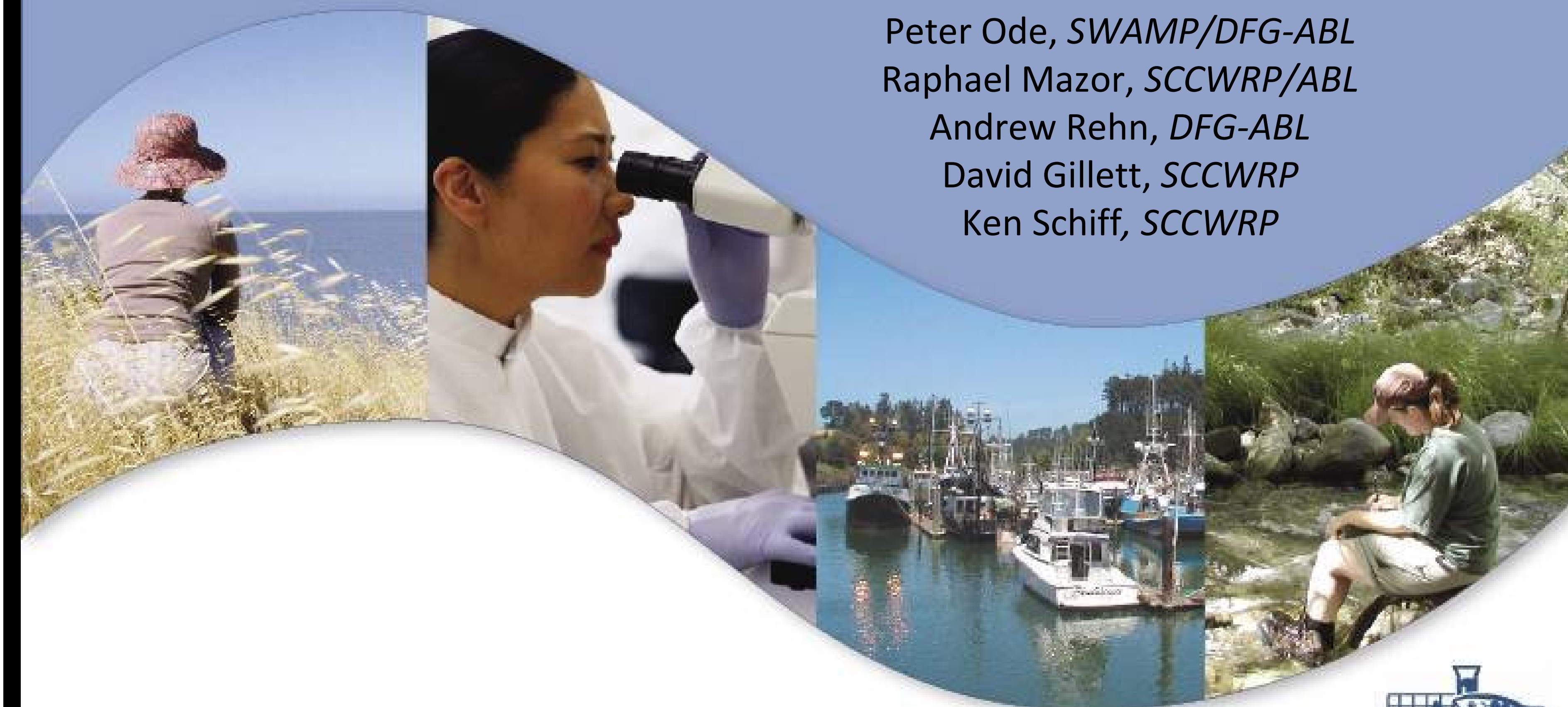
Defining Reference Conditions for CA's Perennial Wadeable Streams

Peter Ode, *SWAMP/DFG-ABL*
Raphael Mazor, *SCCWRP/ABL*
Andrew Rehn, *DFG-ABL*
David Gillett, *SCCWRP*
Ken Schiff*, SCCWRP*



RESOURCES AGENCY CALIFORNIA DEPARTMENT FISH & GAME

Aquatic Bioassessment Laboratory

Water Boards

SWAMP Surface Water Ambient Monitoring Program

Request for feedback

- We've explored many options, now ready to make some decisions and put it together
- We want to know if you think our approaches are technically sound (both generalities and details) and if we've overlooked significant variables or approaches or should explore alternatives
- I'll highlight specific areas where we're looking for direction, but feel free to ask questions as we go along

2

Reference Condition Development

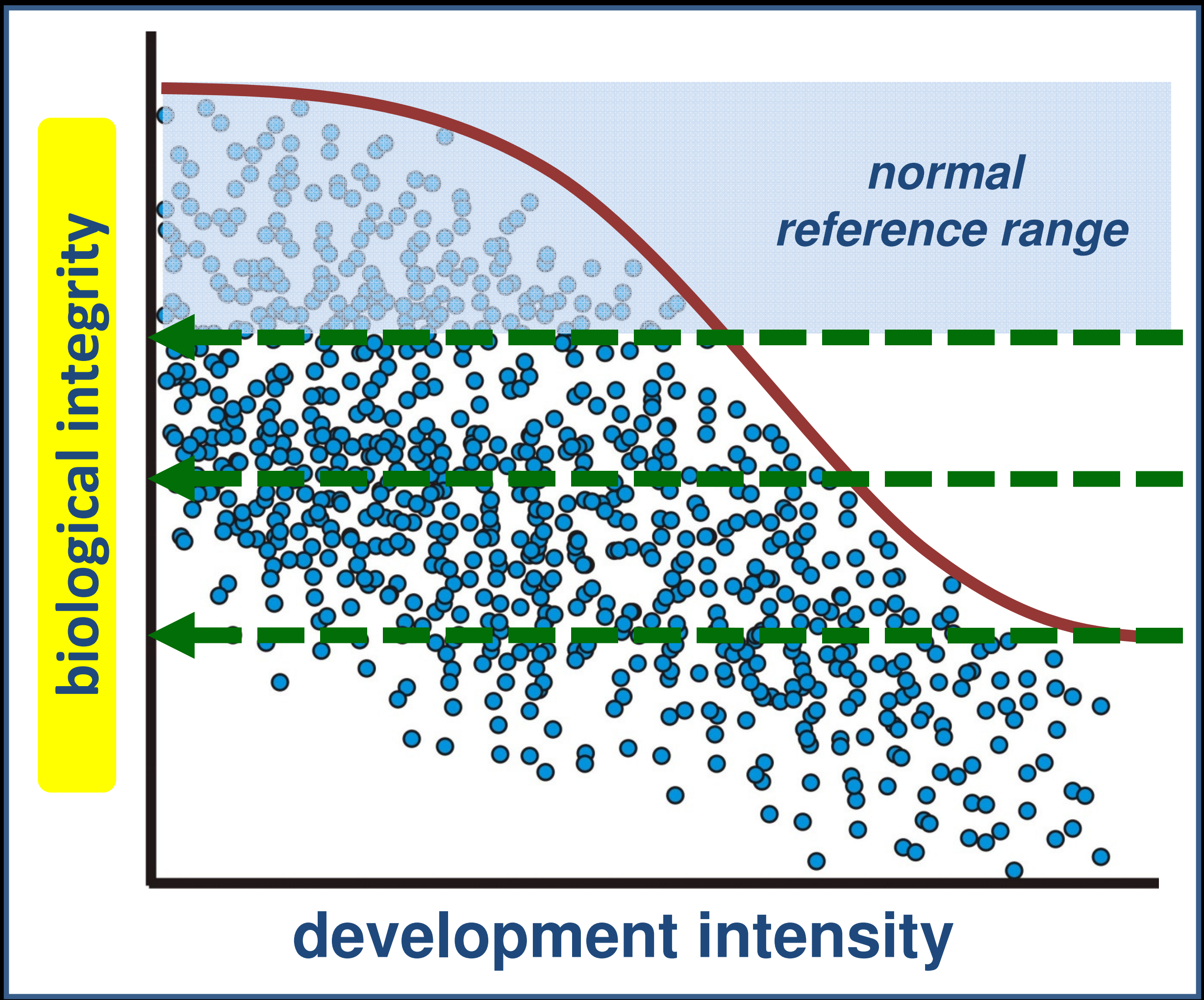
- Background-
  - Objectives and Performance Criteria
  - Review from October
- Development Process
- Evaluating Performance
- Summary of Feedback Questions
- RCMP Beyond Bio-Objectives

SWAMP Surface Water Ambient Monitoring Program

Reference Condition = natural ecological state – defined by sites with low levels of disturbance

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

Reference condition is the foundation of bio-objectives because it provides an **objective** basis for setting biological expectations (standards)



- Reference anchors the upper end of the scoring scale for biological integrity tools (the ruler)
- Reference informs the regulatory bar (impairment threshold)

SWAMP's Reference Condition Management Plan

(Ode and Schiff, adopted March 2009)

RCMP is SWAMP's standardized process for identifying & sampling reference sites for California's wadeable streams

Recommendations for the development and maintenance of a reference condition management program (RCMP) to support biological assessment of California's wadeable streams

Report to the State Water Resources Control Board's Surface Water Ambient Monitoring Program (SWAMP)

Peter Ode, SWAMP Bioassessment Coordinator
Aquatic Bioassessment Laboratory/
Water Pollution Control Laboratory
California Department of Fish and Game
2005 Nimbus Road
Rancho Cordova, CA 95670

Ken Schiff, Deputy Director
Southern California Coastal Water Research Project
3535 Harbor Blvd., Suite 110
Costa Mesa, CA 92626

March 2009

Technical Report 581

SWAMP
Surface Water Ambient Monitoring Program

We're following the plan:

- 2008-2011: Sampling to fill in data gaps
- 2010-2011: Formalizing screening criteria

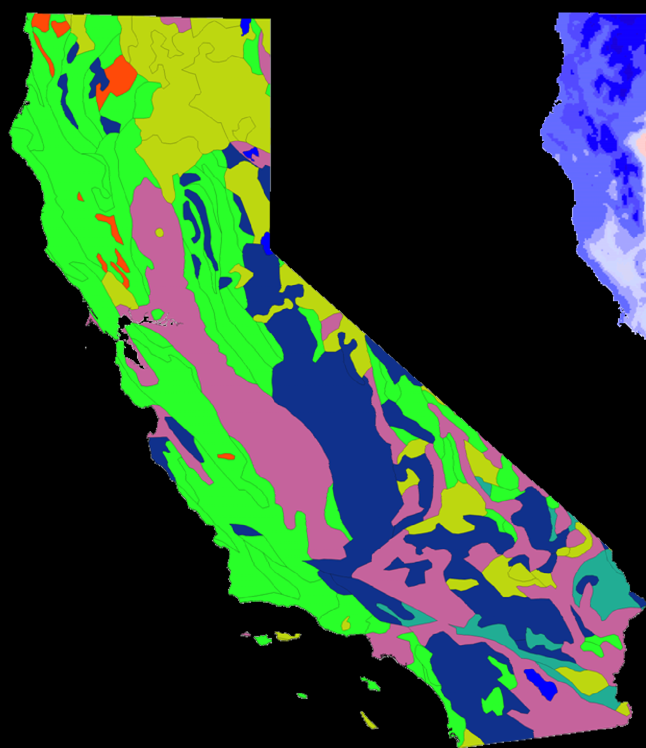
RCMP Philosophy and Objectives

- Use natural condition (or something close to it) as the desired state whenever possible
- Expectations must accommodate CA's diverse ecological and landuse settings, but retain **consistent** meaning throughout the state
- **Objective** = effective and well-supported scoring tools
  - Selection criteria need to support both precision, accuracy and sensitivity of our scoring tools (predicting expectation):
    - Natural state is our anchor, but…
    - Need adequate site density representing all major natural gradients
- The selection process should be transparent

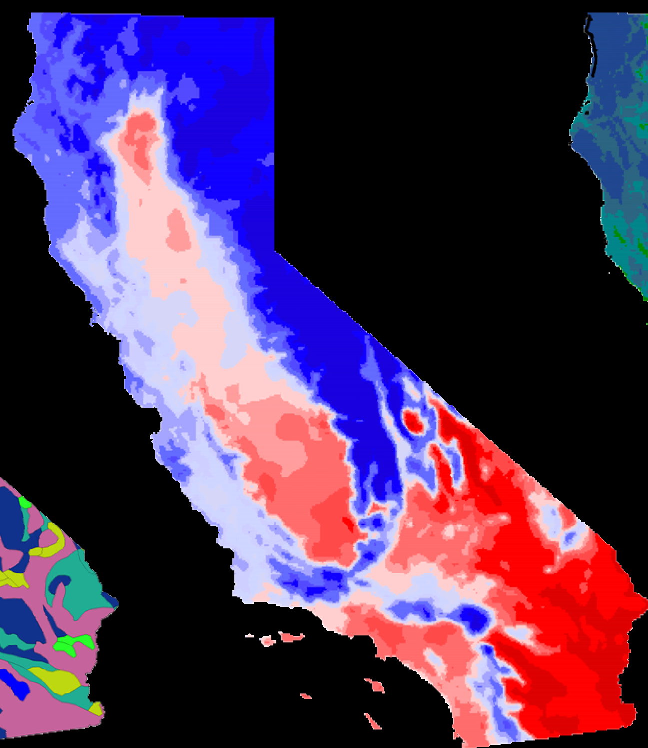
Technical Challenges: California is not Kansas

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

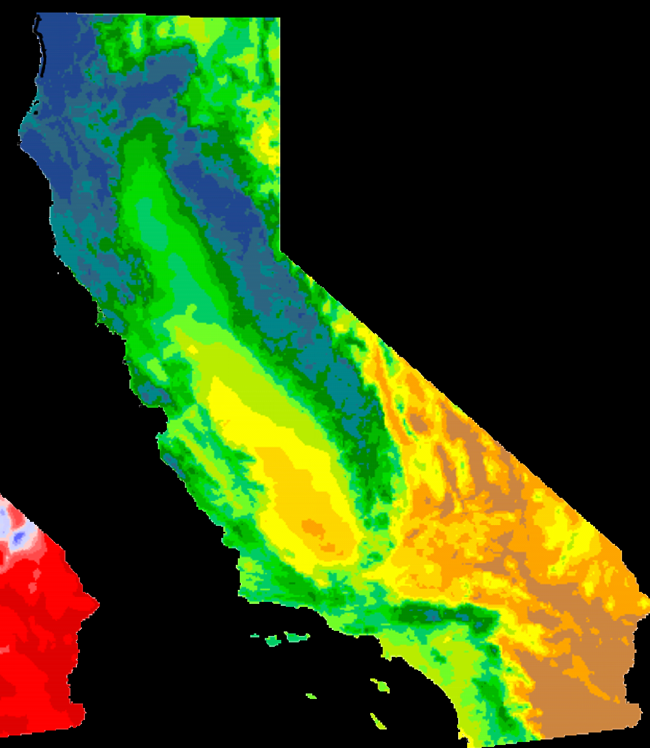
Geology



Temperature



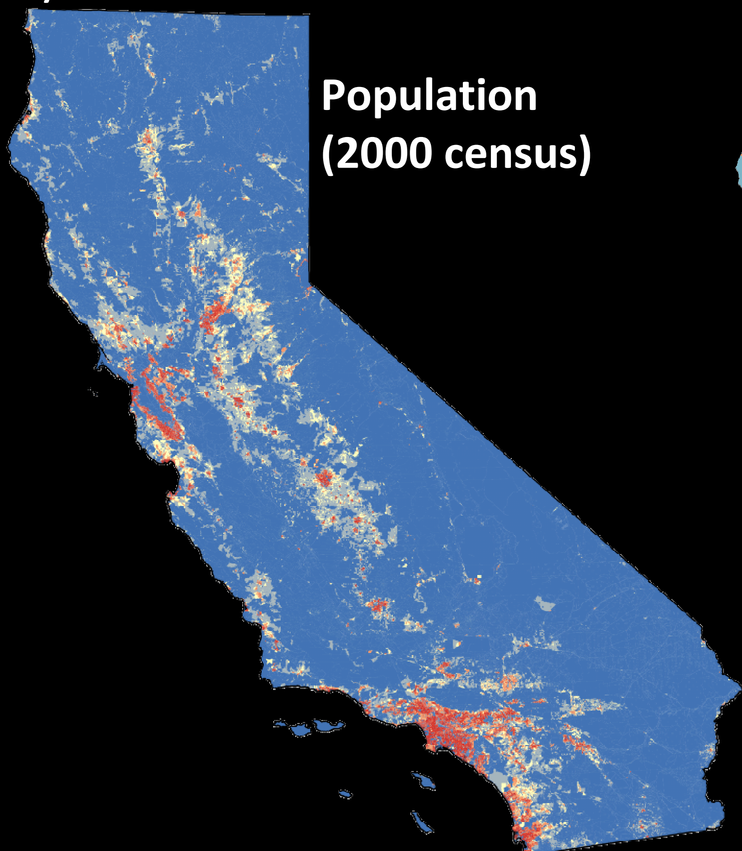
Precipitation



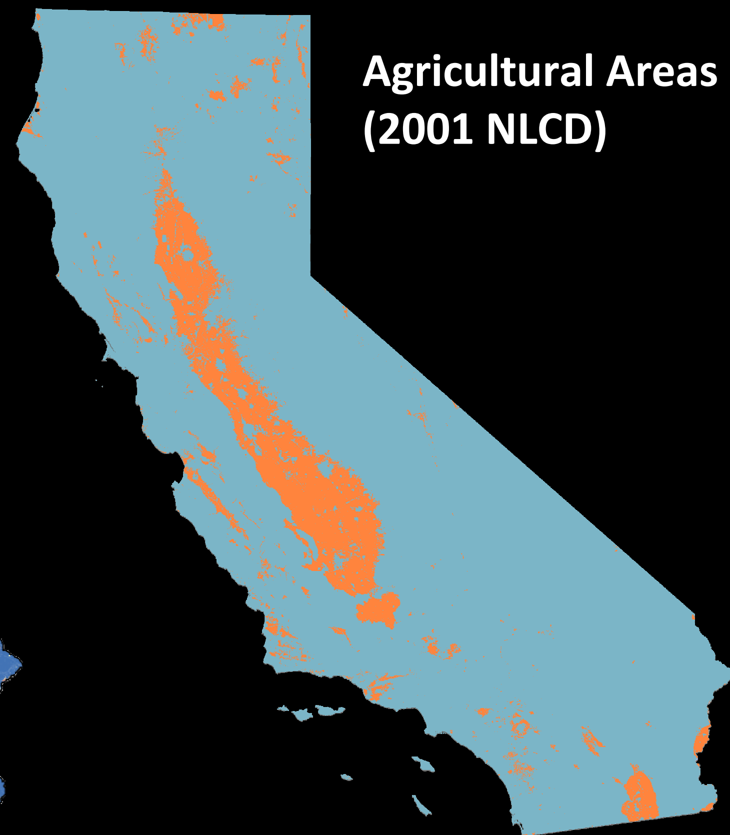
Management of biological variability requires good representation of biology at reference sites **across major gradients = need 100s of sites**

Technical Challenges: California IS Kansas

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



Population (2000 census)

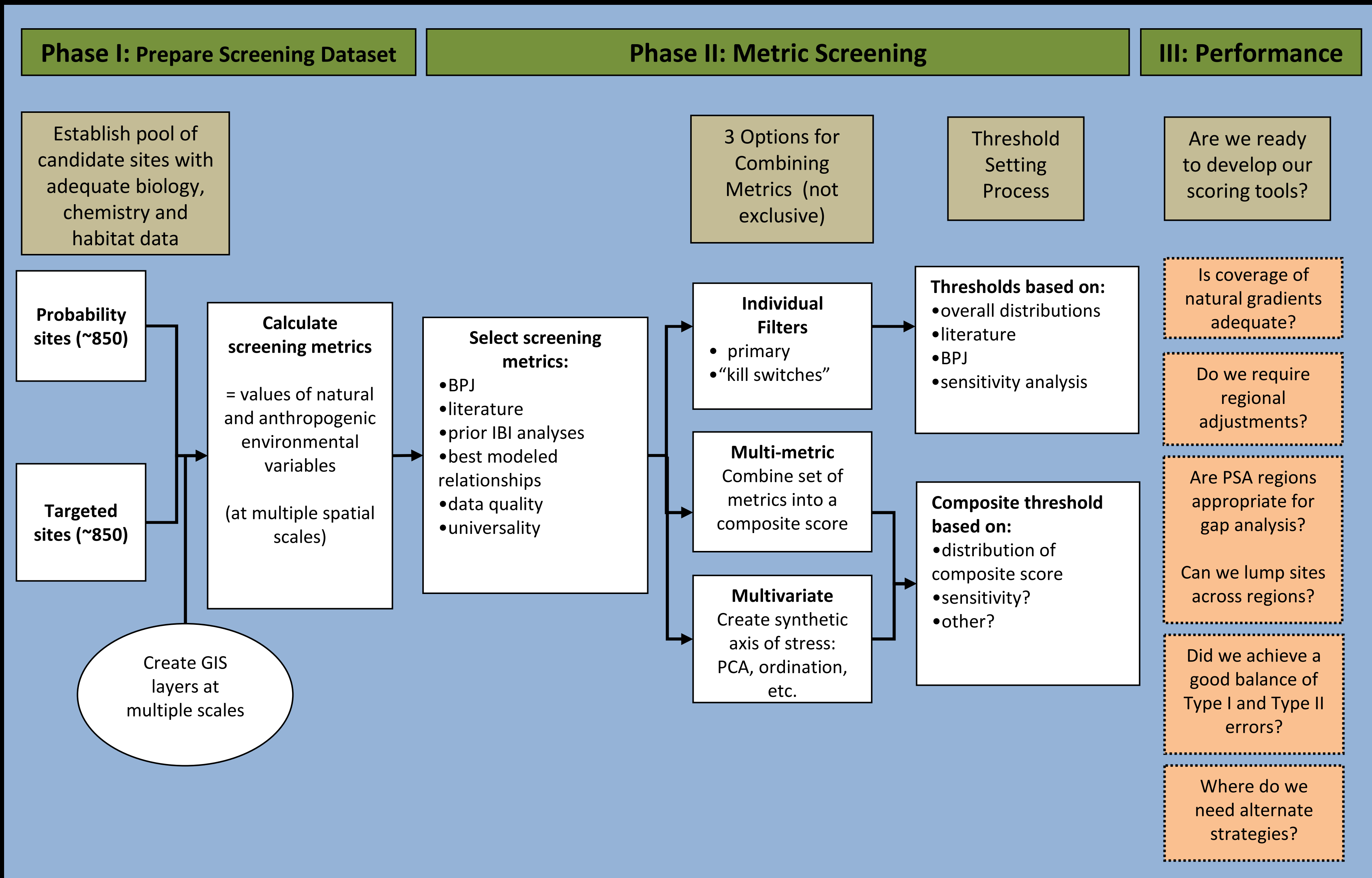


Agricultural Areas (2001 NLCD)

- Extensive human modification complicates the reference selection process because it introduces gaps in representation of natural gradients
- Intense development pressures make some regions unsuited for standard reference approaches

Morning is about STANDARD approaches:

- The standard reference techniques will apply to the vast majority of CA streams
- RCMP Panel recognized that standard process may not be able to work in some settings:
  - Extremely modified settings (e.g., Central Valley)
  - Exceptional classes of streams (e.g., concrete-lined channels)
- We'll deal with alternatives in the afternoon session


Phase I: Prepare Screening Dataset
Phase II: Metric Screening
III: Performance
Establish pool of candidate sites with adequate biology, chemistry and habitat data
Probability sites (~850)
Targeted sites (~850)
Create GIS layers at multiple scales
Calculate screening metrics
= values of natural and anthropogenic environmental variables
(at multiple spatial scales)
Select screening metrics:
•BPJ
•literature
•prior IBI analyses
•best modeled relationships
•data quality
•universality
3 Options for Combining Metrics (not exclusive)
Individual Filters
• primary
•"kill switches"
Multi-metric
Combine set of metrics into a composite score
Multivariate
Create synthetic axis of stress: PCA, ordination, etc.
Threshold Setting Process
Thresholds based on:
•overall distributions
•literature
•BPJ
•sensitivity analysis
Composite threshold based on:
•distribution of composite score
•sensitivity?
•other?
Are we ready to develop our scoring tools?
Is coverage of natural gradients adequate?
Do we require regional adjustments?
Are PSA regions appropriate for gap analysis?
Can we lump sites across regions?
Did we achieve a good balance of Type I and Type II errors?
Where do we need alternate strategies?

Development Process

Phase I: Prepare Screening Dataset

- Assemble a large screening set
- Calculate lots of metrics

Phase II: Develop Screening Process Options

- Select screening metrics
- Set metric thresholds
- Combine multiple metrics

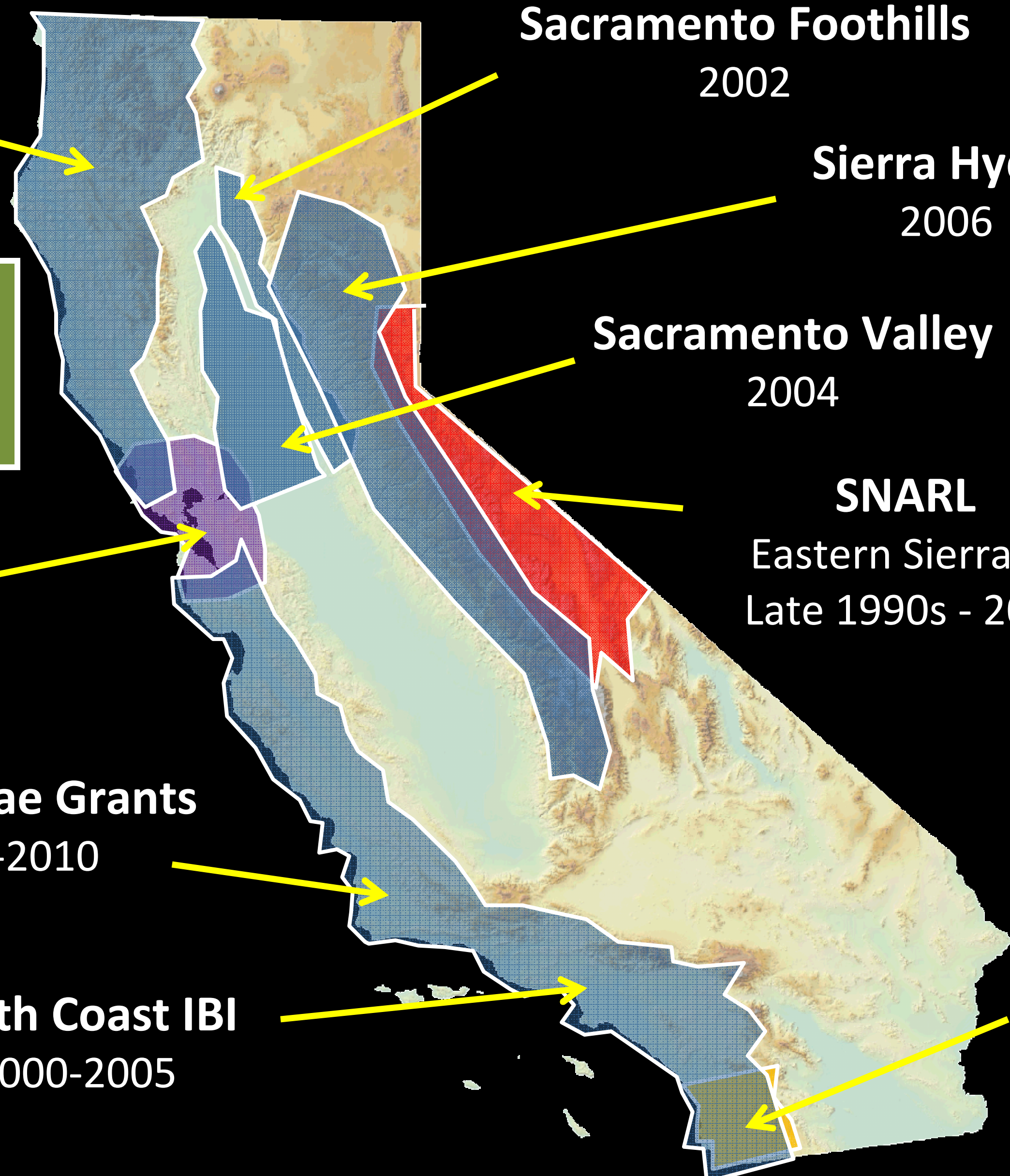
Phase III: Evaluate Performance and Resolve Issues

- Are gradients represented? Where are our gaps?
- Explore biological patterns at reference sites
- Where do we need alternate reference concepts?

Phase IV: Final Reference

- Finalize approaches, metrics and thresholds
- Identify reference sites for bio-objectives development
- Re-evaluate performance characteristics
- Validate with new sites

Significant CA Reference Projects (1997 – 2010)



- USFS (Utah State)
- EMAP
- SWAMP RCMP

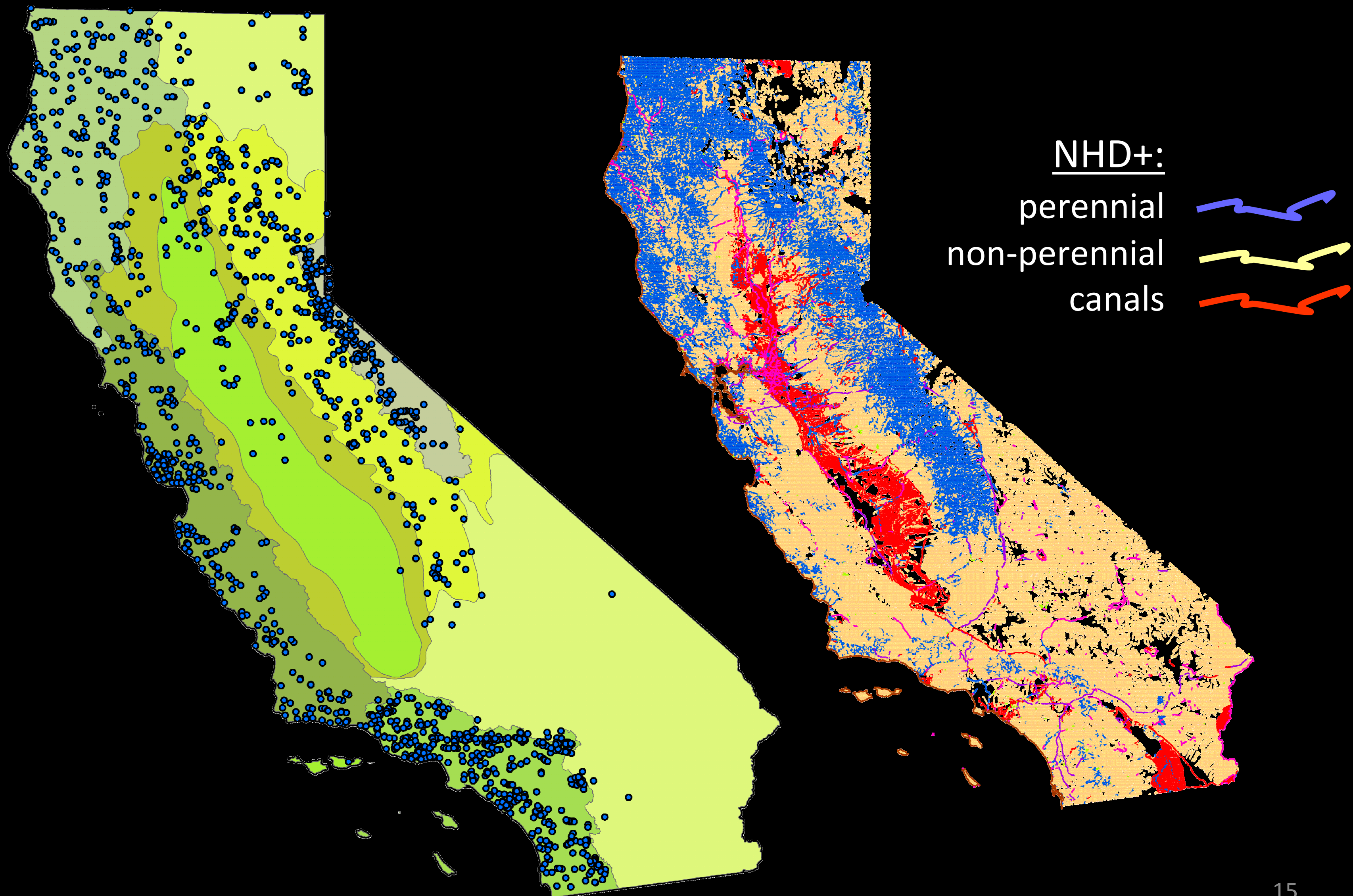
Phase I: Prepare Screening Dataset

(review from October 2010)

Many bioassessment datasets to choose from …
we mostly focused on ones with good local data

Program	Number of Sites	Geographic Distribution	Study Design	Indicators			
				BMI	PHab	Chemistry	Algae
EMAP	230	Statewide	Probabilistic	X	X	X	X
CMAP	200	Statewide	Probabilistic	X	X	X	X
PSA	200	Statewide	Probabilistic	X	X	X	X
USFS/ Utah State	200	Forest Lands	Targeted	X	X		
SMC	200	So Cal	Probabilistic	X	X	X	X
SNARL	100	East Sierra, Cent. Coast	Targeted	X	X	X	X
RWQCBs	>400	Many regions	Targeted and Probabilistic	X	X	X	

Step 1: Assemble Data > 1700 sites
(½ probability/ ½ reference candidates)



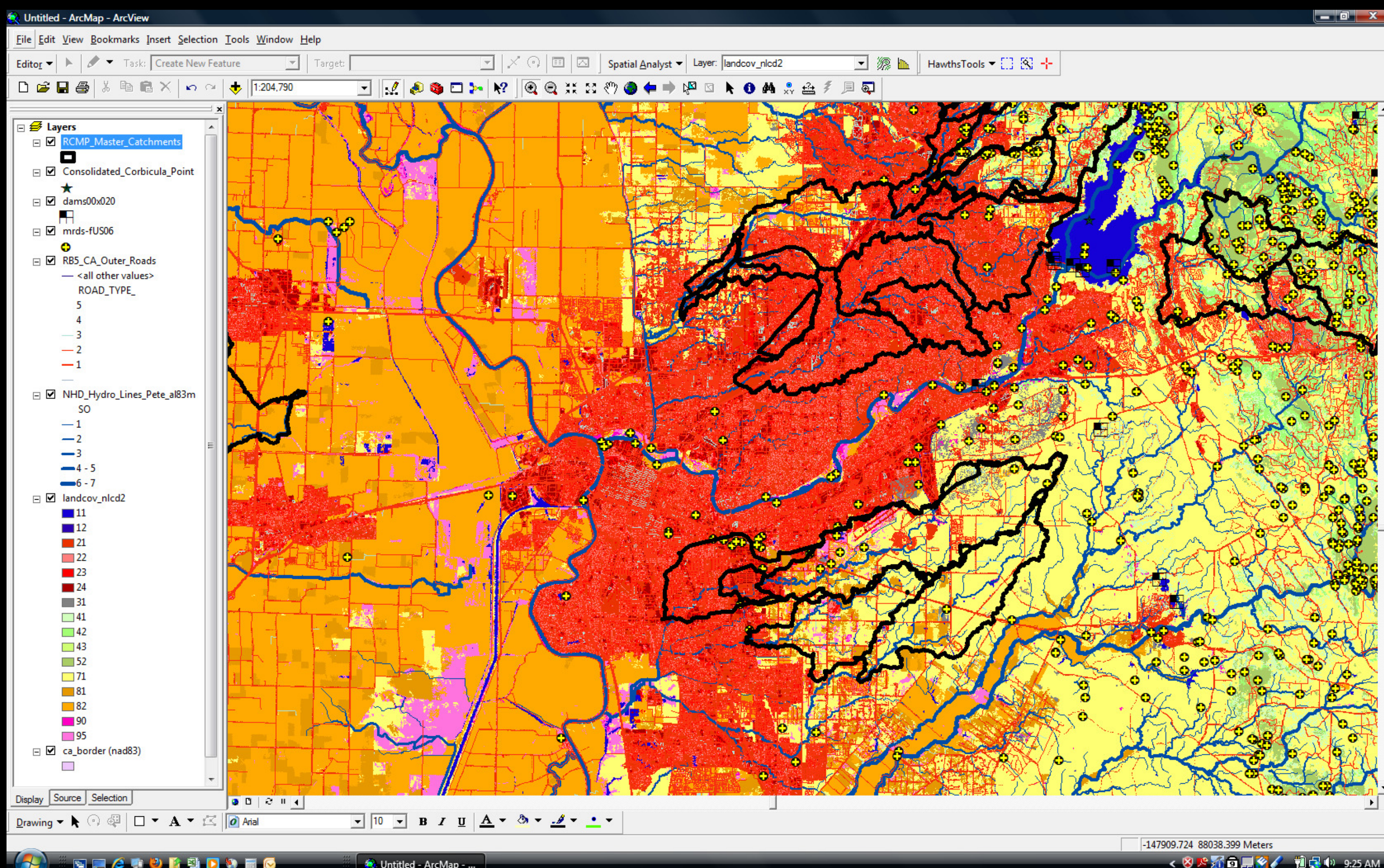
Step 2: Calculate metrics

Lots of GIS data

- Natural gradients
- Stressor gradients

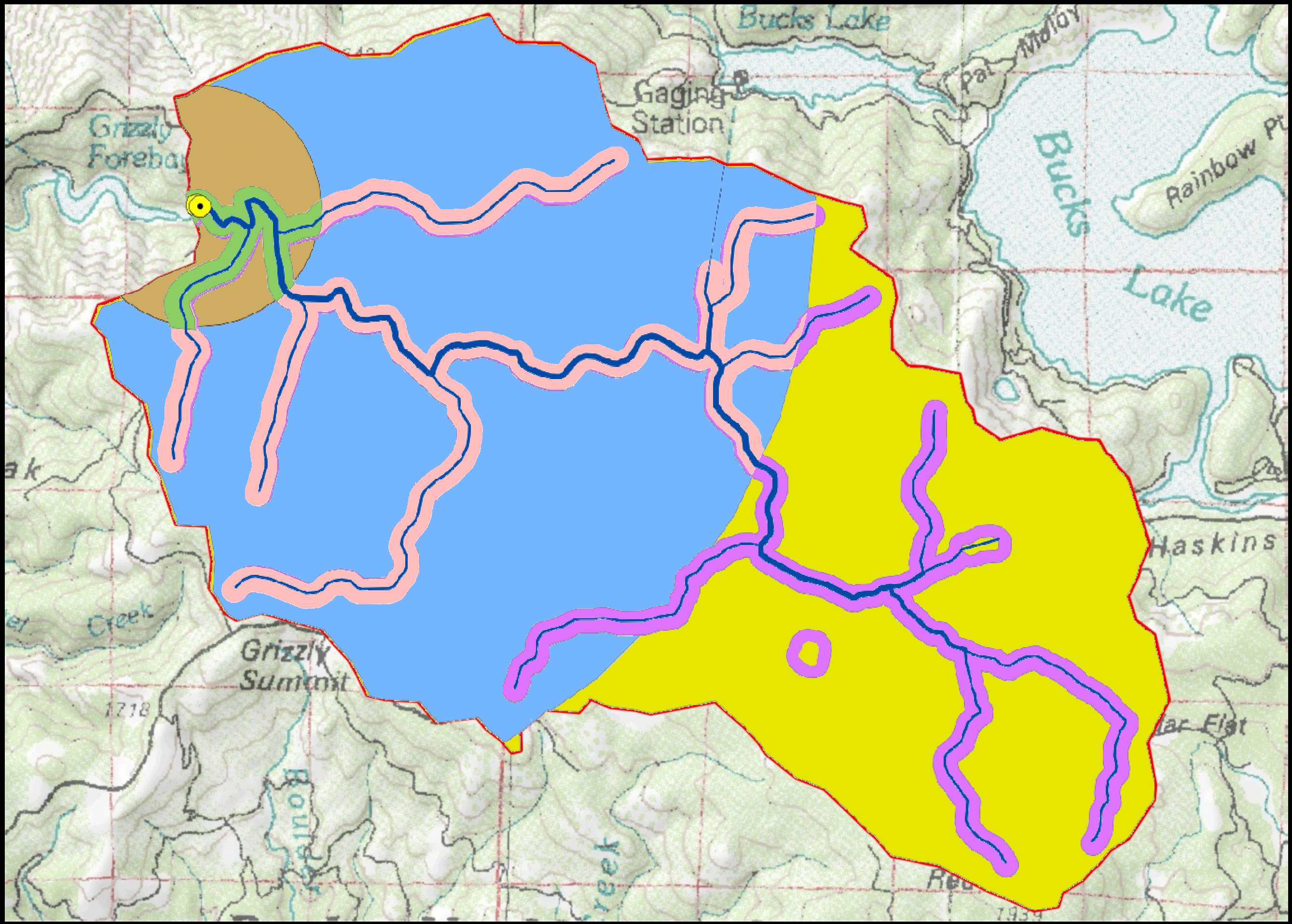
Local condition data

- Chemistry (nutrients, cond, pH, etc.)
- Physical habitat (instream and riparian condition)



Multiple Spatial Scales

Position of stressors in watershed influences their impact



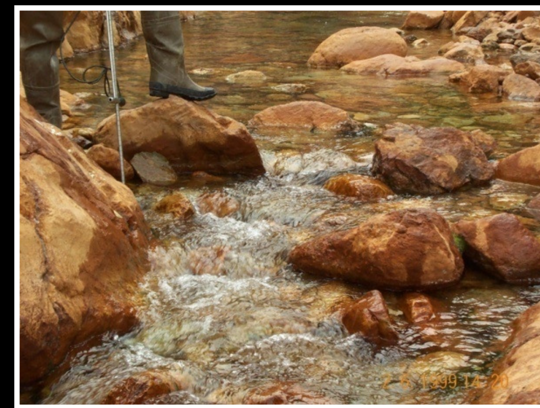
Metric Overview:
station data + natural gradients

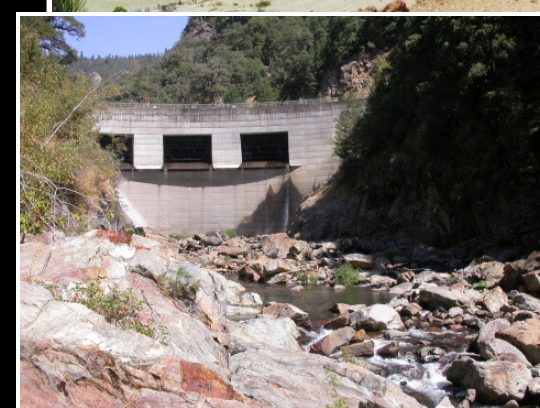
- **Station Data**
  - Regional board, PSA region, county, HUC, stream ID, ownership information
- **Natural Gradients**
  - **POINT DATA:** Coordinates, elevation, climate (PPT/T), ecoregion, stream order, stream volume, stream gradient
  - **BASIN DATA:** area, stream length, basin geology, mineral content

Metric Overview: stressors

(> 170 metrics)

- **Infrastructure**: roads, railroads
- **Population**
- **Hydromodification**
  - manmade channels, canals, pipelines
- **Landuse**
  - NLCD metrics, NLCD change (1992-2001), NLCD % Impervious
  - Timber Harvest, Grazing
- Fire history, dams, mines
- 303d list, NPDES/CWIQS discharges
- Invasive invertebrates, plants





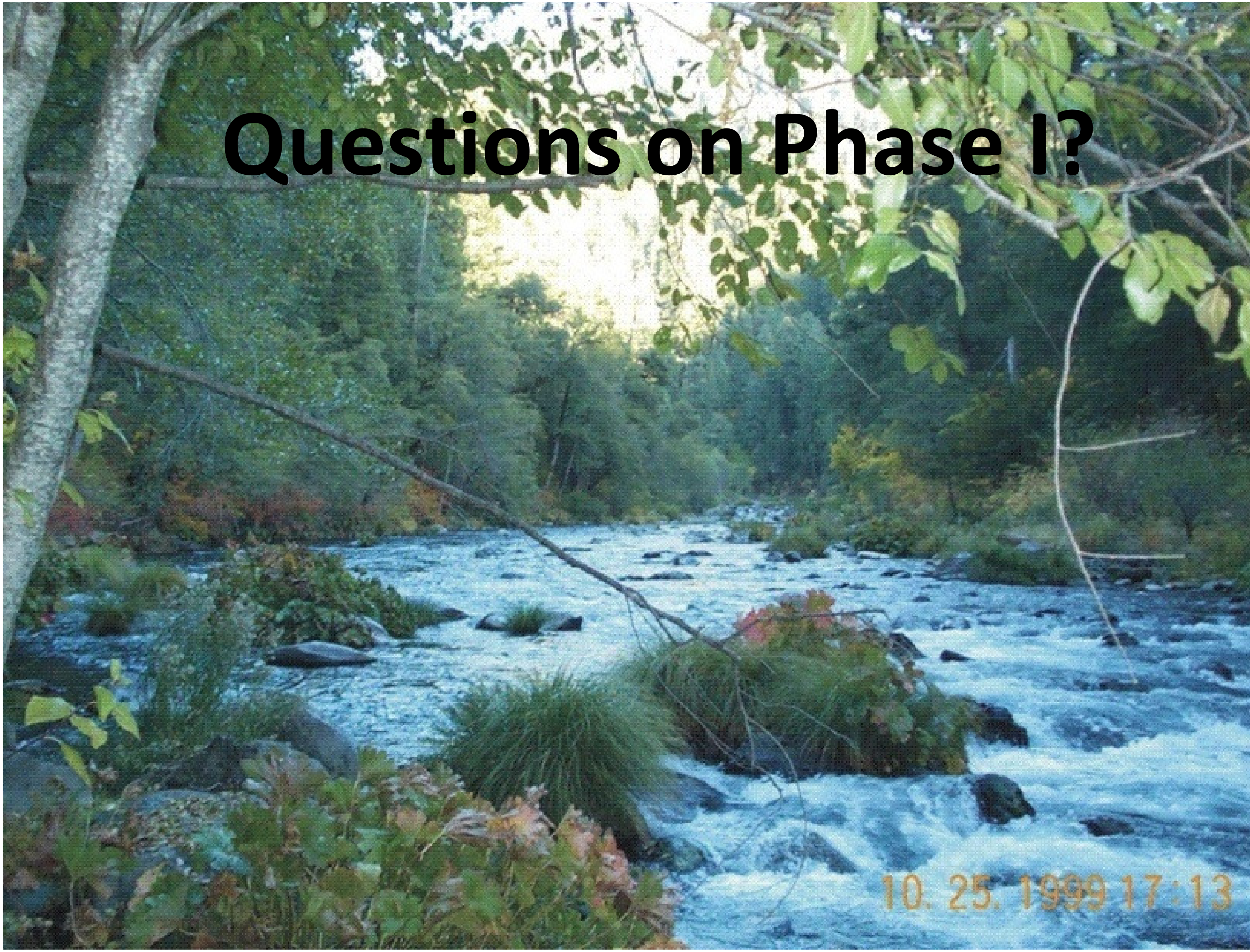
Metric Overview: local condition

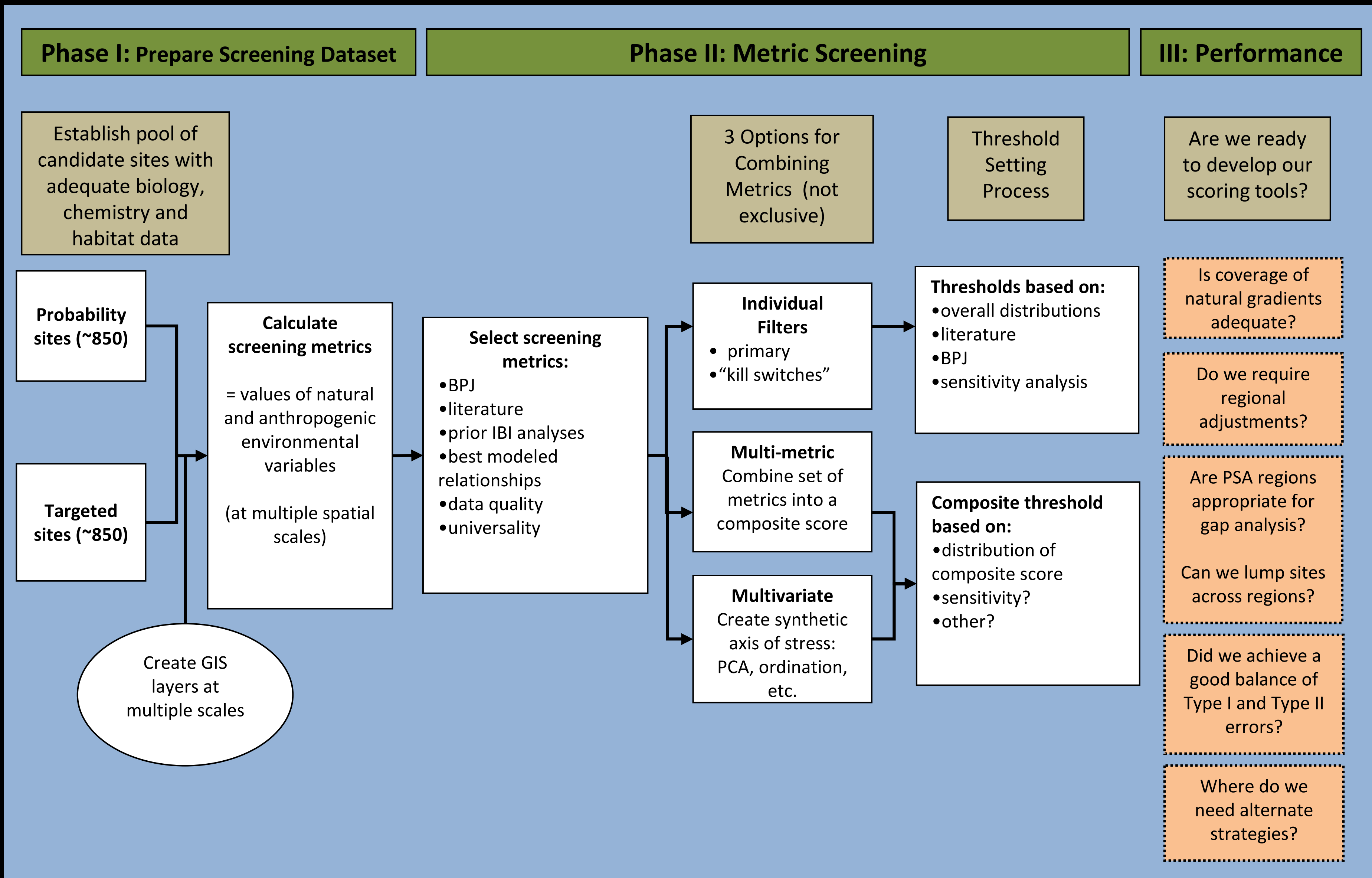
- **Chemistry:** nutrients, conductance, pH, Cl^-, turbidity

- **Habitat** (SWAMP metrics at many sites … similar to EMAP):

  – Riparian condition
  – Instream condition, fines
  – Human disturbance

Questions on Phase I?




Phase I: Prepare Screening Dataset
Phase II: Metric Screening
III: Performance
Establish pool of candidate sites with adequate biology, chemistry and habitat data
Probability sites (~850)
Targeted sites (~850)
Create GIS layers at multiple scales
Calculate screening metrics
= values of natural and anthropogenic environmental variables
(at multiple spatial scales)
Select screening metrics:
•BPJ
•literature
•prior IBI analyses
•best modeled relationships
•data quality
•universality
3 Options for Combining Metrics (not exclusive)
Individual Filters
• primary
•“kill switches”
Multi-metric
Combine set of metrics into a composite score
Multivariate
Create synthetic axis of stress: PCA, ordination, etc.
Threshold Setting Process
Thresholds based on:
•overall distributions
•literature
•BPJ
•sensitivity analysis
Composite threshold based on:
•distribution of composite score
•sensitivity?
•other?
Are we ready to develop our scoring tools?
Is coverage of natural gradients adequate?
Do we require regional adjustments?
Are PSA regions appropriate for gap analysis?
Can we lump sites across regions?
Did we achieve a good balance of Type I and Type II errors?
Where do we need alternate strategies?

Phase II: Combining Metrics

multiple strategies, not mutually exclusive

Filters: Site is removed from candidate pool if any metric exceeds a threshold value

- Primary filters: restrictive threshold used for highly predictive variables;
- "Kill-switches": threshold set as a "backstop"- sometimes less restrictive than primary filter; mostly used for variables with less confidence or weaker association with condition

Composite Approaches:

- Multi-metric: Metrics are combined into an index and sites are removed from pool if index value exceeds a threshold
- Multivariate: Metrics are combined into a multivariate axis of stress and sites are removed from the pool if the axis value exceeds threshold

PHASE II: development and testing

Use prior CA work as basis for creating working versions to explore options

- screening metrics
- combining metrics
- setting thresholds

“Platinum” version: little or no tolerance of activity

“Strawman” version: comparable to prior CA IBI development (or stricter)

Testing metrics

Mostly selected from prior work based on strong discrimination of biological impairment; represent major categories of anthropogenic impacts; a mixture of both landscape and reach-scale variables

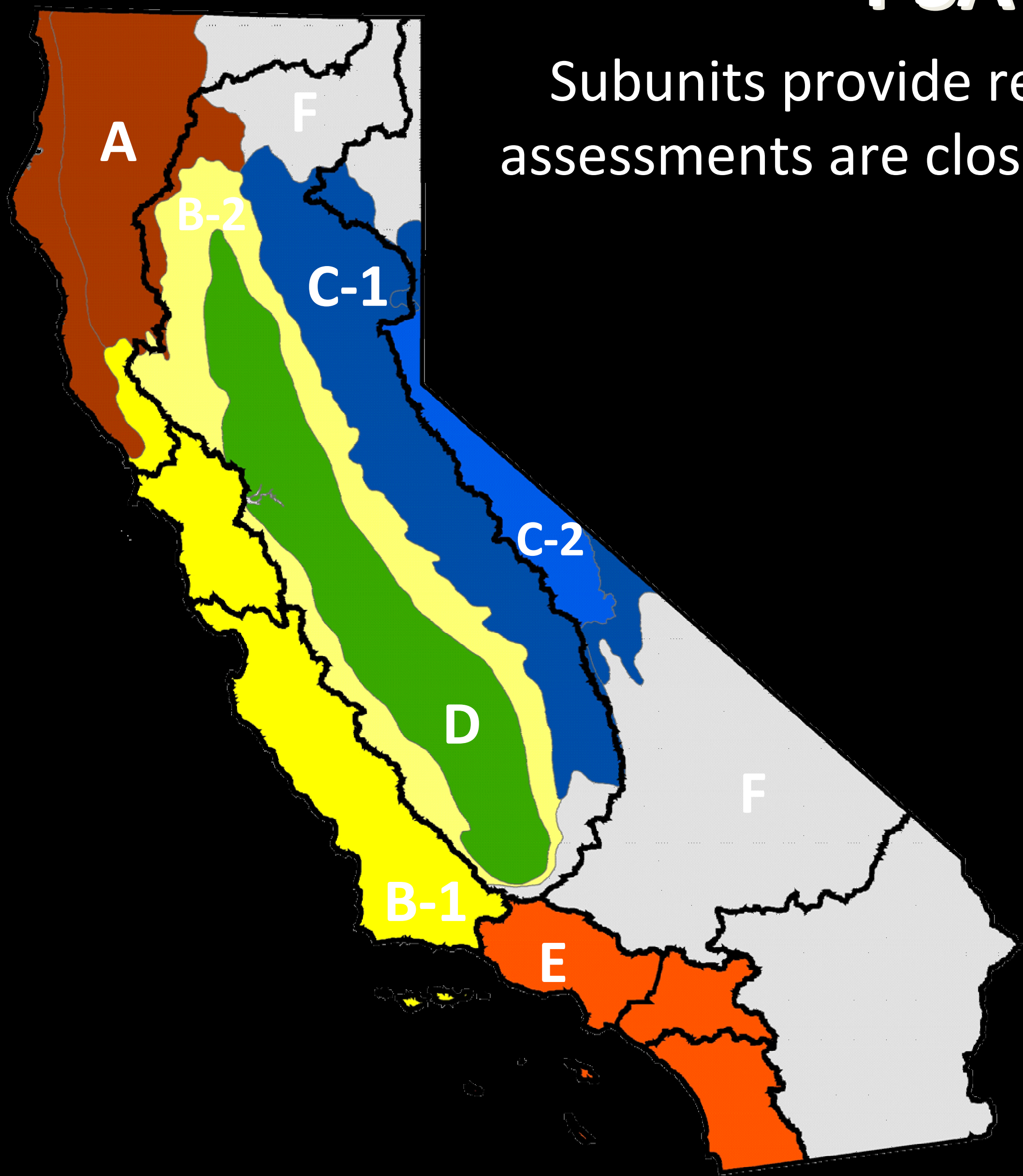
- **Landscape**
  - % Ag, % Urban,
  - Code 21 (recreational veg, roadside veg, rural/exurban development)
  - Road density
- **Upstream modifications**
  - Dams, artificial channels
  - Mines
- **Local disturbance**
  - W1_Hall (EMAP-derived reach-scale index of human disturbance)
  - Water chemistry: Total N, Total P

Thresholds comparable to (or stricter than) other CA values

Metric	"Platinum" (1k, 5k)	Strawman (1k,5k, ws)	South Coast IBI (5k,ws)	North Coast IBI (1k, ws)	Western Sierra IBI	Eastern Sierra IBI
Local Disturbance (W1_Hall)	0	0.5	-	-	-	-
%Ag	0	3, 10	5,5	5,5	5,5	-
%Urb	0	3, 10	3,3	3,3	3,3	-
Watershed Development (%Code 21)	0	10	included in urban	included in urban	included in urban	-
Road Dens (km/km²)	0	1.5/ 3.0	2.0	1.5/ 2.0	2.0	-
Road x-ings (#/km)	-	-	-	-	-	0.2
Pop Density (#/km²)	-	-	150	25/ 50	-	-
TN, TP	1.5/ 0.1	3/ 0.5	-	-	-	-
Dam Storage (MG)	0	10^4	-	-	-	-
Mines	-	0.1	-	-	-	-
% canals&pipes	-	10	-	-	-	-

PSA Regions

Subunits provide regional perspective ….and assessments are closer to the scale of regulation



A= North Coast

B = Oak Chaparral

- **1= Coastal Chaparral**
- **2= Interior Chaparral**

C = Sierra

- **1= Main Sierra**
- **2= Central Lahontan**

D = Central Valley

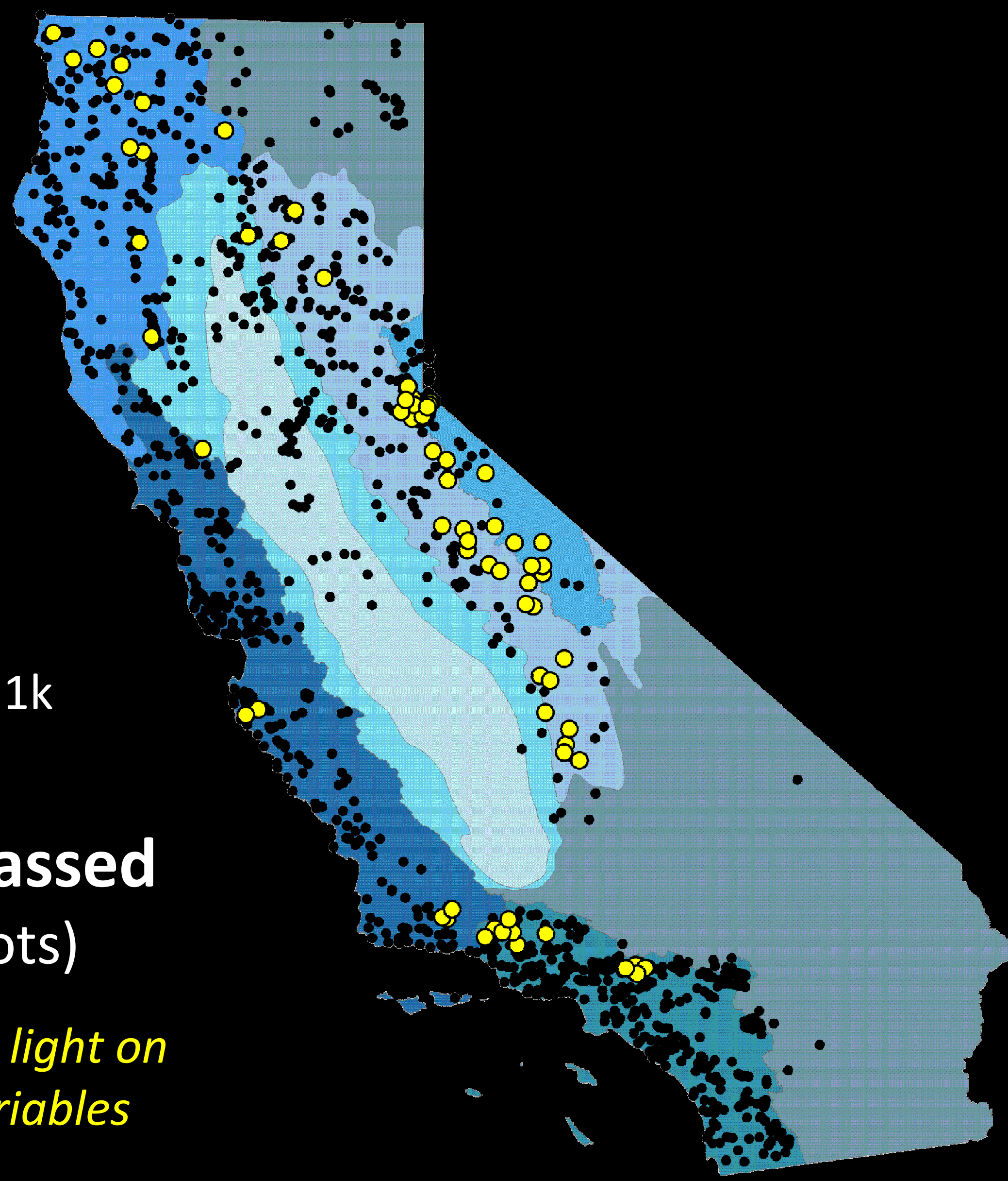
E = South Coast (SMC)

F = Other:

- **Modoc Plateau**
- **Deserts**

Perfect World
"Platinum" Screens

- No tolerance for impacts
- Little wiggle room for natural variability, data errors:
  - TN < 1000 µg/L
  - TP < 50 µg/L
  - W1_Hall = 0
  - Ag, Urban, Code 21 = 0 at 5k, 1k
  - Dams =0



109 sites passed
(**yellow** dots)

-note that we're light on reach scale variables

“Strawman” screens

Relaxed thresholds, close to IBI development

Assumptions:

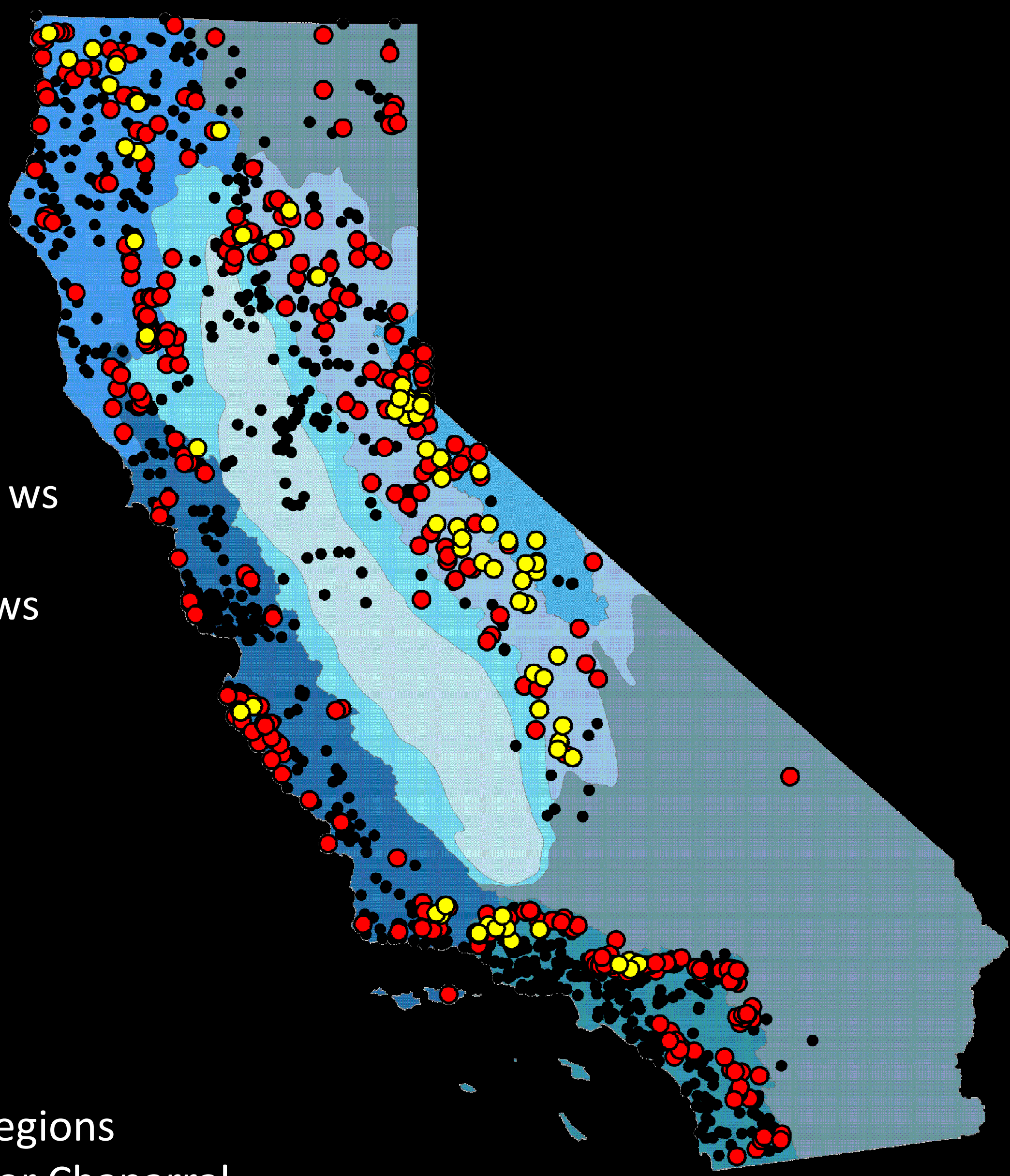
– The relaxation accommodates error in data sources, watershed delineation, but still discriminates true (low stress) reference sites.

– This relaxation has minimal effect on biological response *(tested later)*

Strawman Screens

Thresholds similar or stricter than CA IBI development

- Nutrients: TN < 3000 ug/L or null
  TP < 500 ug/L or null
- W1_Hall < 1.0 or null
- Ag, Urban < 3 at 5k, 1k, <10 at ws
- Code 21 < 10 at 5k, 1k, ws
- Road density < 1.5 at 5k, 1k, 3.0 ws
- Dam storage < 10,000 Mgal
- Riparian mine density < 0.10/km
- % modified channels < 10



586 sites passed (yellow+ red)

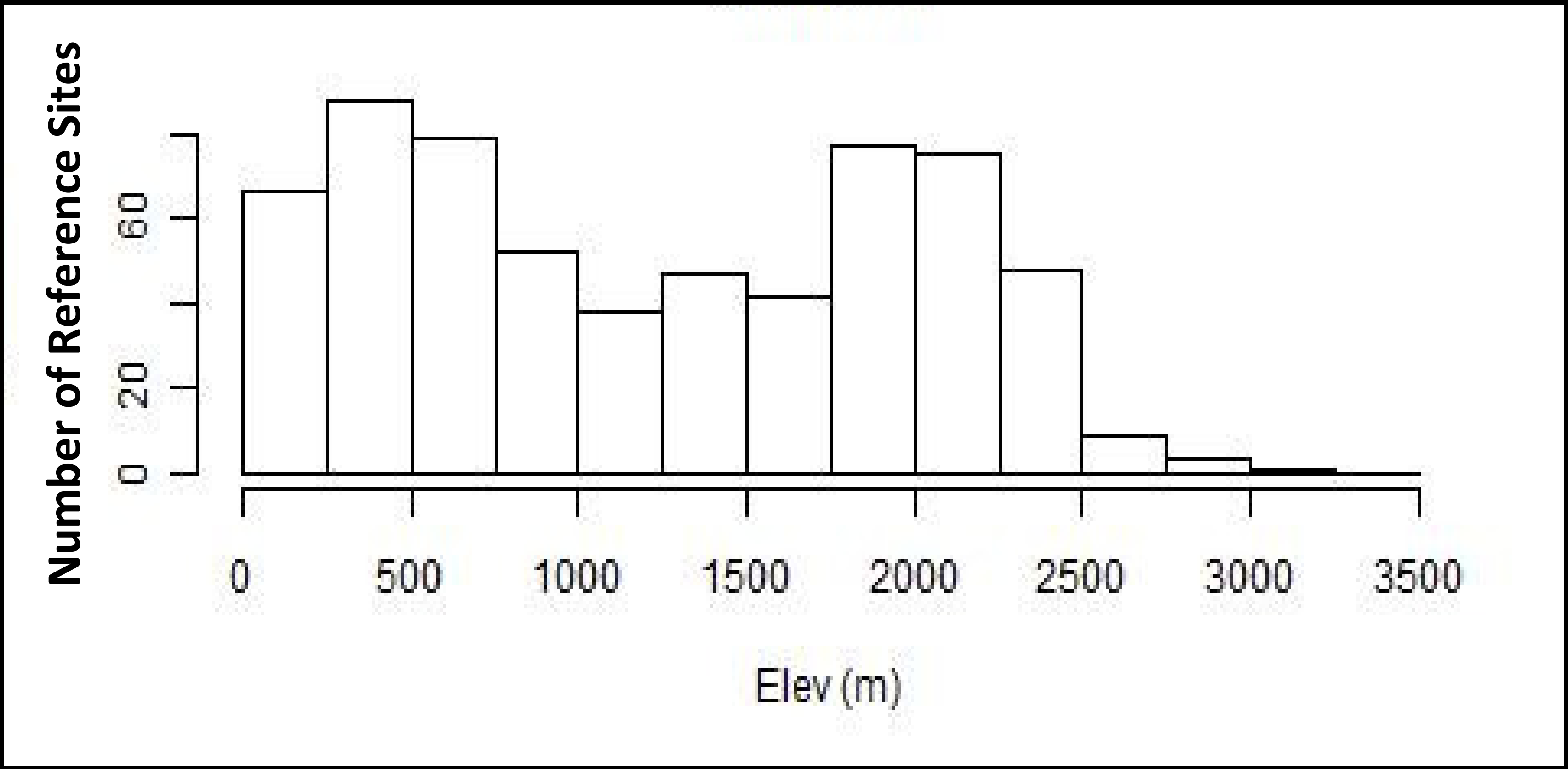
- Good coverage in mountains.
- OK coverage in some chaparral regions
- Few sites in Central Valley, Interior Chaparral

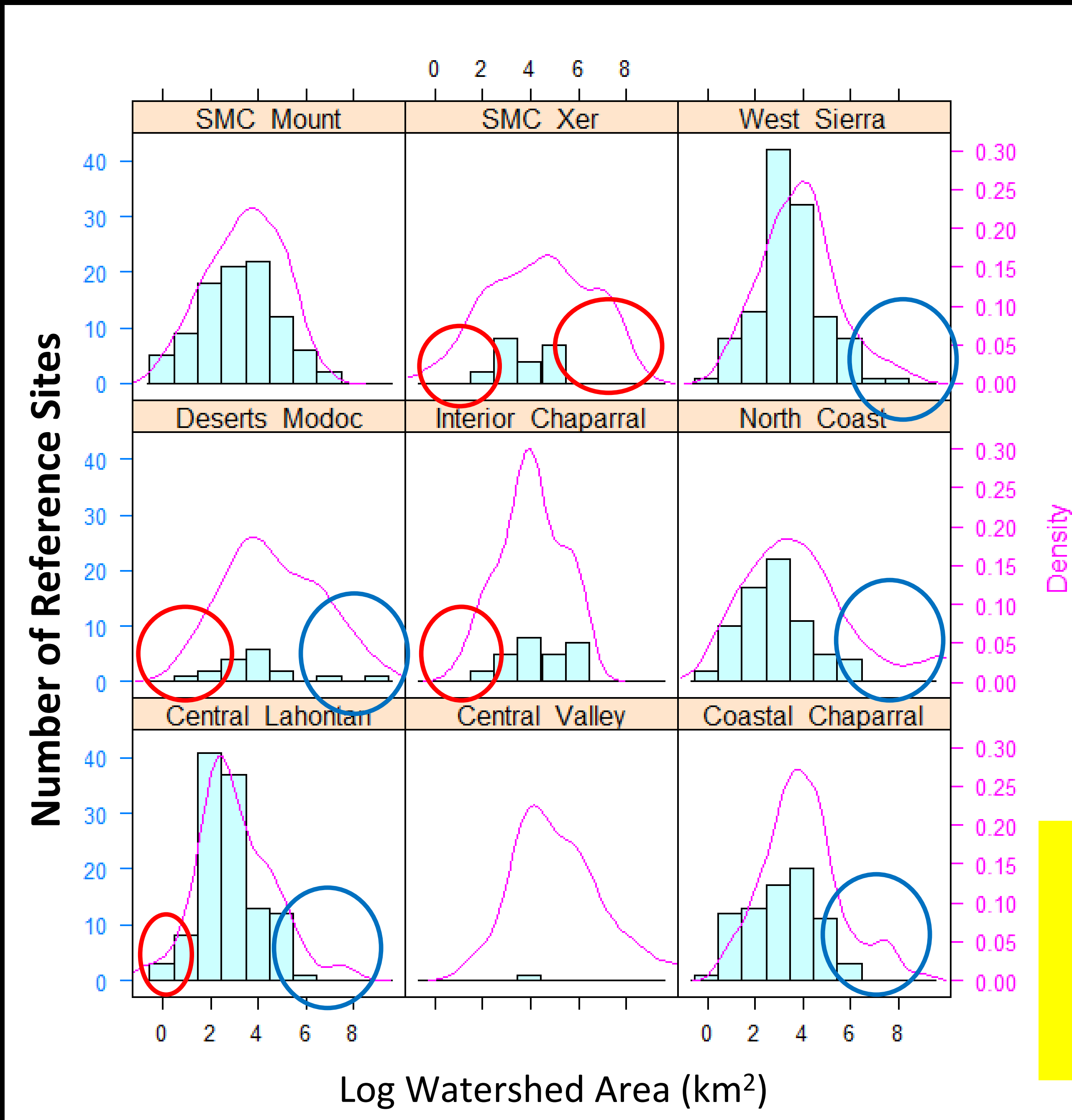
Number of Potential Reference Sites

REGION	Platinum	Strawman
Lahontan	40	119
Central Valley	0	2
Coastal Chaparral	8	80
Interior Chaparral	1	35
Northern Calif	14	73
Southern Calif	11	133
Western Sierra	35	124
Deserts + Modoc	0	20
TOTAL	**109**	**586**

Gradient Analysis:

Do we have adequate coverage of natural gradients?





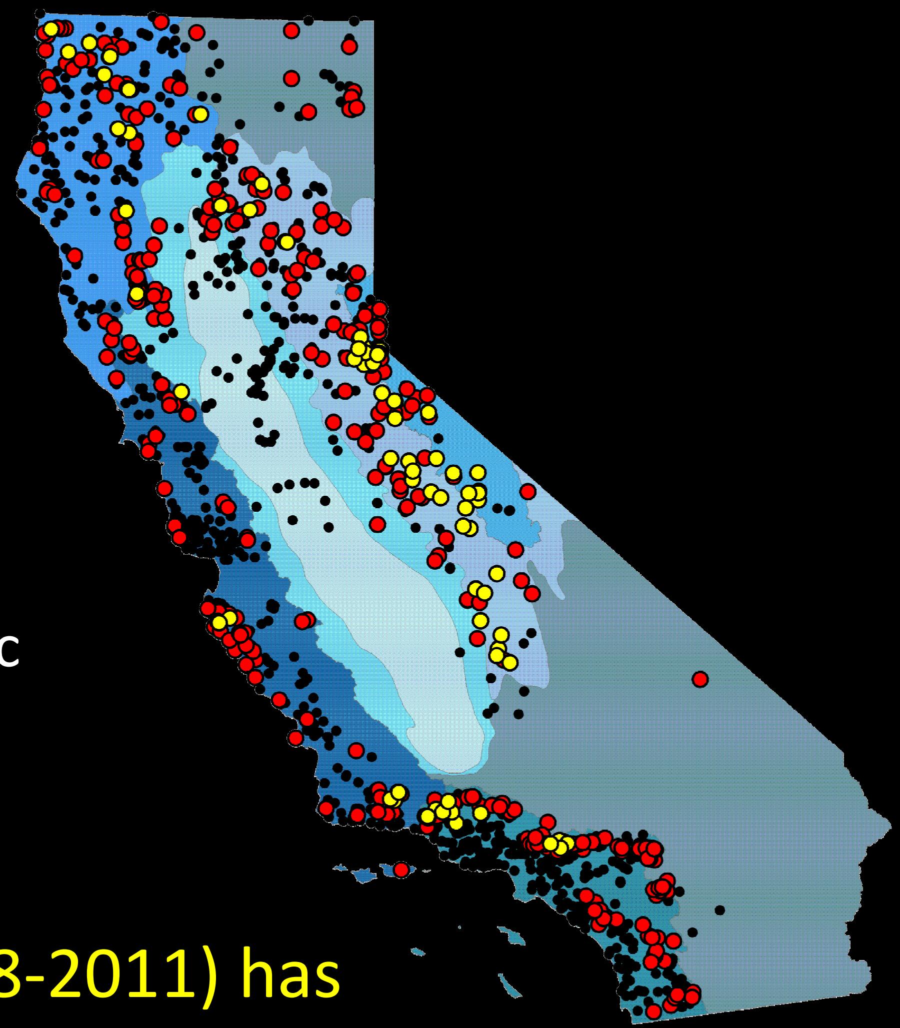
Coverage Gaps

Mostly Non-wadeable

Elevation coverage is strong; low numbers overall in Central Valley, Desert-Modoc, SMC Xeric and Interior Chaparral

REGION	Elevation	Watershed Area	% CaO geology	% N geology	PPT	SedGeo	Slope	Stream Order
North Coast	Good	Few large watersheds (> 10^4 km²)	Good	Good	OK, but few on dry end	Good	Good	Good, but none > 4th
Coastal Chaparral	Good	Good up to 10^6 km²	Good	Good	Good	Good	Good	Good
Interior Chaparral	Good	Low numbers overall (no sites <100 km²)	Good	Good	OK, but few on dry end	Good	Good	Good
Deserts + Modoc	Good	Need more small (< 100 km) and more large watersheds	Good	Good	Ok, but few in wet end	Good	Good	Good
SMC-Mtn	Good	Good	Good	Good	Good	Good	Good	Good
SMC-Xer	Good	Need more small (< 100 km) and more large watersheds	Good	Good	OK, but few on dry end	Good	?	Very few high order streams
West Sierra	OK, but few in low elevation	Good	Good	Good	Good, but light on wet/dry ends	Good	?	OK, but few high order
Central Lahontan	Good	Ok, but few small sheds (<100 km²)	Good	Good	?	Good	Good	Good
Statewide	Good	Good up to 10^6 km², but need more small	Good	Good	OK, but few on dry end	Good – most are 0% or 100%	Good	Good, but few sites > 4th +

Gradient Representation Summary:



Current "strawman" screens produce good coverage with only a few problem areas:

- Central Valley has only 2 sites
- Interior Chaparral and SoCal Xeric have low numbers
- Most gradients are well-represented

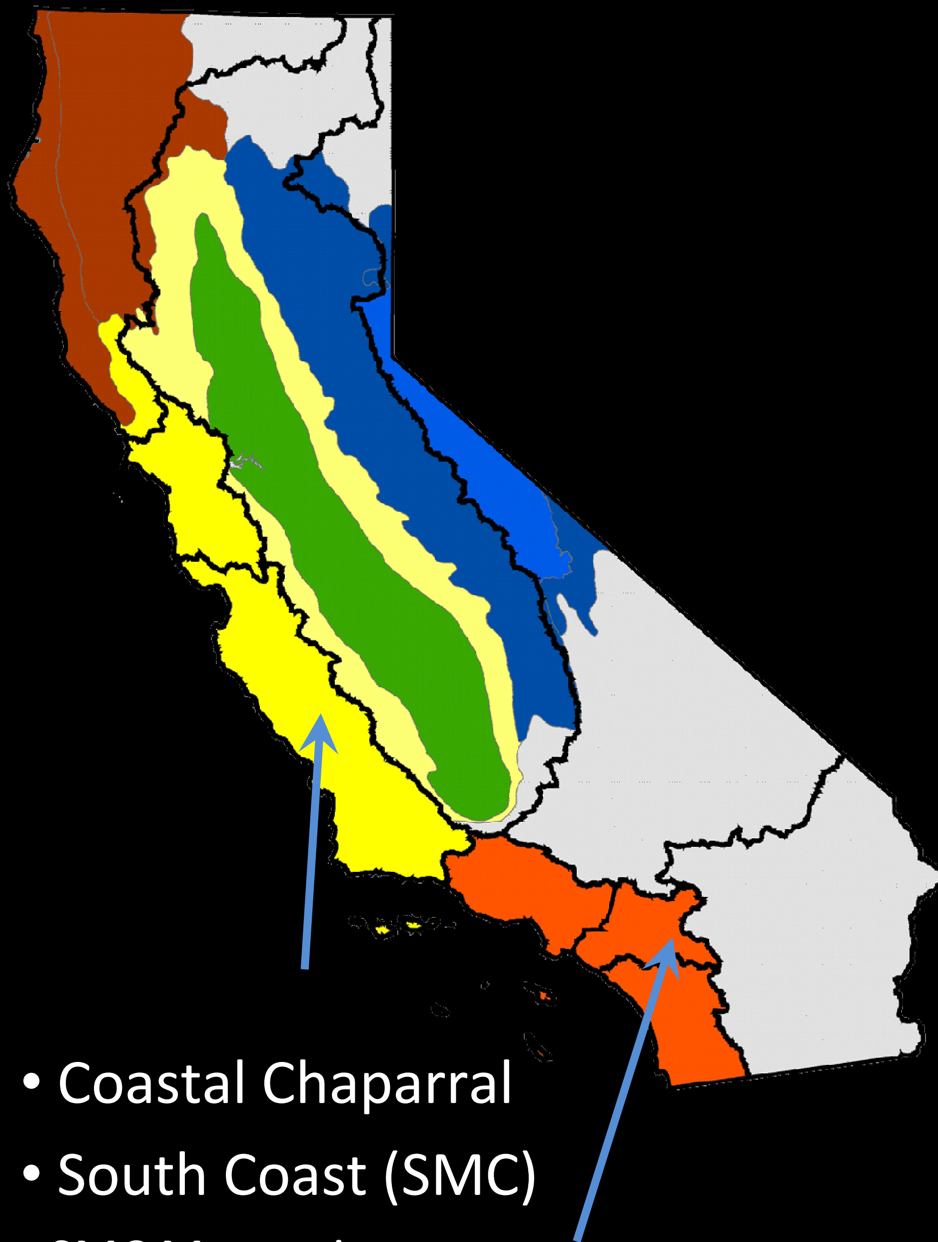
SWAMP RCMP sampling (2008-2011) has been concentrated in low density regions

Pilot Studies

At last meeting, panel recommended that we select one or more regions to test out our development process in order to identify and work through potential implementation issues

Our stakeholders helped us pick two:

- Good range of stressors
- Good data density



- Coastal Chaparral
- South Coast (SMC)

- SMC Mountains
- SMC Xeric

A note about Type I and Type II error

(risks of keeping stressed sites in the reference pool vs. rejecting low stress sites, respectively)

In a perfect world with a large number of undisturbed streams of all types, we could focus exclusively on Type I error.

However, very restrictive criteria result in under-representation of important natural gradients (e.g., platinum). Thus, Type II error (excessive rejection of sites) reduces the performance (i.e., accuracy and precision) of our scoring tools for regulation

Our screening process must balance Type I and Type II error … we want the most restrictive screens that allow us to represent major gradients

- *Vary thresholds with spatial scale*
- *Include "kill switches"*
- *Composite approaches*

Composite Approaches:
Distillation of many stressors into an overall estimate of stress

Can help compensate for Type II error: imperfection in stressor data can lead to inappropriate rejection of sites by single filters

- Accuracy/completeness limitations of GIS layers
- Artifacts of odd spatial configuration of stressors (see road and Code 21 examples)

Probably too lenient by itself for rare, yet important, stressors *(e.g., instream gravel mines, invasive species, dams, etc.)*

Composite Approaches

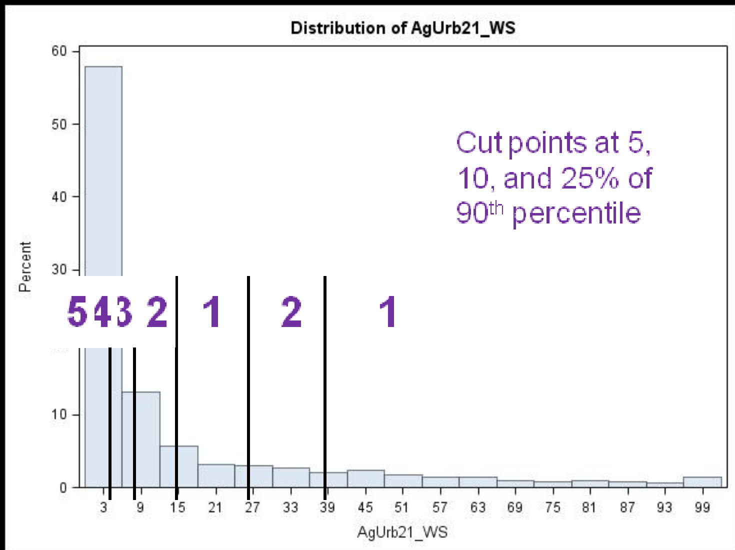
Multi-metric Approach **Analogous to Index of Biotic Integrity (IBI)**

Metrics are combined into an index and sites are removed from pool if index value exceeds a threshold

Multivariate Approach **Analogous to SoCal Marine Benthic Response Index (BRI)**

Metrics are combined into a multivariate axis of stress and sites are removed from the pool if the axis value exceeds a threshold

Multi-Metric:
Metric & Threshold Selection



Index Score	5	4	3	2	1

Percentile Metric (scale)	0	5%	10%	25%
AgUrb21_WS	0.0	2.4	4.8	12.0
ArtificialPaths100kPer_5K	0.0	0.3	0.5	1.3
ArtificialPaths24kPer_5K	0.0	0.8	1.6	4.1
CanalPipe100kPer_5K	0.0	0.4	0.8	2.0
GravelMinesDens_WS	0.0	0.0	0.0	0.0
HousingDens2000_WS	0.0	14.0	28.0	70.0
IMPERVMEAN_WS	0.0	0.6	1.3	3.1
MinesDens_WS	0.0	0.0	0.0	0.1
RDDENSC12_WS	0.0	0.2	0.5	1.2
RDDENSC3_1K	0.0	0.1	0.2	0.5

Metric Selection (2 options)

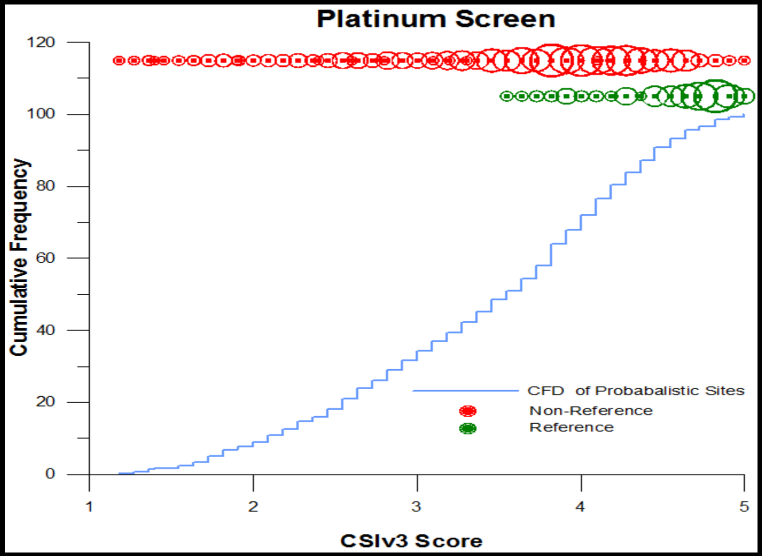
- Either pick *a priori* (e.g., use filter metrics)
- Select based on an objective process (e.g., best stressor-response models)

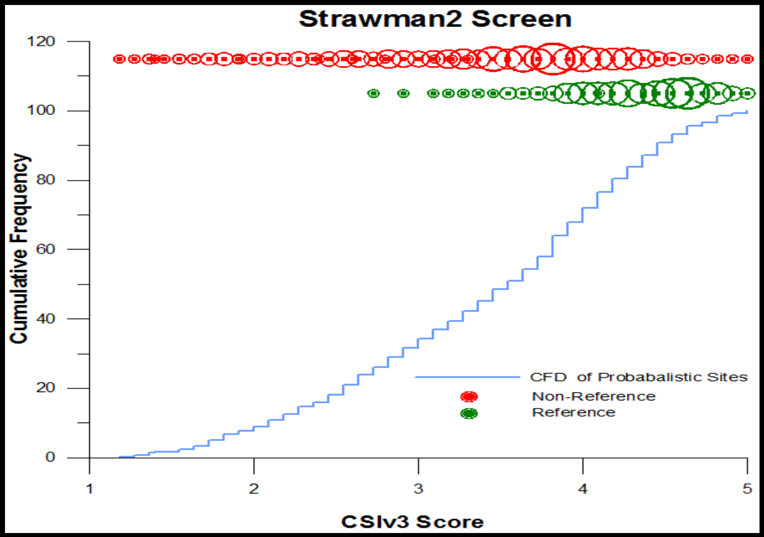
Metric Scoring

– Establish distribution of metric values at probability sites throughout state

– Set scoring cutpoints at percentages of the 90th percentile of each metric's distribution: 0%, 5%, 10% and 25%

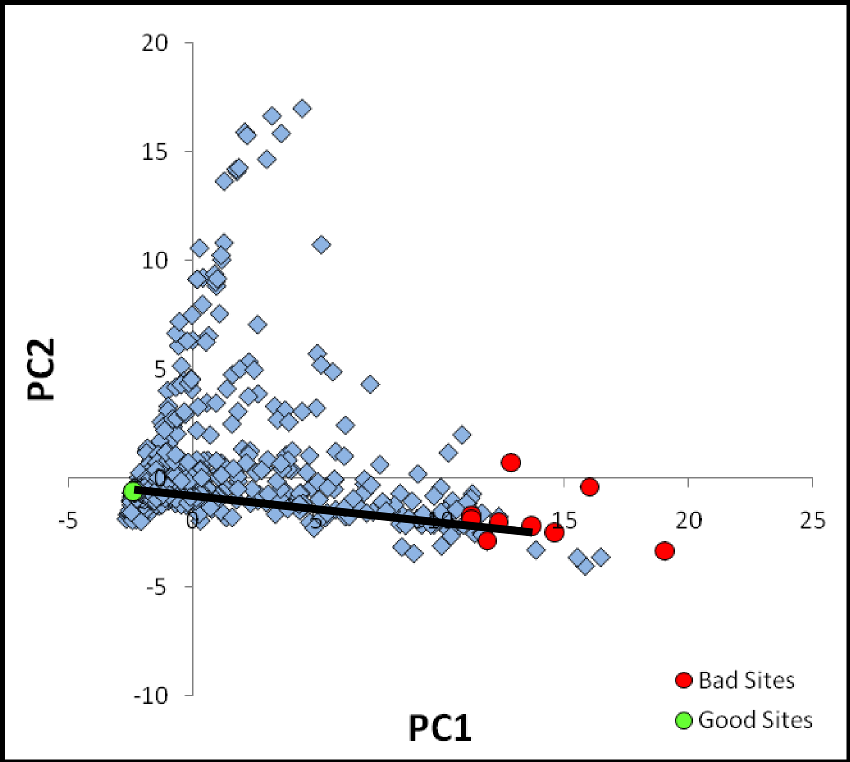
Multi-Metric Stressor Index





- Scores all sites from 1-5
  - Low to high quality
- Evaluated all sites and compared vs. designations from platinum and strawman versions
- Need to create index thresholds for reference designation

Multivariate Ordination



- Designate good and bad sites
  - Based on absence or >95th percentile of stressors
- Use PCA to ordinate all probabilistic sites
- Orient a Good-Bad Axis in PCA space
- Calculate weights of each stressor along the Good-Bad axis
- Identify a screening threshold along PCA axis

Applying Multi-metric/ Multivariate Results

- Axis weights may be a good way to identify good metrics to use for filters (*see also Stressor Modeling*)
- Could help reduce Type II error rates
- Good candidates for combined approach:
  – RCMP panel suggested a MMI + kill switch approach

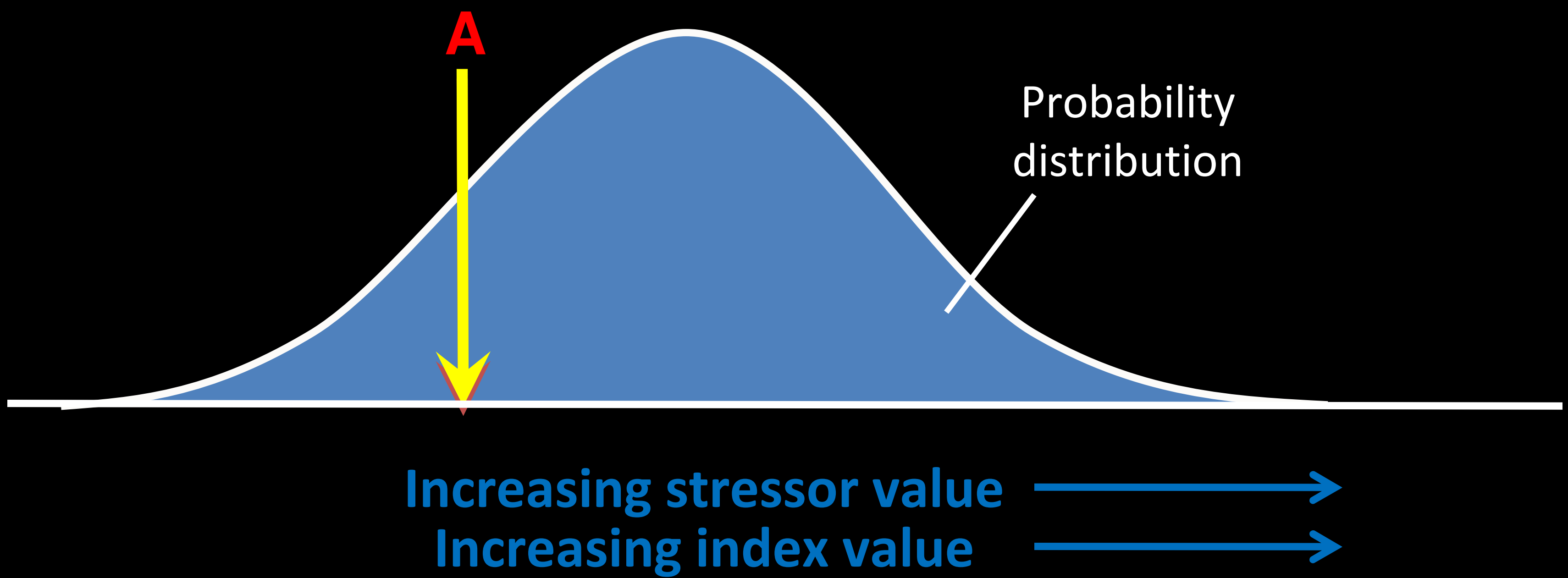
Phase II: Information Used to Inform Threshold Setting Process

(most apply to both filter and composite approaches)

- Distribution-based (statewide or regional)
- Threshold-sensitivity analysis
- Effects-based:
  - Literature values for biological effects
  - Evaluate biological responsiveness directly

Distribution-based

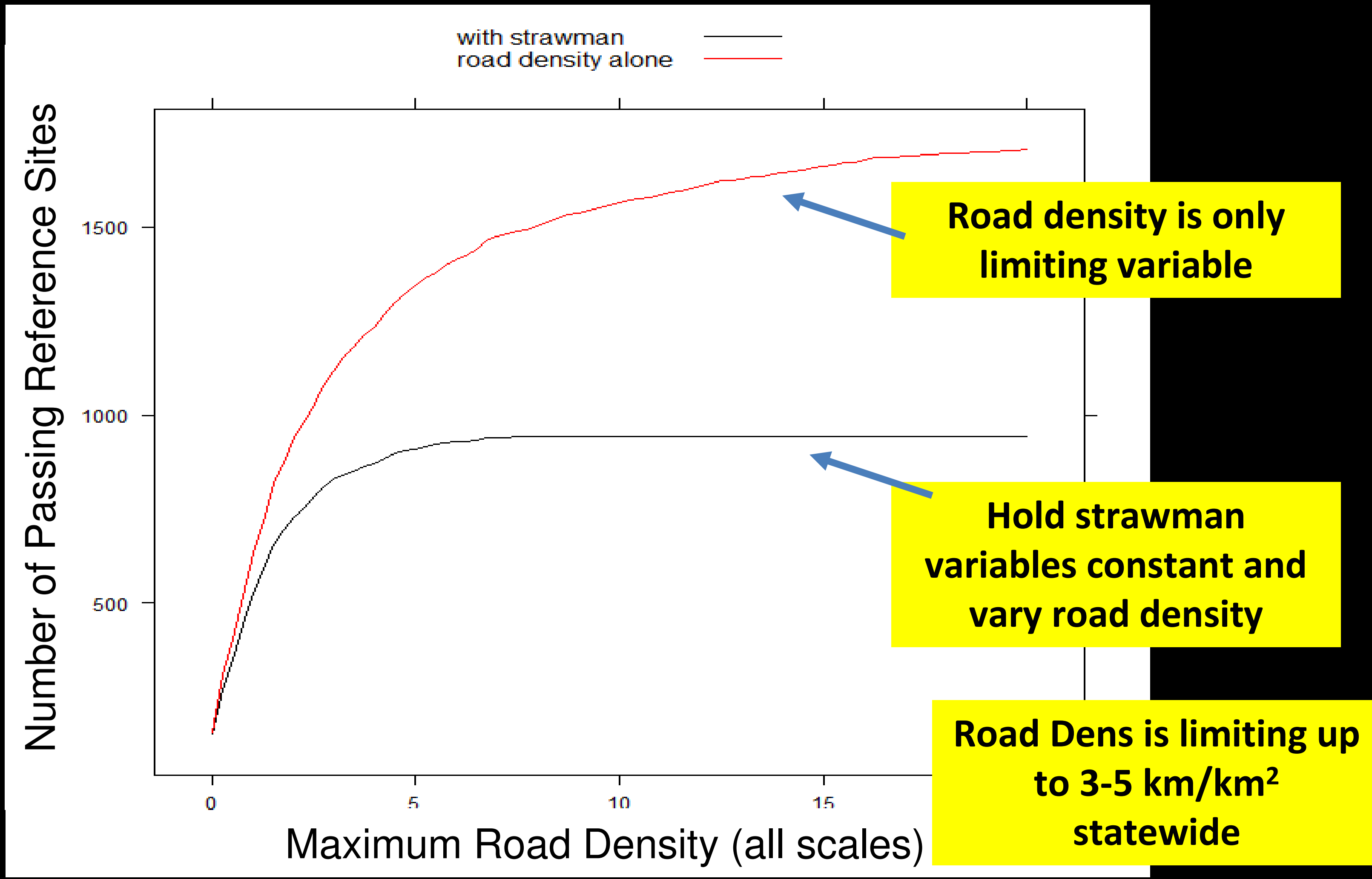
- Stressor thresholds at a fixed proportion of distribution (5th%, 10th %, 25th %, etc.)
- Used a modification of this approach for CA IBI work
- We'll have distributions for all stressors soon



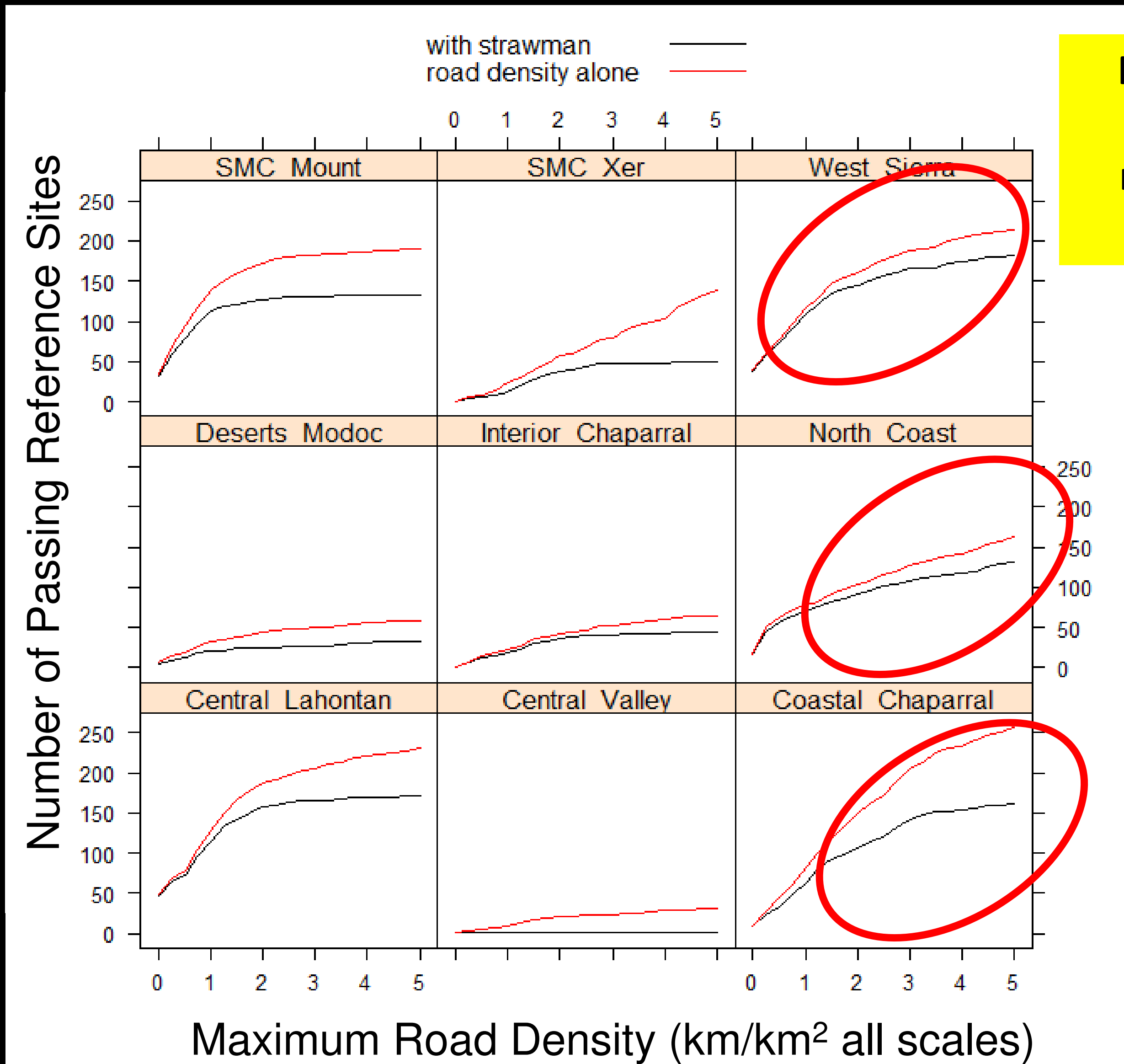
Threshold-Sensitivity:

How many sites do you gain by relaxing thresholds?

Allows us to see where thresholds are limiting

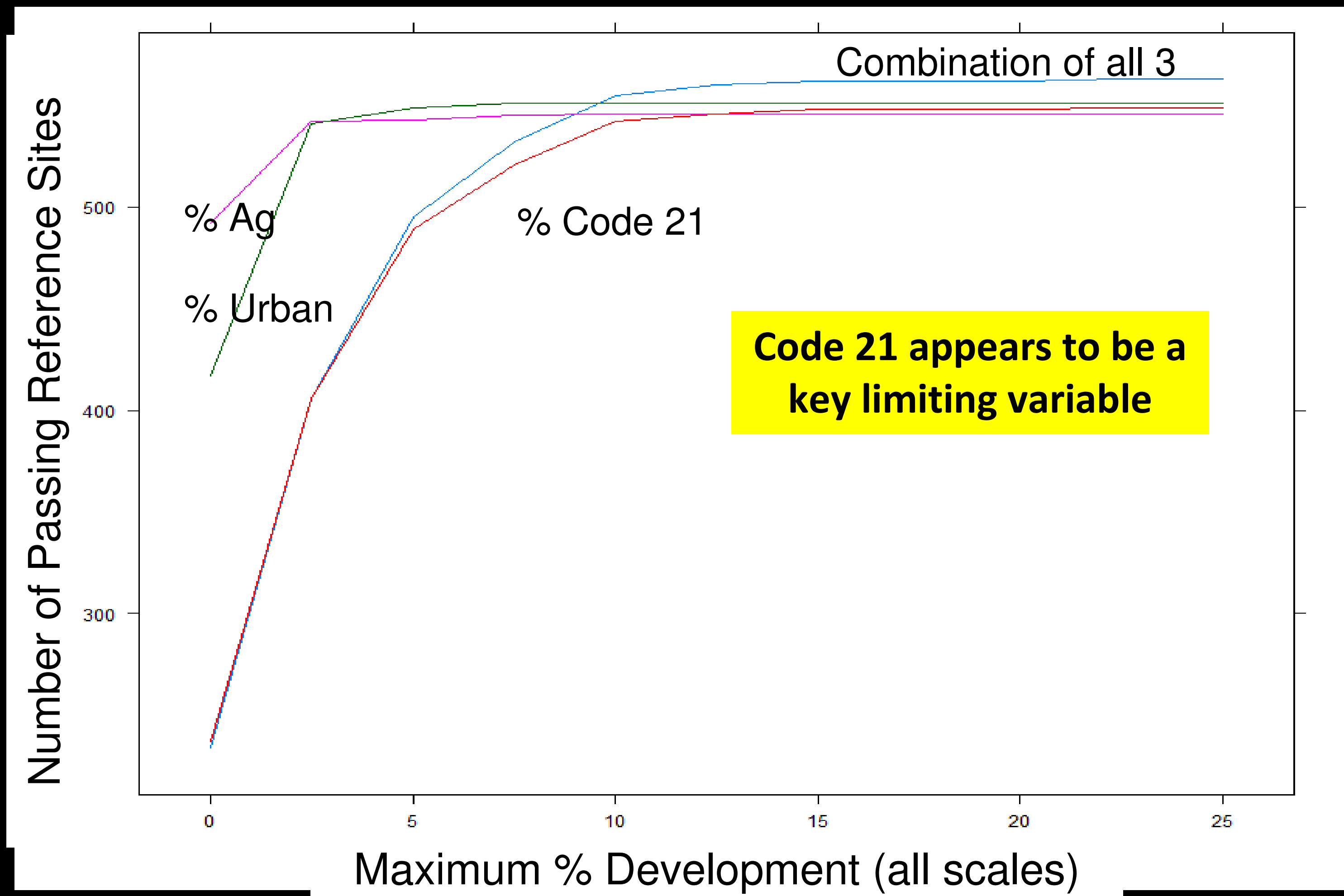


Road Density x Regions



Road density is a major limiting variable for reference sites in some regions

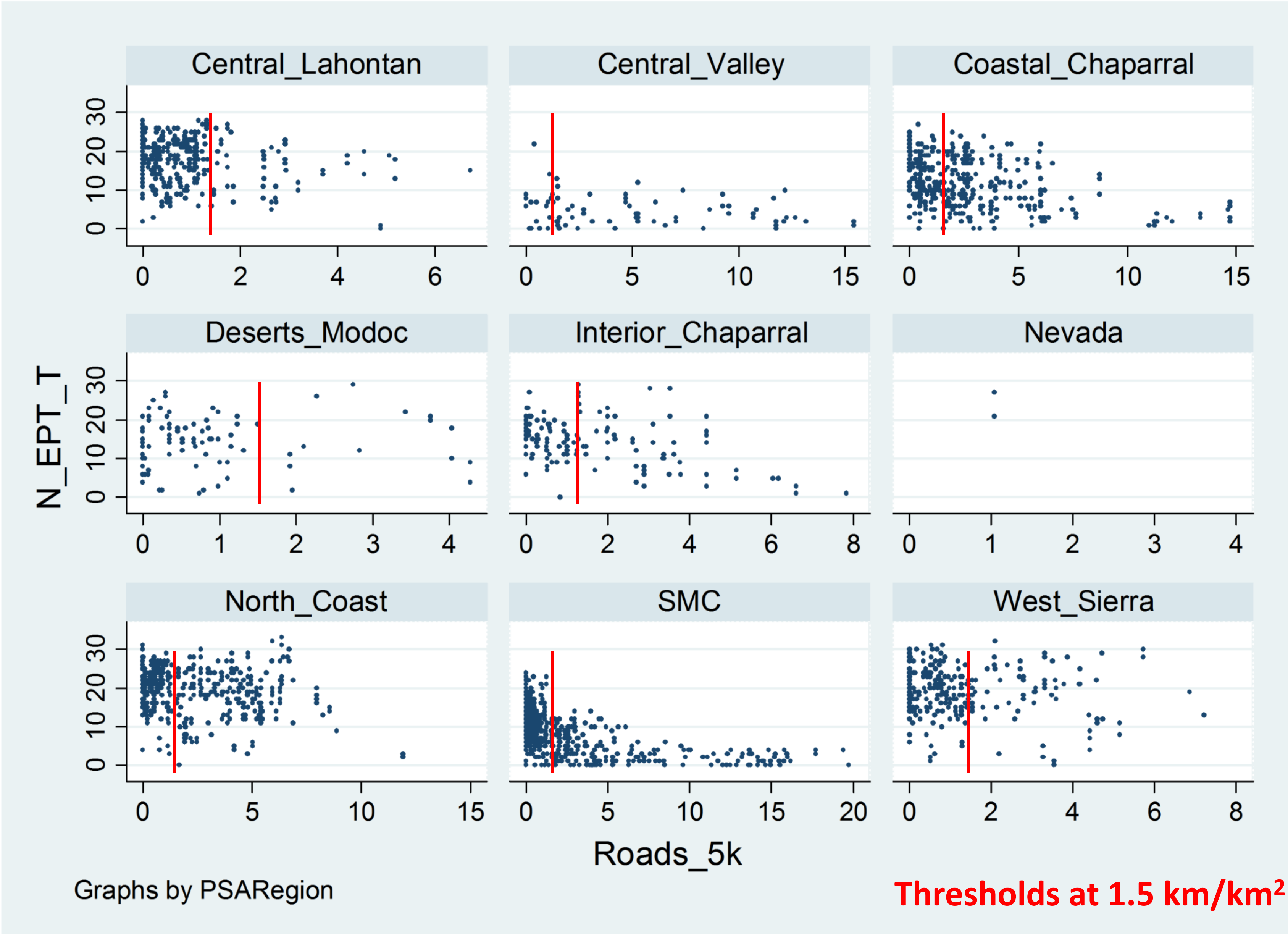
Landscape development variables: hold others constant and vary each independently



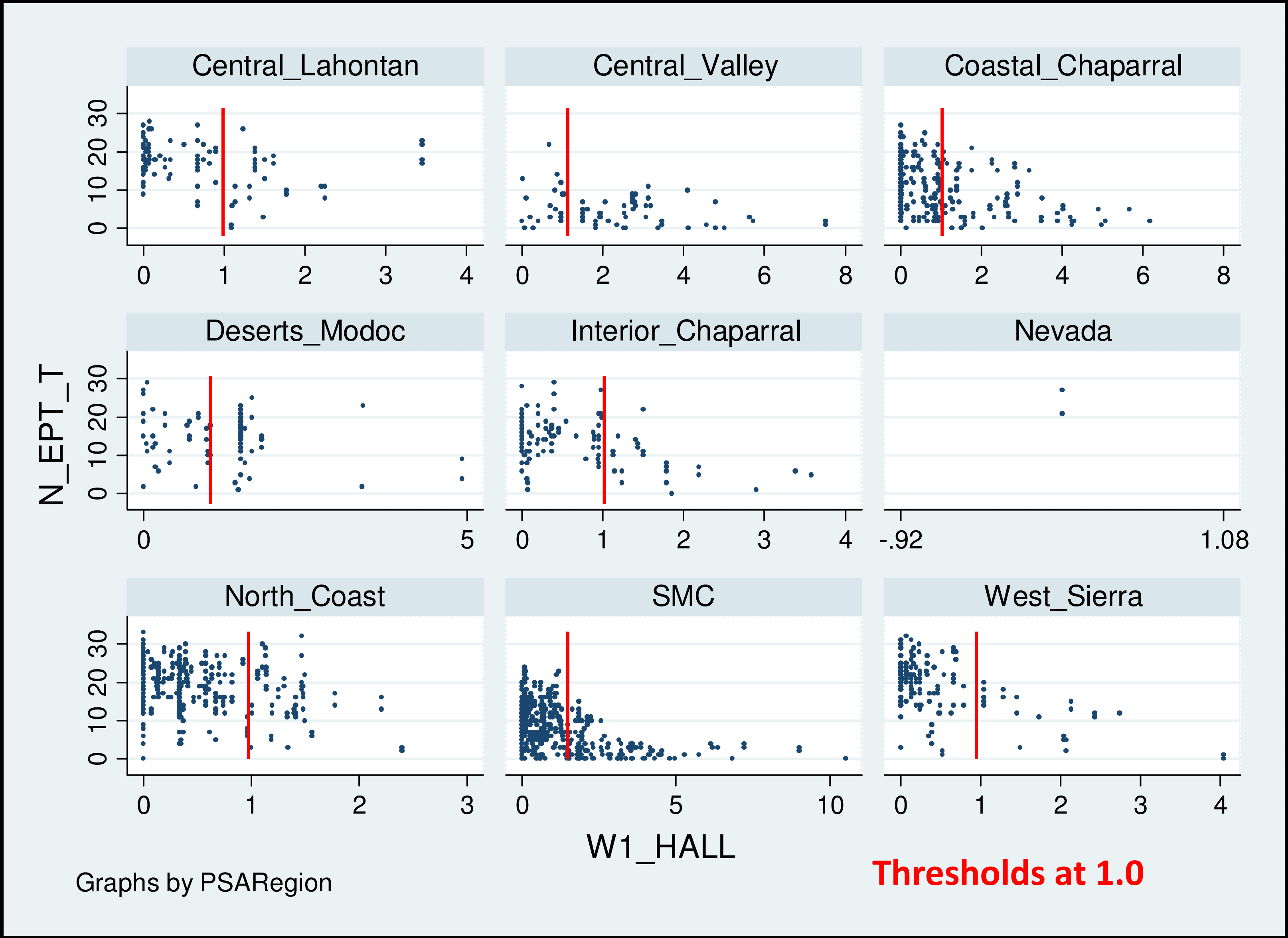
Sensitivity of Biological Metrics to Varying Stressor Thresholds

Since we are using something less than a platinum standard, we'd like some assurance that biological effects associated with our strawman thresholds are minimal

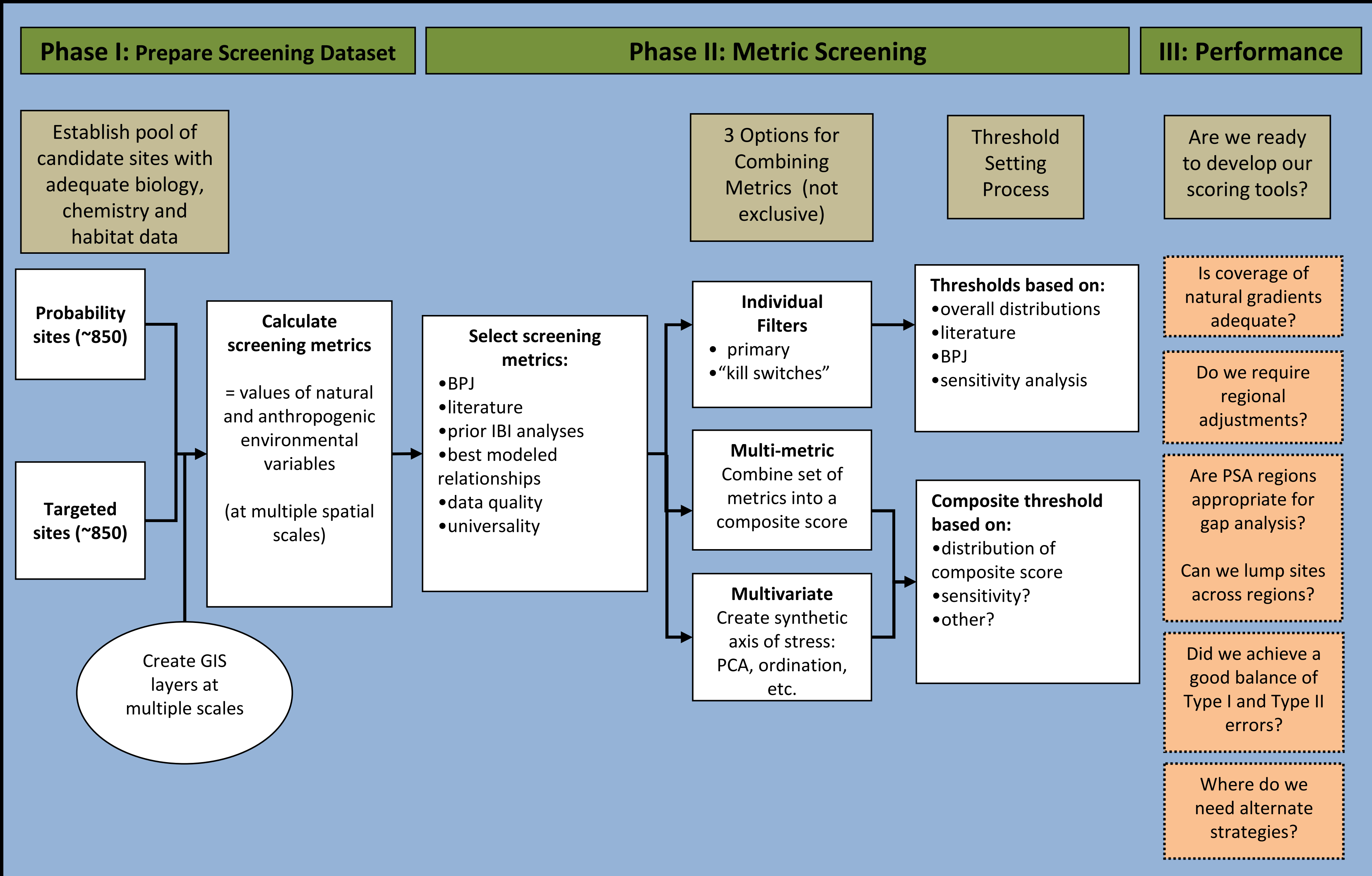
Biology vs. Road Density (km/km^2)



Biology vs. Reach-scale Disturbance




Questions on Phase II?


Phase I: Prepare Screening Dataset
Phase II: Metric Screening
III: Performance
Establish pool of candidate sites with adequate biology, chemistry and habitat data
3 Options for Combining Metrics (not exclusive)
Threshold Setting Process
Are we ready to develop our scoring tools?
Probability sites (~850)
Targeted sites (~850)
Create GIS layers at multiple scales
Calculate screening metrics
= values of natural and anthropogenic environmental variables
(at multiple spatial scales)
Select screening metrics:
•BPJ
•literature
•prior IBI analyses
•best modeled relationships
•data quality
•universality
Individual Filters
• primary
•"kill switches"
Multi-metric
Combine set of metrics into a composite score
Multivariate
Create synthetic axis of stress: PCA, ordination, etc.
Thresholds based on:
•overall distributions
•literature
•BPJ
•sensitivity analysis
Composite threshold based on:
•distribution of composite score
•sensitivity?
•other?
Is coverage of natural gradients adequate?
Do we require regional adjustments?
Are PSA regions appropriate for gap analysis?
Can we lump sites across regions?
Did we achieve a good balance of Type I and Type II errors?
Where do we need alternate strategies?

Phase III: Evaluating Performance

Proximate:

- Did we achieve sufficient coverage?
- Did we account for important variables?

Ultimate:

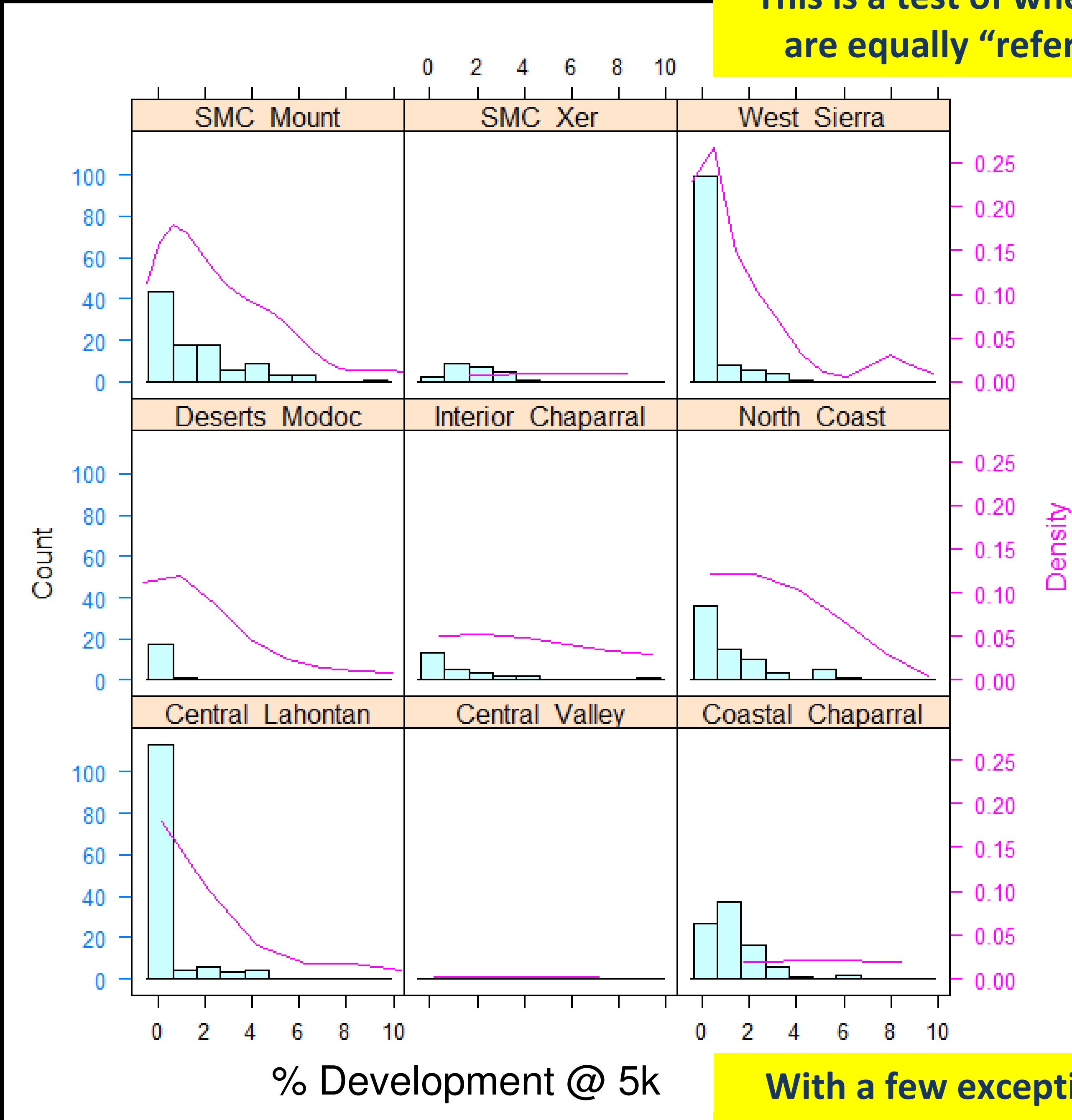
- Have we balanced Type I and II error?
- Does the reference pool support responsive, accurate and precise scoring tools?

Performance Evaluation and Validation

We've explored some of these, and will come back to them in our final questions for the panel

- Is gradient coverage adequate?
  - Do we require regional adjustments? Should we do it anyway?
  - Are PSA regions appropriate/adequate for gradient analysis?
  - Can we lump sites across regions?
- What can biological patterns tell us?
- How can we test if we've balanced Type I and Type II error well-enough?

This is a test of whether our reference sites are equally "referencey" in each region.



With a few exceptions, we're doing pretty well

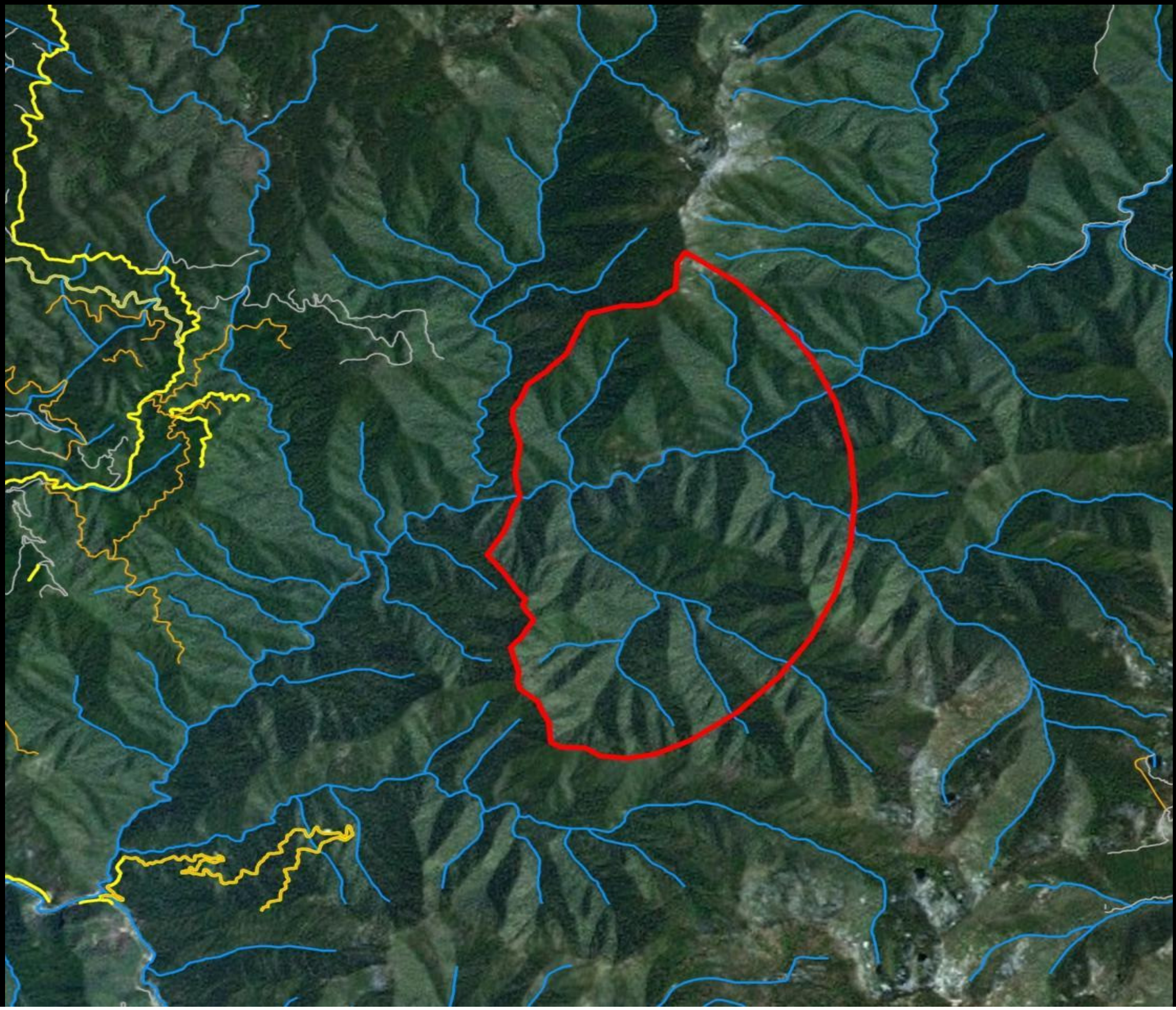
Special Issues:
nutrients, roads and “code21”

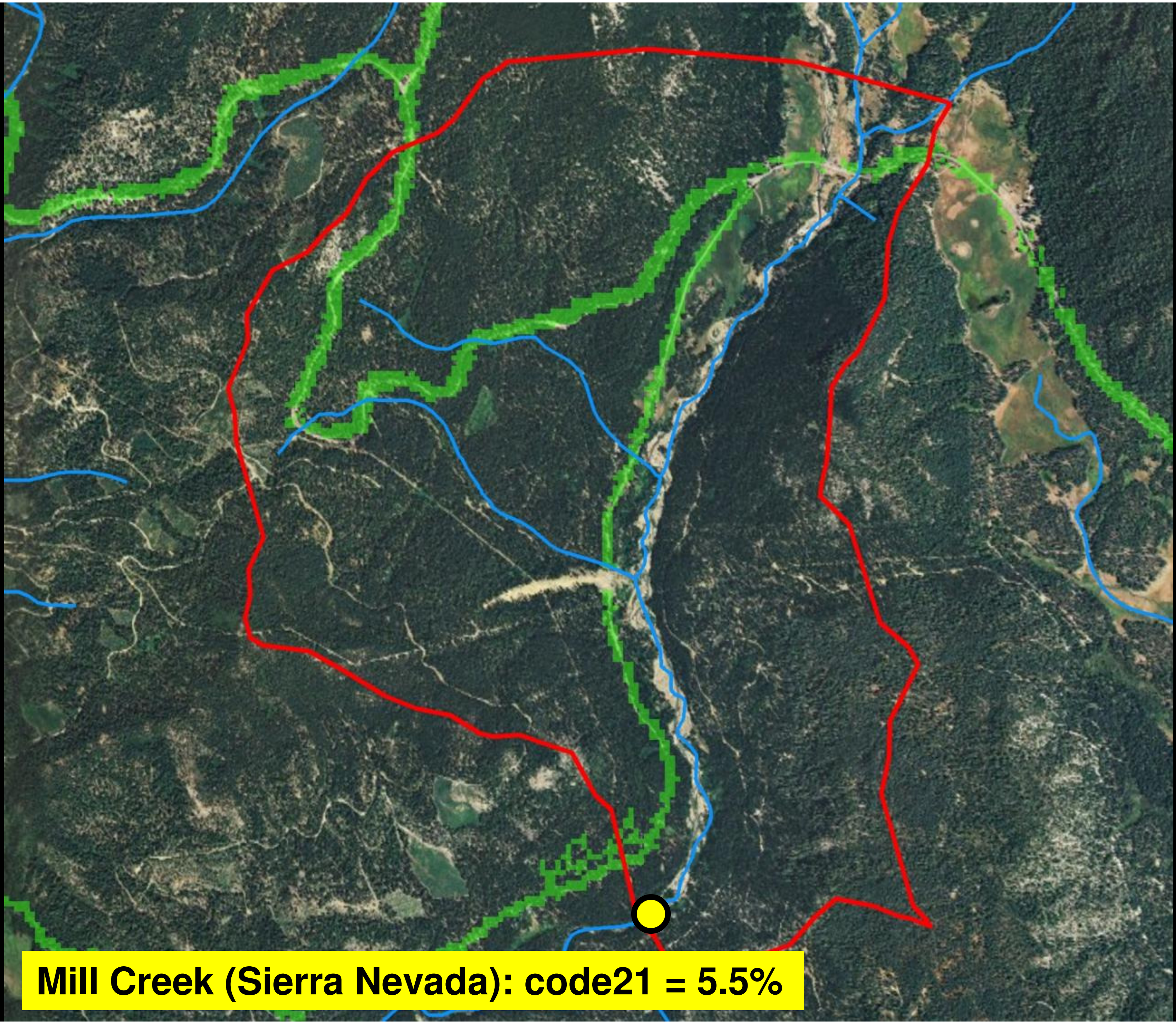
Most variables are fairly straight- forward, but a couple merit special attention (from Stakeholder discussions):

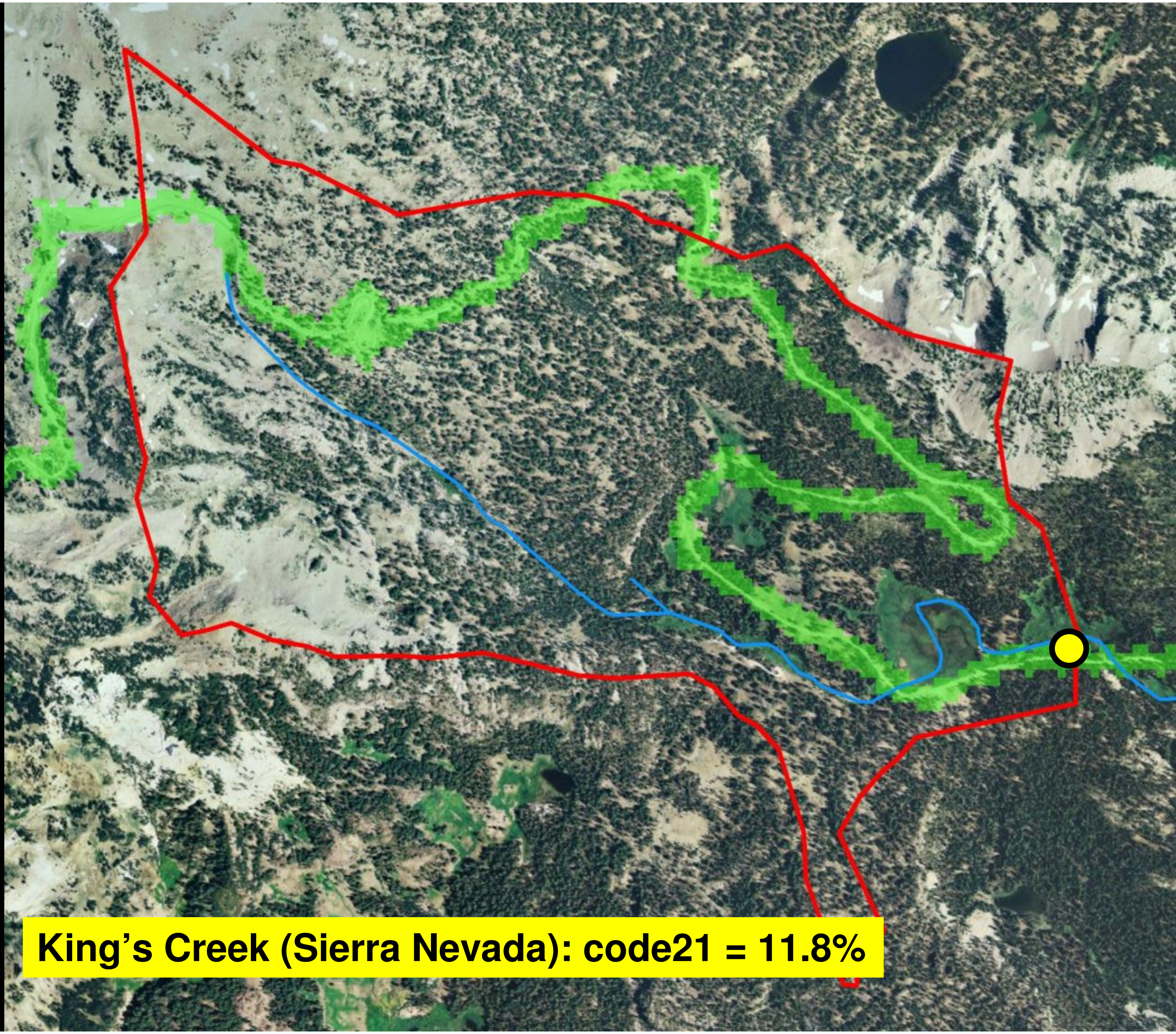
- road density and Code 21 (urban vegetation, roadside vegetation, rural/exurban development)
- nutrients: philosophical questions

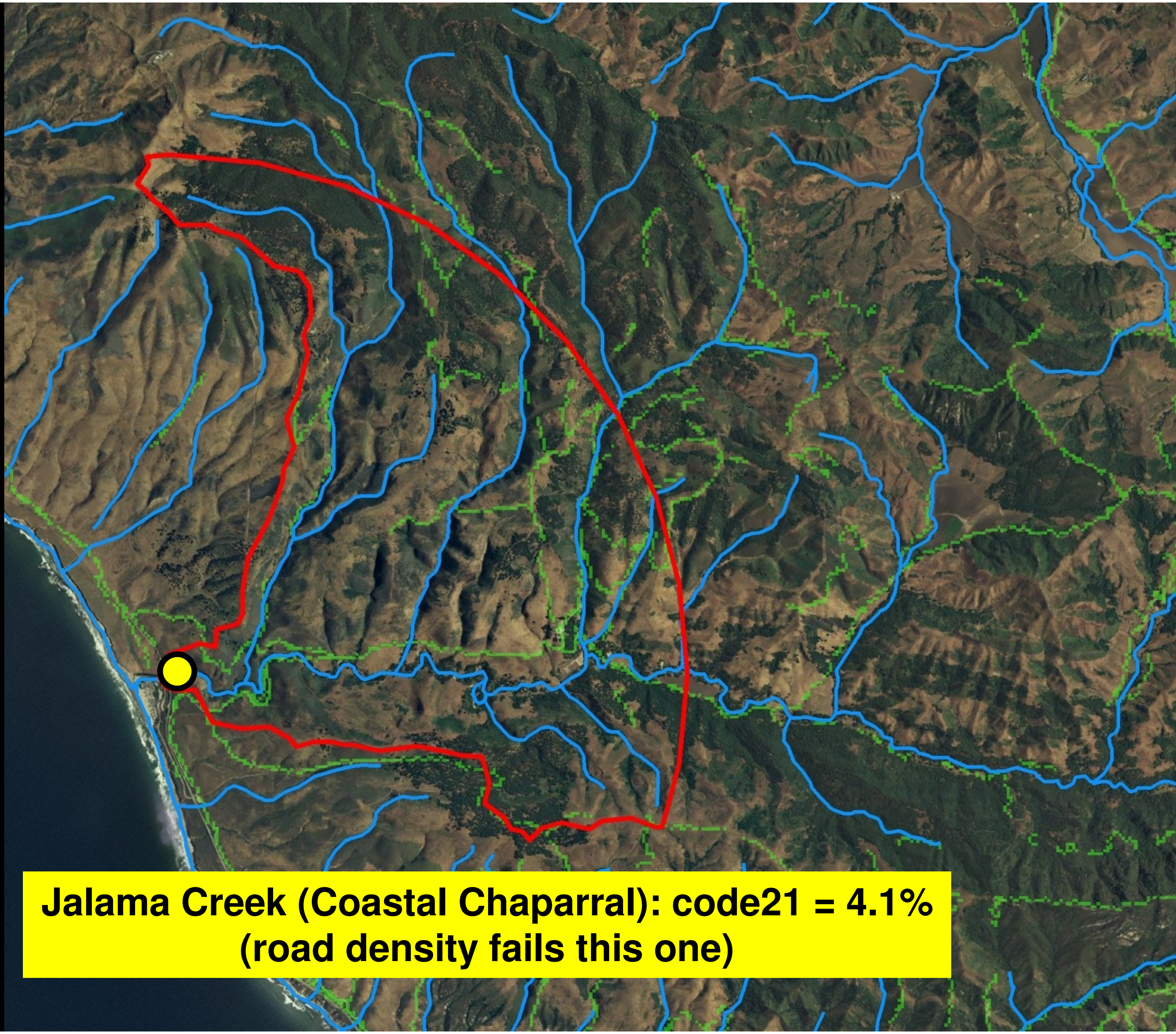
Aerial ortho-photos with overlays of roads and NLCD (2001) Code 21

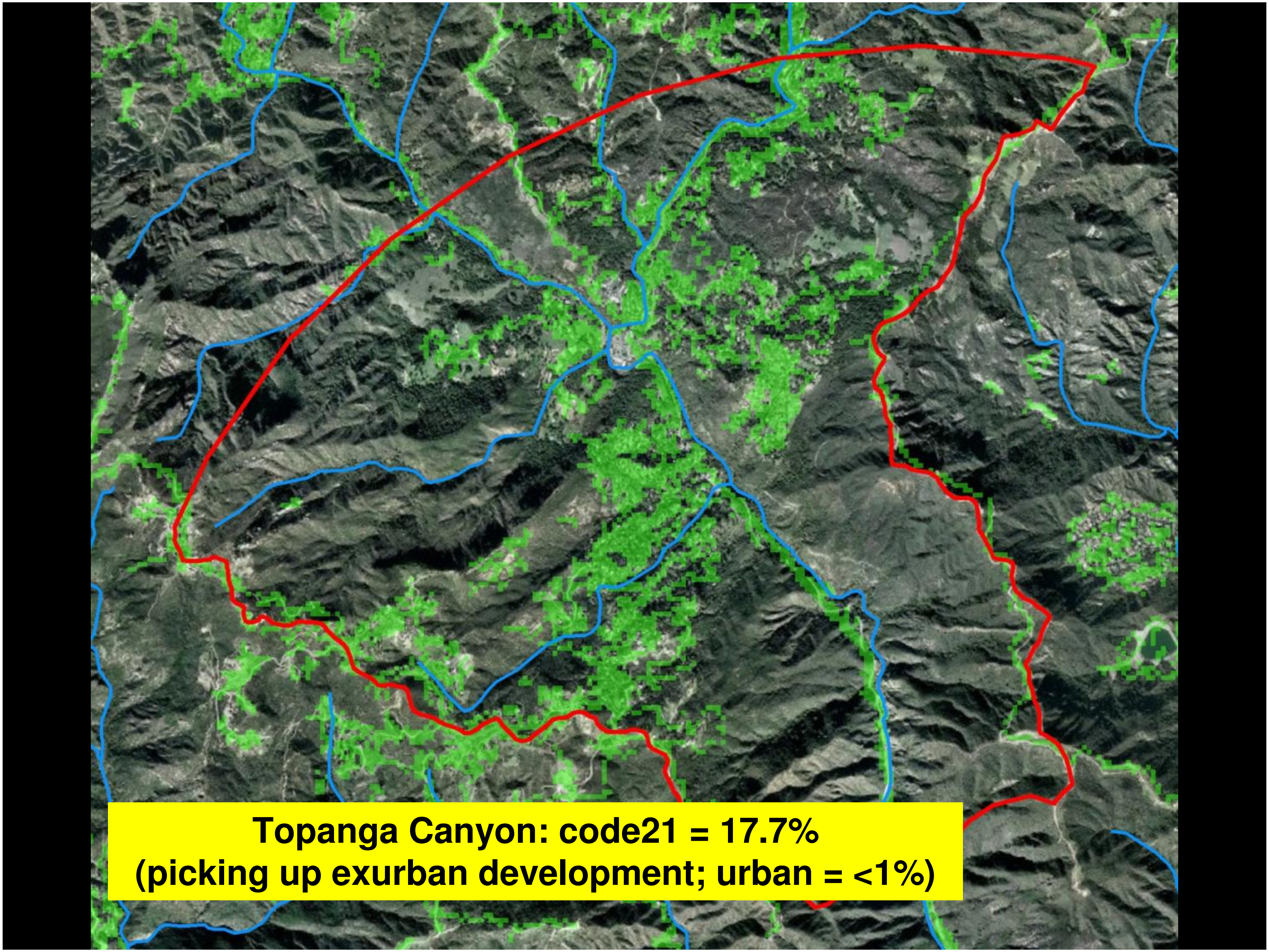
(red line represents watershed within 5k of site)




Mill Creek (Sierra Nevada): code21 = 5.5%


King’s Creek (Sierra Nevada): code21 = 11.8%


Jalama Creek (Coastal Chaparral): code21 = 4.1%
(road density fails this one)


Topanga Canyon: code21 = 17.7%
(picking up exurban development; urban = <1%)

Special Issues:
nutrients, roads and "code21"

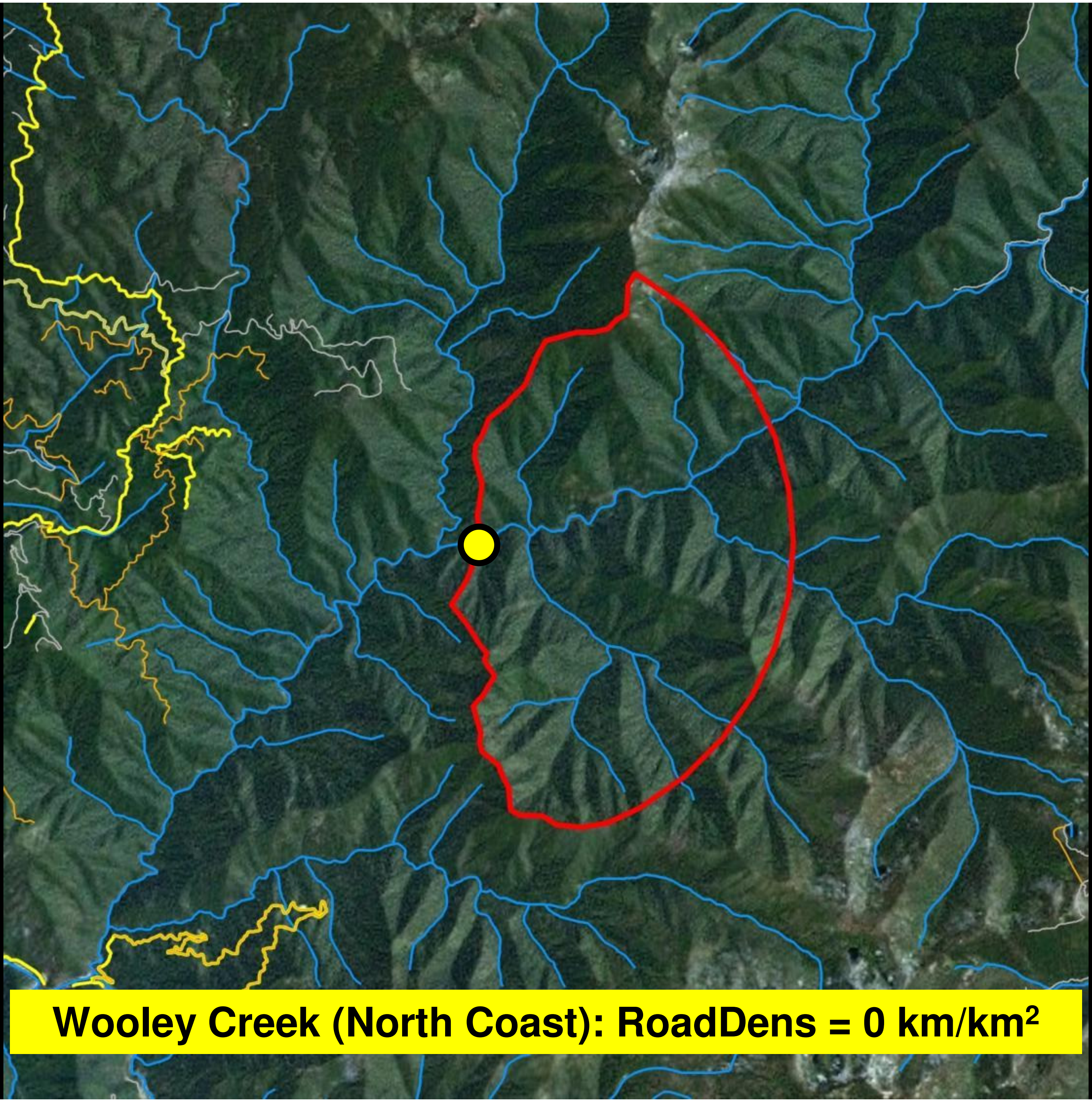
We created a custom road layer that added timber harvest roads (from USFS and CDF records) to a fairly complete base layer

Road classes:

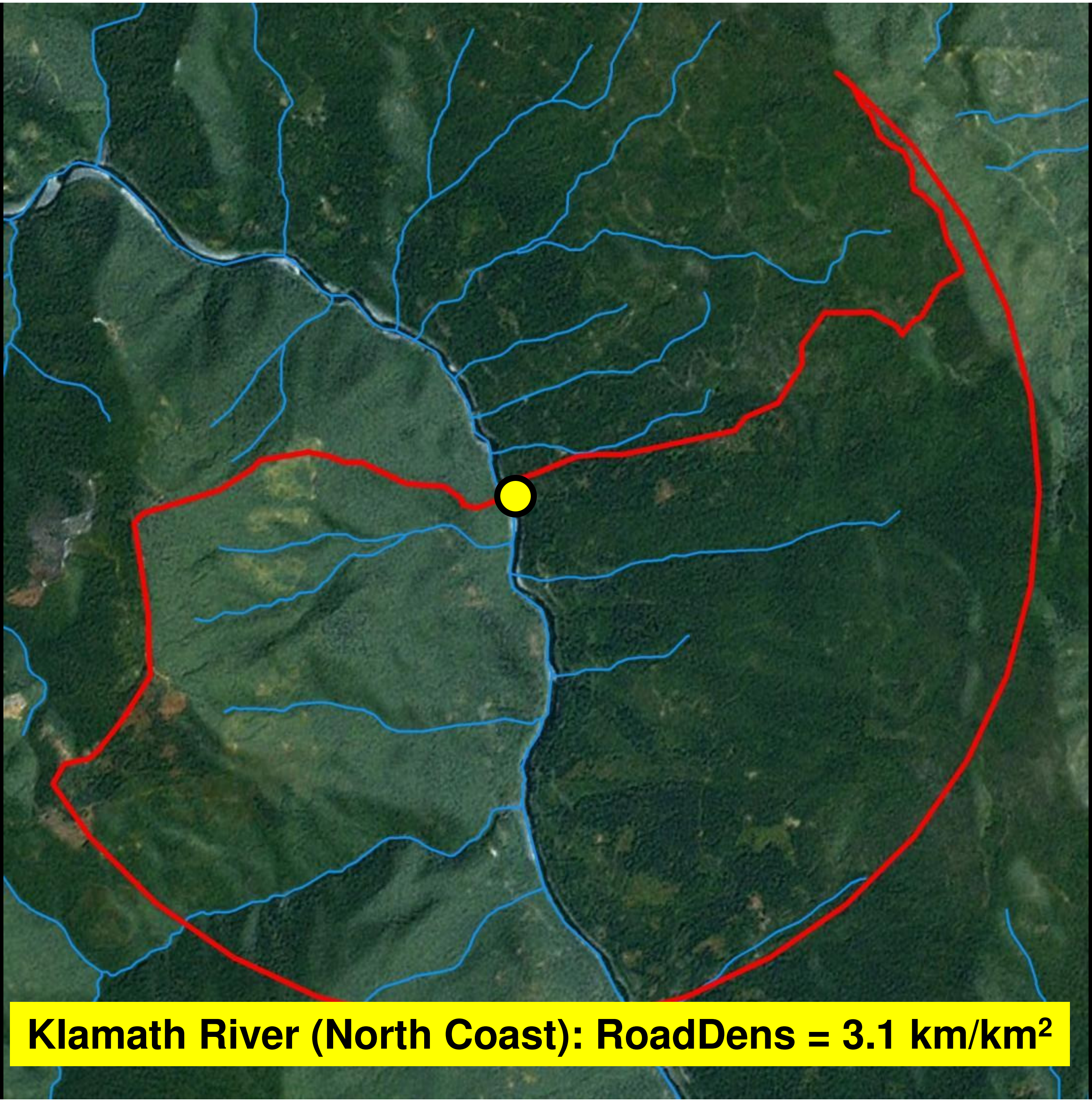
1, 2 = paved roads

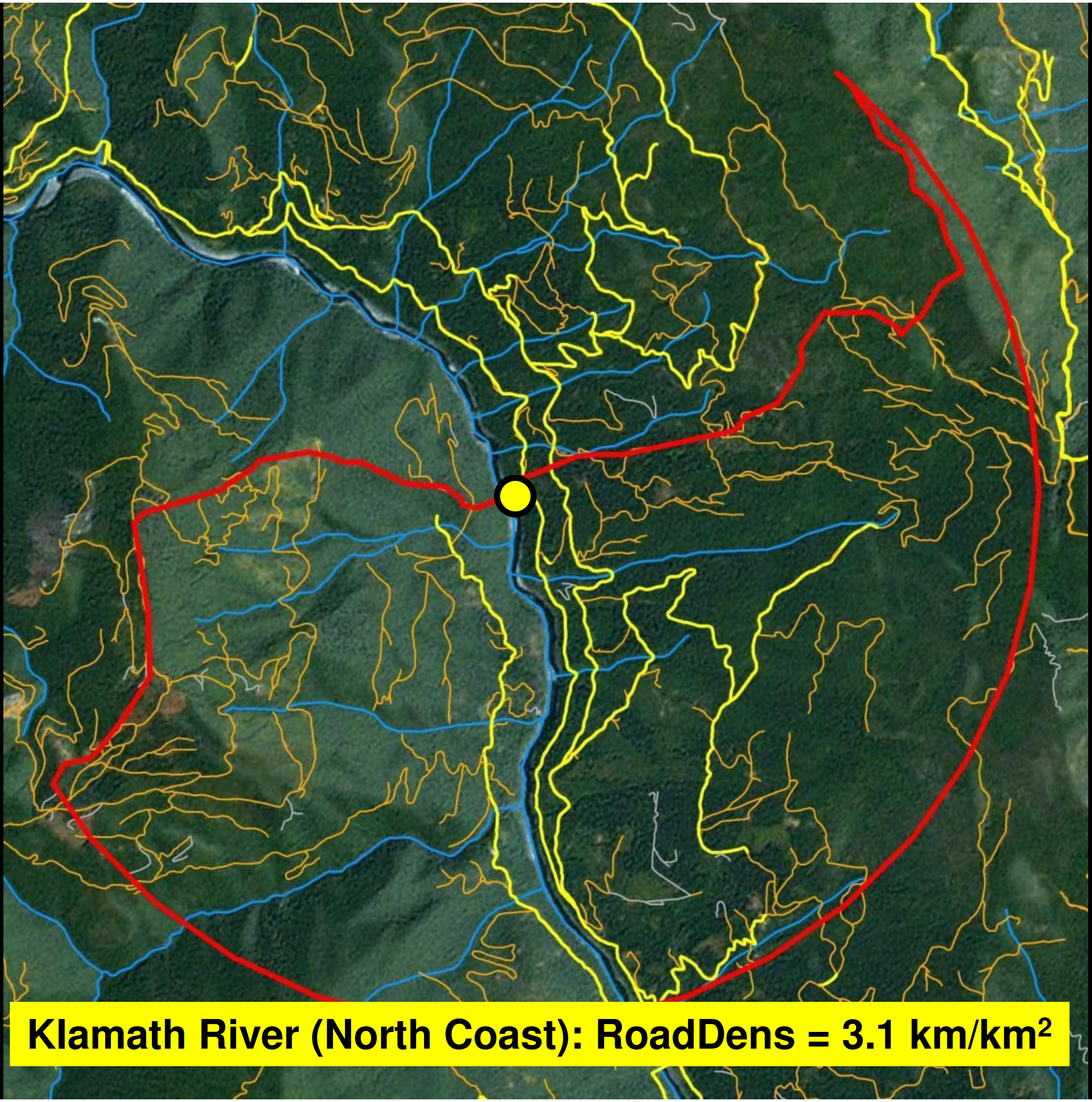
3 = unpaved roads (including logging roads)

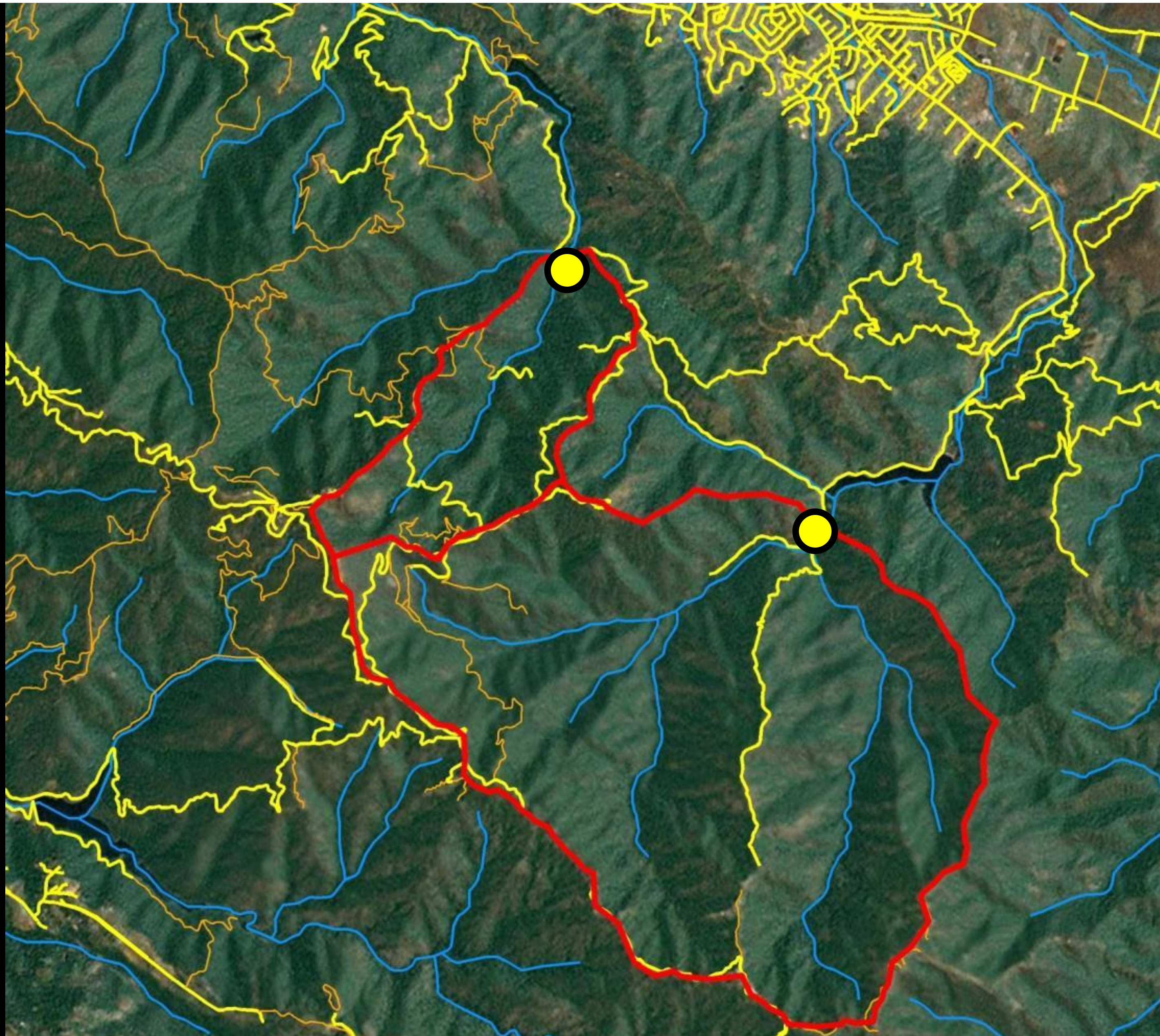
4 = other (trails, 4 x 4 roads)



Wooley Creek (North Coast): RoadDens = 0 km/km^2


Klamath River (North Coast): RoadDens = 3.1 km/km²


Klamath River (North Coast): RoadDens = 3.1 km/km^2



Guadalupe Creek (smaller) and Herbert Creek (Bay Area): RoadDens = 2.0 and 1.1 km/km^2

Special Issues: roads and code 21

Can we raise the bar for road density and code 21 without increasing Type I error very much?

Should we use these variables to make regional adjustments? Statewide adjustments?

Special Issues: nutrients

Should we use nutrient data in our reference screens? If so, how should we choose thresholds?

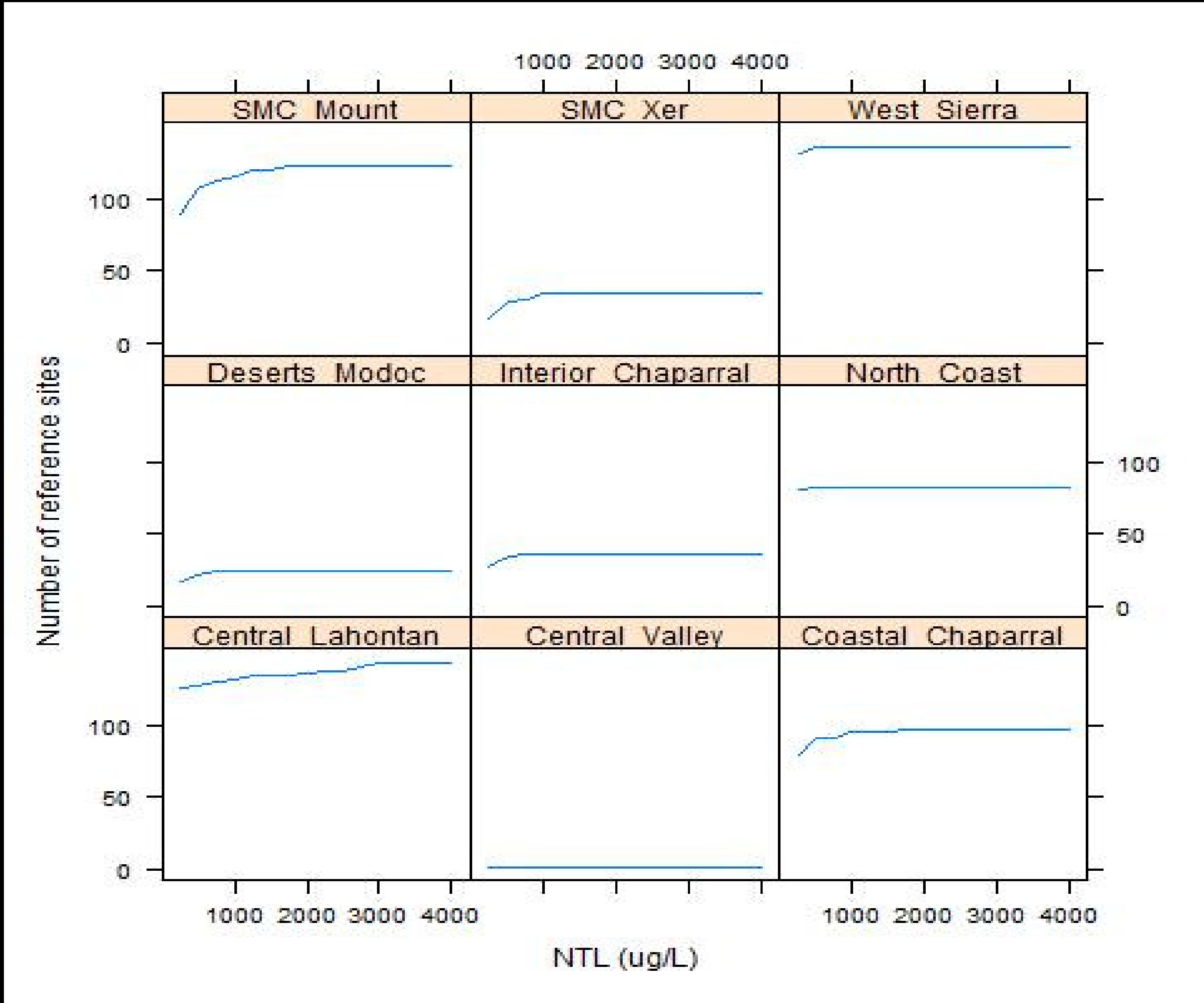
Pros:

- Nutrients are a good surrogate for things we can't measure well (esp. grazing)
- An important limiting factor for biological integrity
- Key focus of current regulatory interest

Cons:

- Effect levels for nutrients are very setting-specific
- Nutrient data (like other WQ parameters) are not always available or reliable

Our nutrient thresholds are generally NOT limiting our site pools

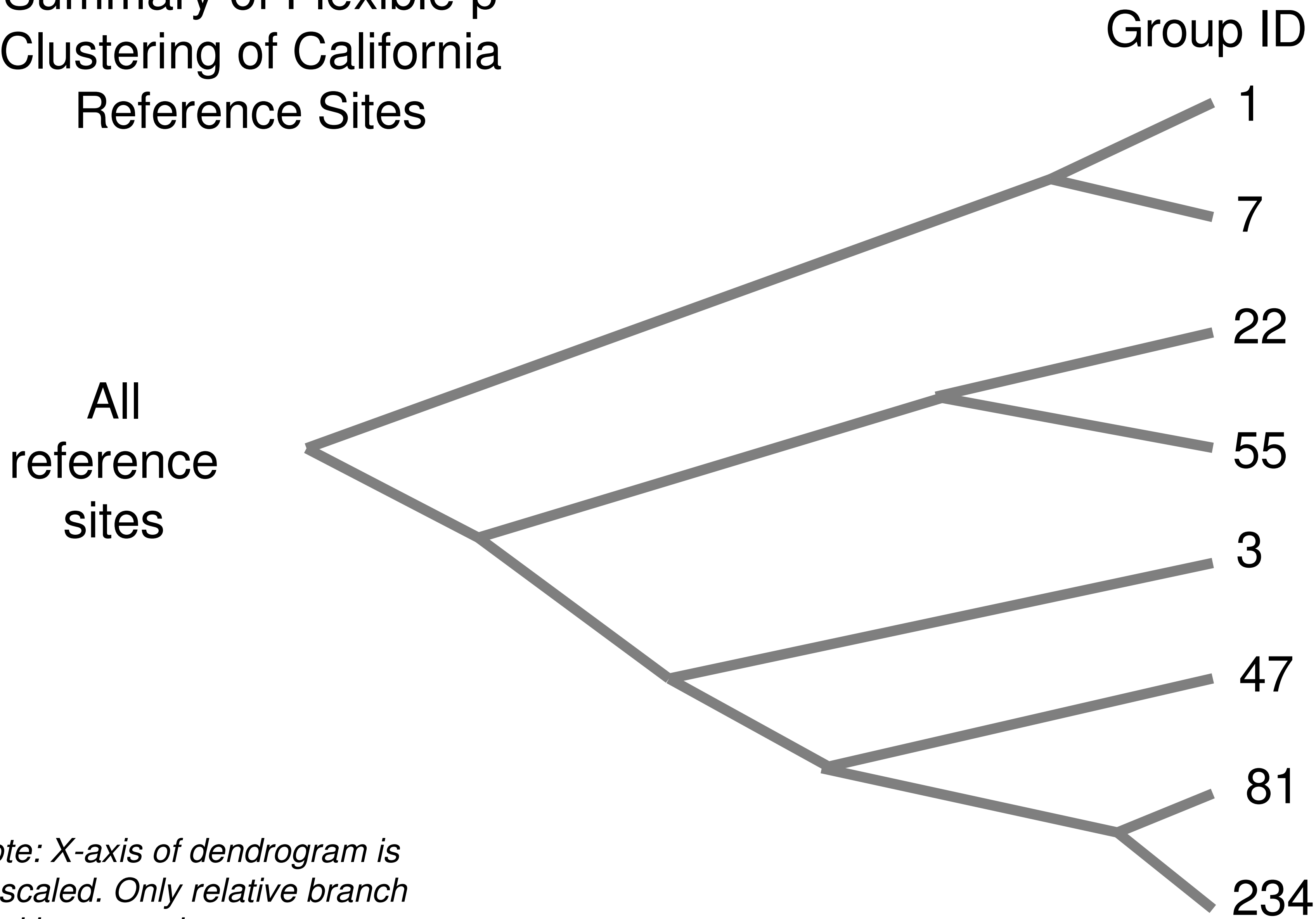


Clustering Analysis:
Why look at biological clusters?

- A supplement to our gradient analyses
- Are PSA regions good for testing gradients
- Tests if we are under-representing biological clusters
- Tells us if we can lump sites across regions

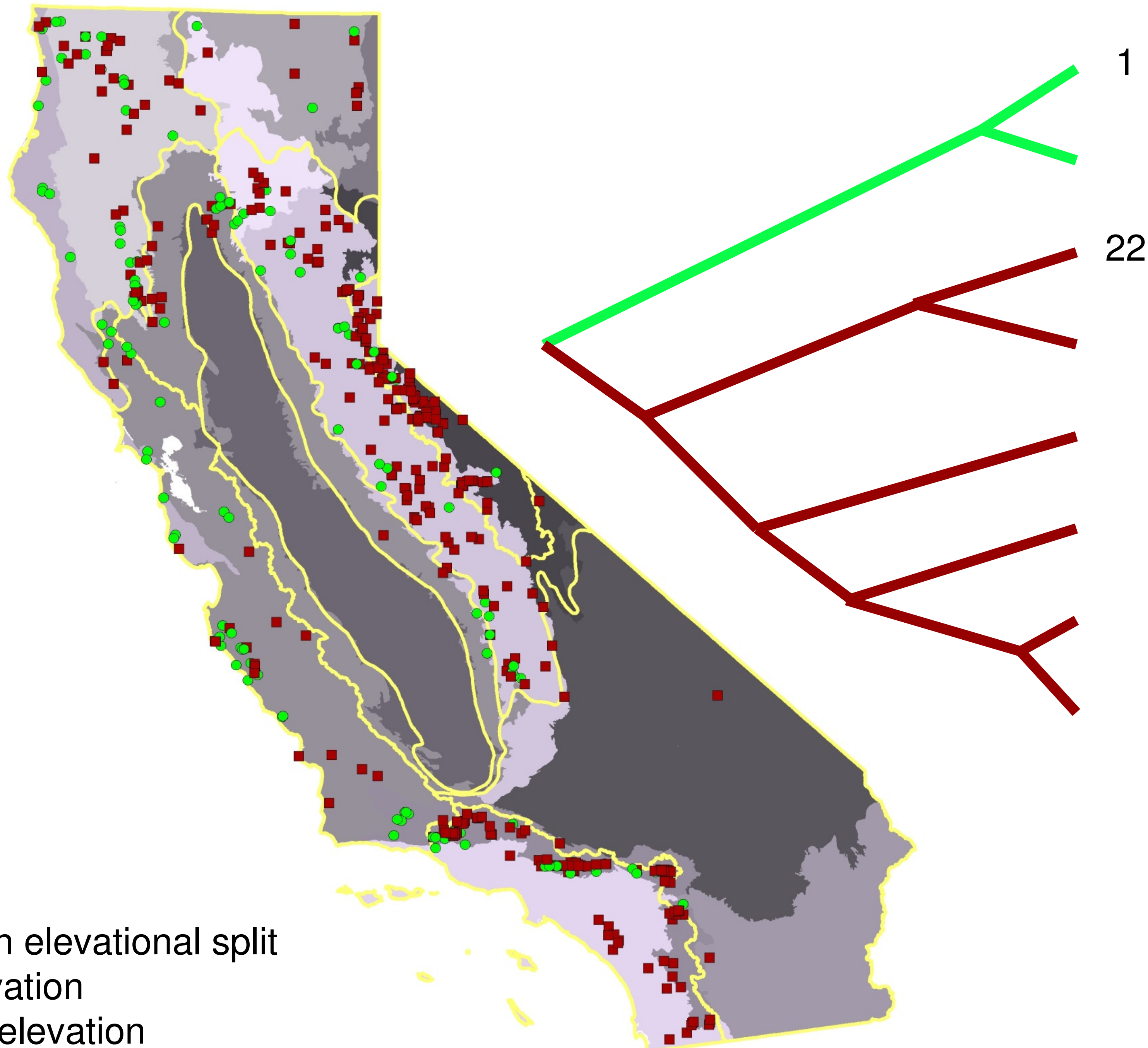
- Preliminary work for modeling expectation
  - Explore ecological trait associations with clusters
  - Identify patterns in natural gradients associated with clusters (i.e., predictors)

Summary of Flexible β Clustering of California Reference Sites



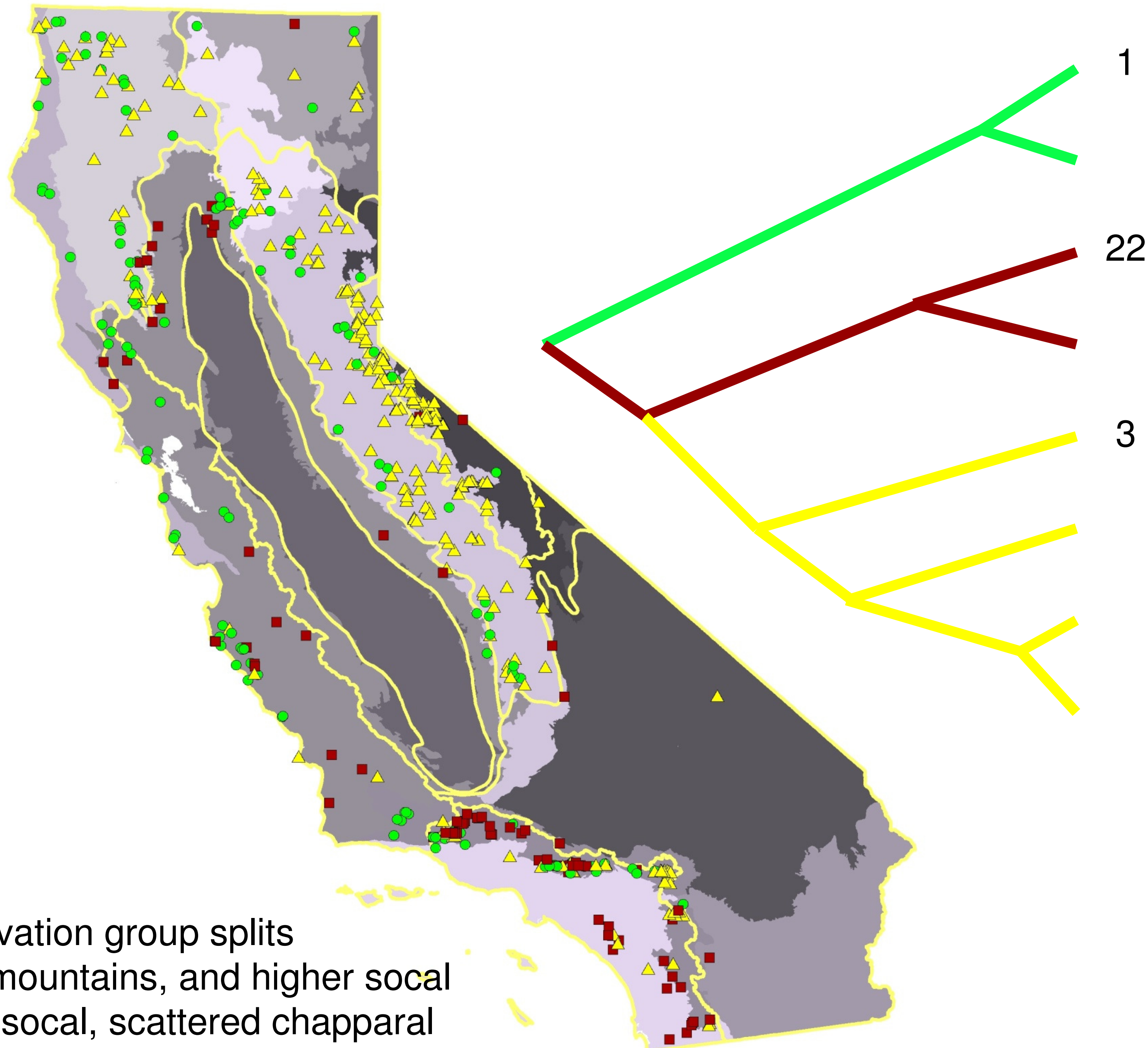
Note: X-axis of dendrogram is unscaled. Only relative branch positions are shown.
β = -0.25. Distance: Bray-Curtis.

2 groups



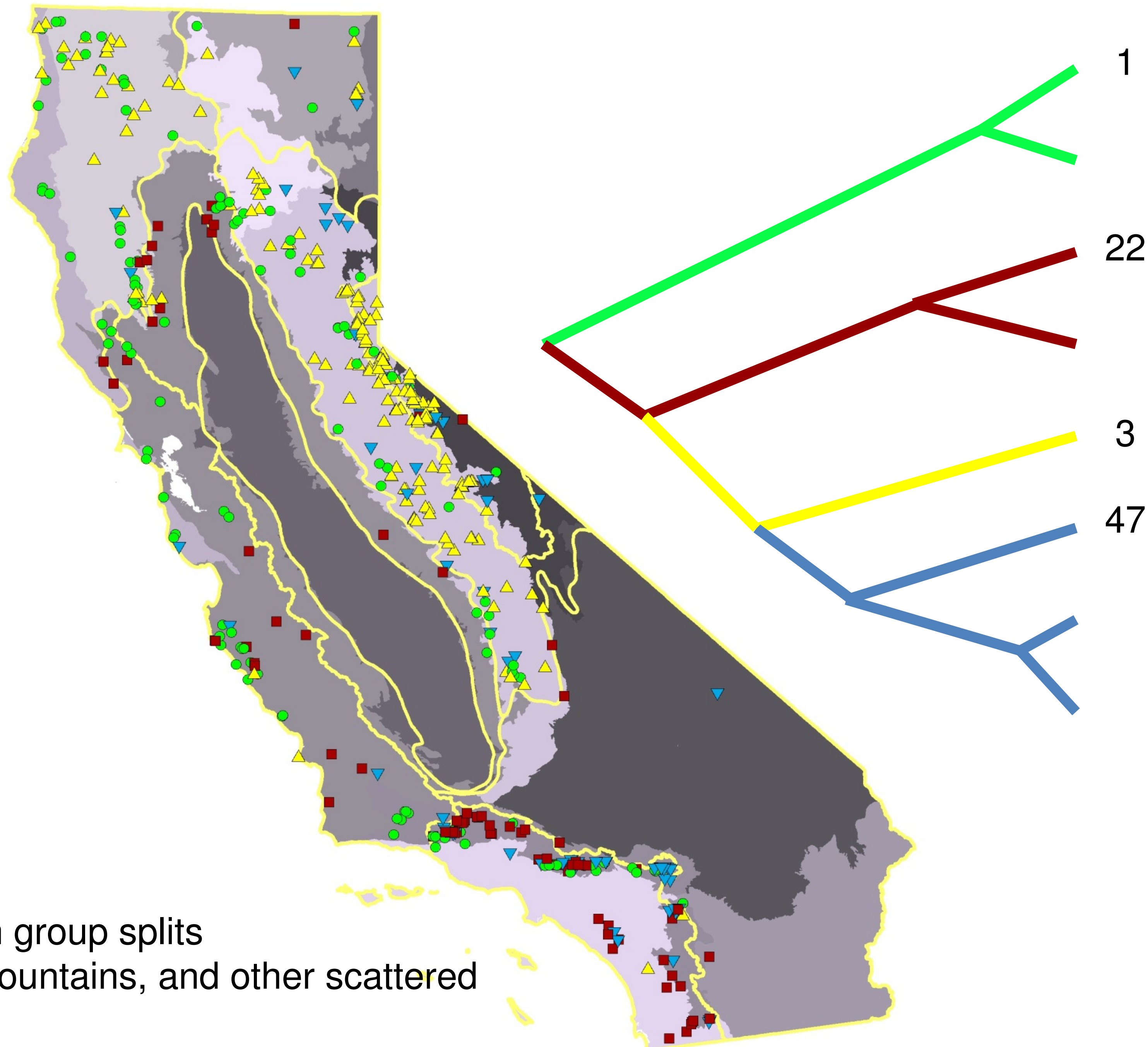
Roughly, an elevational split
1: Low elevation
22: Higher elevation

3 groups



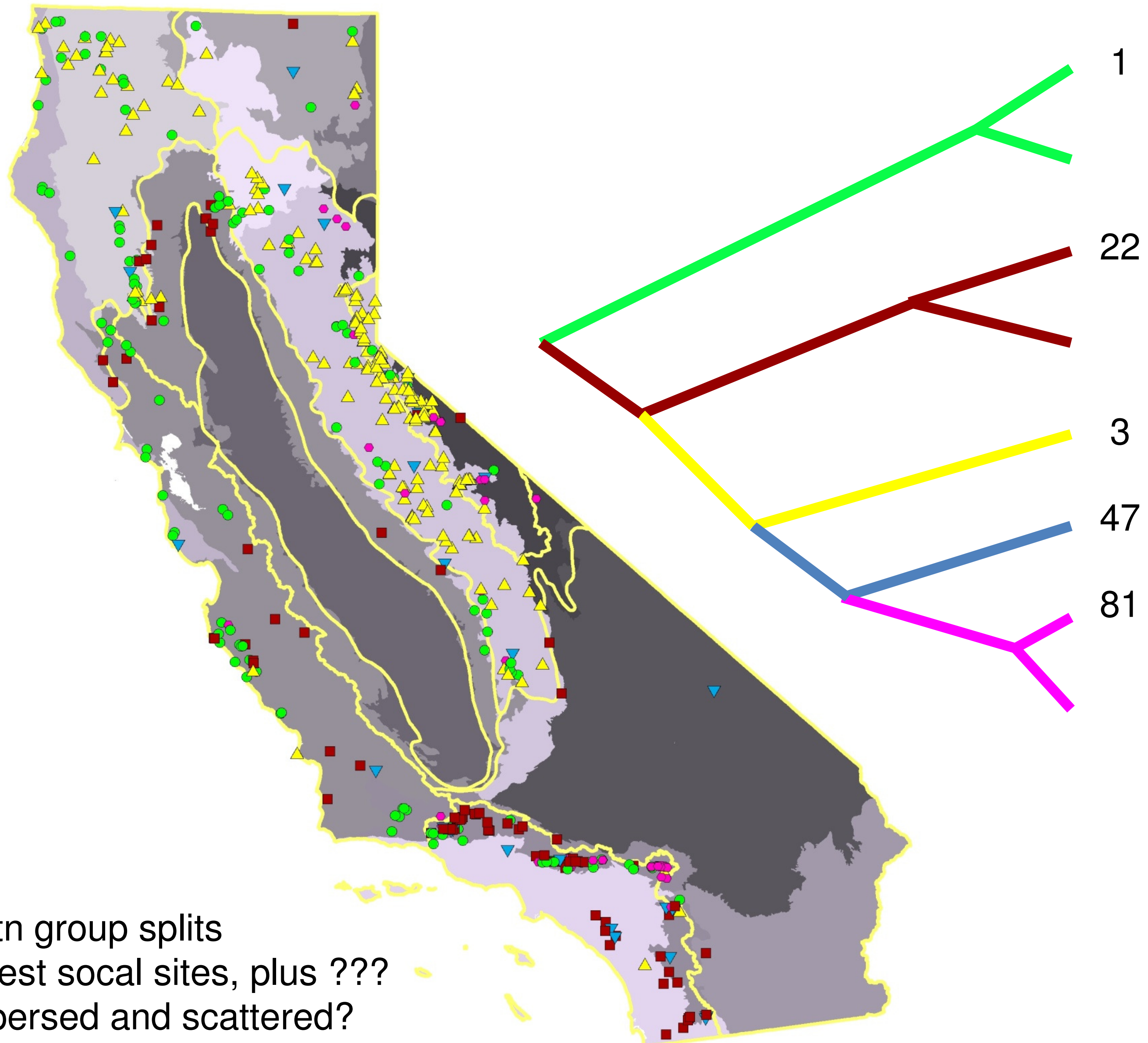
The high elevation group splits
3: Northern mountains, and higher socal
22: Mid elev socal, scattered chapparal (higher?)

4 groups



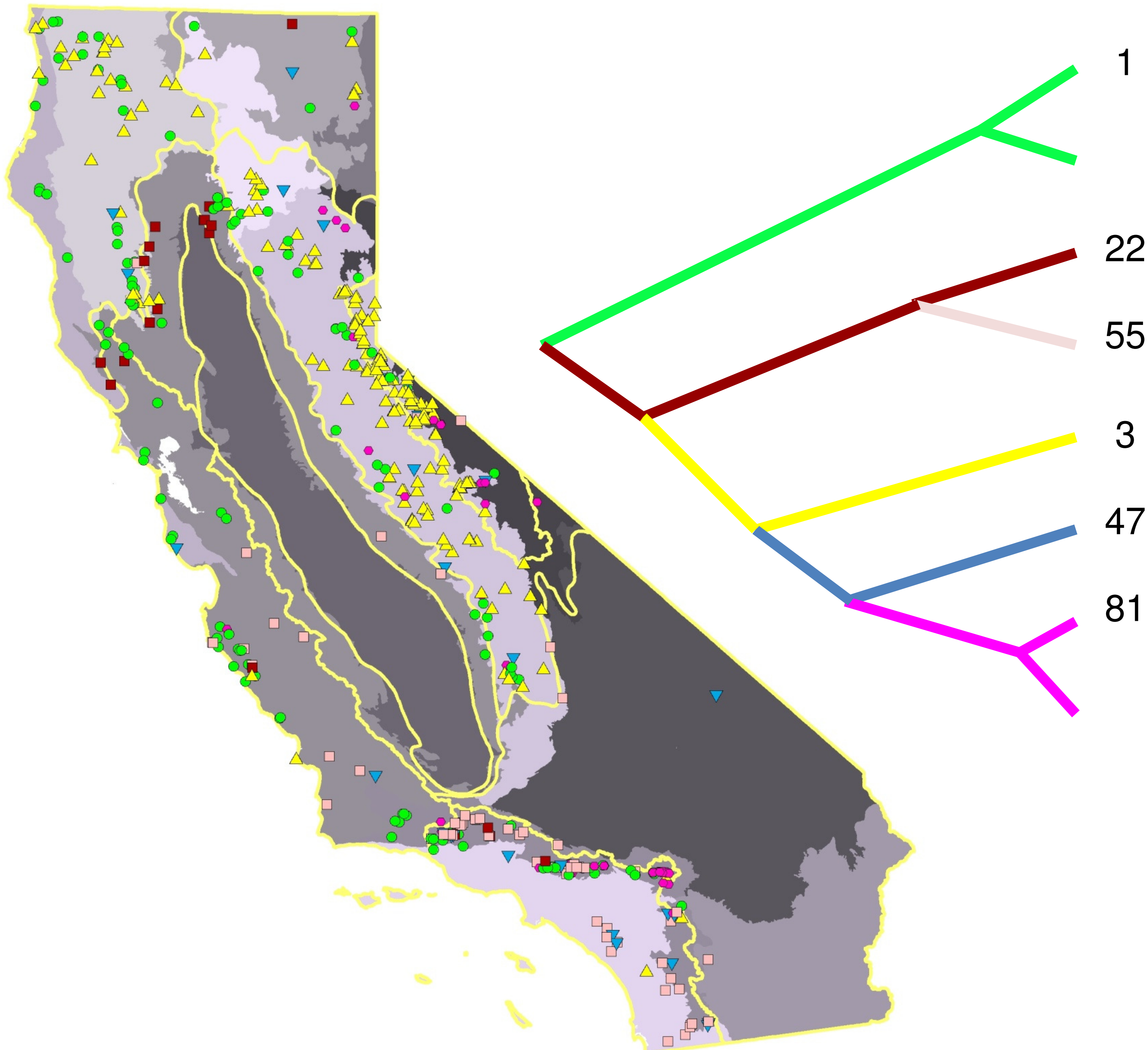
The northern group splits
47: So cal mountains, and other scattered mtns
3: Remaining mountain sites

5 groups

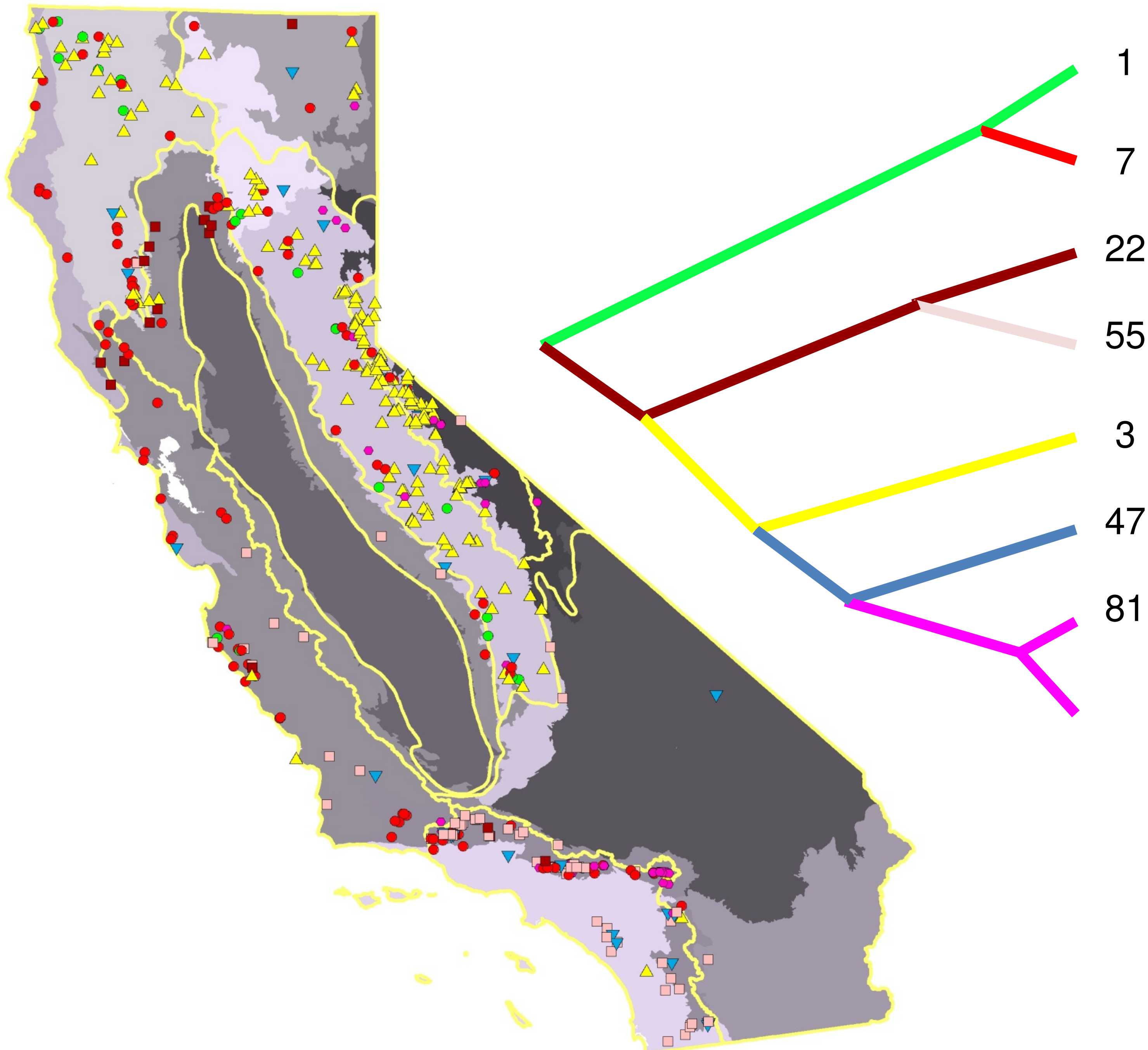


The socal mtn group splits
81: The highest socal sites, plus ???
47: Very dispersed and scattered?

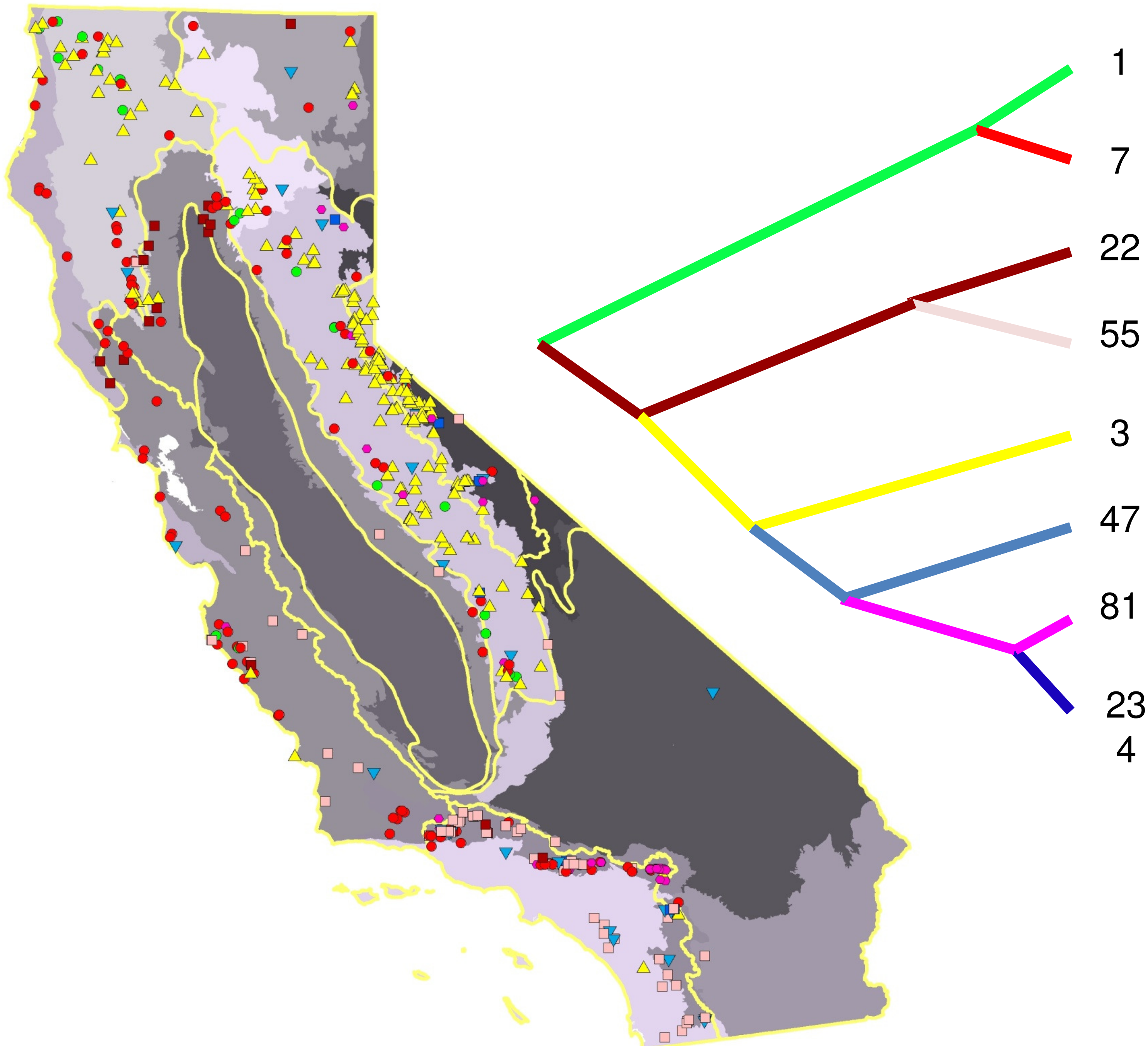
6 groups



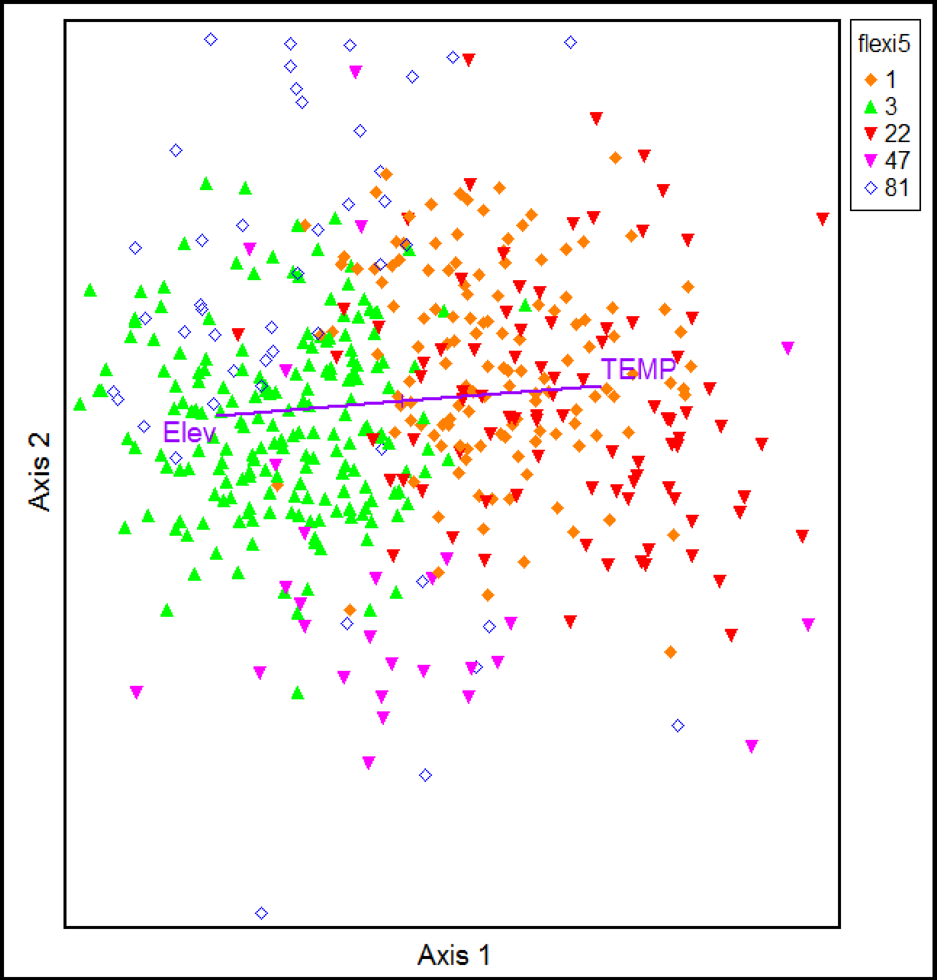
7 groups

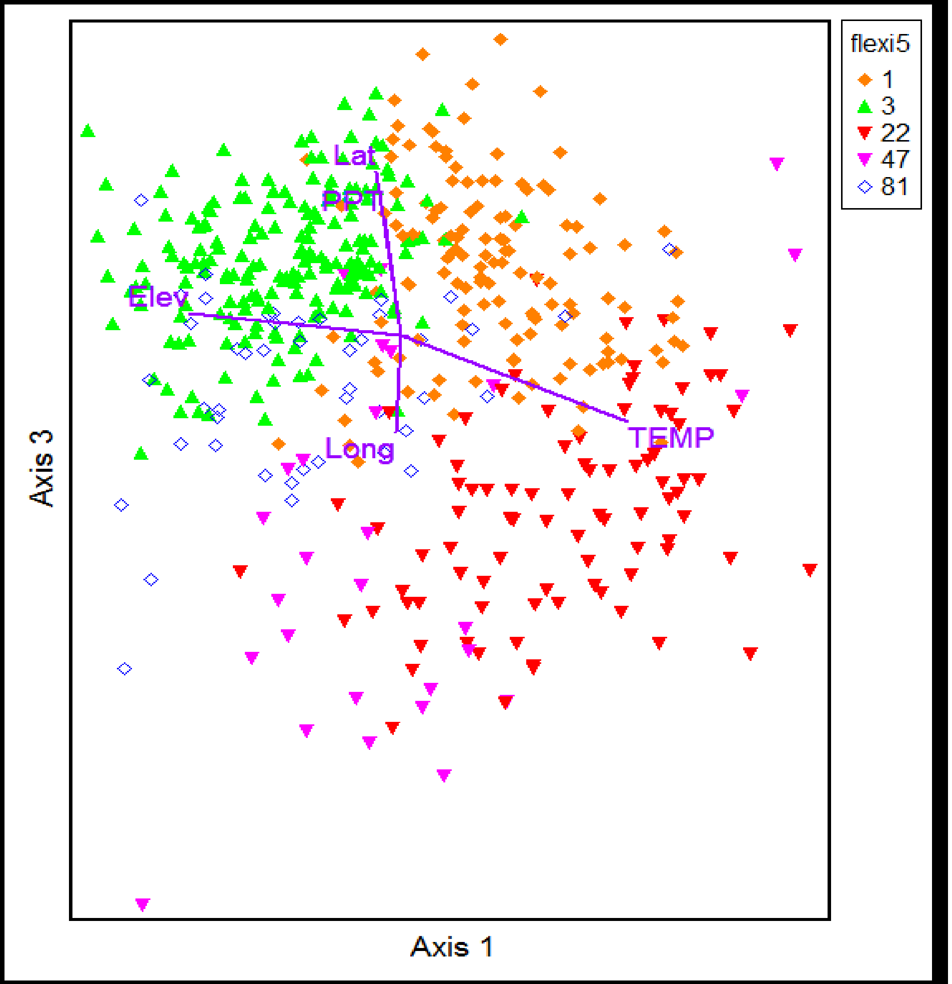


8 groups



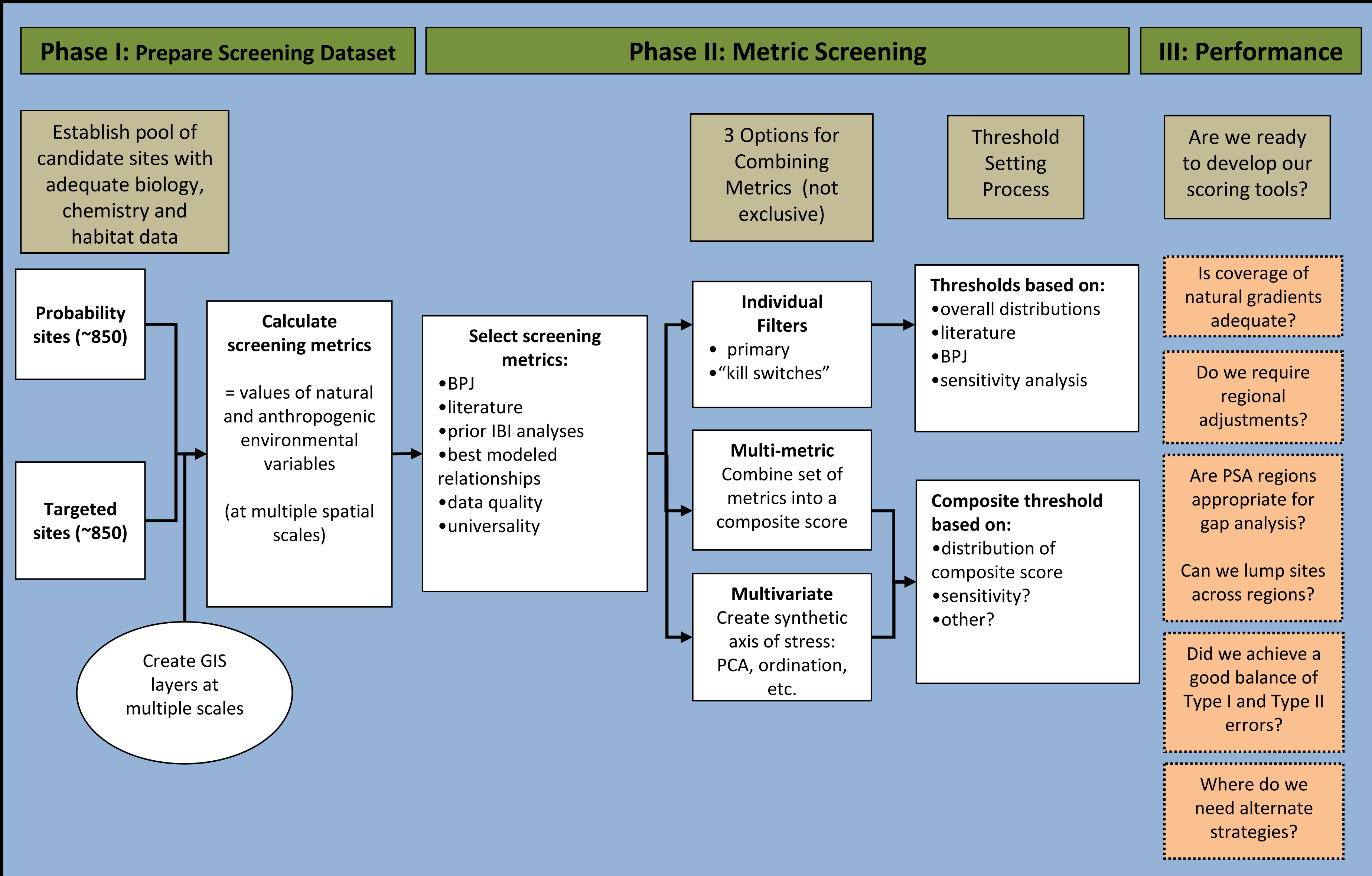
NMS Ordination
(5 groups with key predictors):
Another way to identify important gradients





Clustering: Next steps

- Association of environmental variables
  - Extend to random forests (e.g., thousands of trees)
  - Develop predictive models for novel sites
- Biological characterization of groups
  - Develop trait profiles


Phase I: Prepare Screening Dataset
Phase II: Metric Screening
III: Performance
Establish pool of candidate sites with adequate biology, chemistry and habitat data
3 Options for Combining Metrics (not exclusive)
Threshold Setting Process
Are we ready to develop our scoring tools?
Probability sites (~850)
Targeted sites (~850)
Create GIS layers at multiple scales
Calculate screening metrics
= values of natural and anthropogenic environmental variables
(at multiple spatial scales)
Select screening metrics:
•BPJ
•literature
•prior IBI analyses
•best modeled relationships
•data quality
•universality
Individual Filters
• primary
•“kill switches”
Multi-metric
Combine set of metrics into a composite score
Multivariate
Create synthetic axis of stress: PCA, ordination, etc.
Thresholds based on:
•overall distributions
•literature
•BPJ
•sensitivity analysis
Composite threshold based on:
•distribution of composite score
•sensitivity?
•other?
Is coverage of natural gradients adequate?
Do we require regional adjustments?
Are PSA regions appropriate for gap analysis?
Can we lump sites across regions?
Did we achieve a good balance of Type I and Type II errors?
Where do we need alternate strategies?

A final consideration

Winnowing of the reference pool

- Looks like we can get ~500-650 sites… this level seems to be sufficient to define initial criteria for most streams in the state
- But this number COULD be reduced:
  - Revisions following science panel/stakeholder guidance
  - Removal of sites from the pool as we get more data (reach scale, etc.)
- Want feedback on contingencies in case this happens

Questions for the Advisory Panel (1)

- **Are you satisfied with our metric selection process?**
  - Are we missing significant stressors?
  - Should we include water chemistry (i.e., nutrients)?

- **Which approach (es) do you recommend for combining the reference screening metrics?** (e.g., *single filters, multi-metric, multivariate , kill-switches, combinations?)*

- **What are the most important factors for selecting screening thresholds?**
  - (How) should we adjust the strawman thresholds?

Questions for the Advisory Panel (2)

- **How should we prioritize statewide consistency versus regional flexibility?**
  - Is making regional adjustments (e.g., roads, code21) appropriate?
  - Is it desirable?
  - Ideas for how to deal with low sample size in Interior Chaparral? (e.g., regional adjustments, accept low numbers?)
- **Are our performance measures adequate? Are there better ones for us to try?**
- **Will our reference process be adequate for scoring tool development?**

PHASE IV: What's Next?

Once we receive feedback, we're ready to work through the calculation and testing steps

Verification Steps: aerial imagery screens/ field verification

Final screened set is then ready for developing scoring tools and using in our pilot studies

RCMP Beyond Bio-Objectives

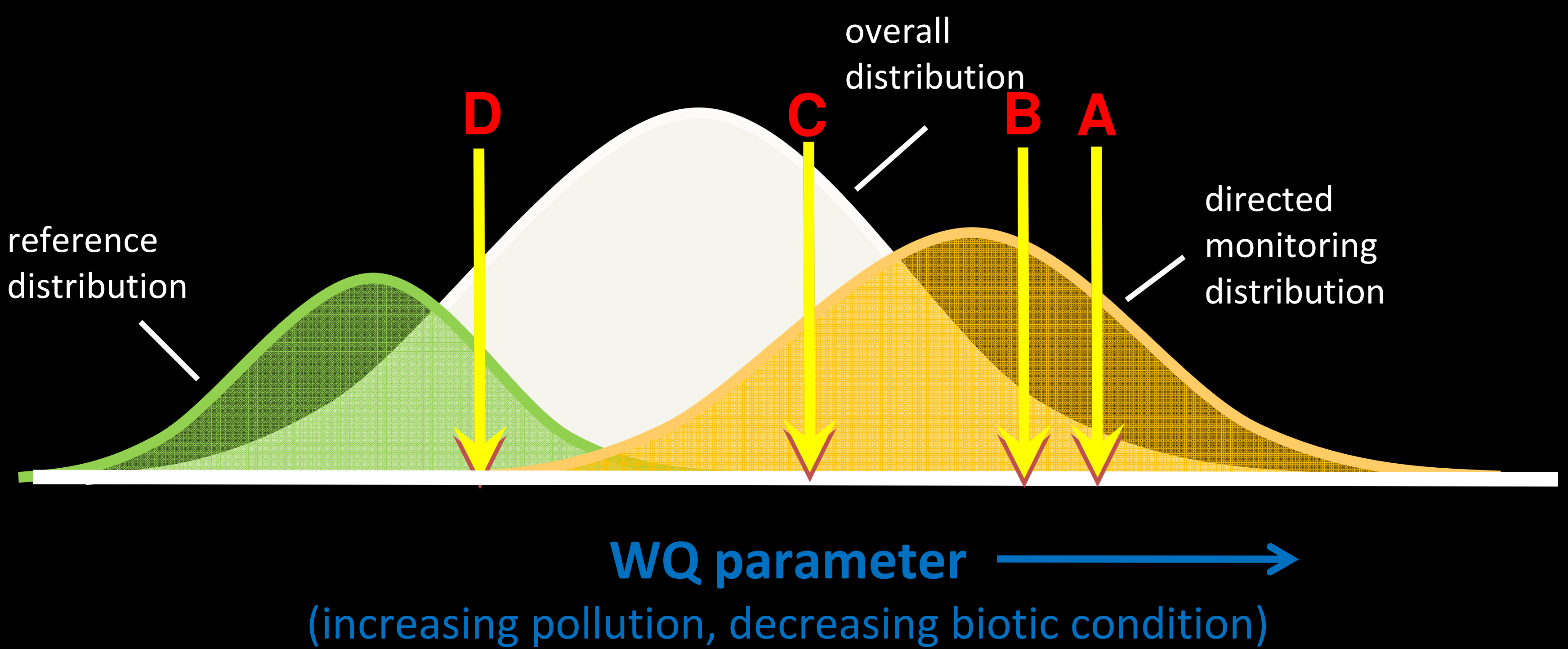
Maintaining the Pool

- Use criteria to find new sites … strengthens ongoing reference condition implementation
- Long-term monitoring

Support for Key Regulatory Applications

- Anti-degradation tool
- Outstanding National Waters
- Setting objective standards for “non-zero” variables

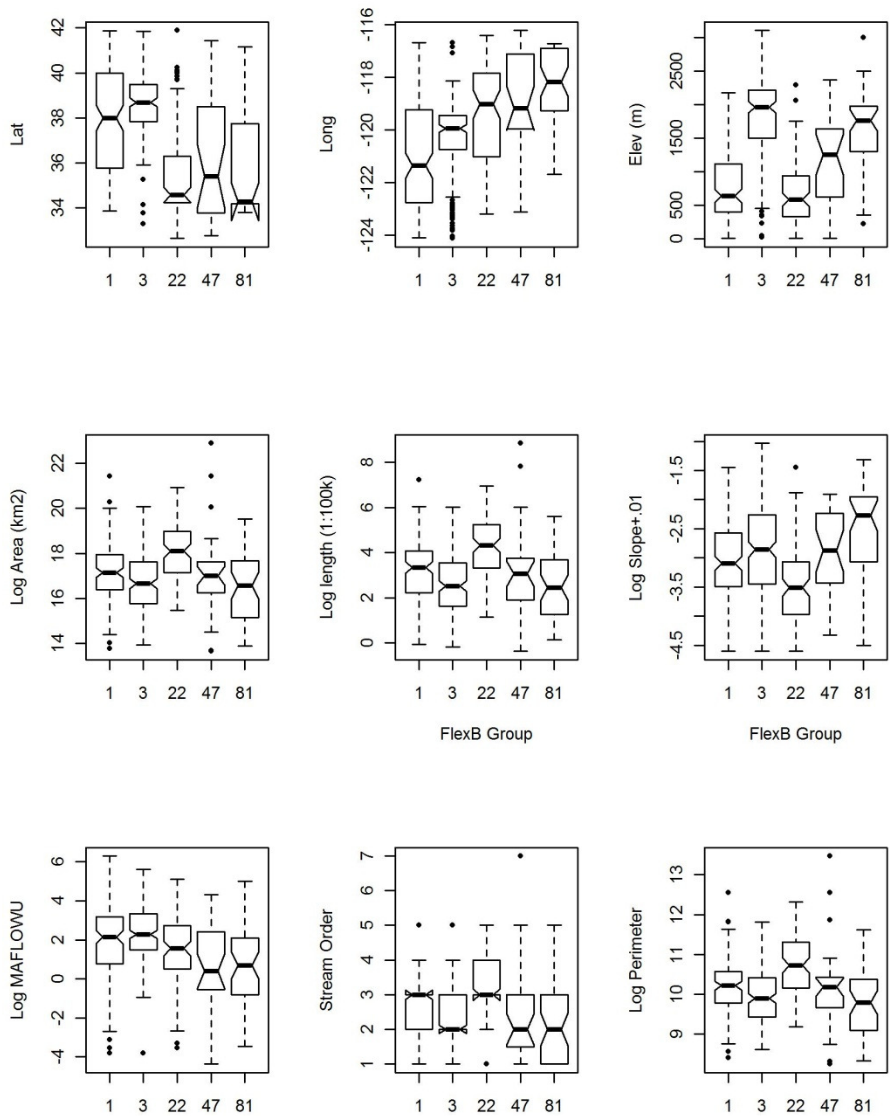
Reference program provides perspective



- Knowledge of the reference distribution can provide objective benchmarks for parameters with non-zero natural values
  - This is especially relevant for ecological endpoints such as bioassessment indicators

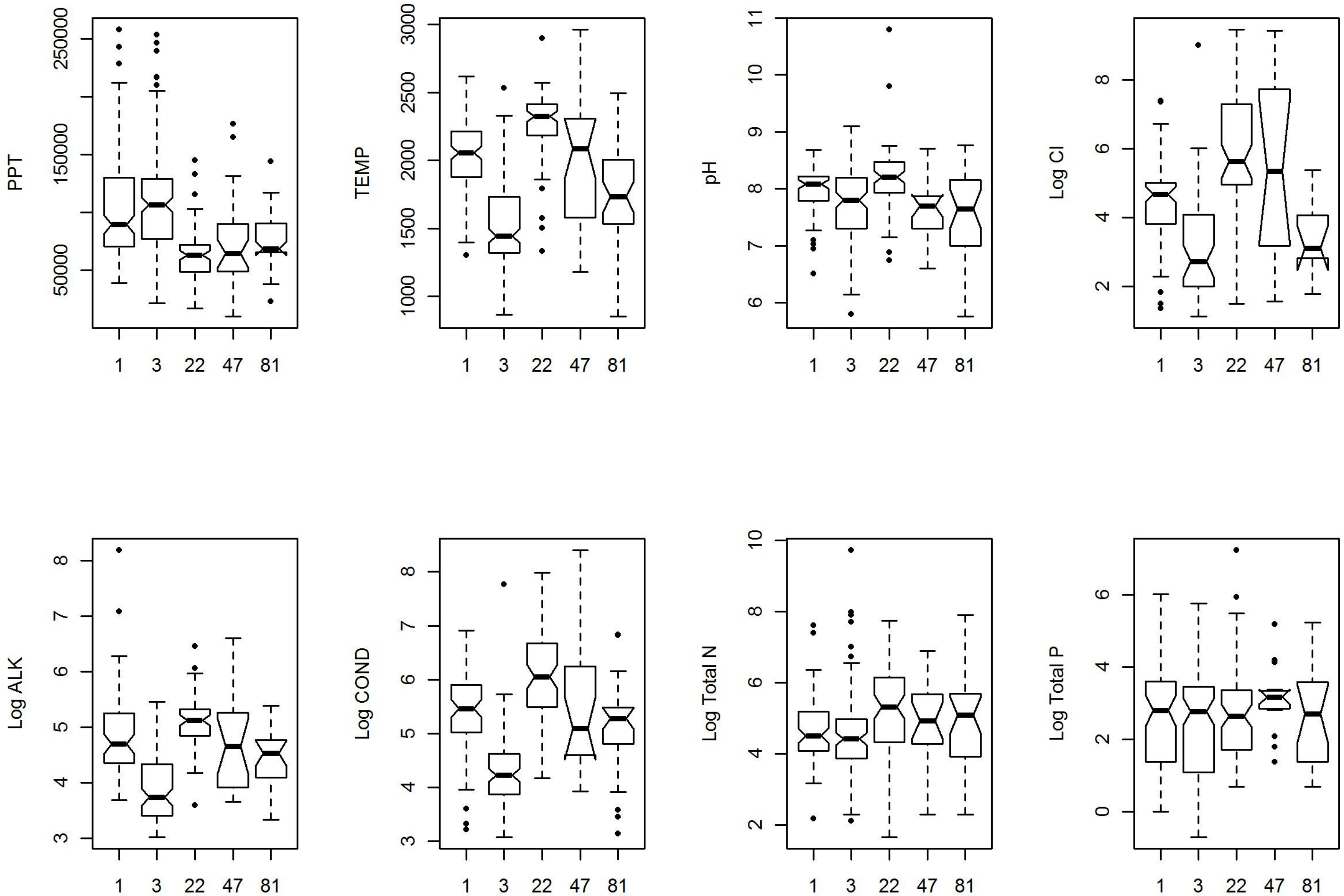
Environmental variability

Geography



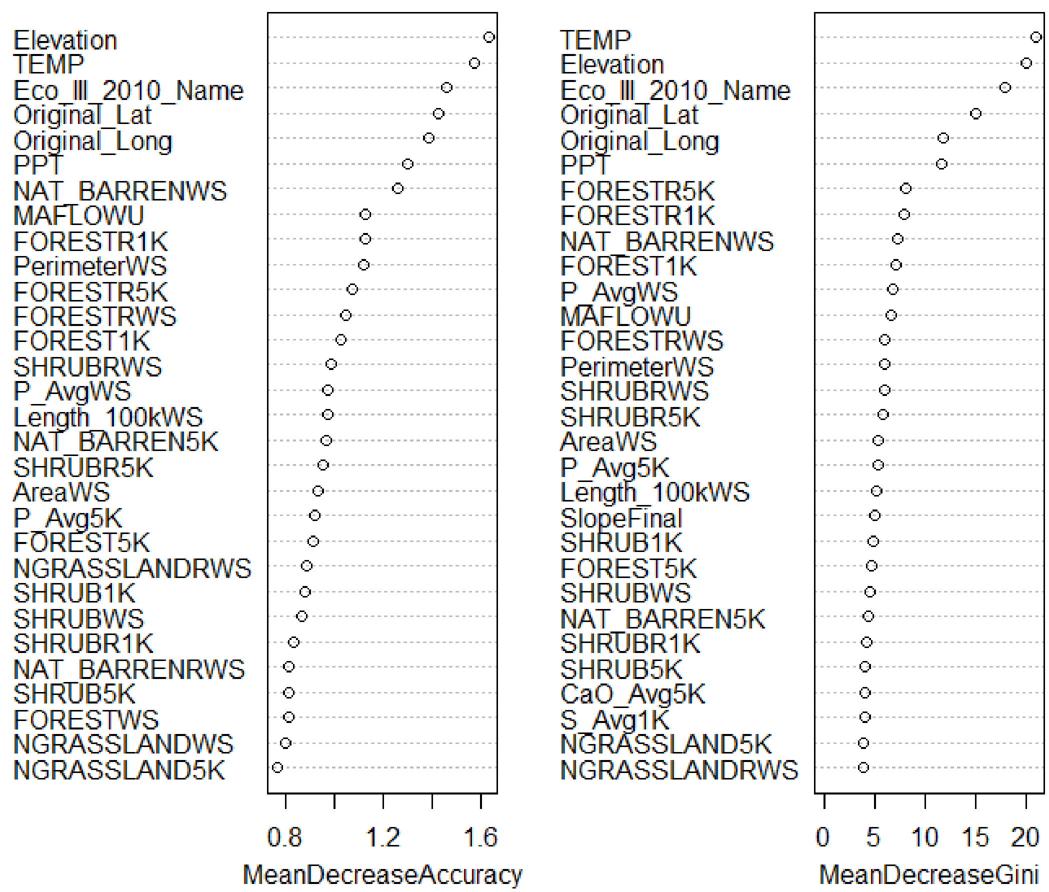
Environmental variability

Climate and WQ

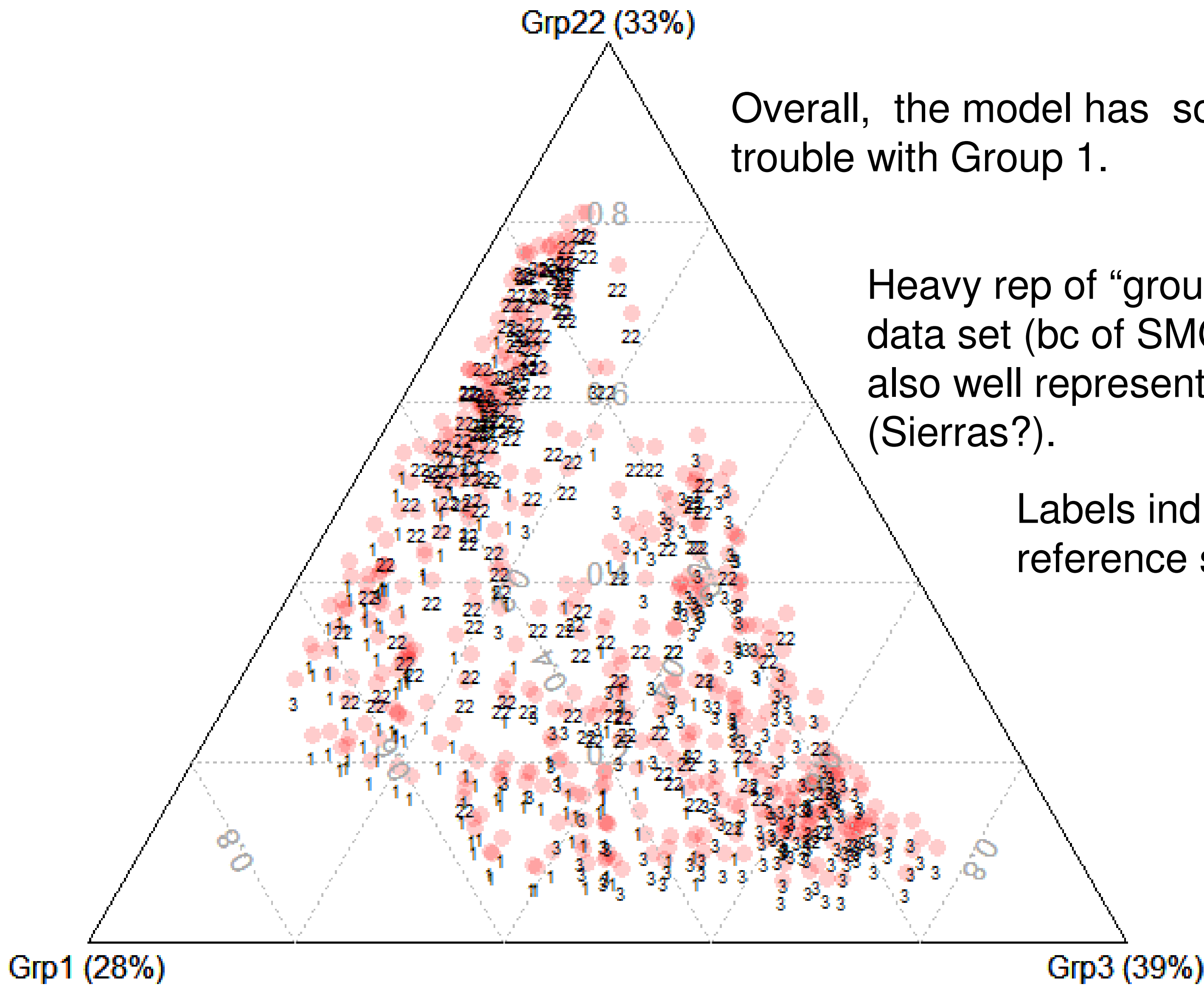


Flex 3: Forest properties (not tree)

- Out-of-bag success rate (=cross validation) : 82%
- Variable importance (not sure if I got the method right):



Predictions of entire data set



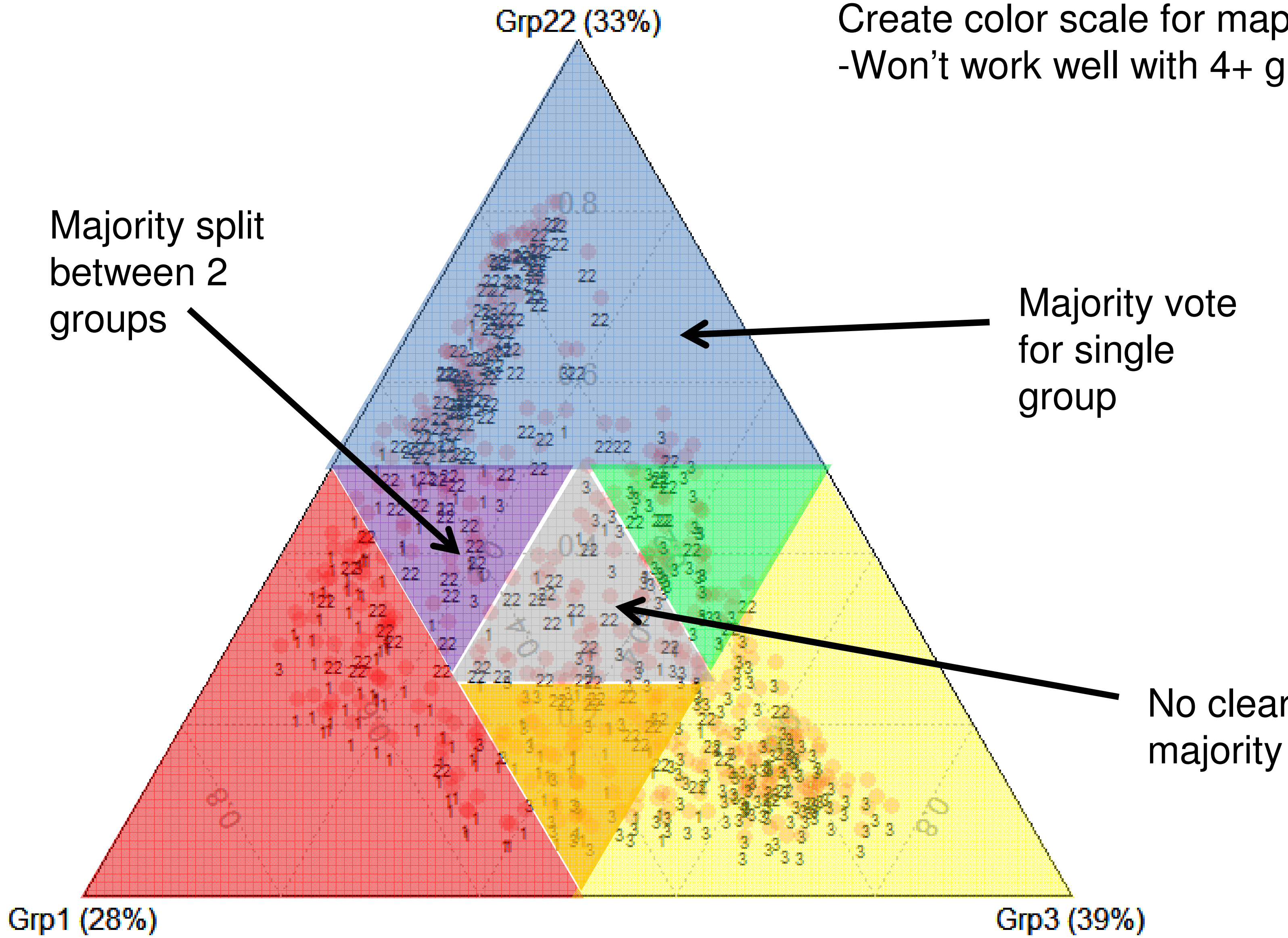
Overall, the model has some trouble with Group 1.

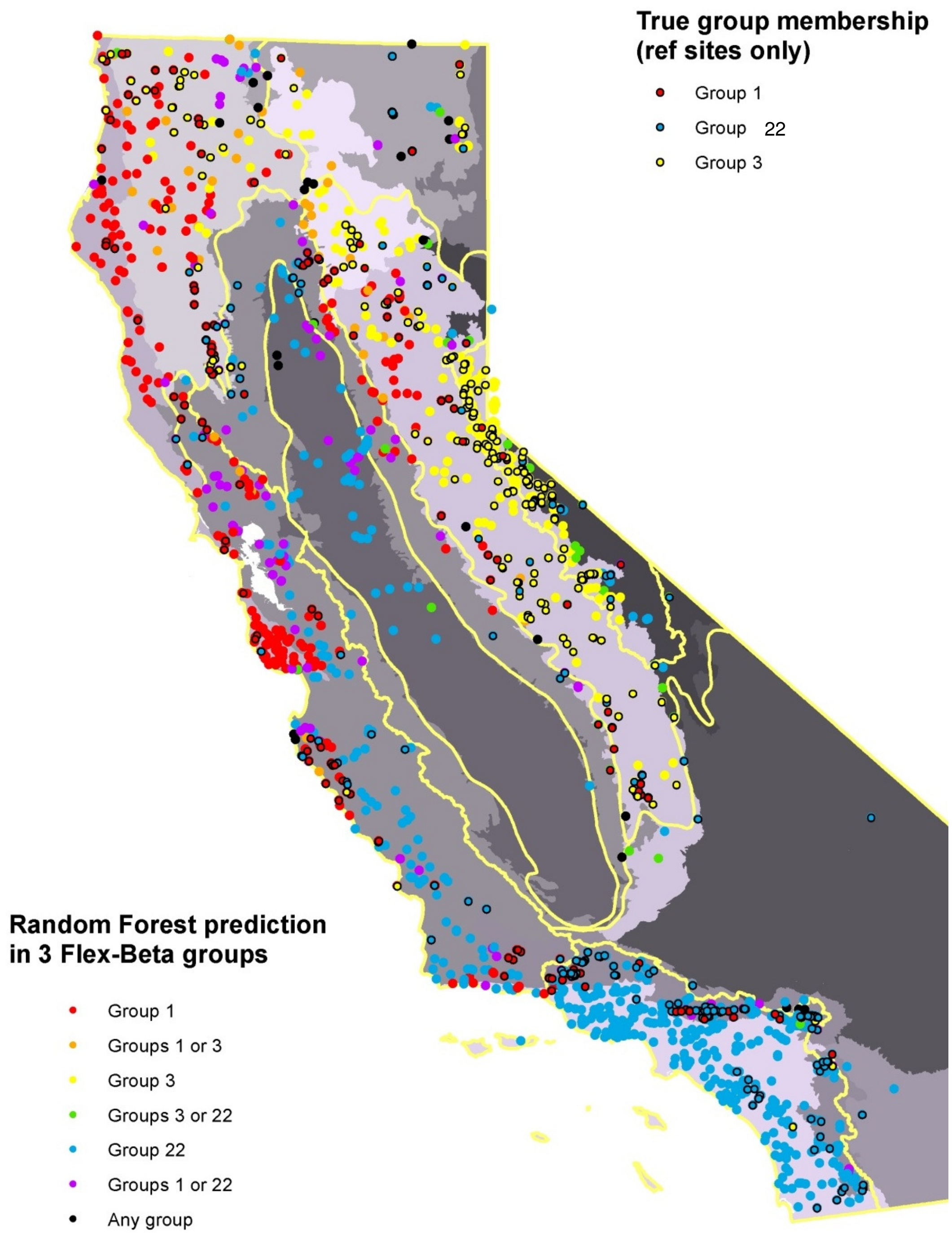
Heavy rep of “group 22” in data set (bc of SMC?). Grp 3 also well represented (Sierras?).

Labels indicate “true” class of reference sites

Predictions of entire data set

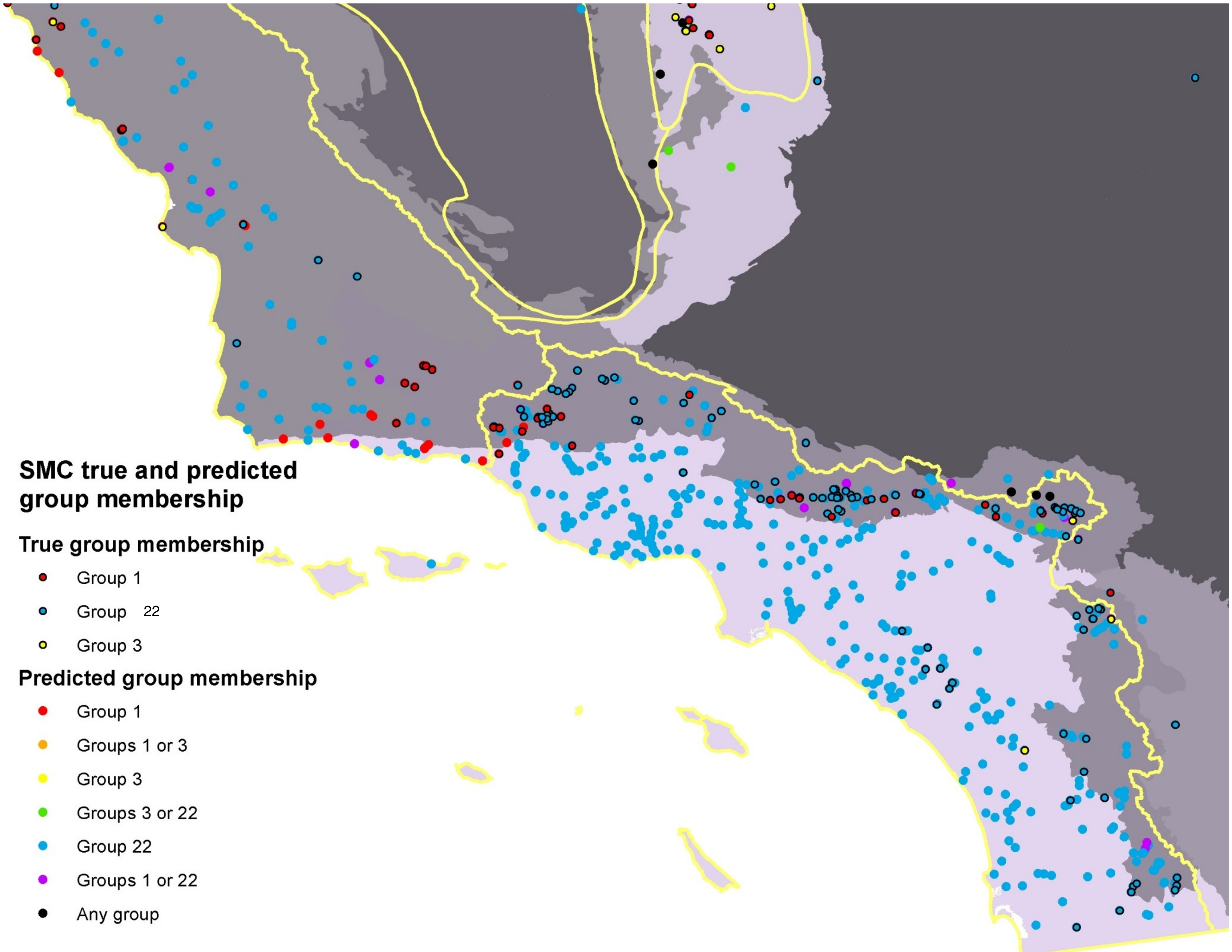
Create color scale for mapping
-Won't work well with 4+ groups!

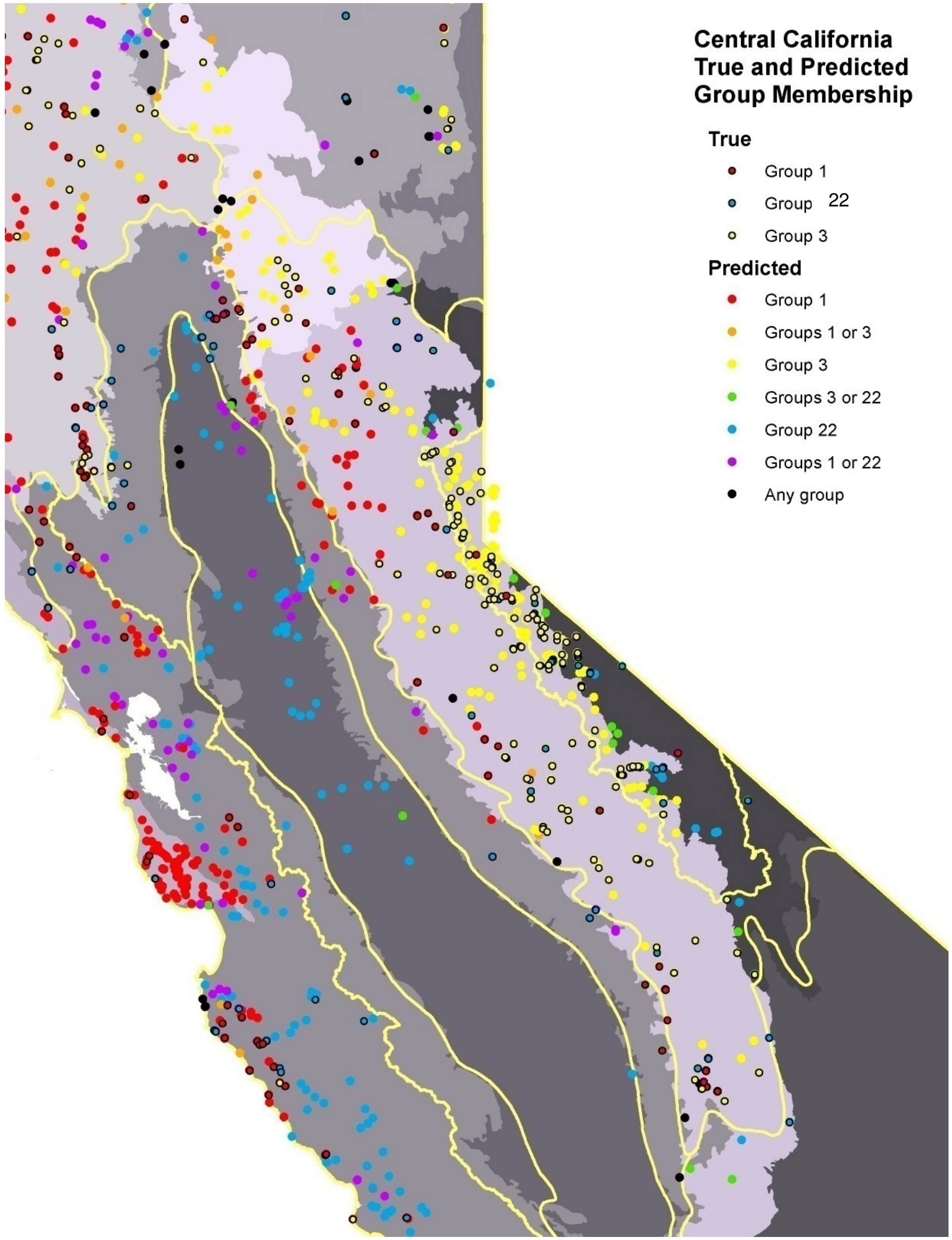


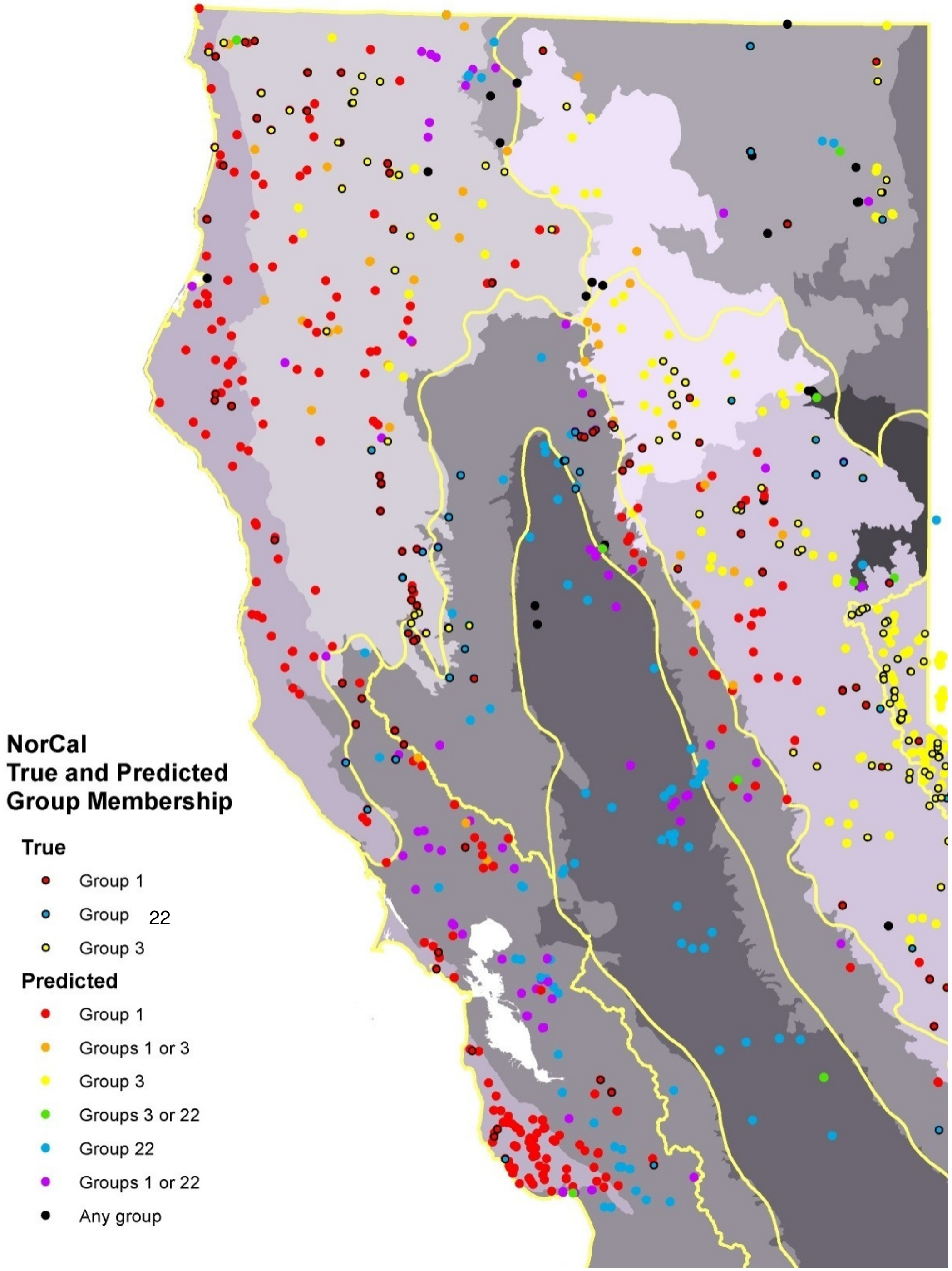


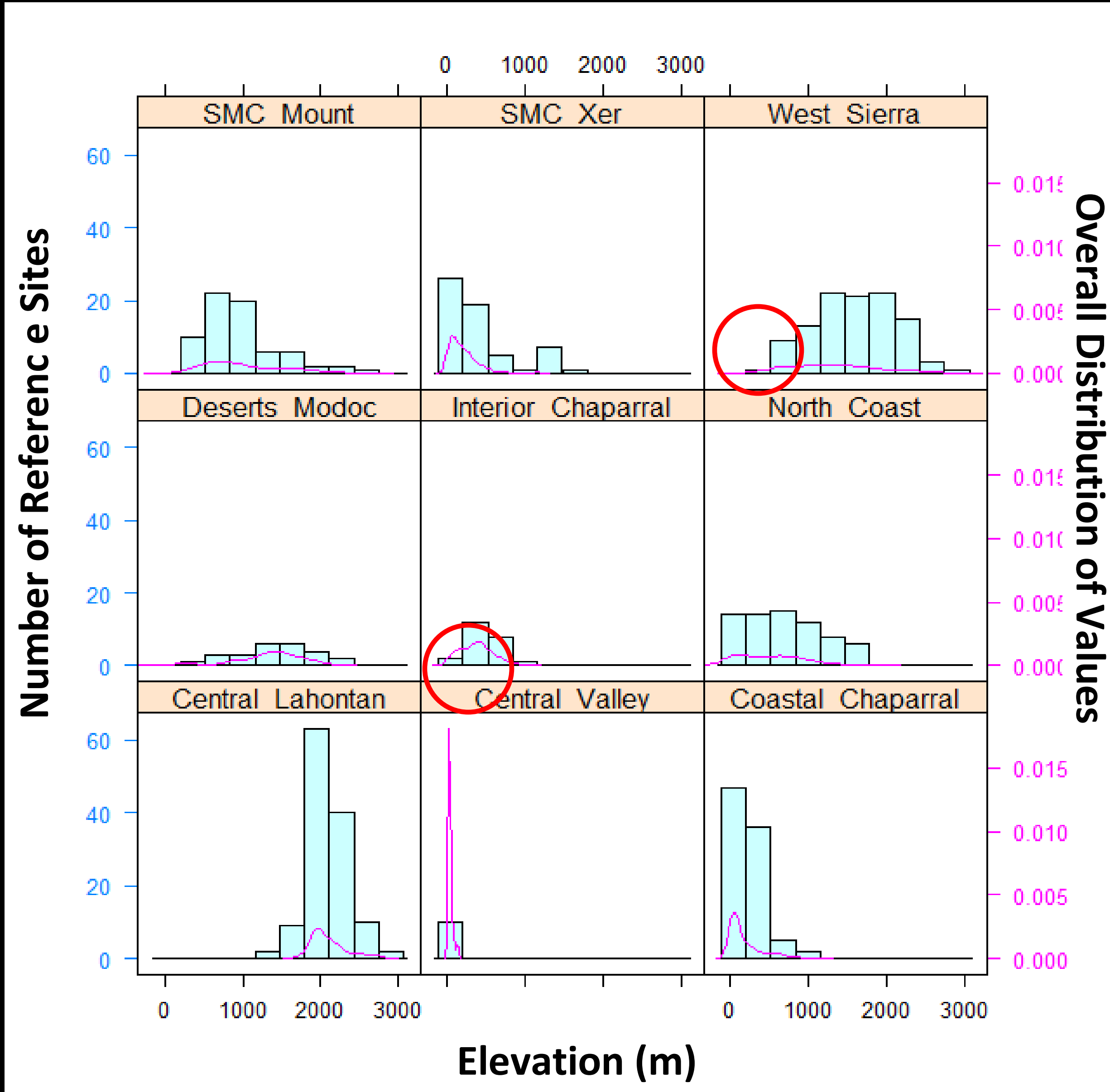
Map of predicted (and true) group membership, based on majority vote by 1000 trees in random forest.

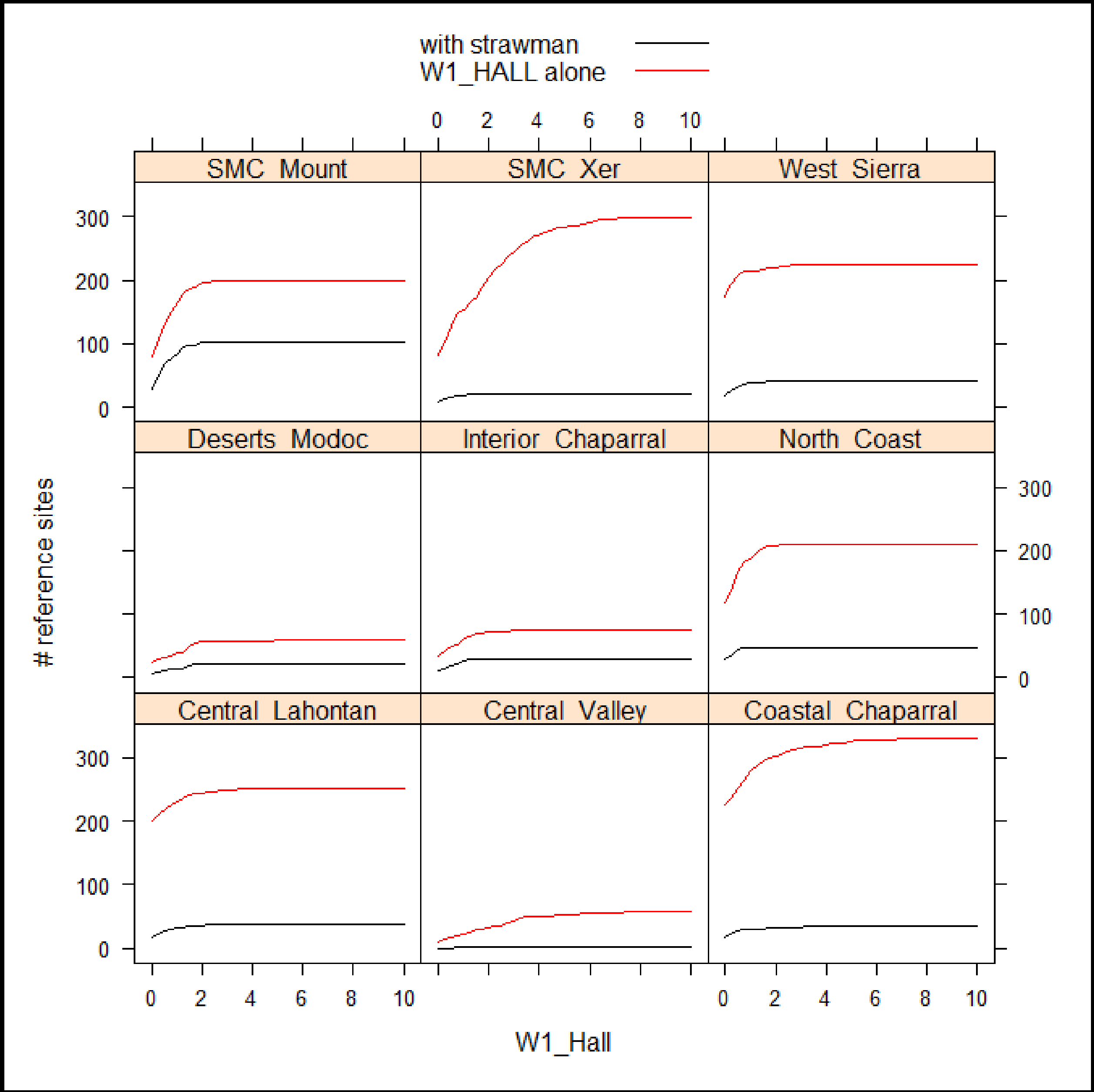
- Some regions dominated by one group:
  - 1: Coast range
  - 3: High Sierras*
  - 22: SMC, Valley, Central Coast
- Lots of "betweeners" in some regions
  - 1/3 in Klamath, Cascades
  - 1/22 in Coastal Chaparral
  - 2/33 in the Sierras.
- Few betweeners in others:
- "Stumpers" are scattered
  - SoCal alpine streams
  - Other clusters? Modoc?


SMC true and predicted group membership
True group membership
Group 1
Group 22
Group 3
Predicted group membership
Group 1
Groups 1 or 3
Group 3
Groups 3 or 22
Group 22
Groups 1 or 22
Any group


Central California
True and Predicted
Group Membership
True
Group 1
Group 22
Group 3
Predicted
Group 1
Groups 1 or 3
Group 3
Groups 3 or 22
Group 22
Groups 1 or 22
Any group


NorCal
True and Predicted
Group Membership
True
Group 1
Group 22
Group 3
Predicted
Group 1
Groups 1 or 3
Group 3
Groups 3 or 22
Group 22
Groups 1 or 22
Any group


SMC Mount
SMC Xer
West Sierra
Deserts Modoc
Interior Chaparral
North Coast
Central Lahontan
Central Valley
Coastal Chaparral
Number of Referenc e Sites
Overall Distribution of Values
Elevation (m)
0
1000
2000
3000
0
20
40
60
0.000
0.005
0.010
0.015


with strawman
W1_HALL alone
SMC Mount
SMC Xer
West Sierra
Deserts Modoc
Interior Chaparral
North Coast
Central Lahontan
Central Valley
Coastal Chaparral
reference sites
W1_Hall
0
100
200
300
0
2
4
6
8
10