Need Survey Support?
Stillings.Garrett@epa.gov

Framework for Restoring Polluted Waters



Visualize CDF Estimates

Choose Cumulative Distribution Function (CDF) Estimate Dataset to Use:

- Upload Estimate Data File
 Current Estimate Data

Choose CDF Analysis file ?

Browse... NLA17TP_CDF.csv

Upload complete

Select Estimate Type ?

Proportion Estimates

Add a Plot Title


Optional

Add Indicator Units

Optional

Define Resource Type/Unit ?

Resource

 Plot Continuous Estimates

Select Indicator

TP

Indicator Threshold (optional)

Select Population

Regional

Add Confidence Limits

Log Scale X-Axis


Add/Remove Subpopulations

Region_01 Region_02 Region_03

Region_04

CDF Estimates ?

NOTE: Plotting and downloading may take a while if there are multiple subpopulations. PLEASE BE PATIENT.

 Download CDF Plot

