Bio-Objectives Technical Update:
Scoring Tool Development and Testing

Technical Update:
Scoring Tool Development and Testing

- **Review of reference work and O/E process**
- **Building the model**
- **Performance Tests**
  - **What we measured and why**
  - **Results**: statewide overview and regional comparisons
- **Recommendations to Science Panel**
- **What's next**

SWAMP Surface Water Ambient Monitoring Program

Objectives:

- Develop scoring tools to objectively assess biological condition of all CA wadeable perennial streams
- Requirement is to balance statewide consistency with regional validity
- Optimize tool based on multiple measures of performance

SWAMP
Surface Water Ambient Monitoring Program

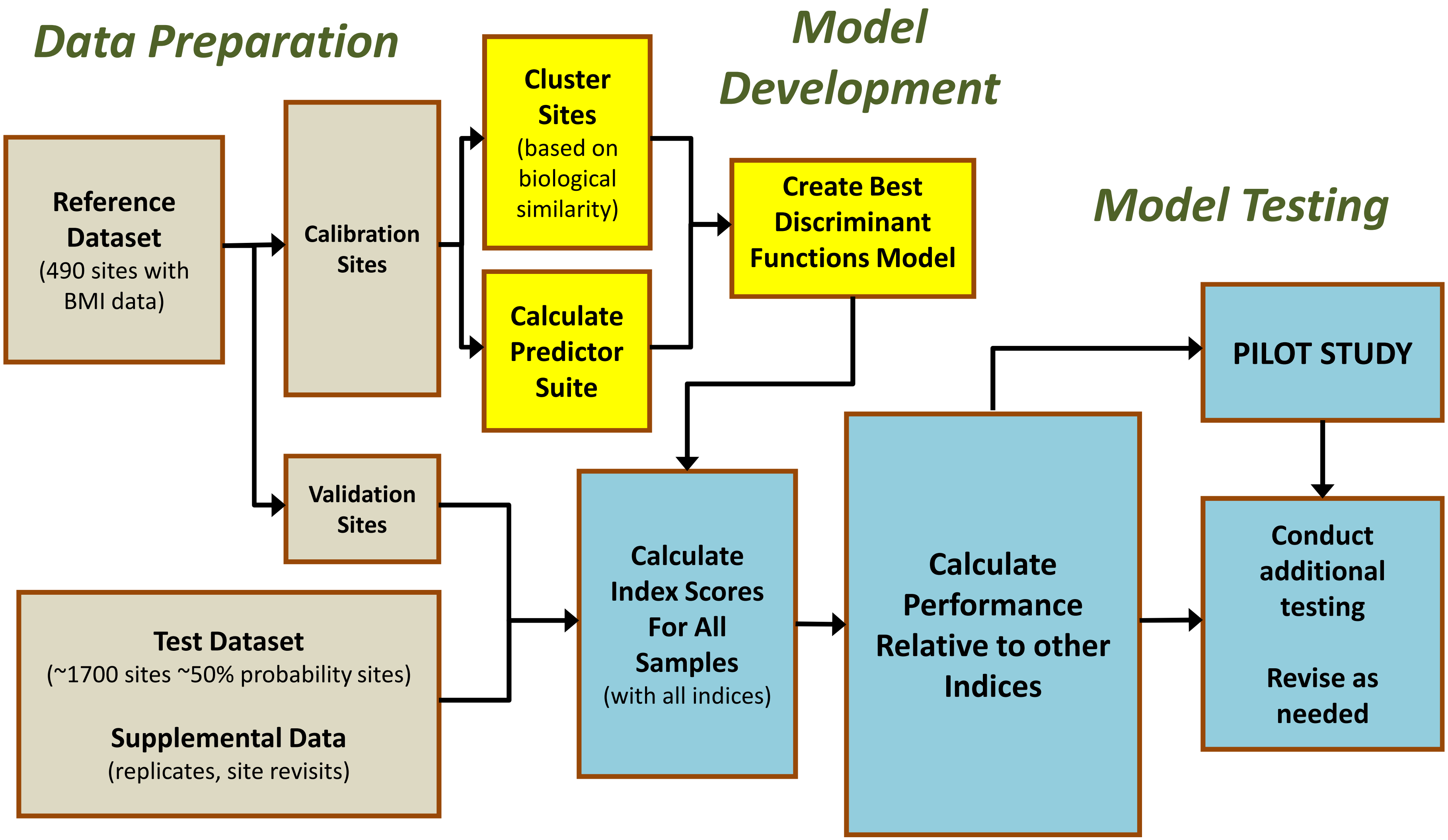
Why Develop A New Tool?

- Existing tools have limitations for statewide application
  - Spatial coverage is limited
  - Reference site definitions not consistent
  - Reference distributions not fully representative

- MMI (IBI) and O/E are both viable approaches; we focused on O/E
  - Designed to predict site-specific expectations, rather than a regional reference average
  - Species loss is a relevant measure of ecological condition
  - Index is amenable to statewide standardization

SWAMP
Surface Water Ambient Monitoring Program

O/E Index Development Process



Scoring Tools Depend on Reference Sites

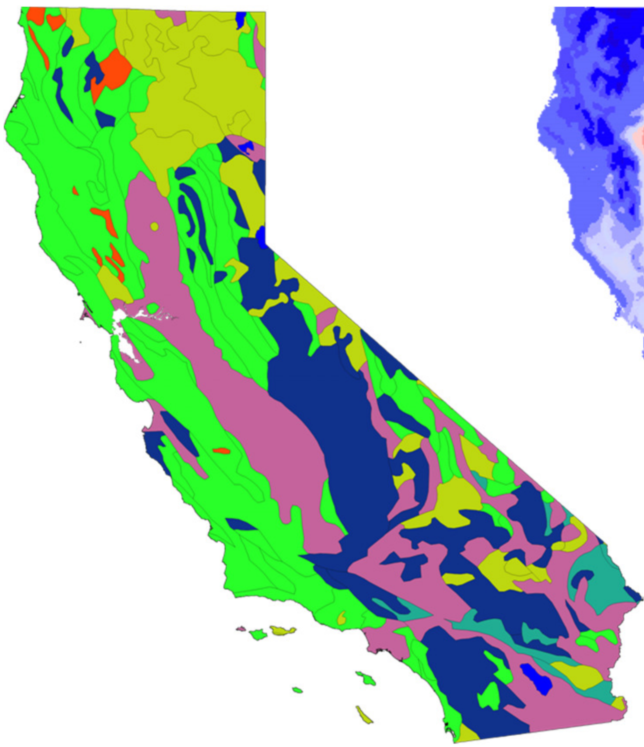
(sites with low levels of disturbance)

"What should the biology look like at a test site?"

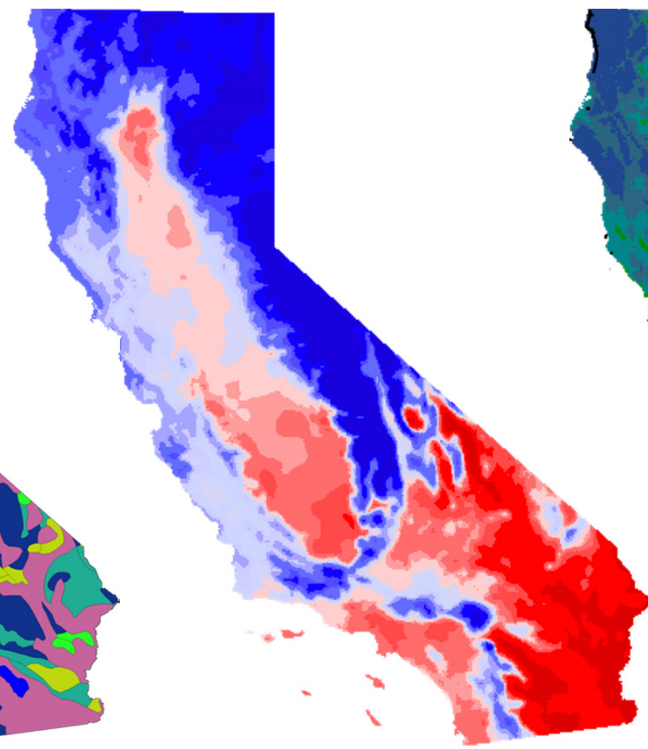
Technical Challenges:

Strong natural gradients result in a large degree of ***natural variation*** *in biological expectations*

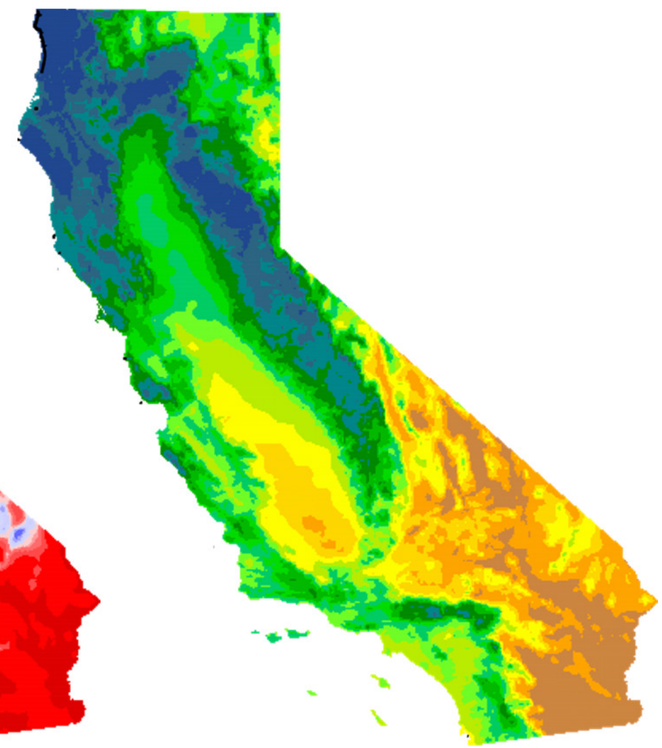
Geology



Temperature



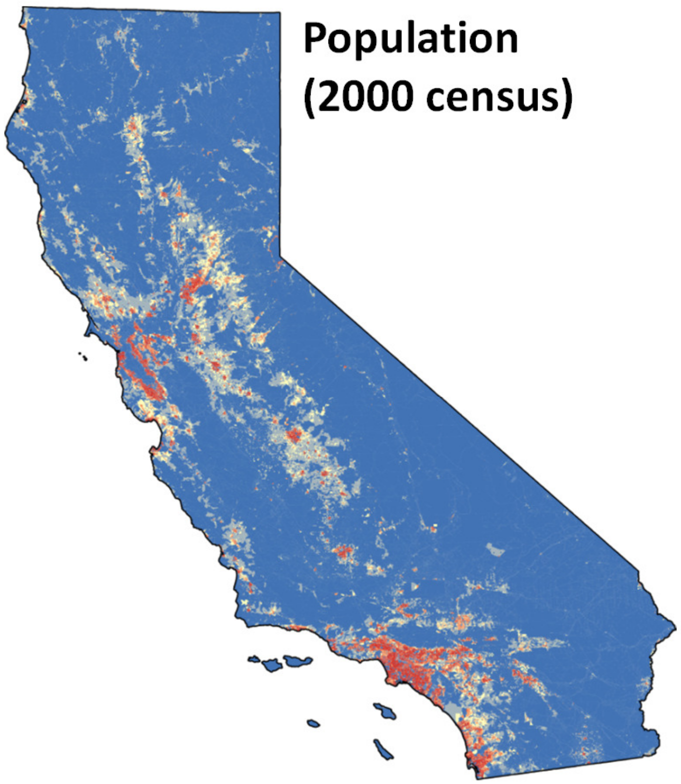
Precipitation

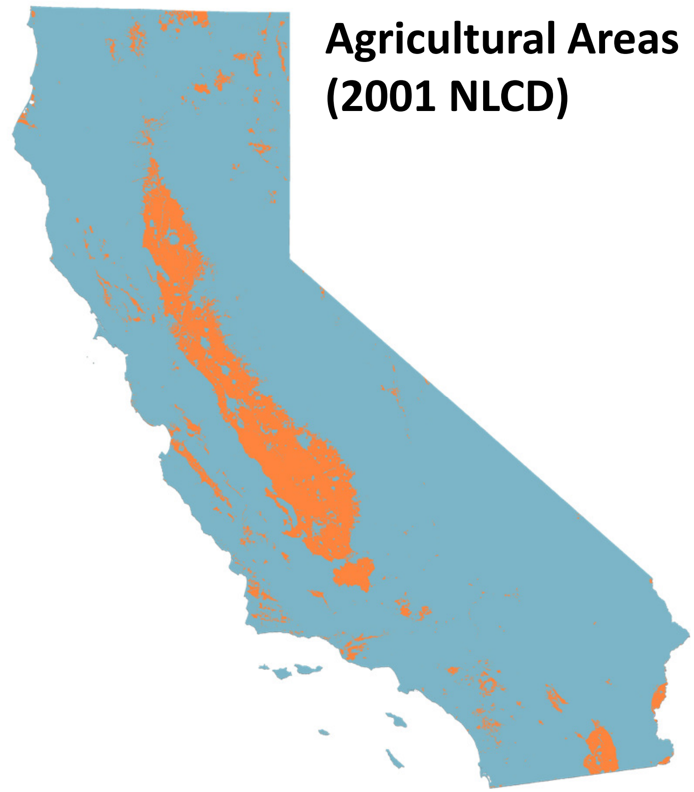


Management of biological variability requires good representation of biology at reference sites **across major gradients = need 100s of sites**

Technical Challenges:

High degree of anthropogenic modification (e.g., impervious surface and intensive agriculture) in some regions





- Extensive modification introduces **gaps in representation** of natural gradients
- Widespread development can make some regions unsuited for standard reference approaches

Reference Criteria for Biological Objectives

Balancing site purity and representativeness

Trade-off: Need to allow limited sources of anthropogenic stress in order to get good representation of all stream types *(this constraint is shared by all bioassessment indices)*

Performance Objectives:

1. Reference pool represents all types of CA streams
2. Biological "quality" is maintained at reference sites

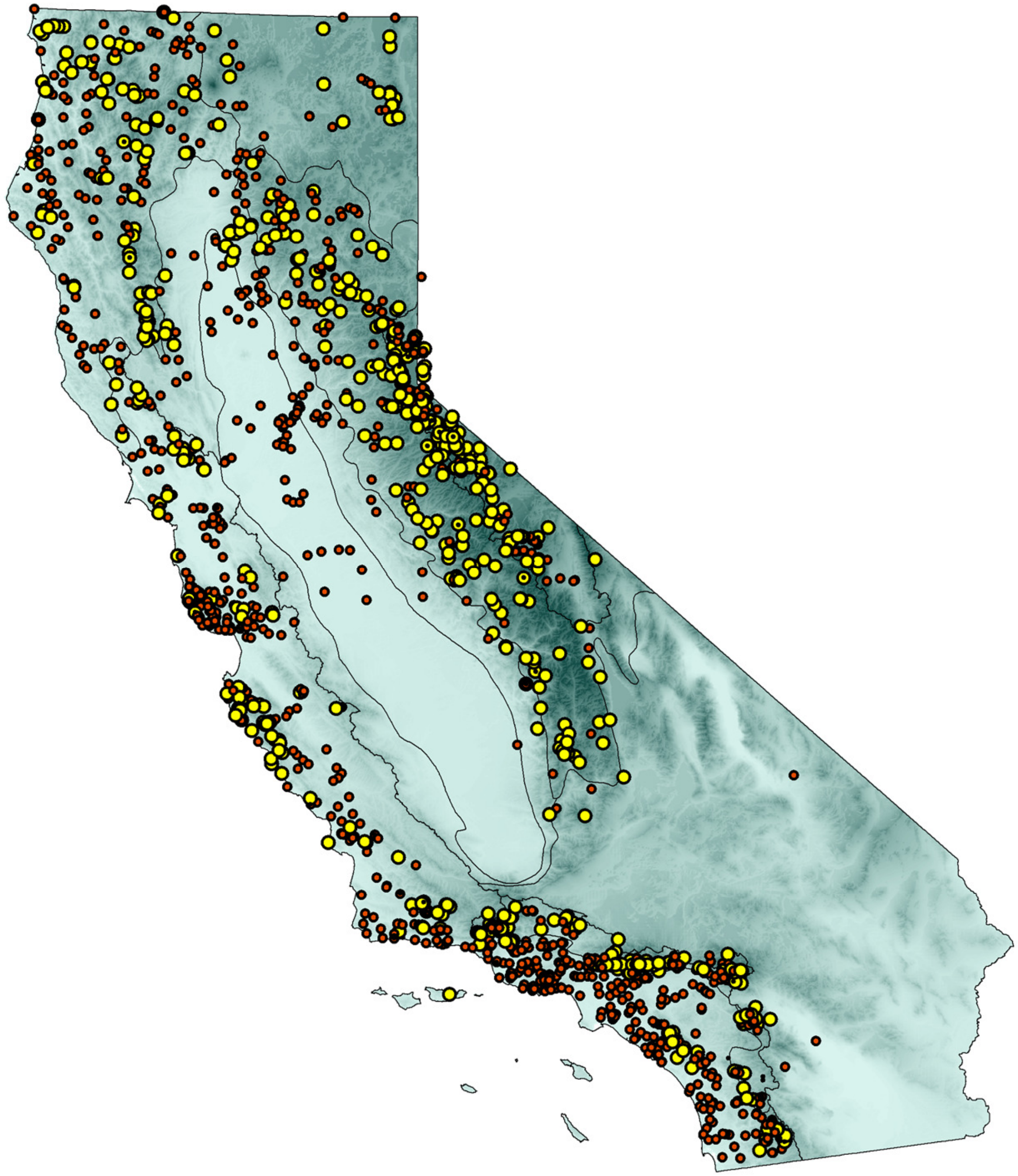
SWAMP Surface Water Ambient Monitoring Program

Thresholds are comparable or stricter than other CA indices and include many more criteria

Metric	2011 Bio-objectives	South Coast IBI (5k,ws)	North Coast IBI (1k, ws)	Current O/Es (Hawkins 2005)
Local Disturbance (W1_Hall)	**1.5**	-	-	riparian vegetation, erosion, grazing
% Agricultural	**3,3,10**	**5**	5	
% Urban	**3,3,10**	**3**	3	
% Ag + Urban	**5,5,10**			
% Code 21	**7,7,10**	**in urban**	in urban	
Road Dens (km/km^2)	**2,2,2**	**2.0**	1.5/ 2.0	
Paved road x-ings (#/ws)	**5/10/50**			
TN, TP (mg/L)	**3.0/ 0.5**	-	-	
Nearest Dams	**>10 km**	-	-	
Active Producing Mines	**0** (5k)	-	-	
% Canals & Pipelines	**10**	-	-	
Gravel Mine Density	**0.1** (r5k)			
Conductivity	**<2000 uS, + <99%, >1%**			
BPJ Screen	**X**	**X**	**X**	**X**

Reference Sites

REGION	n
North Coast	79
Central Valley	1
Coastal Chaparral	87
Interior Chaparral	30
South Coast Mountains	96
South Coast Xeric	22
Western Sierra	131
Central Lahontan	142
Deserts + Modoc	27
TOTAL	**615**



Reference Conditions: Performance Summary

Stream Type Representation - evaluated representation of sites along major natural gradients (elevation, climate, slope, geology, stream size)

- Overall excellent representation in most regions (absent in Central Valley, fewer in SoCal xeric region)
- Some under-representation of very low gradient, large watershed, low elevation settings in Chaparral and South Coast

Biological Integrity

- No significant reduction in biological integrity at reference sites relative to pristine sites

SWAMP Surface Water Ambient Monitoring Program

Observed/ Expected Indices

Developed in UK (Wright and others 1970s-1980s, RIvPACS)
– now widely used worldwide

Species-based approach: Compare number of **observed** ("O") taxa to number of **expected** ("E") taxa

"Expected" taxa at a test site are modeled using predictive modeling techniques

Compare test site to subsets of the reference sites that are physically similar to the test site *(geology, climate, elevation, latitude, etc.)*

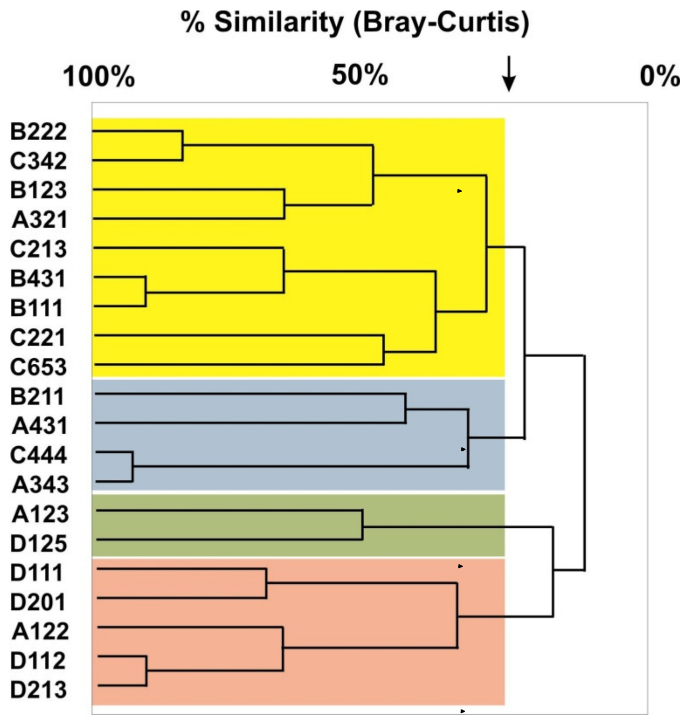
Index score is a direct measure of taxonomic loss

SWAMP
Surface Water
Ambient Monitoring
Program

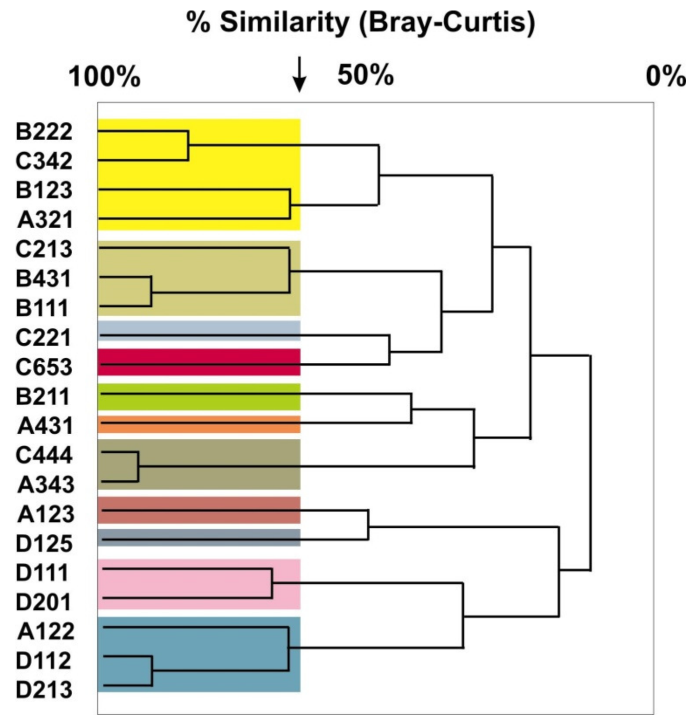
Estimating "E": Step 1

Group reference sites based on biological similarity

Clustering techniques used to identify groups of reference sites with similar species composition



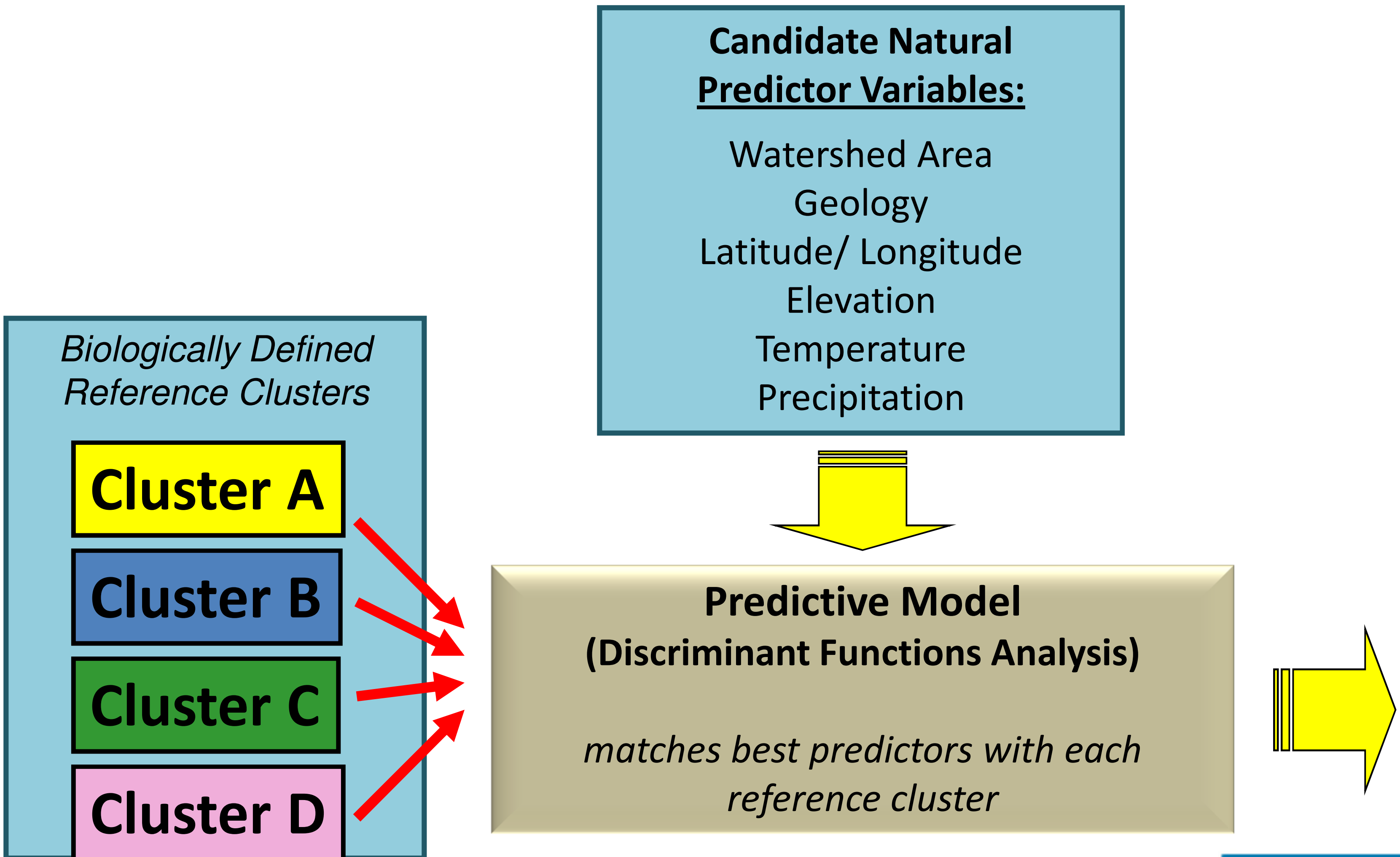
4 classes



11 classes

Estimating "E": Step 2

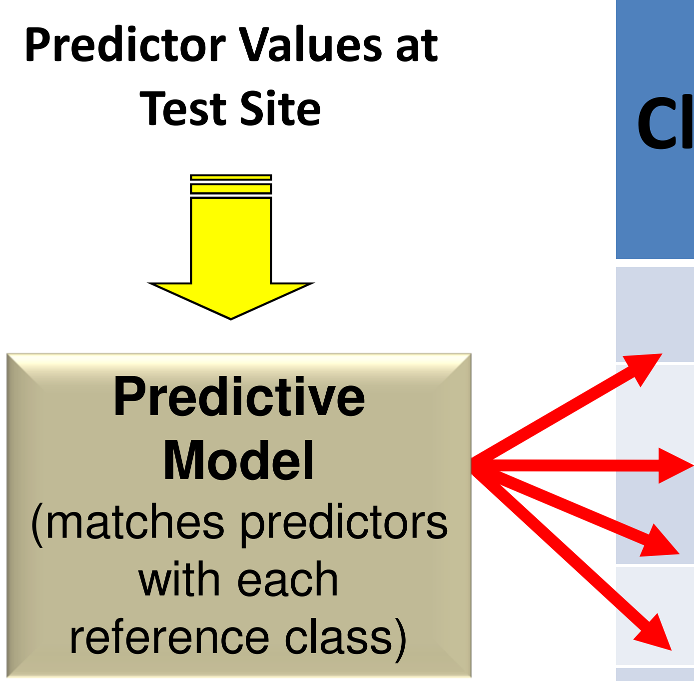
Develop model that will
predict cluster membership for new sites



Estimating “E”: Step 3

Estimate capture probabilities

Use discriminant model output + frequencies of occurrence within each class to estimate **probabilities of capture (PC)** for each taxon at a given site



Cluster	Site’s probability of cluster membership	Frequency of species X (*Farula sp.*) in cluster	Expected contribution to PC
A	0.5	0.6	0.30
B	0.4	0.2	0.08
C	0.1	0.0	0.00
D	0.0	0.0	0.00
Probability of *Farula sp.* being in sample if site is in reference condition			**0.38**

Estimating "E": Step 4

Sum taxon occurrence probabilities estimate the number of native taxa (E) that should be observed (O)

Taxon	pc	O
Atherix	0.70	*
Baetis	0.92	*
Caenis	0.86	
Drunella	0.63	
Epeorus	0.51	*
Farula	**0.38**	
Gyrinus	0.07	
Hyalella	0.00	*
Count	4.07	3

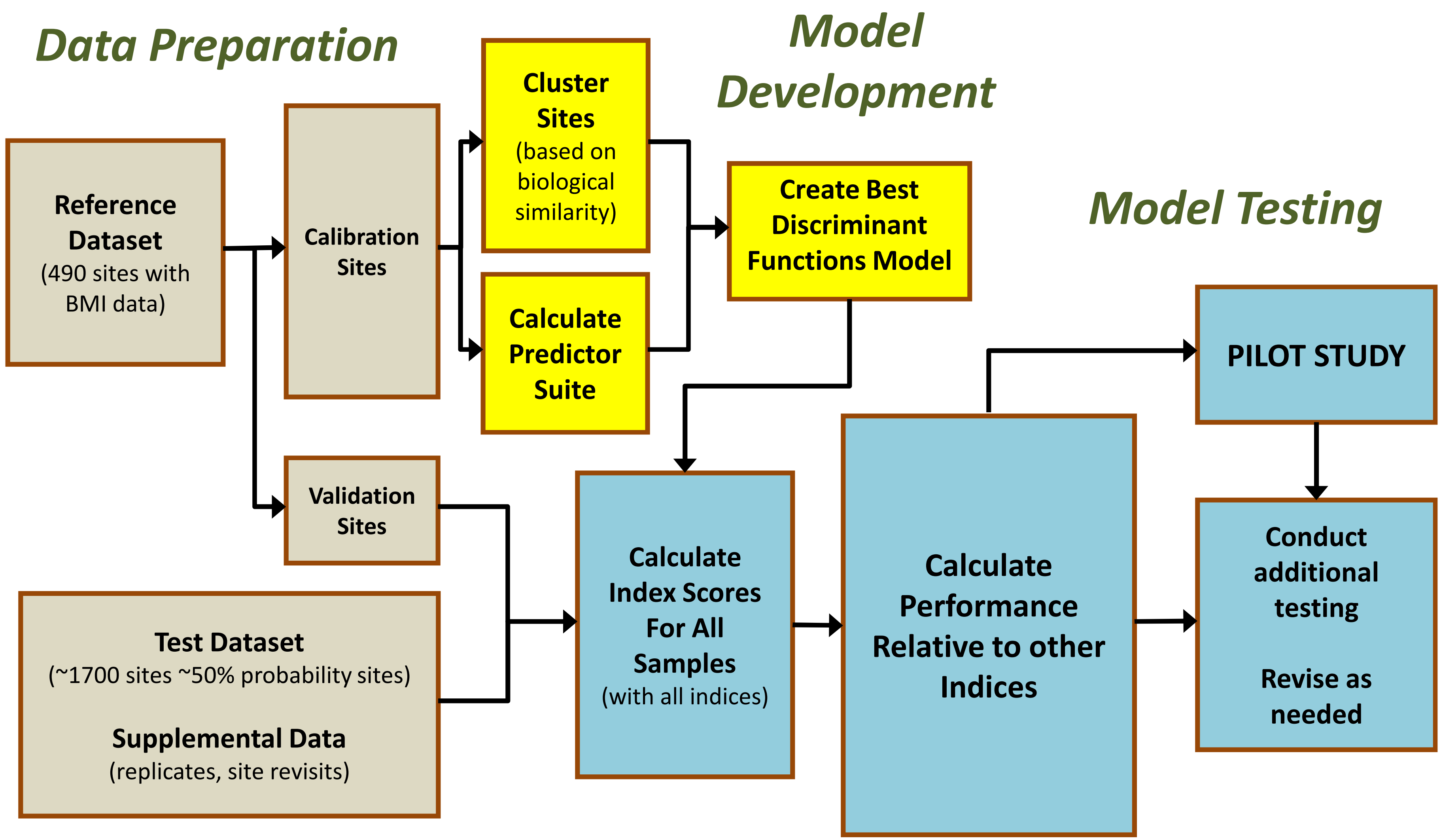
O/E = 3 / 4.07

O/E = 0.74

O/E Score
Indicates proportion of native assemblage present at test site

SWAMP
Surface Water Ambient Monitoring Program

O/E Index Development Process

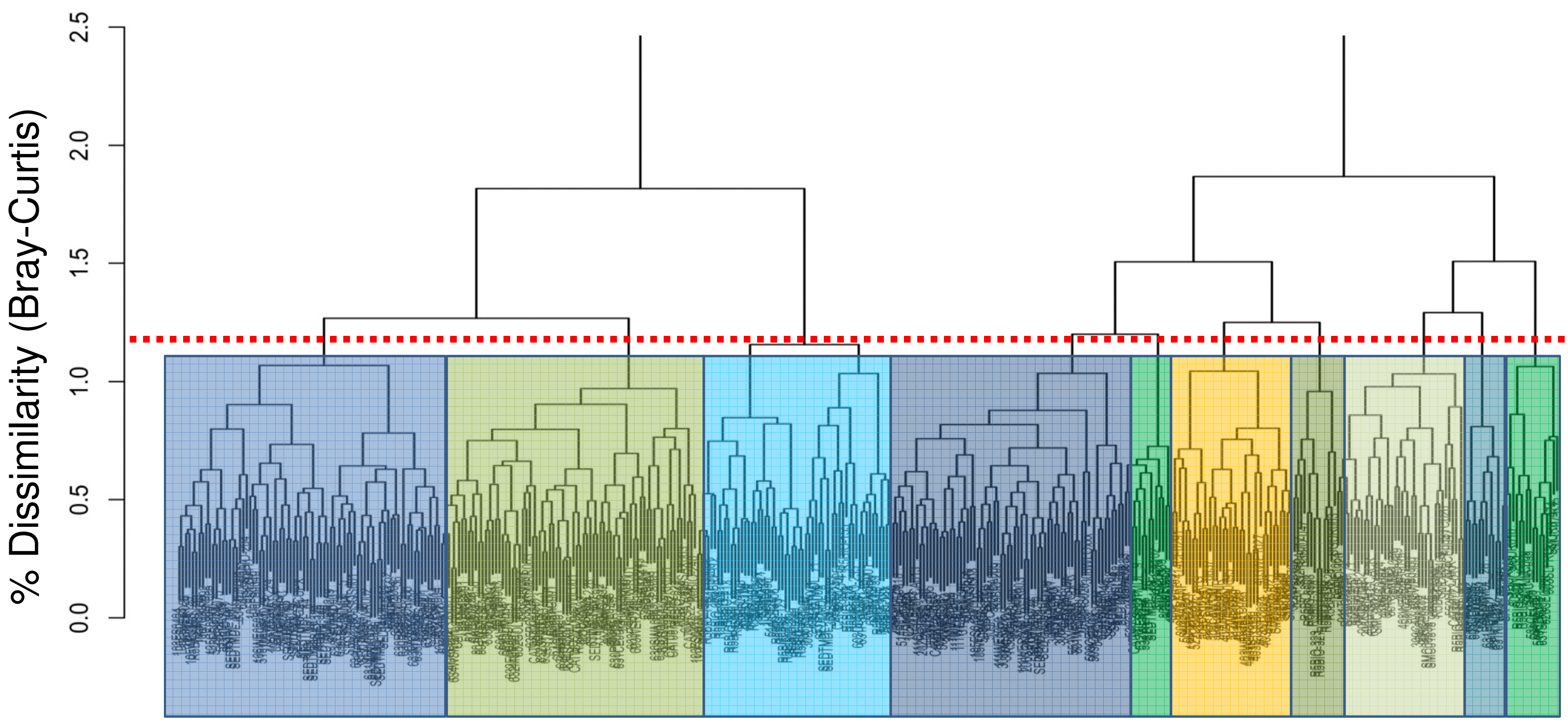


Data Preparation & Initial Decisions

- 615 reference sites identified in reference task
- Taxonomic effort standardized to SAFIT I (a): *mostly genus level IDs, with Diptera: Chironomidae to subfamily*
- 490 sites were suitable for modeling *(i.e., had sufficient BMI counts after removing ambiguous taxa)*
- Prepare 34 **natural** predictor variables
- Split dataset into **calibration** and **validation** sets (80:20, 392 sites in calibration set)

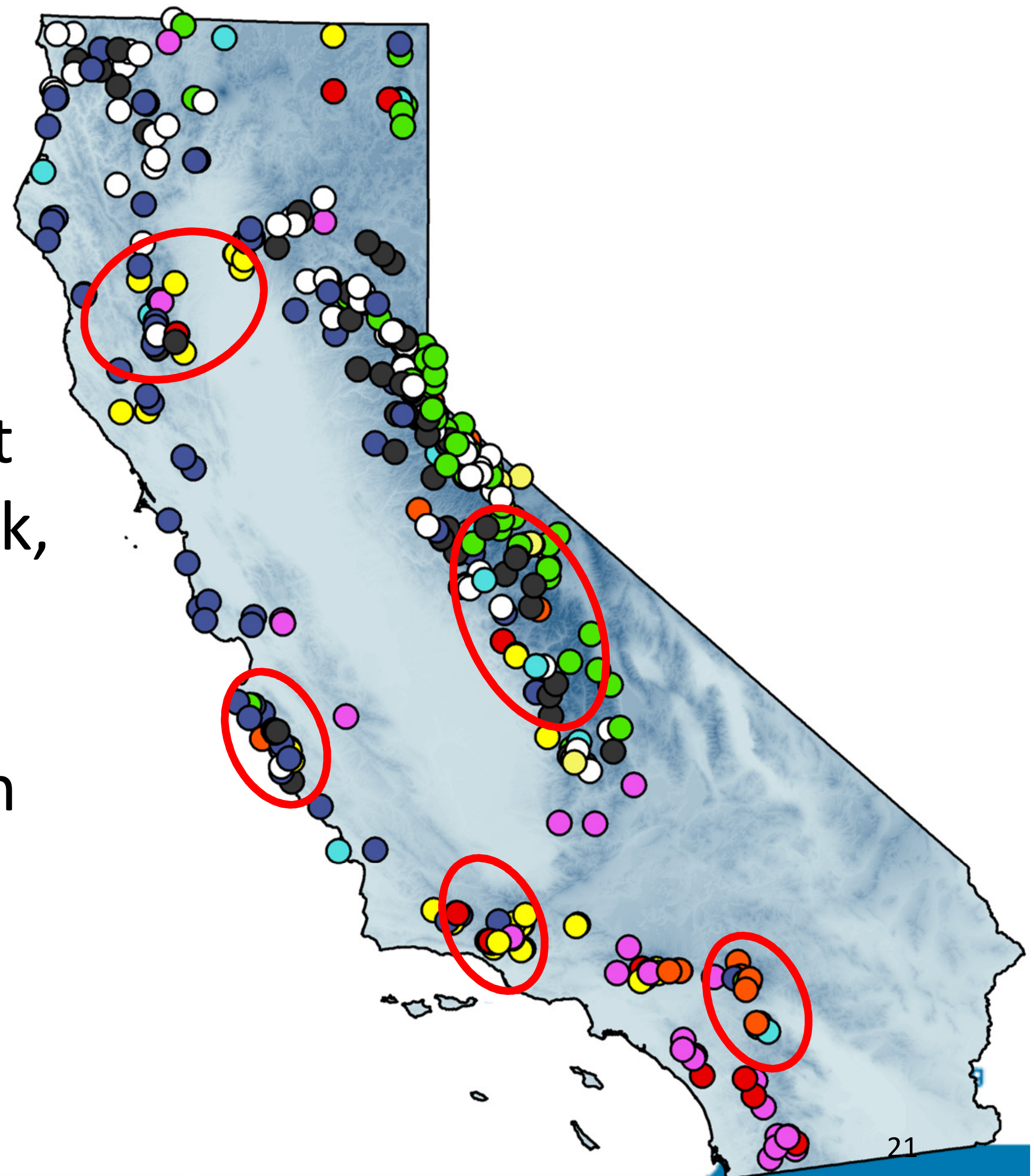
Cluster biological similarity

(Bray-Curtis dissimilarity, flexible-β = -0.25, rare taxa removed if < 5% of sites)

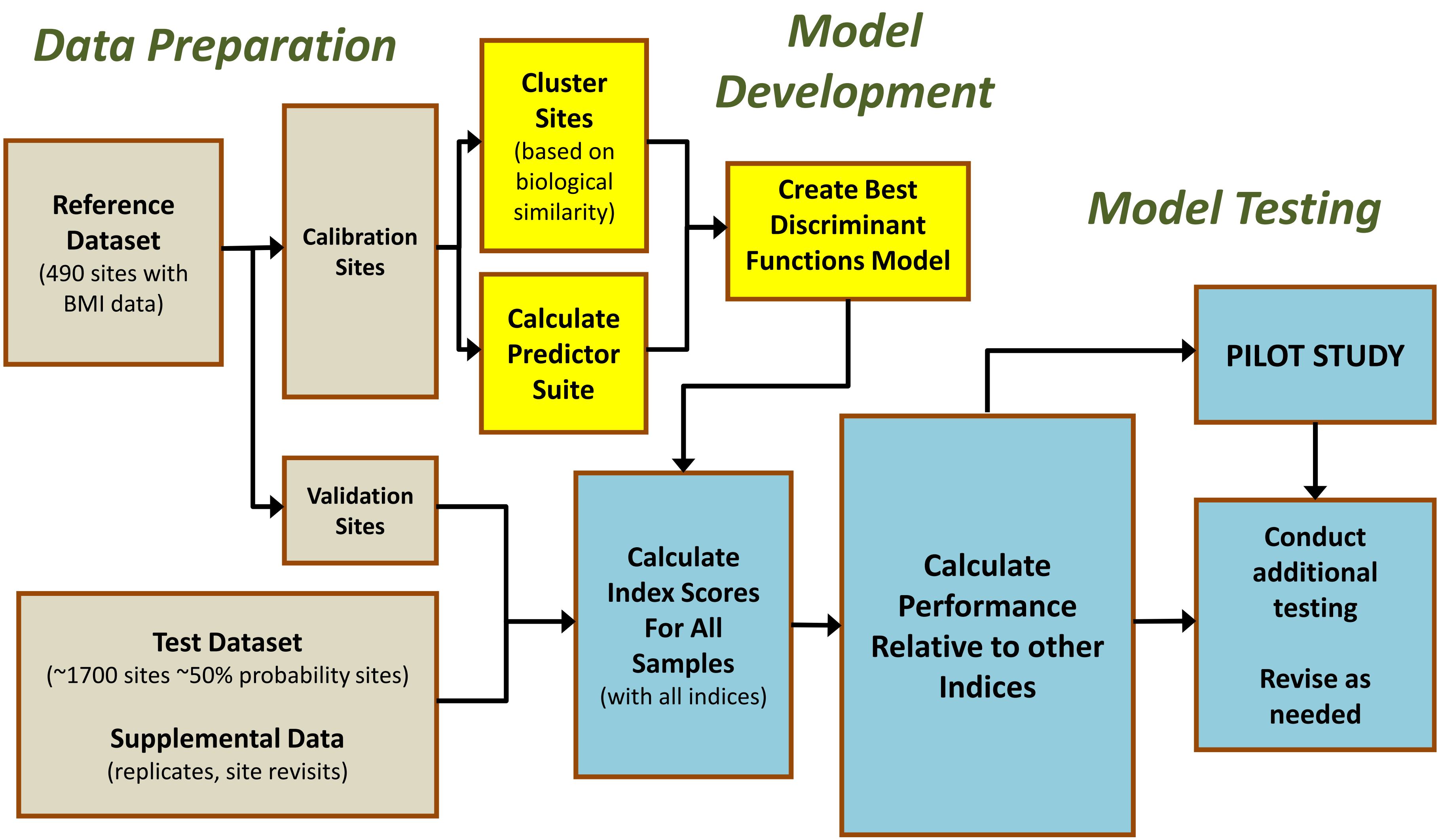


10 biological clusters

- Several large, geographically coherent clusters (e.g., blue, black, green)
- Several pockets of high variability



O/E Index Development Process



Discriminant Functions Model

- Examined all possible subsets of DFA models using 10 predictors (winnowed from 34)
- Explored effects of cluster sizes, RF models, predictor types and numbers, recent climate, etc.
- Best model had 5 predictors. More predictors did not improve model performance

SWAMP
Surface Water Ambient Monitoring Program

Predictors for DFA model

Elevation

Log Watershed Area

Avg PPT (2000 to 2009)

Avg Temp (2000 to 2009)

Log Predicted Conductivity (predicted by conductivity model)

- All predictors are GIS based and can be calculated for novel test sites
- Climate data come from PRISM national data center (Oregon)
- Conductivity predictions come from Olson and Hawkins (in review) model that predicts conductivity at test sites

SWAMP
Surface Water
Ambient Monitoring
Program

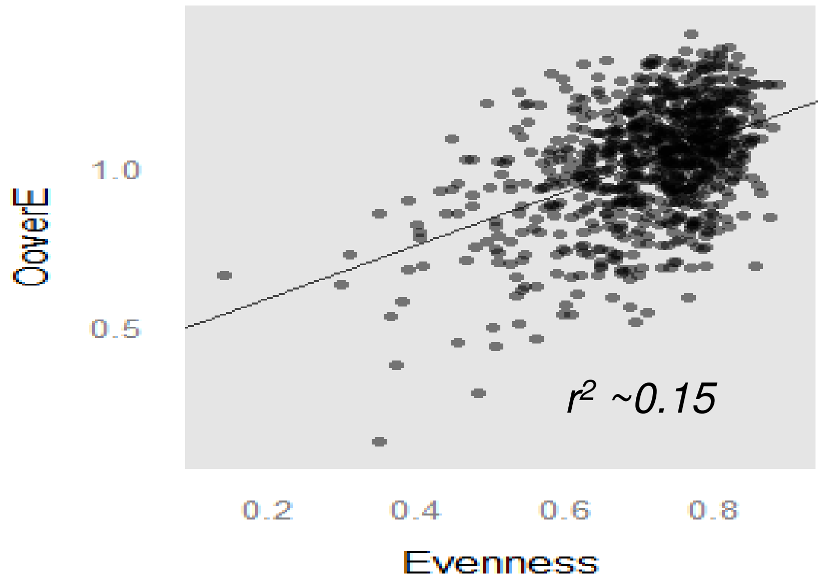
Indices used in comparisons

Name	Description
O/E	O/E index (modeled with 5 predictors)
***O/E_ec**	**O/E index with evenness correction**
O/E_null	O/E index with no predictors (null model)
O/E_null_ec	O/E null model with evenness correction
B-C	Bray-Curtis weighted distance index
B-C_ec	BC with evenness correction
B-C _null	BC null model
B-C_null_ec	BC null model with evenness correction
O/E (2005)	**Current O/E index (Hawkins, 2005; 3 submodels)**
NCIBI	**North Coast IBI**
SCIBI	**South Coast IBI**

Why an evenness correction?

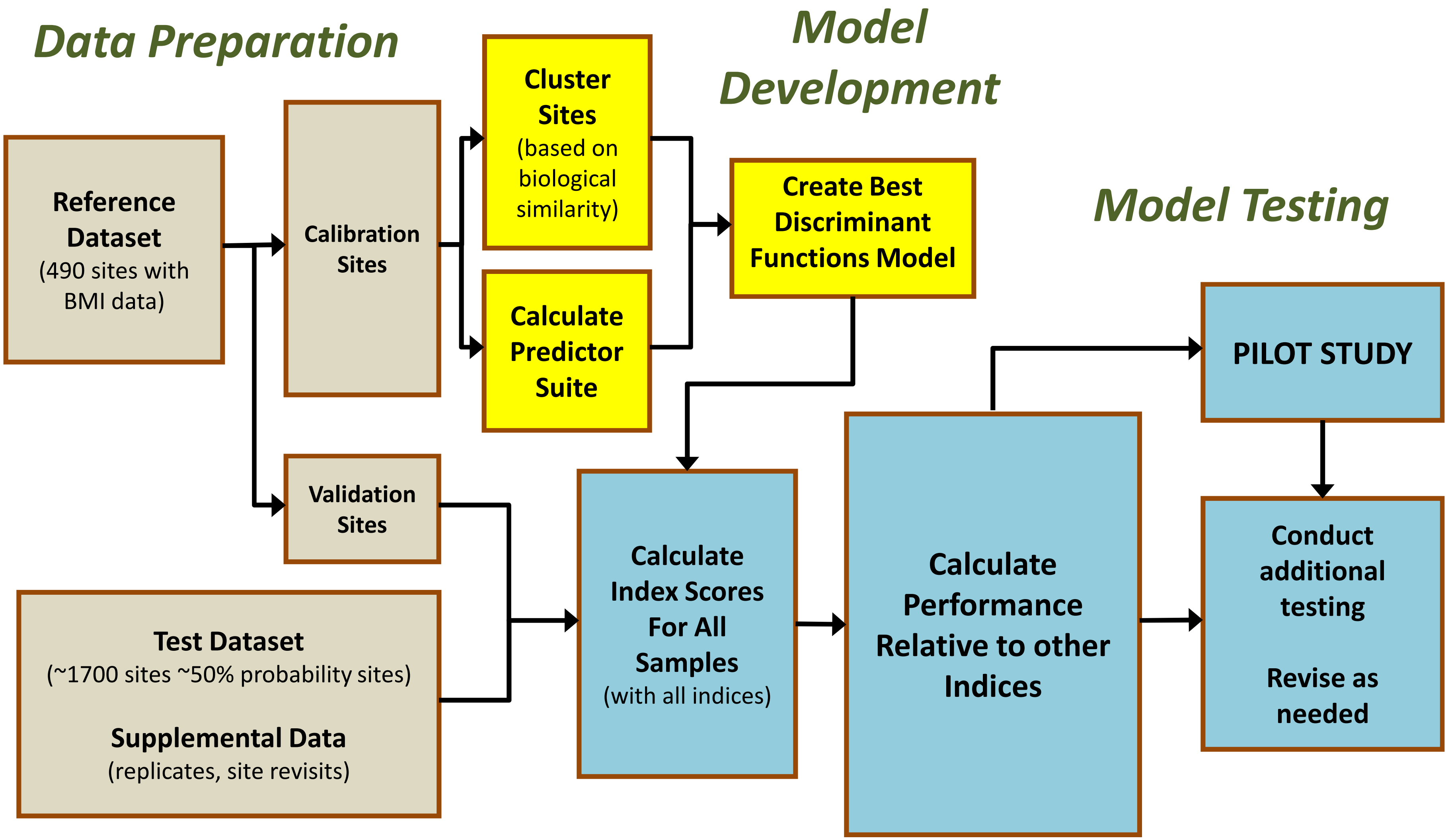
- Diversity is a combination of richness and evenness
- Samples with low evenness can impair our ability to accurately predict richness (a big deal for O/E models)

Correction minimizes the effect of evenness



Taxon	Sample 1	Sample 2
Atherix	10	3
Baetis	11	90
Caenis	12	2
Drunella	9	1
Epeorus	15	1
Farula	13	1
Gyrinus	21	1
Hyalella	9	1
Count	100	100
Richness	9	9

O/E Index Development Process



Scoring Tool Performance Measures

highlights for now, more details at science panel

1. Applicability – the extent of the stream population that can be scored accurately with the index
2. Precision – variability of scores for sites considered to be in similar condition (e.g., reference sites)
3. Accuracy – proximity of score to "true" condition
4. Responsiveness – ability to discriminate impaired sites and sensitivity to gradients of stress
5. Repeatability – similarity of scores for repeated measurements

SWAMP
Surface Water Ambient Monitoring Program

Performance Highlights

- Compare variants of new scoring tools
  - O/E vs Bray-Curtis dissimilarity
  - Clustering vs. no clustering
  - Evenness correction vs. no correction

- Compare new tools with existing scoring tools
  - "Current" O/Es (Hawkins 2005, 3 submodels)
  - SoCal IBI, NorCal IBI

SWAMP
Surface Water Ambient Monitoring Program

Applicability

The extent of the stream population that can be scored accurately with an index

Why do we care? Provides an objective way to evaluate if the environmental setting of a given test site meets the conditions for scoring with an index

How do we measure it?

- Range test: are test site within range of reference predictors (e.g., elevation, watershed area, etc.)
- Distance (in multi-dimensional space) of a test site to the nearest reference cluster

results will be presented at Science Panel

SWAMP Surface Water Ambient Monitoring Program

Precision

variability of scores for sites considered to be in similar condition (e.g., reference sites)

Why do we care?

- Used to establish impairment thresholds (smaller SD means easier to detect deviation from reference)
- Indicates how big a difference the index can detect

How do we measure it?

- Standard deviation of reference sites
- Replicate scoring consistency

SWAMP
Surface Water Ambient Monitoring Program

Precision

standard deviation of reference sites

- Modeled indices are more precise than null indices
- O/E is much more precise than Bray-Curtis
- Evenness-corrected indices are more precise than uncorrected indices

Model	SD	CV
O/E	0.18	0.19
***O/E_ec**	0.17	0.17
O/E_null	0.21	0.21
O/E_null_ec	0.19	0.19
B-C	0.06	0.26
B-C_ec	0.06	0.24
B-C _null	0.07	0.25
B-C_null_ec	0.06	0.23

Responsiveness/ Sensitivity

ability to discriminate impaired sites and sensitivity to gradients of stress

Why do we care?

- Assures that index can detect difference from expected conditions and is responsive across a gradient of stress

How do we measure it?

- Relative strength of discrimination between reference and test sites
- Strength of relationship between index score and gradients of stress

SWAMP
Surface Water Ambient Monitoring Program

Responsiveness:

discrimination between reference and test sites

- Modeled indices are more responsive than null indices
- O/E is equivalent to Bray-Curtis
- Evenness corrected variants are equivalent to uncorrected indices

Model	T-value
O/E	17.6
***O/E_ec**	17.5
O/E_null	12.8
O/E_null_ec	12.1
B-C	16.9
B-C_ec	17.1
B-C _null	14.2
B-C_null_ec	13.9

SWAMP
Surface Water
Ambient Monitoring
Program

Responsiveness

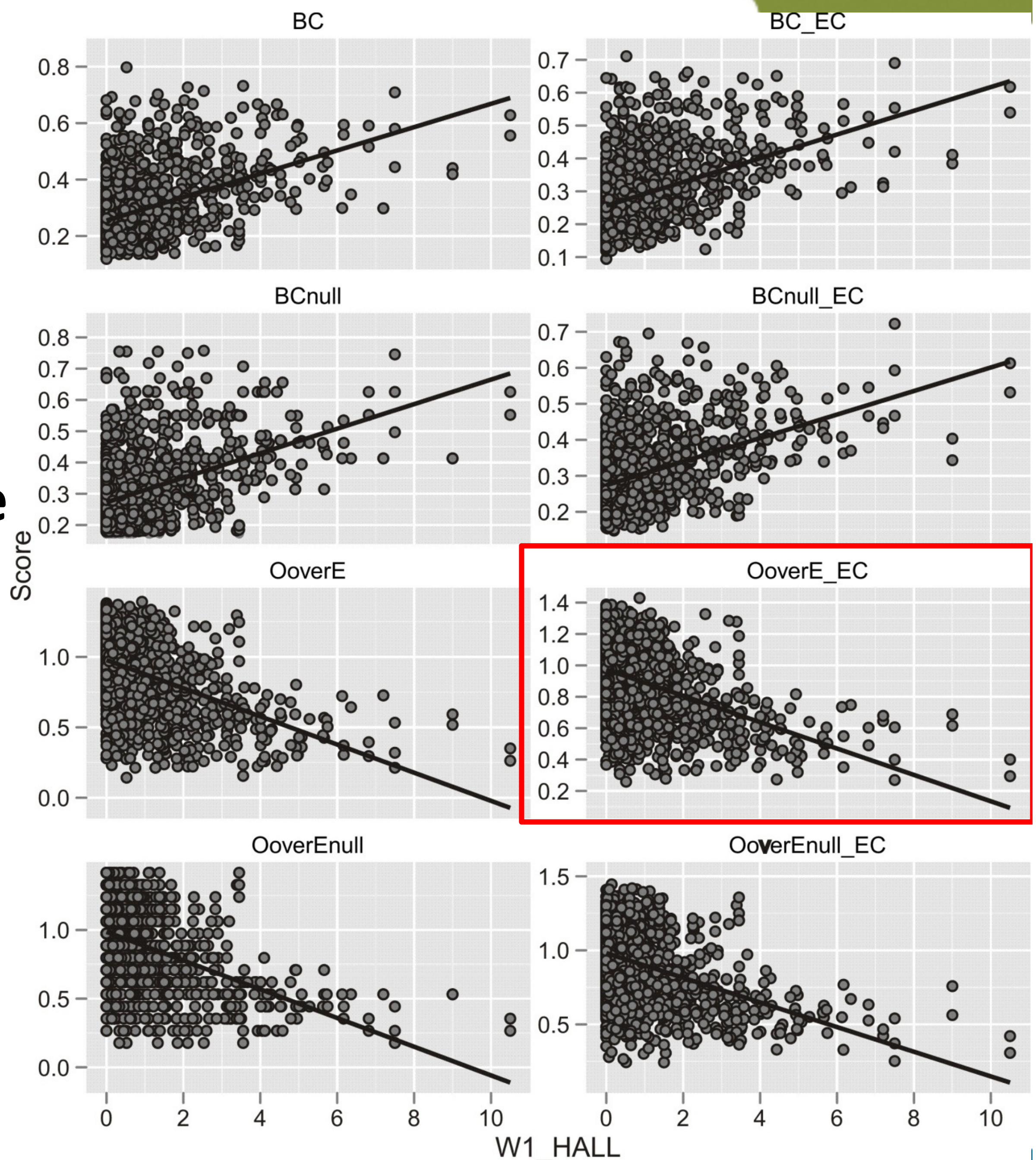
sensitivity to stressor gradients

- **Scores vs. stressor Gradients**
  - Look for “wedge-relationships” (absence of high scores at stressed sites)

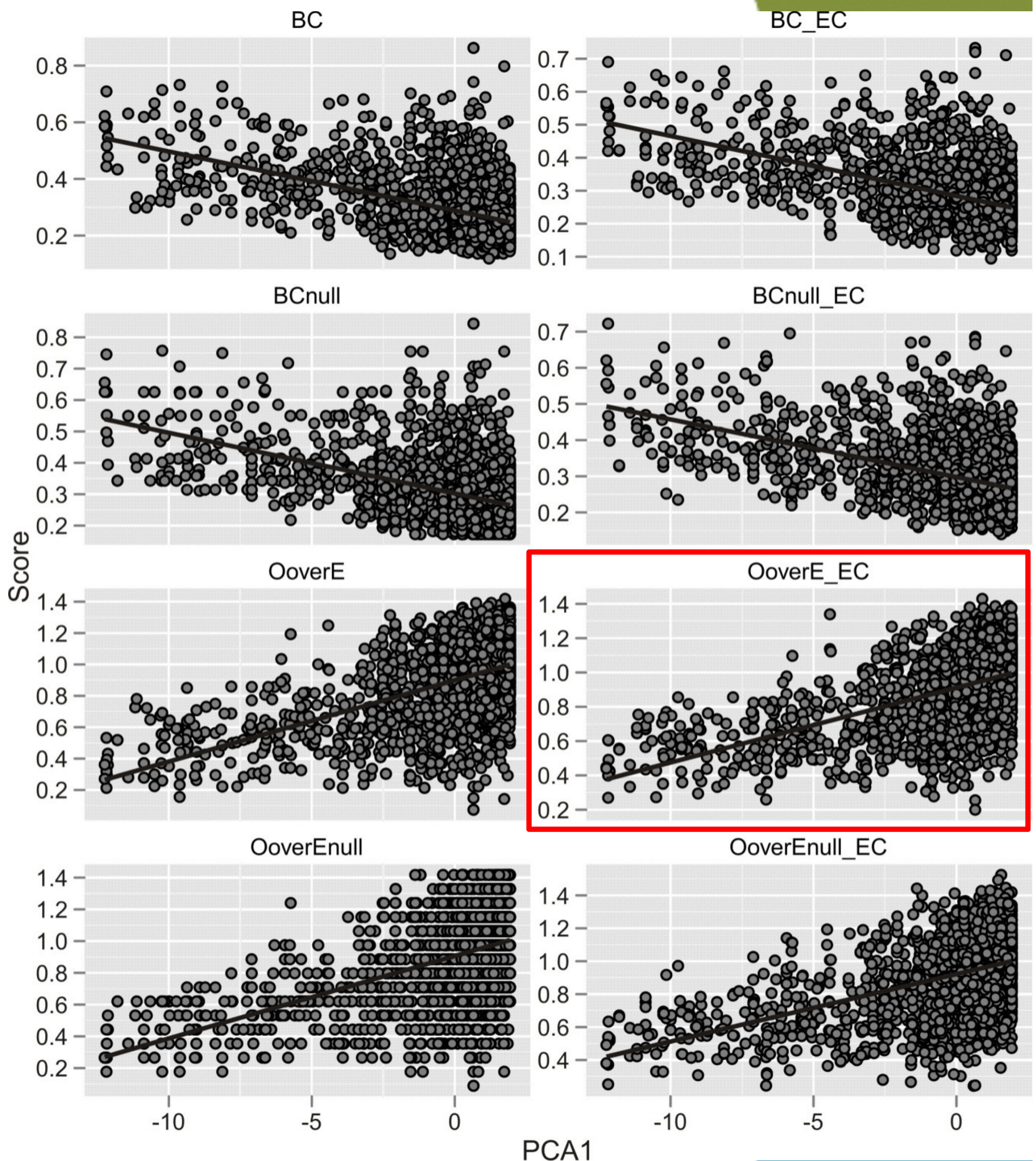
- **Different types of gradients examined**
  - Proximate, mechanistic (metals, pyrethroids, ions)
  - Proximate, non-mechanistic (habitat, nutrients)
  - Ultimate (land cover)
  - Synthetic (PCA axes)

SWAMP Surface Water Ambient Monitoring Program

Responsiveness of new indices to riparian disturbance (W1_Hall)



Responsiveness of new indices to a multivariate composite stressor index (PCA 1)



Accuracy

proximity of score to "true" condition

Why do we care?

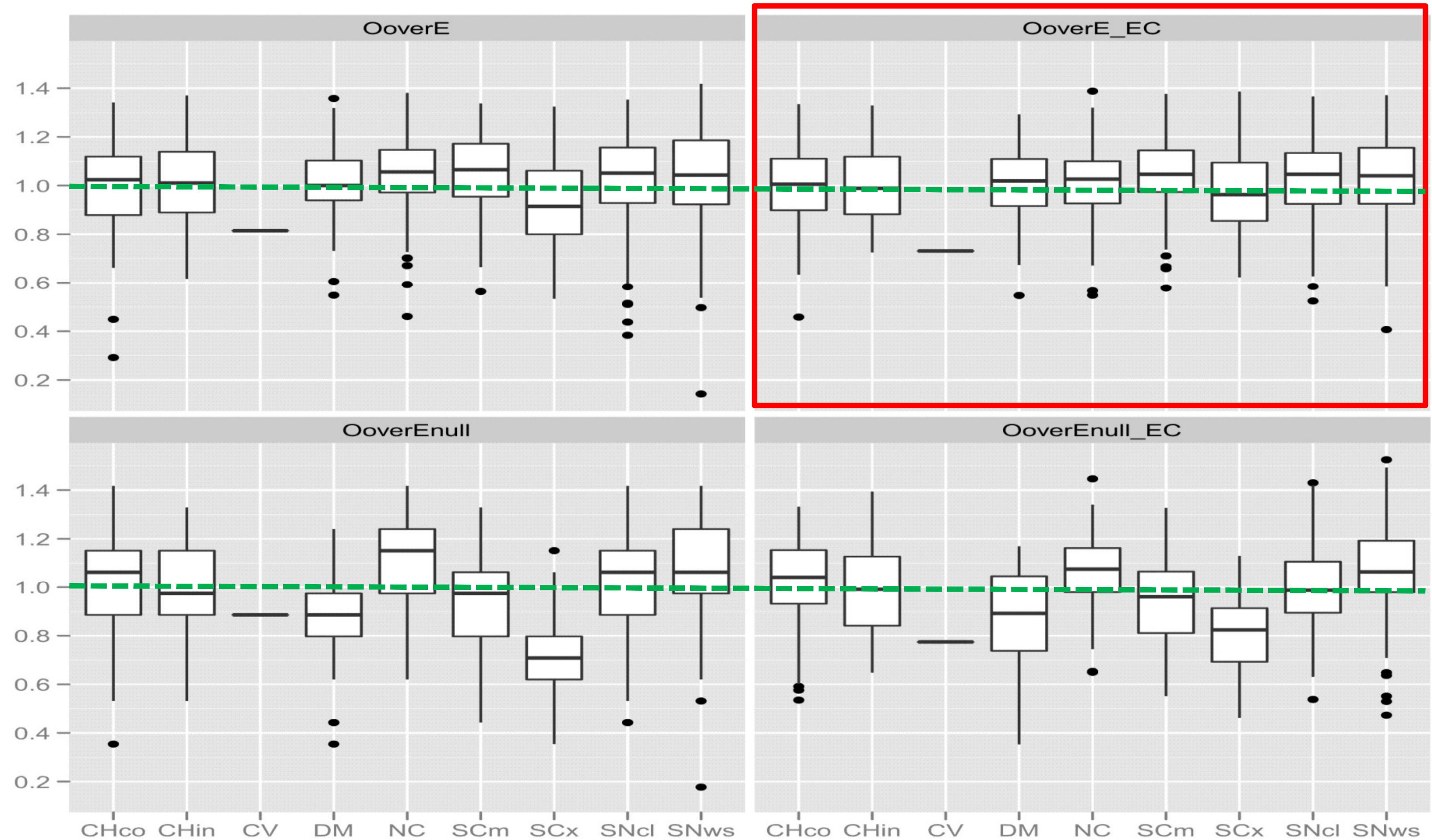
Accurate indices give accurate condition assessments, but direct measures of truth are elusive

How do we measure it? *(indirectly, by looking for bias)*

- Compare scores at ref sites by region
- Compare scores at ref sites vs. natural gradients
- Estimate residual natural variance not explained by scoring tool

SWAMP
Surface Water Ambient Monitoring Program

Regional consistency from a statewide index



- Null models have strong regional biases -- modeling improves this
- Evenness correction makes only slight improvements

Comparisons with Current Tools

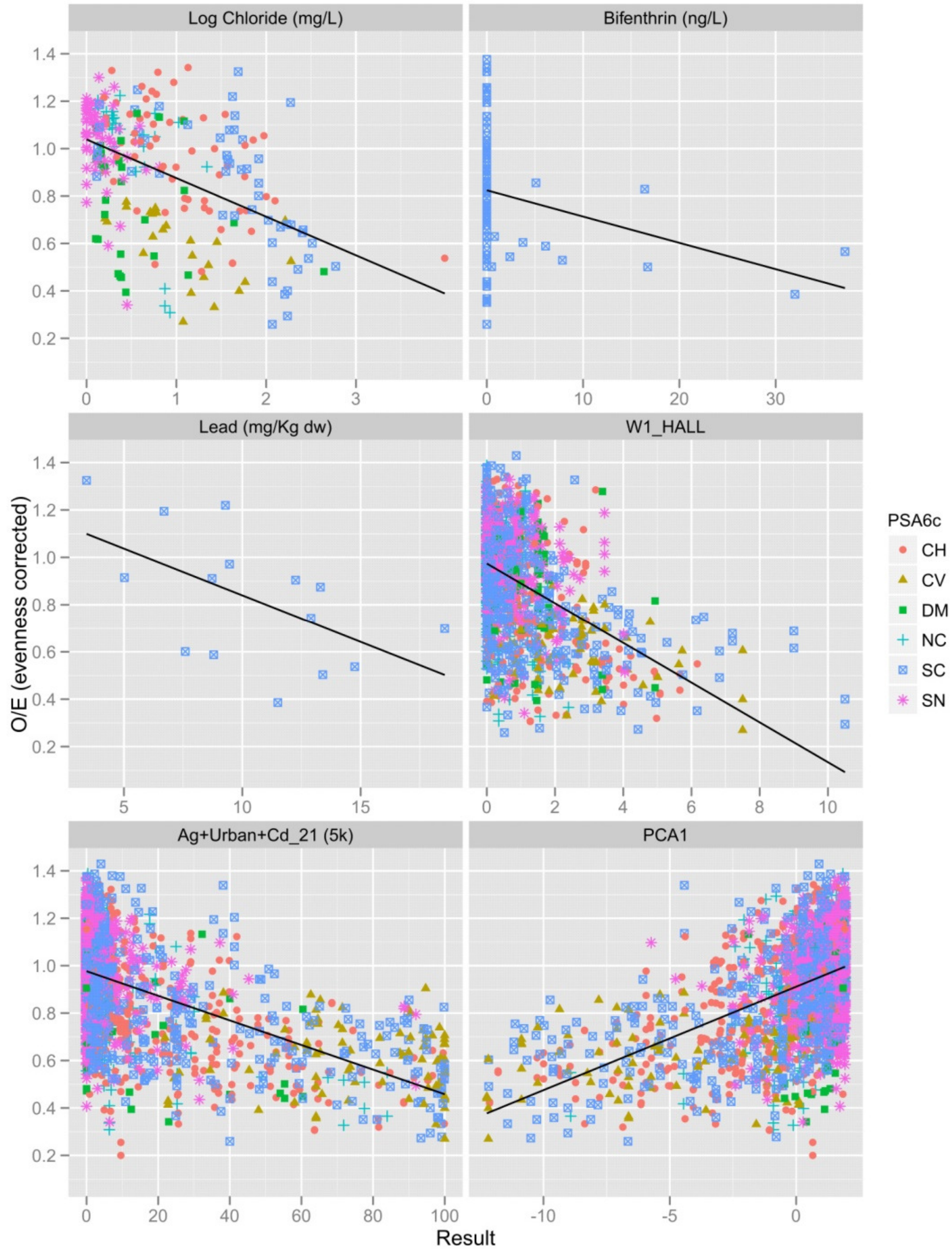
(O/E and IBIs)

INDEX	Precision (sd or CV)			Accuracy (%)		Responsiveness (t-value)
	Reference Calibration	Reference Validation		Residual Natural Variance		Reference v. Test
O/E_ec	0.17	0.16		20		17.5
O/E_original	0.23	0.20		53		14.3
SoCal IBI	0.26	0.16		14		10.5
NorCal IBI	0.17	0.14		31		4.4

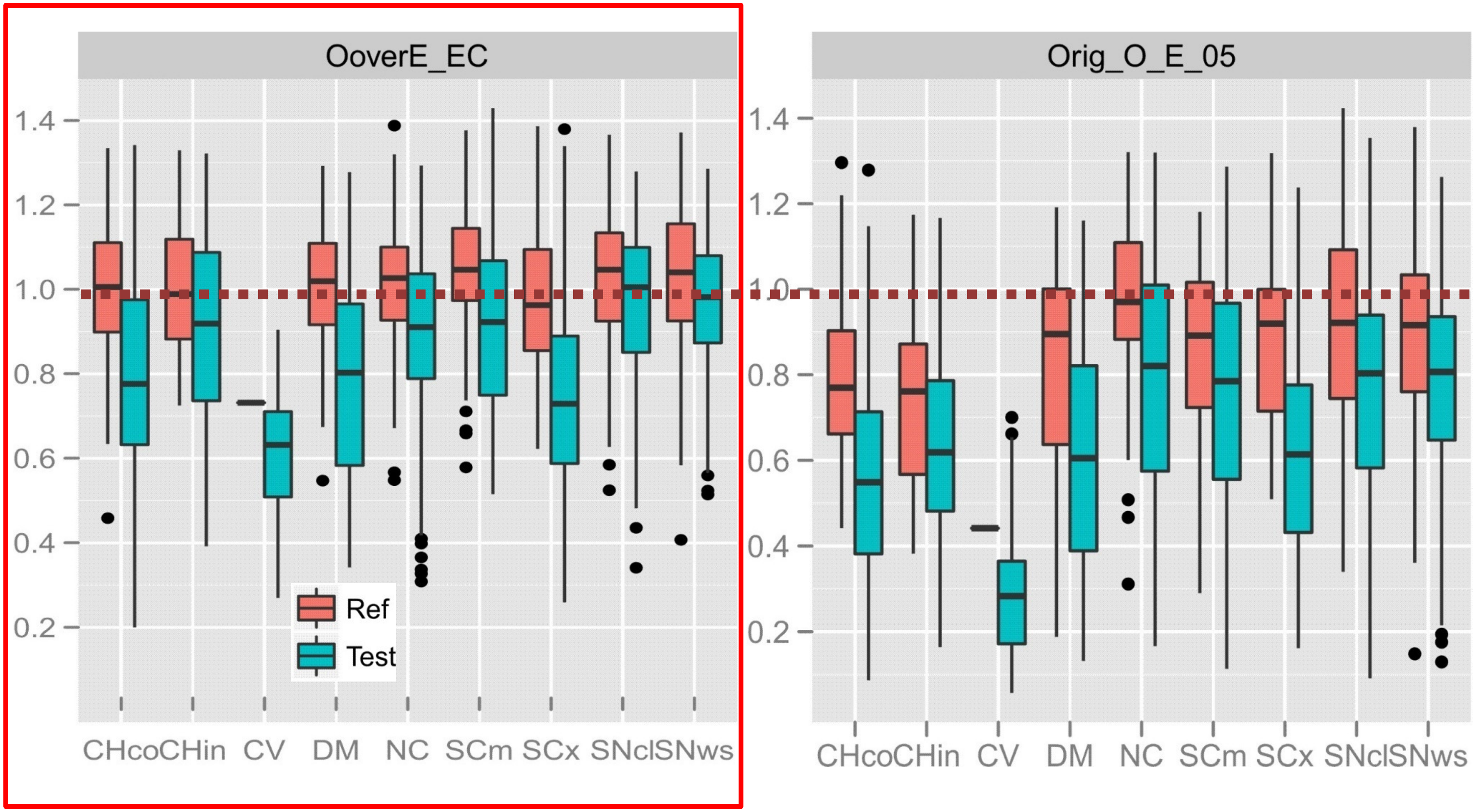
O/E with evenness correction performs as well or better than other indices

Responsiveness

O/E _ec was responsive or very responsive to all gradients we evaluated



Old vs. New O/E Comparisons



- New models have little regional bias and are more precise
- Reference test discrimination is similar, but strong overall bias

Performance Summary

New indices:

- O/E with evenness correction is as good or better than other index variants

Comparisons with old indices:

- Better precision
- Better accuracy
- Better discrimination of test – reference
- New O/E scores higher than old O/E and IBIs

SWAMP
Surface Water Ambient Monitoring Program

Recommendations to Science Panel

- New O/E index performs well
- Want to explore some patterns we see in our performance measures
- Current focus is on optimization of scoring tool and exploring implications for different applications (e.g., influence of temporal variability, recent climate)

SWAMP
Surface Water Ambient Monitoring Program

What's Next

testing in progress: results presented at Science Panel

Precision (consistency tests)

- Consistency of assessment at true replicates
- Long-term (inter- and intra-annual) consistency

Accuracy (bias)

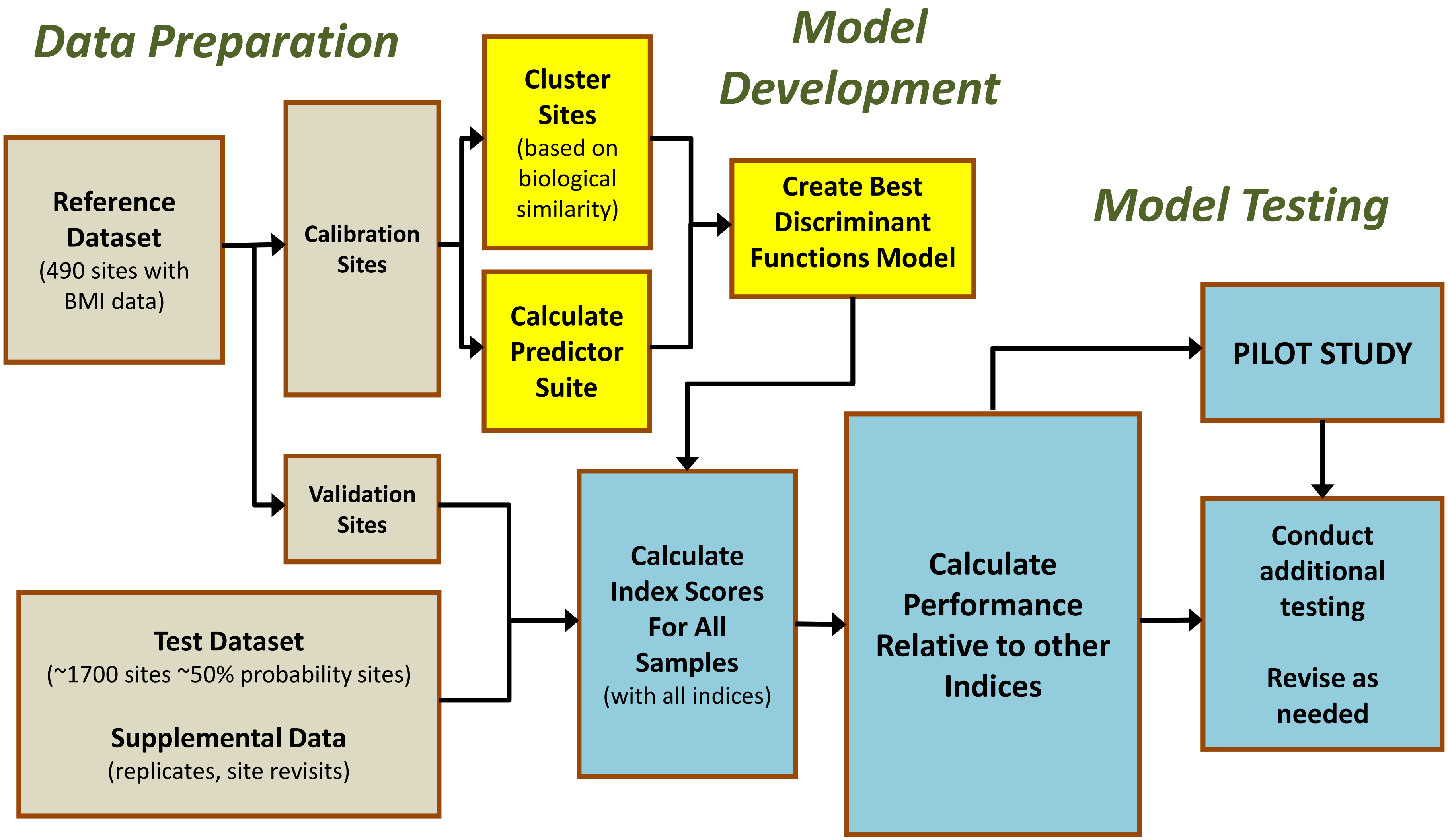
- Explore sources and implications of differences between old and new scoring tools, including separation of natural and anthropogenic sources
- Explore effects of recent climate and temporal variability

Responsiveness

Applicability

SWAMP Surface Water Ambient Monitoring Program

O/E Index Development Process



[redacted]

[redacted]