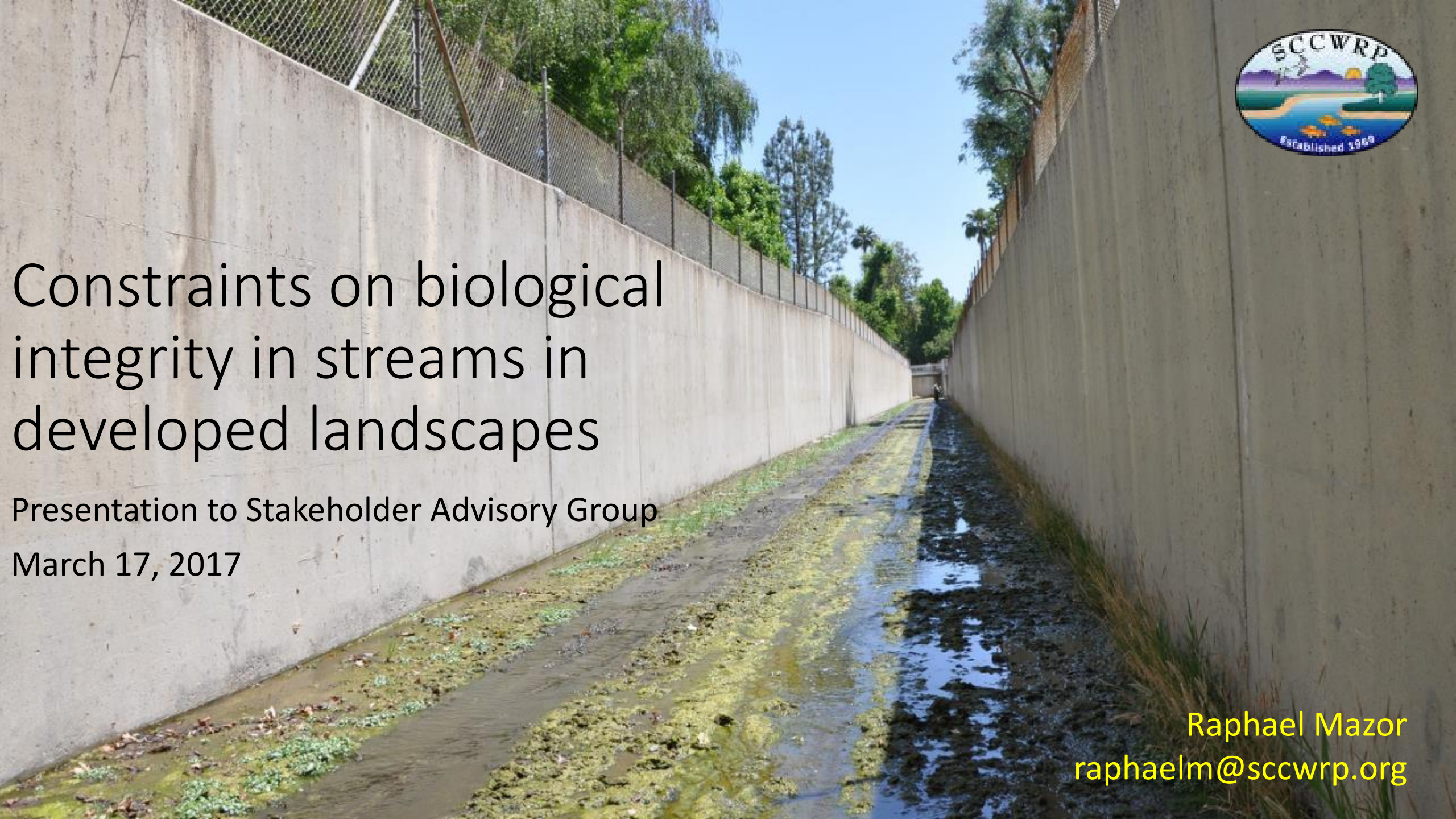


Constraints on biological integrity in streams in developed landscapes

Presentation to Stakeholder Advisory Group

March 17, 2017

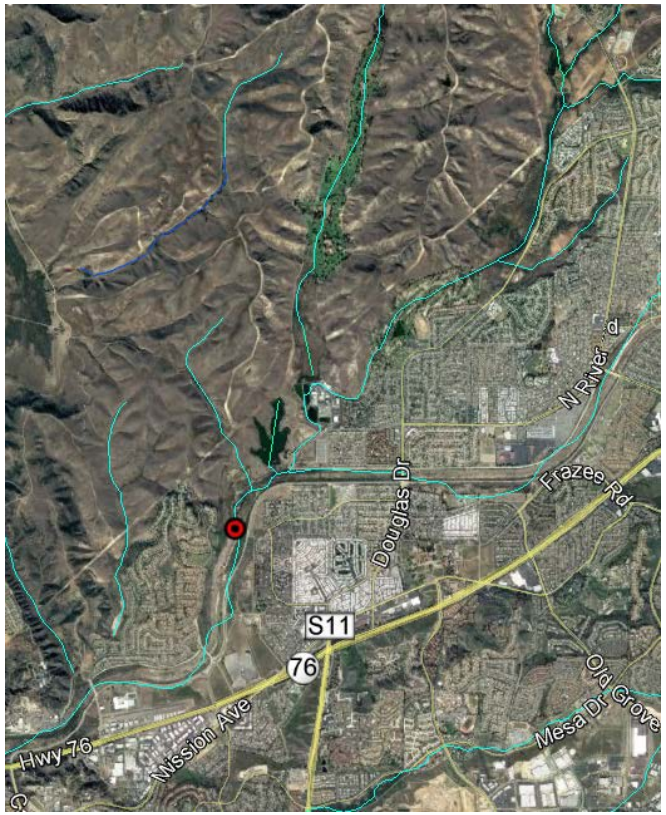
Raphael Mazor
raphaelm@sccwrp.org

Background and goals

- The Water Boards are committed to exploring options for managing streams with constrained biological integrity
  - E.g., different priorities or timeframes for improvements
  - “Alternative thresholds” unlikely
- Management options will be discussed during policy development, but may not be set within this policy.
- We will develop one way of screening streams that may be constrained by landscape development.
  - Statewide screening based on GIS
  - Field visits and other data may also play a role
  - Screening is a starting point, not the final word.

Two ways to identify constrained streams: Channels vs Landscapes

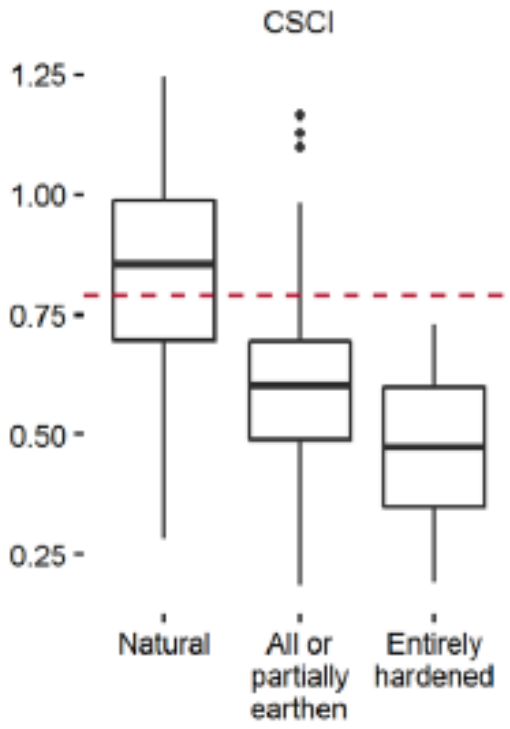
Modified channel

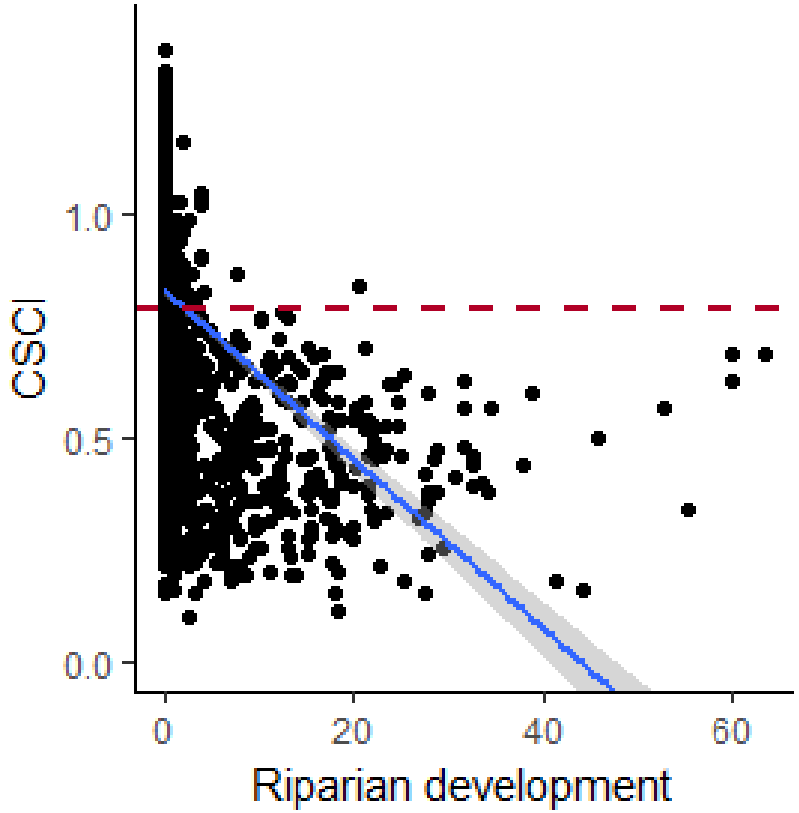


Developed landscape

- Field determination vs. GIS
- Harder to map channel mod
- Channel mod may define the problem too narrowly
- Both approaches have strengths, but landscape approach is better for screening and statewide application

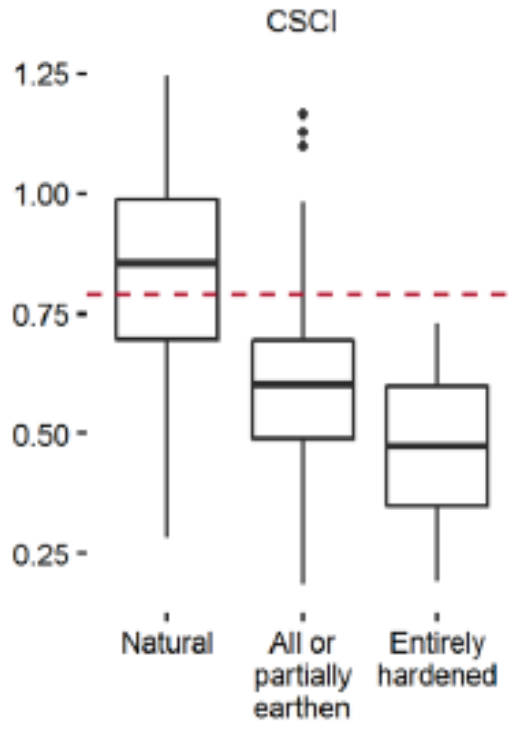
Development can constrain biological integrity

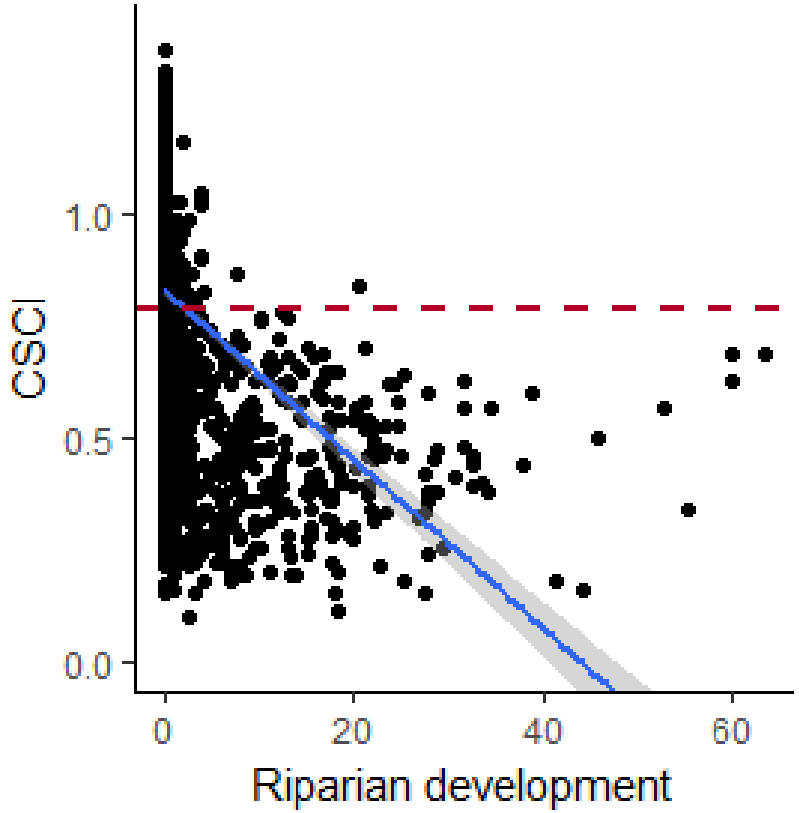


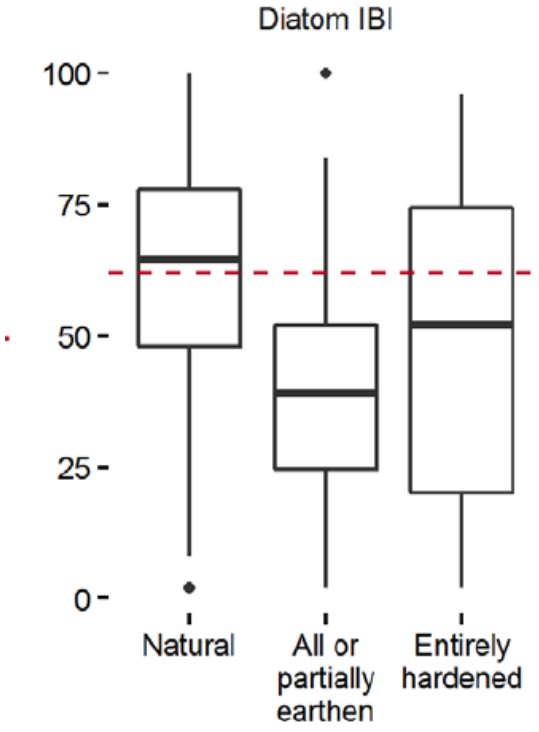


High scores (above threshold) rarely, if ever, seen in certain stream types

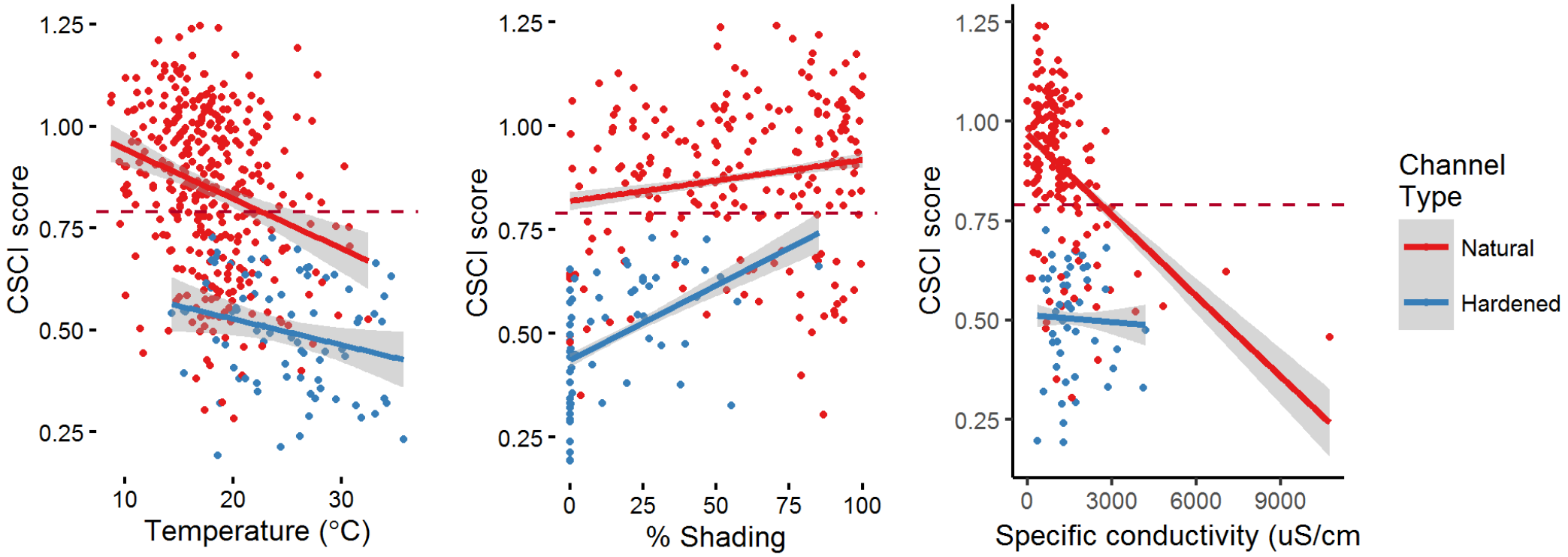
Development can constrain biological integrity (bugs moreso than algae)







Dampened response to WQ gradients



Improving WQ may not protect bio-integrity

Tentative definition of developed landscapes

Landscapes where developed land uses are likely to limit CSCI scores

(…and ASCI scores)

Approach

- Build a model to predict ranges of CSCI scores associated with land use gradients
  - Select land use parameters (e.g., urban or ag land cover)
  - Use national STREAMCAT database of watershed characteristics: Easy statewide applicability
  - Quantile random forest: Provides *range* of likely CSCI scores in different landscapes
- Identify landscapes where statewide “default” assessment endpoints are *unlikely* to be met

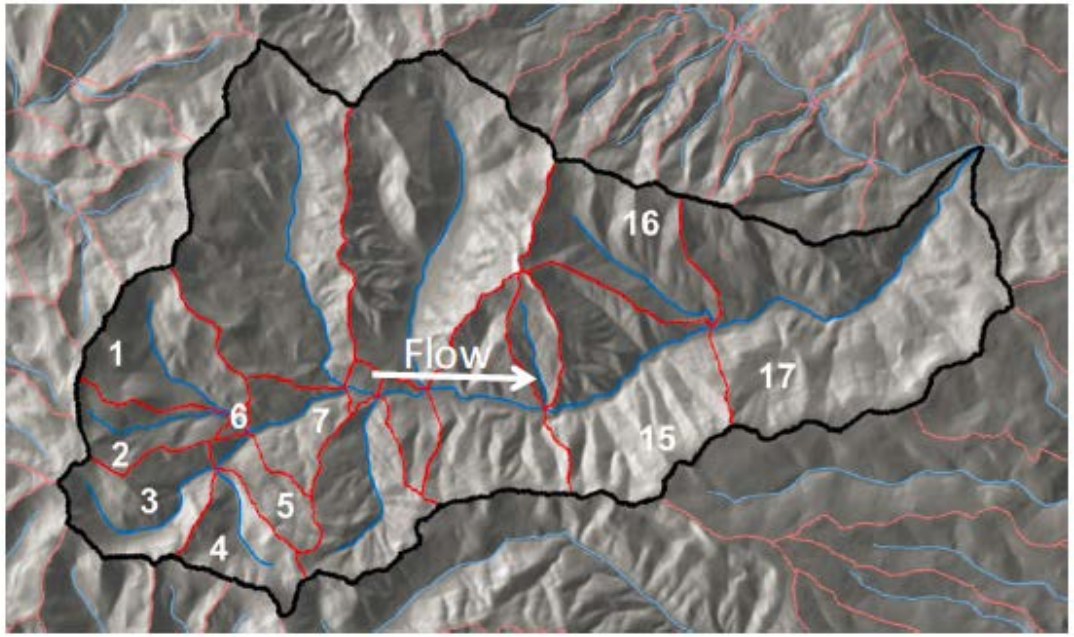
Three key factors in modeling decisions

1. Model development: What kinds of variables should we include?
2. Model application: What thresholds to use for identifying likely “high” or “low” scoring streams?
3. Model application: What likelihoods for defining “likely” or “unlikely”?

- Tech team is evaluating these decisions with Regulatory Advisory Group on an iterative basis
- We’d like to provide you with impacts of these decisions so you can provide feedback to the Water Board

Predictor data source: STREAMCAT

- Nearly all stream segments from NHD+ (1:100k scale) represented
- Lots of data calculated for each watershed and catchment
  - Metrics also calculated for 100-m riparian buffers
- STREAMCAT makes it easy to explore statewide landscape models on a large scale



Types of data in STREAMCAT

- Natural variables (e.g., geology, climate, watershed area)
  - These DON'T affect CSCI scores! No need to include in models.
- Stressor variables
  - These DO affect CSCI scores
  - Some reflect transient impacts (e.g., pesticide)
  - Some reflect long-term impacts (e.g., landcover)
  - Some are debatable, especially in rural settings (e.g., roads, dams, imperviousness, mines)
- Different variables are good for different models and applications
  - Identify landscape-constraints? Only long-term stressors
  - Identify likely high-scoring streams? Long-term and transient stressors are appropriate
- Tech team has preliminary classifications, currently being vetted with Regulatory Advisory Group

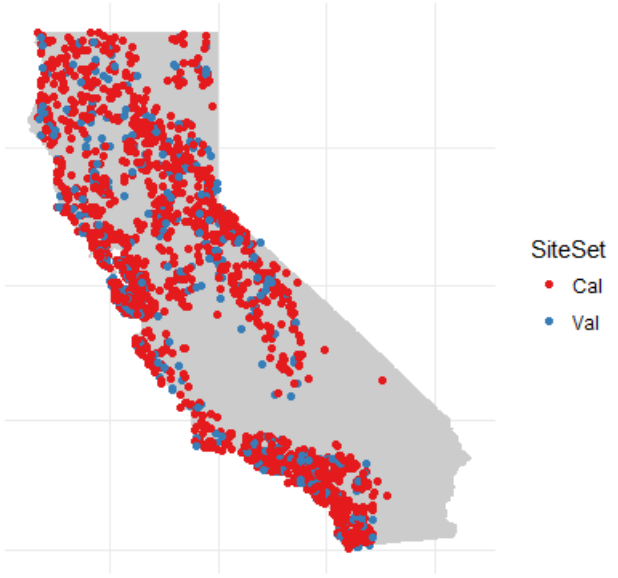
Channelization/Armoring

- Poorly characterized in STREAMCAT (NHD sources are not very reliable)
- Statewide, NHD-registered data not available
- Many armored streams are captured by other variables (e.g., riparian landcover)
- May be better addressed after landscape-scale screening with field data (e.g., SMC-type analyes)

Building the models

Preliminary work:

- 3252 sites, split 80% calibration 20% validation
  - Stratified by 6 PSA regions
  - Each region further stratified into thirds by imperviousness
- Where multiple samples are available, only one selected for modeling
- “Kitchen sink” models, for exploratory purposes:
  - All relevant predictors (next slide) included at multiple spatial scales.
  - These are overfit models! Refinement, validation is the next step.



PSA6	Bottom third	Top third
CH	0.14	2.03
CV	0.55	9.54
DM	0.07	0.17
NC	0.04	0.11
SC	0.29	6.41
SN	0.07	0.22

Types of variables we may include in models

Simple	Moderate	Complex
Urban land cover (NLCD 2011) Ag land cover (NLCD 2011) Canal density (NHD+)	All CDLmin variables Mine density Dam density and storage Road density Road crossings	All CDL and CDLmin variables Impervious surfaces (NLCD 2006) Fertilizer applications Pesticide applications (1997) Non-native veg cover Forest loss Fire perimeters Aerial deposition of N, S
Just a few “permanent” stressors. Best for identifying constraints?	Includes some “debatable” stressors.	Includes “transient” stressors. Best for predicting CSCI scores?

In general…

- Not a big difference among models
- All models performed similarly well (pseudo-R^2 was always ~0.6)
- Variables that occur in rural areas (e.g., low-density urban, ag, road density, non-native veg cover) are more influential than variables that are restricted to heavily developed areas (e.g., high-density urban)

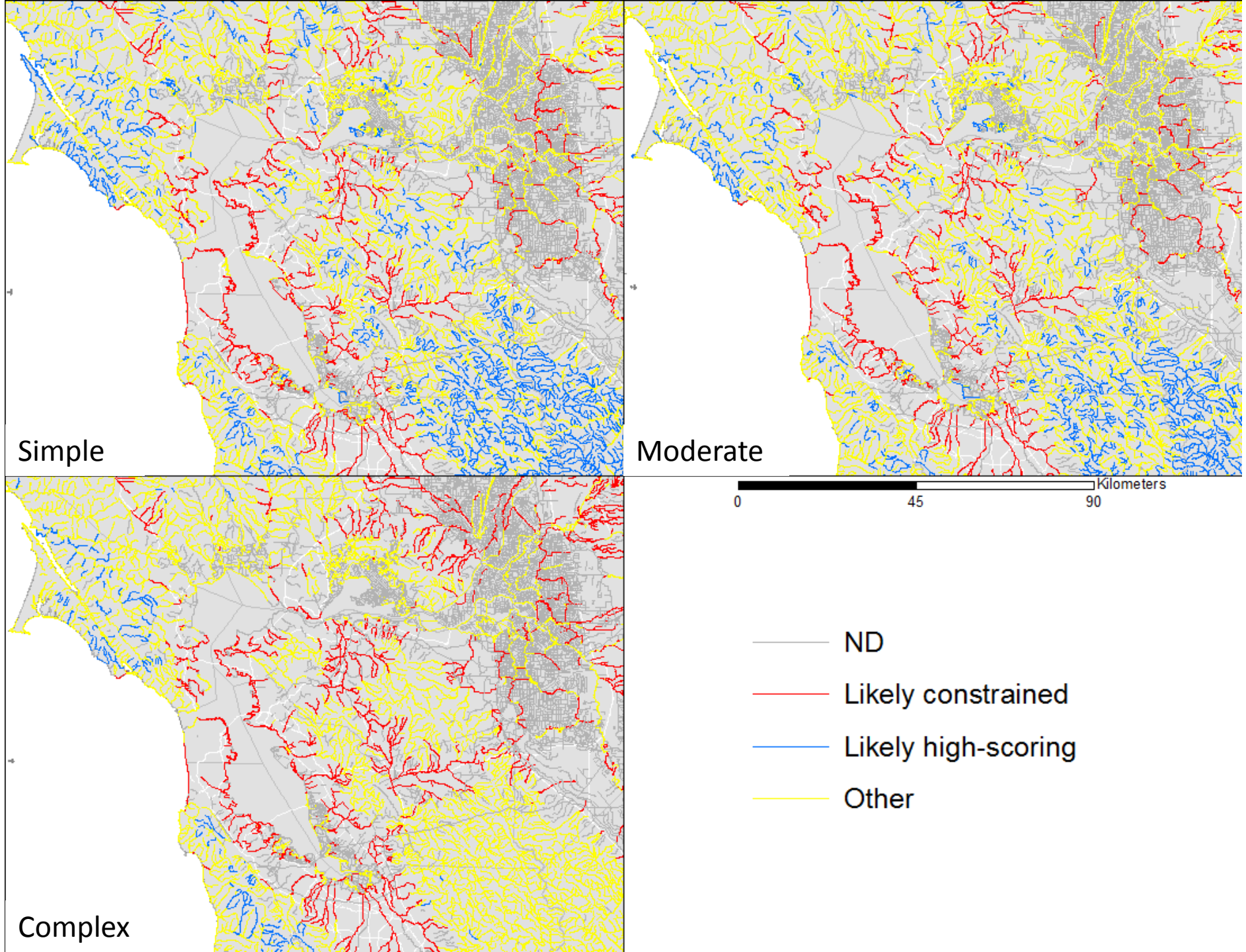
What are the outcomes of these models?

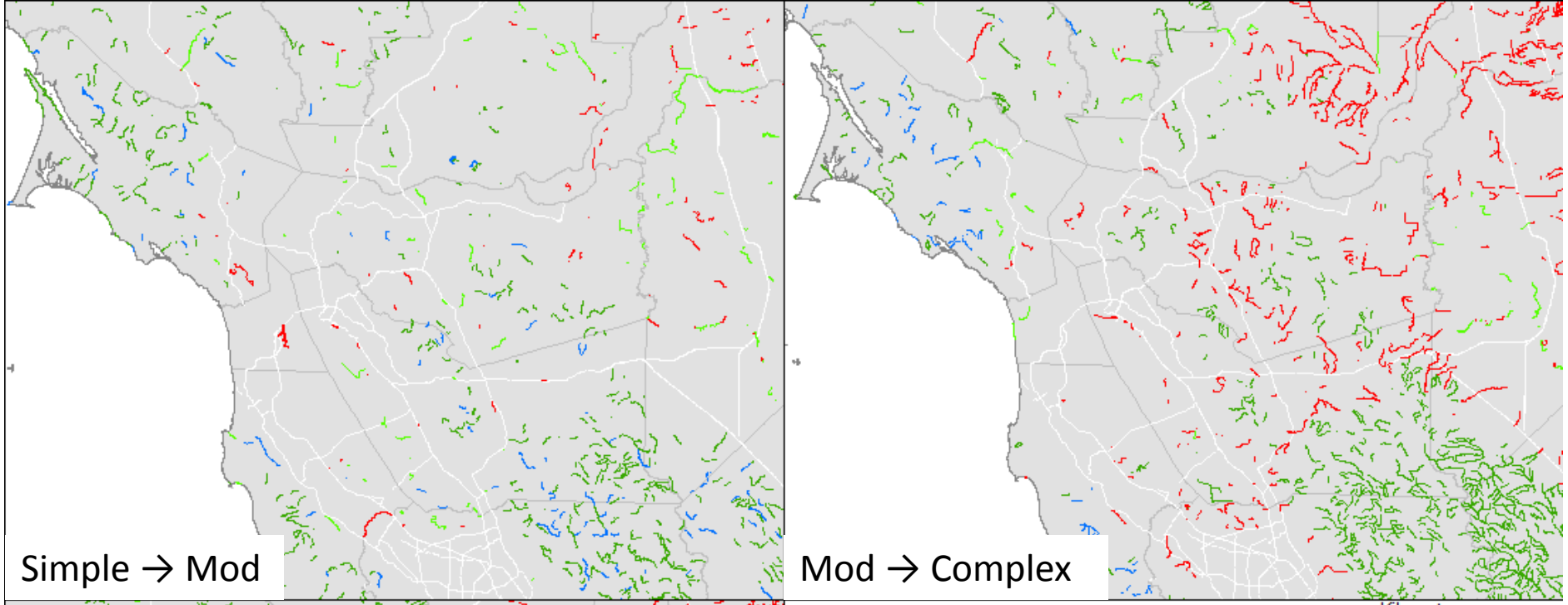
- Maps show three models, with a simple classification scheme:
  - Likely constrained: <10% chance of scores over 0.79
  - Likely high-scoring: <10% chance of scores under 0.79
  - Other: All other streams where data were sufficient to run the model
  - ND: Insufficient data in STREAMCAT to run the model
- Maps are based on one model, but you can use multiple models:
  - E.g., simple models for constrained streams, complex models for high-scoring streams
- While more complex models identify the greatest number of constrained streams/lowest number of high-scoring streams, this can be changed with different classification schemes.

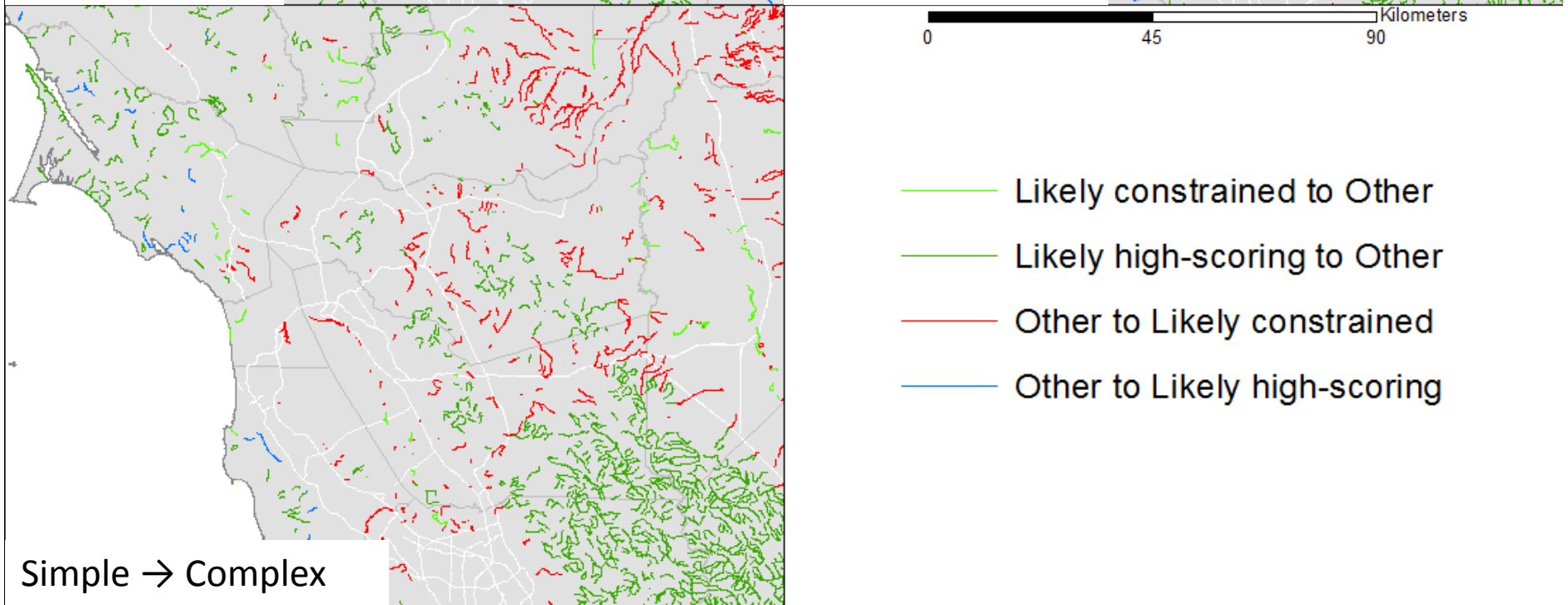
We want a classification scheme that reflects our assumptions/values, not one that produces the map we like best

Example maps

- Maps showing 3 classifications for the Bay Area
    1. Likely low-scoring (constrained)
    2. Likely high-scoring
    3. Other
    4. Not determined
- Maps showing disagreements among models in the Bay Area
    - Simpler model vs more complex model
        1. Likely constrained to other
        2. Likely high scoring to other
        3. Other to likely constrained
        4. Other to likely high scoring


Simple
Moderate
Complex
Kilometers
0
45
90
ND
Likely constrained
Likely high-scoring
Other


Simple → Mod
Mod → Complex
Kilometers
0
45
90


Simple → Complex
Likely constrained to Other
Likely high-scoring to Other
Other to Likely constrained
Other to Likely high-scoring

Next steps

- Vet plan with science panel (4/19)
- Refine and validate models (now through May)
  - Simplify and test models with validation data
  - Repeat with ASCI (Late Summer)
- Produce and distribute maps/data (May)
- Discuss outcomes with advisory groups (Summer)
- Produce report (Late Summer/Fall)

Three key factors in modeling decisions

1. Model development: What kinds of variables should we include?
2. Model application: What thresholds to use for identifying likely “high” or “low” scoring streams?
3. Model application: What likelihoods for defining “likely” or “unlikely”?

- Tech team is evaluating these decisions with Regulatory Advisory Group on an iterative basis
- We’d like to provide you with impacts of these decisions so you can provide feedback to the Water Board

Questions?