DEVELOPMENT OF ENVIRONMENTAL FLOW MODELS IN THE MINNESOTA RIVER WATERSHED

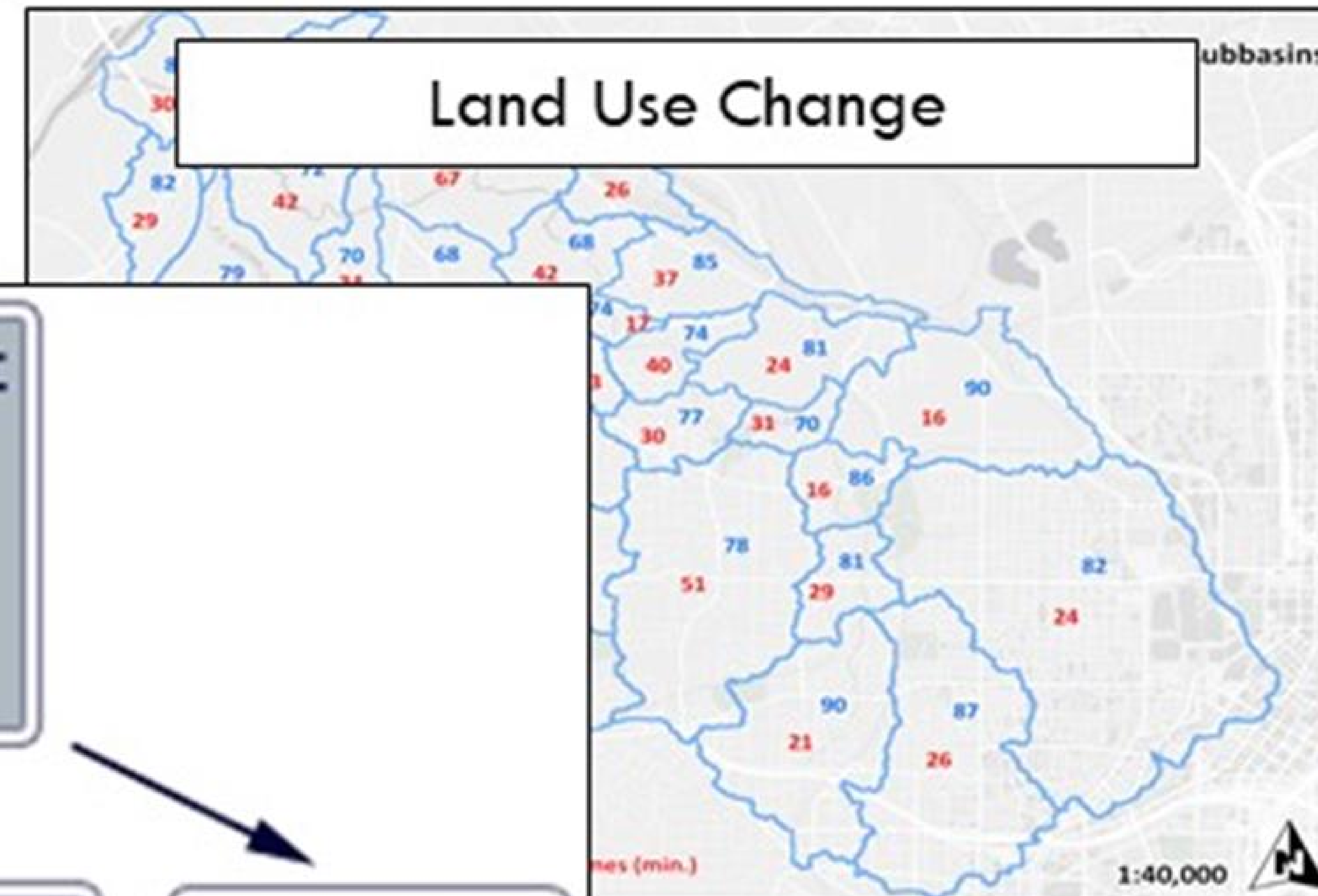
Co-authors:
Chuck Theiling & Michael Dougherty

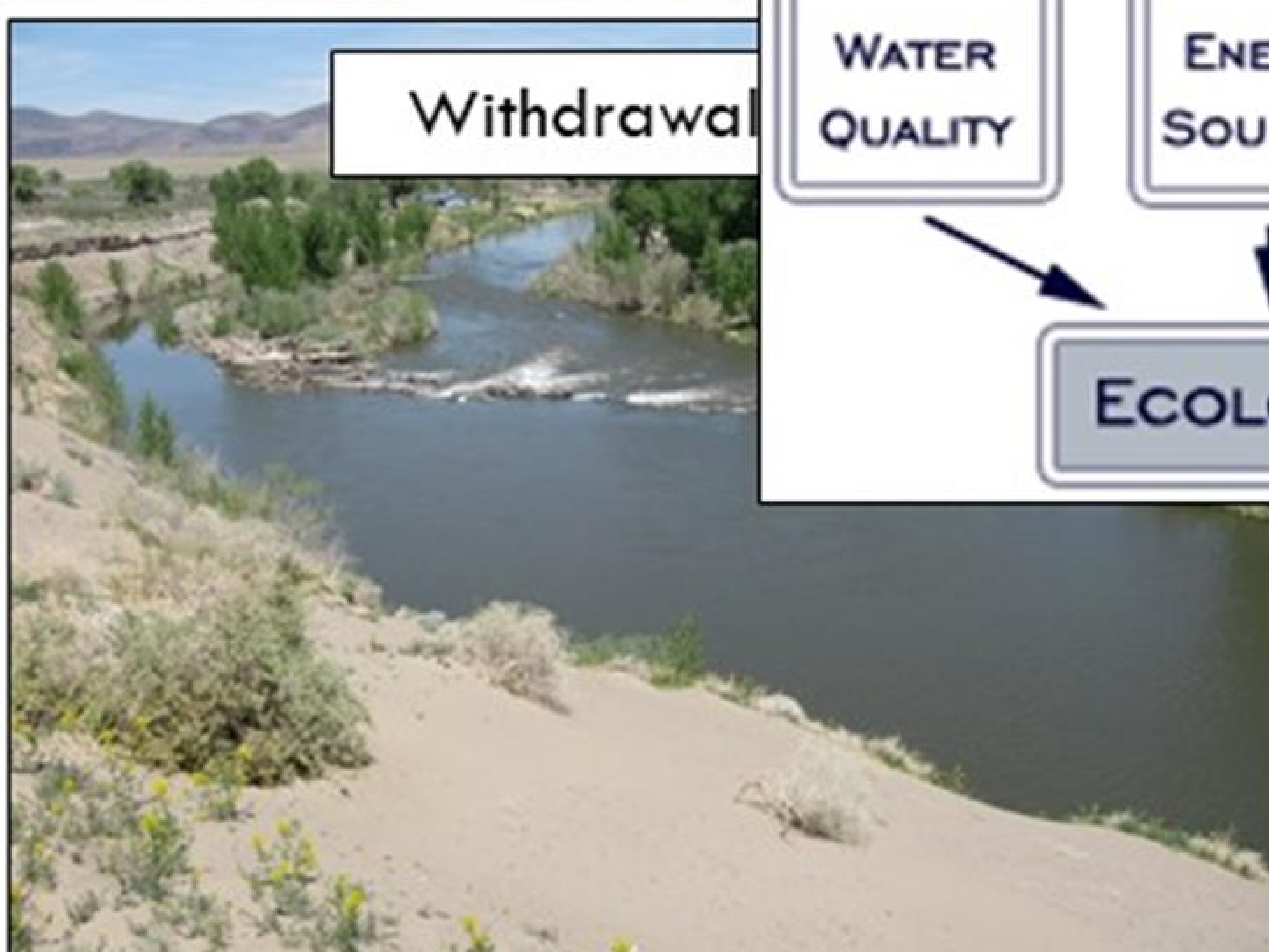
Kind funders:
USACE St. Paul District + Minnesota Pollution Control Agency
USACE Ecosystem Management and Restoration Research Program (EMRRP)

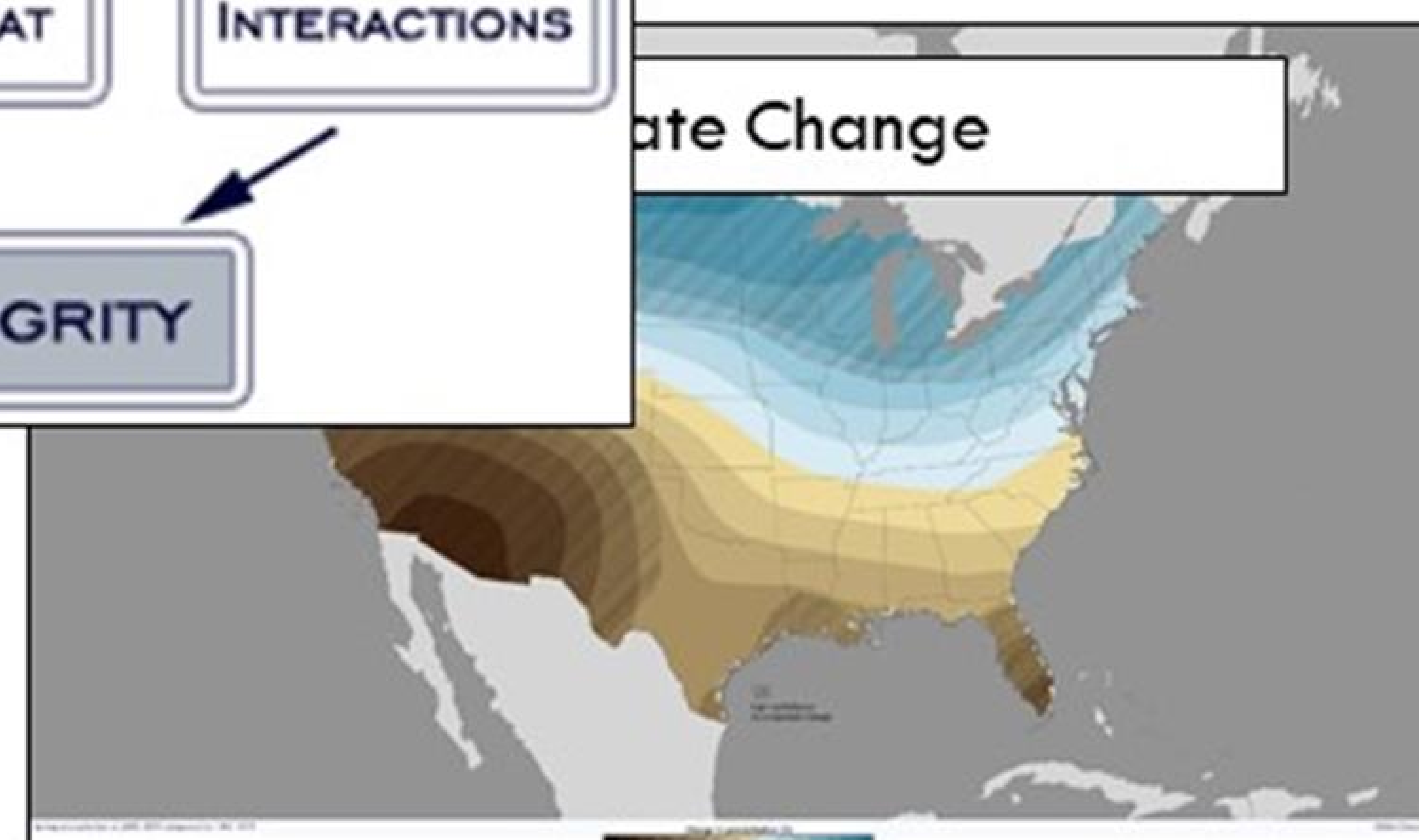
Kyle McKay, Ph.D., P.E.
September 2018

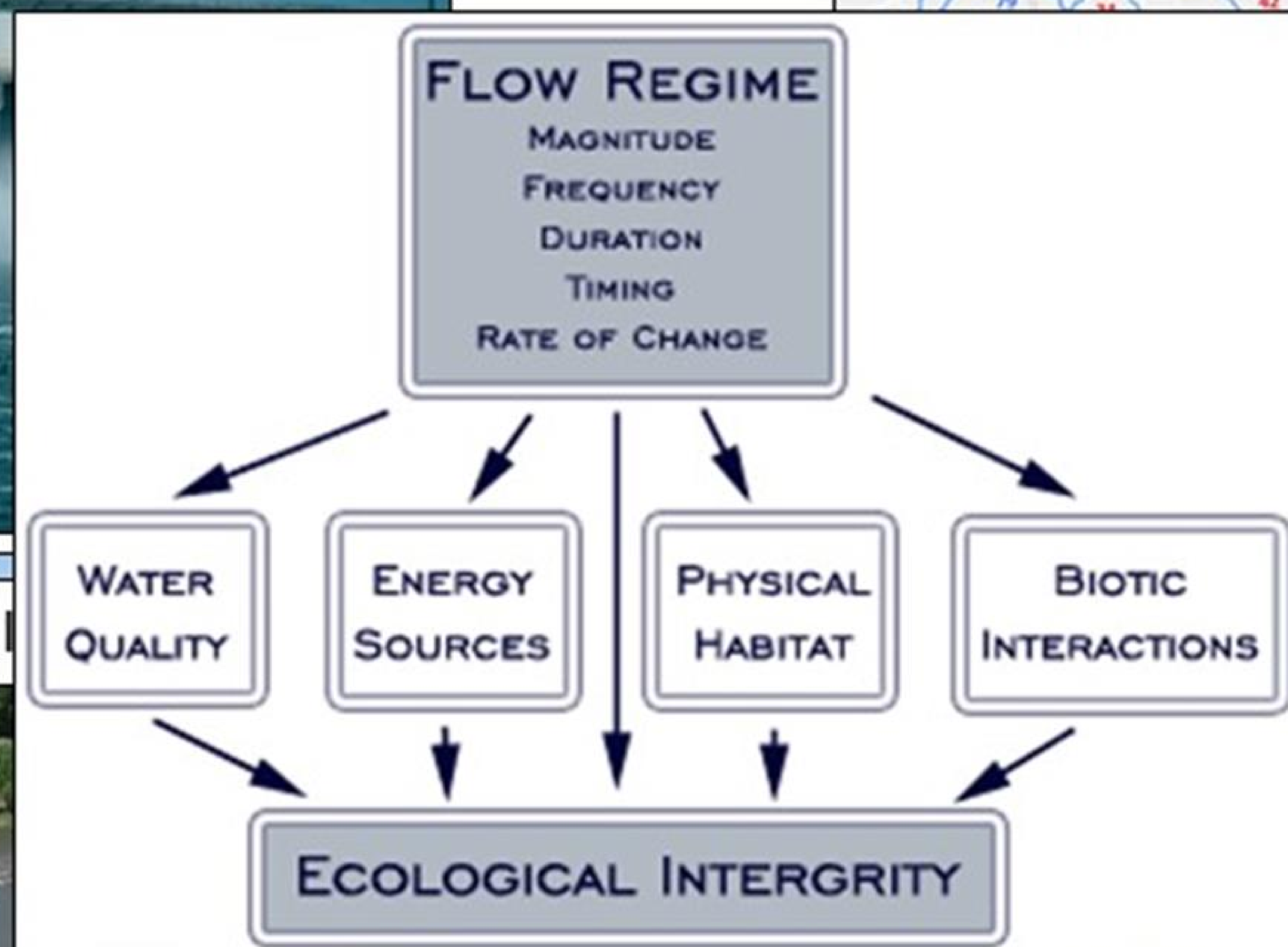
USACE Webinar Series
Ecological Modeling Project

How much water does a river need?
Environmental Flows





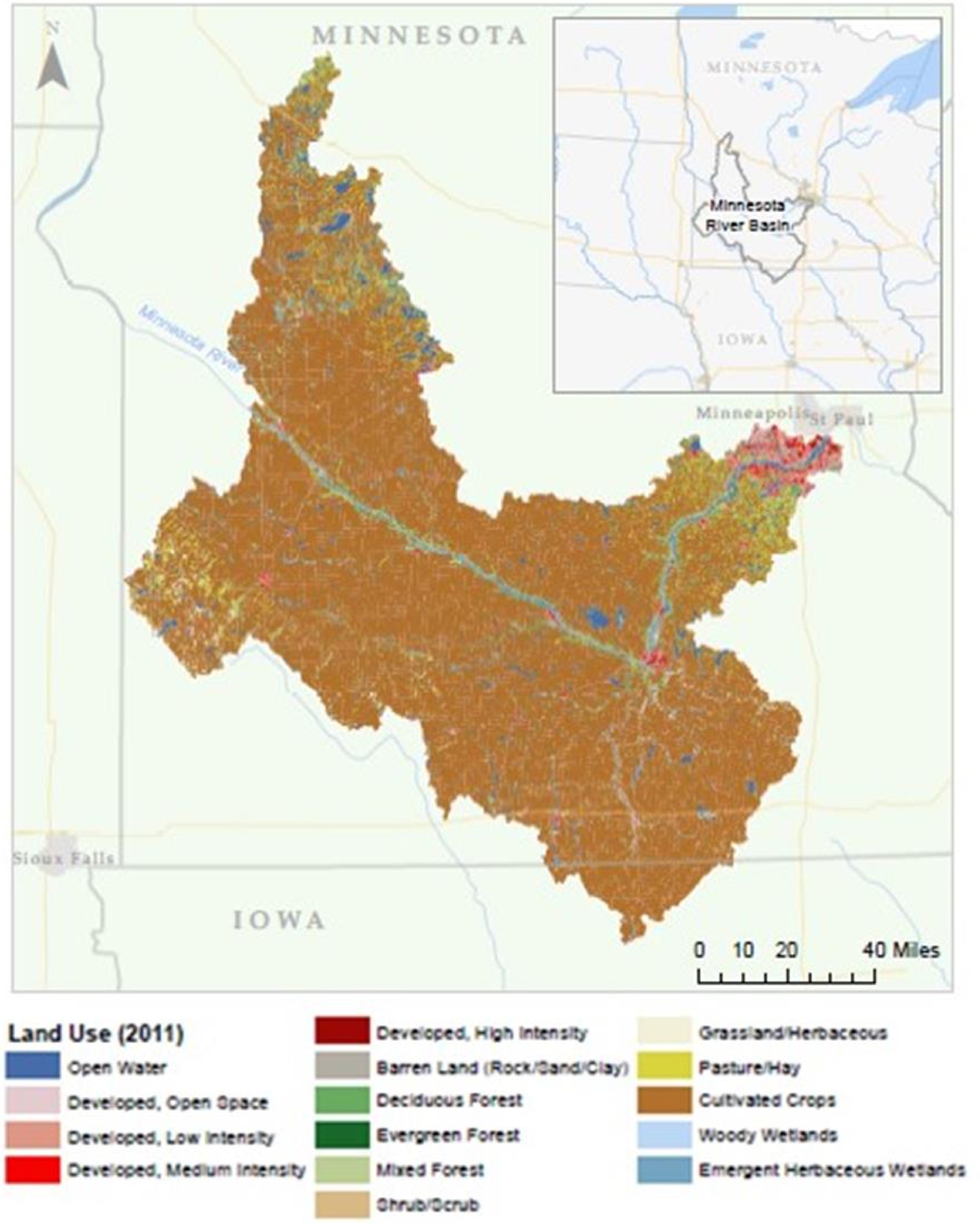




Study Site:
Minnesota River Basin

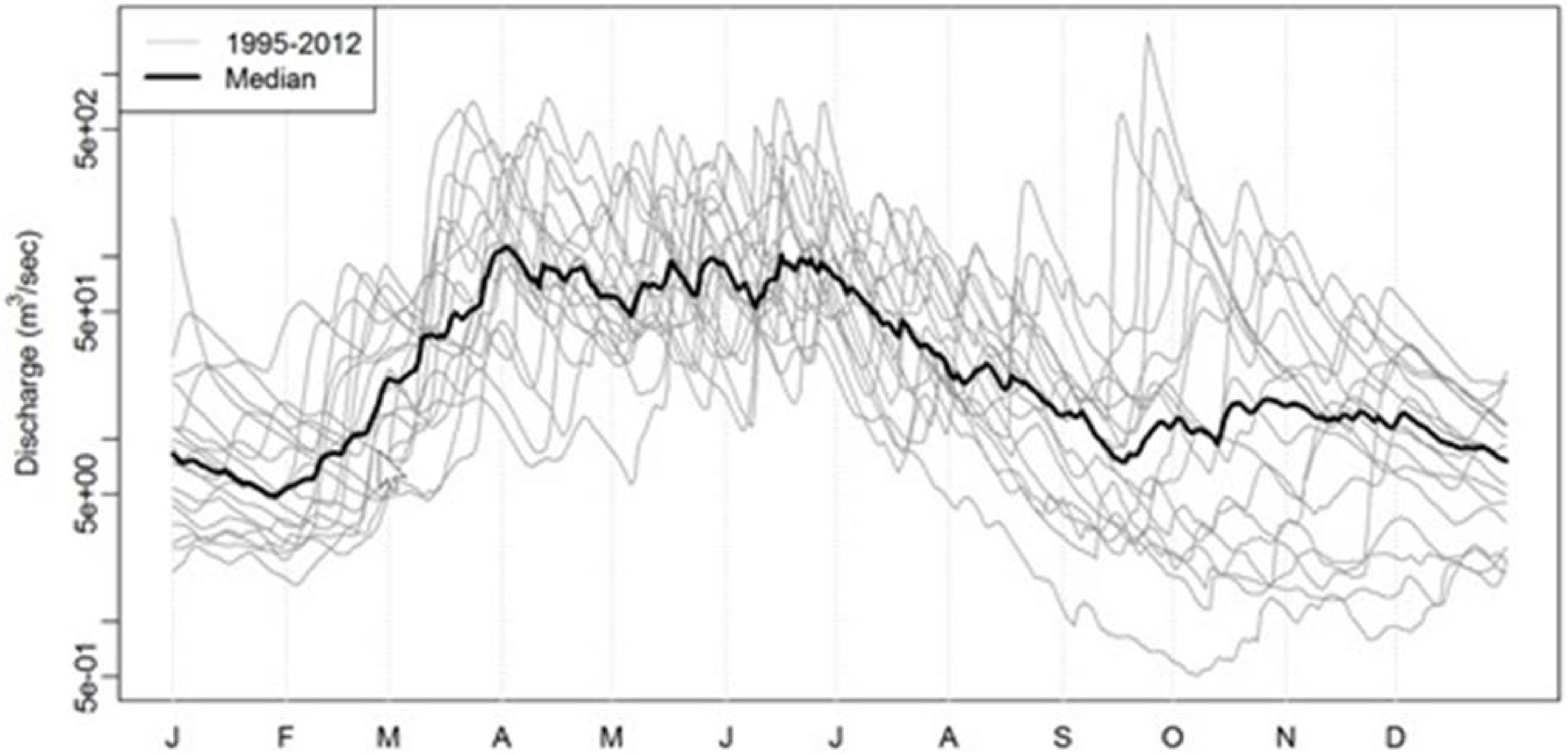
Minnesota River Basin

- Large drainage area (43,900 sq km)
- Historically dominated by poorly drained soils, tall-grass prairies, and extensive wetlands
- Extensive land use conversion and agricultural drainage
- High aquatic biodiversity
- State-declared management priority



Hydrologic Model (HSPF)

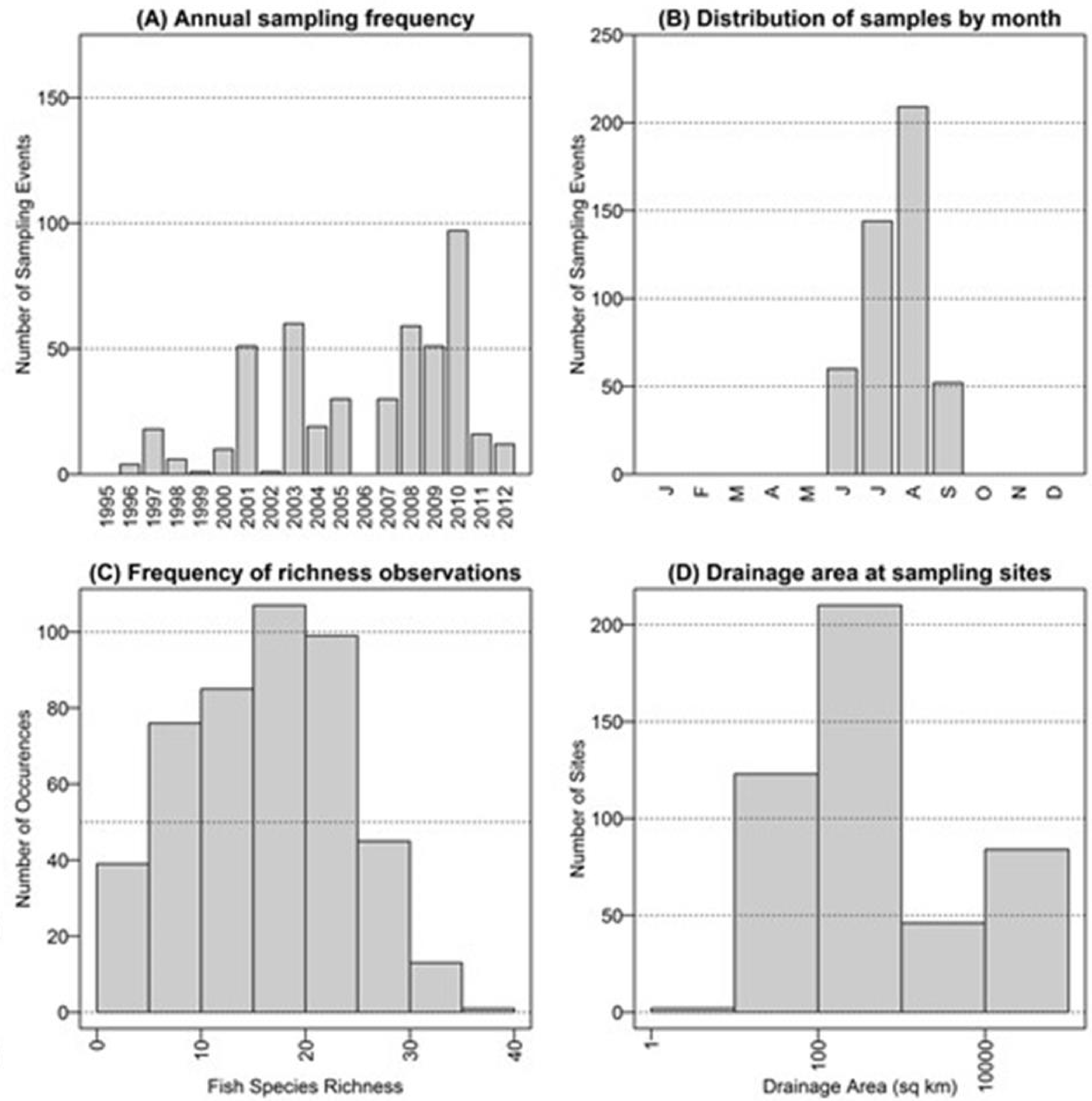
- Hydrologic Simulation Program Fortran (HSPF)
  - Model domain in use since ~1990
- Calibrated with observed streamflow (9 primary gages)
  - "Good" calibration over a 17-year window ($R^2 > 0.85$ and NSE > 0.75)
- Simulated flows for 1995-2012
  - Hydrologic and water quality outcomes
  - Sub-daily time-scales (aggregated to daily)
  - 1,016 locations throughout the watershed (i.e., "pour points").



Example Hydrograph
Blue Earth Basin
(RCH870)
Existing Condition

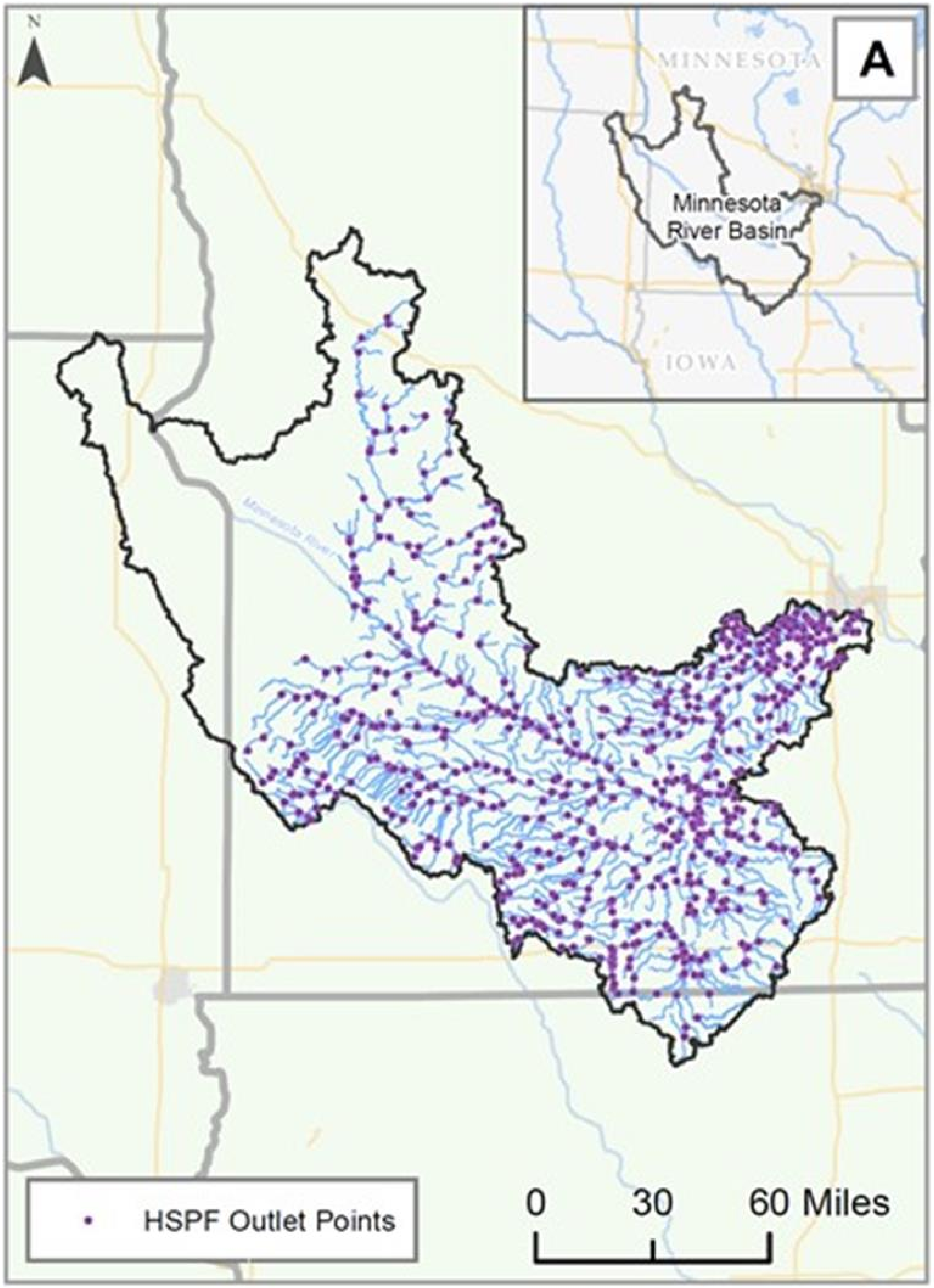
Fish Sampling

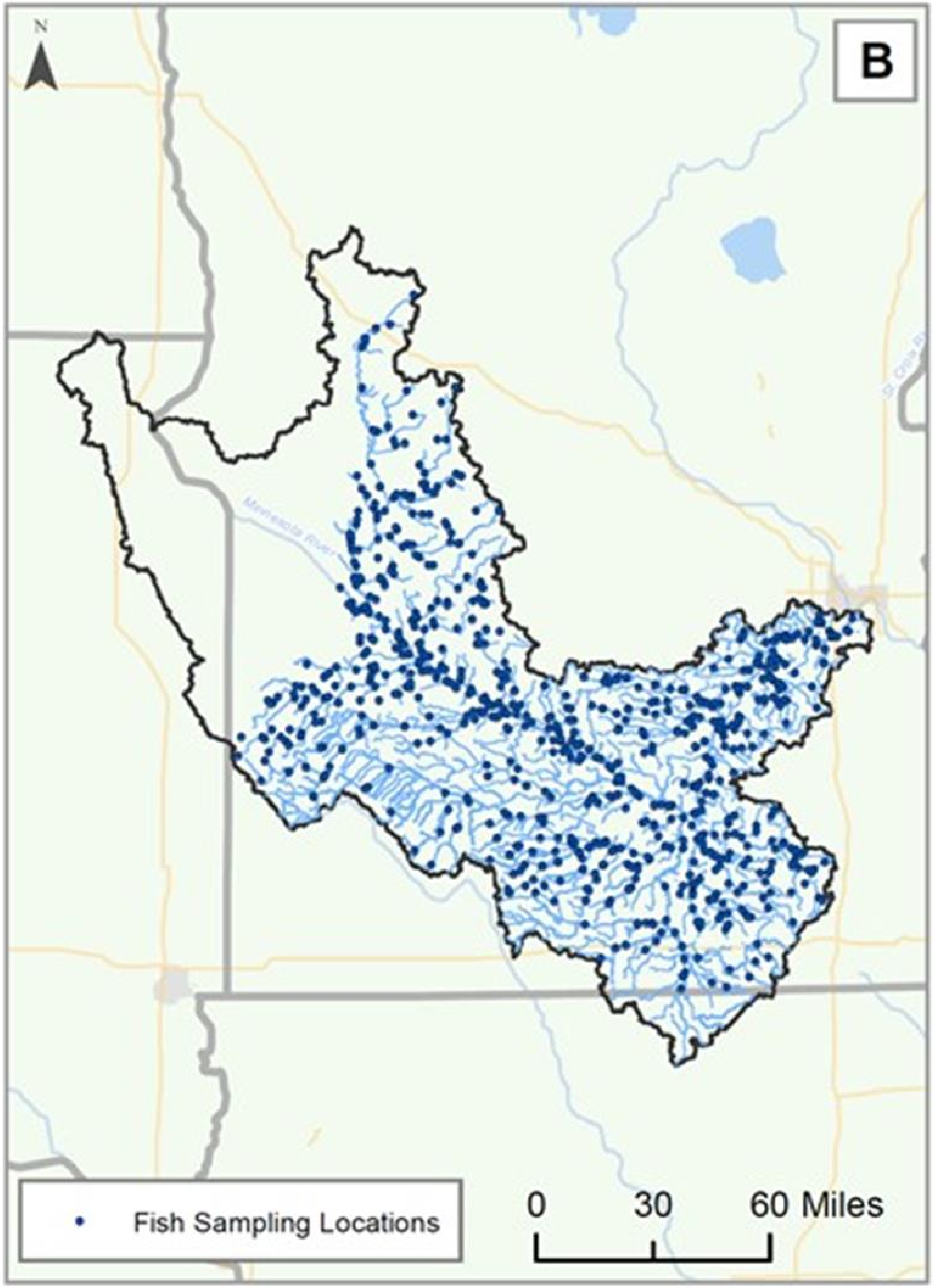
- MPCA biological monitoring
  - Sampled by electrofishing from 1996 to 2012 (639 visits)
  - Only richness data used here
- Hydrologic model used to estimate daily streamflow at sampling sites
  - Area-weighted gage-transfer
- Fish sites were eliminated if:
  - No nearby HSPF node
  - Sampling outside of 1995-2012
  - HSPF and sampling drainage areas were more than 25% different
  - Problems with HSPF calibration
- Final richness data set
  - 463 sampling events / 314 locations
  - Watersheds size = 6 to 43,000 km^2
  - Mean discharge = 0.05 to 349 m^3/s



Data provided by Scott Niemela and John Sandberg (MN-PCA)

Hydrology and Fish Sampling Sites





Map by Mary Windsor (MVR)

Research questions

- How does hydrology affect fish communities?
  - Model development
- How could changes in land use alter hydrology and subsequently affect fish communities?
  - Model application
- What can we learn from the Minnesota River case study about ecological model development?
  - Model reflection

Model Development

Model Development: Three competing philosophies

1. Hydrology as a proxy for ecology
   - Common assumption in eflows
   - Avoids messy flow-ecology relationships
   - May not reflect ecological outcomes
2. Deductive model
   - Hypothesis / principles driven approach
   - Managers may be more inclined to use understandable models
   - May not have the highest predictive power
3. Inductive model
   - May have higher predictive capability
   - "Black box" in terms of visualizing the mechanisms

Model-1: Hydrologic Change

Seven fundamental daily streamflow statistics

- Mean
- Coefficient of variation
- Skewness
- Kurtosis
- Auto-regressive lag-one (AR1) correlation coefficient
- Seasonal amplitude
- Seasonal phase shift

What is the relative change in all seven statistics?

Figure: Archfield et al. (2013, River Research)

Model-2:
Deductive, Linear Model

- What variables really matter?
- Dimensional analysis
  - OLD technique used to define scaling relationships

Variable	Basis for inclusion	Symbol	Units
Fish community richness	Number of fish taxa at a given site	R	1
Drainage area	Richness shown to scale with watershed size due to species-area relationships	A	L^2
Channel width	Surrogate for habitat quantity and geomorphic change	W	L
Channel depth	Surrogate for floodplain accessibility, light availability, and energy sources	D	L
Mean discharge	Hydrologic change from agricultural drainage has been shown to increase river flow across a range of flow frequencies	Q_m	L^3 / T
Gravitational constant	Consistent scaling parameter for dimensional analysis	g	L / T^2

Model-2:
Deductive, Linear Model

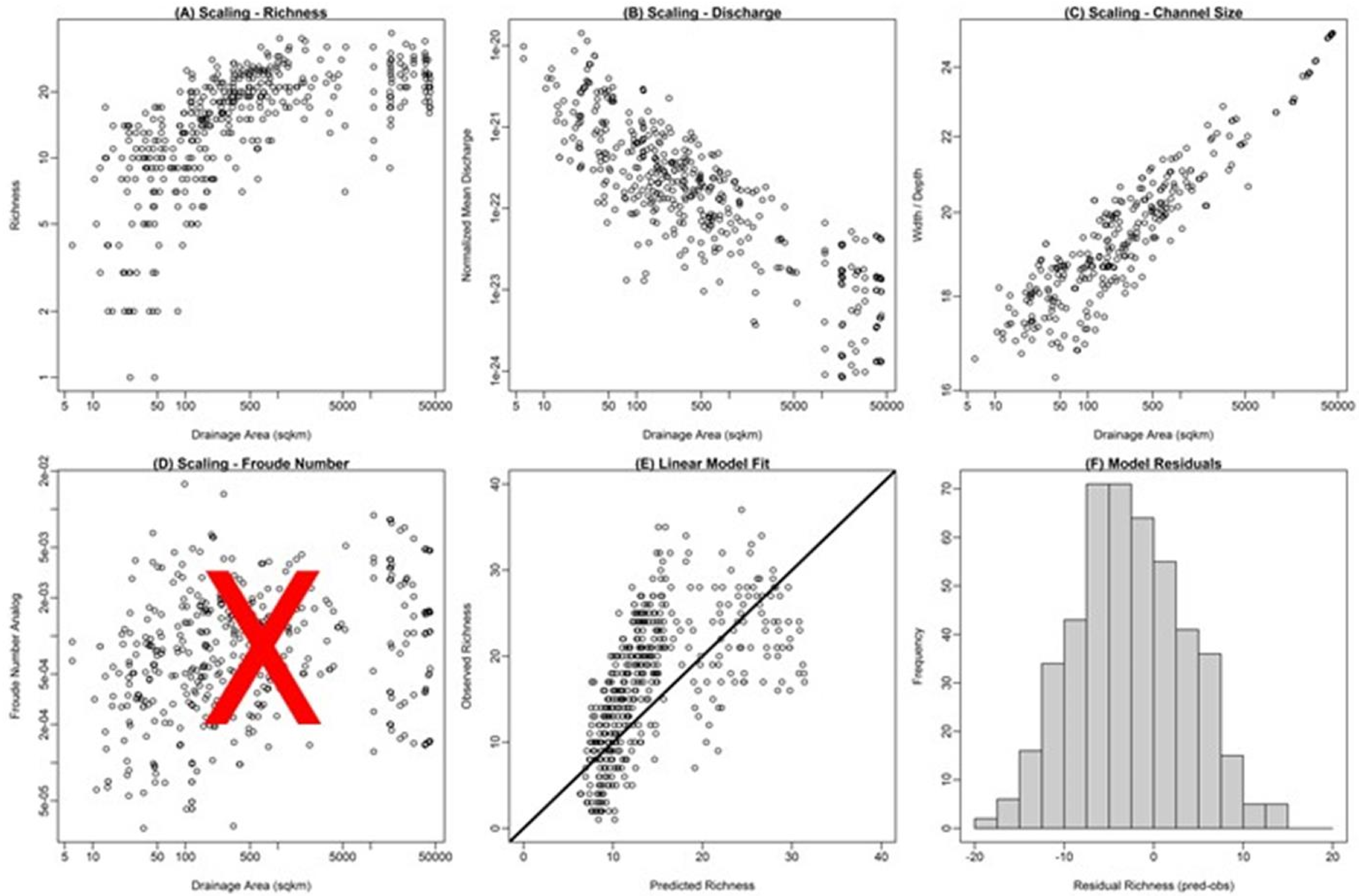
- Buckingham Pi Theorem
  - 6 variables – 2 dimensions = 4 dimensionless quantities
  - Method of repeating variables was used to normalize all variables
  - Dimensionless terms were combined and rearranged

$$R = f\left(\frac{{Q_m}^2}{gA^{2.5}}, \frac{W}{D}, \frac{{Q_m}^2}{gW^2D^2}\right)$$

- Term 1: Relative discharge for a given basin size
- Term 2: width-to-depth ratio (common geomorphic metric)
- Term 3: Froude number (an important hydrodynamic variable)

$$\ln R = \beta_1 + \beta_2 \ln\left(\frac{{Q_m}^2}{gA^{2.5}}\right) + \beta_3 \ln\left(\frac{W}{D}\right) + \beta_4 \ln\left(\frac{{Q_m}^2}{gW^2D^2}\right)$$

Model-2:
Deductive, Linear Model



$$\ln R = -8.543 + -0.0422 \ln\left(\frac{Q_m{}^2}{gA^{2.5}}\right) + 3.035 \ln\left(\frac{W}{D}\right)$$

$R^2 = 0.39$
(Froude term removed)

Model-3: Inductive, Machine Learning Model

- Inductive, data-driven models
  - Address challenges such as: "mixed" data types, missing values, robustness to outliers, and hierarchical interactions of ecological processes
- Boosted Regression Trees (caret / gbm packages in R)
  - Potential co-variates (135): drainage area, bankfull channel width, and bankfull channel depth, bunches of hydrologic covariates in the sample year, the year prior to sampling, and the long-term record
  - Model parameterization (boring details)
    - 10-fold cross-validation and bag fraction of 0.8
    - Interaction depth = 4, Number of trees = 450, Shrinkage rate = 0.01
  - Variable relative importance was used to select a subset of the most predictive co-variates to construct a final BRT

Model-3: Inductive, Machine Learning Model

<u>Full Model</u>

$R^2 = 0.81$

<u>Top Six Variables</u>

Long-term median

Long-term mean

Width

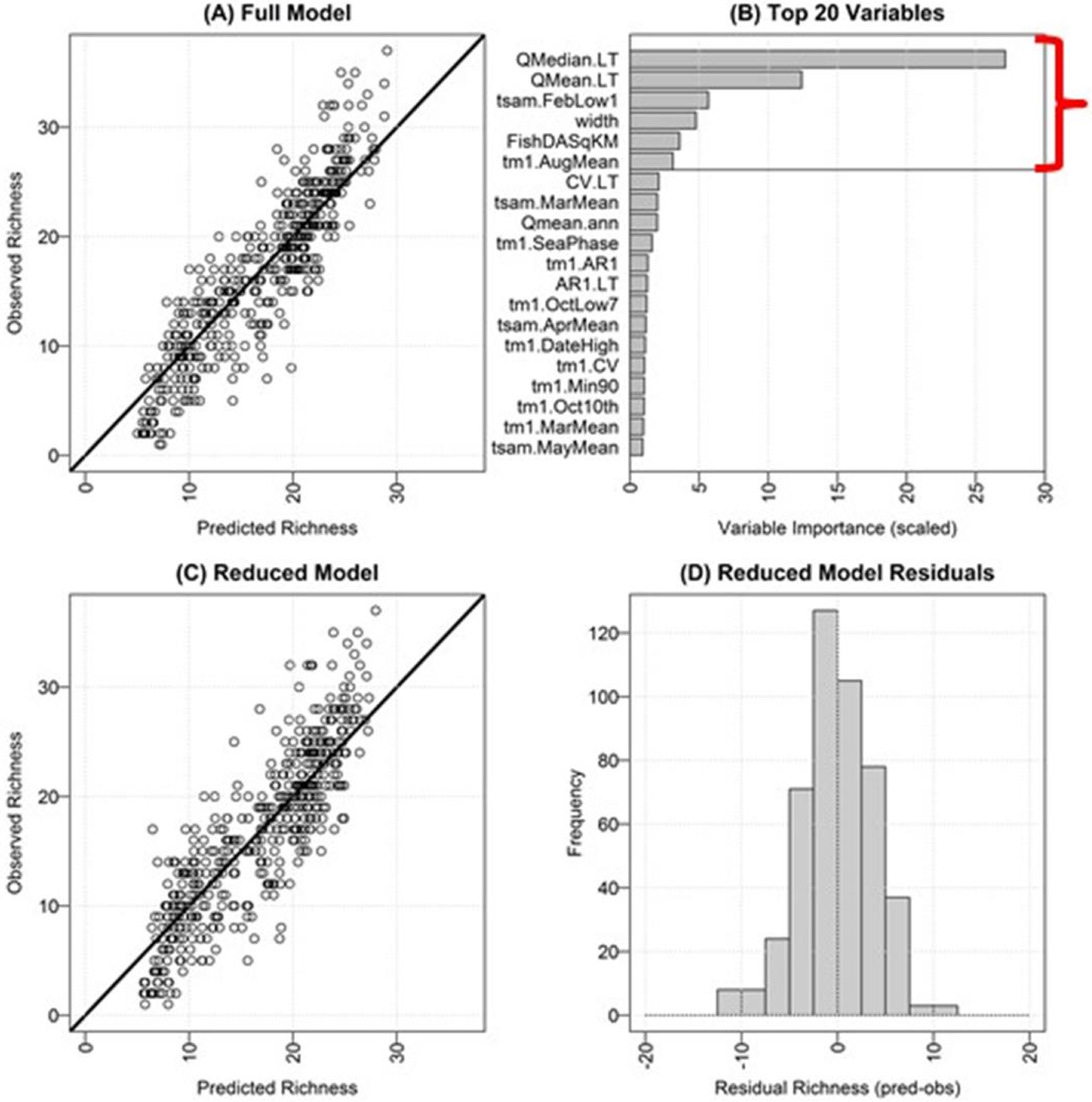
Drainage Area

Feb 1-day low

Aug mean (t-1)

<u>Reduced Model</u>

$R^2 = 0.75$



Model Application

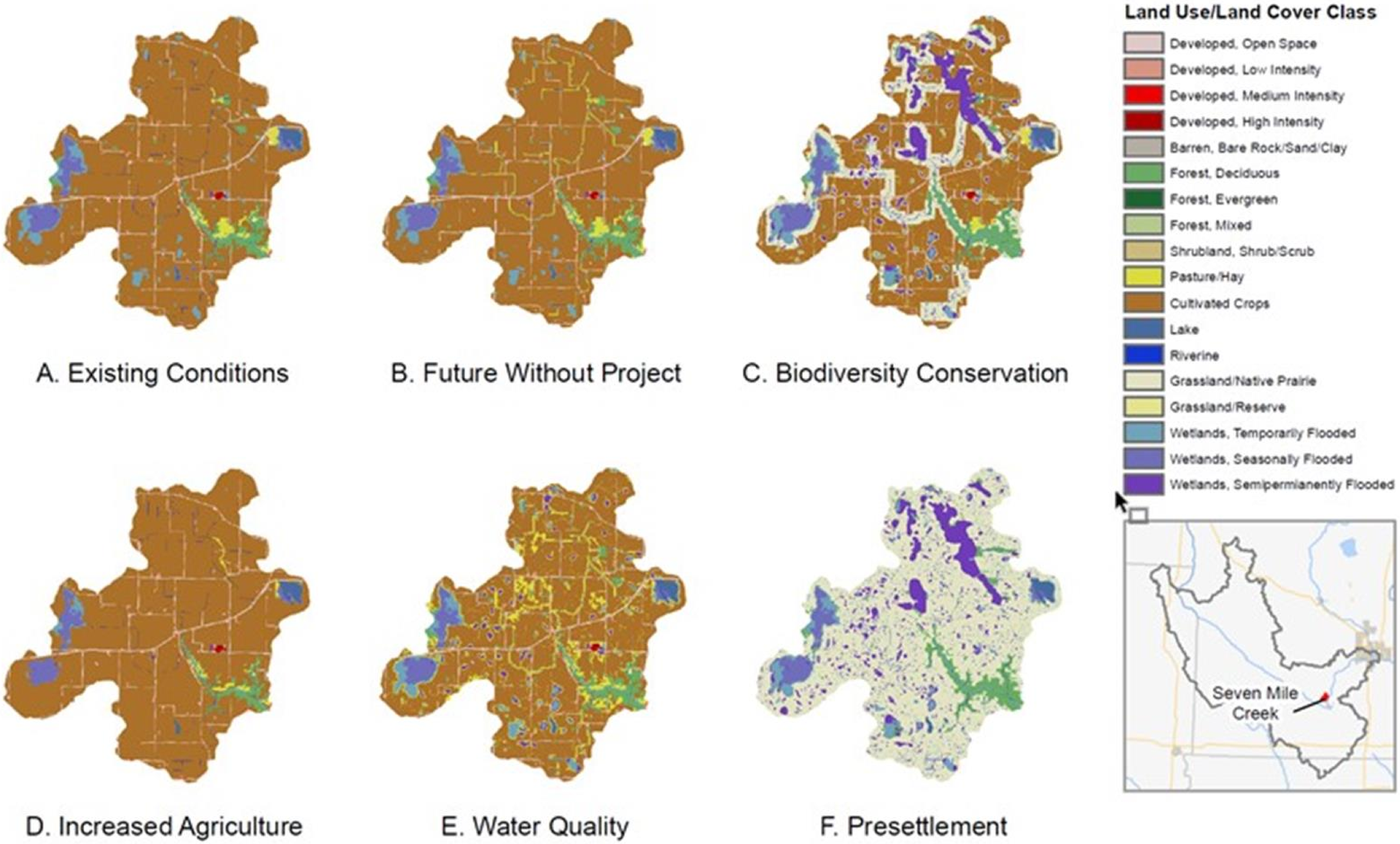
(ongoing project = all results subject to change)

Forecasting Land Use Change: Six Scenarios

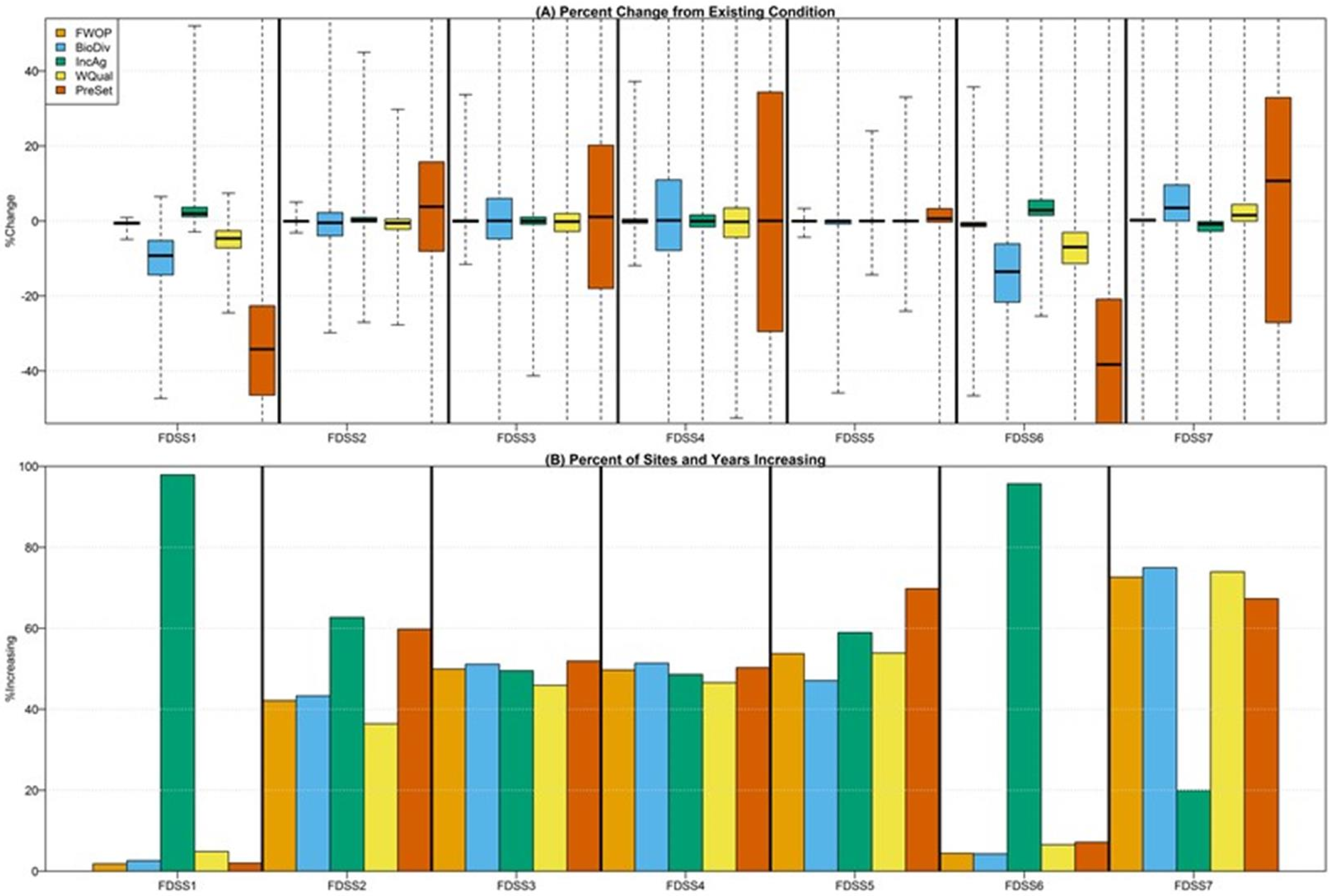
Scenario	Description
Existing Condition (ExCond)	What's out there right now Conditions during fish sampling
Future Without Project (FWOP)	Minor change from ExCond Includes enforcement of a riparian buffer ordinance
Biodiversity Conservation (BioDiv)	Maximize outcomes for biodiversity and conservation
Agricultural expansion (IncAg)	Maximizes agricultural outcomes to the greatest extent practical (given existing laws and policies)
Water quality improvement (WQual)	Maximizes Best Management Practices for water quality
Pre-settlement conditions (PreSet)	Mimics historical land uses in the basin (which is unrealistic relative to current economic uses) Generalized point of comparison for ecological potential

Forecasting Land Use Change: Six Scenarios

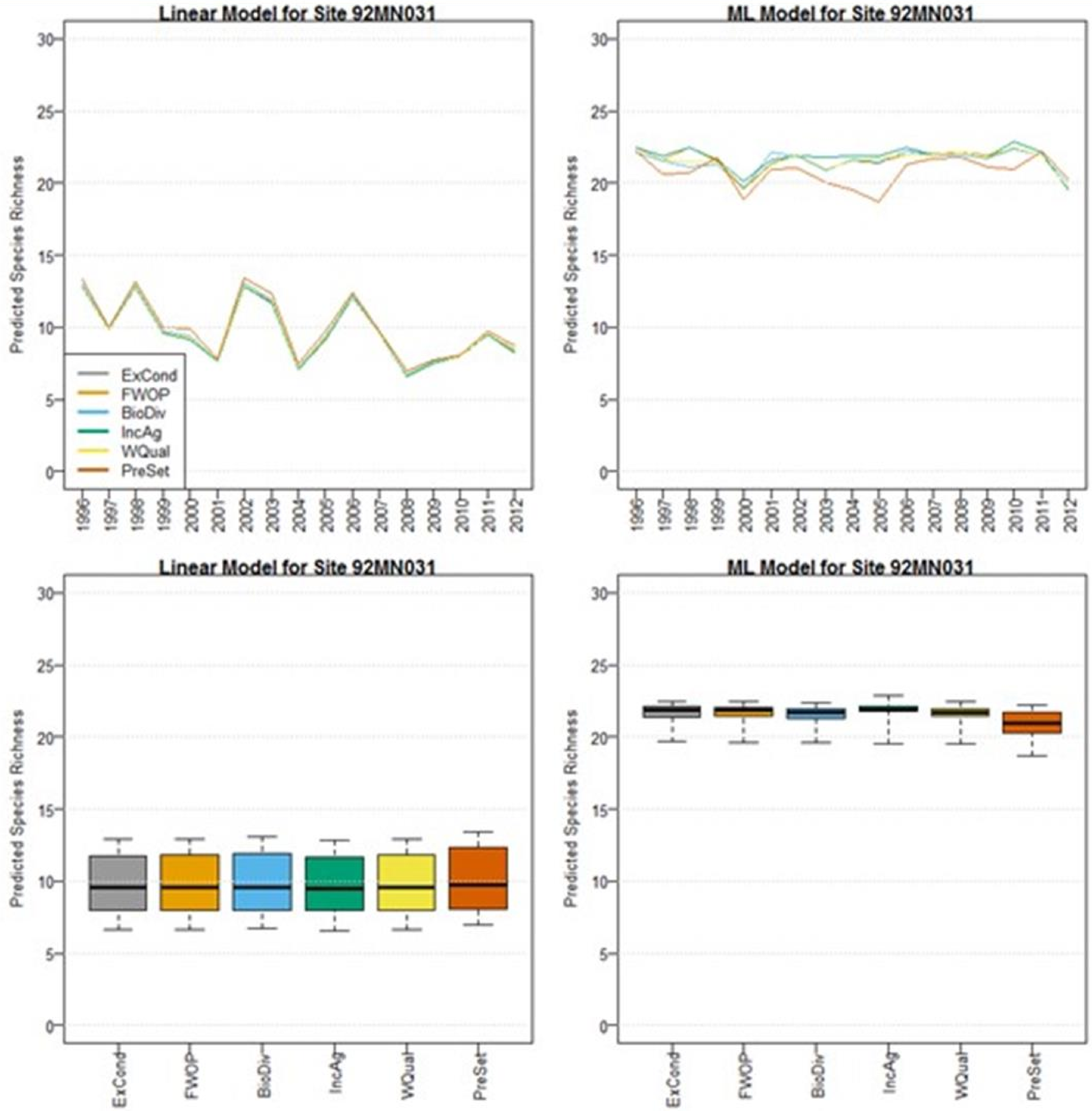
Seven Mile Creek Alternative Landscape Scenarios



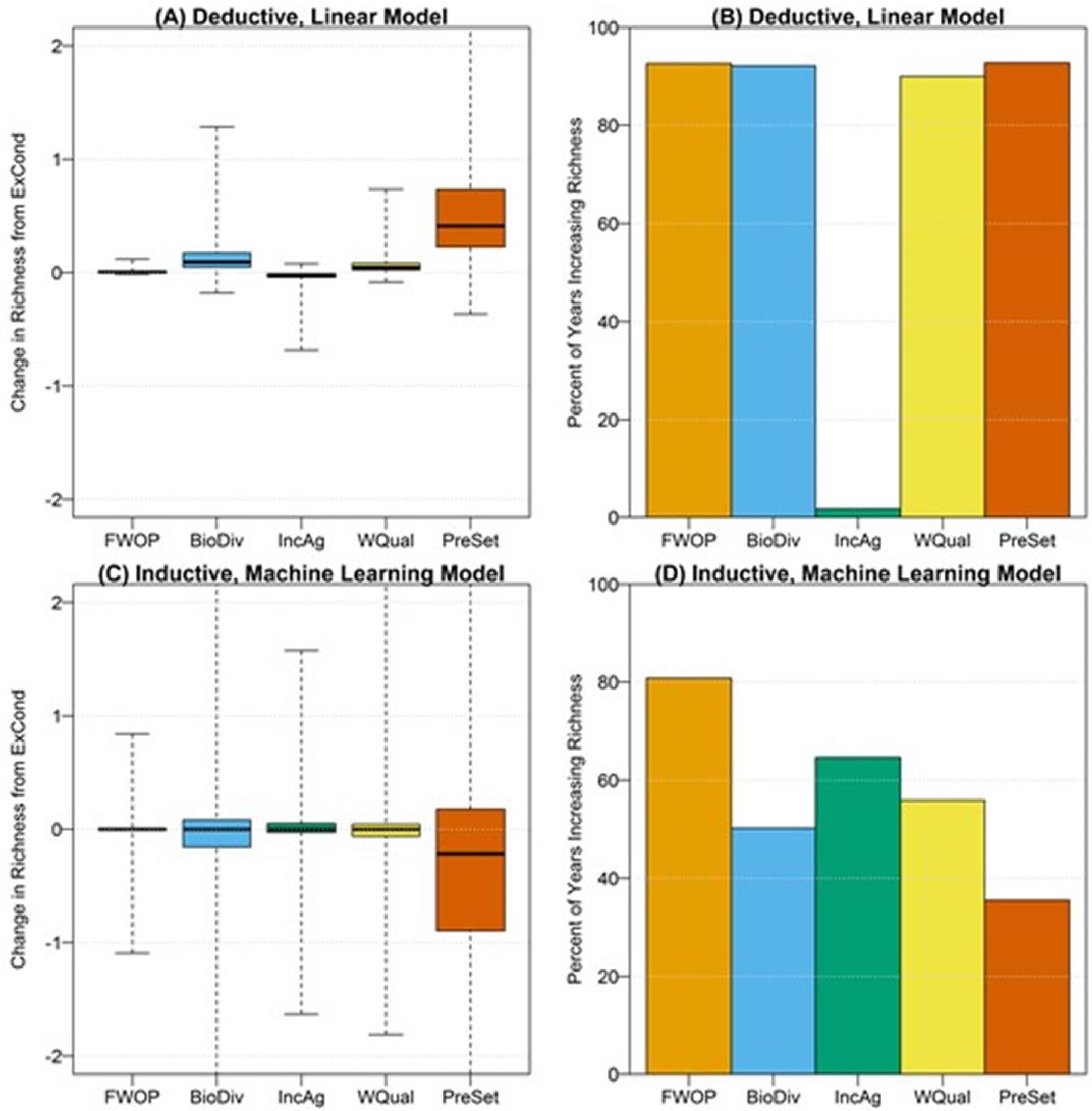
Model Application: Hydrologic Outcomes (Model-1)



Forecasting Land Use Change: Ecological Outcomes (Models 2/3)



Model Application: Ecological Outcomes (Models 2/3)



What did we learn in the Minnesota River Basin?

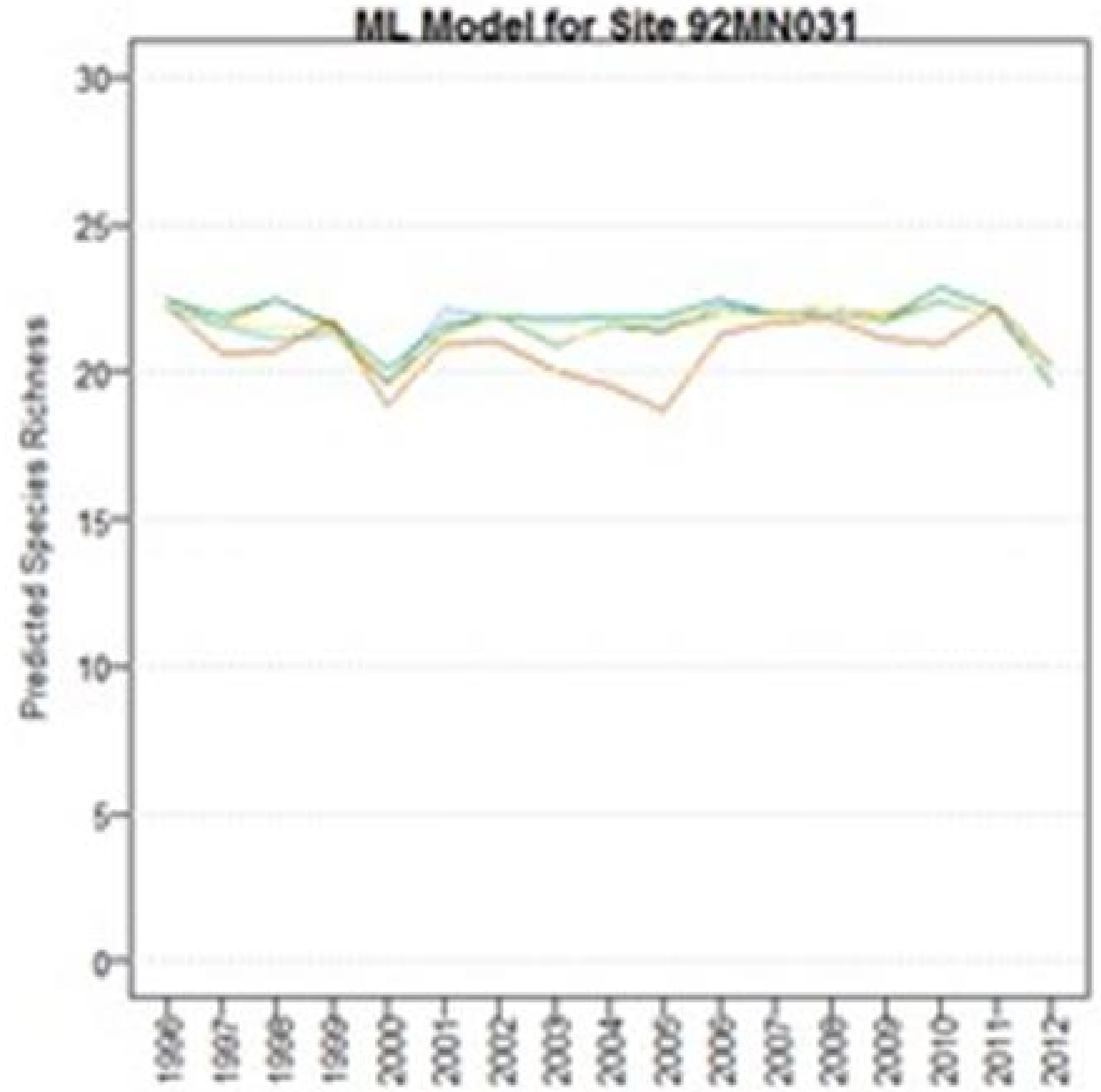
Land Use Scenario	Flow-Mean	Flow-CV	Flow-Skew	Flow-Kurt	Flow-AR1	Flow-SeaAmp	Flow-SeaPhase	Richness-Linear	Richness-ML
Future without actions	↓					↓	↑	↑	↑
Biodiversity	↓					↓	↑	↑	
Increased Agriculture	↑	↑				↑	↓	↓	↑
Water quality	↓	↓				↓	↑	↑	
Pre-Settlement	↓				↑	↓	↑	↑	↓

All changes relative to the Existing Condition
Only the changes >±10% are shown here

Reflections on Modeling

What did we learn about eflows?

- Response to central tendency or extreme?
  - Time series simulation was crucial to understanding ecological effects of land use
  - Still sorting through this use of the data…
- State variables (e.g., richness) are a good start
  - Rate variables would be better.
  - Need to couple spatially distributed data sets with long-term data sets at a few sites...



Alternative Ecological Metrics

Community metrics (e.g., IBI)

Taxa-specific response (e.g., occupancy)

Survival and reproduction rates

Biomass, growth, or production data

…

A brief philosophical tangent about competing models

Epistemology describes "a set of values concerning truth or validity"

- What type of knowledge is sought?
- What methods are appropriate?
- What kinds of data constitute evidence?
- Is replication necessary?
- What are the fundamental assumptions of your discipline?
- What bias is inherent in a given approach?
- How are questions asked?
- How are problems framed?
- …

What is "significance"?

Lay usage of significance: important or meaningful

Statistical significance: a low probability that observed measurements or values of interest occurred by chance.

Statistical, but not biological significance: experimental results indicate a finding, but biologically meaningless

USACE policy significance: institutional, public, or technical factors indicating potential "federal interest" in an ecosystem

Definition, questions, and examples from: Eigenbrode et al. (2007, *Bioscience*), Raymond et al. (2010, *J of Env Man*), and Donovan et al. (2015, *SAGE Open*)

What did we learn from competing model?

- Modern tools allow for development of many models
  - Don't have to prescribe a way of thinking (epistemology)
- Similar variables emerged from all analyses
  - Maybe we understand the key driving variables
- Weight-of-evidence can lead to confidence and robustness in management actions
  - Do we make the same recommendation regardless of model choice?
  - Requires condensing diverse lines of evidence efficiently

Collaborators & Acknowledgements

Chuck Theiling & Michael Dougherty

Scott Niemela & John Sandberg

USACE St. Paul District

Contact Information

Kyle McKay

601-415-7160

kyle.mckay@usace.army.mil