ESTIMATION OF TOTAL RED SNAPPER ABUNDANCE IN LOUISIANA AND ADJACENT FEDERAL WATERS

Study Design
and Responses to SSC Comments

OBJECTIVES OF THIS PRESENTATION

- Clearly describe the study design
- Explain the rationale and events that resulted in the decisions regarding our final approach
- Along the way, discuss how and why we think major issues for potential bias were avoided; also, how less important issues may still exist

ABUNDANCE

- This study was designed for model-based inference of Red Snapper abundance from data obtained by field surveys for two separate responses:
  - Total fish density (TFD) from hydroacoustic surveys
  - The proportion of this TFD that were Red Snapper (PropRS) from surveys using submerged rotating video cameras (SRVs)

AGE, LENGTH, AND GROWTH

- Age, length, and growth were collected from fish at the same sites chosen to sample for abundance estimation
- Vertical hook and line sampling was used for discrete habitat types
- Bottom longlines were uncharacterized bottom (UCB) habitat

SAMPLING DESIGN

- The primary objective of the site selection process was to choose samples representative of the population while reducing cost to within budgetary constraints
- As such, the selection process sacrificed randomness for some habitat types
- Site selection was noninformative; i.e., it was not influenced by our preconceived notion of Red Snapper distribution
- However, some sites were selected purposively to ensure representation of certain habitats and opportunistically to reduce costs

STRATIFICATION

HORIZONTAL

- Three longitudinal regions:
  - West
  - Central
  - East
- Ultimately three depth zones:
  - 10-25 m
  - 25-45 m
  - 45-150 m

HABITAT TYPE

- Initially, five habitat types:
  - Standing platforms
  - Natural banks
  - Pipeline crossings
    - Ultimately, not used but it influenced other site selections
  - Artificial reefs
  - UCB

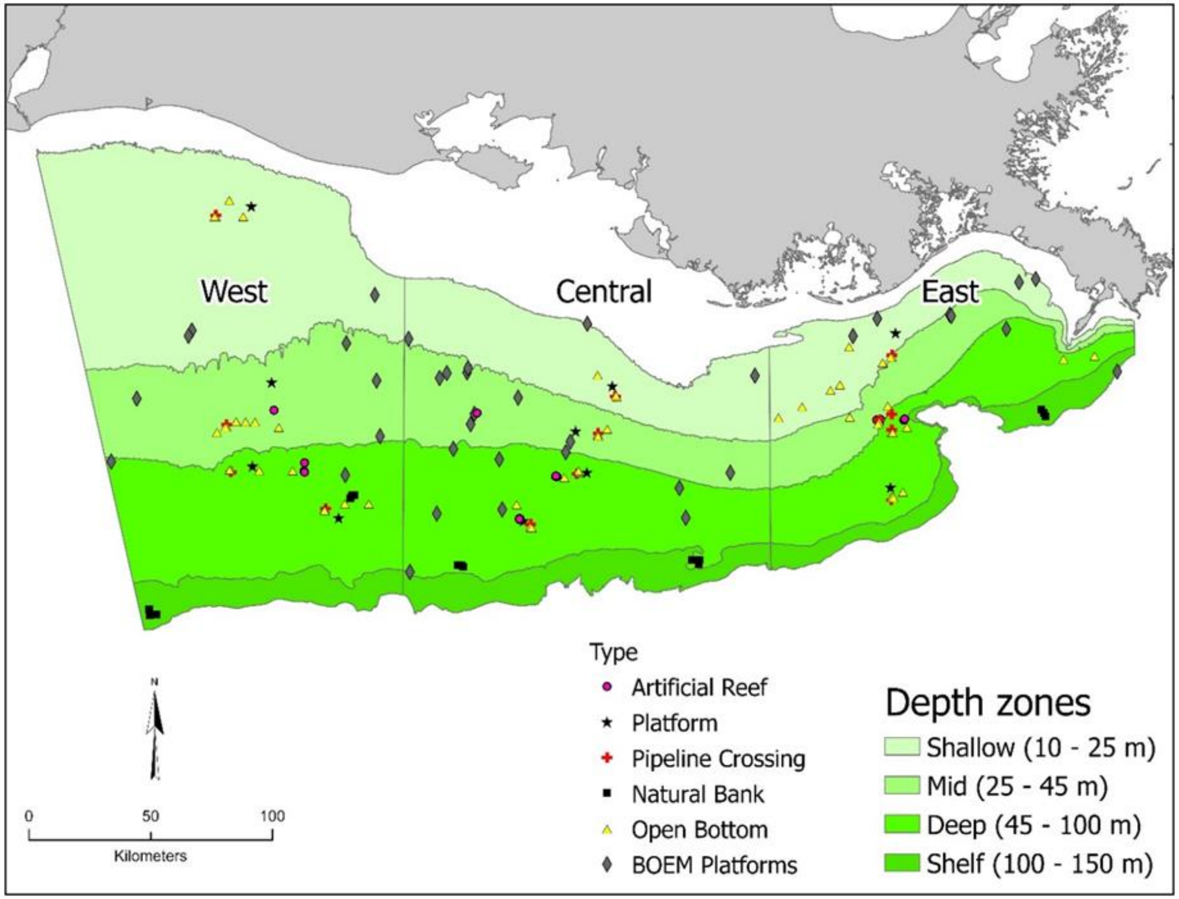
VERTICAL

- At each site, both hydroacoustic and SRV samples were collected at 10 m vertical depth bins
- Thus, multiple samples were collected at nearly all sites by these gear types
- Hook and line sampling could not be parsed into vertical depth bins
- Before statistical analysis, we calculated the distance from the center of each bin to the bottom and converted this categorical variable into a continuous variable we termed meters from the bottom (MFB)

ALLOCATION OF SITES

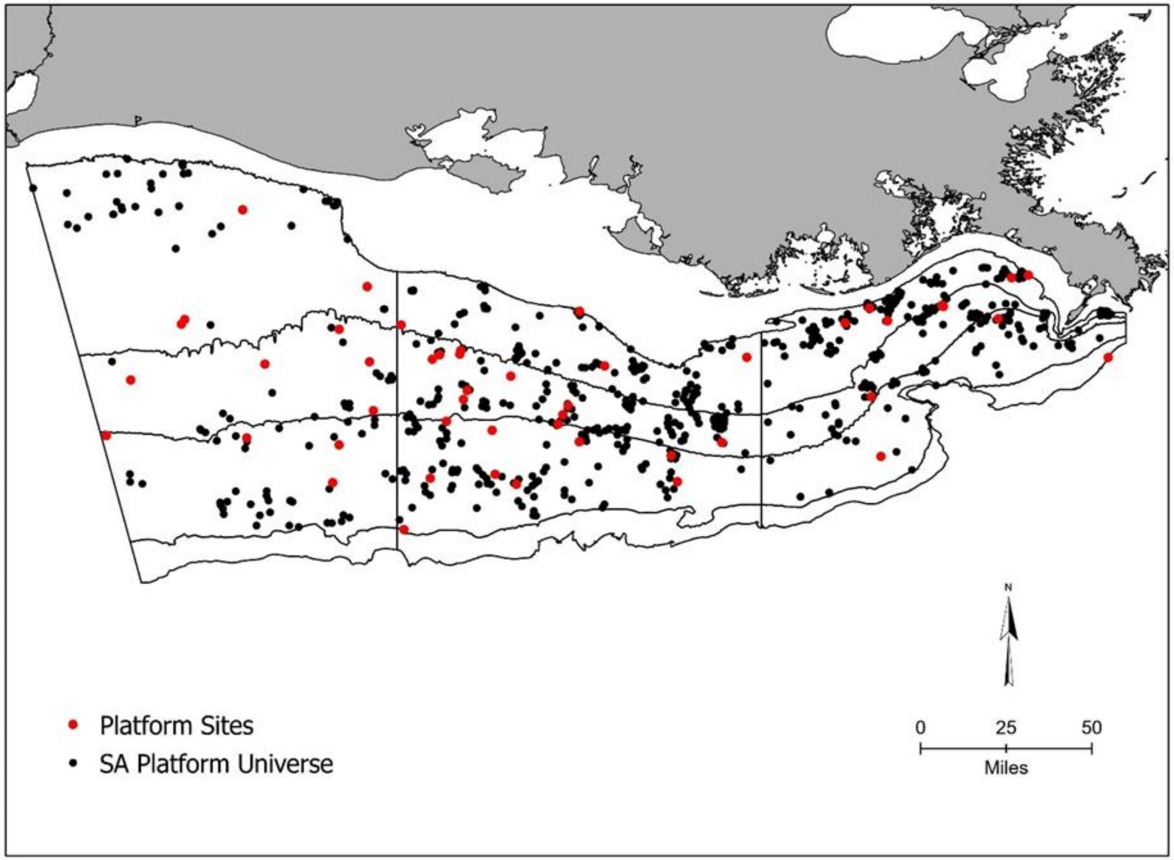
		ArtificialReef	NaturalBank	OpenBottom	PipelineCrossing	Platform
West	Shallow			3	1	4
	Mid	1		6	1	5
	Deep	3	6	6	2	4
Central	Shallow			2	1	4
	Mid	1		2	1	10
	Deep	6	6	4	2	10
East	Shallow			7	1	5
	Mid	2		3	1	3
	Deep	2	3	6	3	3

ALLOCATION OF SITES

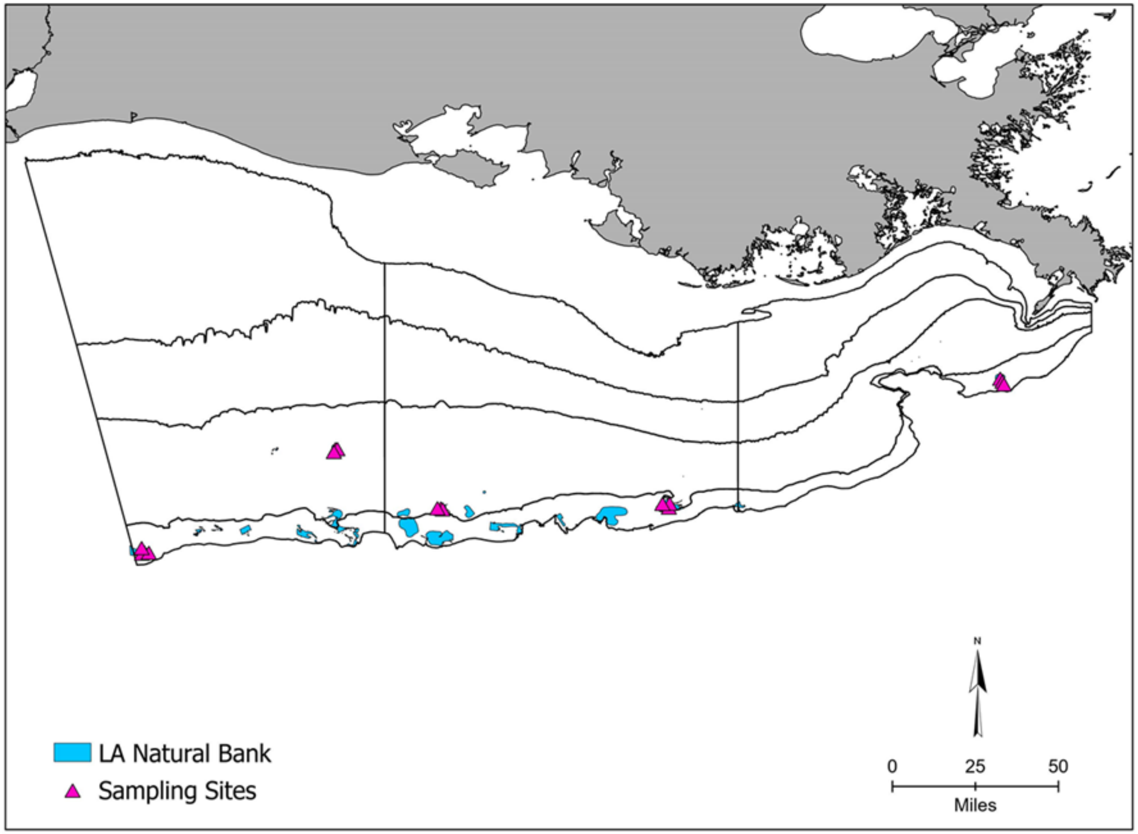


SITE SELECTION

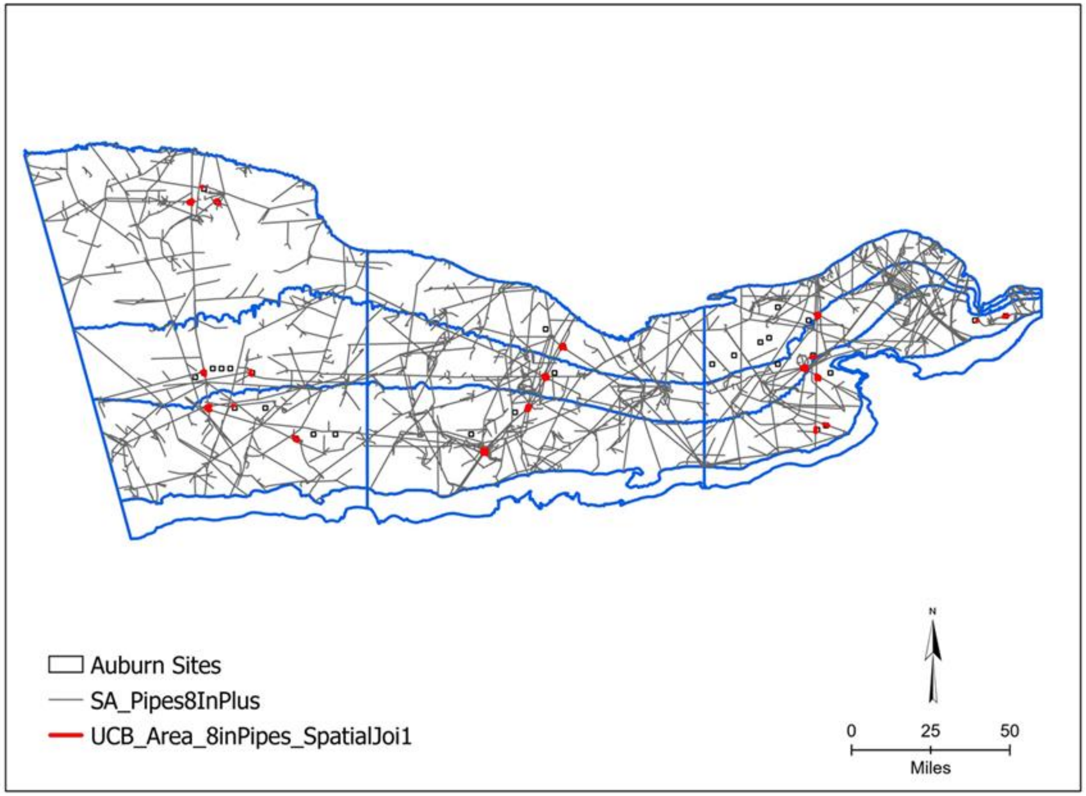
STANDING PLATFORMS



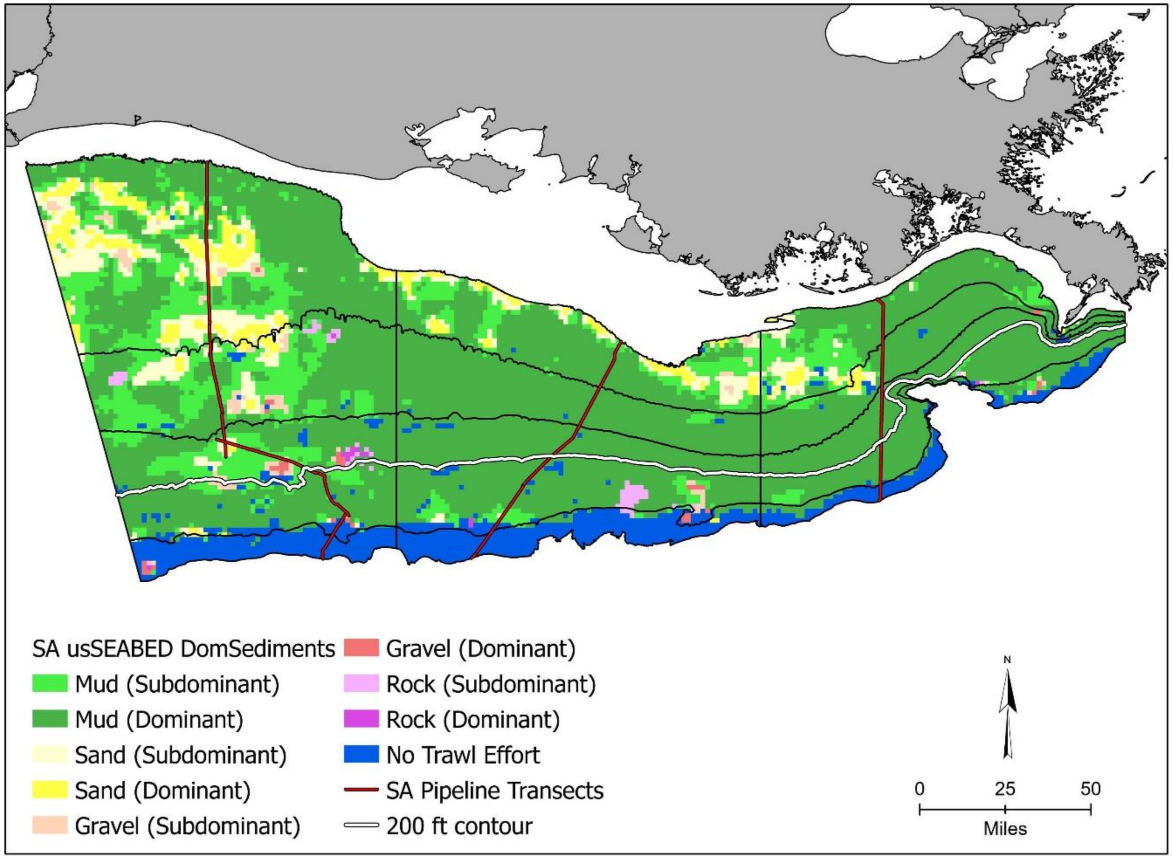
NATURAL BANKS



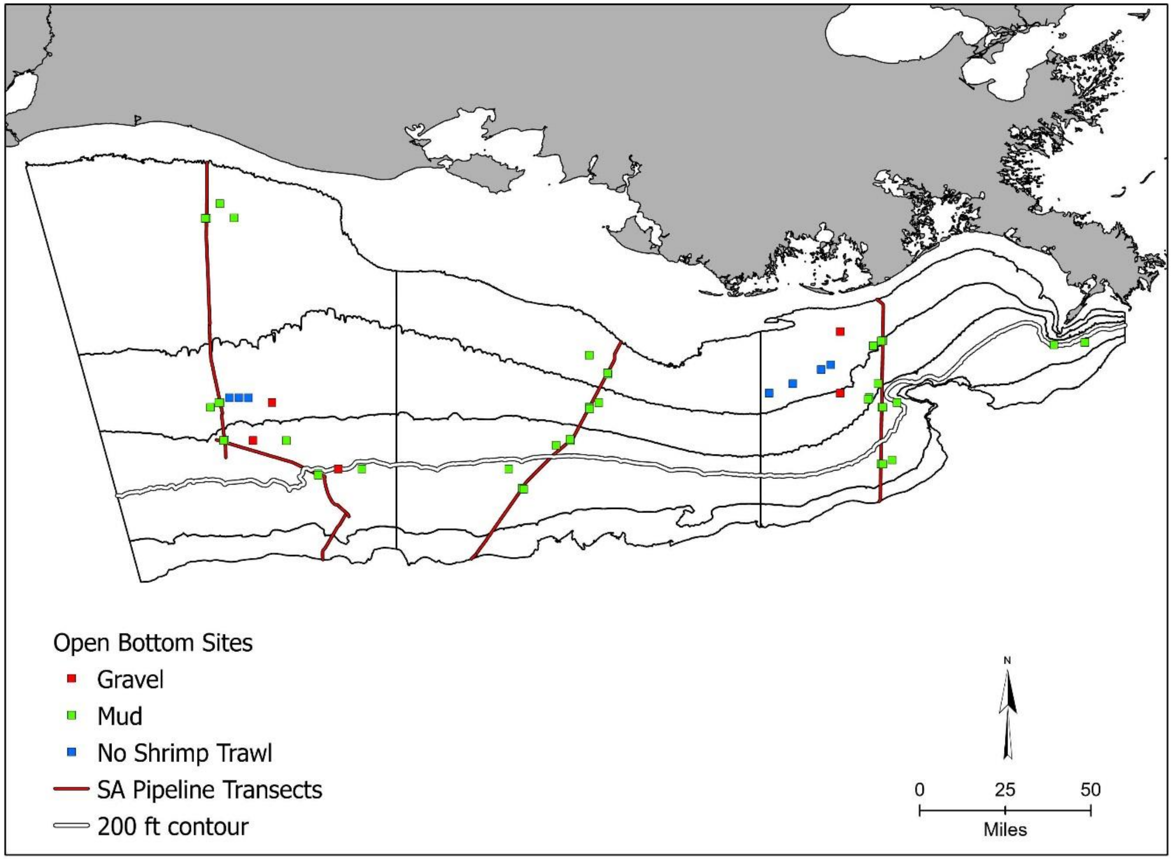
PIPELINE CROSSINGS



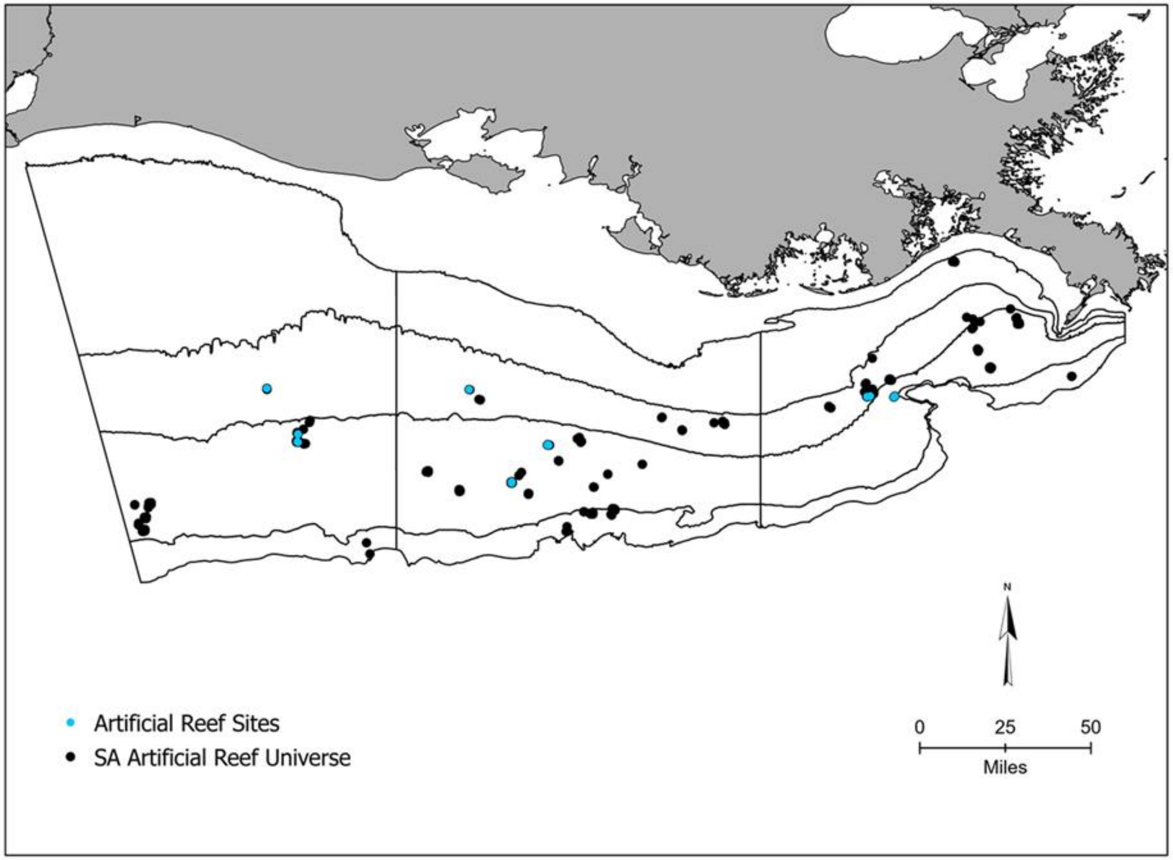
OUR UNDERSTANDING OF SUBSTRATE TYPE



UCB


Open Bottom Sites
Gravel
Mud
No Shrimp Trawl
SA Pipeline Transects
200 ft contour
N
0
25
50
Miles

ARTIFICIAL REEFS



CHOOSING AN INFERENTIAL FRAMEWORK

INFERENCE FROM SAMPLES TO THE POPULATION

- Two approaches:
  - Design-based inference: J. Neyman and E. Pearson
  - Model-based inference: R.A. Fisher
- The distinction between the two and how sampling design is related to model specification is routinely omitted by researchers

DESIGN-BASED INFERENCE

- Values of the sampled units themselves are held to be fixed, and the variance around any statistic based on these samples comes from randomness in the selection process
- Sampling design must be clearly defined and accounted for to ensure unbiased design-based inference
- Researcher is unconcerned with the conditions causing unit values to vary from one to the next. If the sample was random, then the conditions controlling their values were observed randomly, and are therefore representative of what the population experienced

MODEL-BASED INFERENCE

- Sample selection is held to be fixed while values of the sample units themselves are not, and their randomness comes from a stochastic process
- Unbiased inference is possible from nonrandomly collected samples chosen purposive or opportunistically under certain conditions

MODEL-BASED INFERENCE REQUIREMENTS

- Formulate a statistical model with terms for all important variables and their interactions
- Assume a parametric distribution from which the error term was randomly generated, which in turn renders a random dependent variable
- Meet Fisher's "conditionality principle" that can be compromised under three circumstances

MODEL-BASED INFERENCE
Conditionality Principle

- If sampling units were stratified before selection, then terms for these strata should be included in the model
- Units are sometimes clustered into groups that are each sampled multiple times. Must include a random effect to account for correlation among within site/subject samples
- When the selection of sampling units is influenced by the unit values… Selection of samples was informative. Difficult to fix!

WHICH ONE TO CHOOSE?

- Design-based inference—advantages:
  - Avoids the subjectivity that comes with assuming a distribution for the response, appropriate specification of a model, and correct conditioning on all selection and design variables
  - Conditioning often swallows degrees of freedom
  - Can be prone to error if relevant selection and auxiliary variables are unknowingly omitted

WHICH ONE TO CHOOSE?

- Design-based inference—disadvantages:
  - Empirical random sampling is not always possible
  - Important auxiliary variables and their interactions cannot be addressed if sample size is limited
- We chose to use…

POST-PROCESSING OF DATA

DELETING/POOLING SITES

- Pipeline crossings didn't work out
- Some sites of like habitat were pooled if they were <400 m apart
- One platform was <400 m from an artificial reef and was deleted. We had more platform sites, so…

MODEL SPECIFICATION

INDEPENDENT VARIABLES

- All stratifying variables
- Meters from the bottom
- DO and Salinity
- Site as a random effect to account for nonindependence for multiple observations at a site
- Platform complexity—not used; intractable like pipeline crossings
- Substrate type—not used; GIS layer too coarse to accurately determine for a given site
- Temperature—not used; collinear with DO and Salinity

MODELING PROPORTION OF RED SNAPPER FROM SRV SURVEYS

```
PropRS <- gam(RS/Total ~ Region + DZ + HabType +
              s(DO, k=5, bs="tp", m=1) + s(Salinity, k=5, bs="tp", m=1) +
                  s(MFB, by=HabType, bs="tp", k=5, m=1) +
                  s(Site, bs="re"),
              family = binomial(link="logit"), weights=Total, method="REML",
              optimizer=c("outer","newton"), data=SRVdata)
```

MODELING TFD FROM HYDROACOUSTIC SURVEYS

```
TFD <- gam(TFD ~ Region + DZ + HabType +
        s(DO, k=5, bs="tp", m=1) + s(Salinity, k=5, bs="tp", m=1) +
            s(MFB, by=HabType, bs="tp", k=5, m=1) +
            s(Site, bs="re"),
        family=tw(a=1.01, b=1.99, link="log"), method="REML",
        optimizer=c("outer","newton"), data=HydroData)
```

PREDICTING RED SNAPPER ABUNDANCE

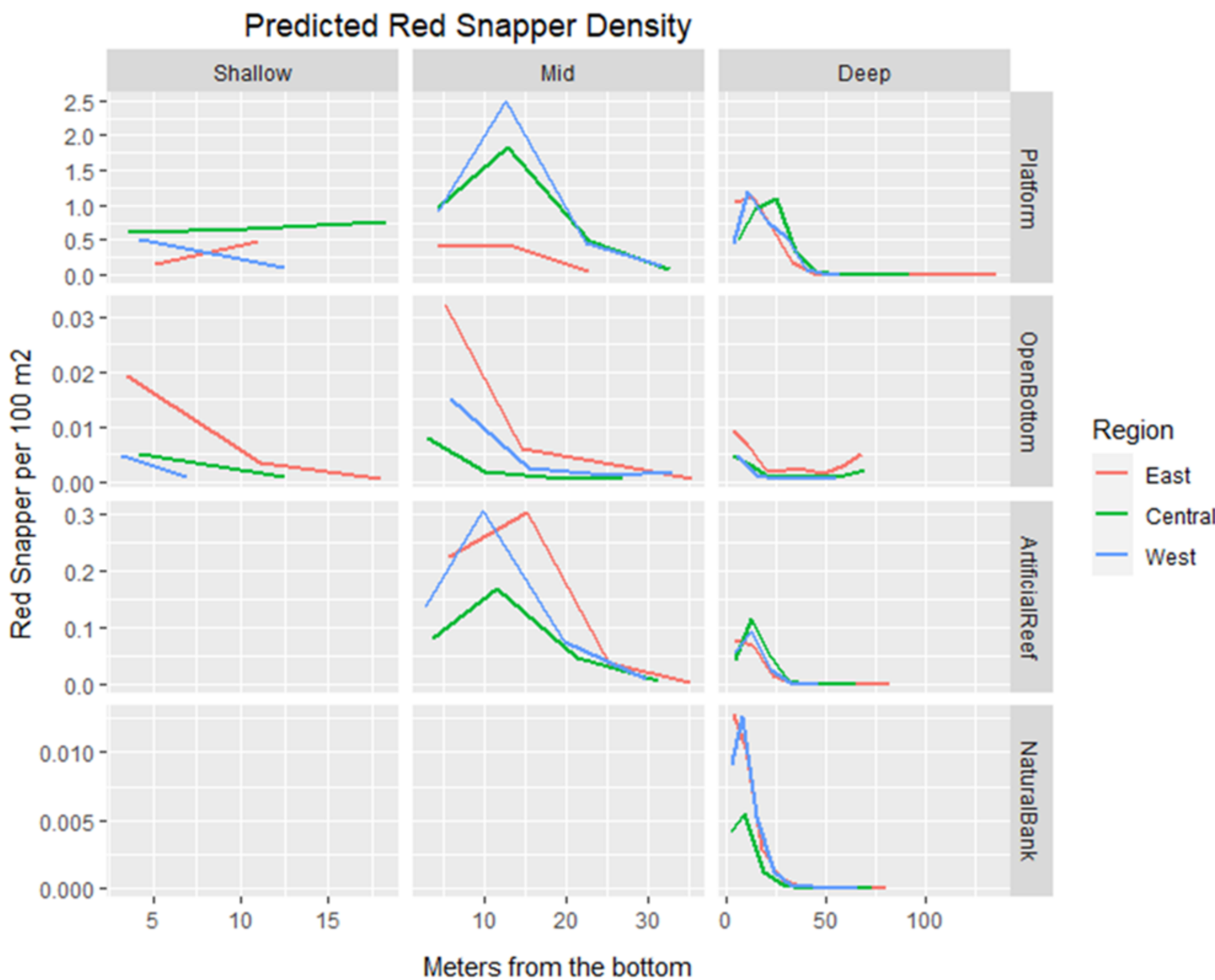
- Multiplied outputs from both models after conditioning on pertinent variables
  - Fixed the random effect of Site to zero
  - Used average values of covariates observed for each combination of strata

PREDICTING RED SNAPPER ABUNDANCE

- Multiplied outputs from both models after conditioning on pertinent variables
  - Fixed the random effect of Site to zero
  - Used average values of covariates observed for each combination of strata

RESULTS

PREDICTIONS BY VERTICAL DEPTH



SUMMING OVER DEPTH

Habitat type	Region	Depth zone	Red Snapper per 100 m² Mean	LCL	UCL	Red Snapper per km² (*) or per structure (**)	LCL	UCL	Area km² (*) or structure count (**)	Total abundance	Subtotal (% of overall)	LCL	UCL
Natural Banks*	West	Deep	0.03	0.01	0.06	291	134	634	180	52,393	118,647	66,390	212,036
	Central		0.01	0.00	0.03	115	47	279	521	59,923	(1%)		
	East		0.03	0.01	0.08	272	95	783	23	6,331			
Platforms**	West	Shallow	0.61	0.19	1.92	257	81	811	62	15,911	727,210	545,780	968,953
		Mid	4.00	1.89	8.44	1,689	800	3,565	25	42,226	(9%)		
		Deep	3.07	1.81	5.23	1,299	763	2,210	52	67,528			
	Central	Shallow	2.04	0.88	4.74	861	370	2,002	118	101,624			
		Mid	3.37	1.77	6.40	1,423	748	2,705	133	189,220			
		Deep	2.95	1.76	4.94	1,247	744	2,088	117	145,852			
	East	Shallow	0.64	0.17	2.38	268	72	1,005	182	48,854			
		Mid	0.88	0.30	2.58	373	128	1,089	58	21,615			
		Deep	3.02	1.38	6.60	1,275	583	2,790	74	94,381			
Artificial reefs**	West	Mid	0.53	0.21	1.36	408	238	698	5	2,038	86,954	67,068	112,737
		Deep	0.18	0.08	0.39	135	85	213	121	16,329	(1%)		
	Central	Mid	0.30	0.11	0.79	233	134	404	35	8,144			
		Deep	0.22	0.10	0.50	170	108	269	160	27,279			
	East	Mid	0.57	0.19	1.68	440	236	819	57	25,056			
		Deep	0.16	0.06	0.42	127	74	218	64	8,109			
UCB*	West	Shallow	0.01	0.00	0.02	56	15	218	10,268	579,014	7,444,780	5,440,478	10,187,477
		Mid	0.02	0.01	0.05	207	90	479	5,297	1,096,405	(89%)		
		Deep	0.01	0.00	0.02	84	46	153	7,162	599,056			
	Central	Shallow	0.01	0.00	0.03	62	14	267	4,407	271,732			
		Mid	0.01	0.00	0.03	112	40	314	3,760	420,083			
		Deep	0.01	0.01	0.03	149	85	261	7,511	1,120,980			
	East	Shallow	0.02	0.01	0.08	234	69	799	3,058	716,409			
		Mid	0.04	0.02	0.11	420	156	1,130	2,327	978,324			
		Deep	0.03	0.02	0.05	319	185	550	5,213	1,662,776			
											8,377,591	6,365,225	11,026,166

NOT TALKED ABOUT

- Measurement bias with SRV sampling
- Age and growth bias from hook selectivity