POSSIBLE LESSONS FOR CEER-GOM FROM EMAP

N. Scott Urquhart
STARMAP Program Director
Department of Statistics
Colorado State University
CONTEXT FOR COMMENTS

◆ SPACE-TIME AQUATIC RESOURCES MODELING AND ANALYSIS PROGRAM = STARMAP

⇒ FUNDED BY EPA’s STAR PROGRAM, AS IS CEER-GOM (==> “SIBLING” PROGRAMS)

⇒ STARMAP IS TO USE EMAP AS A DATA SOURCE AND CONTEXT

⇒ NSU = STARMAP PROGRAM DIRECTOR @ CSU

● 10 YEARS OF COLLABORATION WITH EMAP
● 40 YEARS AS STATISTICIAN WORKING WITH ECOLOGISTS
LESSONS - A FEW

1. STATISTICS DOES NOT HAVE ALL OF THE TOOLS YOU NEED.

2. A USEFUL INDICATOR SHOULD APPLY ACROSS A WIDE RANGE OF CONDITIONS

3. YOU DO NOT KNOW WHAT YOUR DATA WILL BE USED FOR 20 YEARS FROM NOW
LESSON 1: STATISTICS DOES NOT HAVE ALL OF THE TOOLS YOU NEED

→ ECOLOGICAL/ENVIRONMENTAL RESEARCH PRODUCES SITUATIONS FOR WHICH APPROPRIATE STATISTICAL PROCEDURES DO NOT EXIST

- **IF A STATISTICAL APPROACH DOES NOT “MAKE SENSE” CHALLENGE YOUR STATISTICIANS TO FIND SOMETHING WHICH FITS YOUR SITUATION.**

- **EX: CLASSICAL SAMPLING THEORY IS BASED ON A LIST; IT LEADS TO ANSWERS OF LIMITED VALUE WHEN APPLIED TO SAMPLING STREAMS. OUTGROWTH: THIS HAS LED TO AN EXTENSION OF SAMPLING THEORY TO COVER CONTINUOUS SAMPLING FRAMES.**
LESSON 1:
STATISTICS DOES NOT HAVE ALL OF THE TOOLS YOU NEED
CONTINUED

◆ EPA HAS RECOGNIZED THIS!

⇒ EPA’s STAR PROGRAM INVESTS ~$2.5M/YEAR TOWARD DEVELOPMENT OF SOLUTIONS TO THIS LIMITATION

● COLORADO STATE UNIVERSITY: SPACE-TIME AQUATIC RESOURCE MODELING AND ANALYSIS PROGRAM (STARMAP). DIRECTOR = NSU

● OREGON STATE UNIVERSITY: DESIGN-BASED/MODEL-ASSISTED SURVEY METHODOLOGY FOR AQUATIC RESOURCES. DIRECTOR = DON STEVENS
LESSON 1:
STATISTICS DOES NOT HAVE ALL OF THE TOOLS YOU NEED
CONTINUED II

◆ EPA HAS RECOGNIZED THIS!

⇒ EPA’s STAR PROGRAM INVESTS ~$2.5M/YEAR TOWARD DEVELOPMENT OF SOLUTIONS TO THIS LIMITATION

• ...

• UNIVERSITY OF CHICAGO. DIRECTOR: CENTER FOR INTEGRATING STATISTICAL AND ENVIRONMENTAL SCIENCE. DIRECTOR = MICHAEL STEIN
EPA’s REQUEST FOR APPLICATIONS (RFA)

◆ CONTENT REQUIREMENTS

→ RESEARCH IN STATISTICS
  - DIRECTED TOWARD USING, IN PART, DATA GATHERED BY PROBABILITY SURVEYS OF THE “EMAP-SORT.”
→ TRAINING OF “FUTURE GENERATIONS” OF ENVIRONMENTAL STATISTICIANS
→ OUTREACH TO THE STATES and TRIBES

◆ ADMINISTRATIVE REQUIREMENT
MAJOR ADMINISTRATIVE REQUIREMENT

- “... EACH OF THE TWO PROGRAMS ESTABLISHED WILL INVOLVE COLLABORATIVE RESEARCH AT MULTIPLE, GEOGRAPHICALLY DIVERSE SITES.”
- CLOSE COOPERATION BETWEEN TWO PROGRAMS
- CSU and OSU SUBMITTED A PAIR OF COORDINATED PROPOSALS
EPA’s REQUEST FOR APPLICATIONS (RFA) - continued III

- THE TWO PROGRAMS:
  - DESIGN-BASED/MODEL ASSISTED SURVEY METHODOLOGY - @ OSU
  - SPATIAL AND TEMPORAL MODELING, INCORPORATING HIERARCHICAL SURVEY DESIGN, DATA ANALYSIS, MODELING - @ CSU

- CHECK ON THE RFA @

  http://es.epa.gov/ncerqa/rfa/aquastat01.html
RESPONSE to RFA from CSU

◆ INSTITUTIONS:

► COLORADO STATE UNIVERSITY
  ● STATISTICS
  ● NATURAL RESOURCES ECOLOGY LAB
  ● FOREST SCIENCE
  ● BIORESOURCE AND AGRICULTURAL ENGINEERING

► OREGON STATE UNIVERSITY (PROGRAM 1, too)

► SOUTHERN CALIFORNIA COASTAL WATER RESEARCH PROJECT

► WATER QUALITY TECHNOLOGY, INC
CSU PROPOSAL - CONTENT

◆ 1. COMBINING ENVIRONMENTAL DATA
◆ 2. LOCAL INFERENCE
◆ 3. DEVELOPING AQUATIC INDICATORS
◆ 4. OUTREACH
◆ 5. ADMINISTRATION/COORDINATION
CSU PROPOSAL - APPROACH

- TAKE EXISTING SETS OF
  - PROBABILITY &
  - NON-PROBABILITY DATA
- START WORKING WITH THE DATA WITH A PERSPECTIVE OF DRAWING INFERENCES
  - IDENTIFY ISSUES WE DON’T KNOW HOW TO HANDLE
  - HAVE POST-DOCS AND PRE-DOCTORAL STUDENTS CONDUCT RESEARCH ON THESE TOPICS
WHAT IS DISTINCTIVE ABOUT “AQUATIC RESOURCES”?

◆ THEY ARE THINGS LIKE

- STREAMS
- RIVERS
- WETLANDS
- LAKES & PONDS
- ESTUARIES
- PRAIRIE POTHOLES
- NEAR COASTAL OCEANIC WATERS
FOR MOST AQUATIC RESOURCES,

- THERE ARE MANY “SMALL” ONES
- PROGRESSIVELY FEWER AS THEY GET BIGGER
- INTEREST, BIOLOGICAL & SOCIETAL, TENDS TO STAY CONSTANT OR EVEN INCREASE WITH SIZE
- SIMPLE RANDOM SAMPLING WOULD SELECT MOSTLY “SMALL” ONES, FEW “BIG” ONES.

IMPLICATION:

- UNEQUAL PROBABILITY SAMPLING
WHAT IS DISTINCTIVE ABOUT "AQUATIC RESOURCES"?
CONTINUED II

- Spatial statistics tends to focus on two-dimensional space.
- Streams and rivers essentially amount to one-dimensional objects in two-space.
- But much landscape information is complete coverage in two-space.

◆ CHALLENGE:

- Merge these perspectives.
DISTINCTIVE EMAP PERSPECTIVE

◆ DEFINE THE POPULATION OF INTEREST

➔ CONDUCT A PROBABILITY SURVEY OF IT

- CAREFULLY DEFINE THE SAMPLING FRAME
- VARIABLE PROBABILITY SELECTION OF SITES, BUT WITH SPATIAL BALANCE
- CAREFULLY DEFINE RESPONSES TO BE EVALUATED
- TRAIN FIELD CREWS WELL
- MANAGE DATA WITH CARE AND AN “AUDIT TRAIL”
- LEARN FROM PAST MISTAKES, THROUGHOUT
FUTURE NEEDS - STATES & TRIBES

- STATES AND TRIBES MUST REPORT ON THE CONDITION OF ALL "WATERS" UNDER THEIR JURISDICTION

- A REQUIREMENT OF SECTION 305b OF THE CLEAN WATER ACT

- RESULTS IN BIANNUAL REPORT TO CONGRESS

- STARTING IN 2004 THE RECOMMENDED STANDARDS WILL CHANGE TO BEING BASED ON PROBABILITY SAMPLING

- OUTREACH PROJECT OPPORTUNITY!
Lesson 2: A USEFUL INDICATOR SHOULD APPLY ACROSS A WIDE RANGE OF CONDITIONS

◆ CONDITIONS SHOULD INCLUDE

  ➔ SPACE

  ➔ TIME

  • IDENTIFY ITS APPROPRIATE TIME WINDOW, IF LIMITED

  ➔ PHYSICAL/BIOLOGICAL CONDITIONS

  ➔ ENVIRONMENTAL QUALITY

◆ EVALUATION SITES SHOULD NOT BE IN A “CORNER” RELATIVE TO SUCH FEATURES
INDICATORS SHOULD APPLY ACROSS A WIDE RANGE OF CONDITIONS

◆ FROM YOUR PROPOSAL IT APPEARS THAT MOST OF YOUR STUDIES WILL BE CONDUCTED IN

→ GALVESTON BAY
→ MOBILE BAY
→ APALACHICOLA BAY

◆ THESE DIFFER PRIMARILY (?) AS A CONSEQUENCE OF FRESHWATER INPUT
POSSIBLE SPATIAL LIMITATIONS
of
CEER-GOM

GALVESTON
BAY

MOBILE
BAY

APALACHICOLA
BAY
EVALUATION SITES SHOULD NOT BE IN A “CORNER” RELATIVE TO IMPORTANT FEATURES

◆ OBSERVATIONS:

➔ HAVE YOU CONFOUNDED HIGH POLLUTION WITH LOW FRESHWATER INPUT?

➔ CONSIDER EVALUATING PROMISING INDICATORS OVER A WIDER SPATIAL DOMAIN IN THE LATTER YEARS OF THE PROGRAM
DEALING WITH LOCAL VARIATION

IN THE PRESENCE OF SUBSTANTIAL LOCAL VARIATION

MANY SCIENTISTS CONCENTRATE ON GETTING PRECISE LOCAL DETERMINATIONS

INSTEAD, CONSIDER COLLECTING MATERIAL OVER SOME SPACE

- THEN MIX (COMPOSITE) THE LOCALLY COLLECTED MATERIAL
- DO LABORATORY EVALUATIONS ON A SUBSAMPLE FROM THE WELL MIXED COMPOSITE
- THIS USES PHYSICAL AVERAGING
Lesson 3:

YOU DO NOT KNOW WHAT YOUR DATA WILL BE USED FOR 20 YEARS FROM NOW

◆ POPULAR PERSPECTIVE - WE “KNOW” LOTS ABOUT THE “ENVIRONMENT”

◆ REALITY: GOOD AQUATIC DATA IS SCARCE

- SPATIALLY EXTENSIVE
- OVER A REASONABLE TIME SPAN
- WELL DOCUMENTED PROCEDURES
- WELL TRAINED CREWS
- CAREFULLY EXECUTED STUDIES
- DATA PUBLICLY AVAILABLE
THE VALUE OF “METADATA”

◆ DATA

⇒ WITHOUT CONTEXT ARE NUMBERS
  ● NEARLY WORTHLESS TO OTHERS
  ⇒ DATA WITH CONTEXT IS INFORMATION
  ● CAN BE VALUABLE TO OTHERS

◆ CONTEXT IS CALLED METADATA
ASSOCIATE METADATA WITH ALL DATA

◆ USE IT TO DOCUMENT
  ➔ SITE SELECTION AND LOCATION
  ➔ FIELD PROTOCOLS FOR GATHERING
    • DATA & MATERIAL
  ➔ LABORATORY METHODS
  ➔ QUALITY ASSURANCE/QUALITY CONTROL
  ➔ METHODS USED TO DEAL WITH
    • NONDETECTS, MISSING OR LOST DATA, ETC
THANK YOU FOR YOUR ATTENTION