## Reference Conditions

Objective:

- Represent all types of perennial streams in California
- Ensure biological integrity at reference sites


## Reference Sites

| REGION | $n$ |
| :---: | :---: |
| North Coast | 79 |
| Central Valley | 1 |
| Coastal Chaparral | 87 |
| Interior Chaparral | 30 |
| South Coast <br> Mountains | 96 |
| South Coast Xeric | 22 |
| Western Sierra | 131 |
| Central Lahontan | 142 |
| Deserts + Modoc | 27 |
| TOTAL | 615 |



## Gradient

## Representation

- Overall excellent representation in most regions
- Central Valley and South Coast (xeric only) very underrepresented
- Very low gradient , large watershed, low elevation settings slightly underrepresented in
Chaparral/ S. Coast


Predicted Conductivity ( $\mu \mathrm{S} / \mathrm{cm}$ )



## Biological Integrity

- "Heatmap" of biological variation related to various stressors in both the reference population (left) and across all sites (right)


## - Anthropogenic

 sources of variation were generally low in the reference poolPer Agriculture 1k Per Agriculture 5k Per Agriculture WS Per Urban 1k Per Urban 5k Per Urban WS Per Ag + Urban 1k Per Ag + Urban 5k Per Ag + Urban WS Per Code 21 1k Per Code 21 5k Per Code 21 WS Per Impervious 1k Per Impervious 5k Per Impervious WS Road density 1k Road density 5k Road density WS Road crossings 1k Road crossings 5k Road crossings WS Per Grazing 1k Per Grazing 5k Per Grazing WS Population density 1k Population density 5k Population density WS Dam Distance Dam Count
Canal Distance Per Canal


$\begin{array}{lll}0.0 & 0.2 & 0.4\end{array}$
$r^{2}$

## Reference Comments

1. Add more technical detail:
"...provide more detail on the process for identifying metrics, thresholds, or kill switches, and the differences among these"
2. Define limits to bio-objective application
"Define how far the traditional approach can be used in reference-poor regions like the Central Valley and South Coast Xeric"
3. Explain how reference data will be used
"It is not clear how reference data will be used in a regulatory context"

## Predictive Models:

(Observed/ Expected Models)
Developed in UK (Wright and others 1970s-1980s, RIvPACS), adapted in Australia (AusRivAS) and US (Chuck Hawkins, Utah State... source of most of these slides)

Species-based approach: Compare number of observed ("O") taxa to number of expected ("E") taxa
"Expected" taxa derived from predictive modeling techniques

## Estimating "E"

## Step 1. Classify 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 class membership for new sitesBiologically Defined Reference Clusters:


| Cluster Natural |
| :---: |
| Predictor Variables: |
| Watershed Area |
| Geology |
| Latitude/ Longitude |
| Elevation |
| Temperature |
| Precipitation |

## 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

| Predictor <br> Variables | Cluster | Site's probability of cluster membership | Frequency of species $X$ (Farula sp.) in cluster | $\begin{aligned} & \text { Expected } \\ & \text { contribution } \\ & \text { to PC } \end{aligned}$ |
| :---: | :---: | :---: | :---: | :---: |
|  | A | 0.5 | 0.6 | 0.30 |
| Predictive | B | 0.4 | 0.2 | 0.08 |
| Model <br> (matches predictors |  | 0.1 | 0.0 | 0.00 |
| with each reference class) | 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 of taxon occurrence probabilities is an estimate of the number of native taxa ( E ) that should be observed (0)

| Taxon | pc | 0 |
| :---: | :---: | :---: |
| Atherix | 0.70 | $*$ |
| Baetis | 0.92 | $*$ |
| Caenis | 0.86 |  |
| Drunella | 0.63 |  |
| Epeorus | 0.51 | $*$ |
| Farula | 0.38 |  |
| Gyrinus | 0.07 |  |
| Hyalella | 0.00 | $*$ |
| E | 4.07 | 3 |

$$
\begin{aligned}
& \mathrm{O} / \mathrm{E}=3 / 4.07 \\
& \mathrm{O} / \mathrm{E}=0.74
\end{aligned}
$$

O/E (scaled 0.0 to 1.0): represents proportion of native assemblage present at test site

## Step 2: Cluster biological similarity

(Bray-Curtis dissimilarity, flexible- $\beta=-0.25$, rare taxa removed if < 2.5\% of sites)


## Scoring Tool Performance Measures

1. Precision
2. Accuracy
3. Responsiveness
4. Repeatability
5. Applicability
