EMAP SYMPOSIUM- 2004
Theme 2 – Session 2:
RELATING FINDINGS FROM 305(b) TO 303(d)

Session Chair
Scott Urquhart
Department of Statistics
Colorado State University

DESIGN-BASED APPROACHES FOR RELATING FINDINGS FROM 305(b) TO 303(d)

- What is "Design Based"?
  - ➔ Statistical inferences rest on the probability structure incorporated in the selection of the sample
- Design Based, but Model Assisted
  - ➔ Design based, but incorporate models, like various kinds of regression models
- Model Based
  - ➔ Statistical inference rests on assumed models
    - *Perhaps well defended*

DAN McKENZIE:
ONE OF THE THEME ORGANIZERS
IS NOT HERE

- **I Knew What He planned to Say in His Intro to the Theme Yesterday**
  - ➔ **I had planned to build on what he said!**
  - ➔ **It is hard to build without a foundation!**
  - ➔ **So I'm going present five of his slides.**

EMAP's Guiding Figure

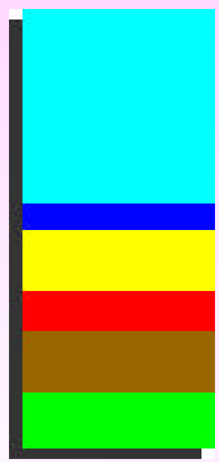
Status & Association Questions

Status

Associations

Extent
of
Resource

(number,
length,
area)



Good
%

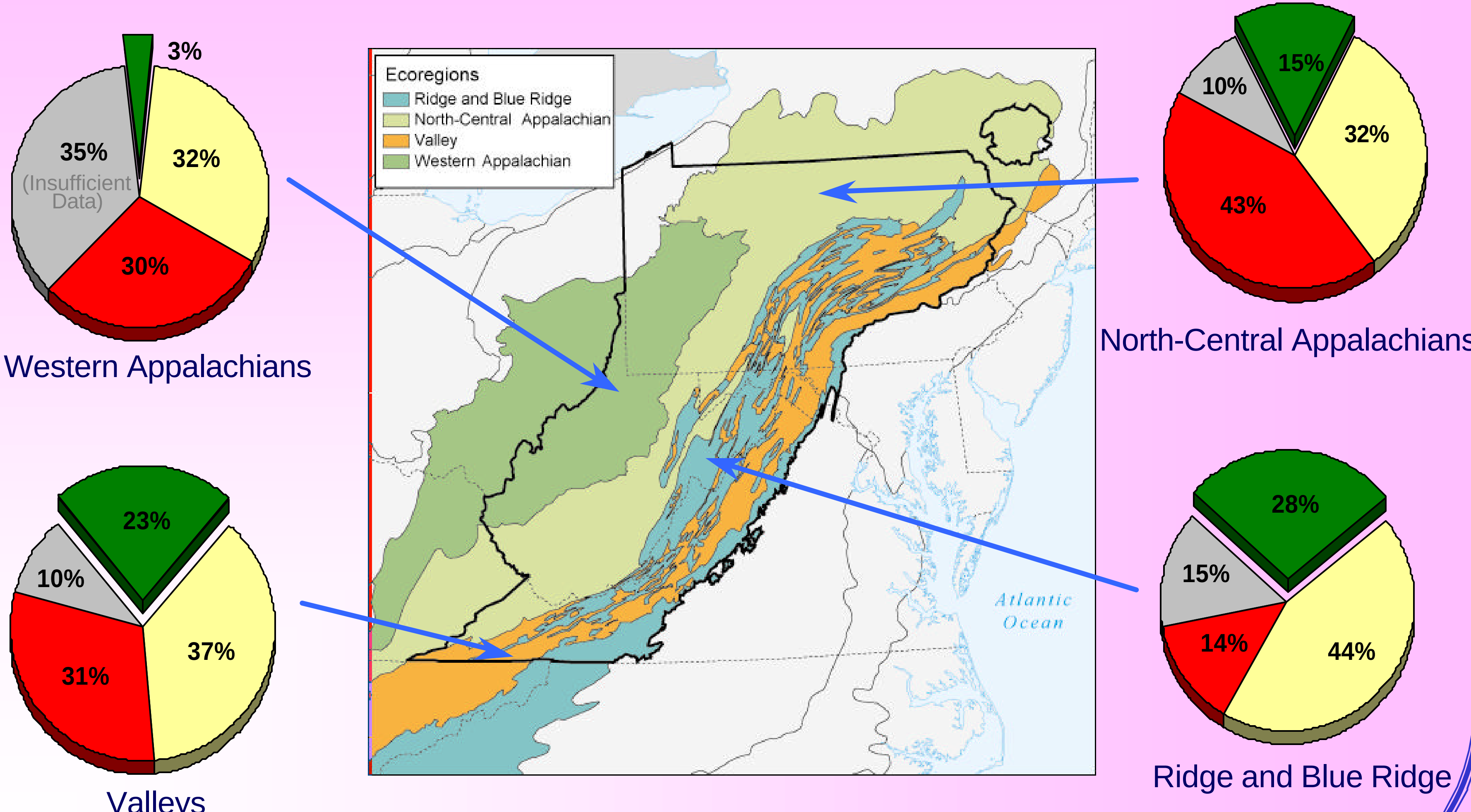
Poor
%

- Nominal
- Unknown
- Acidity
- Toxicity
- Eutrophication
- Habitat

FROM DAN McKENZIE – LABELED "DRAFT – DRAFT"

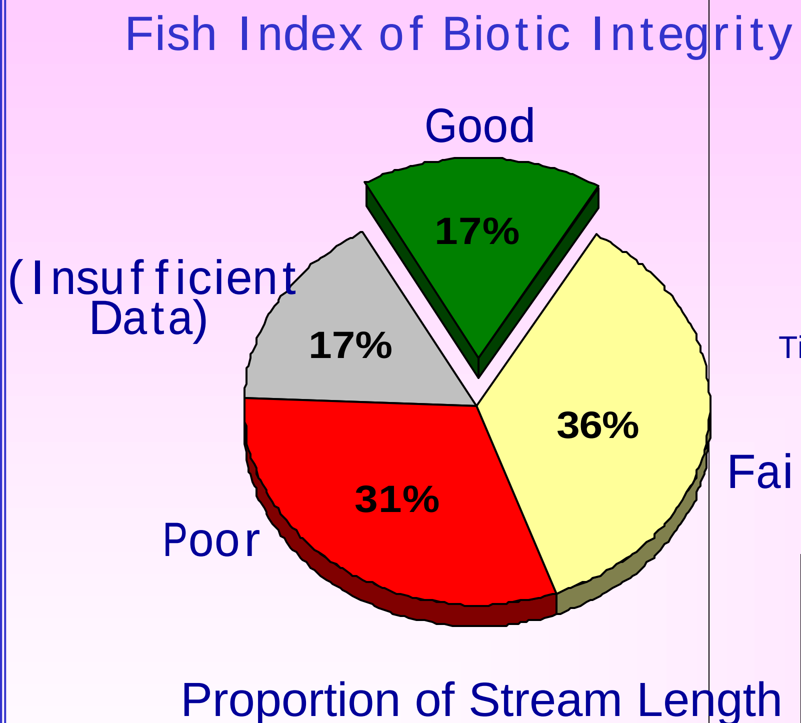
Geographic Targeting

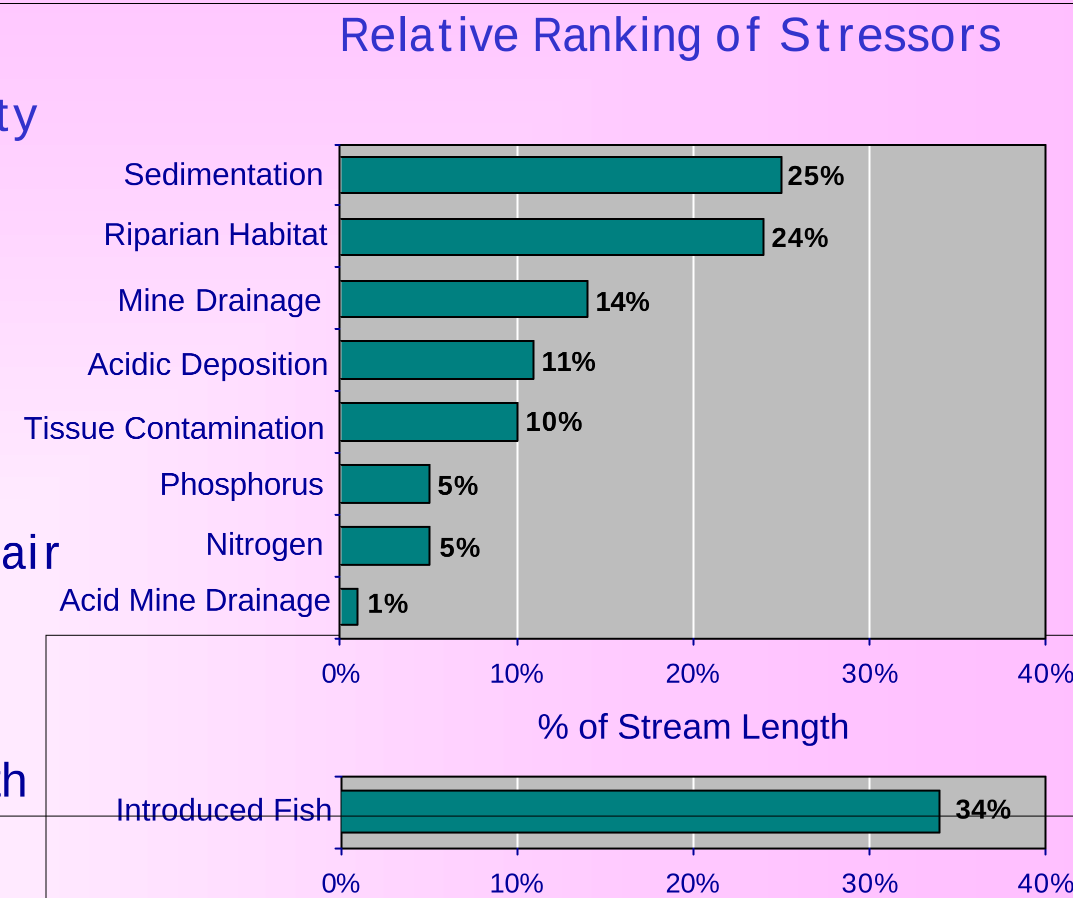
Where does Fish IBI suggest problems?



FROM DAN McKENZIE – LABELED "DRAFT – DRAFT"

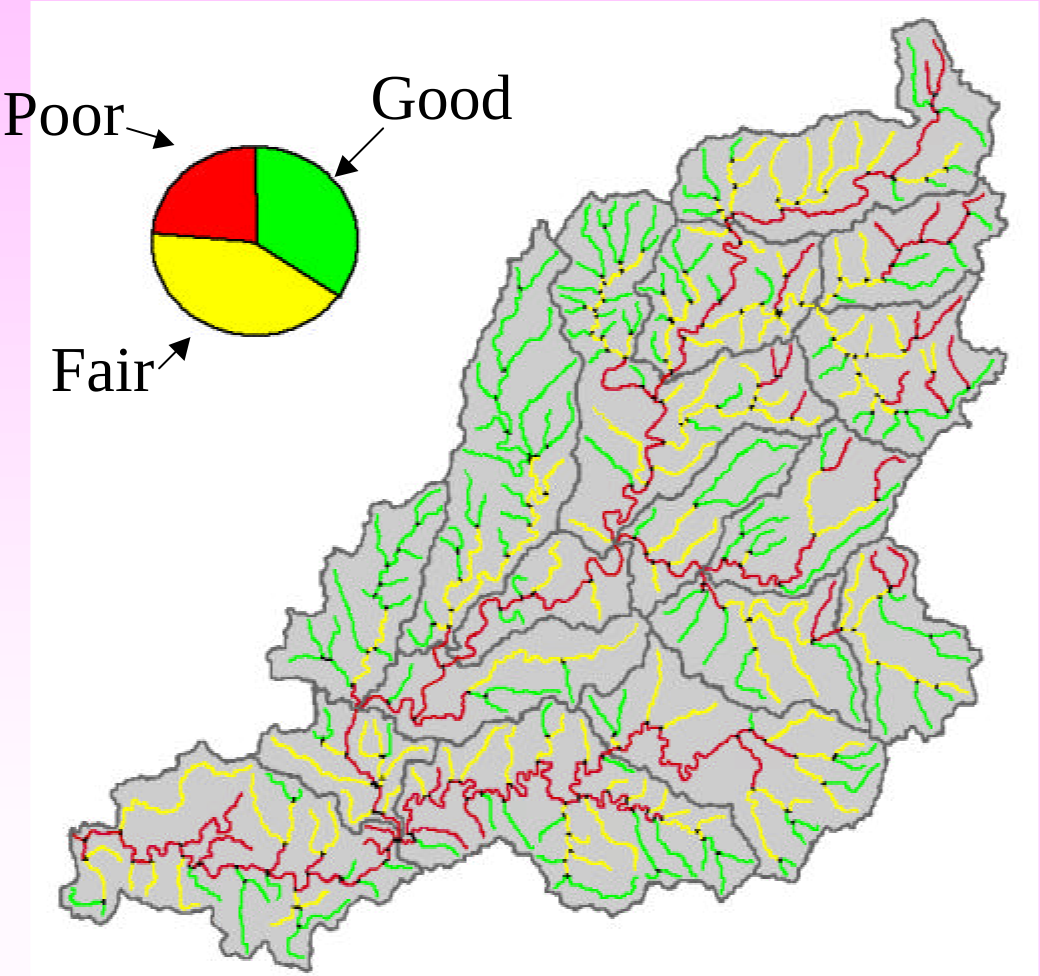
EMAP Probability Survey Example Results (complex)





FROM DAN McKENZIE – LABELED "DRAFT – DRAFT"

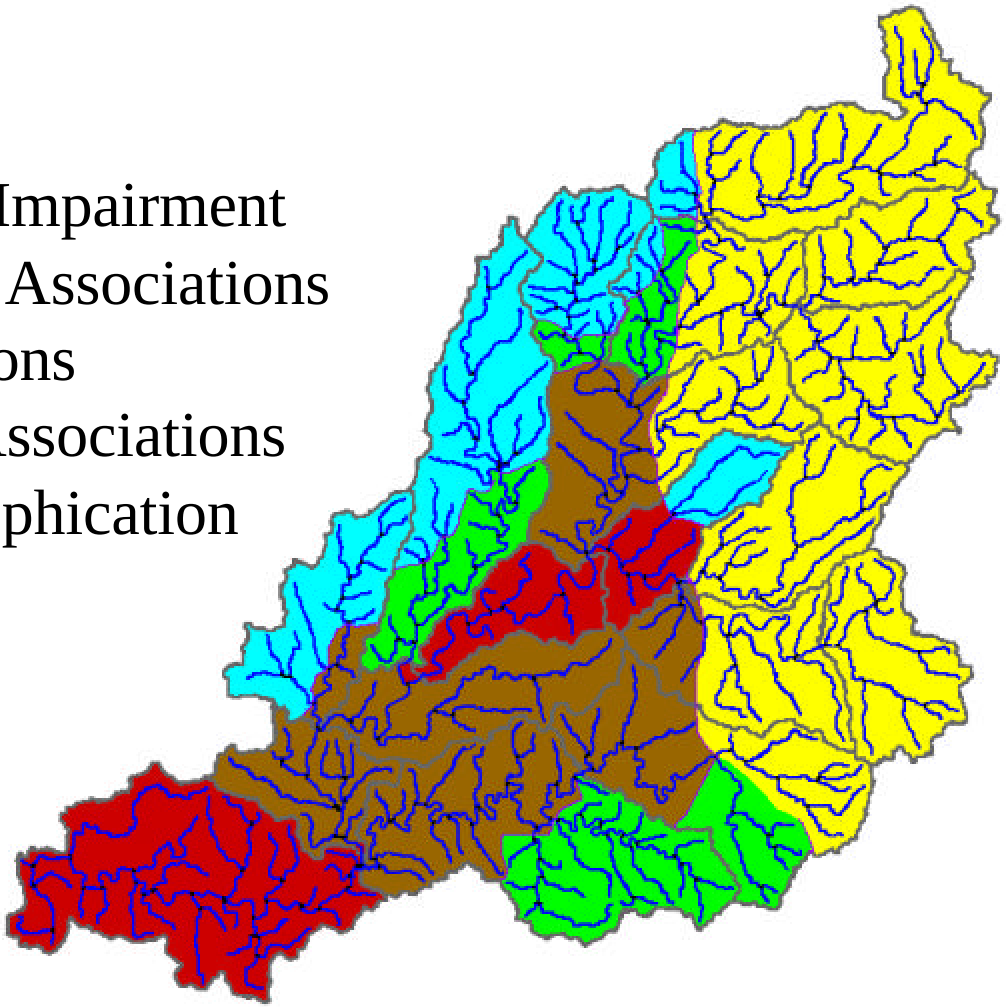
Example: Extending EMAP Status
Estimated IBI Condition at Reach Scale



FROM DAN McKENZIE – LABELED "DRAFT – DRAFT"

Potential Areas for Target Surveys

- High Prob. Non-Impairment
- Riparian Habitat Associations
- Acidic Associations
- Eutrophication Associations
- Toxicity & Eutrophication Associations



FROM DAN McKENZIE – LABELED “DRAFT – DRAFT”

EMAP SYMPOSIUM- 2004
Theme 2 – Session 2:
RELATING FINDINGS FROM 305(b) TO 303(d)

- Linking CWA Sections 305(b) and 303(d) – statistical Perspective.
  - → Overview - Scott Urquhart, Colorado State University
  - → A Role for Small Area Estimation -F. Jay Breidt, Colorado State University
  - → Estimating Power to Detect Trends – Brian R. Gray, USGS

A STATISTICAL PERSPECTIVE
ON
LINKING SECTIONS 305(b) AND 303(d)
OF
THE CLEAN WATER ACT

N. Scott Urquhart

SENIOR RESEARCH SCIENTIST
DEPARTMENT OF STATISTICS
COLORADO STATE UNIVERSITY

EMAP Affiliate

SPACE-TIME AQUATIC RESOURCE MODELING and ANALYSIS PROGRAM (STARMAP)

STARMAP FUNDING

Space-Time Aquatic Resources Modeling and Analysis Program

The work reported here today was developed under the STAR Research Assistance Agreement CR-829095 awarded by the U.S. Environmental Protection Agency (EPA) to Colorado State University. This presentation has not been formally reviewed by EPA. The views expressed here are solely those of presenters and STARMAP, the Program they represent. EPA does not endorse any products or commercial services mentioned in these presentation.

This research is funded by

U.S.EPA – Science To Achieve Results (STAR) Program

Cooperative Agreement # CR - 829095

PATH for TODAY
GETTING FROM
305(b) SURVEYS TO 303(d) TMDLs

- Spatial-Temporal Modeling for Aquatic Systems
- A Conceptual Model for Linking Two Sorts of Data: Probability Survey & Other Sites Where STARMAP fits in
  - ➔Spatial-temporal modeling for aquatic systems
  - ➔Relevant current STARMAP research
  - ➔Learning materials for aquatic monitoring
    - *Poster: 6 – 8pm, Wednesday, Bellview Ballroom*
- How YOU Can Help STARMAP Develop Tools to Help YOU
- Discussion/Questions

SPATIAL-TEMPORAL MODELING
for
AQUATIC SYSTEMS

- Spatial-temporal Modeling = ???
  - ➔ Most statistical techniques taught in graduate statistical methods courses assume observations are uncorrelated.
  - ➔ REALITY = Nearby things often are more alike than things far apart – regardless of context
    - *This is spatial correlation*
  - ➔ So what should we do?
    - *Design studies to minimize the impact of spatial correlation – EMAP is set up this way*
      - Good use of resources for summaries & estimating relationships
    - *Capitalize on the spatial correlation to get reliable forecasts of nearby response values*
  - ➔ Add time to the mix for spatial-temporal modeling
- How to Pull All of This Together FOR 305(b)/303(d)?

AVAILABLE INFORMATION

("ASSUMPTIONS")

- A Response of Interest
- A Probability Sample In A Region {305(b)}
- Some Purposefully Chosen Points in the Region
- Spatially Intensified Points Near Some of the Points
- Predictors at Whatever Density Desired, Like Landscape (GIS)

STRATEGY TO CONSIDER

- 1. Estimate Response/Predictor Relationship
- 2. Estimate the Spatial Relationship
  - ➔ Semivariogram
- 3. Estimate the Response/Predictor Values for a Dense Set of Points
- 4. Use Spatial Interpolation to Combine Forecasted Response Values With Observed Values
- 5. Get Confidence Bounds on the Combined Estimates

1. ESTIMATE RESPONSE/PREDICTOR RELATION

- **Estimate the Relation Between the Response and the Predictors Using:**
  - **(a) The probability selected points, and separately**
  - **(b) The purposefully selected points**
  - **Combine the two estimates?**
    - ***If (a) & (b) don't differ very much, combine them***
    - ***If they differ substantially, use (a)***
      - **Reason – by differing, the biases in the purposefully selected points affect the estimated relation, while the probability selected points represent the whole region.**
  - **Denote the resulting estimate as $f_1(s)$, where s represents a point in (two-dimensional) space**

LIMITATIONS OF APPROACHES

- **Many Investigators Have Unreasonable Expectations for**
  - **Remotely sensed variables (GIS generated data)**
    - ***Good for extent – like land use classes, but …***
    - ***Aerially sensed features see the surface***
      - **Even only the canopy top**
    - ***Much flowing water has been underground at some point in its transit from precipitation to its eventual resting place***
    - ***Variables like land classes may predict from 50% or even 70% down to 10% of the variation in some interesting chemical indicators.***
  - **Spatial Statistics (to be discussed next)**
    - ***We'll return to this***

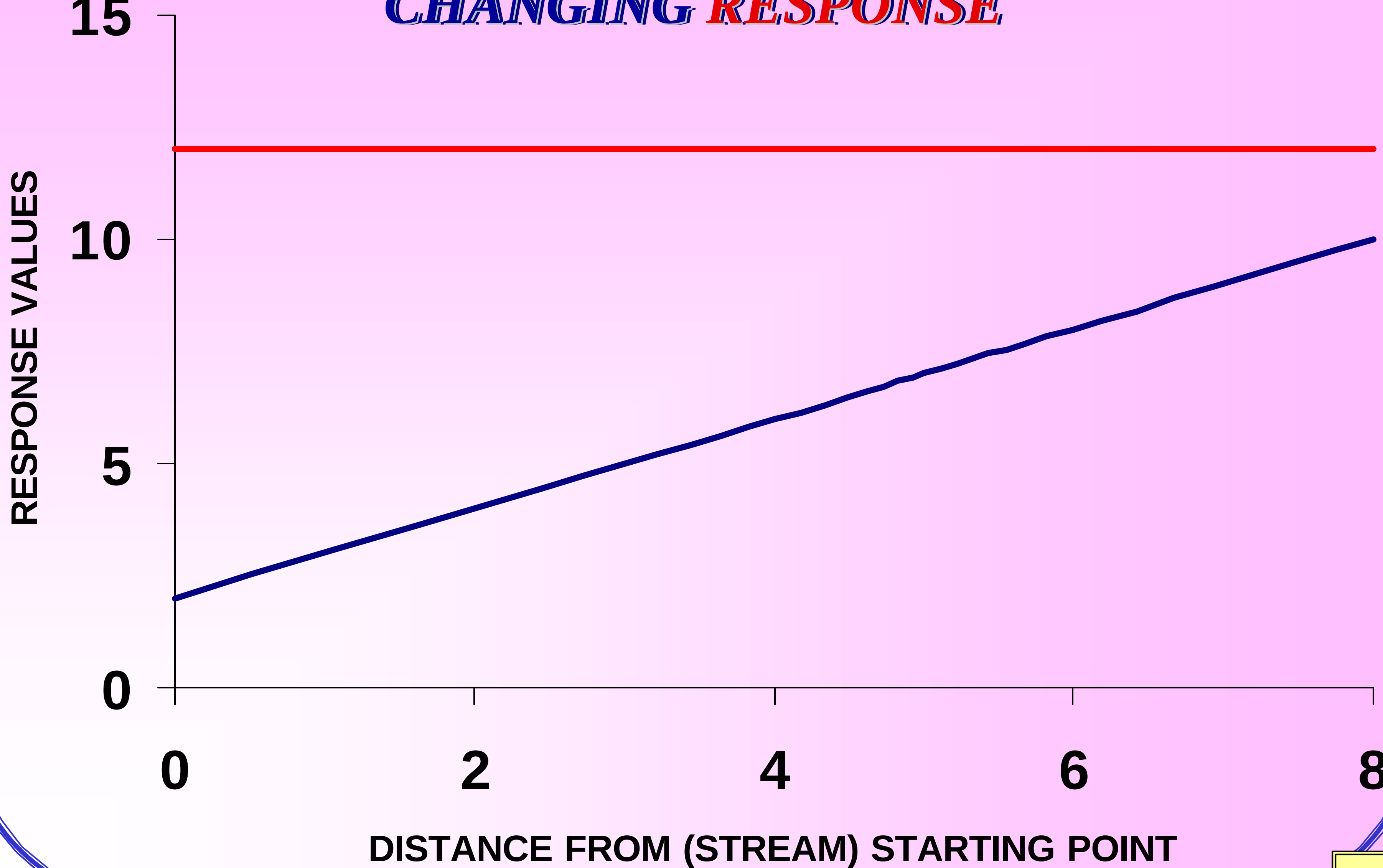
2. ESTIMATE THE SPATIAL RELATIONSHIP

- Use All Of The Available Relevant Data to Estimate the Semivariogram, $g(h)$,
- But Especially Rely on the Intensified Set of Points.
  - ➔ Spatial statistics usually mesures distance “as the bird flies”, but
  - ➔ Consider measuring distance along the stream/river network
    - *STARMAP has active work in this area*

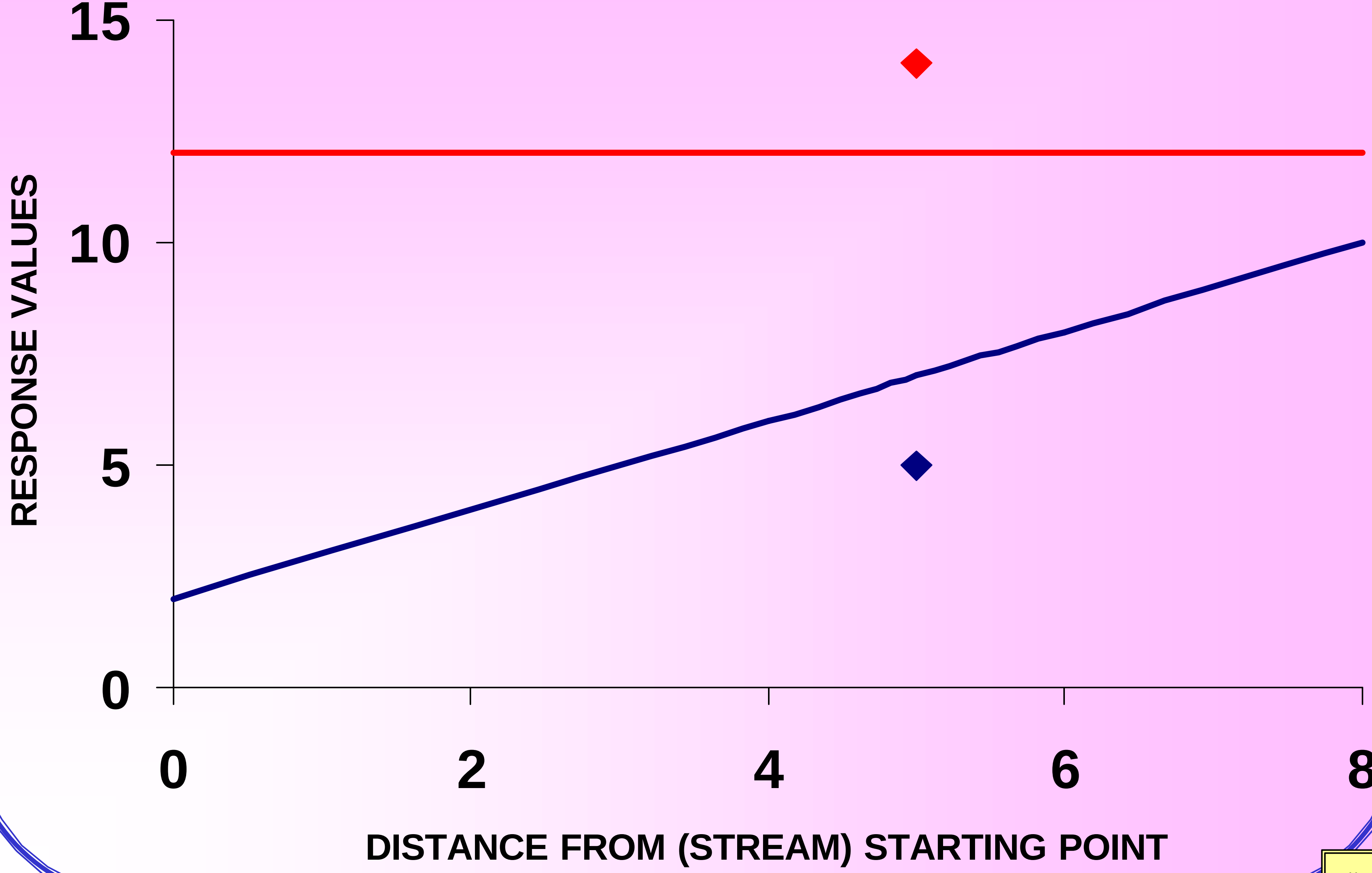
3. ESTIMATE THE RESPONSE/PREDICTOR RELATION FOR A DENSE SET OF POINTS

- A Dense Set of Points Might Be Every Kilometer Along the Stream/River Network
  - ➔Along a particular part of the stream network the result might look like what is shown on the next slide
    - *This shows only a small local part of the functions*
    - *This sort of representation should extend across the entire stream/river network*

CONSTANT RESPONSE
or a
CHANGING RESPONSE



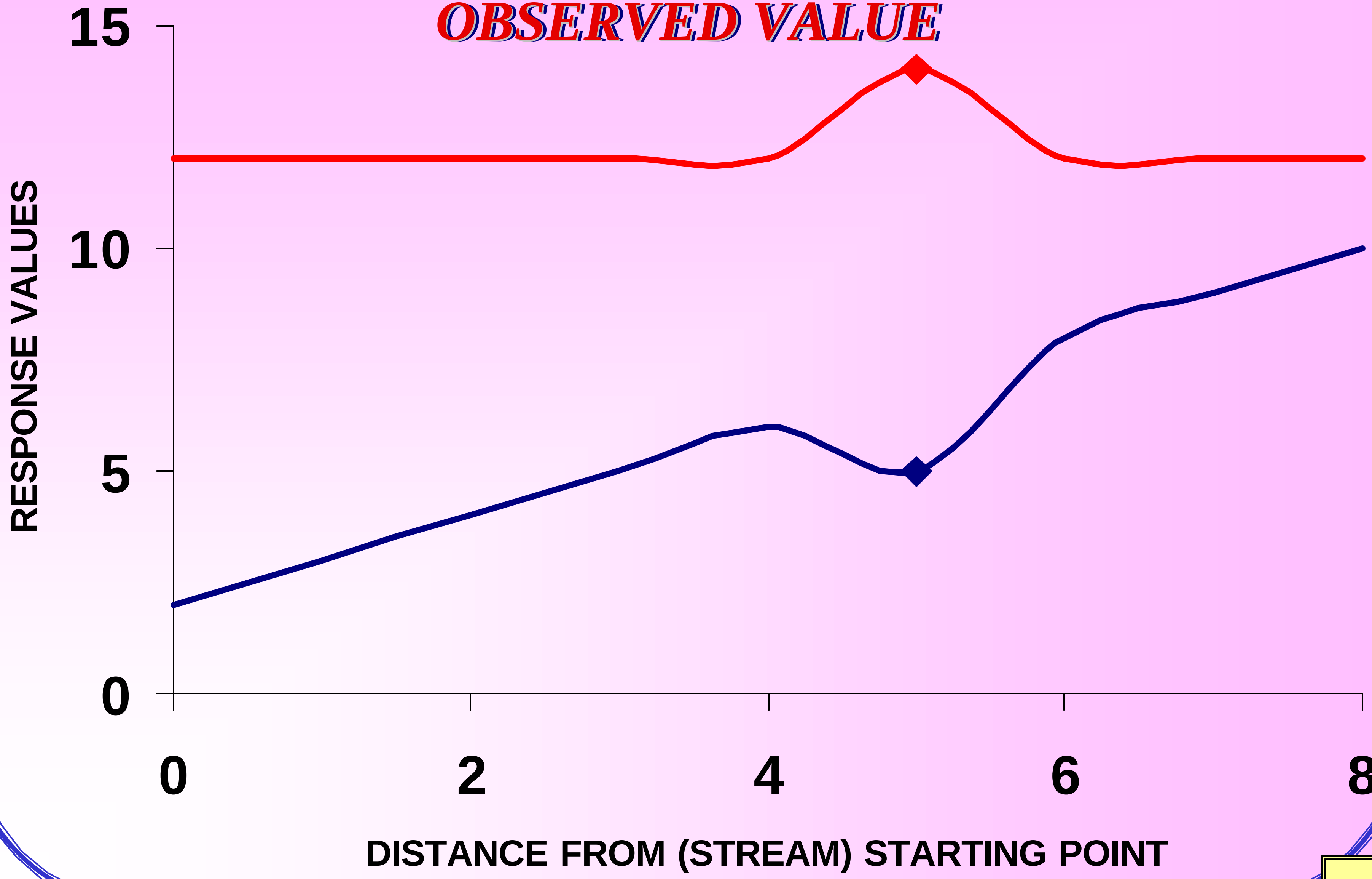
DO WHAT IF THE OBSERVED DATA DOESN'T MATCH THE PREDICTIVE RELATION?



4. A ROLE FOR SPATIAL INTERPOLATION

- If a Legitimate Observation is Below the Predictive Relation, It is Likely Nearby Points are Also. Make Use of This Expected Relation.
- Use Spatial Interpolation, of Which Kriging is An Example, to Smooth the Relation Through the Observed Point and Back to the Less Informed General Relation
  - ➔ Perhaps take a weighted average between the predicted value and the observed or spatially interpolated value
    - = *“shrinkage” estimate*

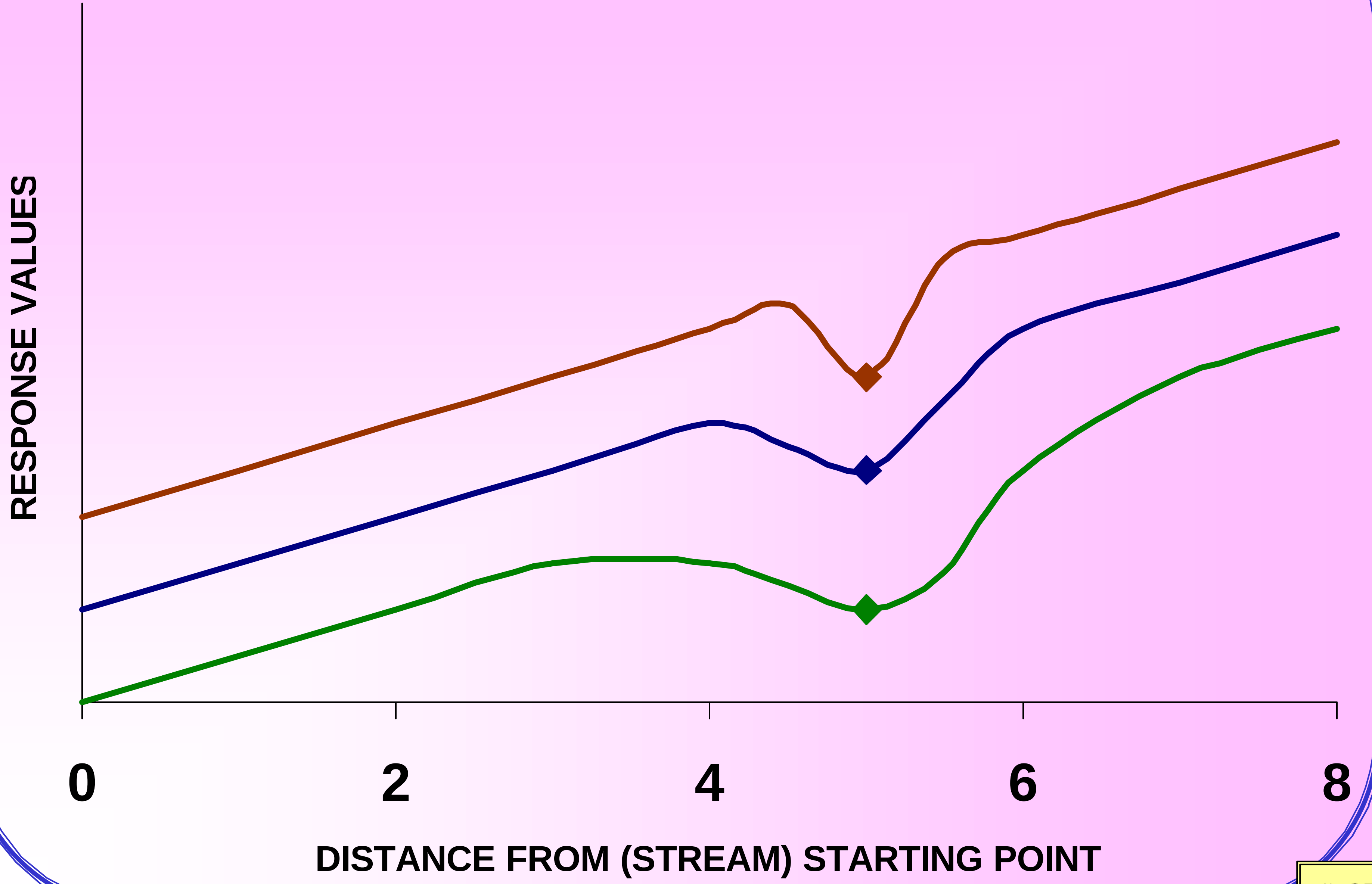
CHANGING RESPONSE WITH SPATIAL INTERPOLATION THROUGH OBSERVED VALUE



OPEN QUESTION

- How Far Should the Spatial Interpolation Extend?
  - ➔ What difference would that make?
  - ➔ See the next figure
- This is an Open Question for Now.

DIFFERENT INTERPOLATION RANGES



LIMITATIONS OF APPROACHES

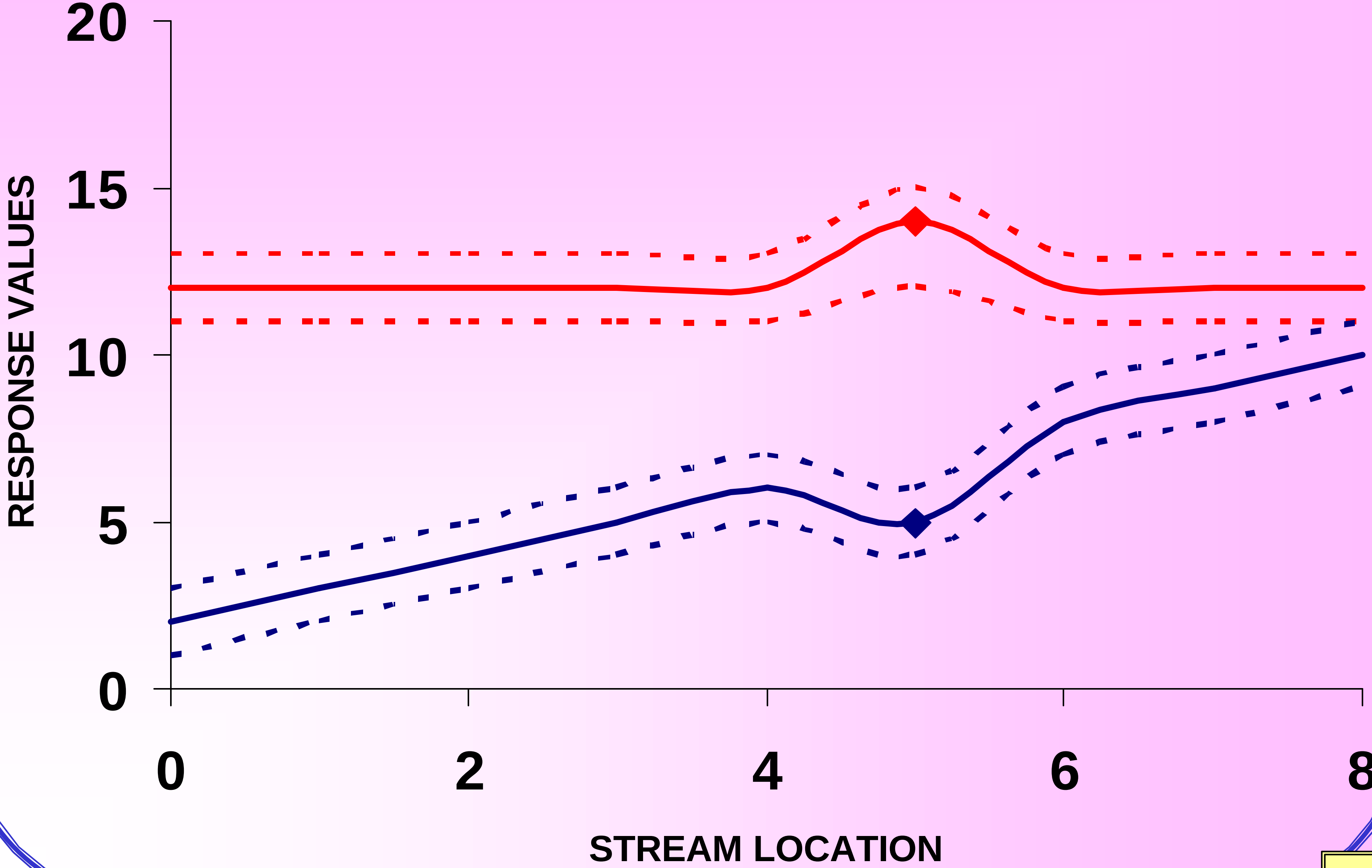
(second look)

- **Many Investigators Have Unreasonable Expectations for**
  - **Remotely sensed variables (GIS generated data)**
    - ***Discussed earlier***
  - **Spatial Statistics**
    - ***After accounting for habitat-type variables, aquatic responses may not exhibit much spatial correlation***
    - ***Certainly true in some forest situations***

RELATION TO CWA 303(d) IS?

- Wherever the Forecasted Response Exceeds the Standard, Go Check for Possible Violation
  - ➔ "Exceed" could be either high or low, depending on the response

5. FORECASTED RESPONSE VALUES WITH CONFIDENCE BOUNDS



RELATION TO CWA 303(d) IS?

- **Wherever the Forecasted Response Exceeds the Standard, Go Check for Possible Violation**
  - **"Exceed" could be either high or low, depending on the response**
- **Better Yet, Get a Confidence Bound on the Forecasted Response**
  - **Examine locations which exceed the confidence bound, rather than the forecasted response only.**
    - ***Way to allocate scarce resources***
  - **Width of confidence bounds will vary depending on how good the information is for the various points**

RELEVANT CURRENT STARMAP RESEARCH

- **Overall Objective: Develop and Disseminate Statistical Methods**
  - ➔ **Spatial/temporal/survey-related modeling**
  - ➔ **Relevant to aquatic monitoring**
  - ➔ **Next talk illustrates some of this perspective**
- **Current Research: How Should EMAP-type Sampling Be Intensified to Estimate Spatial Correlation:**
  - ➔ **Current context – City of San Diego and Southern California Coastal Water Research Project (SCCWRP)**
    - ***Accurate maps of environmental measures around SD's oceanic sewage outfall***

RELEVANT CURRENT STARMAP RESEARCH

(continued)

- Learning Materials for Aquatic Monitoring
  - → See poster

WHAT CAN YOU DO FOR STARMAP?

- ❖ **That Will Benefit Your Interests?**
  - ➔ **Do you have, or know of, aquatic environmental data sets which**
    - *Are dense along a stream or river network?*
    - *Like every 100m to 2 km*
    - *n = 100^{+} - hopefully without major habitat changes*
    - *If so, talk to me about them before we leave here*
  - ➔ **Look at the learning materials**
    - *Feedback on the interface*
    - *Poster: 6 – 8pm, Wednesday, Bellview Ballroom*
    - *Statistical topics you would like to be included – forms at poster display*
    - *Have access to studies which might be turned into case studies?*

QUESTIONS ARE WELCOME