How Do You Validate a Watershed Model Without Calibration Data?

Brian E. Haggard
Arkansas Water Resources Center
University of Arkansas System
Division of Agriculture

OUTLINE

- My interest in watershed models
- The pyramid of watershed modeling
- How do you validate a watershed model without calibration data?
  - regionalization of watershed model parameters
  - use monitoring data collected outside modeling time frame
- How have we expanded monitoring efforts in northwest Arkansas?
- The use of models… defines data needs
- Last, I will keep it short to get us back on track or out early!

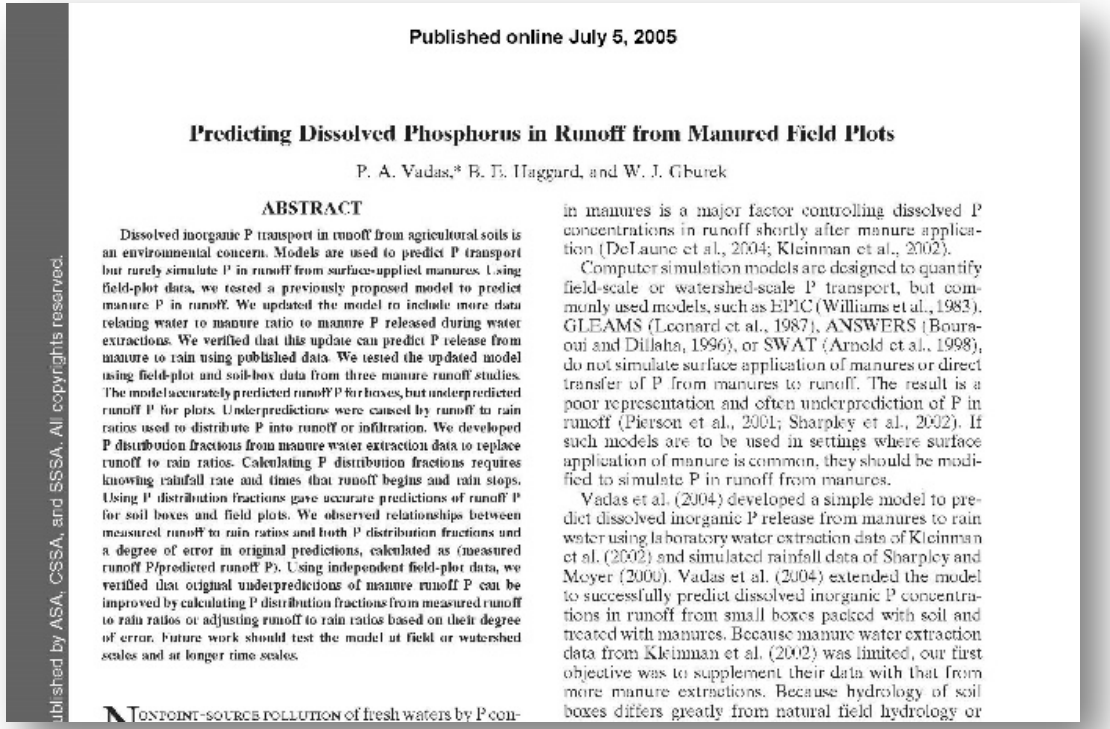
1. My interest in watershed models…

- First, I am not a watershed modeler *per se* but I did stay at a Marriott
- But, I like to:
  - find weaknesses in watershed models, or poke holes in them
  - collaborate with modelers to ‘fix’ watershed models
- So, I will go through some collaborative examples of ‘fixing’ watershed models
  - surface manure applications
  - in-stream nutrient cycling

At times, you might have to ‘trick’ watershed models into simulating things…

SURFACE APPLIED MANURES

Journal of Environmental Quality



Published online July 5, 2005

Predicting Dissolved Phosphorus in Runoff from Manured Field Plots

P. A. Vadas,* B. E. Haggard, and W. J. Gburek

ABSTRACT

Dissolved inorganic P transport in runoff from agricultural soils is an environmental concern. Models are used to predict P transport but rarely simulate P in runoff from surface-applied manures. Using field-plot data, we tested a previously proposed model to predict manure P in runoff. We updated the model to include more data relating water to manure ratio to manure P released during water extractions. We verified that this update can predict P release from manure to rain using published data. We tested the updated model using field-plot and soil-box data from three manure runoff studies. The model accurately predicted runoff P for boxes, but underpredicted runoff P for plots. Underpredictions were caused by runoff to rain ratios used to distribute P into runoff or infiltration. We developed P distribution fractions from manure water extraction data to replace runoff to rain ratios. Calculating P distribution fractions requires knowing rainfall rate and times that runoff begins and rain stops. Using P distribution fractions gave accurate predictions of runoff P for soil boxes and field plots. We observed relationships between measured runoff to rain ratios and both P distribution fractions and a degree of error in original predictions, calculated as (measured runoff P/predicted runoff P). Using independent field-plot data, we verified that original underpredictions of manure runoff P can be improved by calculating P distribution fractions from measured runoff to rain ratios or adjusting runoff to rain ratios based on their degree of error. Future work should test the model at field or watershed scales and at longer time scales.

NONPOINT-SOURCE POLLUTION of fresh waters by P con-

in manures is a major factor controlling dissolved P concentrations in runoff shortly after manure application (DeLaune et al., 2004; Kleinman et al., 2002).

Computer simulation models are designed to quantify field-scale or watershed-scale P transport, but commonly used models, such as EPIC (Williams et al., 1983), GLEAMS (Leonard et al., 1987), ANSWERS (Bouraoui and Dillaha, 1996), or SWAT (Arnold et al., 1998), do not simulate surface application of manures or direct transfer of P from manures to runoff. The result is a poor representation and often underprediction of P in runoff (Pierson et al., 2001; Sharpley et al., 2002). If such models are to be used in settings where surface application of manure is common, they should be modified to simulate P in runoff from manures.

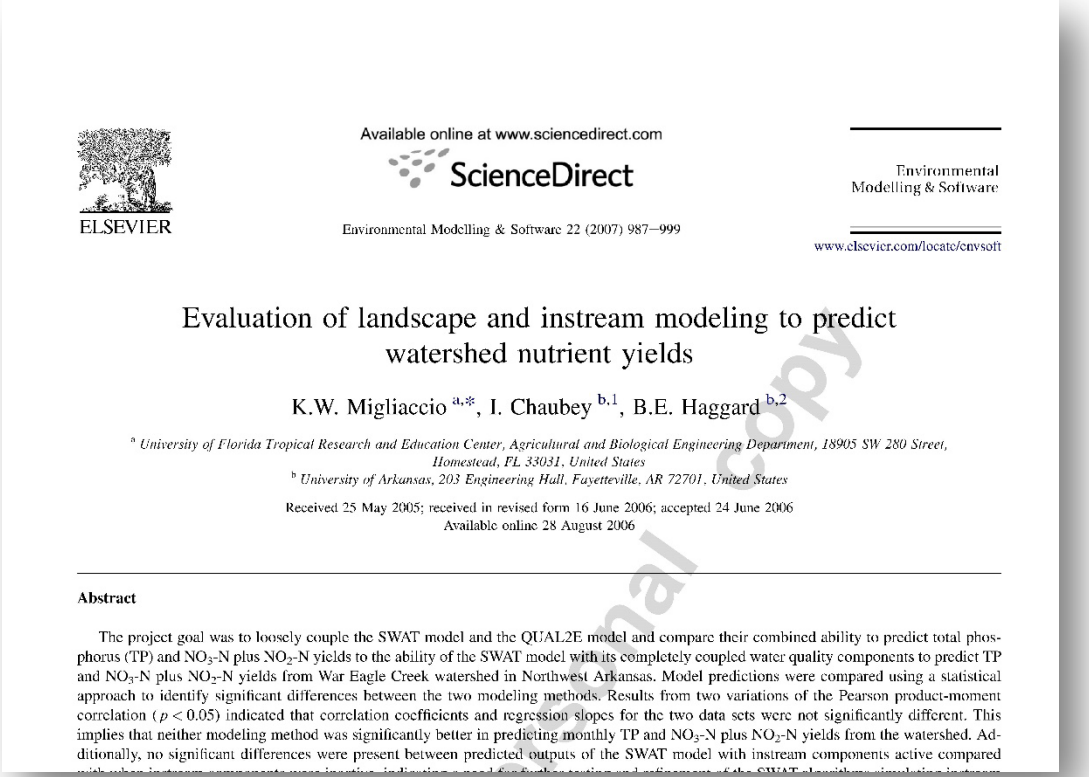
Vadas et al. (2004) developed a simple model to predict dissolved inorganic P release from manures to rain water using laboratory water extraction data of Kleinman et al. (2002) and simulated rainfall data of Sharpley and Moyer (2000). Vadas et al. (2004) extended the model to successfully predict dissolved inorganic P concentrations in runoff from small boxes packed with soil and treated with manures. Because manure water extraction data from Kleinman et al. (2002) was limited, our first objective was to supplement their data with that from more manure extractions. Because hydrology of soil boxes differs greatly from natural field hydrology or

ublished by ASA, CSSA, and SSSA. All copyrights reserved.

Other times, you have to evaluate watershed model components…

IN-STREAM NUTRIENT CYCLING

How does QUAL2E function in SWAT?



ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

Environmental Modelling & Software 22 (2007) 987–999

Environmental Modelling & Software

www.elsevier.com/locate/envsoft

Evaluation of landscape and instream modeling to predict watershed nutrient yields

K.W. Migliaccio [a,*], I. Chaubey [b,1], B.E. Haggard [b,2]

[a] University of Florida Tropical Research and Education Center, Agricultural and Biological Engineering Department, 18905 SW 280 Street, Homestead, FL 33031, United States
[b] University of Arkansas, 203 Engineering Hall, Fayetteville, AR 72701, United States

Received 25 May 2005; received in revised form 16 June 2006; accepted 24 June 2006
Available online 28 August 2006

Abstract

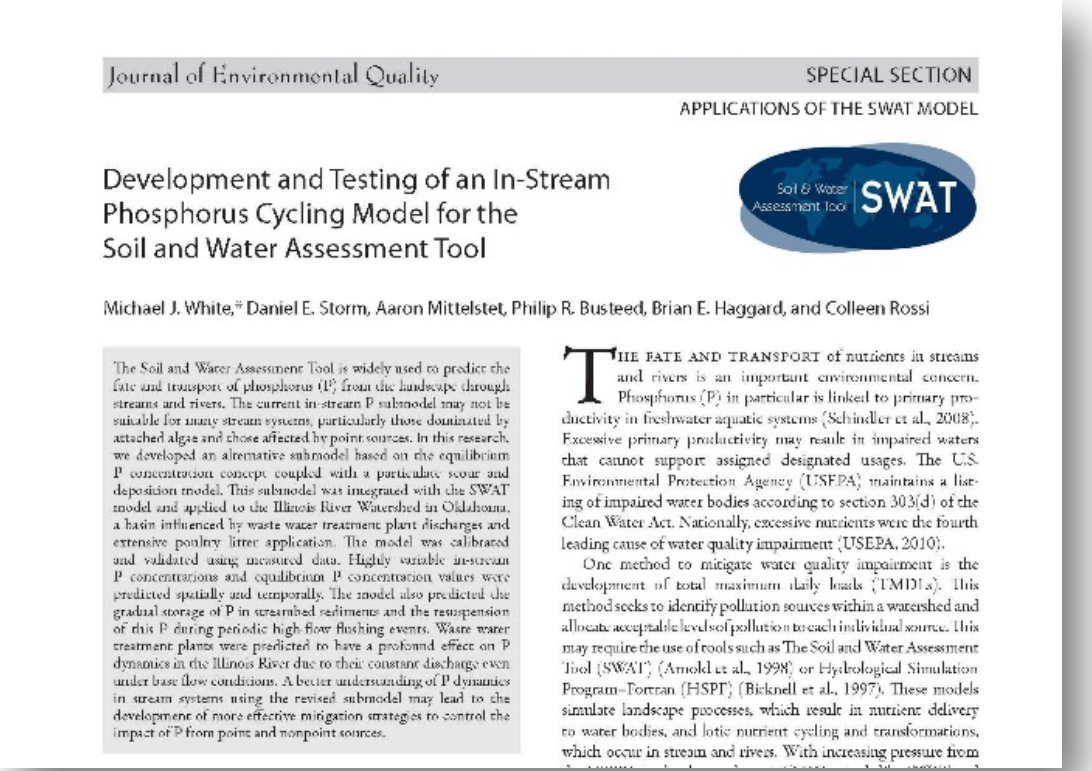
The project goal was to loosely couple the SWAT model and the QUAL2E model and compare their combined ability to predict total phosphorus (TP) and NO_3-N plus NO_2-N yields to the ability of the SWAT model with its completely coupled water quality components to predict TP and NO_3-N plus NO_2-N yields from War Eagle Creek watershed in Northwest Arkansas. Model predictions were compared using a statistical approach to identify significant differences between the two modeling methods. Results from two variations of the Pearson product-moment correlation ($p < 0.05$) indicated that correlation coefficients and regression slopes for the two data sets were not significantly different. This implies that neither modeling method was significantly better in predicting monthly TP and NO_3-N plus NO_2-N yields from the watershed. Additionally, no significant differences were present between predicted outputs of the SWAT model with instream components active compared

Personal copy

Then, you might have to develop external modules for watershed models…

PHOSPHORUS-SEDIMENT INTERACTIONS

Stream Phosphorus Cycling Model SWAT



Journal of Environmental Quality SPECIAL SECTION
APPLICATIONS OF THE SWAT MODEL

Development and Testing of an In-Stream Phosphorus Cycling Model for the Soil and Water Assessment Tool

Michael J. White,* Daniel E. Storm, Aaron Mittelstet, Philip R. Busteed, Brian E. Haggard, and Colleen Rossi

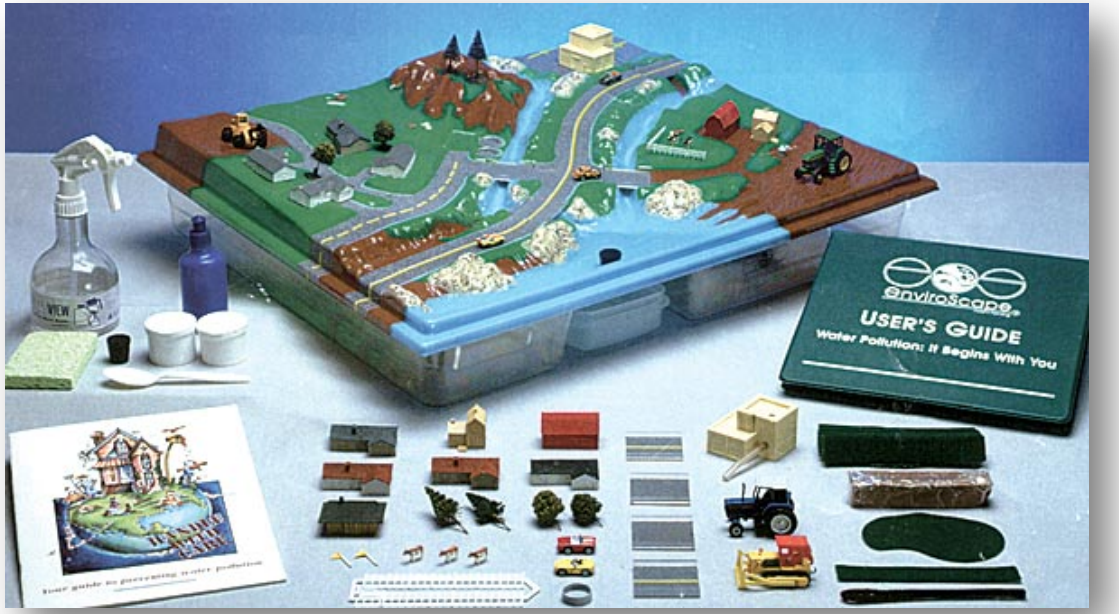
The Soil and Water Assessment Tool is widely used to predict the fate and transport of phosphorus (P) from the landscape through streams and rivers. The current in-stream P submodel may not be suitable for many stream systems, particularly those dominated by attached algae and those affected by point sources. In this research, we developed an alternative submodel based on the equilibrium P concentration concept coupled with a particulate scour and deposition model. This submodel was integrated with the SWAT model and applied to the Illinois River Watershed in Oklahoma, a basin influenced by waste water treatment plant discharges and extensive poultry litter application. The model was calibrated and validated using measured data. Highly variable in-stream P concentrations and equilibrium P concentration values were predicted spatially and temporally. The model also predicted the gradual storage of P in streambed sediments and the resuspension of this P during periodic high-flow flushing events. Waste water treatment plants were predicted to have a profound effect on P dynamics in the Illinois River due to their constant discharge even under base flow conditions. A better understanding of P dynamics in stream systems using the revised submodel may lead to the development of more effective mitigation strategies to control the impact of P from point and nonpoint sources.

The fate and transport of nutrients in streams and rivers is an important environmental concern. Phosphorus (P) in particular is linked to primary productivity in freshwater aquatic systems (Schindler et al., 2008). Excessive primary productivity may result in impaired waters that cannot support assigned designated usages. The U.S. Environmental Protection Agency (USEPA) maintains a listing of impaired water bodies according to section 303(d) of the Clean Water Act. Nationally, excessive nutrients were the fourth leading cause of water quality impairment (USEPA, 2010).

One method to mitigate water quality impairment is the development of total maximum daily loads (TMDLs). This method seeks to identify pollution sources within a watershed and allocate acceptable levels of pollution to each individual source. This may require the use of tools such as The Soil and Water Assessment Tool (SWAT) (Arnold et al., 1998) or Hydrological Simulation Program–Fortran (HSPF) (Bicknell et al., 1997). These models simulate landscape processes, which result in nutrient delivery to water bodies, and lotic nutrient cycling and transformations, which occur in stream and rivers. With increasing pressure from

So, I like to find weaknesses or problems in watershed models… then help ‘fix’ them.

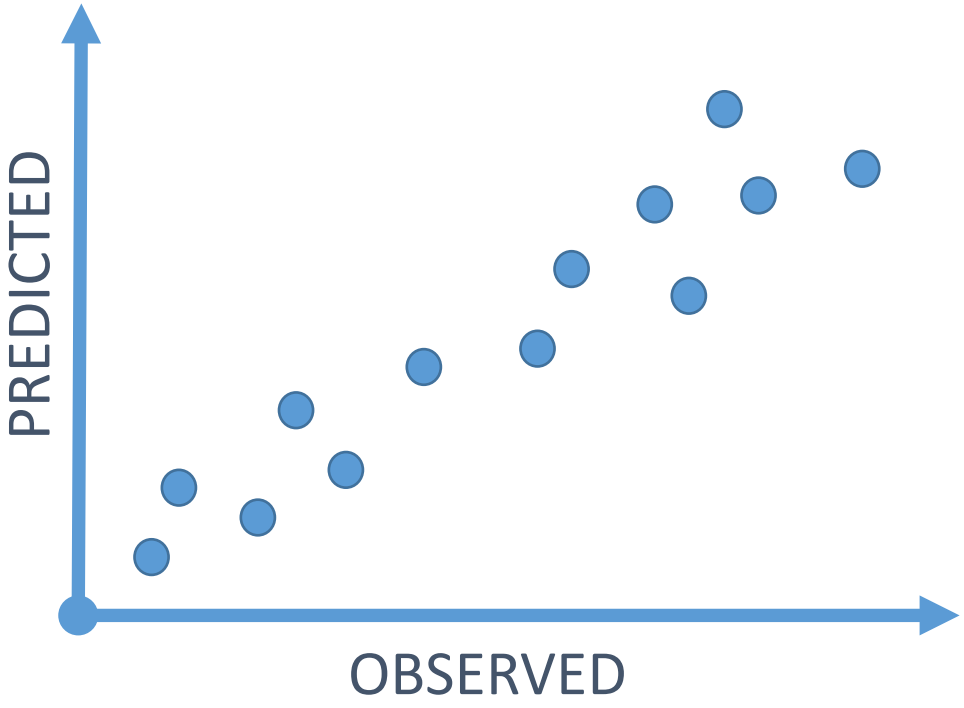
ALL MODELS ARE WRONG, BUT SOME ARE USEFUL – George Box



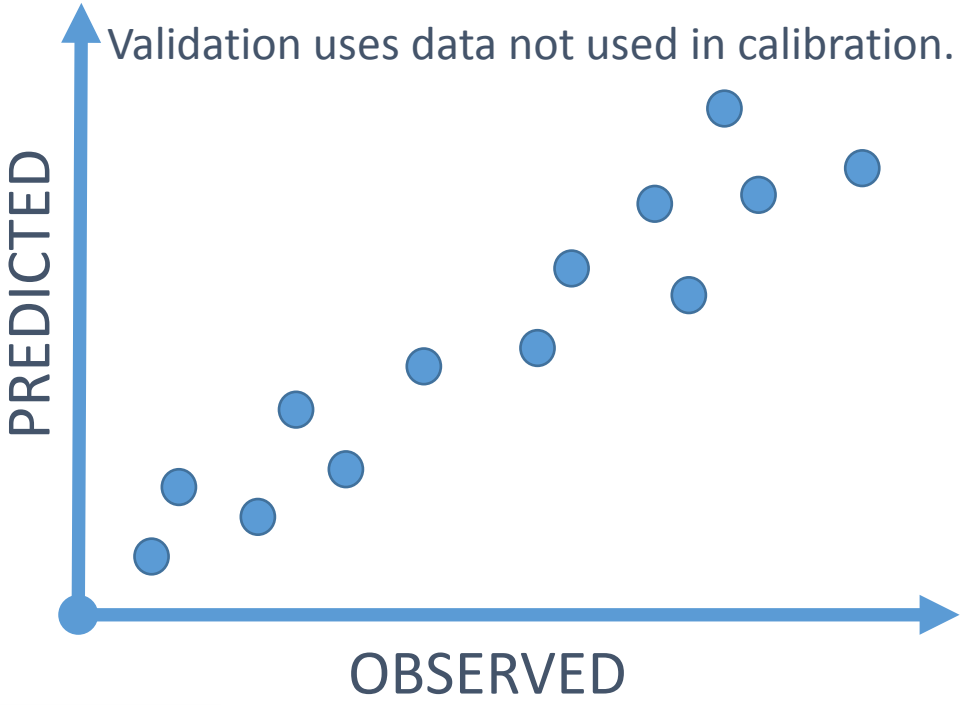
Collaborators have included:

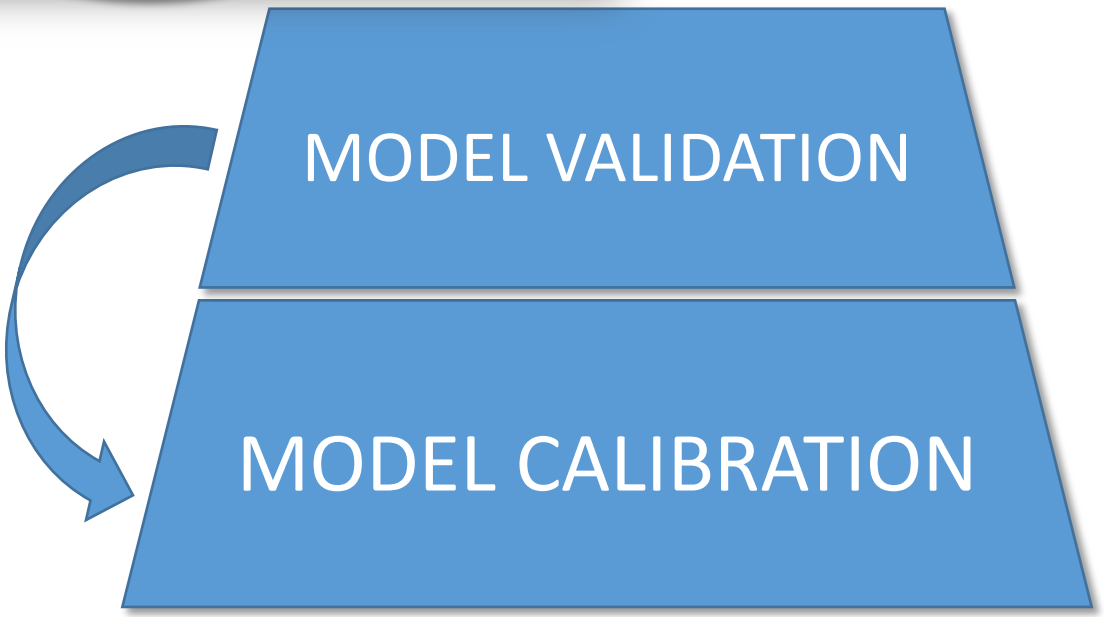
- Indrajeet Chaubey (SWAT/QUAL2E)
- Reed Green (CE-QUAL-W2)
- Kati Migliaccio (SWAT/QUAL2E)
- Dan Storm (SWAT/BATHTUB)
- Pete Vadas (SWAT)
- Mike White (SWAT)

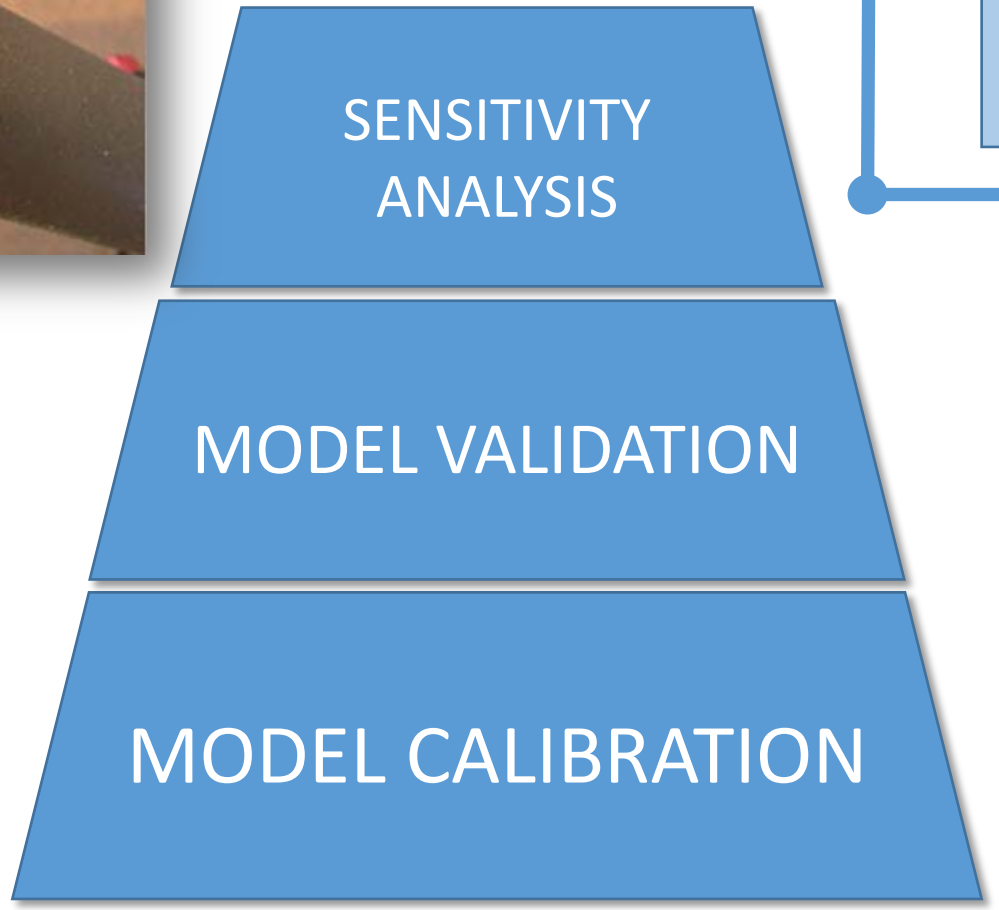
2. The watershed modeling pyramid…

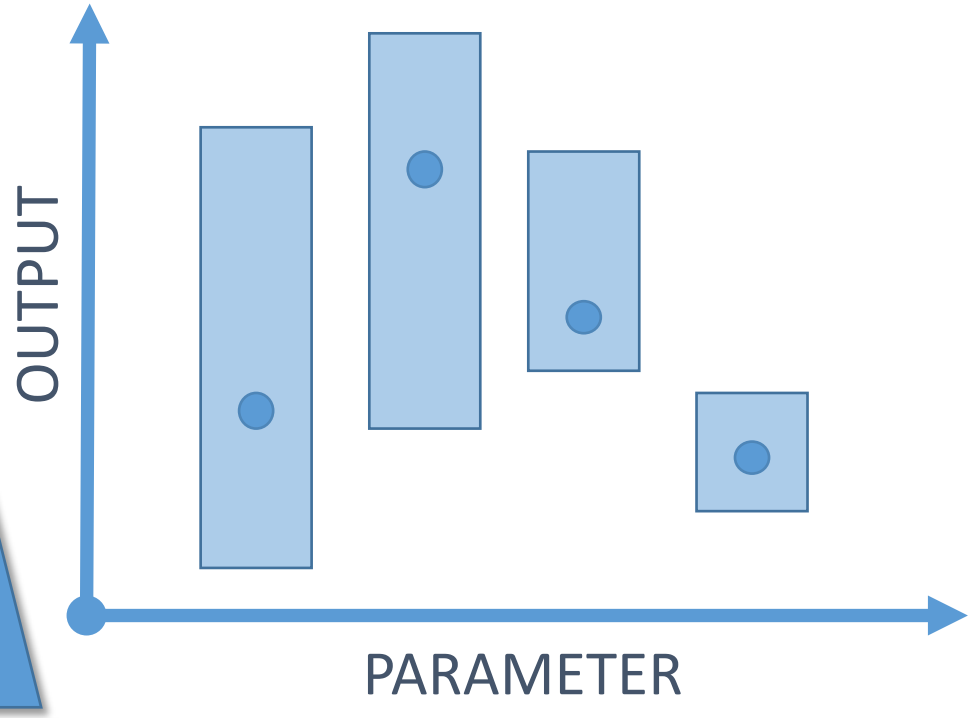


MODEL CALIBRATION

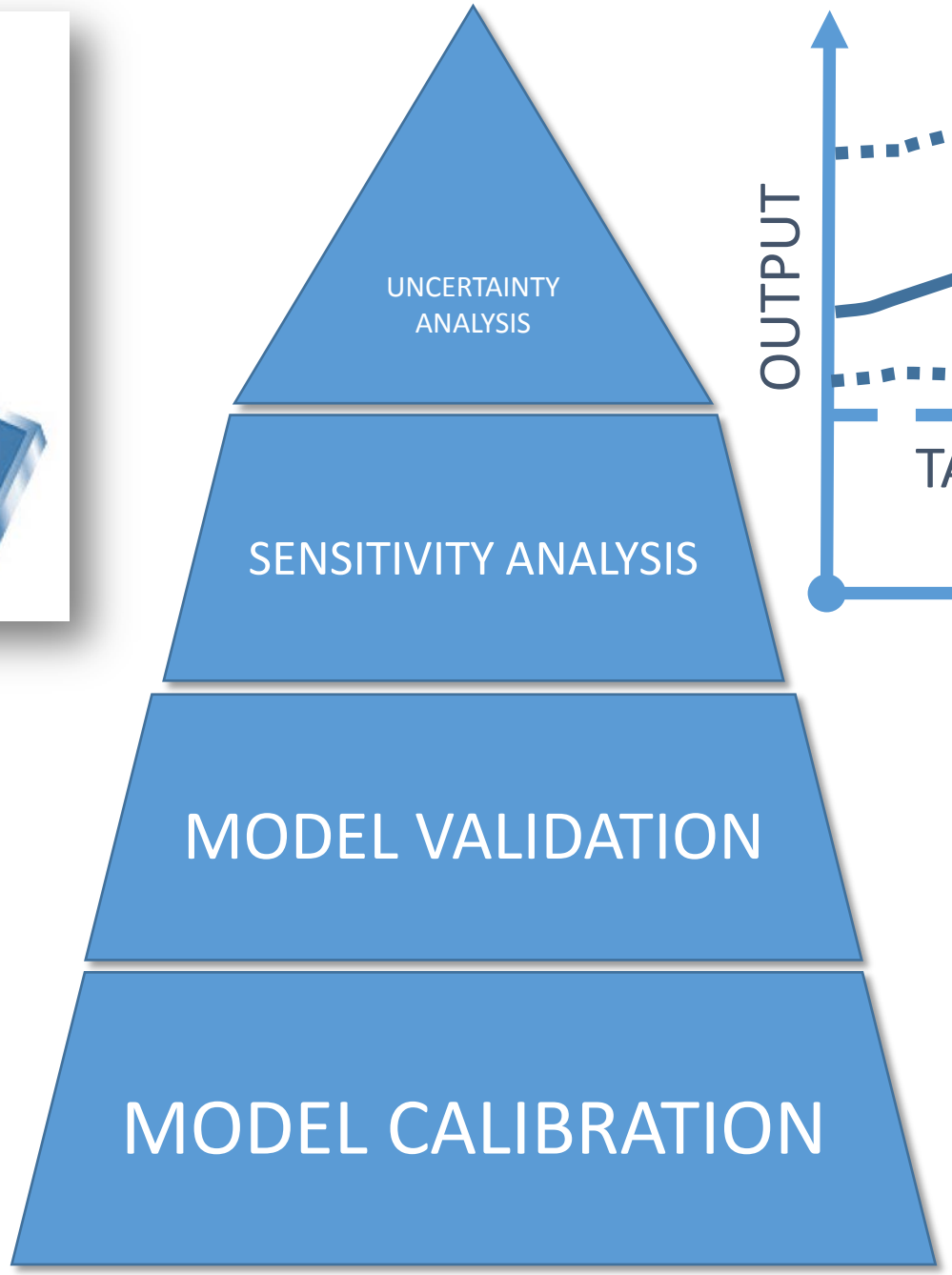

Validation uses data not used in calibration.
PREDICTED
OBSERVED

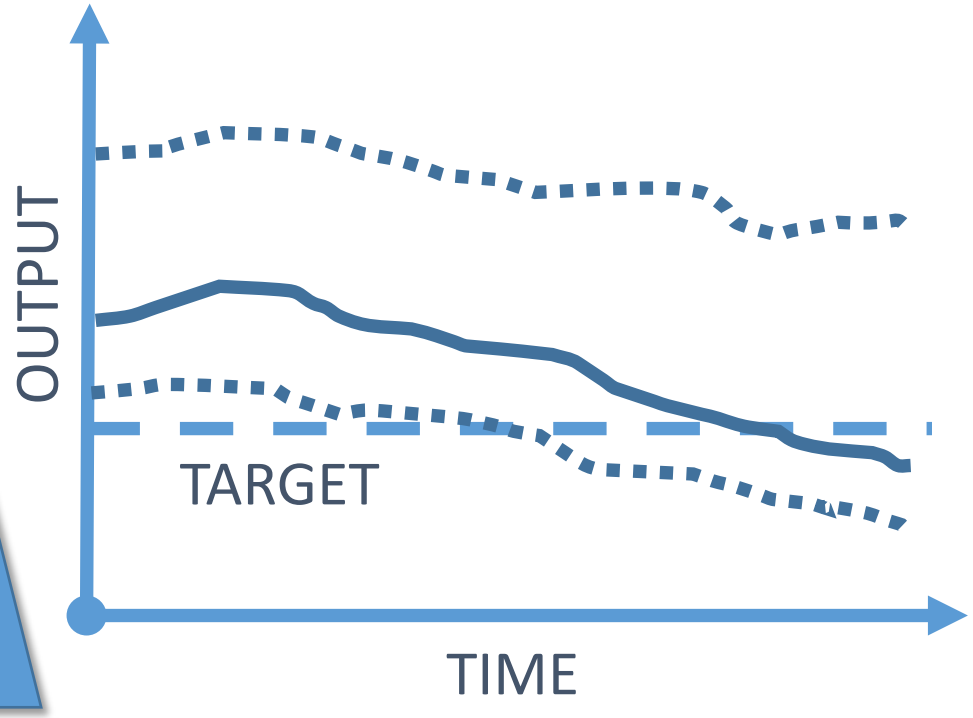

MODEL VALIDATION
MODEL CALIBRATION


SENSITIVITY ANALYSIS
MODEL VALIDATION
MODEL CALIBRATION


OUTPUT
PARAMETER

Success


UNCERTAINTY ANALYSIS
SENSITIVITY ANALYSIS
MODEL VALIDATION
MODEL CALIBRATION

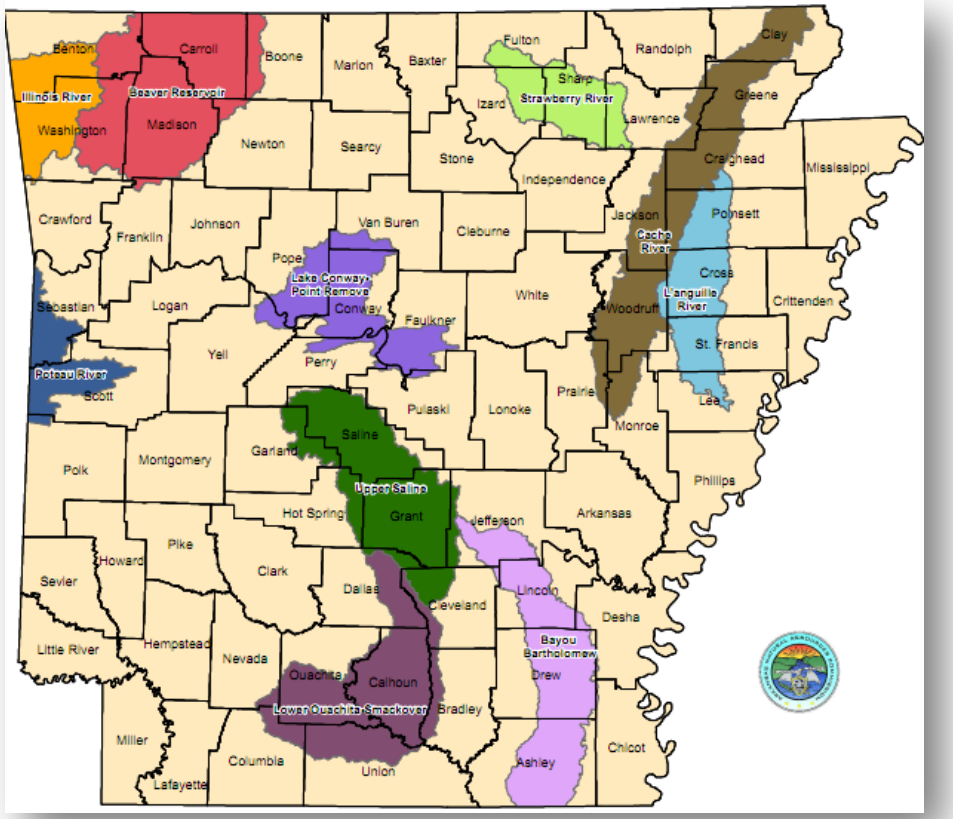

OUTPUT
TARGET
TIME

3. How do you calibrate a watershed model without calibration data?

REGIONALIZATION OF WATERSHED MODEL PARAMETERS

- Maybe, better termed generalization of parameters
- Thought is that model parameters are related to watershed characteristics
  - conceptually, makes sense…
- This process is relatively successful for hydrology
  - maybe, less so for N, P and sediment

But, we need to be able to model places where we don't have data – right?



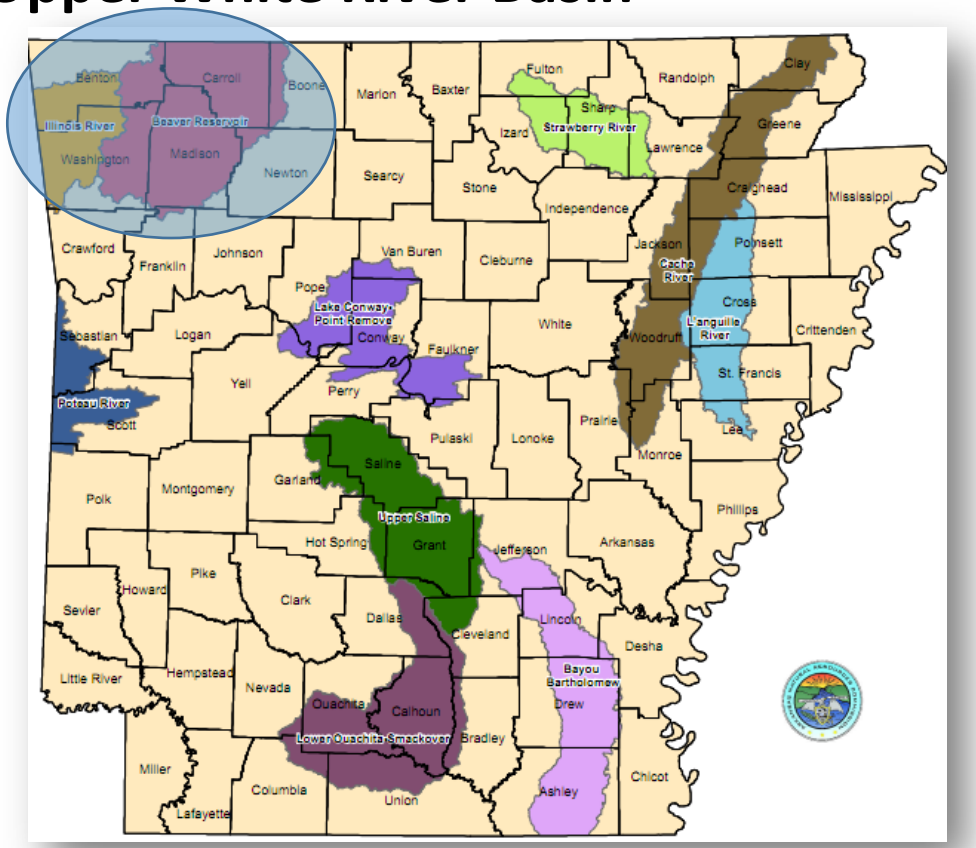
In Arkansas, one state agency wanted to use SWAT to help prioritize subwatersheds…

THE REASON WAS TO HELP THIS AGENCY TARGET ITS 319 FUNDS

- The first two watersheds were modeled in 2009, and these are:
  - well-known
  - data rich, relatively speaking
  - been modeled multiple times
- The SWAT models were calibrated and validated…
  - HUC 8 or 10 level
- But, the subwatersheds were prioritized at the HUC 12 level…
  - very little available data

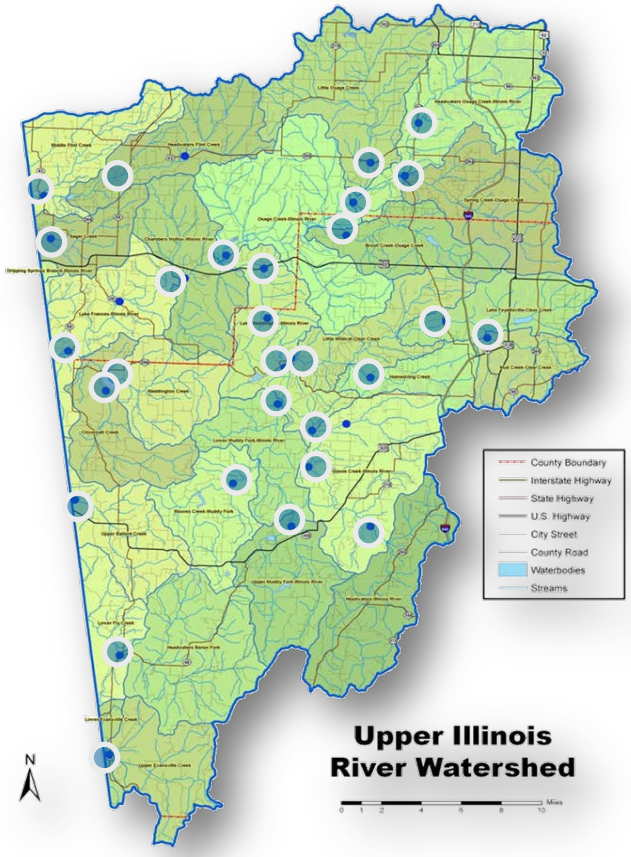
Upper Illinois River Watershed

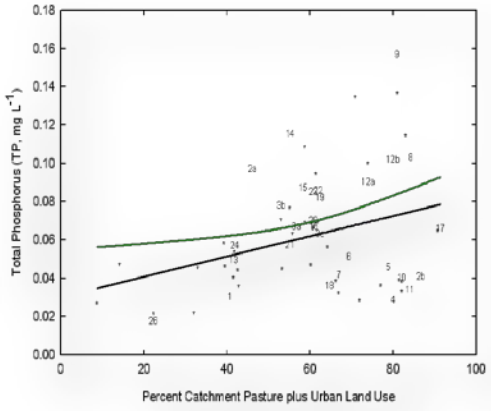
Upper White River Basin

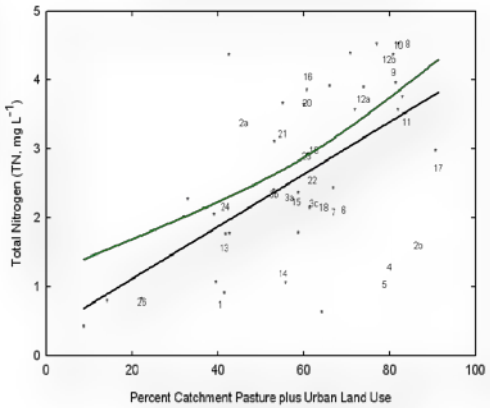


Or, some thought HUC 12 level data was not available… but there was some monitoring.

THERE WAS A SEPARATE PROJECT, MONITORING WQ AT HUC 12 OUTLETS







But, the first round of priorities did not match between SWAT and monitoring.

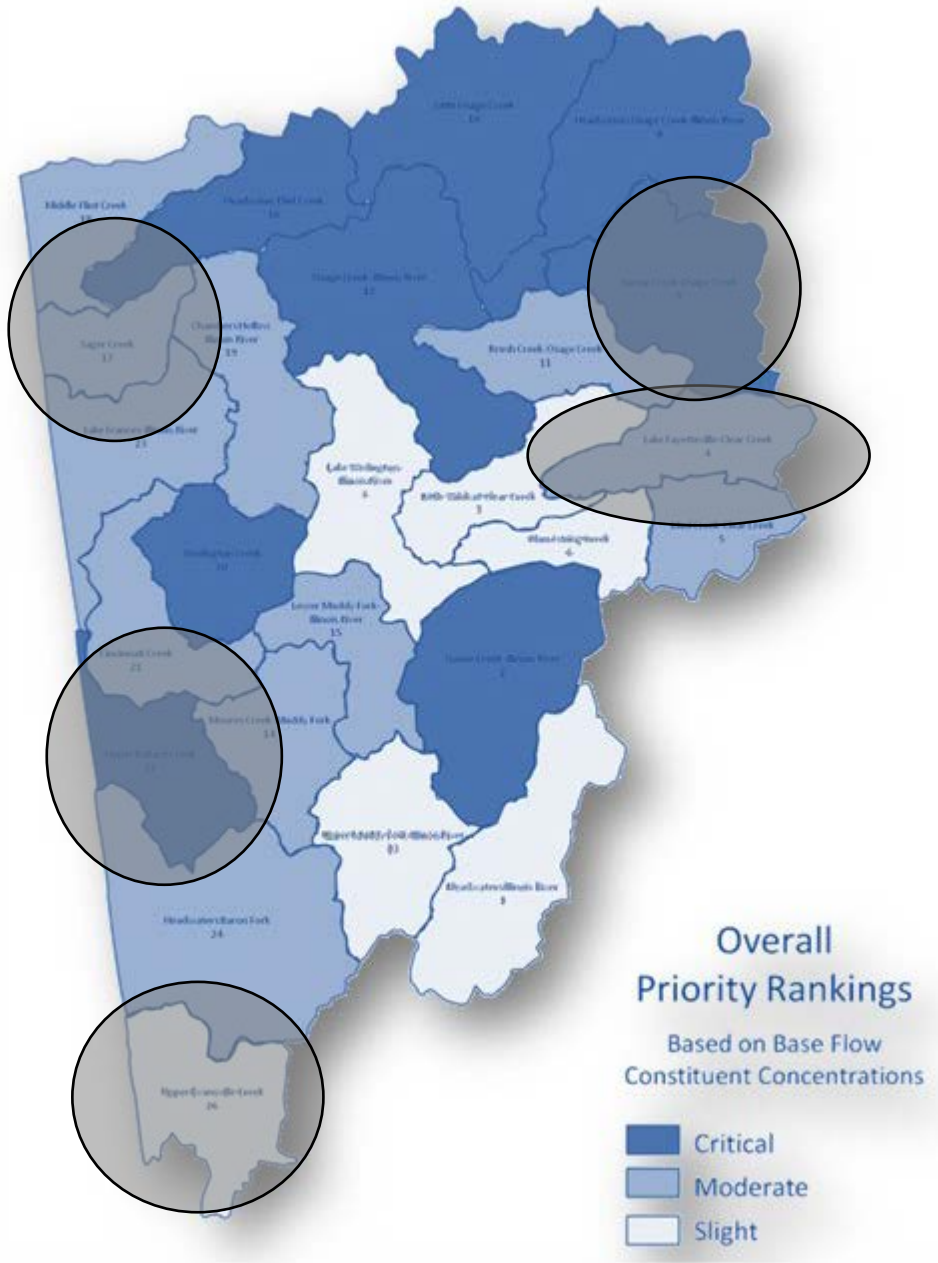
- For example, the HUC 12s:
  - dominated effluent discharge were not P priorities
  - with mostly forest were not low priorities (N and sediment)
- Phrases that agitate watershed model skeptics include:
  - model has 10 site-yrs of data (at one site!), where has monitoring one year
    - does 29 sites then equal 29 site-yrs?
  - There is uncertainty in monitoring data (based on Harmel et al., 206)
    - doesn't this propagate even more uncertainty in model output

So, this plus additional (available) data was used to re-calibrate the model.

THE COMPARISON WAS STILL NOT PERFECT, BUT MUCH BETTER (80% AGREEMENT).

- Let's remember:
  - monitoring priorities based on base flow conditions, which were related to storm events
  - modeling focused on nonpoint sources, e.g. runoff loads
- But, this little exercise combined monitoring and modeling in next project, and then the next…

PRIORITIES BASED ON WQ DATA



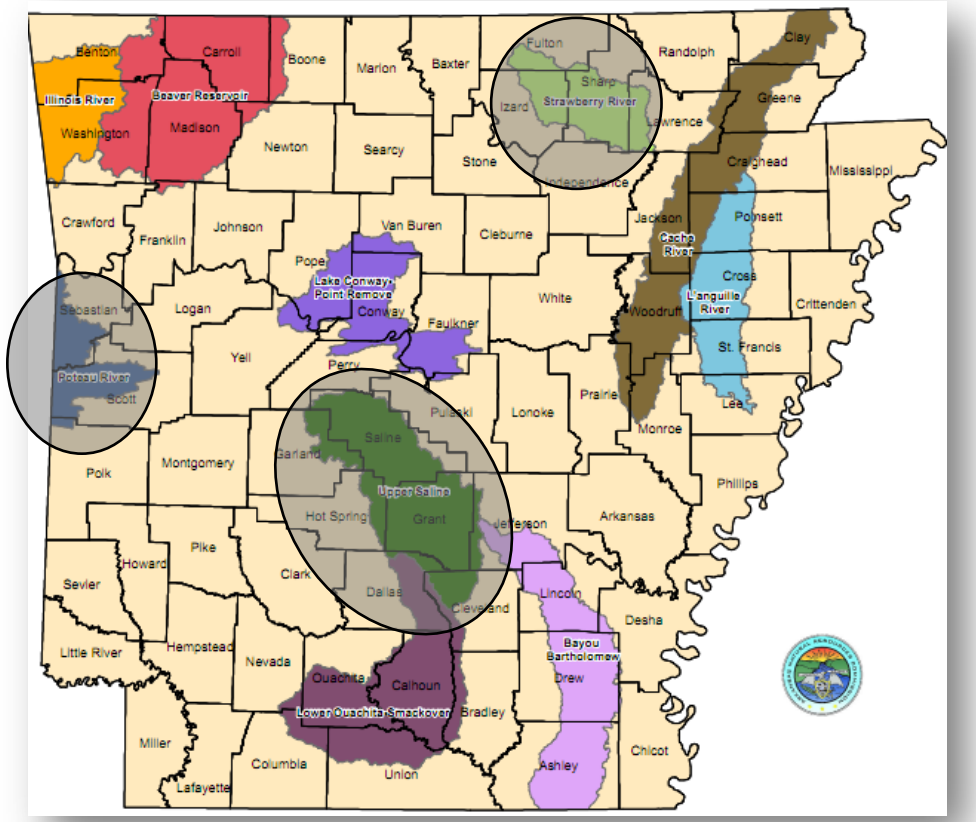
The next watersheds were more data limited?

- There was at least one (up to 3) USGS discharge site per HUC 12
  - although one was stage only, which required developing a rating curve
- Only one watershed had any available WQ data
  - during the modeling period
- Model inputs and parameters were generalized from existing models
- So, we decided to monitor WQ at the same time watershed model was being developed.

POTEAU RIVER WATERSHED

UPPER STRAWBERRY WATERSHED

UPPER SALINE WATERSHED



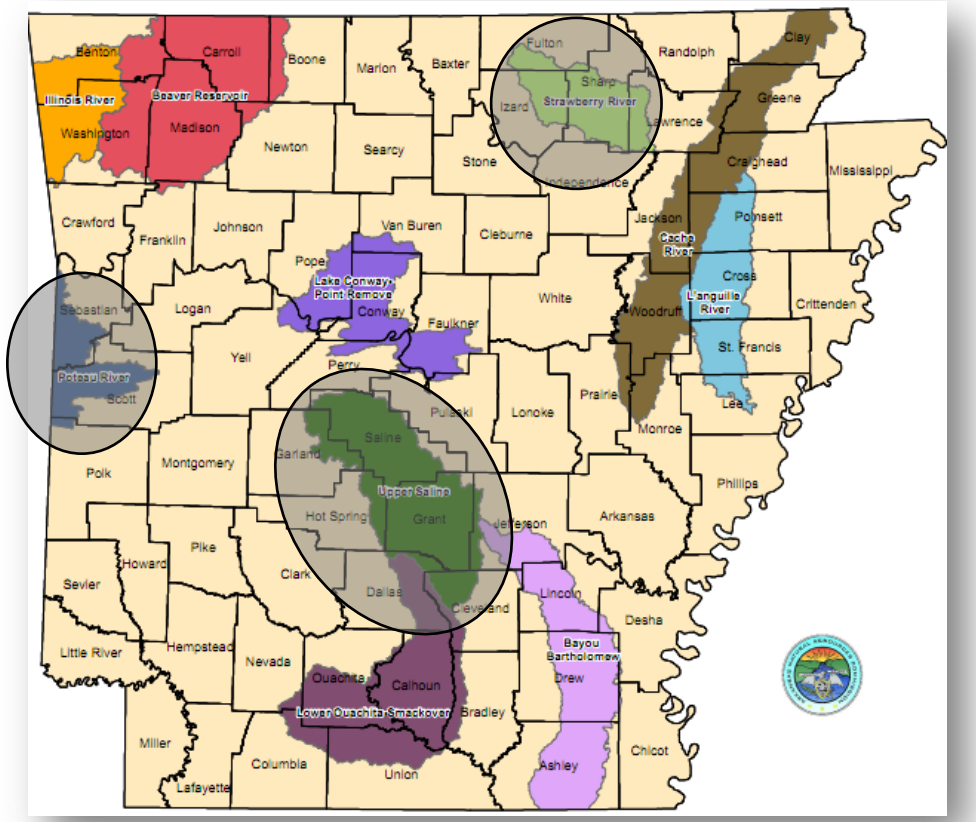
The monitoring included:

- Collecting 36 samples at existing USGS discharge...
  - targeting rise, peak and fall
  - LOADEST Software
  - annual and monthly loads
    - N, P and sediment
    - projecting forward and backward
- Collecting water samples across at least 20 HUC 12 outlets targeting…
  - land use and management gradient
  - headwaters to large rivers
  - seasonal variation (i.e., monthly)

POTEAU RIVER WATERSHED

UPPER STRAWBERRY WATERSHED

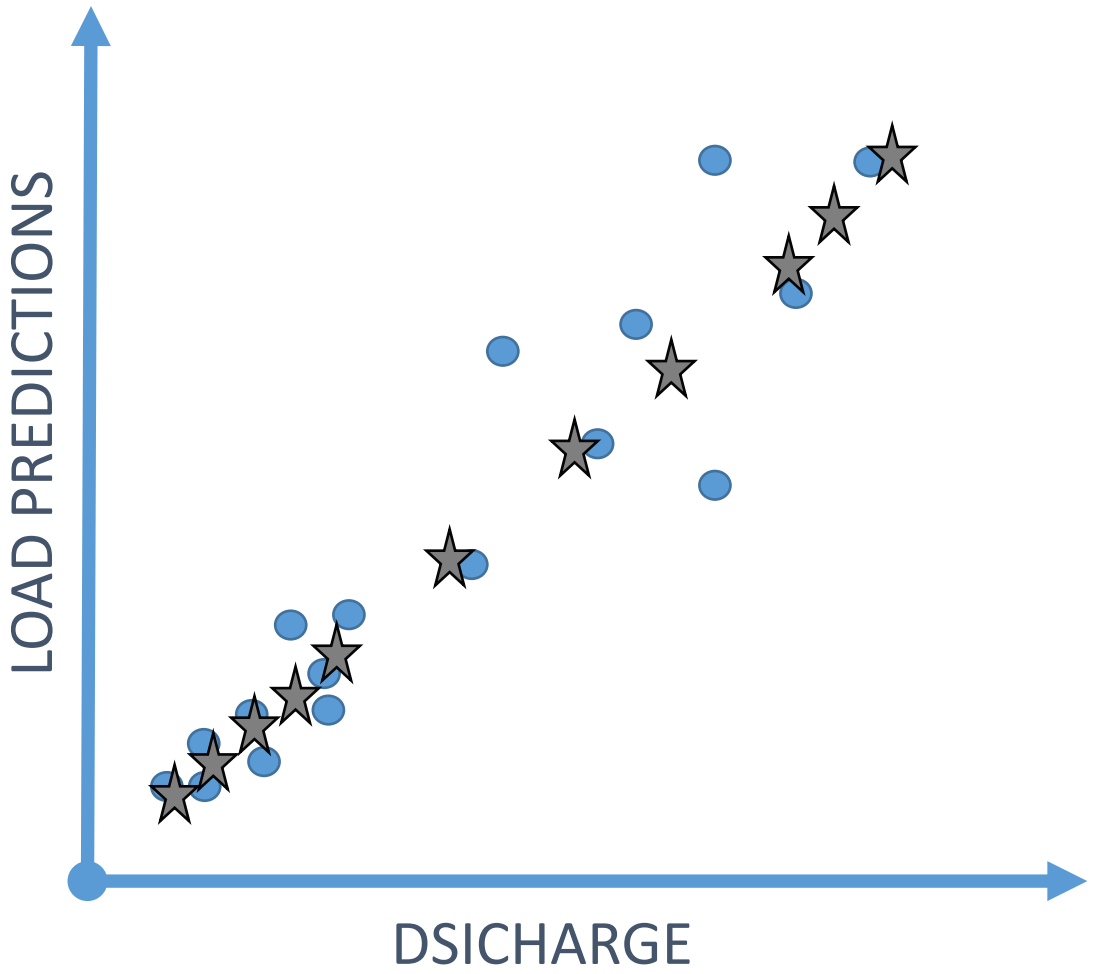
UPPER SALINE WATERSHED



Alright, so how did we use this data to validate the watershed model, i.e. SWAT?

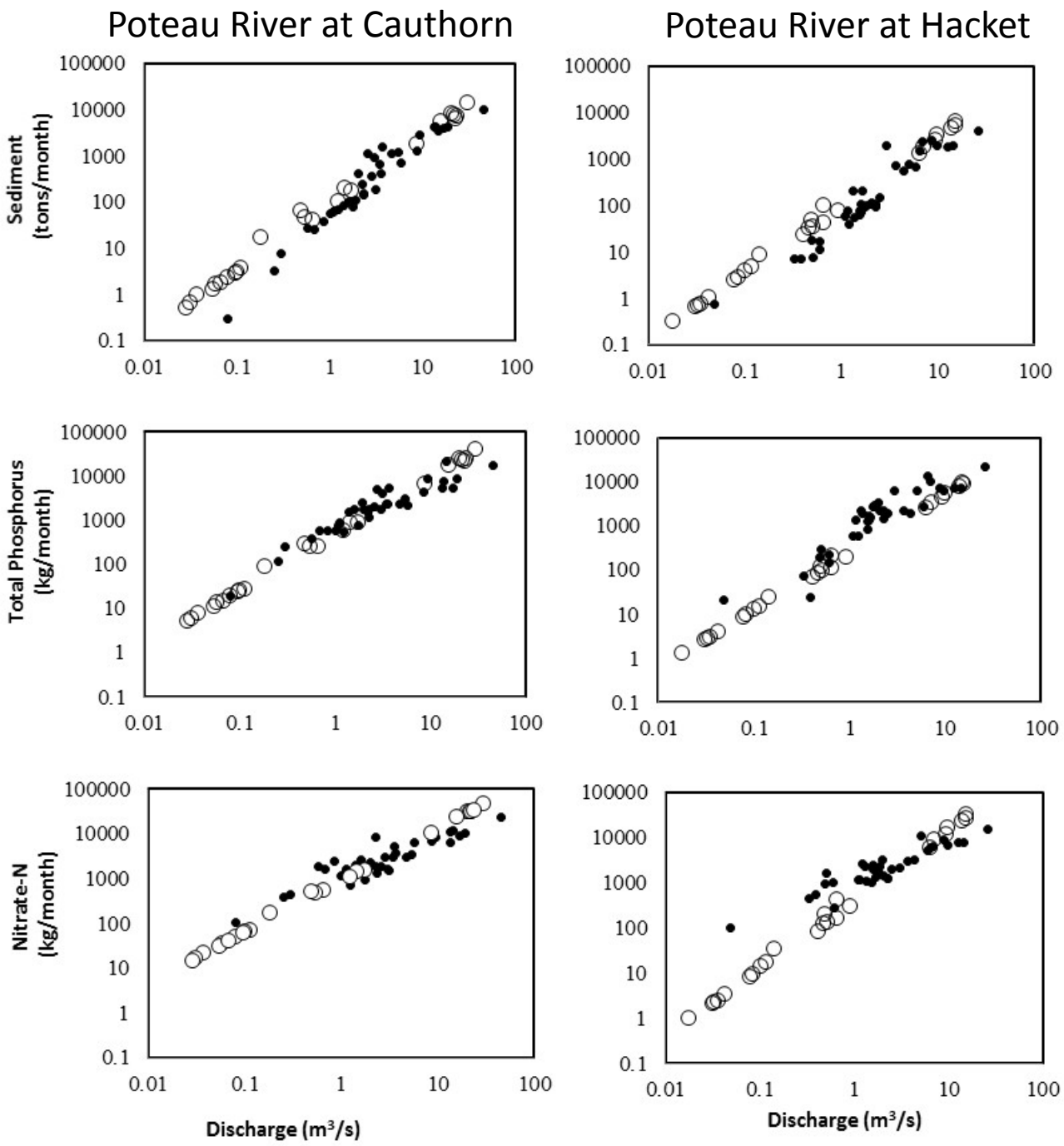
MAYBE, BETTER STATED TO IMPROVE OUR CONFIDENCE IN MODEL OUTPUT

- We compared loads, but wait the modeling and monitoring period did not overlap…
- We assumed the relationship between load and flow was stationary (*Hirsch says bad idea*)
- We wanted to see the estimated loads and model output falling on the same trend line



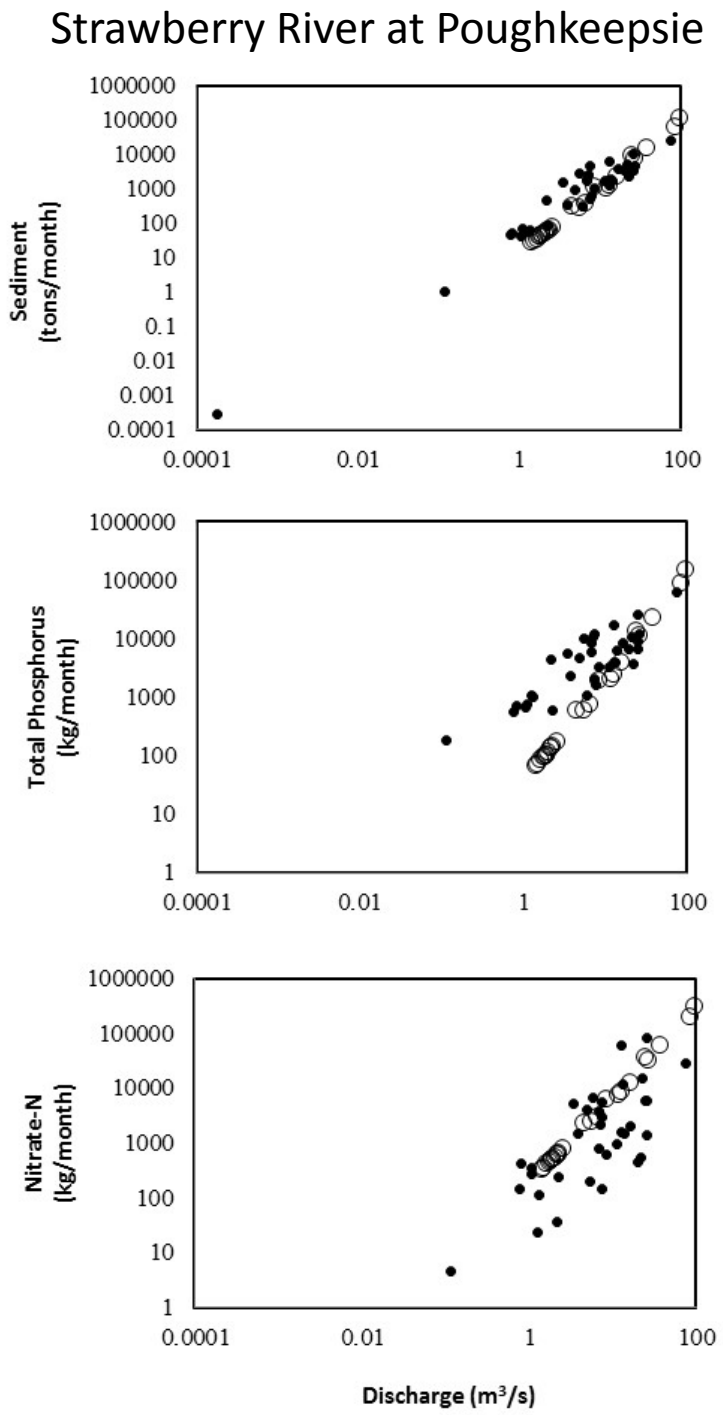
POTEAU RIVER WATERSHED

- We had excellent agreement between load estimates from collected data and model
  - following discharge relationship
- But, we kind of cheated in this watershed because there was a little USGS data available
  - for model calibration
- However, that gave us confidence that this approach might work



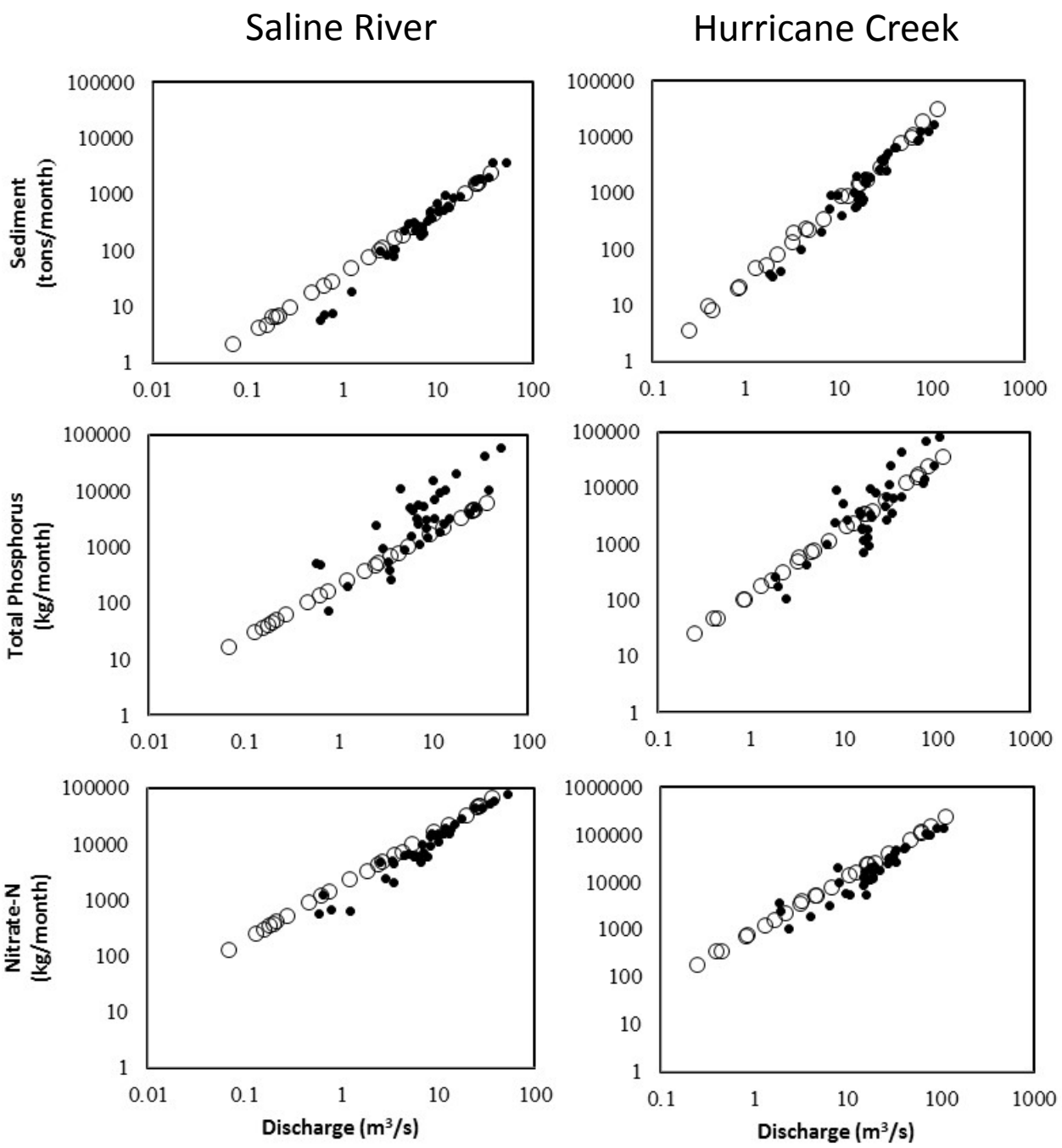
STRAWBERRY RIVER WATERSHED

- The agreement here was not as strong following the discharge relationship
  - sediment was good
  - P not so much
  - N was alright
- So, our confidence in this watershed was limited
  - remember only one USGS stage recording site, no discharge
  - especially for P

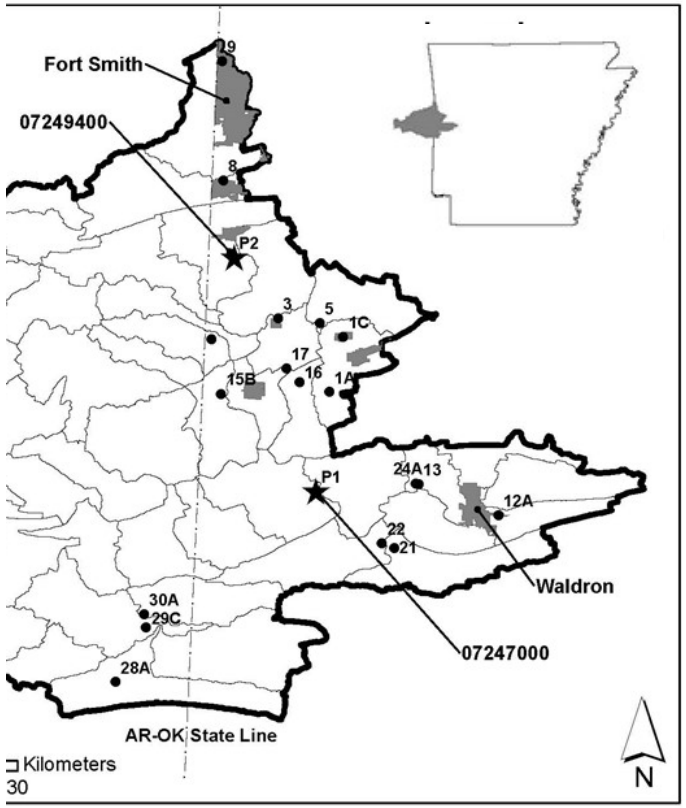


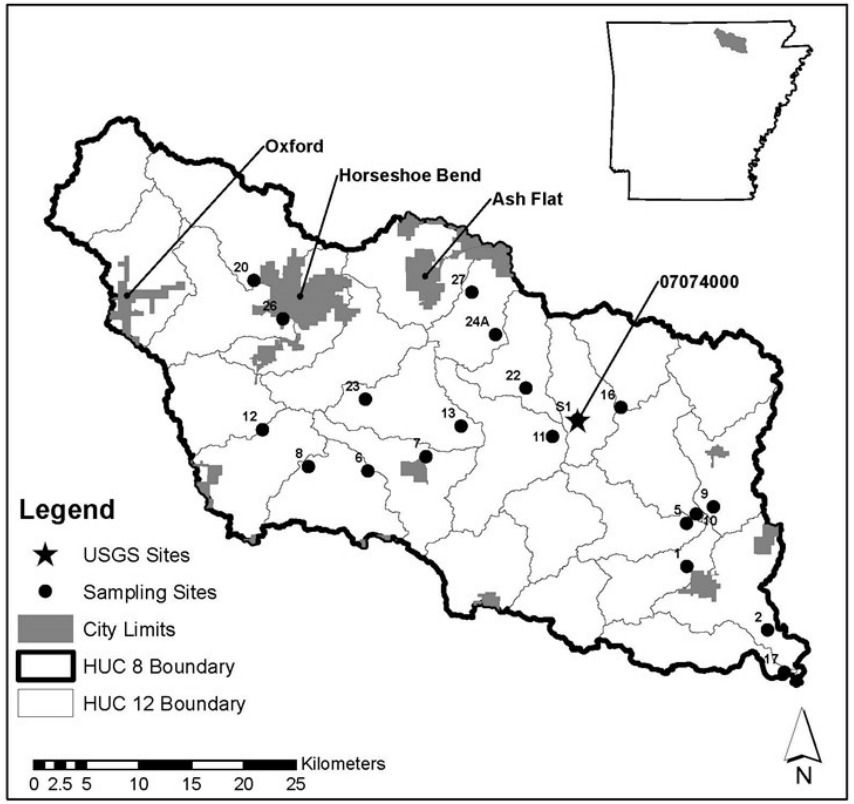
UPPER SALINE RIVER WATERSHED

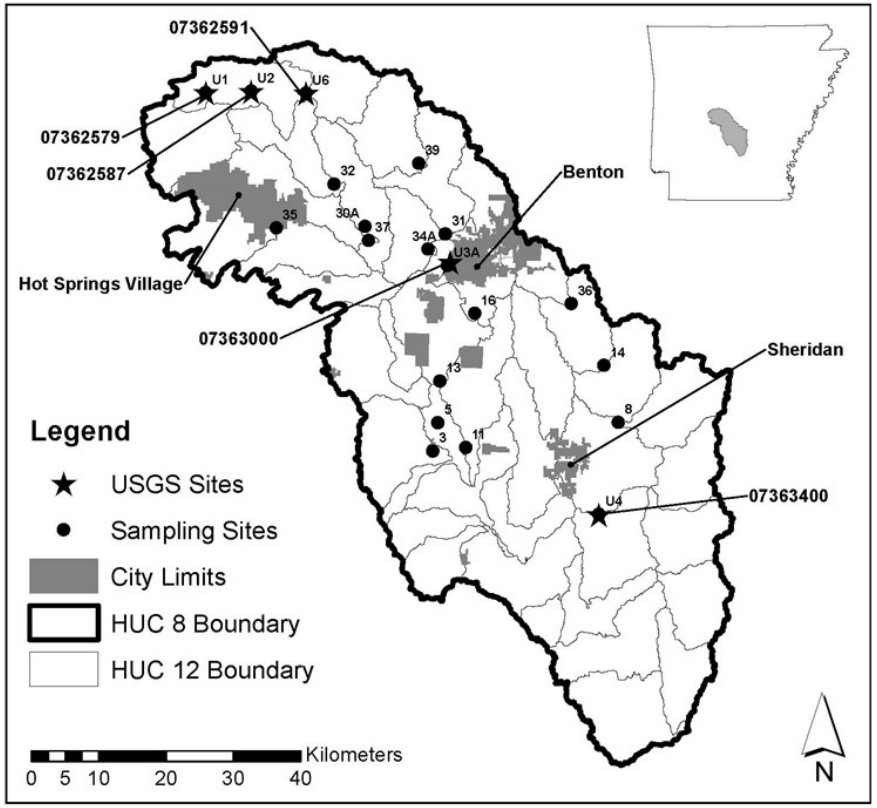
- We had excellent agreement between load estimates from collected data and model
  - following discharge relationship
  - maybe some divergence with sediment loads at low flow
- However, this comparison gave us confidence that in the load estimates from SWAT
  - at these sites…



We also had the HUC 12 level monitoring… The intent with this data was to evaluate how the model predicted concentrations during base flow spatially across the watersheds.





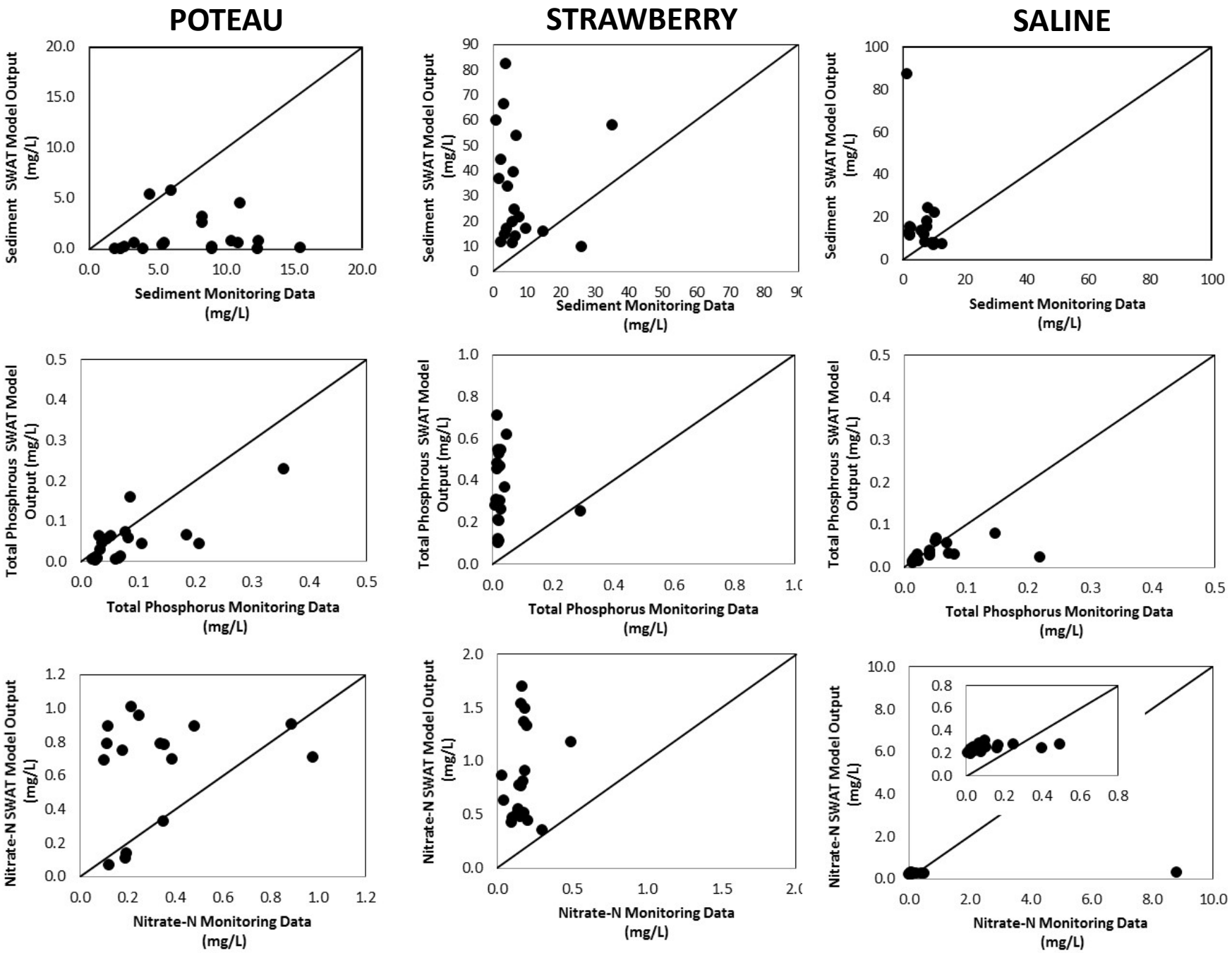


Since our intent was to prioritize (low to high), we compared the site ranks in a watershed: (Spearman Ranks)

- **POTEAU:** sediment not significant, but N and P ranks were correlated
- **STRAWBERRY:** not good across all three
- **SALINE:** sediment not significant, but N and P ranks were correlated

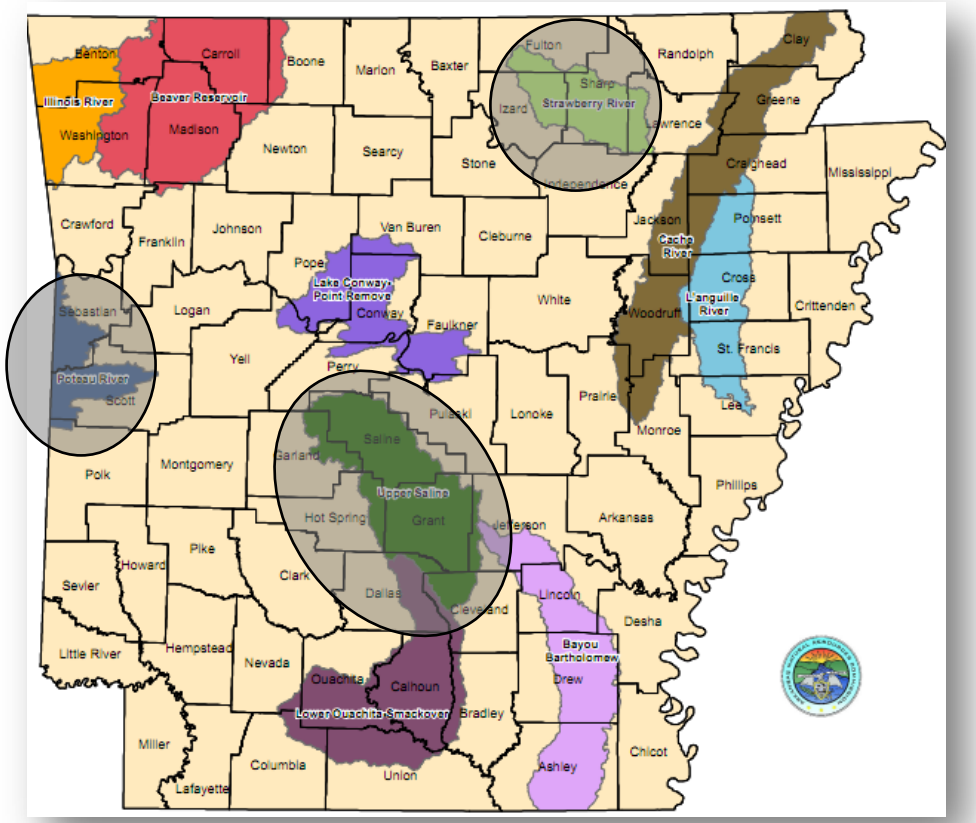
We need to remember:

- time frames not same
- model average verse monthly sample average



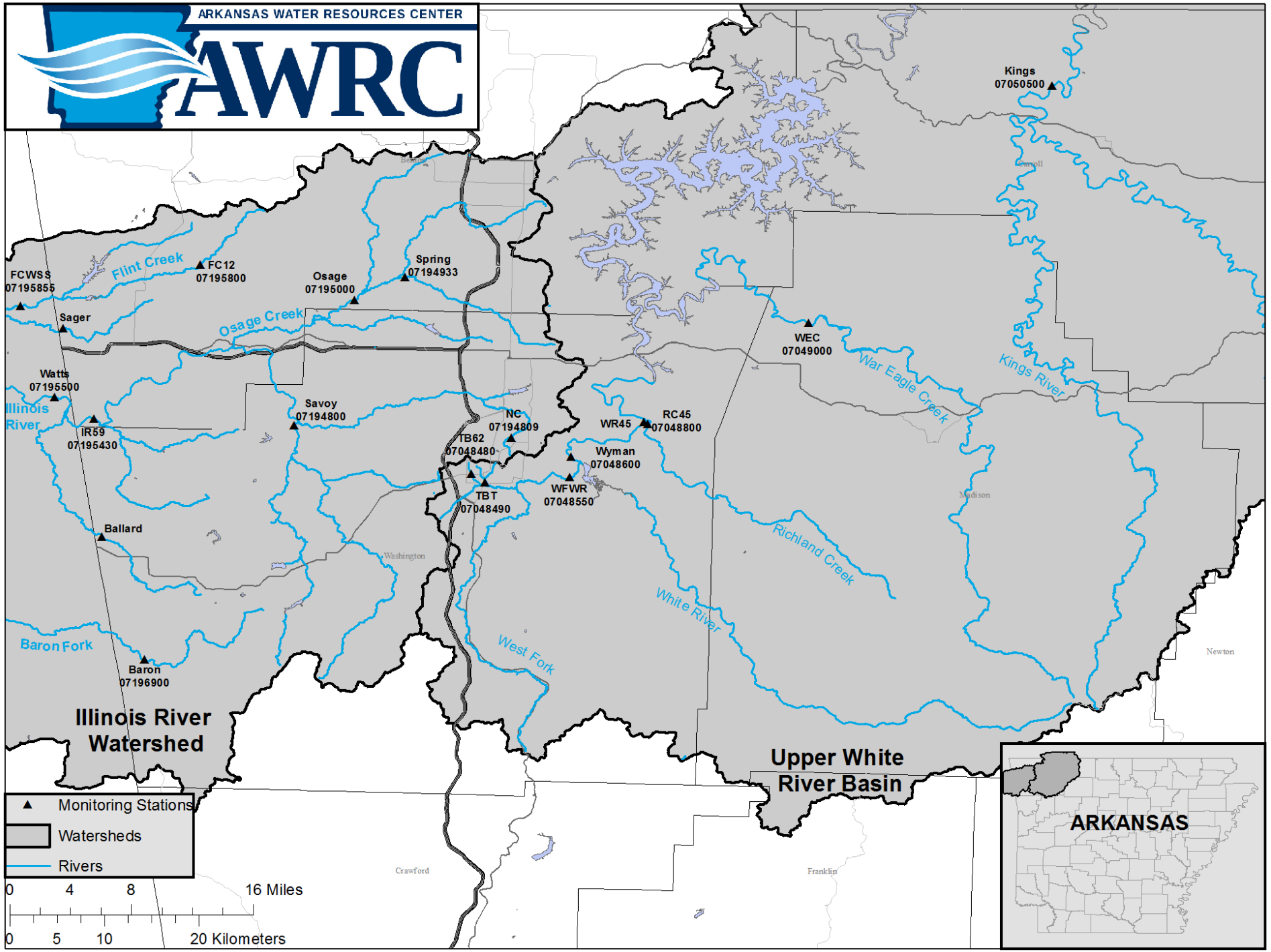
What did having the data outside of the modeling period allow us to do?

- We were able to increase our confidence in the load predictions at sites where USGS discharge or stage was available
  - across all watersheds
- We were also able to increase our confidence in the subwatershed prioritization based on SWAT
  - except at the one location, Strawberry River Watershed

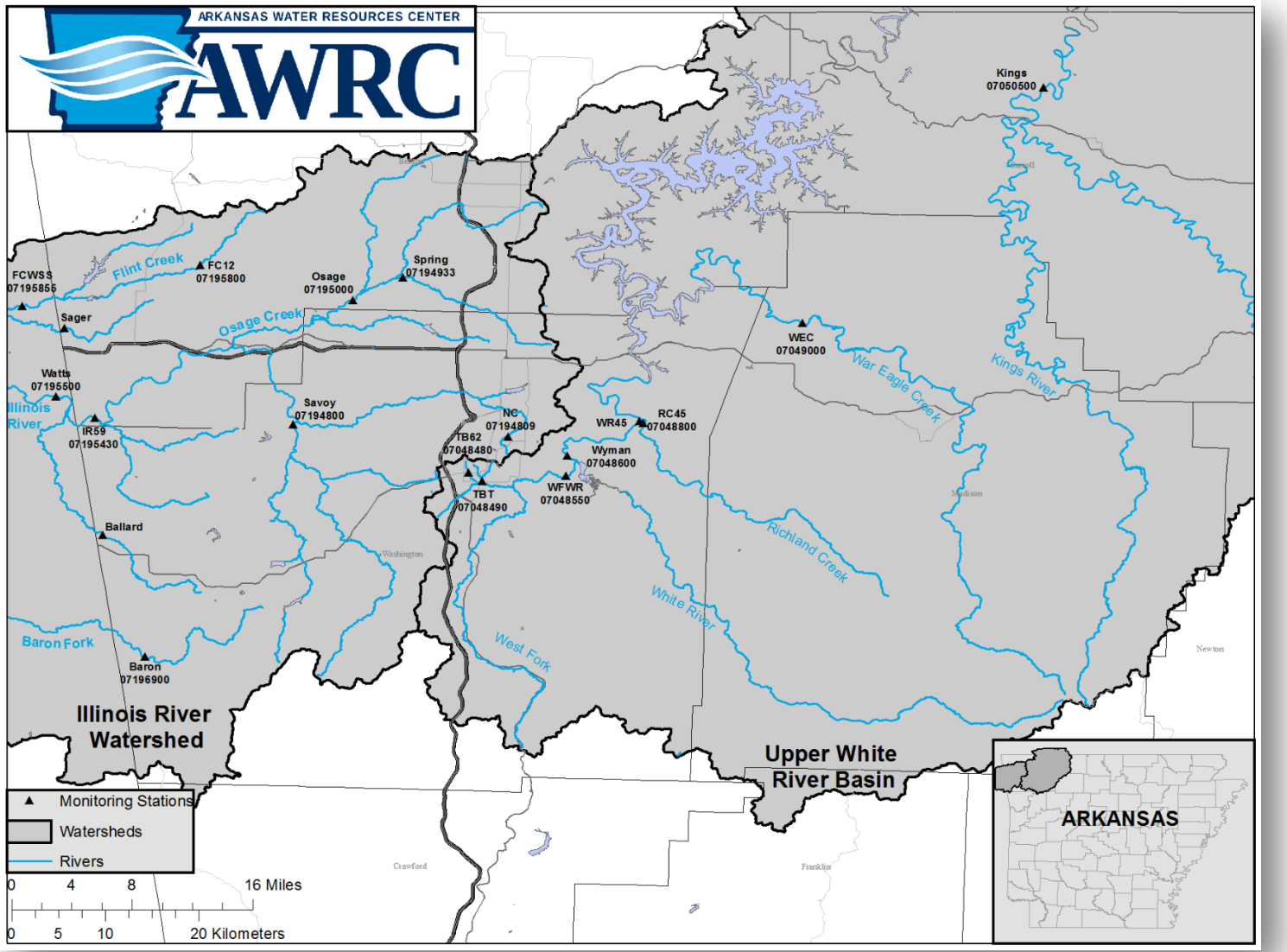


4. How have we expanded monitoring efforts in northwest Arkansas?

We are collecting approximately 46 water samples per site.



5. The use of models… defines data needs



- HUC 8 LEVEL

We can get away with typical available data

- HUC 12 LEVEL

You need to have data from smaller basins, across land uses, topography, etc.

We also need to consider what processes are important

- in-stream storage and release
- changes over time

QUESTION
S?