

Comparability of Biological Assessments Derived from Predictive Models of Increasing Geographic Scope



Peter Ode

California Department of Fish and Game

Charles Hawkins

Utah State University

Alan Herlihy

Oregon State University

John Van Sickle

US EPA-ORD

Daren Carlisle

USGS-NAWQA

Lester Yuan

US EPA-ORD

Now that we have predictive models that cover large geographic areas like WSA and EMAP's models ...

...can we use these large models equally well for local assessments?

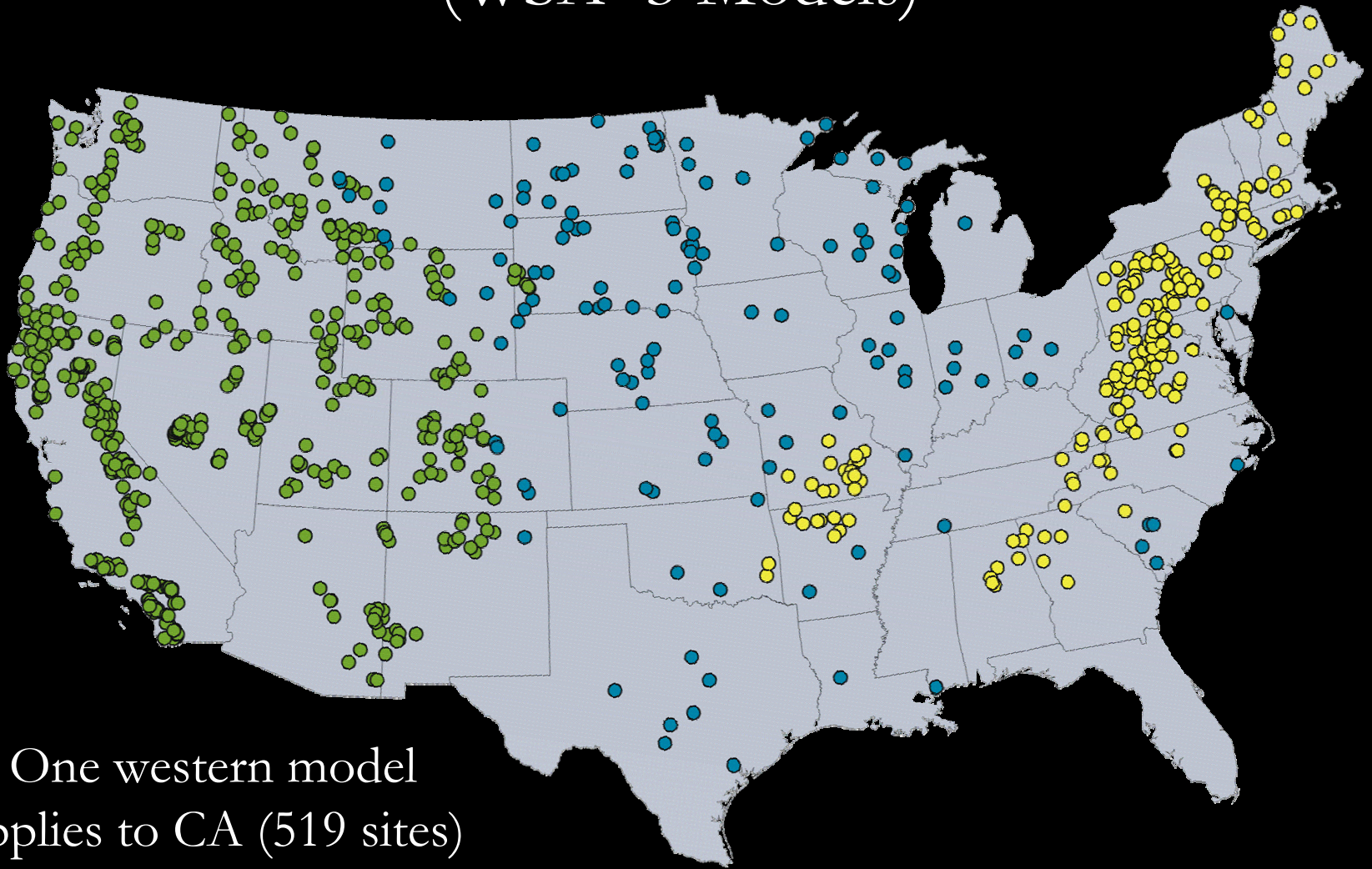
This potential is very attractive, but larger models have some limitations that may restrict their value for regional assessments...

QUESTION: What are the consequences of using models derived at different spatial scales...

- for regional assessments?
- for site-specific assessments?

APPROACH: compare the scores derived from larger models with scores from recently created CA models, using set of test sites from CA

Model 1. National Wadeable Streams Assessment (WSA -3 Models)

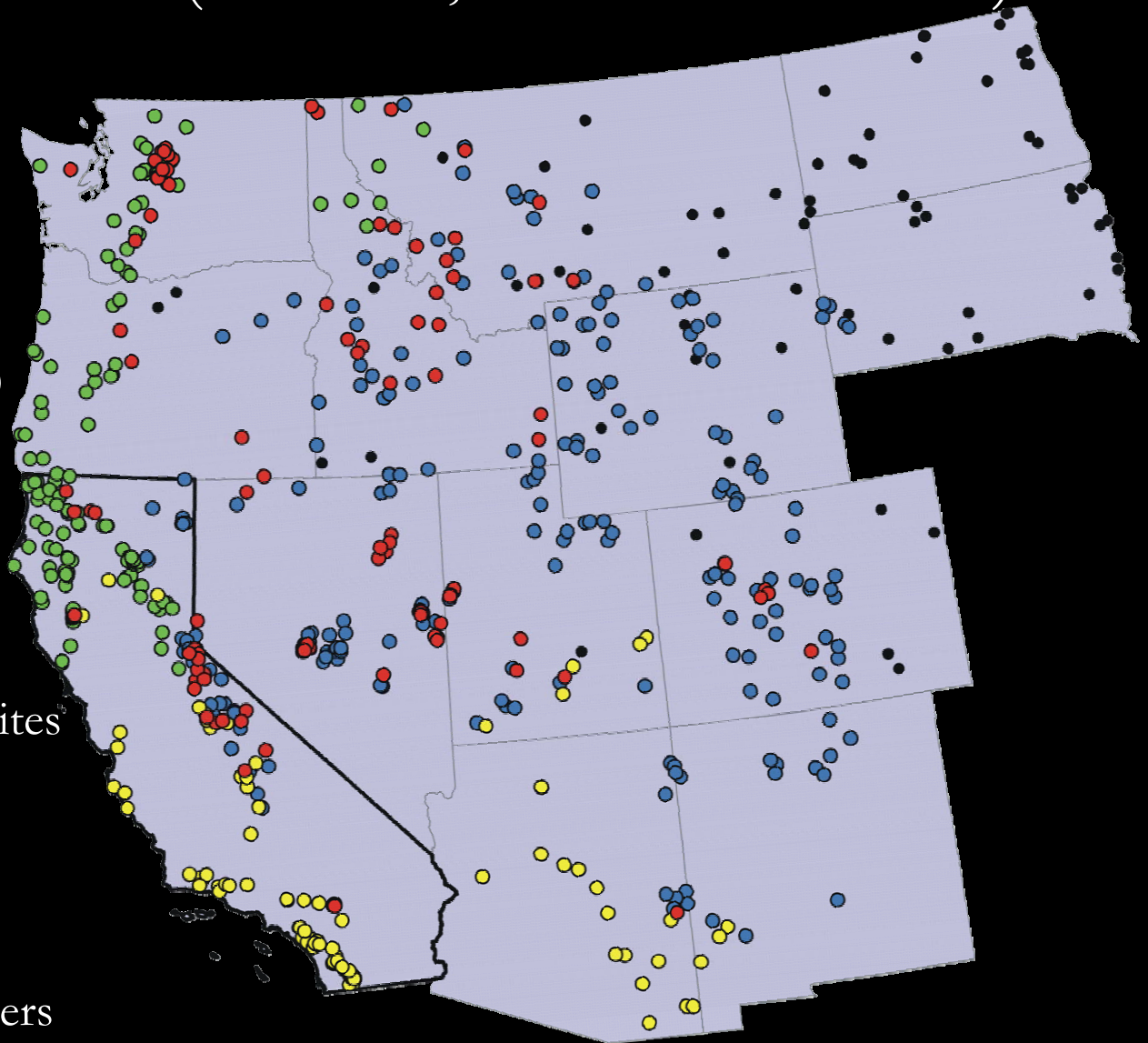


Models 2a, 2b. Environmental Monitoring and Assessment Program, Western Pilot (WEMAP, 629 sites-209 in CA)

5 separate models
(represented by
colors) gave better
performance than a
single western model
(4 apply to California)

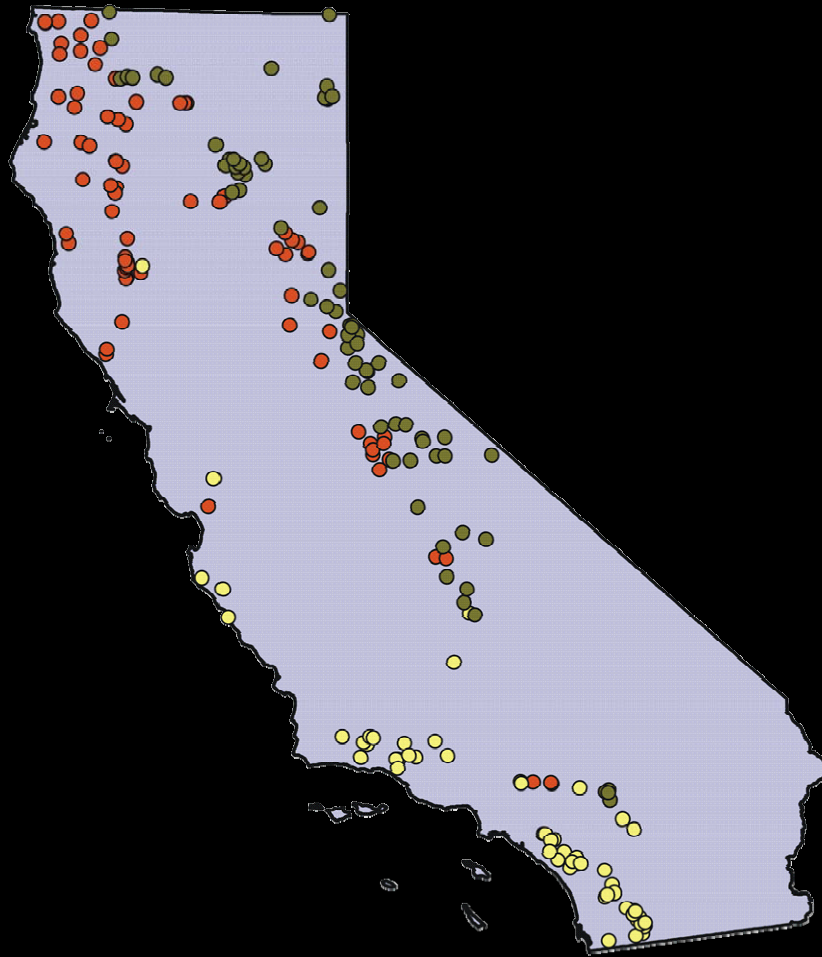
Two WEMAP versions:

1. WEMAP (null): no
division of reference sites
within the 5 models
2. WEMAP (full): Some
models with reference
sites divided into clusters



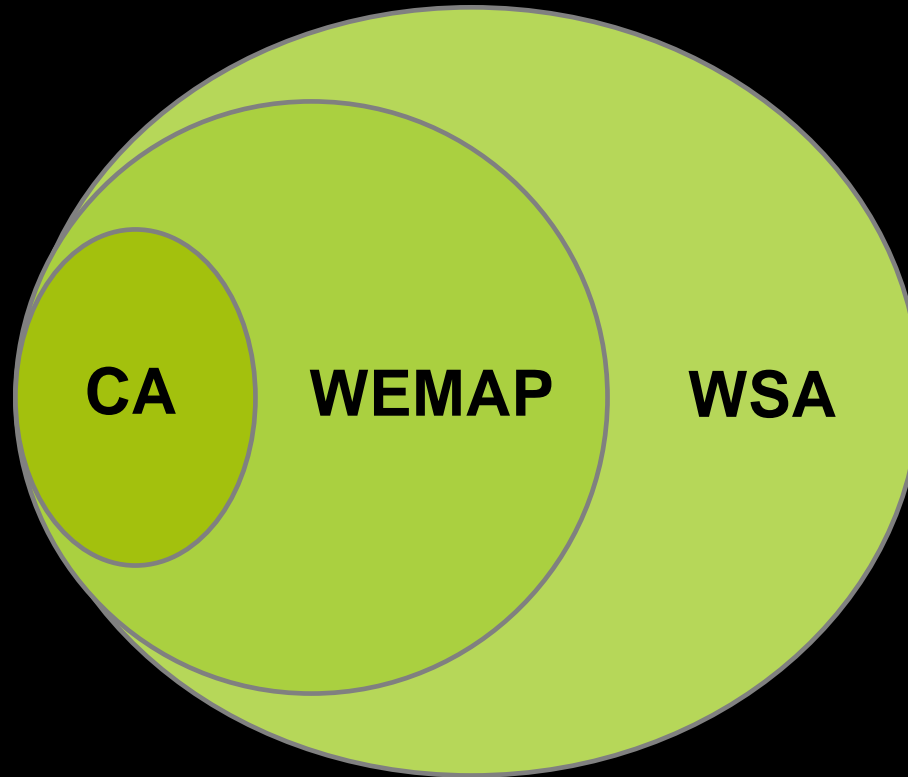
Model 3. California Models (206 sites)

3 models performed
better than a single
statewide model



Spatial Relationships of Model Reference Sites

Increasing geographic range ➡



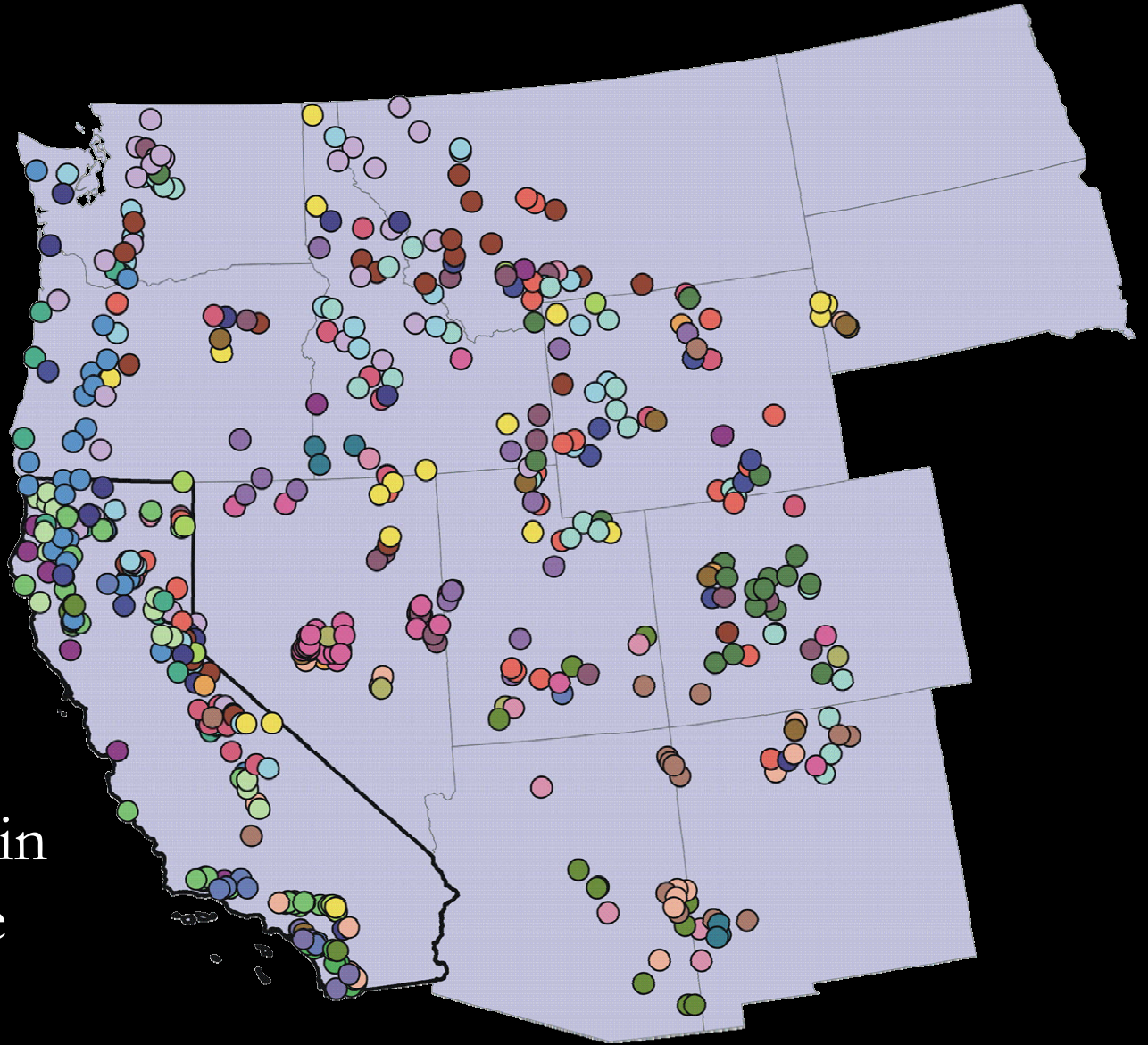
As we expand the spatial scale of the models, we're expanding the geographic area from which we combine reference sites into the clusters used to predict "E".

The larger models contain a smaller proportion of CA reference sites and thus, are increasingly influenced by the characteristics of sites outside the state.

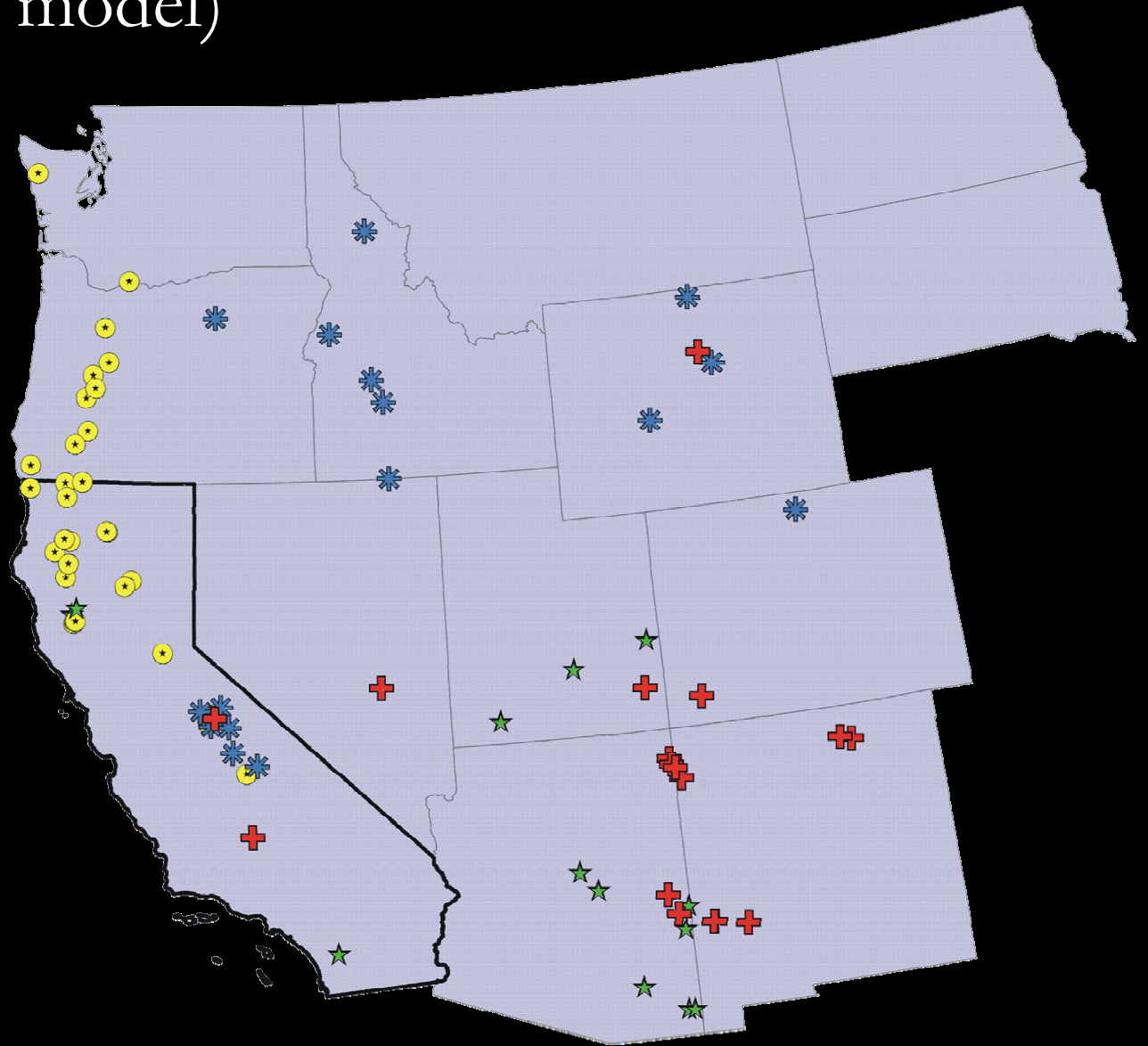
Scale Example: Wadeable Streams Assessment (WSA)

Western Model
= 30 clusters
(24 in CA)

Great variability in
geographic range
of clusters....

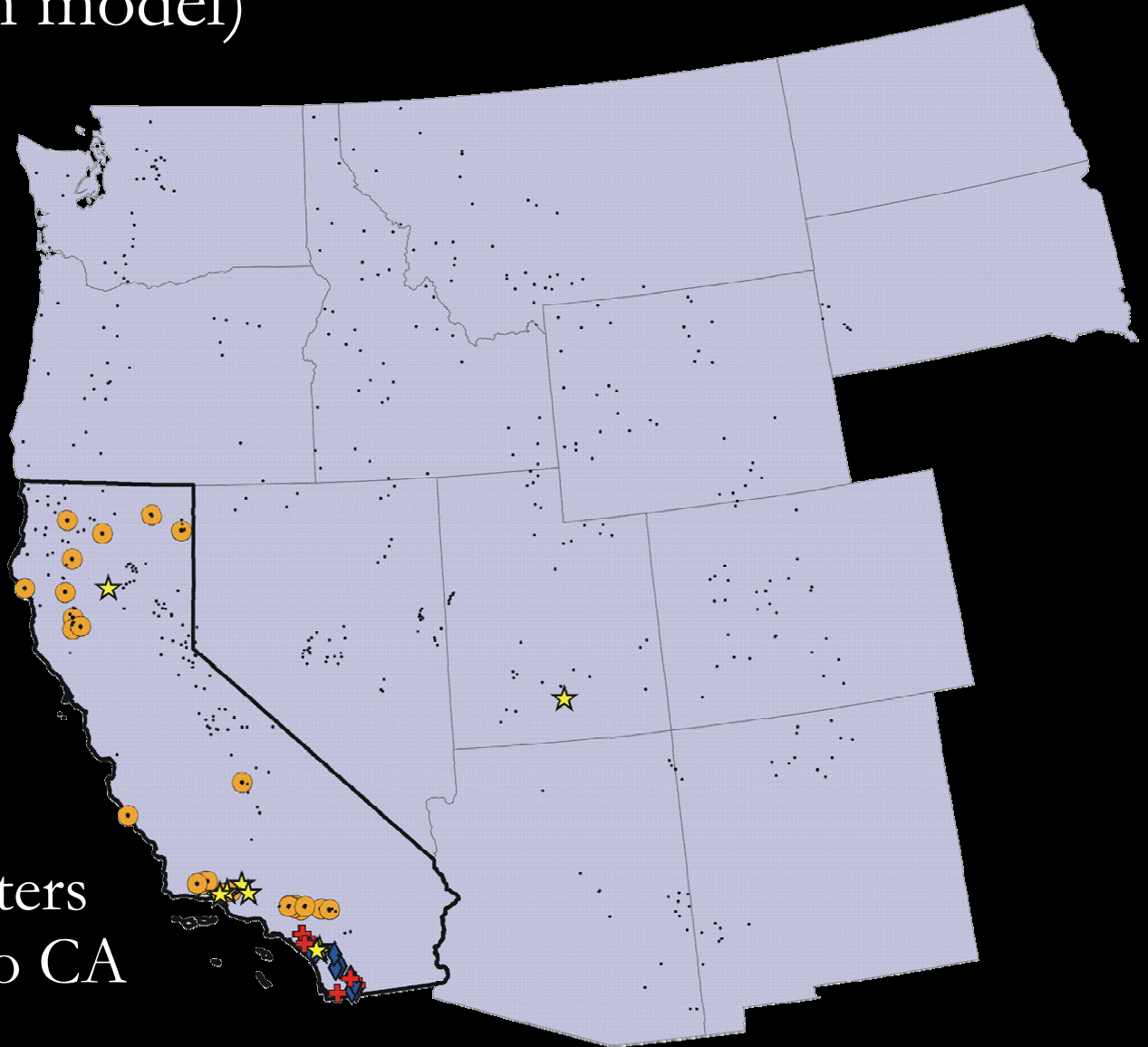


National Wadeable Streams Assessment (WSA, western model)



Most clusters are
widespread...

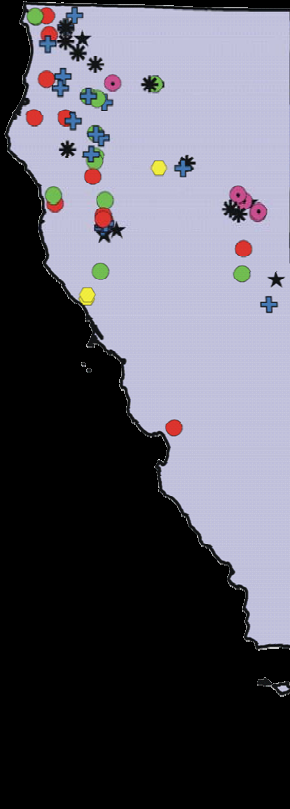
National Wadeable Streams Assessment (WSA –western model)



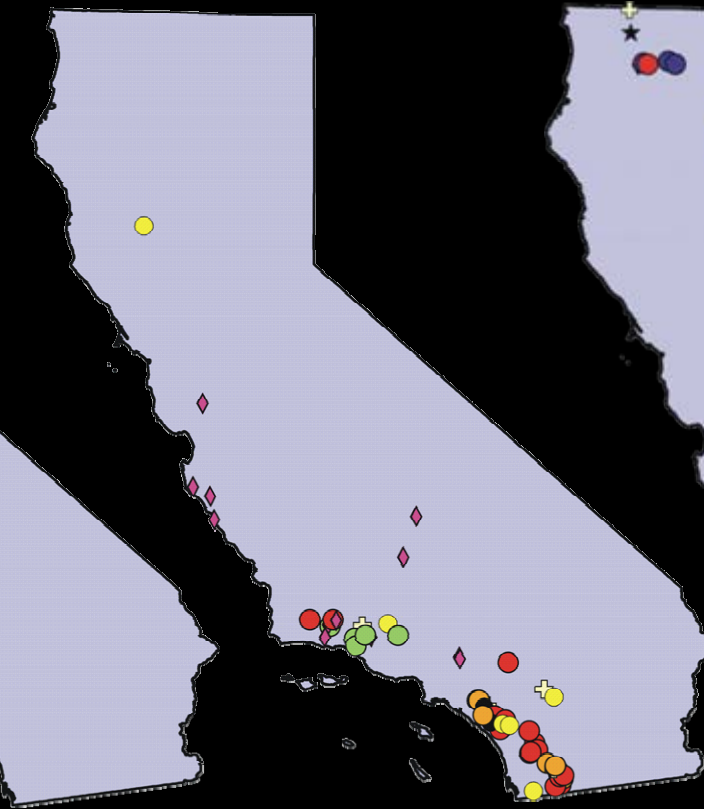
... but a few clusters
are restricted to CA

California Models

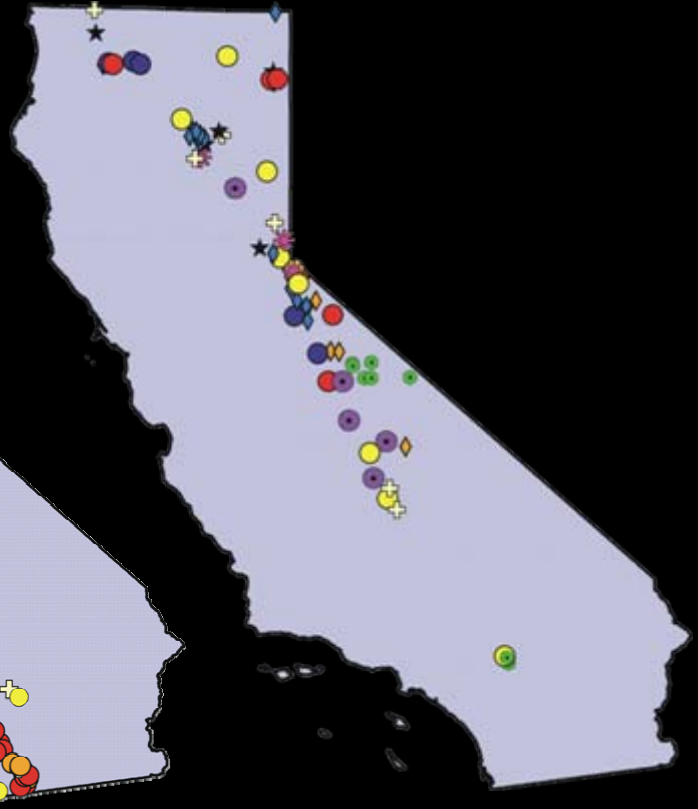
Model 1
Cool and Wet



Model 2
Dry and Warm



Model 3
Cold and Mesic



Clusters occur at much smaller spatial scales than in the larger models...
...tends to increase similarity in taxonomic composition and predictor variables

Predictor variables have a lot of overlap,
but each sub-model has unique combinations

California Predictors	
Group 1	Watershed Area
	Longitude
	Latitude
	Temperature
Group 2	Longitude
	Precipitation
Group 3	Watershed Area
	Temperature

WEMAP Predictors	
Group 2	Watershed Area
	Longitude
	Elevation
	Precipitation
All other CA groups use null models (no predictors)	

WSA Predictors
Watershed Area
Longitude
Day of Year
Min. Temperature
Elevation
Precipitation
Slope

Predictors are a key to understanding model differences

1. Models vary in both in the specific predictors and in the geographic range of the predictor gradients....
2. Predictors that work for larger geographic areas may miss or under-represent regionally important gradients...
3. Variation in predictor association within a taxonomic group (e.g., species within a genus) that occurs across the geographic range of a model can also influence model accuracy and precision...

... do these factors affect performance?

The Test Dataset: 128 CA non-reference sites (none used in model development)

107 WEMAP Probabilistic
Sites
+ 21 USFS Sites

ANALYSES:

- Use test set to compare larger models to CA models
- Assume CA models represent the “truth”
- Score sites with all 4 models
- Compare precision, sensitivity, accuracy and bias

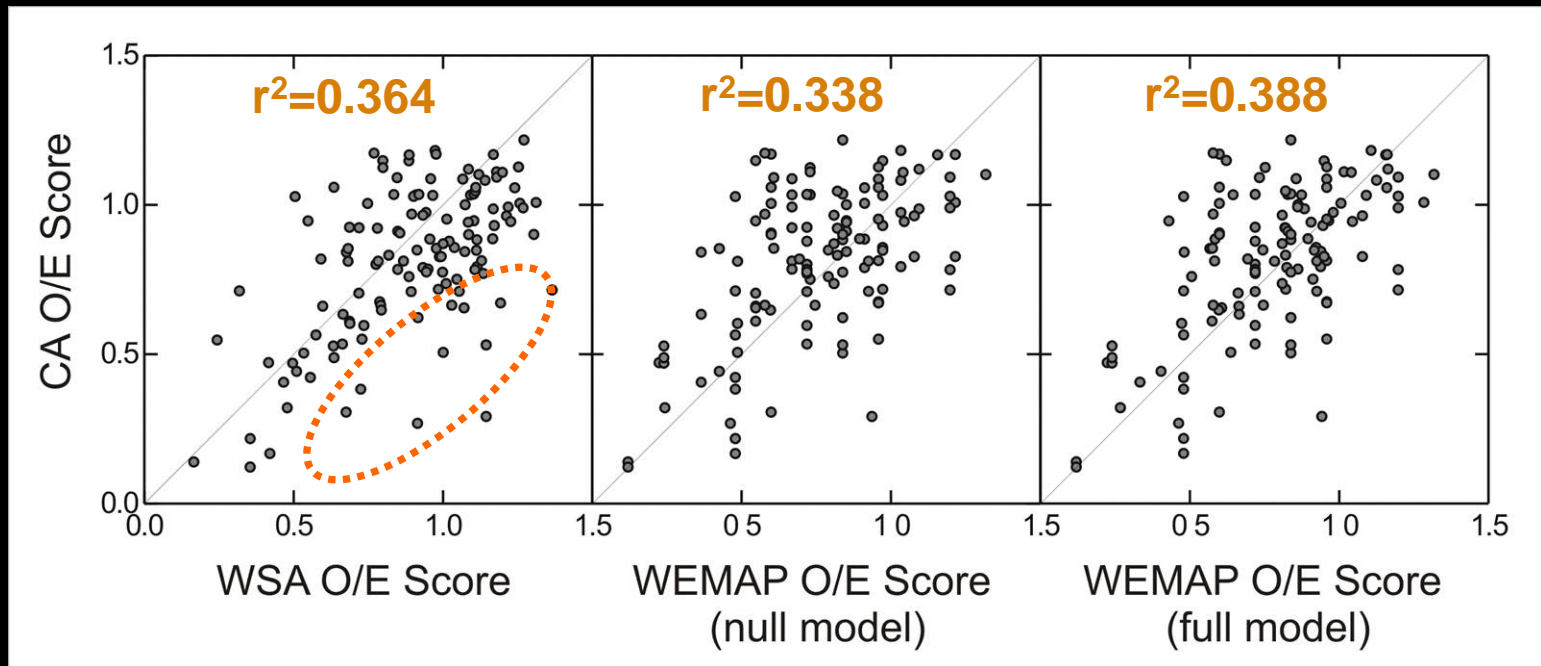


“how sensitive are the models?”
(standard deviations)

California		WEMAP (null)		WEMAP (full)		WSA	
Group 1	0.13	Group 1	0.388	Group 1	0.243	Western	0.198
Group 2	0.17	Group 2	0.196	Group 2	0.150		
Group 3	0.16	Group 3	0.200	Group 3*	0.200		
		Group 4	0.220	Group 4*	0.220		
		Group 5	0.170	Group 5*	0.170		

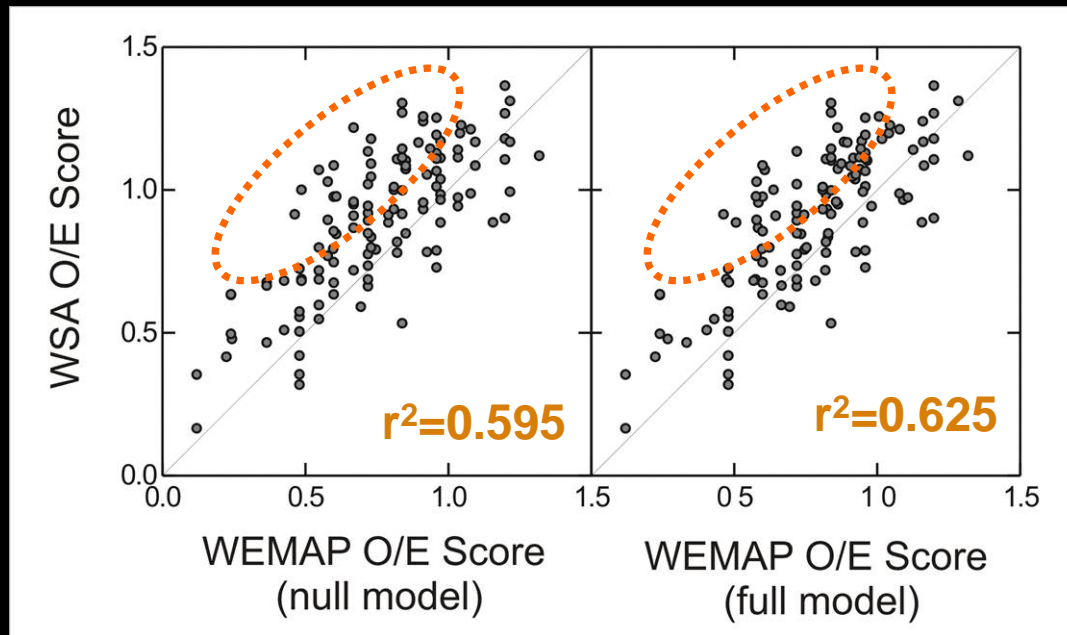
- California models are more precise than either WEMAP or WSA
- WEMAP (full) slightly more precise than WEMAP (null) or WSA

“do the models give sites the same score?”
(correlations with CA Models)



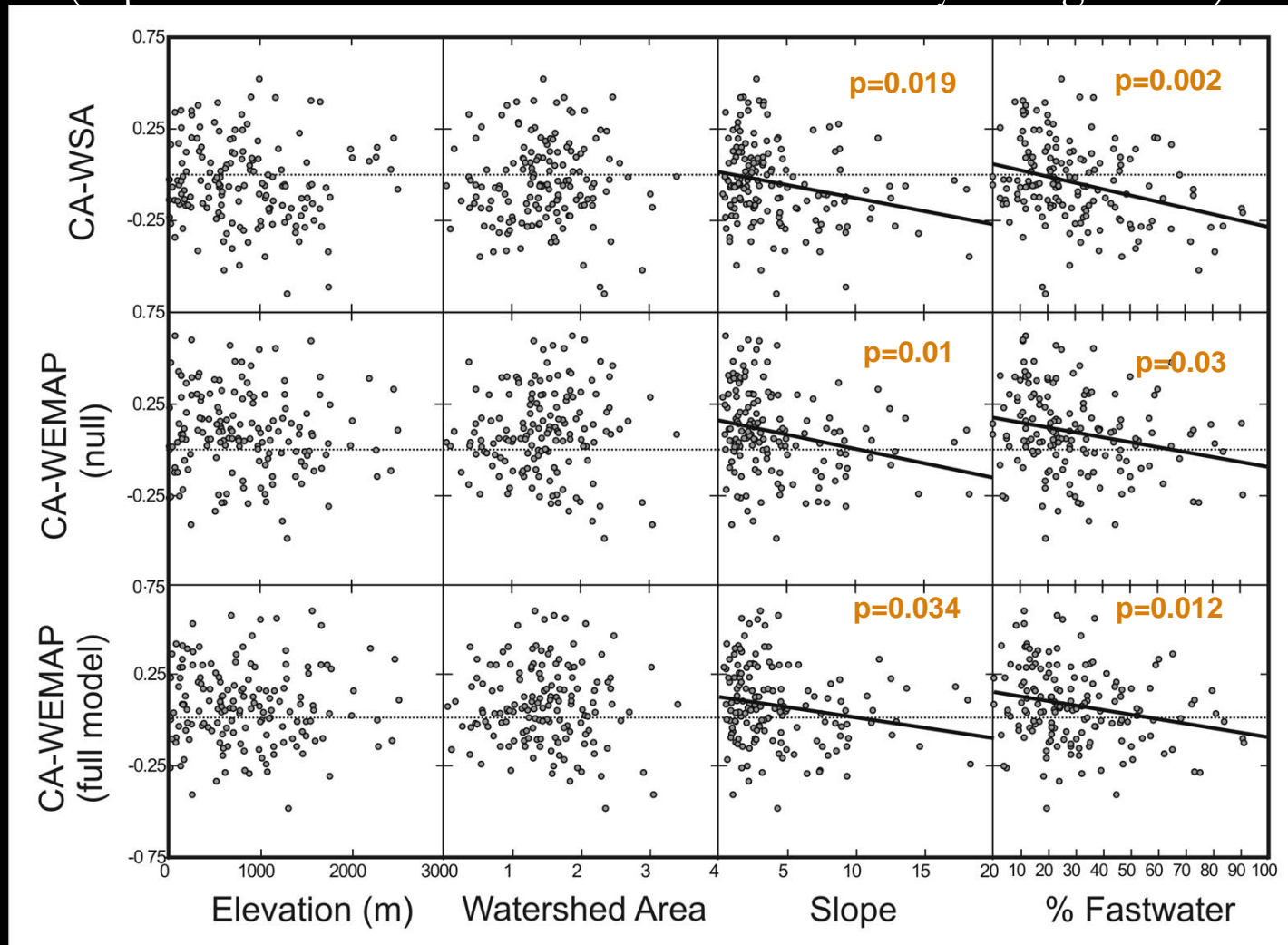
Weak correlations, some evidence of bias toward higher scores under WSA model

Correlations between the two larger models



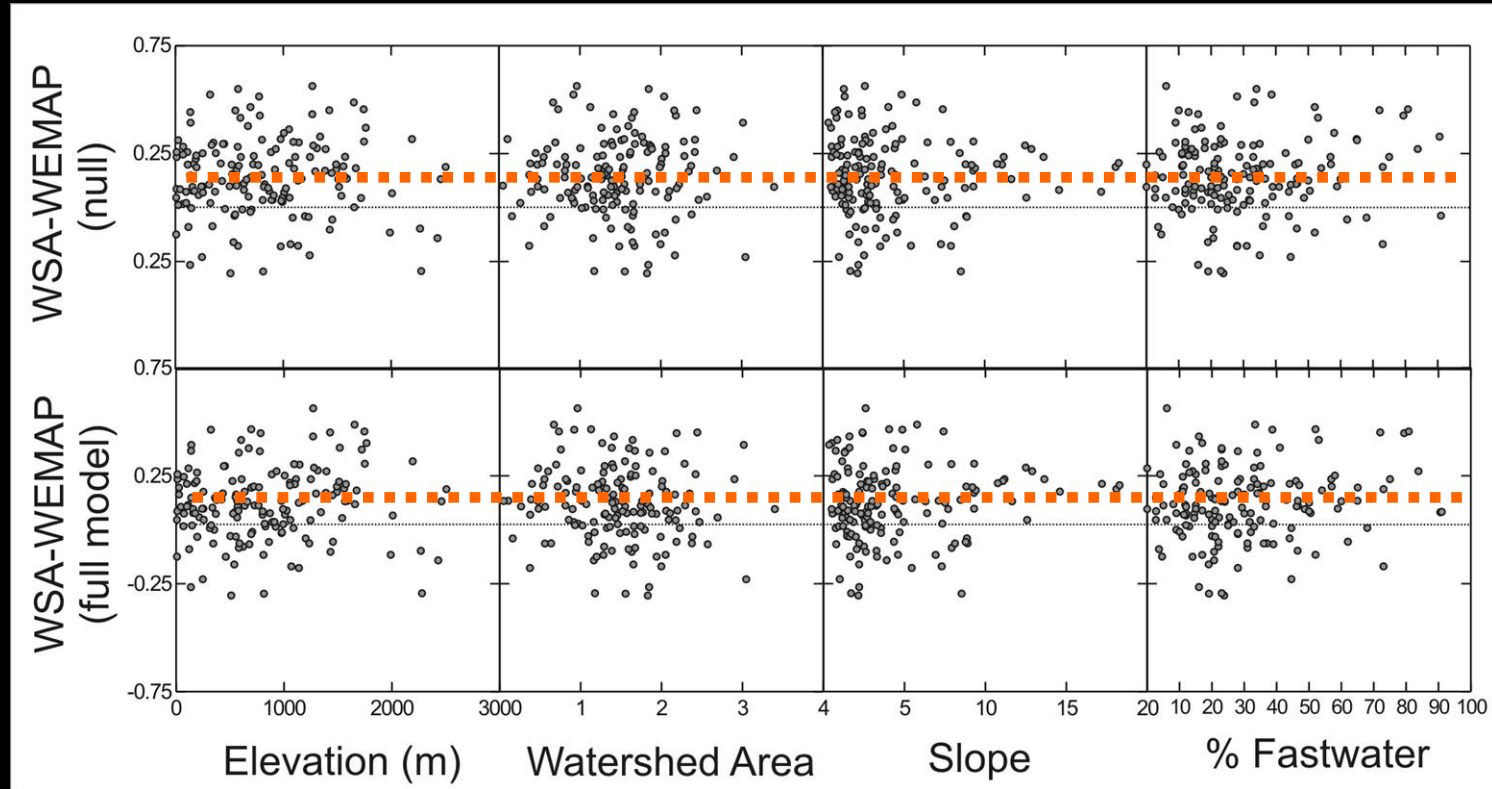
Stronger correlations, but prominent bias in WSA model

“do differences in model scores vary with
key environmental gradients”
(expect a flat line if score differences don't vary with gradient)

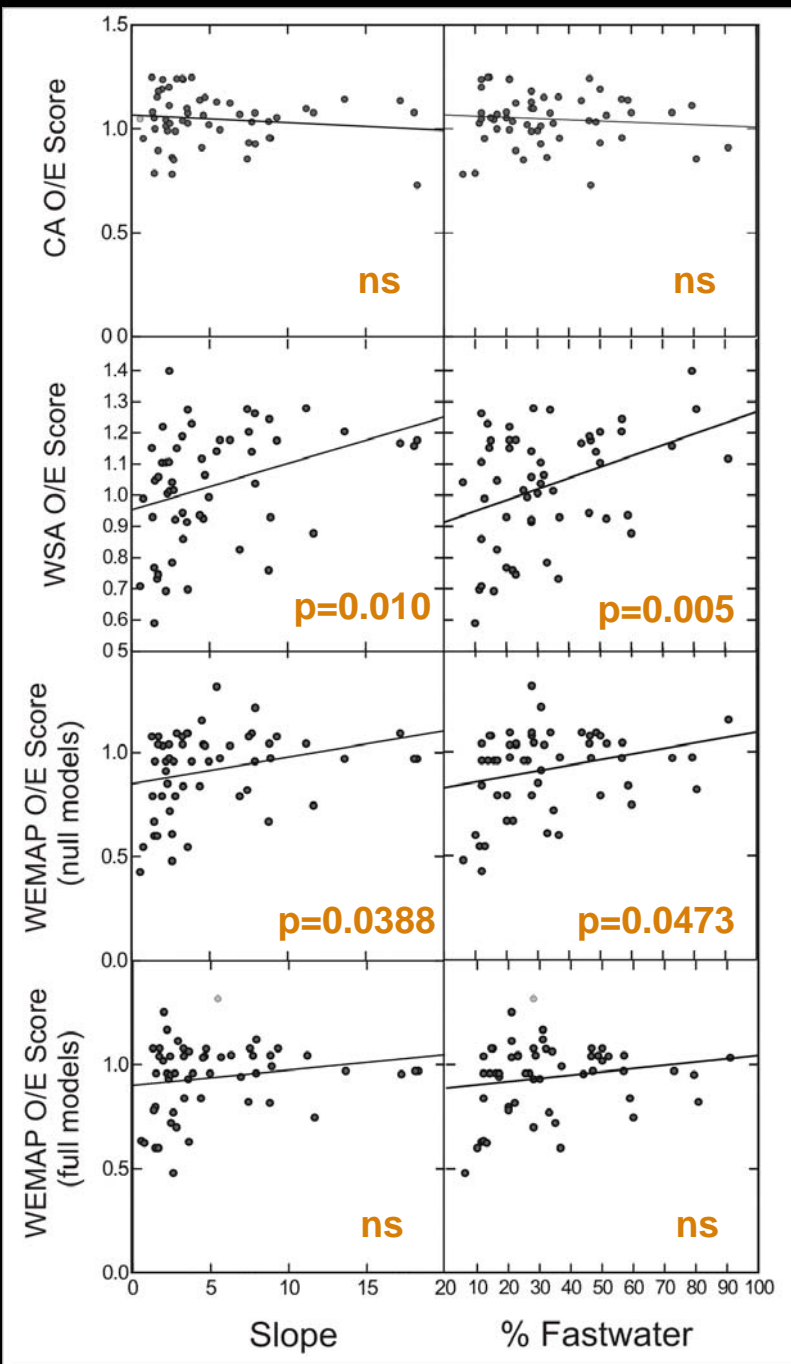


Difference between larger models and CA models tends to
vary with increasing slope and % fastwater

Model Bias vs. Environmental Gradients (large models)



WSA model tends to score sites $\sim .15$ units higher than WEMAP models, but no bias with gradients...
... are both large models failing to account for these gradients or is CA getting it wrong?



O/E scores for **reference sites** versus environmental gradients under the 4 models (expect a flat line if models account for gradients)

- CA models clearly not affected by these gradients
- Both WSA and WEMAP (null, but not full) had significant relationships with slope and % fastwater that were not accounted for by the models

Impairment Decisions: Impaired (I) or Not Impaired (NI)

Use an impairment threshold (mean – 2sd for all models) to compare number of **false negatives** and **false positives** relative to the expectation of the CA models

		CA Model 1 (n=59)		CA Model 2 (n=44)		CA Model 3 (n=25)		TOTALS (n=128)		ALL
		I	NI	I	NI	I	NI	I	NI	
CA ("truth")	I	16	-	17	-	6	-	39	-	39 I
	NI	-	43	-	27	-	19	-	89	89 NI
WSA	I	7	1	7	2	0	1	14	4	18 I
	NI	9	42	10	25	6	18	25	85	110 NI
WEMAP (null)	I	10	4	10	3	4	2	25	9	34 I
	NI	5	39	7	24	2	17	14	80	94 NI
WEMAP (full)	I	12	6	10	3	4	3	26	12	38 I
	NI	4	37	7	24	2	16	13	77	90 NI

- WSA generally very forgiving (misses 2/3 of impaired sites, few false +)
- WEMAP misses 1/3 of impaired sites, but ~ same number of false + as false -

Summary of Results

1. WEMAP models were similar

- Full models had slightly better correlations with other models
- Slightly less bias in impairment decisions

2. Measures of Precision/Sensitivity:

- Larger models appear to have similar precision (weak correlations with CA models, but OK with each other)
- Both models have higher sd values than CA (takes more species loss to detect impairment)
- **Assessments:** WSA model strongly underestimates impairment, WEMAP (null) slightly underestimates impairment

Summary of Results (continued)

3. Measures of Accuracy/ Bias:

- WSA tends to overestimate site quality relative to all others (sometimes by a lot)
- WSA and WEMAP tend to increasingly underscore sites with increasing slope and increasing fast-water habitats relative to CA models

Does it matter?

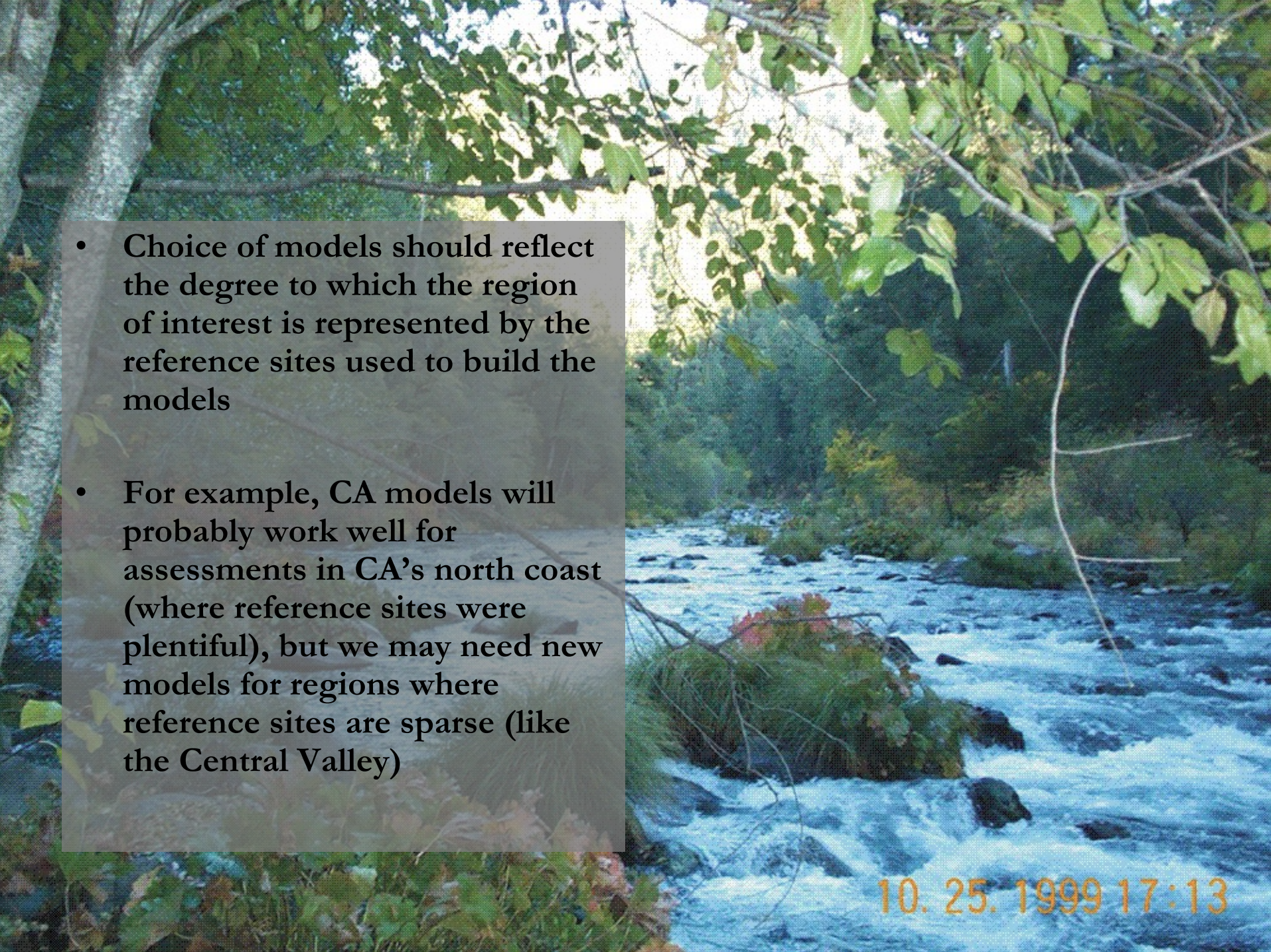
“ it depends on what question you’re asking”

Regional Assessments- Accuracy (lack of bias) is more important than precision... (we can make up for low precision by looking at large numbers of samples)

- WEMAP and WSA may be interchangeable (with a correction factor?)
- Whether larger models are appropriate for smaller assessments (e.g., state condition assessments) will depend on the strength of the bias with environmental gradients

Site Specific Monitoring... where both accuracy and precision are important, this is pretty strong evidence that we still need local models ...

WSA and WEMAP models are not appropriate because they get it wrong too often

- 
- A photograph of a river flowing through a forest. The river is in the foreground, with white water rapids and rocks. The background is a dense forest with green trees. A semi-transparent text box is overlaid on the left side of the image, containing two bullet points.
- Choice of models should reflect the degree to which the region of interest is represented by the reference sites used to build the models
 - For example, CA models will probably work well for assessments in CA's north coast (where reference sites were plentiful), but we may need new models for regions where reference sites are sparse (like the Central Valley)

10. 25. 1999 17:13

Acknowledgements

Field Crews:

- EMAP-West
- USFS/ Utah State University
- CA Aquatic Bioassessment Laboratory (ABL, Shawn McBride, Michael Dawson, Jennifer York)

Invertebrate Taxonomists:

- EcoAnalysts
- Utah State Bug Lab
- ABL (Doug Post, Dan Pickard, Joe Slusark, Brady Richards)

Funding and Data Sources

- USFS, Joseph Furnish - CA models
- EPA-Office of Research and Development, WEMAP and WSA

