



Using R to Analyze Data from Probabilistic Monitoring in Oklahoma

Jean Lemmon

Background

- 2000 - Ambient monitoring program known as the Small Watershed Rotating Basin Monitoring Program
 - 414 USGS 11-digit HUC basins collated into 11 planning basins
 - Sampled every five weeks for two years with statewide coverage of all sites in five years
- 2008 – Probabilistic component
 - 50 sites in each of 5 major basins – 250 sites in State monitored in 5 years
 - Sites taken in order from master site list generated by Tony Olsen with the EPA Corvallis Lab
 - Water chemistry, fish, macroinvertebrates, and habitat monitored at same time as Rotating Basin fixed sites
- 2013 - Data analysis using R Open Source (Free) Software – scripts supplied by EPA



Probabilistic Site List

- 150 sites generated for 50 targeted sites
- Start with #1 and work down list using sites in order until 50 are available
- Possible reasons for not monitoring a site:
 - Landowner permission denied
 - Completely inaccessible
 - Dry
 - In the middle of an impoundment
- Data collected
 - Chemistry and bacteria
 - Habitat
 - Fish
 - Benthic macroinvertebrates (both winter and summer)



R Software

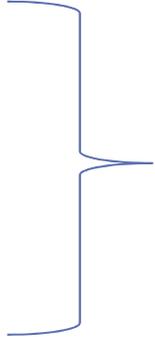
- Evaluating probabilistic data
- Define “Good”, “Fair”, and “Poor” conditions for:
 - Stressors (chemical and physical parameters)
 - Indicators (fish/bugs)
- Re-weight the sites actually monitored
- Run data through R open software (FREE) programs
 - Extent Estimates (% of stream length)
 - Relative Risk (strength of association between elevated stressor and degraded biota)
 - Attributable Risk (estimate of stressor effect on indicator)

Why do this?

- Extent Estimates give the percent of stream length in “Good,” “Fair,” or “Poor” condition
- Relative Risk assesses the relative importance of multiple stressors (Which stressors should be the major focus for remediation, restoration or protection?)
- Attributable Risk estimates the percent of improvement (if it were possible to completely address a stressor)



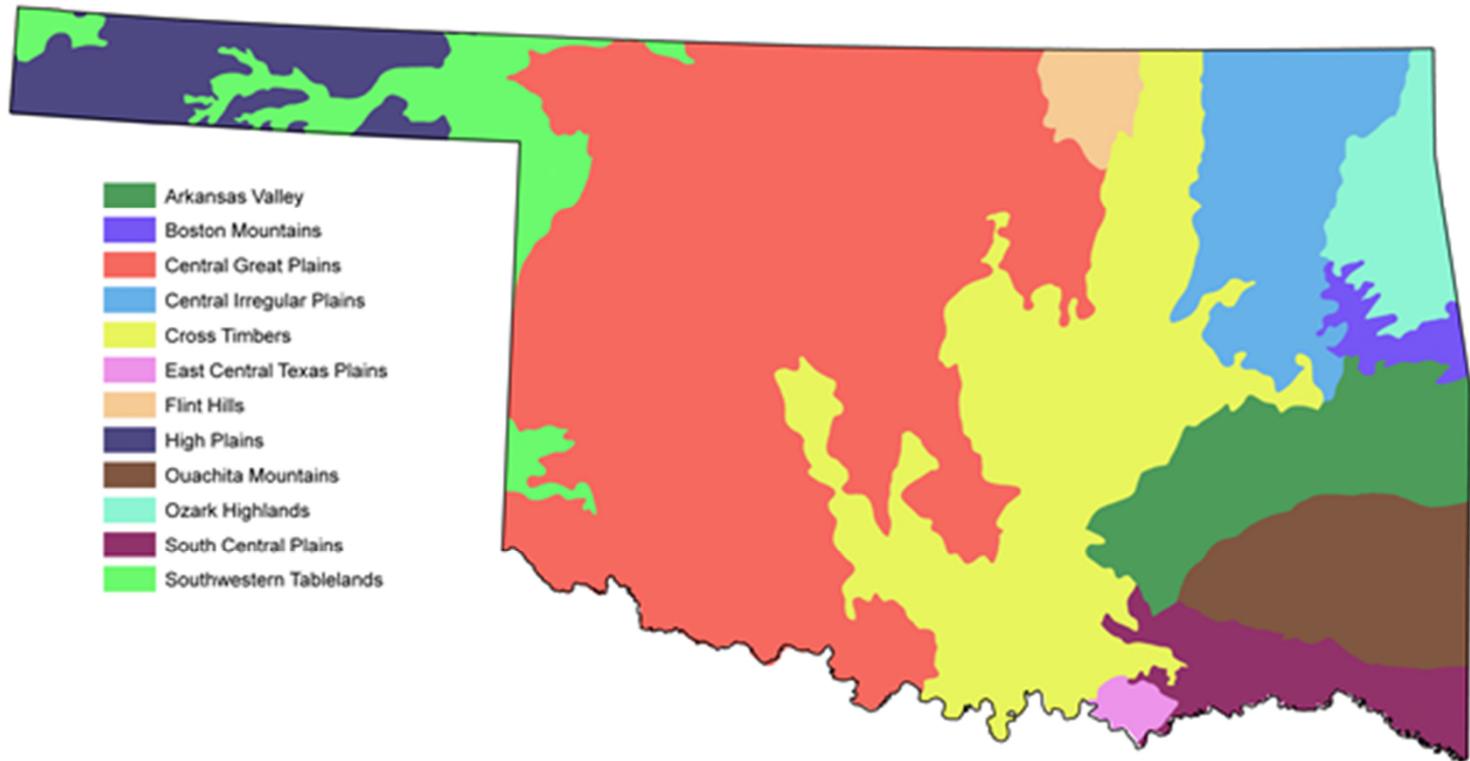
Stressors (chemical and physical parameters)

- Total Nitrogen
 - Total Phosphorus
 - Conductivity
 - Turbidity
- 
- Identified by EPA as most common problems
- Other chemical parameters
 - Habitat observations
 -

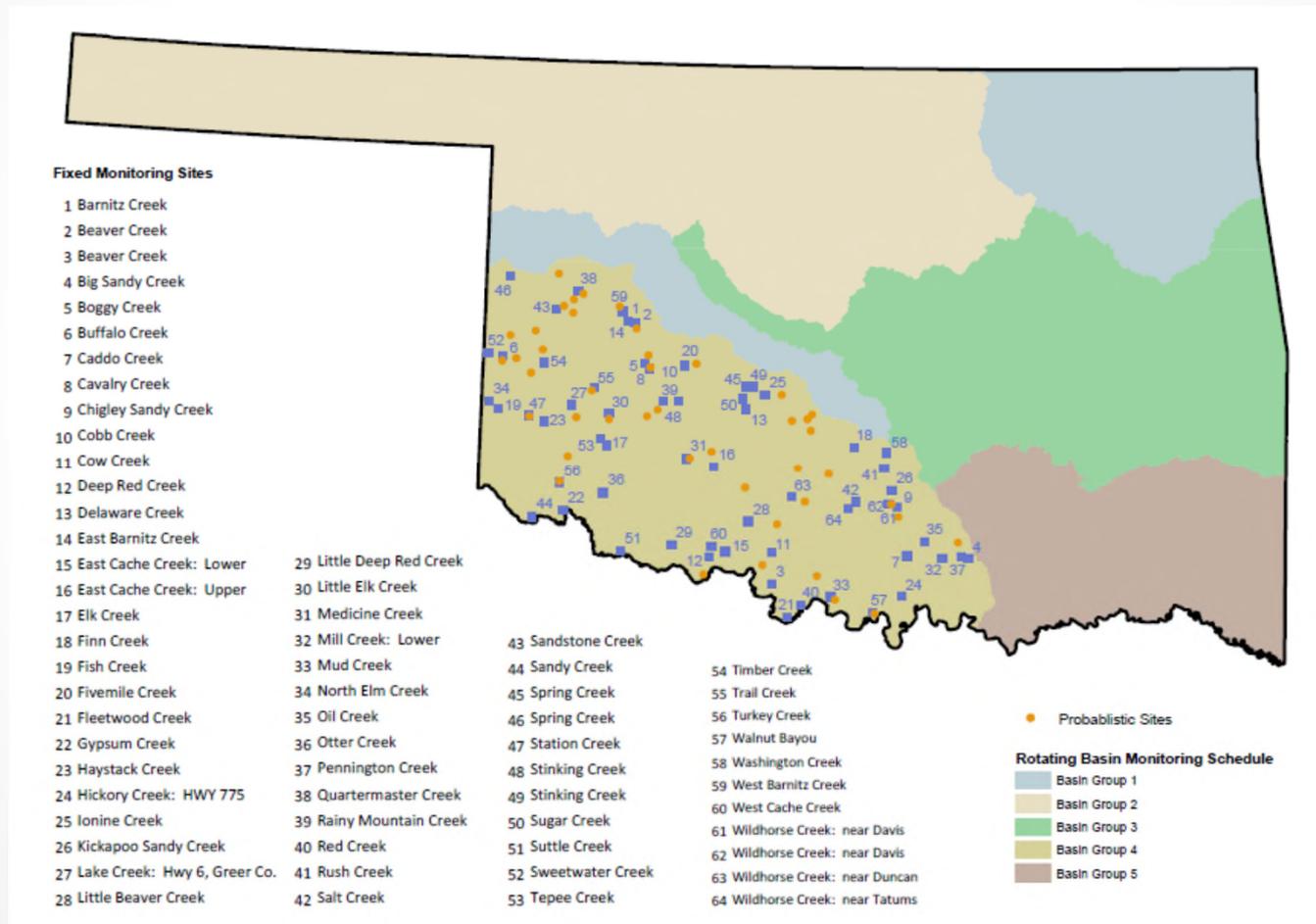
Indicators

- Fish Index of Biotic Integrity (IBI) calculated using:
 - Number of species
 - Number of sensitive benthic species
 - Number of sunfish species
 - Number of intolerant species
 - Proportion tolerant individuals
 - Proportion insectivorous cyprinid individuals
 - Proportion individuals as lithophilic spawners
- Benthic Macroinvertebrate IBI calculated using:
 - Taxa richness
 - Modified HBI (using NC tolerance values)
 - EPT/Total
 - EPT Taxa
 - % Dominant 2 Taxa
 - Shannon-Weaver

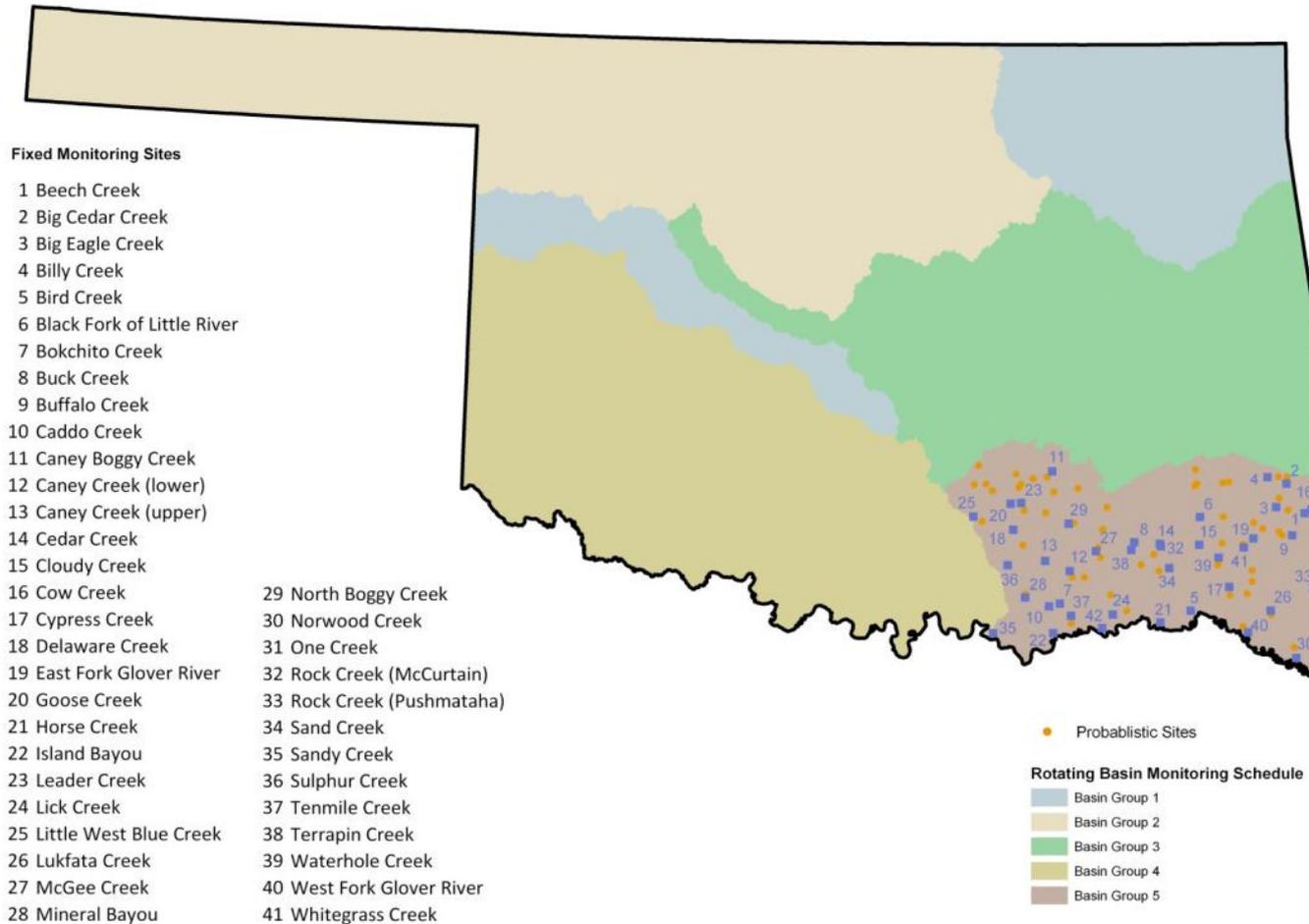
Oklahoma Ecoregions

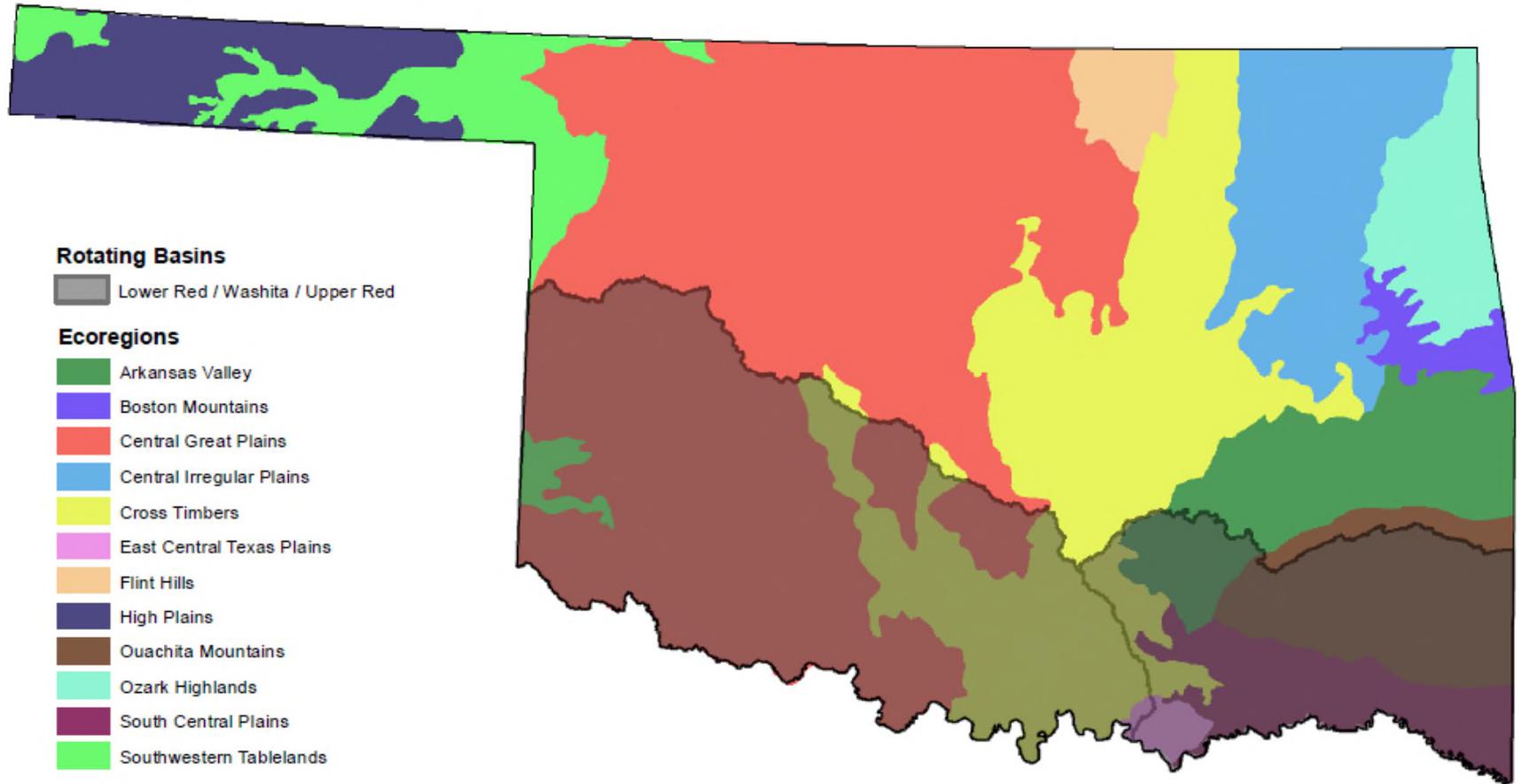


Upper Red/Washita Basin



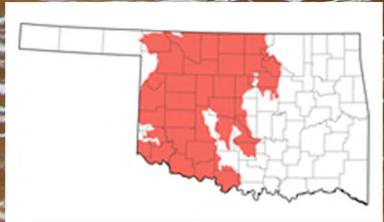
Lower Red River Basin





Southwestern Tablelands





Central Great Plains



Cross Timbers

Arkansas Valley





Ouachita Mountains

South Central Plains

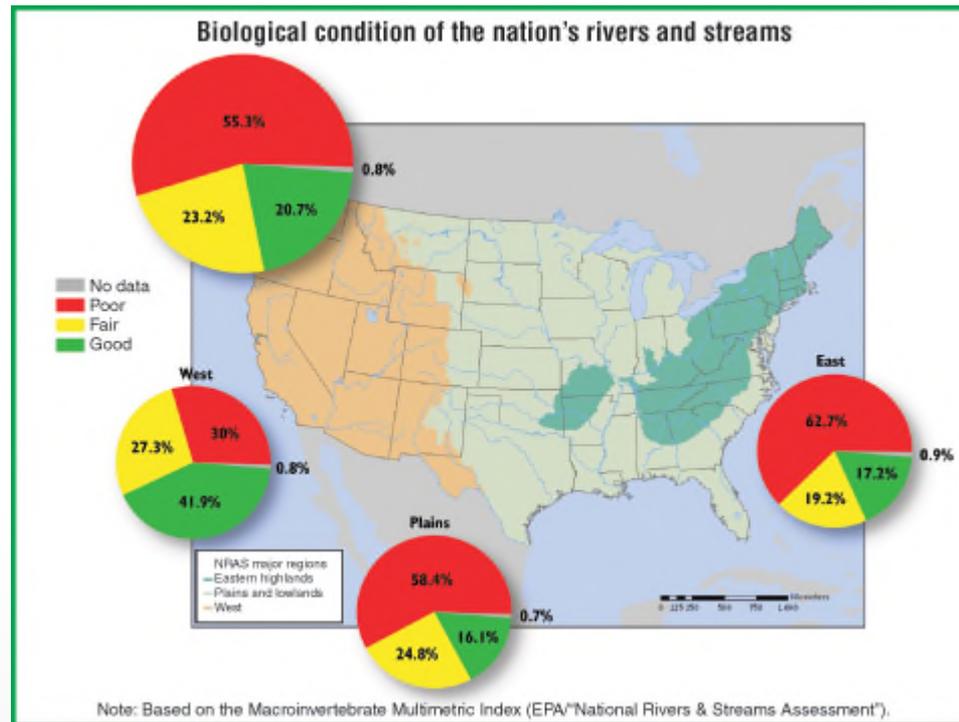


Define “Good, Fair, Poor”

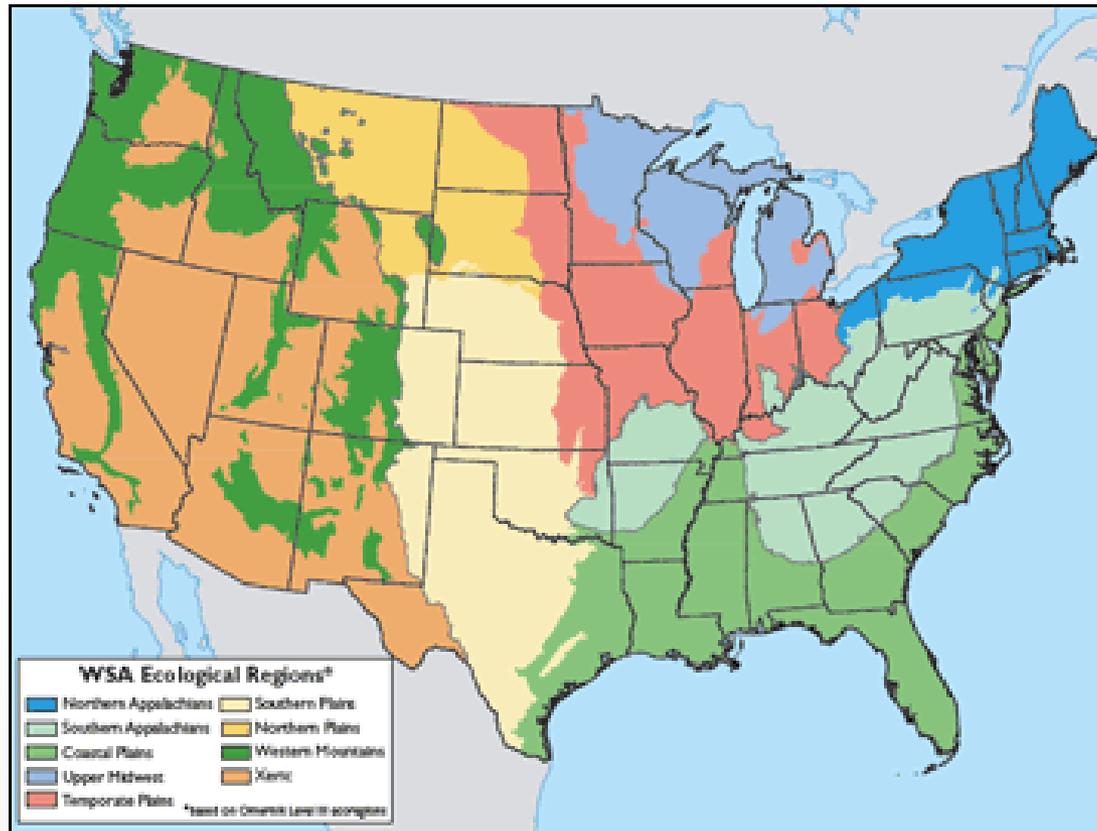
- EPA National Rivers and Streams Assessment (NRSA)
- EPA Wadeable Streams Assessment (WSA) by ecoregion
- % of Oklahoma Conservation Commission (OCC) high quality reference conditions
- Comparison with 1 or 2 times the standard deviation of OCC high quality reference conditions



EPA National Rivers & Streams Assessment (NRSA)



EPA Wadeable Streams Assessment (WSA)



OCC HQ Reference

- High quality reference sites chosen in each ecoregion because they have a healthy biological community and good habitat
- Percent of high quality reference conditions
- Comparison with 1 or 2 times the standard deviation of OCC high quality reference conditions
- Is a condition “Poor” if it does not meet State of Oklahoma standards?



Total Nitrogen Poor/Fair

Ecoregion	NRSA (mg/L)	WSA (mg/L)	OCC HQ Ref (2 * StdDev)
Southwest Tablelands	1.570	1.050	
Central Great Plains	1.570	1.600	3.422
Cross Timbers	1.570	0.900	1.699
Arbuckle Uplift	1.570	1.500	1.699
South Central Plains	2.078	0.750	0.879
Ouachita Mountains	0.535	0.450	0.834
Arkansas Valley	0.535	0.683	3.057

The average OCC high quality reference site does not always meet the NRSA and WSA conditions.

Total Phosphorus

Poor/Fair

Ecoregion	NRSA (mg/L)	WSA (mg/L)	OCC HQ Ref (2 * StdDev)
Southwest Tablelands	0.095	0.055	
Central Great Plains	0.095	0.130	0.24
Cross Timbers	0.095	0.110	0.145
Arbuckle Uplift	0.095	0.050	0.145
South Central Plains	0.108	0.070	0.091
Ouachita Mountains	0.024	0.025	0.047
Arkansas Valley	0.024	0.060	0.168

The average OCC high quality reference site does not always meet the NRSA and WSA conditions.

Conductivity Poor/Fair

Ecoregion	NRSA ($\mu\text{S}/\text{cm}^2$)	WSA ($\mu\text{S}/\text{cm}^2$)	OCC HQ Ref (2 * StdDev)
Southwest Tablelands	2000	2300	
Central Great Plains	2000	2925	3839
Cross Timbers	2000	1000	855
Arbuckle Uplift	2000	1000	1009
South Central Plains	1000	500	304
Ouachita Mountains	1000	500	97
Arkansas Valley	1000	500	2973

The average OCC high quality reference site does not always meet the NRSA and WSA conditions.

Turbidity Poor/Fair

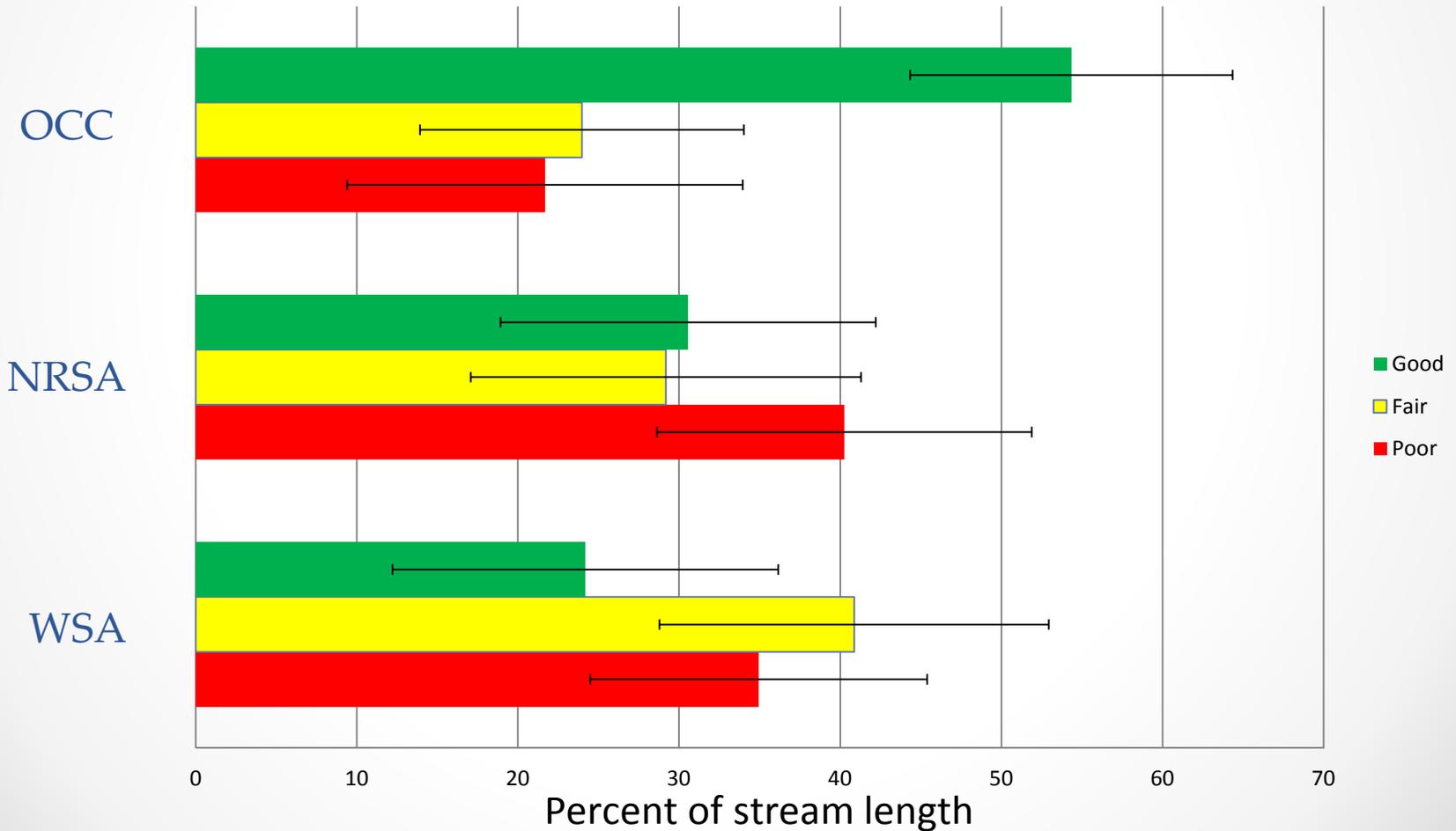
Ecoregion	WSA (NTU)	OCC HQ Ref (2 * StdDev)
Southwest Tablelands	20	
Central Great Plains	45	38
Cross Timbers	40	38
Arbuckle Uplift	7	21
South Central Plains	20	19
Ouachita Mountains	10	18
Arkansas Valley	10	28

The average OCC high quality reference site does not always meet the NRSA and WSA conditions.

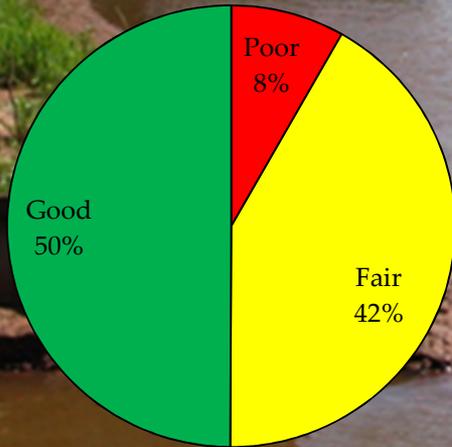
Chemistry Comparison

siteID	Ecoregion	Total N			Total P			Conductivity			Turbidity	
		Total Ni (NRSA)	Total Ni (WSA)	Total Ni (OCC)	Total P (NRSA)	Total P (WSA)	Total P (OCC)	Conductivity (NRSA)	Conductivity (WSA)	Conductivity (OCC)	Turbidity (WSA)	Turbidity (OCC)
OKR09730-003	Cross Timbers	Fair	Poor	Good	Poor	Fair	Fair	Poor	Poor	Poor	Poor	Poor
OKR09730-005	Central Great Plains	Poor	Poor	Poor	Poor	Poor	Fair	Poor	Poor	Fair	Poor	Poor
OKR09730-008	Central Great Plains	Poor	Poor	Good	Poor	Poor	Fair	Poor	Poor	Fair	Good	Good
OKR09730-012	Central Great Plains	Fair	Fair	Good	Good	Good	Good	Fair	Fair	Good	Good	Good
OKR09730-014	Central Great Plains	Fair	Fair	Good	Good	Good	Good	Poor	Poor	Fair	Good	Good
OKR09730-019	Central Great Plains	Good	Good	Good	Poor	Fair	Good	Fair	Fair	Good	Good	Good
OKR09730-020	Central Great Plains	Fair	Fair	Good	Fair	Good	Good	Fair	Fair	Good	Good	Good
OKR09730-025	Central Great Plains	Fair	Good	Good	Fair	Good	Good	Poor	Poor	Fair	Good	Good
OKR09730-026	Central Great Plains	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good	Good
OKR09730-027	Cross Timbers	Good	Good	Good	Fair	Fair	Good	Good	Good	Good	Poor	Poor
OKR09730-031	Cross Timbers	Fair	Fair	Good	Fair	Fair	Good	Fair	Poor	Poor	Poor	Poor
OKR09730-041	Central Great Plains	Fair	Fair	Good	Poor	Poor	Fair	Fair	Fair	Good	Good	Good
OKR09730-045	Central Great Plains	Fair	Fair	Good	Poor	Poor	Poor	Poor	Poor	Poor	Poor	Poor
OKR09730-046	Cross Timbers	Fair	Poor	Fair	Poor	Poor	Fair	Good	Fair	Fair	Fair	Good
OKR09730-047	Cross Timbers	Fair	Poor	Good	Fair	Fair	Good	Good	Good	Good	Poor	Poor

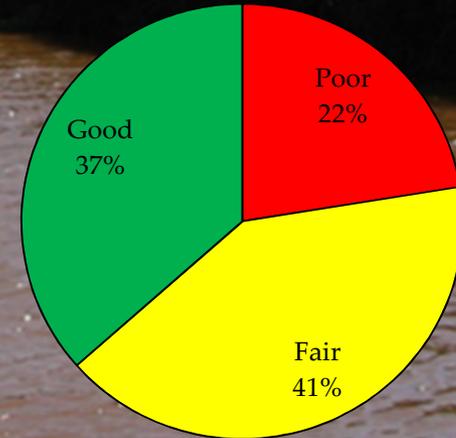
Conductivity Extent Estimate



Upper Red / Washita



Macroinvertebrate Condition
(% of Total Stream Length)



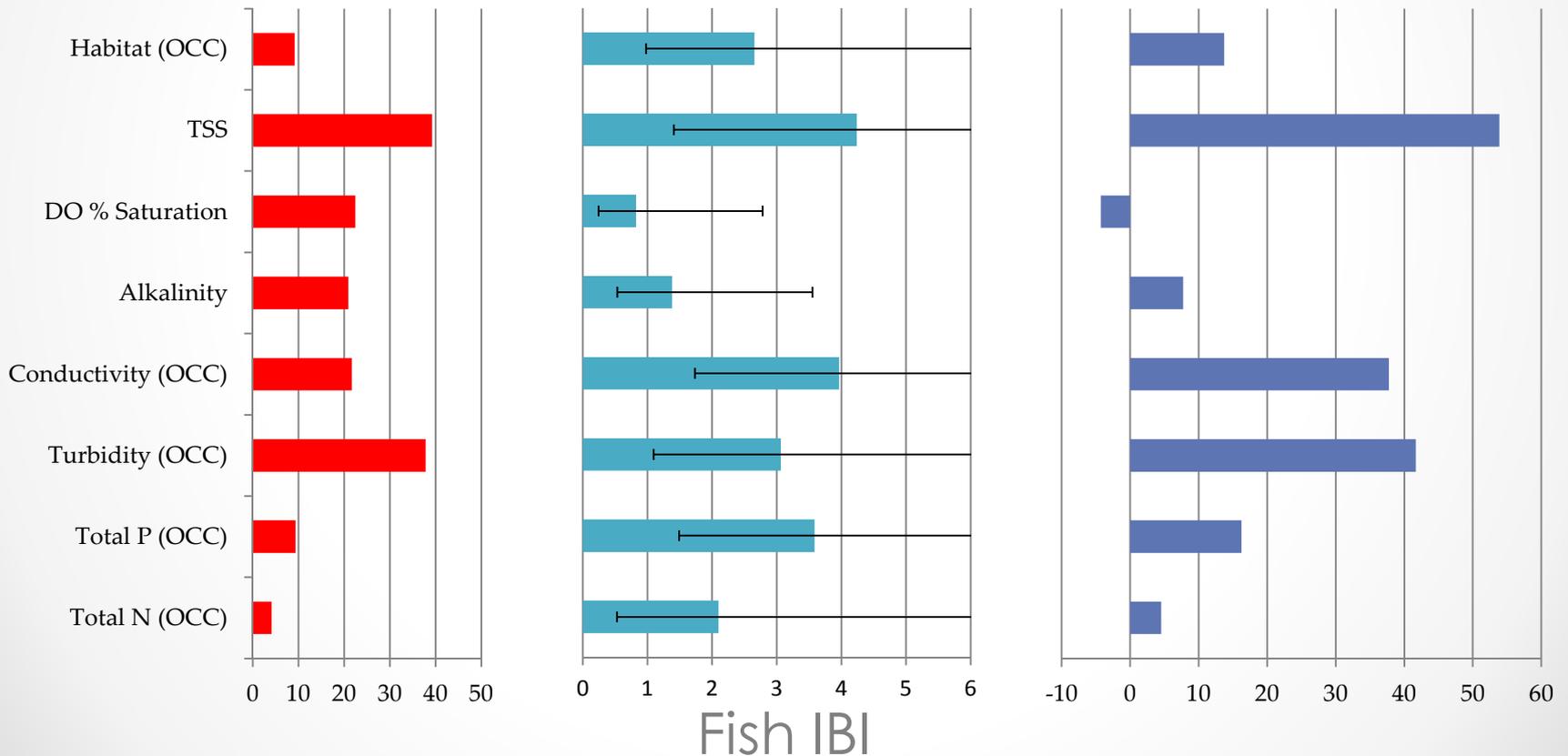
Fish Condition
(% of Total Stream Length)

Upper Red/Washita

Extent Estimates

Relative Risk

Attributable Risk



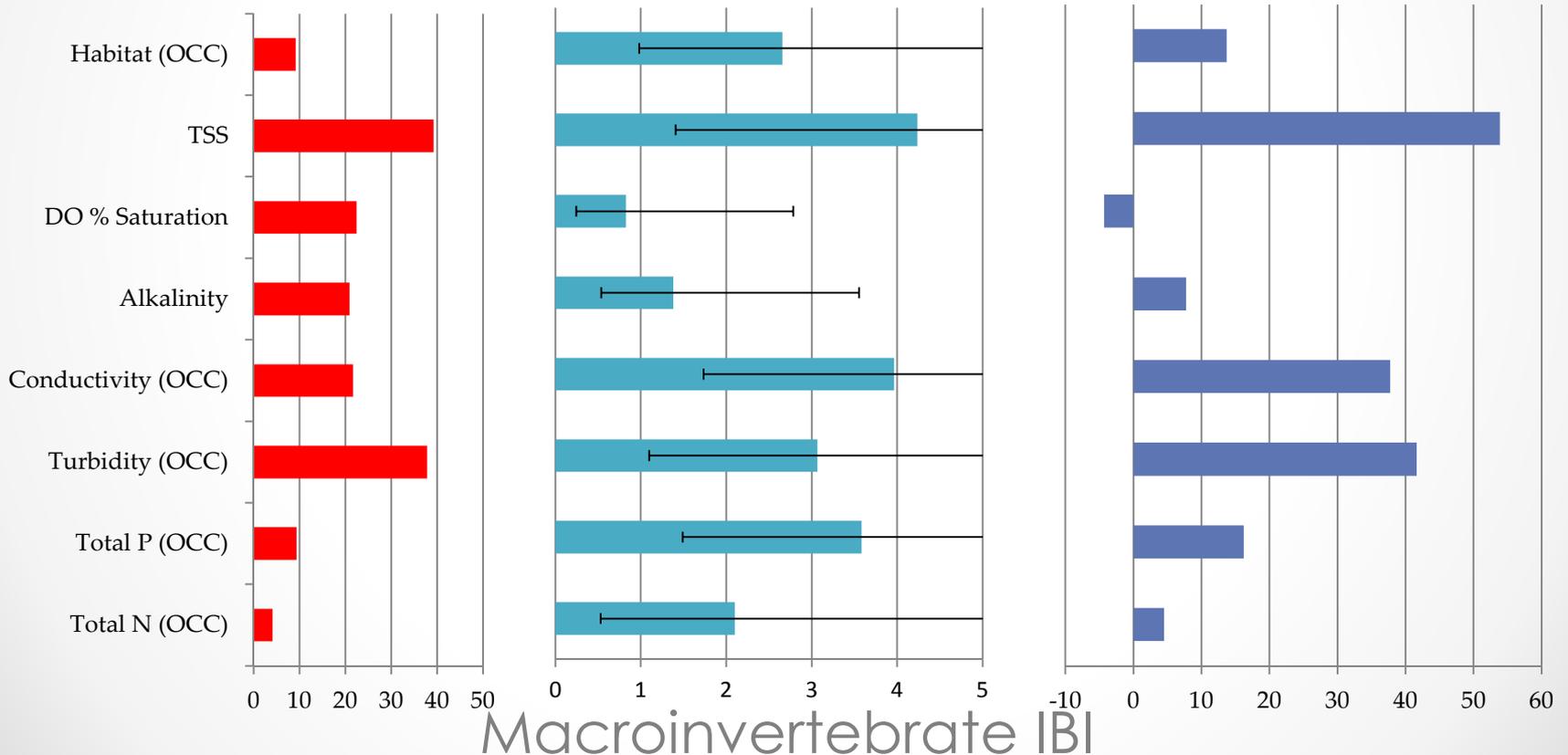
(Extent stream length in Poor condition = 22.5%)

Upper Red/Washita

Extent Estimates

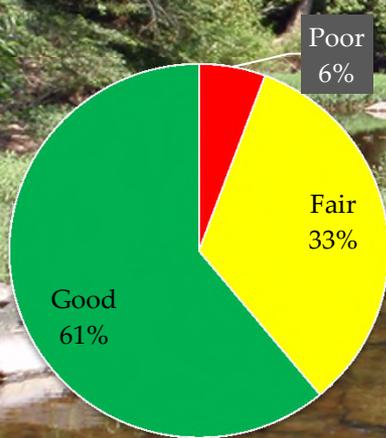
Relative Risk

Attributable Risk

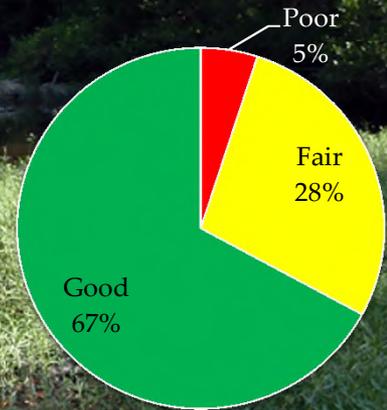


Macroinvertebrate IBI
(Extent of stream length in Poor condition = 8.2%)

Lower Red River Basin



Macroinvertebrate Condition
(% of Total Stream Length)



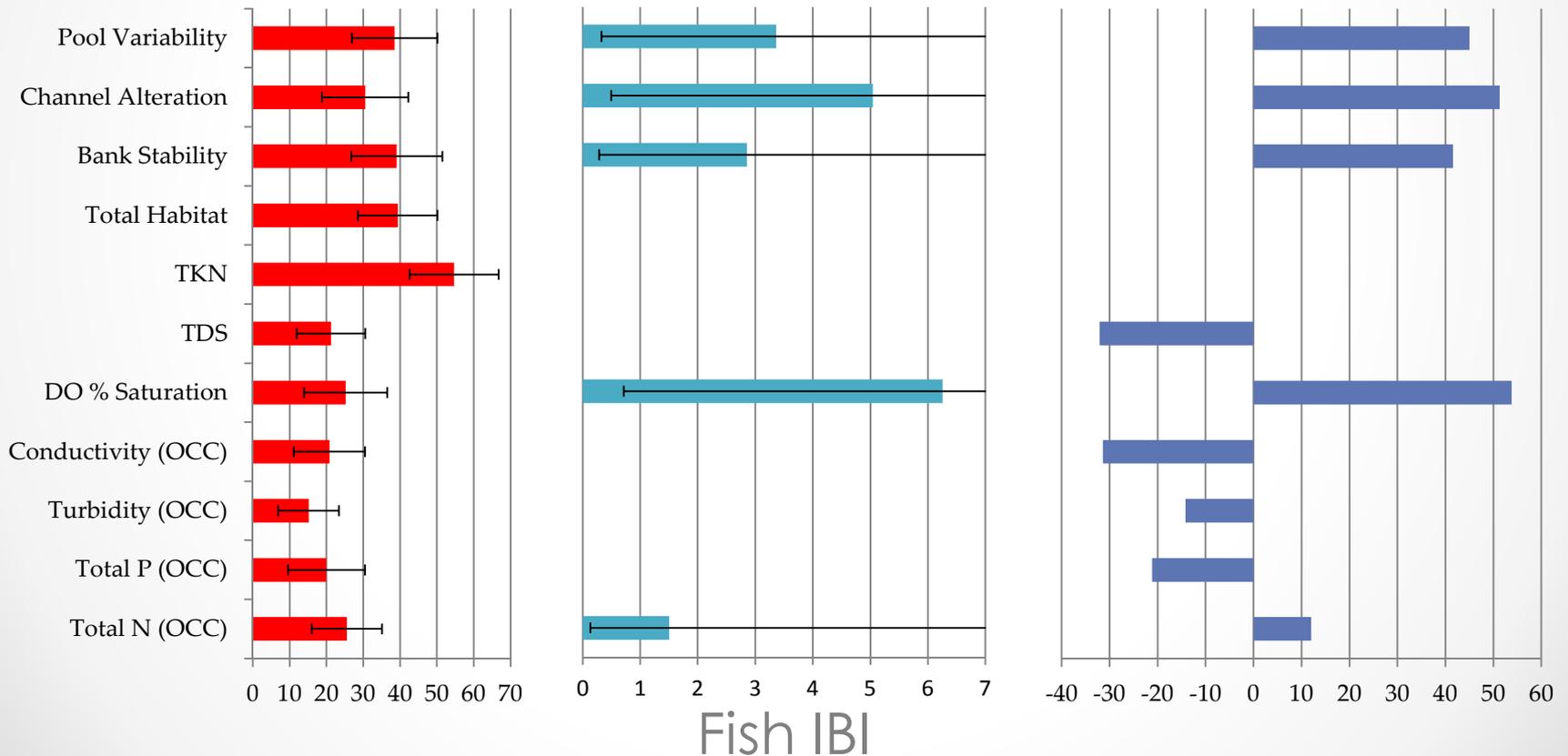
Fish Condition
(% of Total Stream Length)

Lower Red River Basin

Extent Estimates

Relative Risk

Attributable Risk



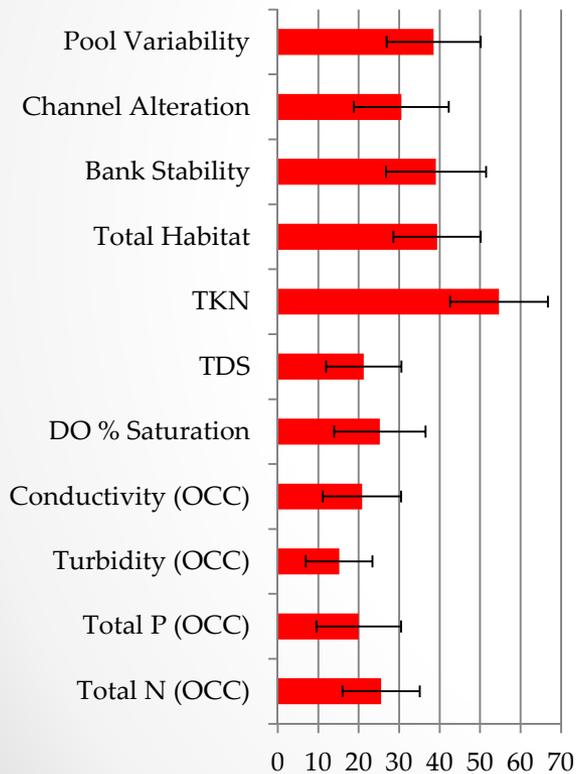
(Extent of stream length in Poor condition = 5.0%)

Lower Red River Basin

Extent Estimates

Relative Risk

Attributable Risk



What happened?

Macroinvertebrate IBI

(Extent of stream length in Poor condition = 5.68%)

Lower Red River Basin Relative Risk

➤ `summary(subset(all.dat.rr, select=c('MIBI_Cond', 'TP_NRSA_Cond')));`

MIBI_Cond
NotPoor:26
Poor : 1
NA's :25

TN_NRSA_Cond
NotPoor:26
Poor :23
NA's : 3

MIBI_Cond
NotPoor:26
Poor : 1
NA's :25

TP_NRSA_Cond
NotPoor:27
Poor :22
NA's : 3

TP_EcoR_Cond
NotPoor:29
Poor :22
NA's : 1

TP_NRSA_Cond

Conductivity_EcoR
NotPoor:43
Poor : 8
NA's : 1

Conductivity_OCC
NotPoor:40
Poor :11
NA's : 1

Turbidity_OCC
NotPoor:35
Poor :18
NA's : 1

Conductivity_OCC
NotPoor:43
Poor : 8
NA's : 1



Conclusions

Upper Red River / Washita Basin

- Top three stressors
 - Total suspended solids
 - Turbidity
 - Conductivity
- Landuse is predominantly cropland and rangeland
- Possible areas of focus to protect our streams
 - Riparian area protection
 - Alternative water supply and keeping livestock out of the streams

Conclusions

Lower Red River Basin

- Top stressors
 - Dissolved Oxygen
 - Channel alteration
 - Bank stability
 - Pool variability
- Landuse is primarily natural forest and silviculture
- Possible areas of focus to protect our streams
 - Stream bank protection
 - Riparian area protection

But drought!



R Software

- Relatively easy to use
- Challenge: Identifying “Good, Fair, Poor” conditions
- Helpful hints:
 1. Work in Excel or Minitab or a familiar program
 2. Save the file as CSV
 3. Use R scripts
 4. Save the results as CSV
 5. Work in Excel or Minitab or a familiar program to graph the results

Conclusions

- Probabilistic monitoring is useful for a quick and relatively inexpensive method to identify stressors putting streams at risk (visit each site once or twice)
- Ambient monitoring provides much more information and allows identification of high quality reference conditions (visit each site twenty times in two years)

A serene landscape photograph of a calm lake at dawn. The water is still, reflecting the sky and the surrounding trees. A thick layer of mist hangs over the water's surface. In the background, a dense forest of trees is visible, with some showing early autumn colors. The word "Questions?" is written in a white, elegant serif font across the center of the image. The overall mood is quiet and contemplative.

Questions?

Questions?

No time for questions!

