Empirically Modeling Freshwater Ecosystems to Address Management Problems

John Olson

Desert Research Institute

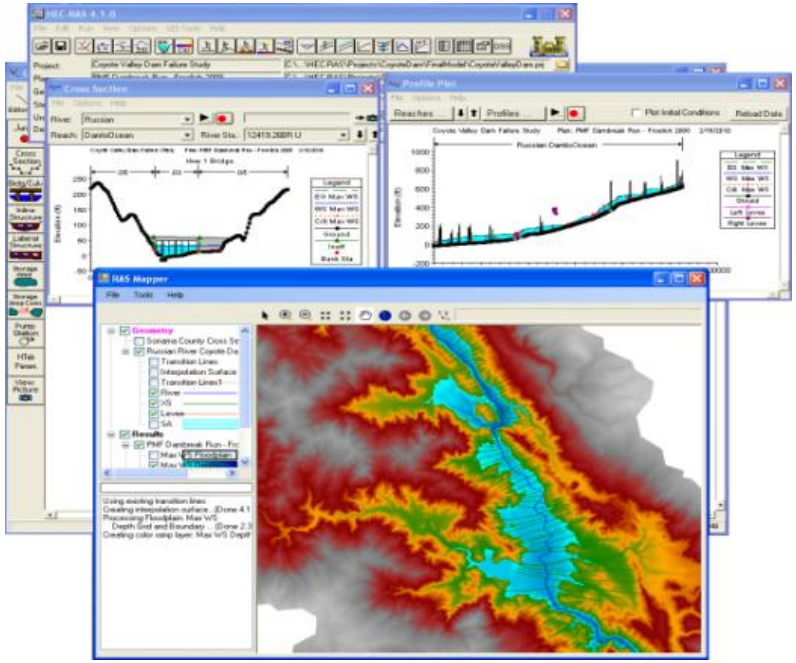
13 November2014

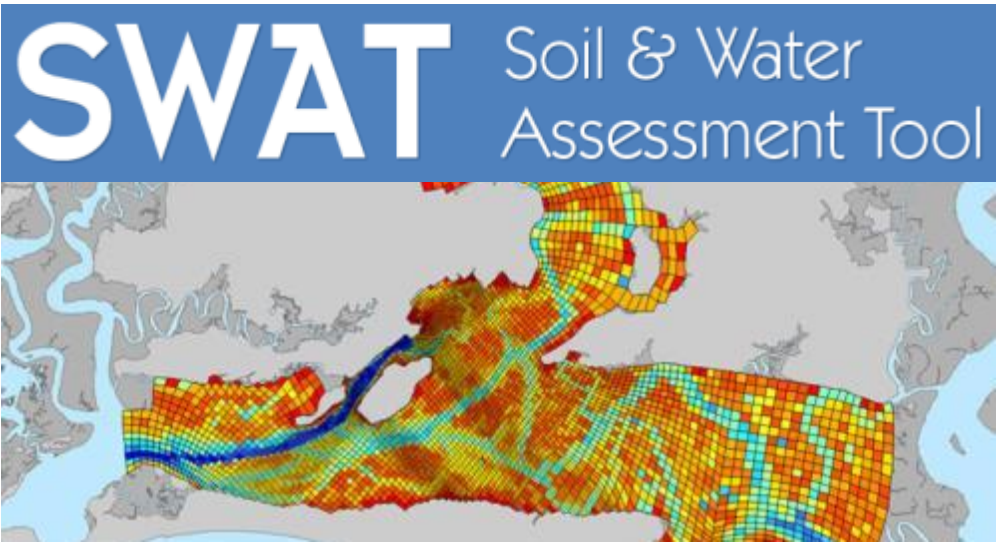
Modeling helps us:

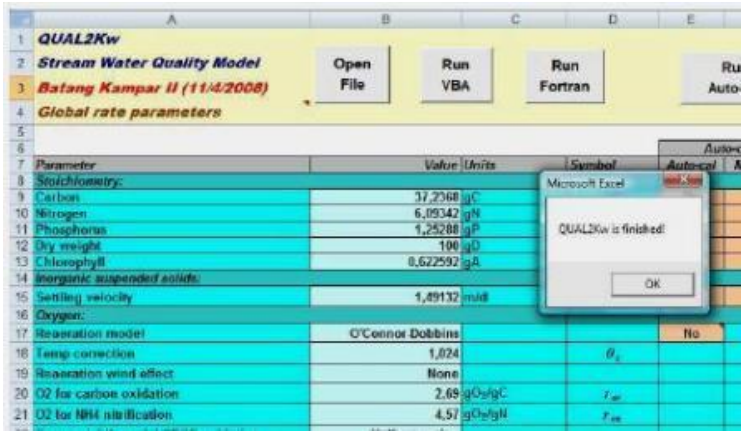
- understand
- predict

Process based:

- HEC-RAS (COE)
- SWAT (USDA)
- QUAL2K (USEPA)





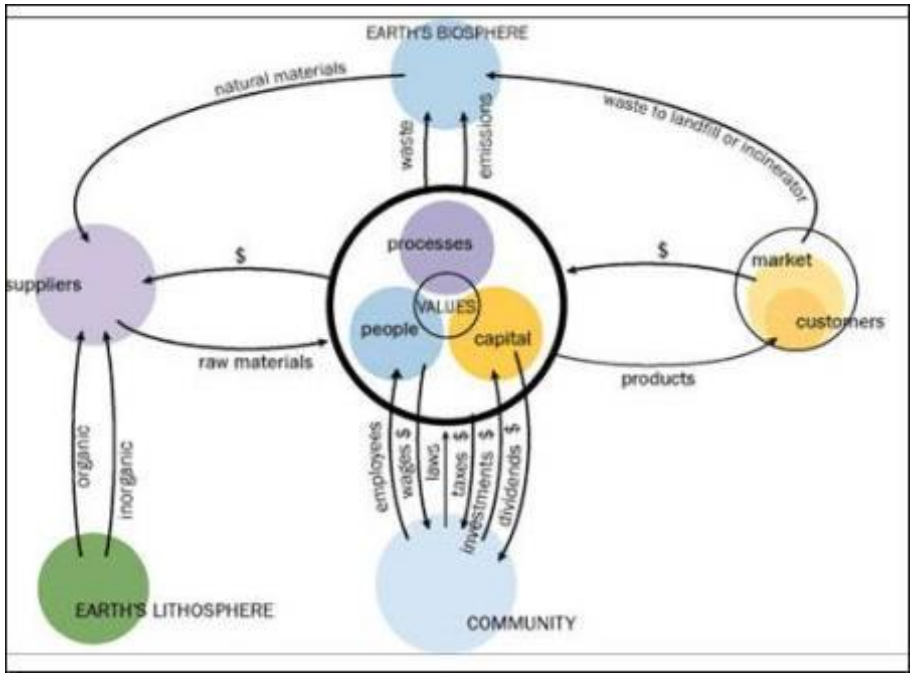


Modeling helps us:

- understand
- predict

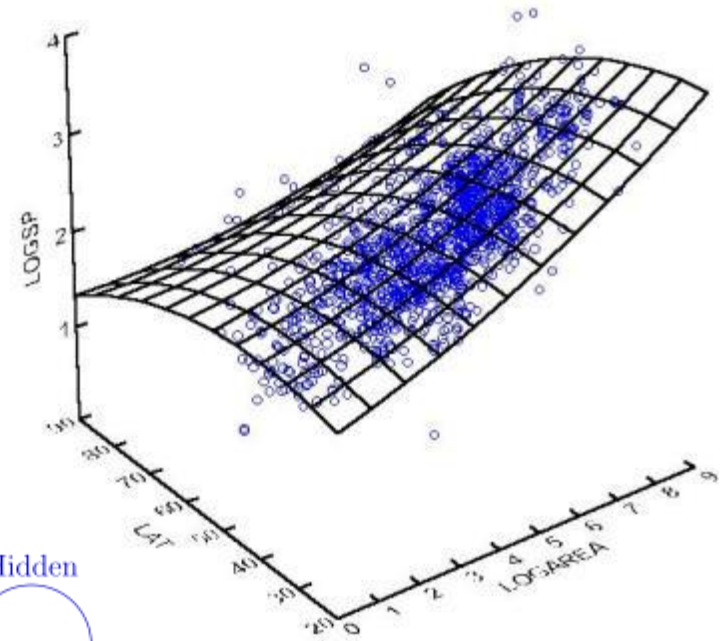
Limitations

- Complexity/open systems
- Large spatial scales
- Calibration/equifinality



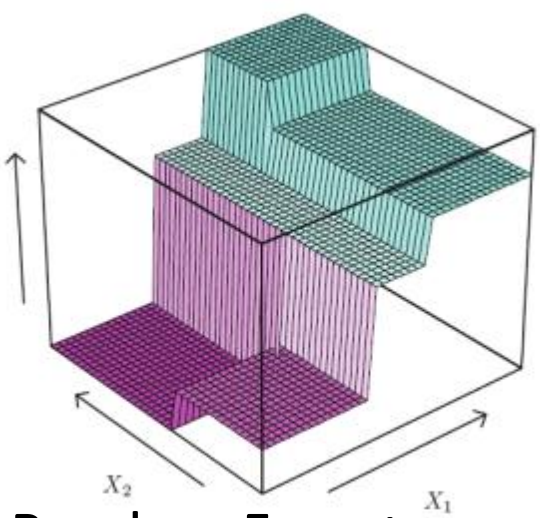
Modeling helps us:

- understand
- predict

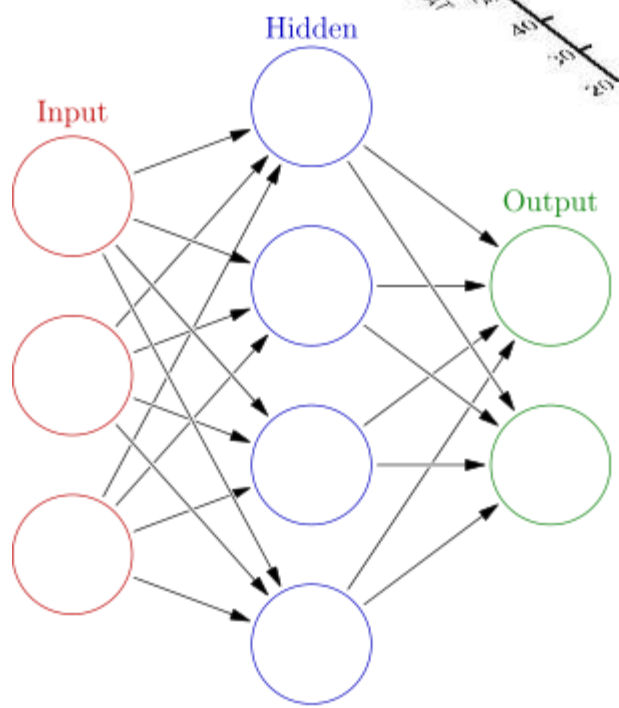


Empirical:

- Regression
- Machine Learning



Random Forests



Artificial Neural Networks

Chemistry

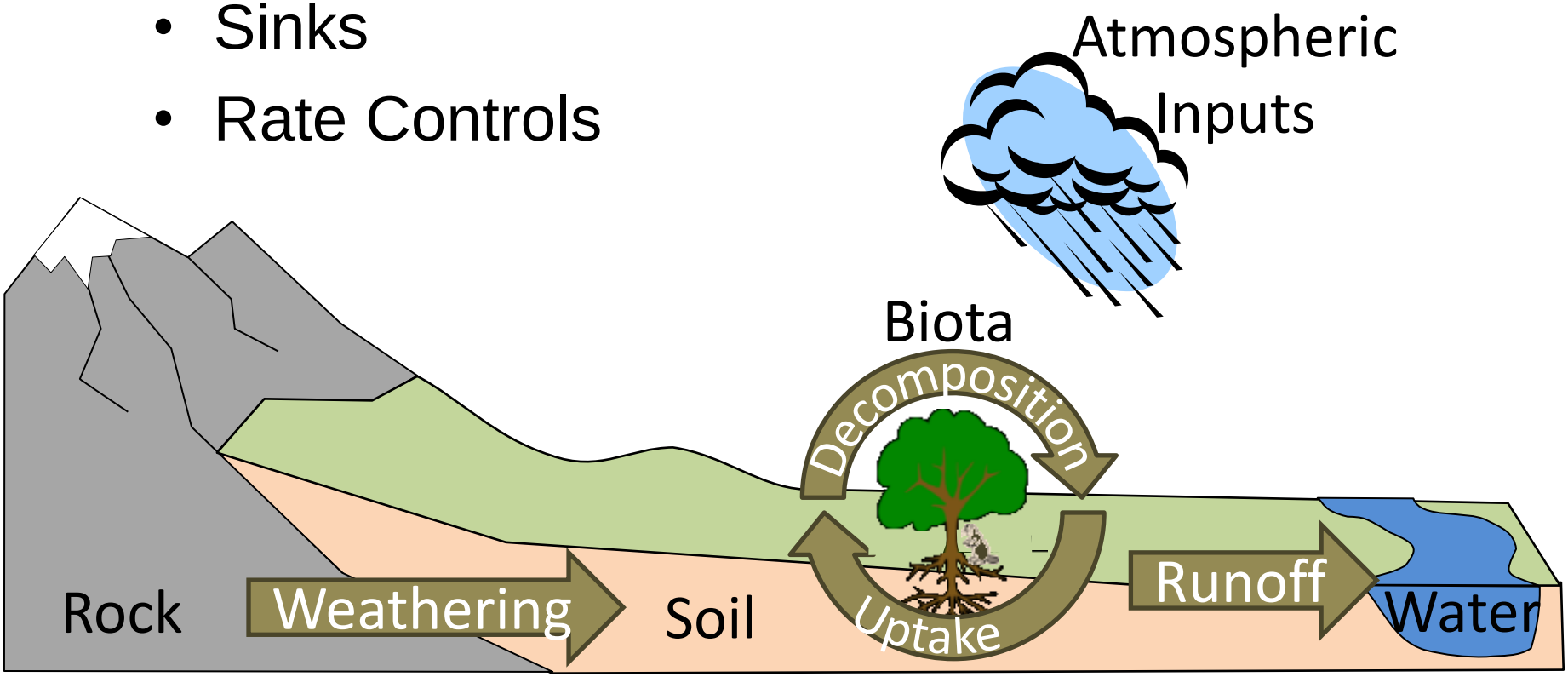
- Total Nitrogen (TN)
- Total Phosphorus (TP)
- Electrical Conductivity (EC)
- Alkalinity (Alk)
- Major Ions (Ca, Mg, SO_4)

- What are reasonable impairment criteria?
- What are realistic restoration goals?

Predicting Natural Stream Chemistry

Should be able to empirically model water chemistry from estimates of:

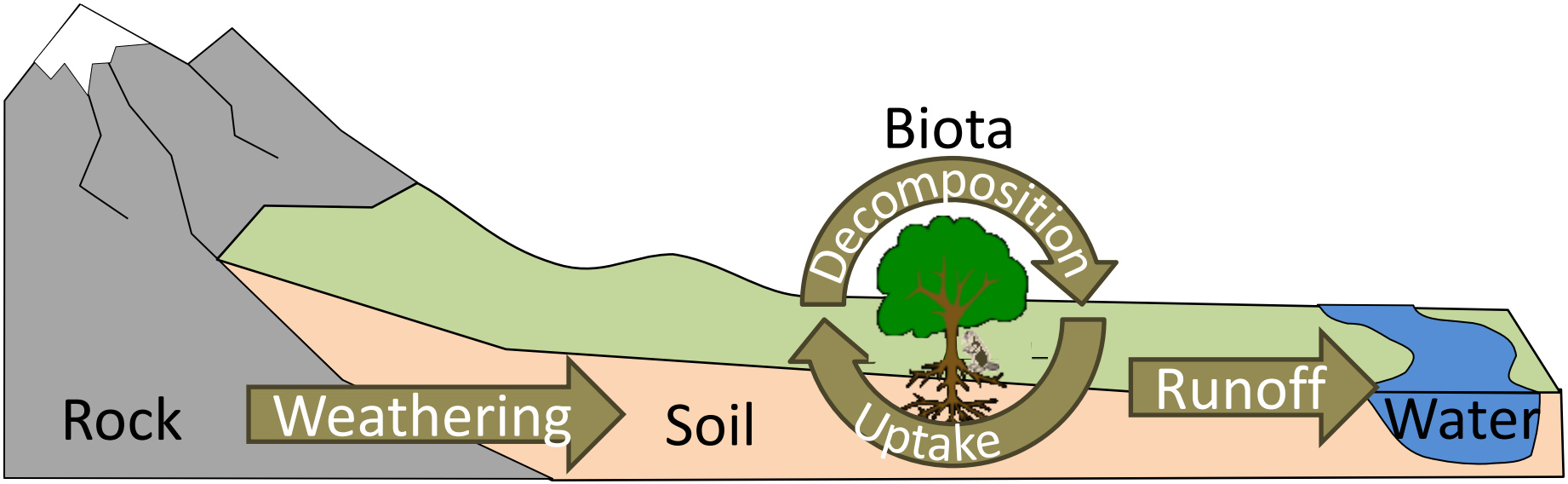
- Sources
- Sinks
- Rate Controls



Predicting Natural Stream Chemistry

Approach:

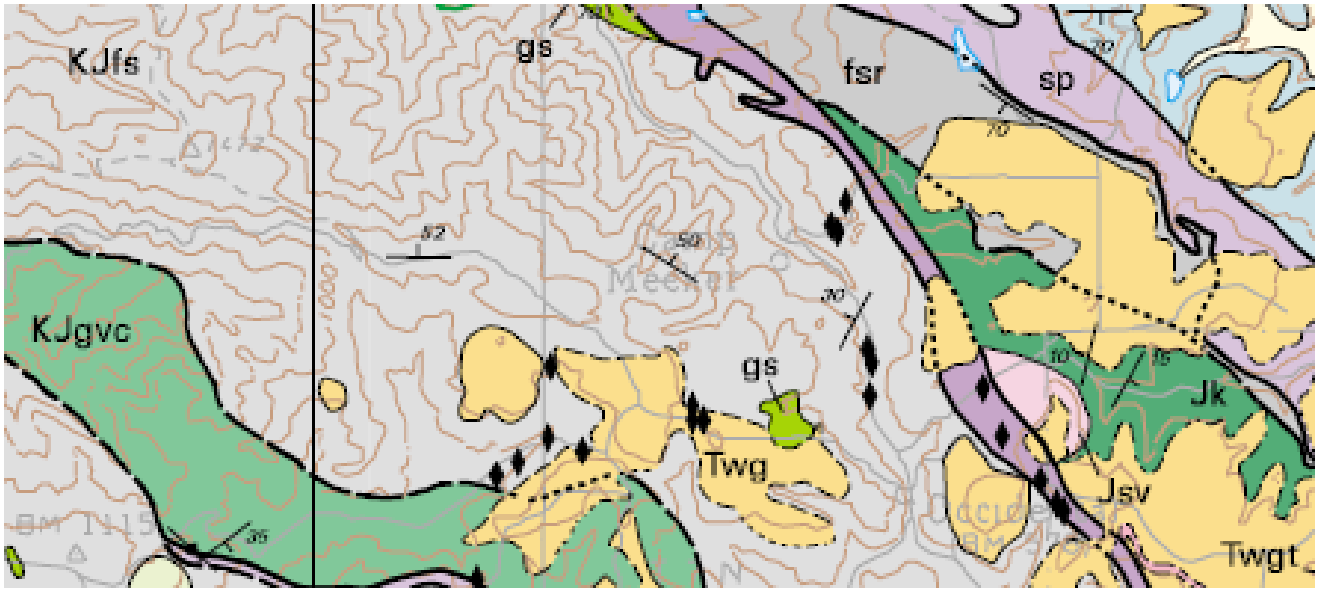
- Measure catchment environment at sites with minimal human impacts
- Develop empirical model relating environment to observed stream chemistry



Predicting Natural Stream Chemistry

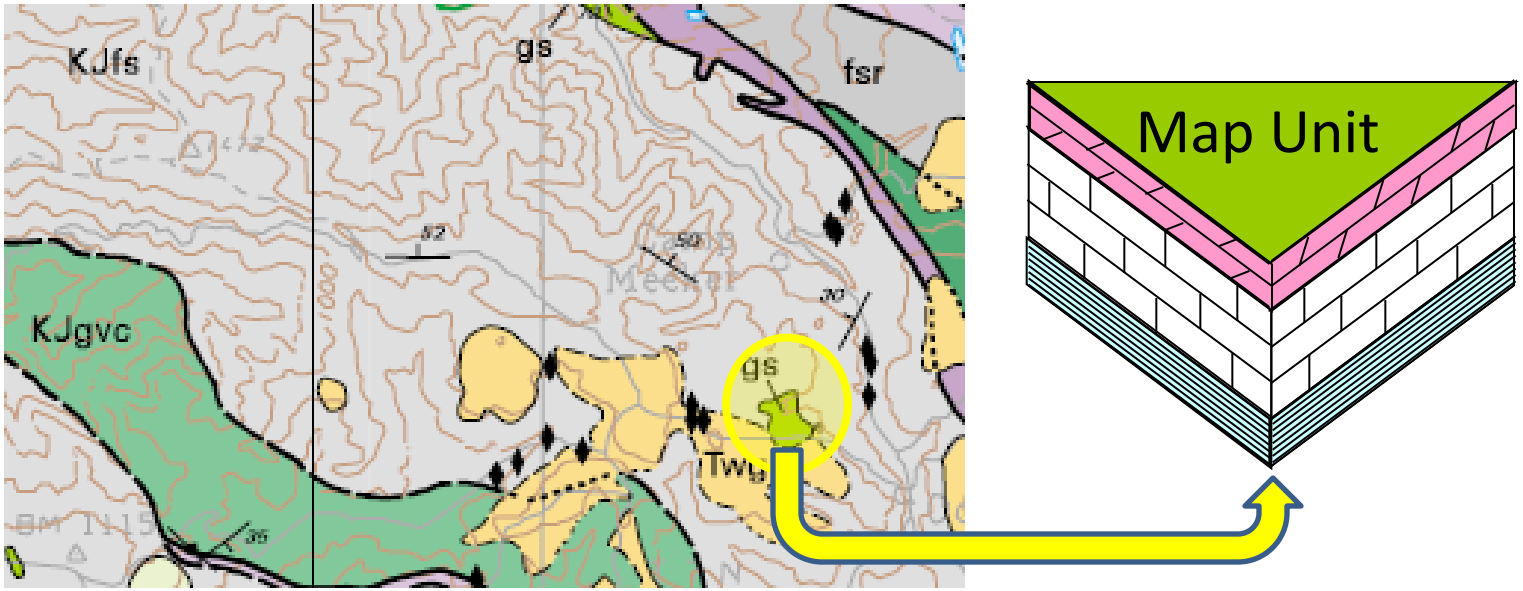
Problem: Geologic data (maps) are categorical

- Many categories
- Categories based on time, structure, formative process, not chemical or physical attributes
- Categories often mixtures of many rock types

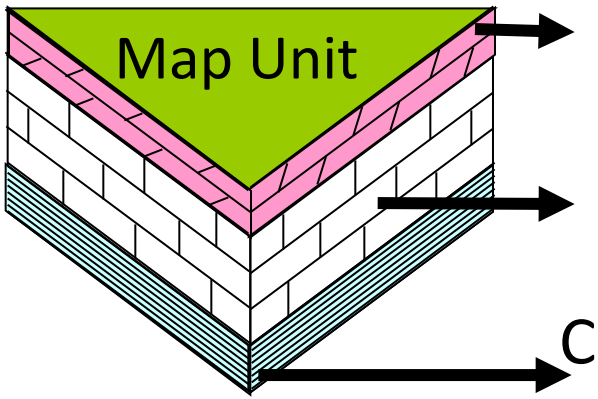


Characterizing Geology

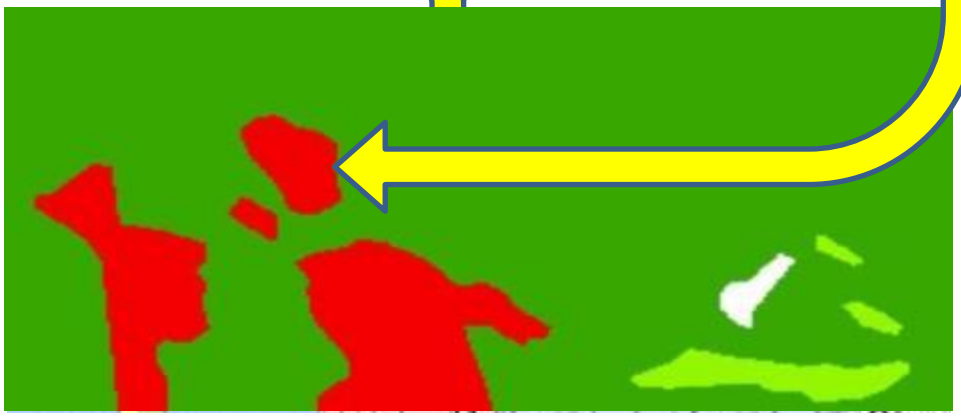
Based on most complete coverage with finest resolution lithology: Preliminary Integrated Geologic Map Databases for the United States, (1:500,000 scale)

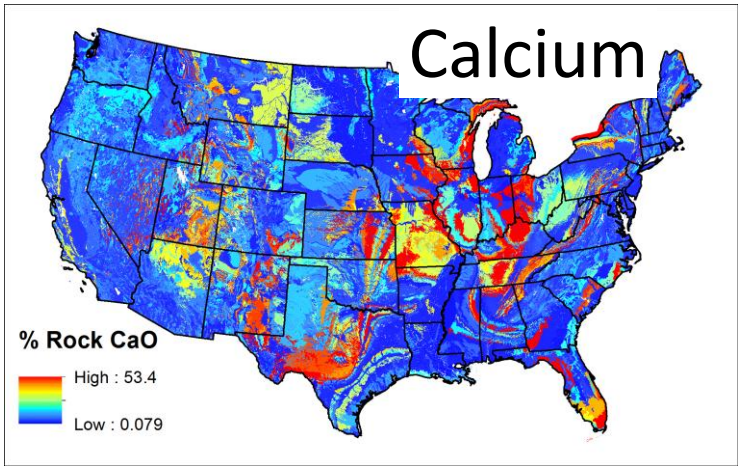


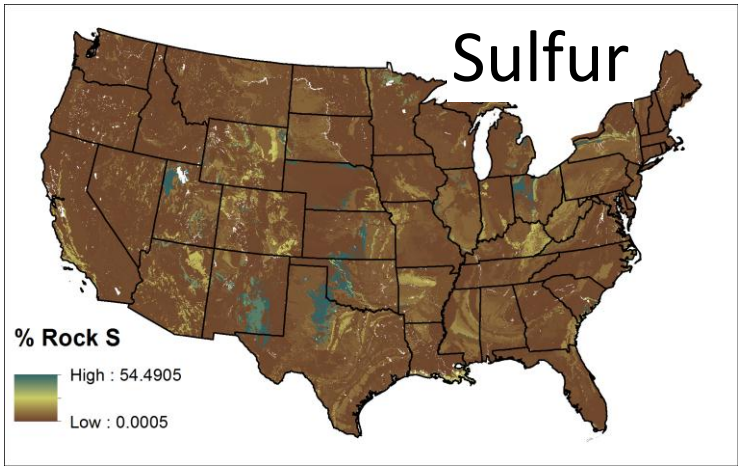
Characterizing Geology

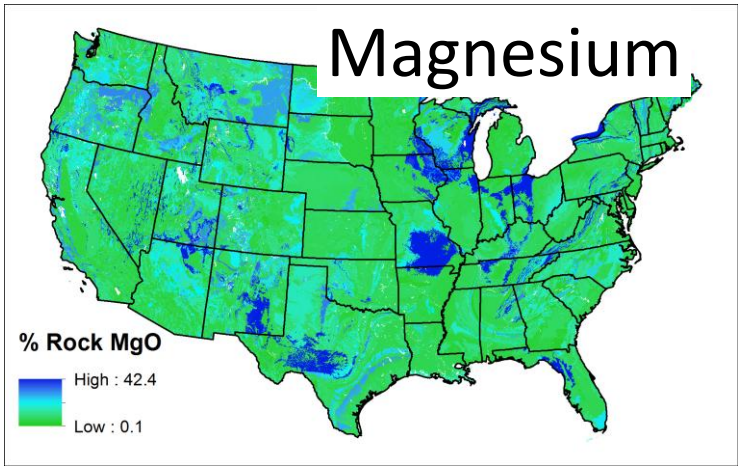


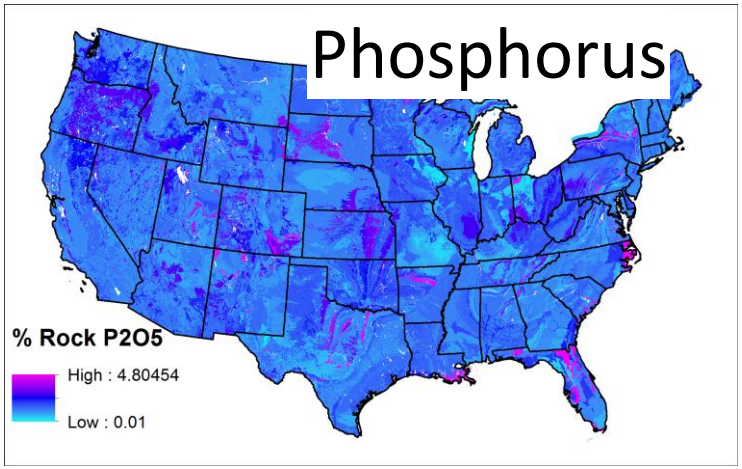
Lithology % CaO	Prevalence (Weight)	Weighted Average
Dolostone: 29.0	Minor (0.2311)	38.1 % CaO
Limestone: 46.0	Major (0.7119)	
Calcareous Shale 23.6	Minor (0.2311)	

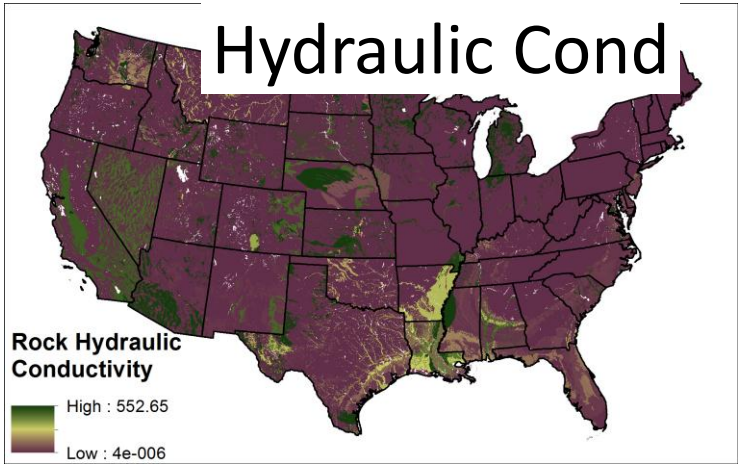


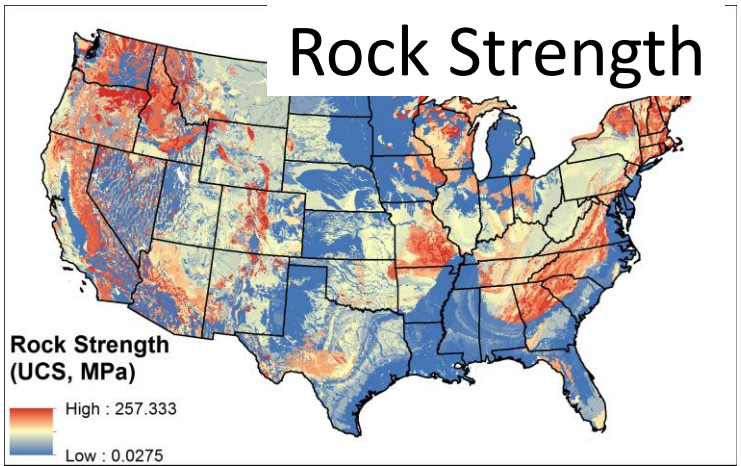

Calcium
% Rock CaO
High : 53.4
Low : 0.079

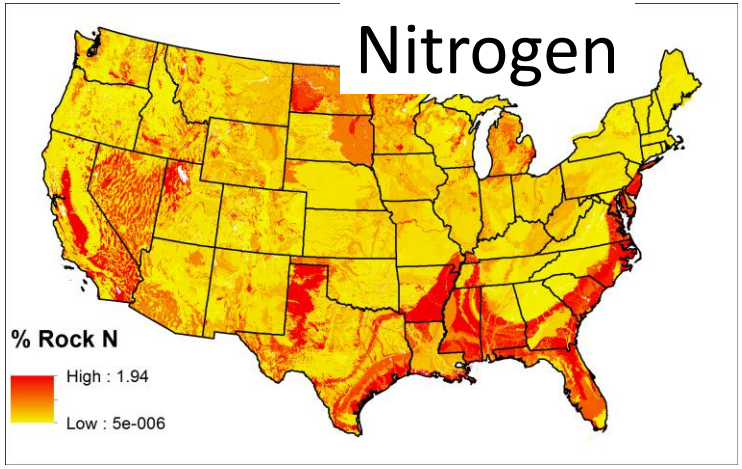

Sulfur
% Rock S
High : 54.4905
Low : 0.0005

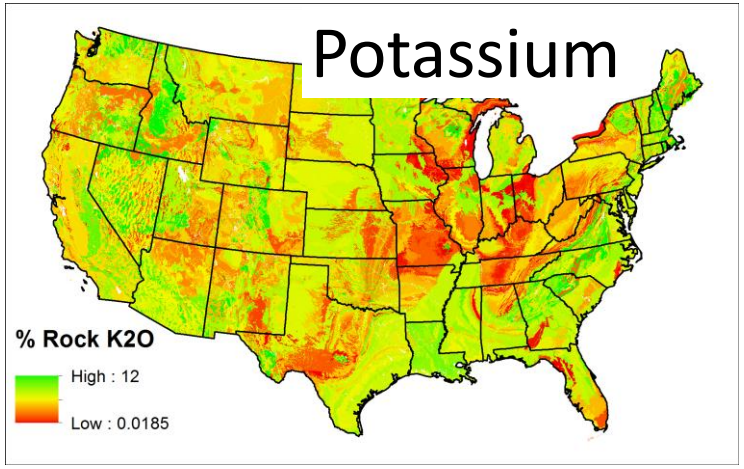

Magnesium
% Rock MgO
High : 42.4
Low : 0.1

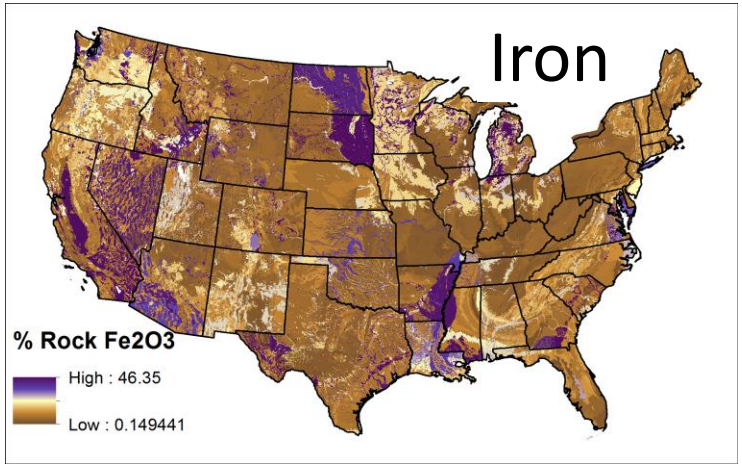

Phosphorus
% Rock P2O5
High : 4.80454
Low : 0.01

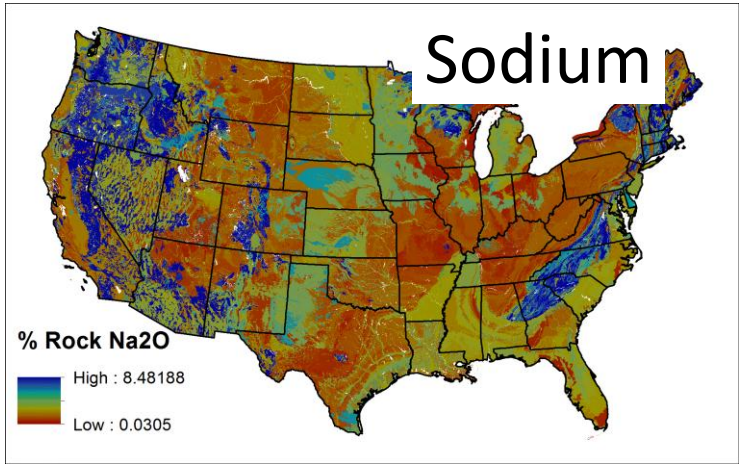

Hydraulic Cond
Rock Hydraulic Conductivity
High : 552.65
Low : 4e-006

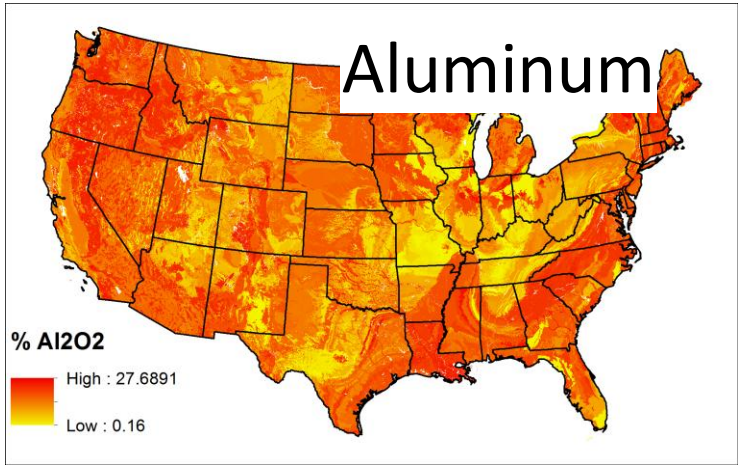

Rock Strength
Rock Strength (UCS, MPa)
High : 257.333
Low : 0.0275

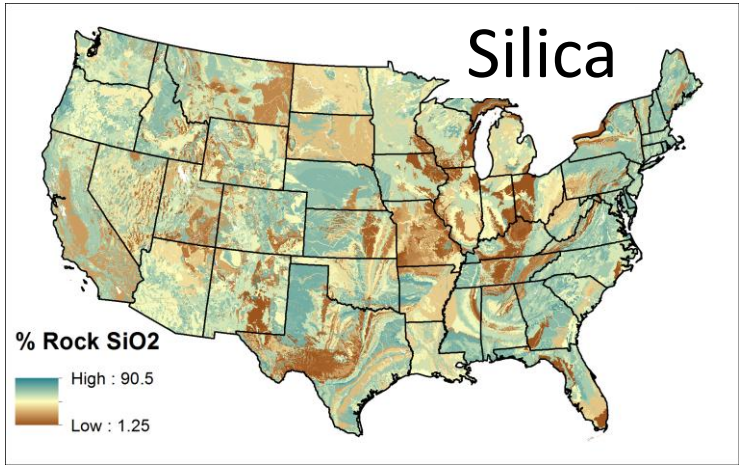

Nitrogen
% Rock N
High : 1.94
Low : 5e-006


Potassium
% Rock K2O
High : 12
Low : 0.0185

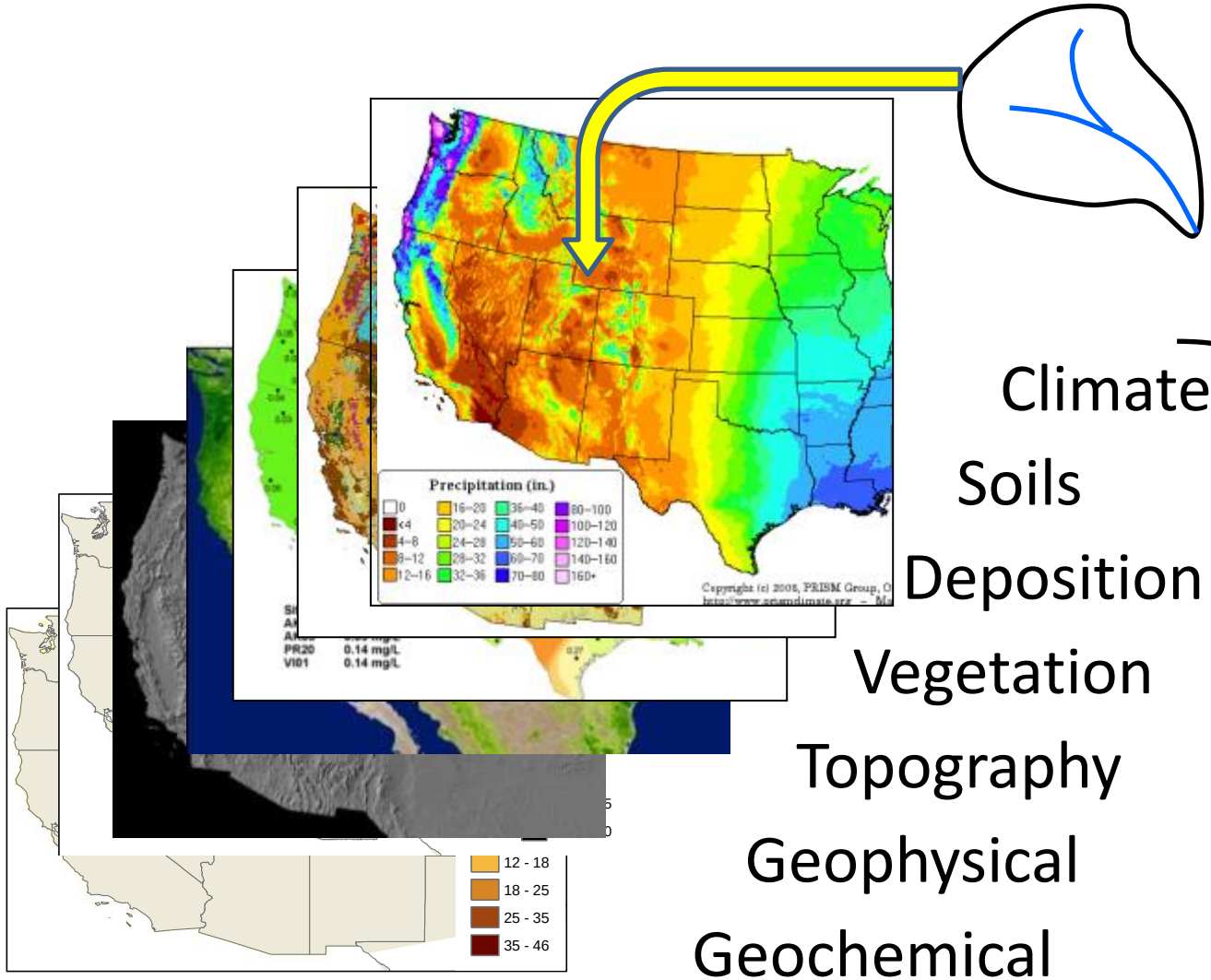

Iron
% Rock Fe2O3
High : 46.35
Low : 0.149441


Sodium
% Rock Na2O
High : 8.48188
Low : 0.0305


Aluminum
% Al2O2
High : 27.6891
Low : 0.16


Silica
% Rock SiO2
High : 90.5
Low : 1.25

Measure Catchment Attributes



Climate

Soils

Deposition

Vegetation

Topography

Geophysical

Geochemical

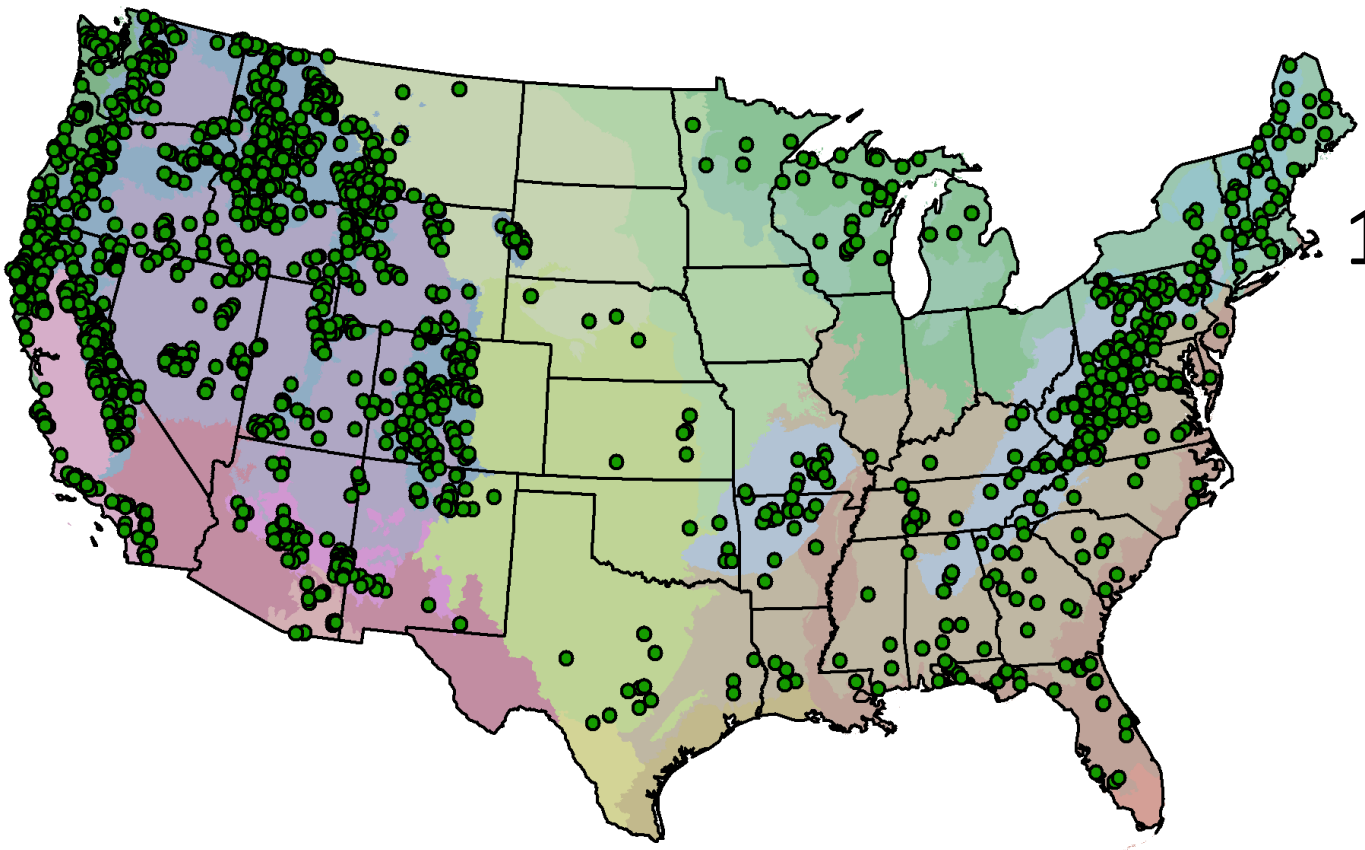
Sites

Attributes

Data

Developing Water Chemistry Models

Screened 2449 reference sites identified and sampled for water chemistry by multiple agencies



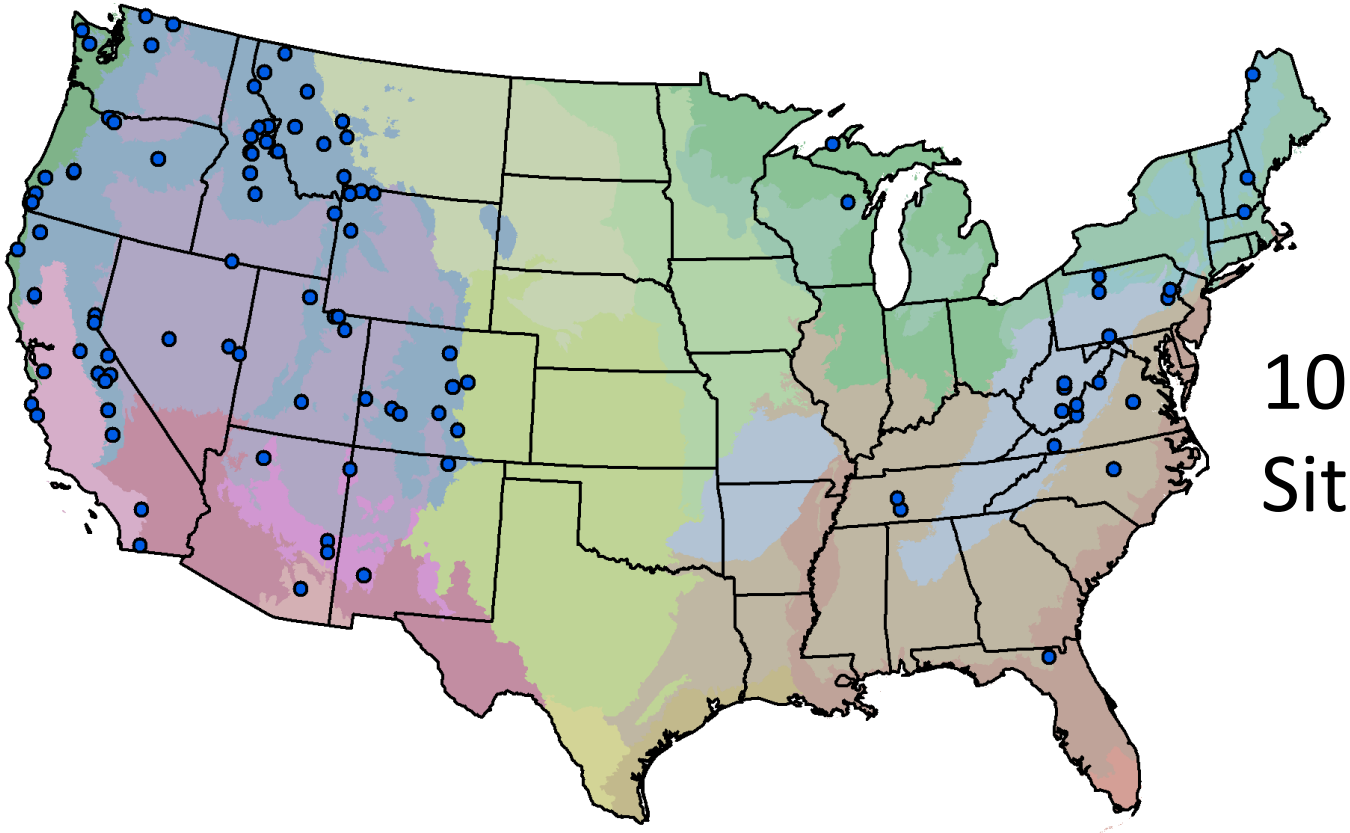
1975 sites

- 1935 EC
- 1768 ANC
- 1225 Ca
- 923 SO_4
- 1274 TP
- 1058 SiO_2

- Samples made at “base-flow”

Developing Water Chemistry Models

Withheld 5% of samples from each level II ecoregion for external validation data

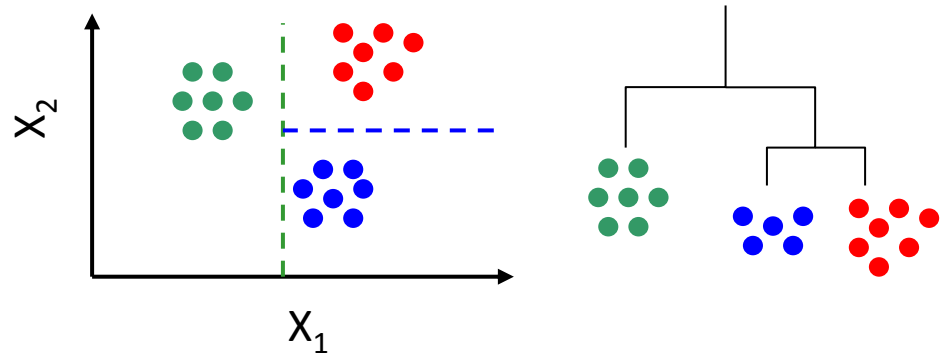


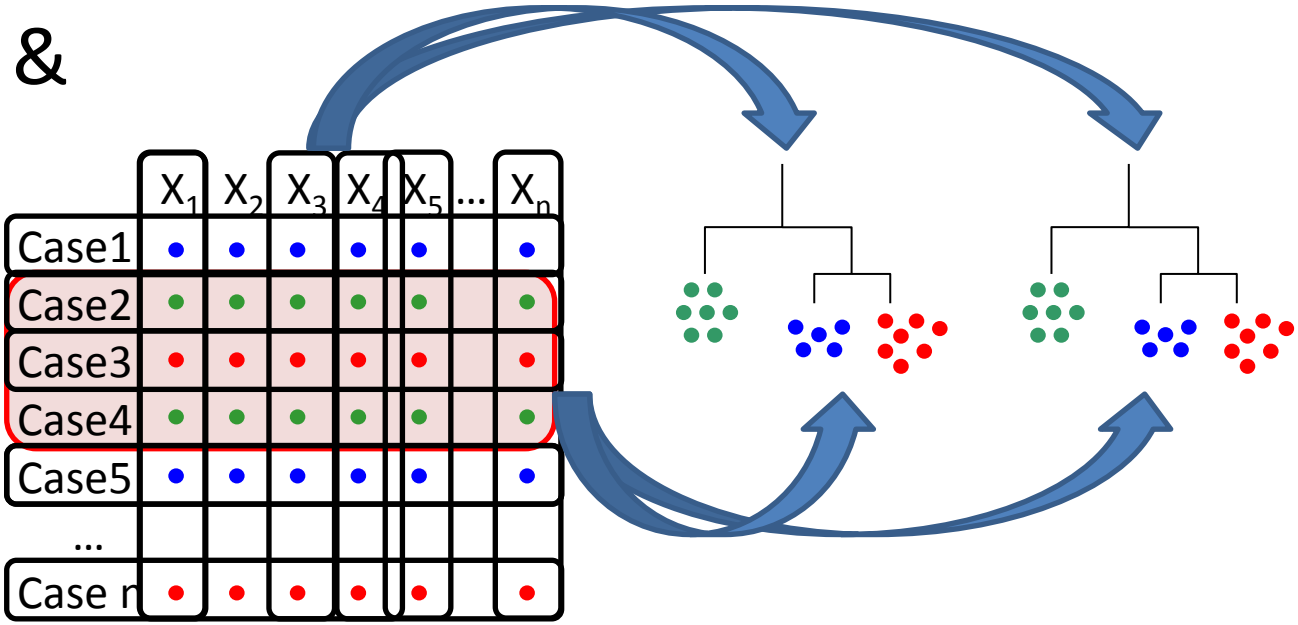
100 Validation Sites

Developing Water Chemistry Models

Random Forest (RF) Modeling

- Recursive partitioning creates regression trees (CART)
- Trees built on random subsets of data & predictors
- Unused data used to evaluate

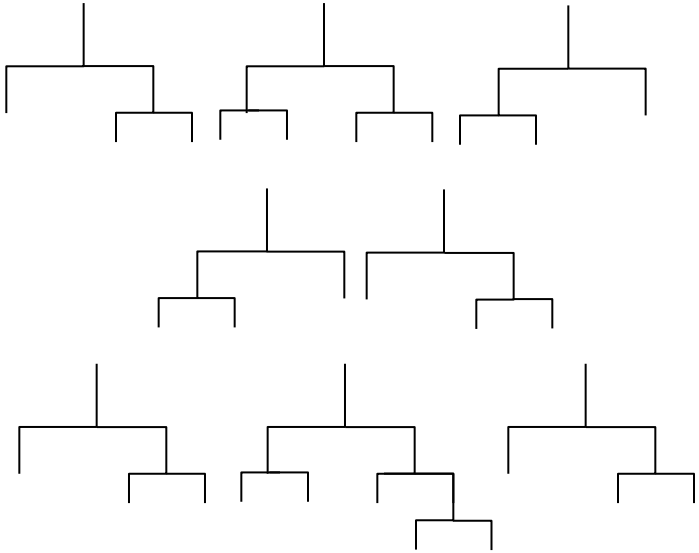




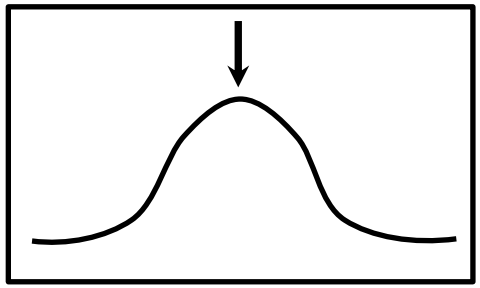
Developing Water Chemistry Models

Random Forest (RF) Modeling

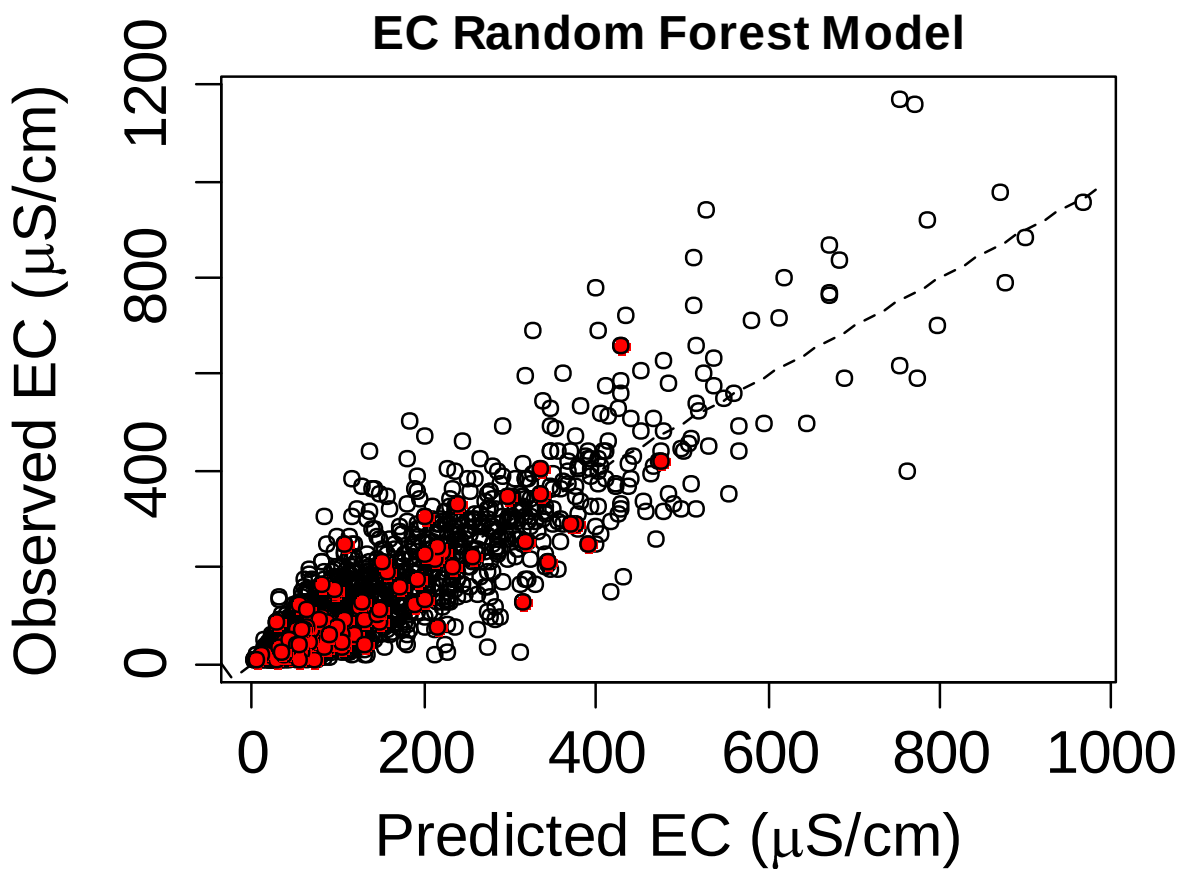
- Build 1000s of trees
- Predicts outcomes by averaging across trees



- Advantages:
  - Does not overfit
  - Effectively models non-linear data
  - Incorporates interactions
  - But cannot extrapolate!



Results: EC Model



Training Data:

n=1835

NSE = 0.80

r^2 = 0.80

RMSE = 65 µS/cm

Validation Data:

n = 100

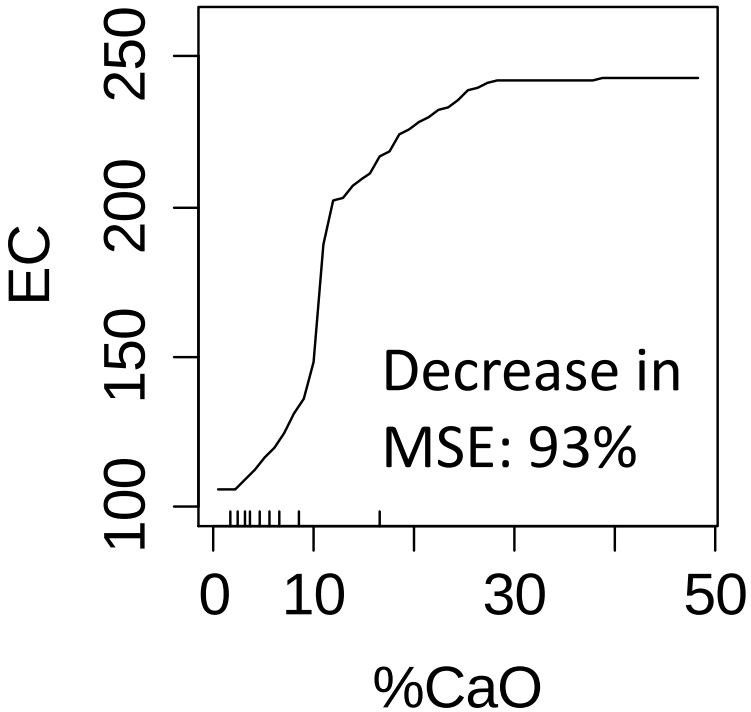
NSE = 0.74

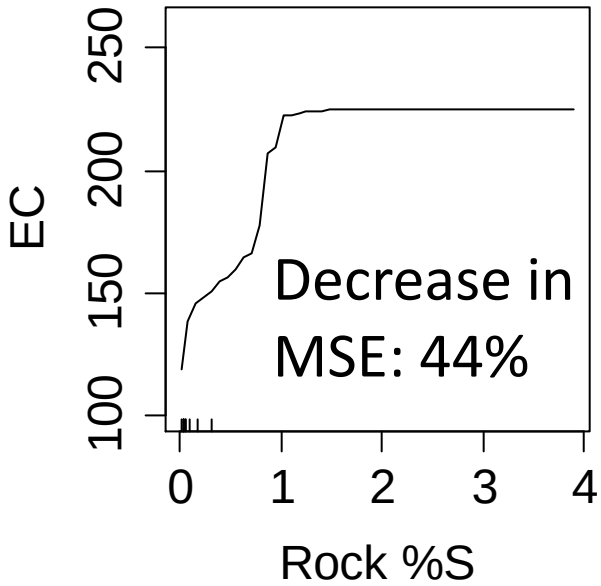
r^2 = 0.76

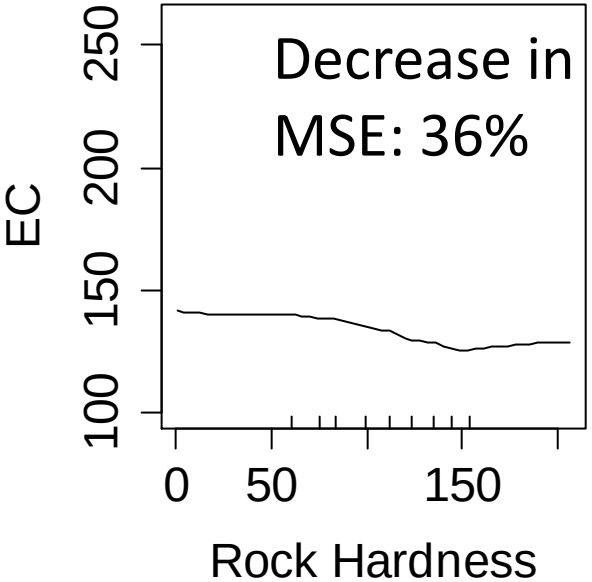
RMSE = 56 µS/cm

Training Data Validation Data 1:1 Line

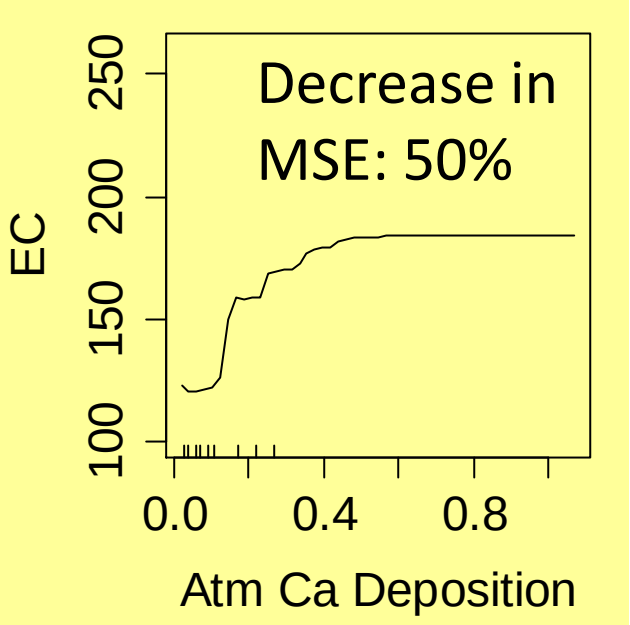
Results: EC Model - Geology Predictors

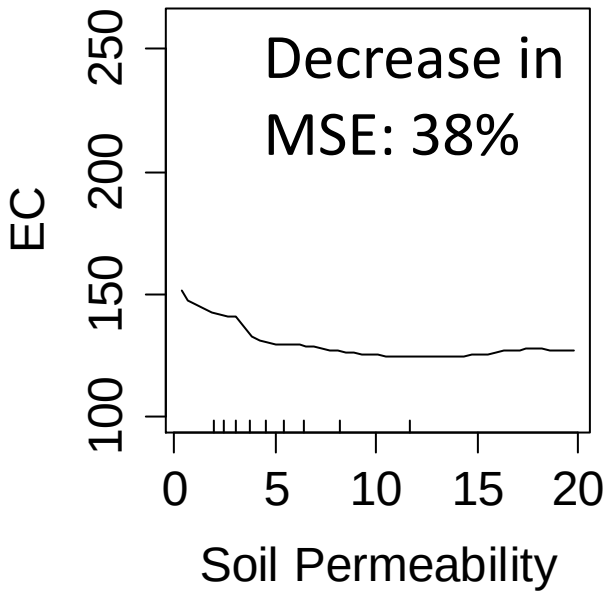


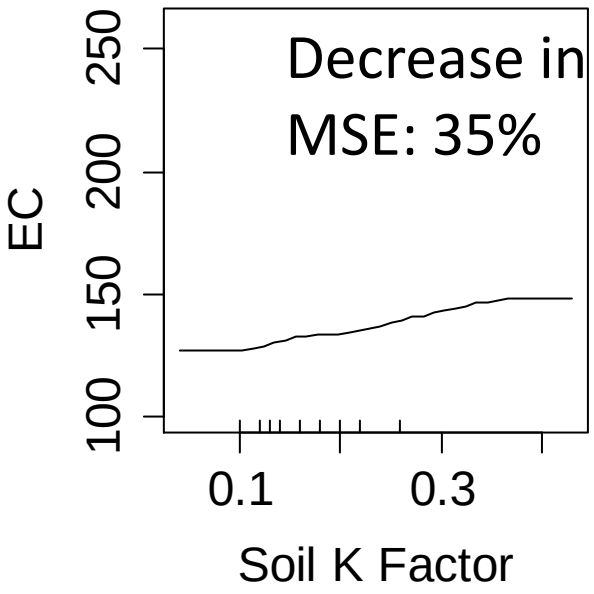


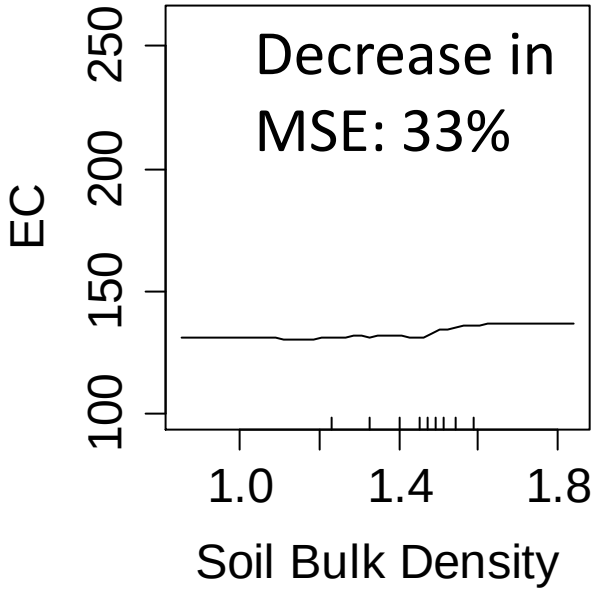


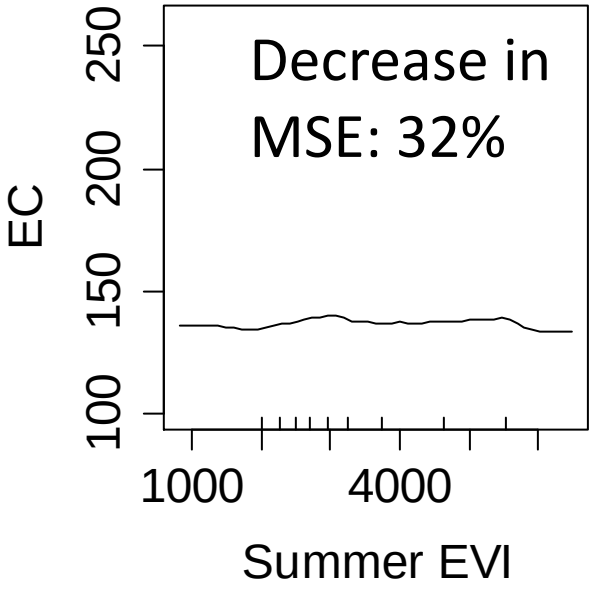
Results: EC Model – Other Predictors

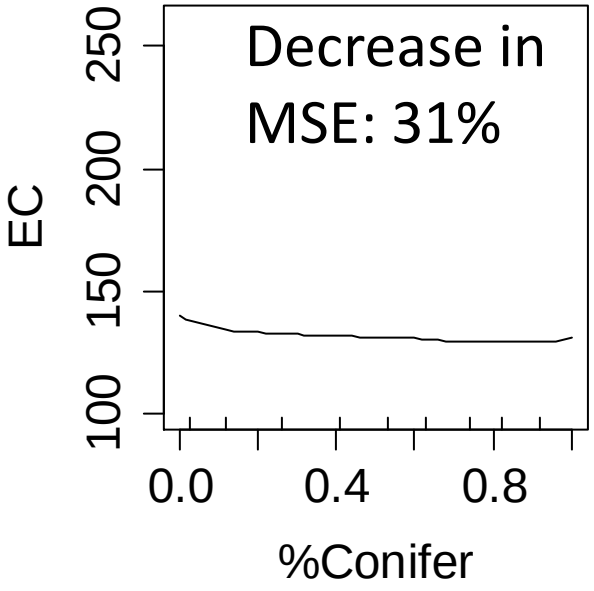




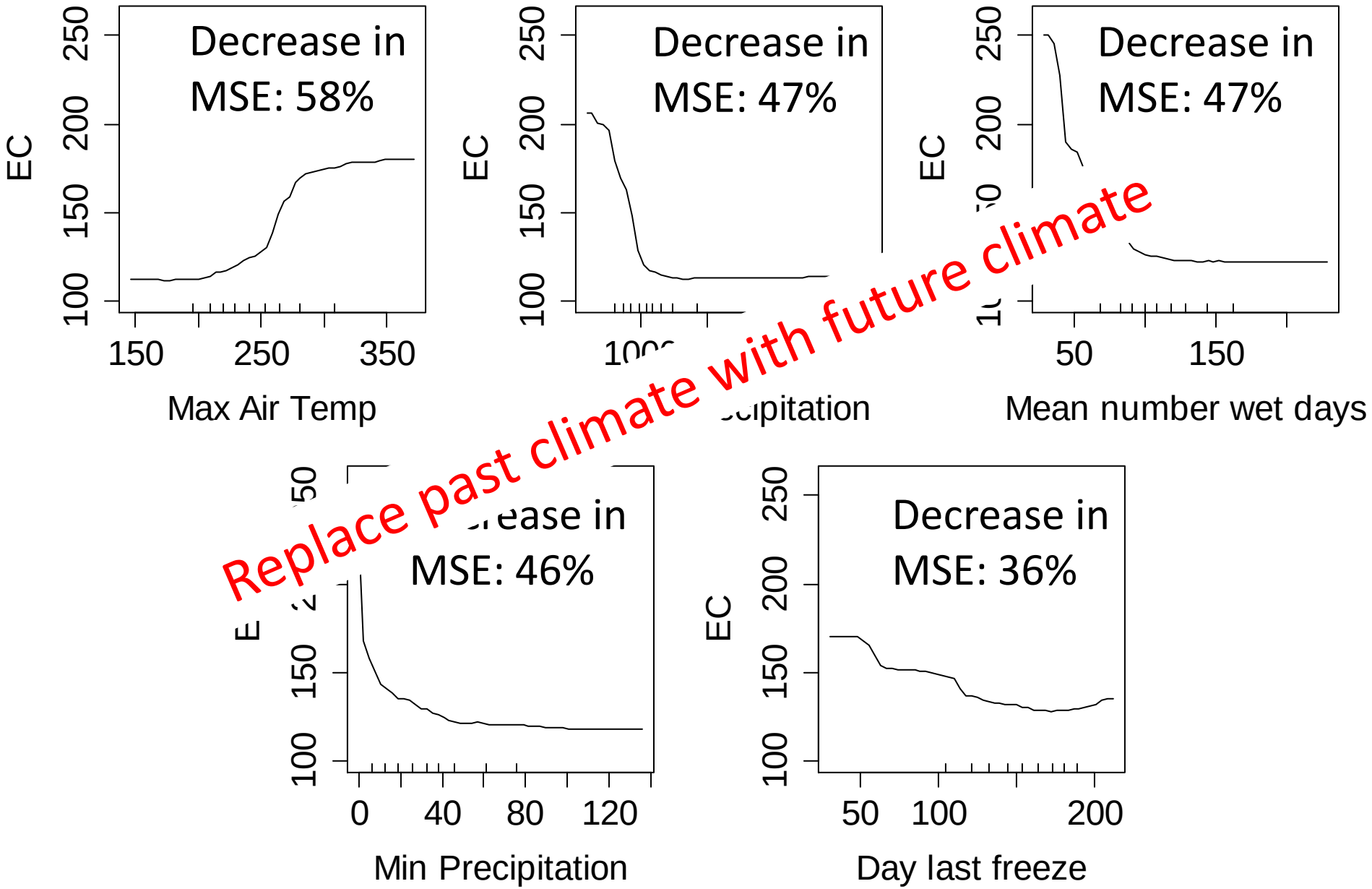




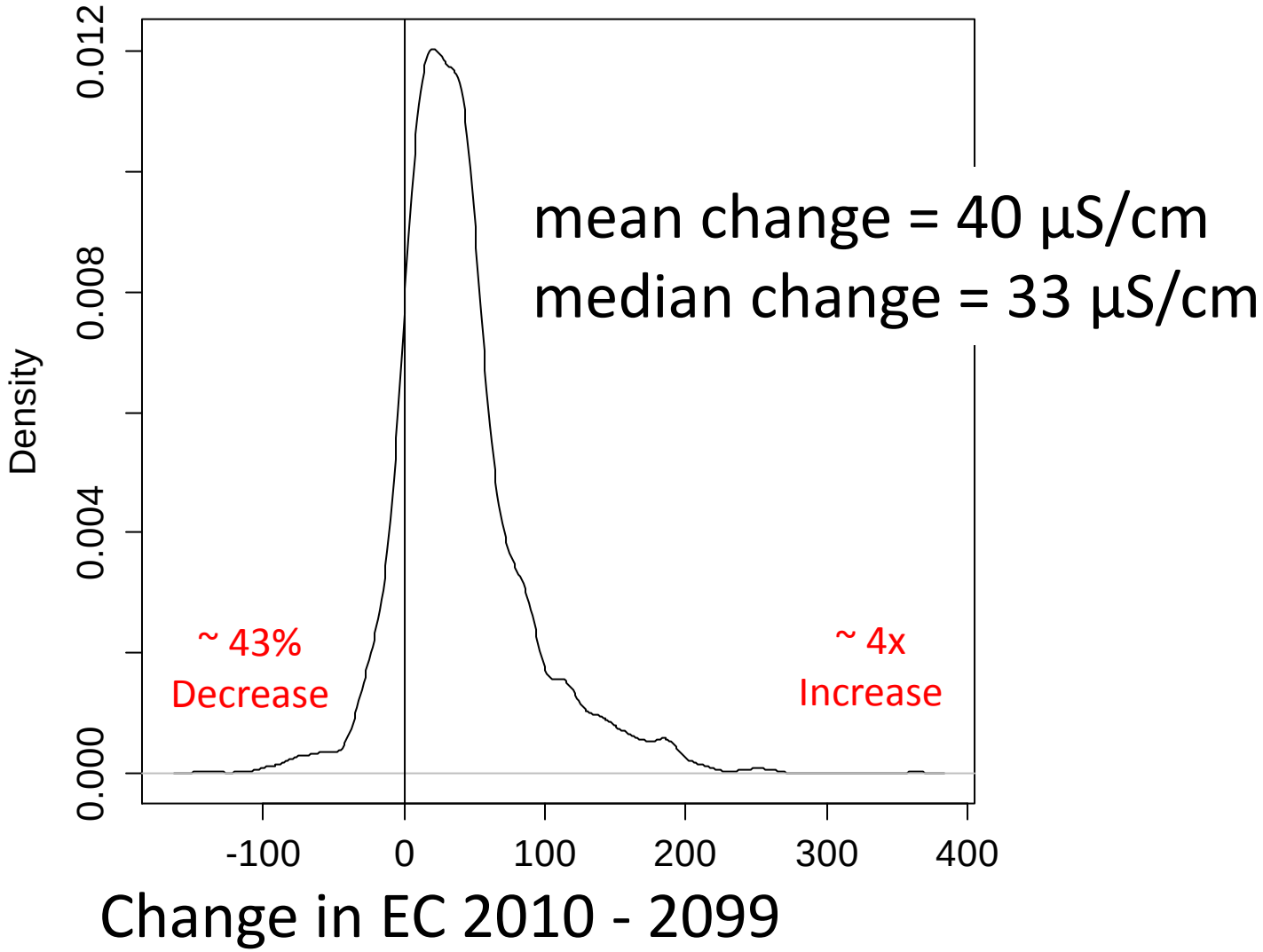




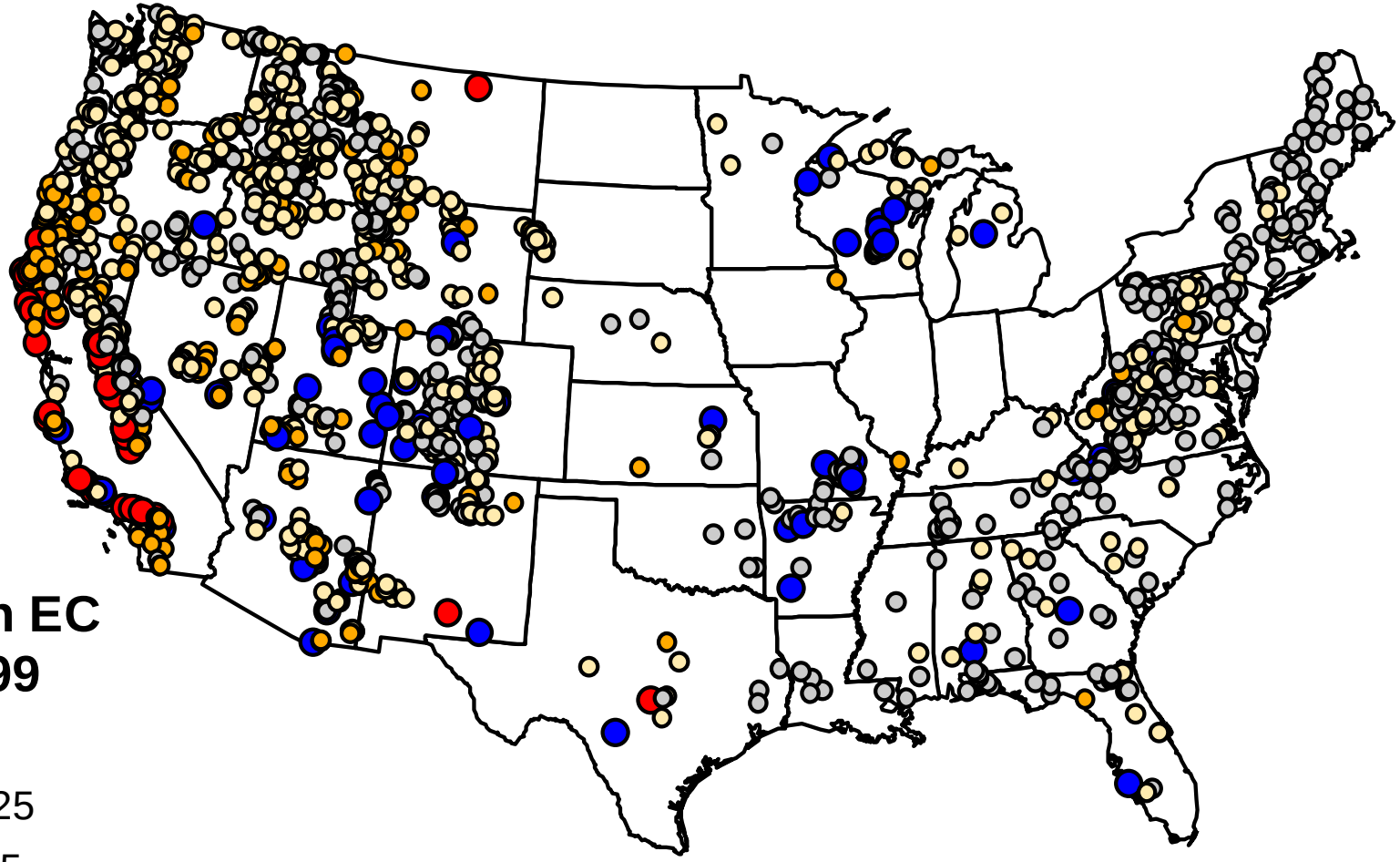
Results: EC Model – Climate Predictors



Effects of Climate Change on EC



Effects of Climate Change on EC



Change in EC
2010 - 2099

- < -25
- -25 - 25
- 25 - 75
- 75 - 150
- >150

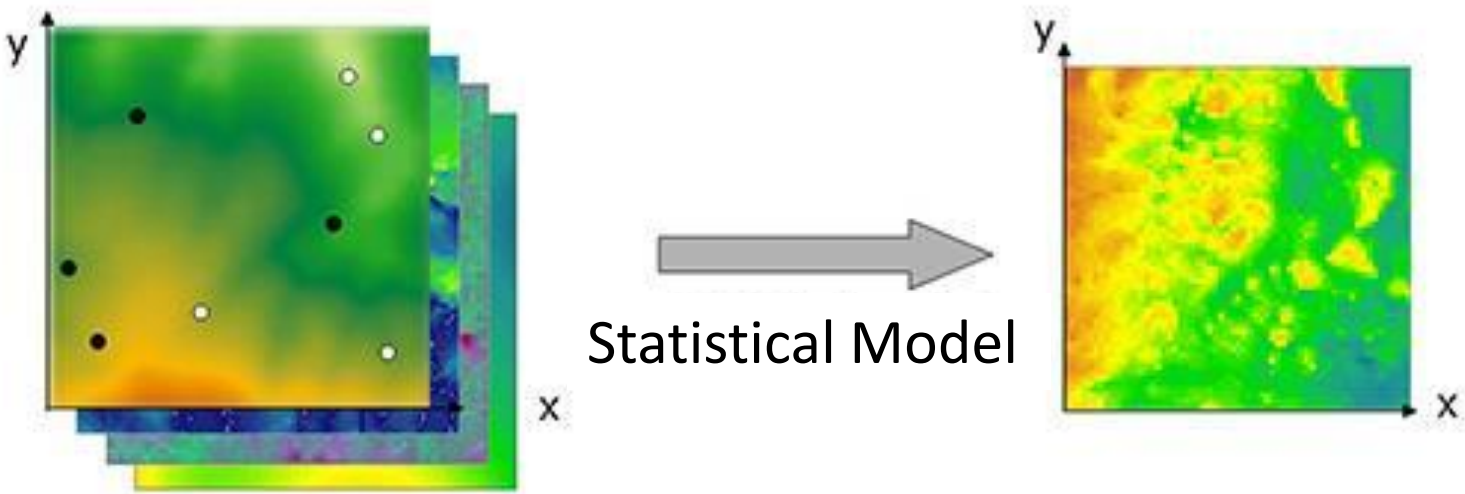
Can we empirically model distributions of species of concern?

Threatened and Endangered

Invasive and Nonnative

Can we empirically model distributions of species of concern?

Species Distribution Models (SDMs), AKA Niche Models



Field records & maps of environment

Probability species is present

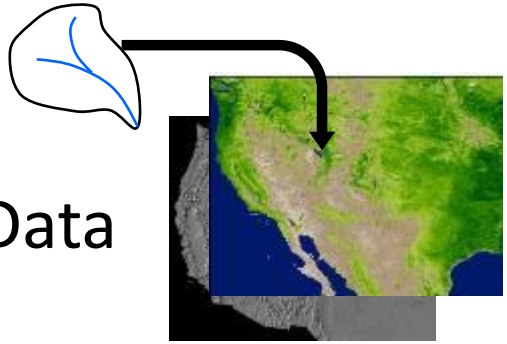
NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

E-DNA & Genetically ID'ed samples

- Simple inexpensive sampling protocol
- Works well with rare & cryptic spp
- High detection rates (> 0.8)

NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

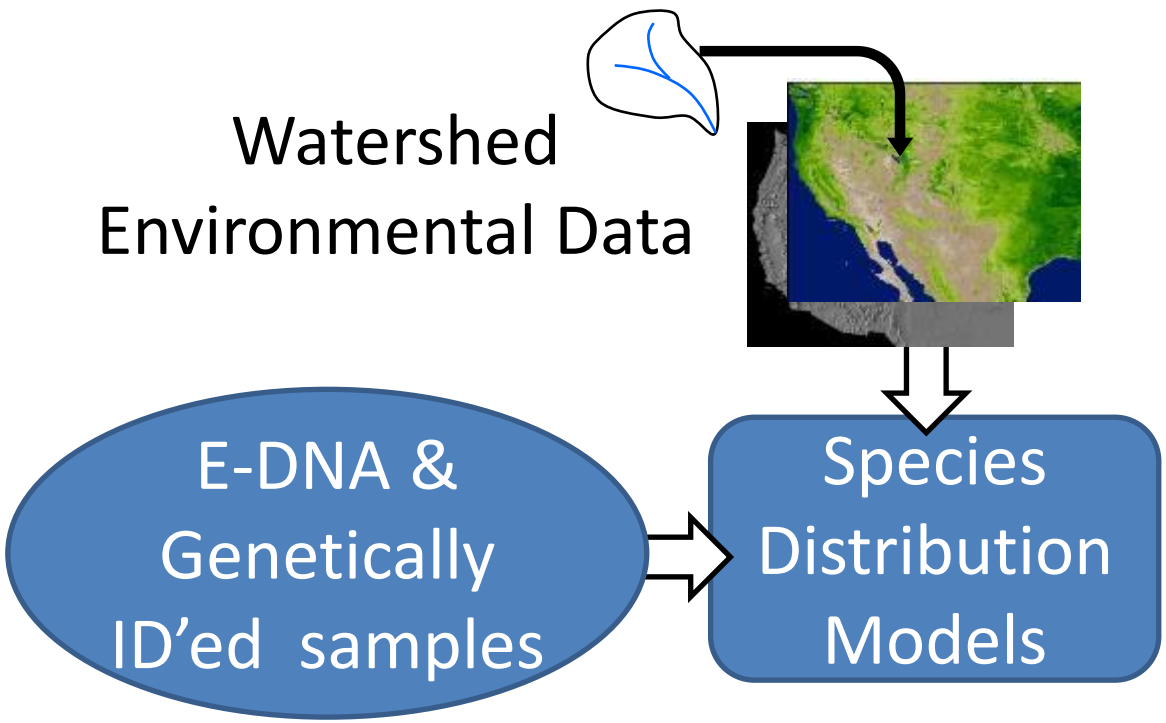
Watershed
Environmental Data



E-DNA &
Genetically
ID'ed samples

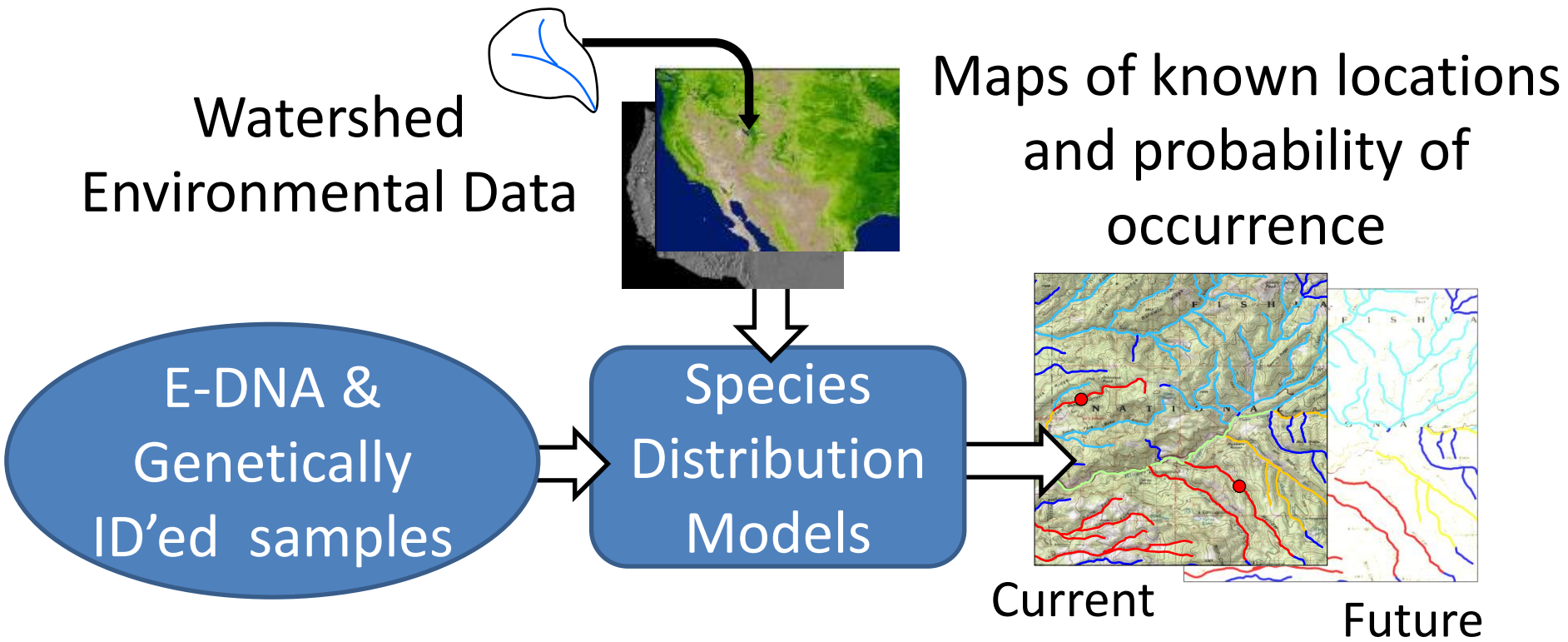
From satellite :
Temperature, ET, & vegetation
From other sources:
Geology, topography, hydrology, atmospheric deposition, etc.

NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

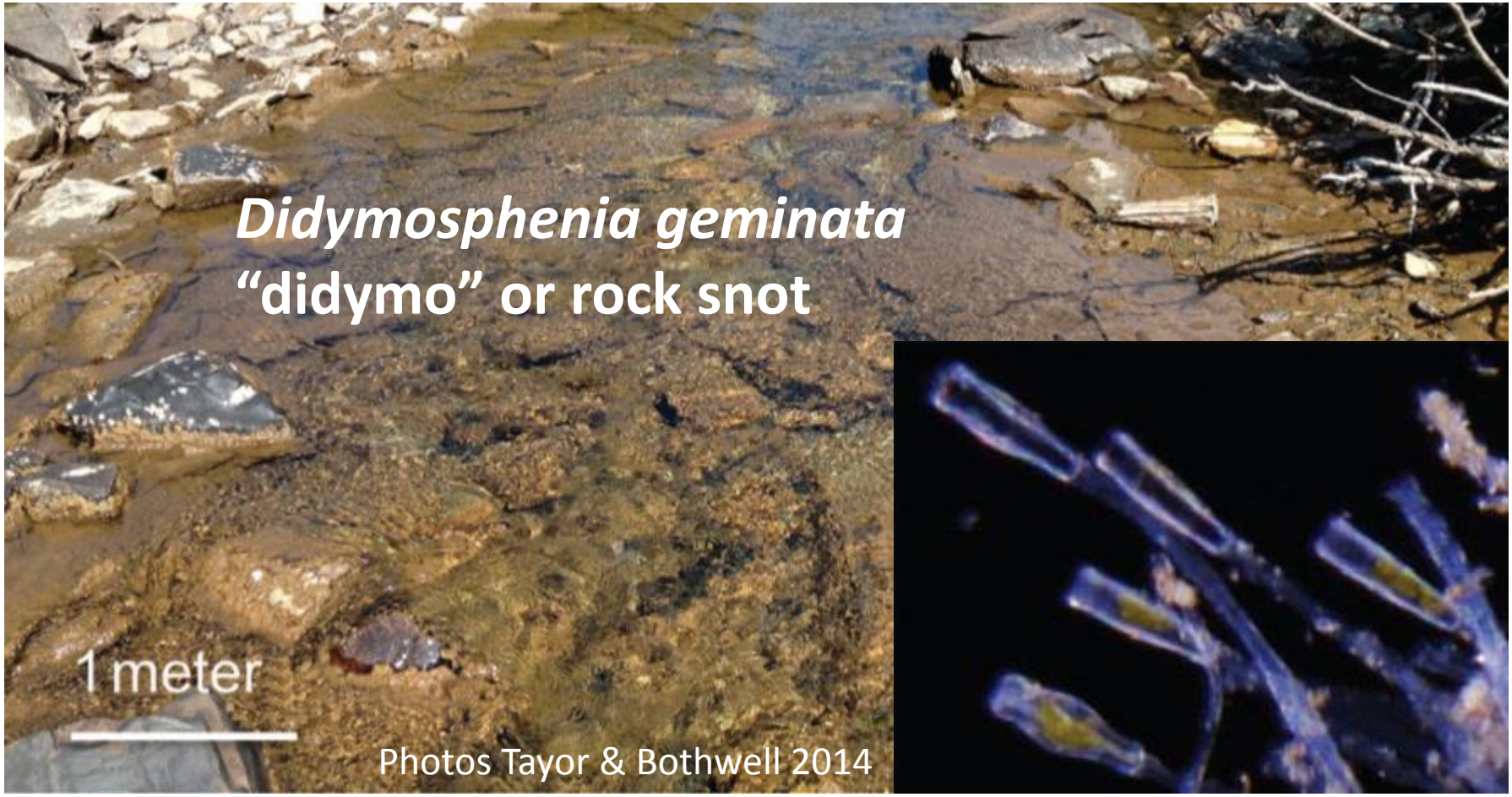


Random Forests
Maximum Entropy

NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

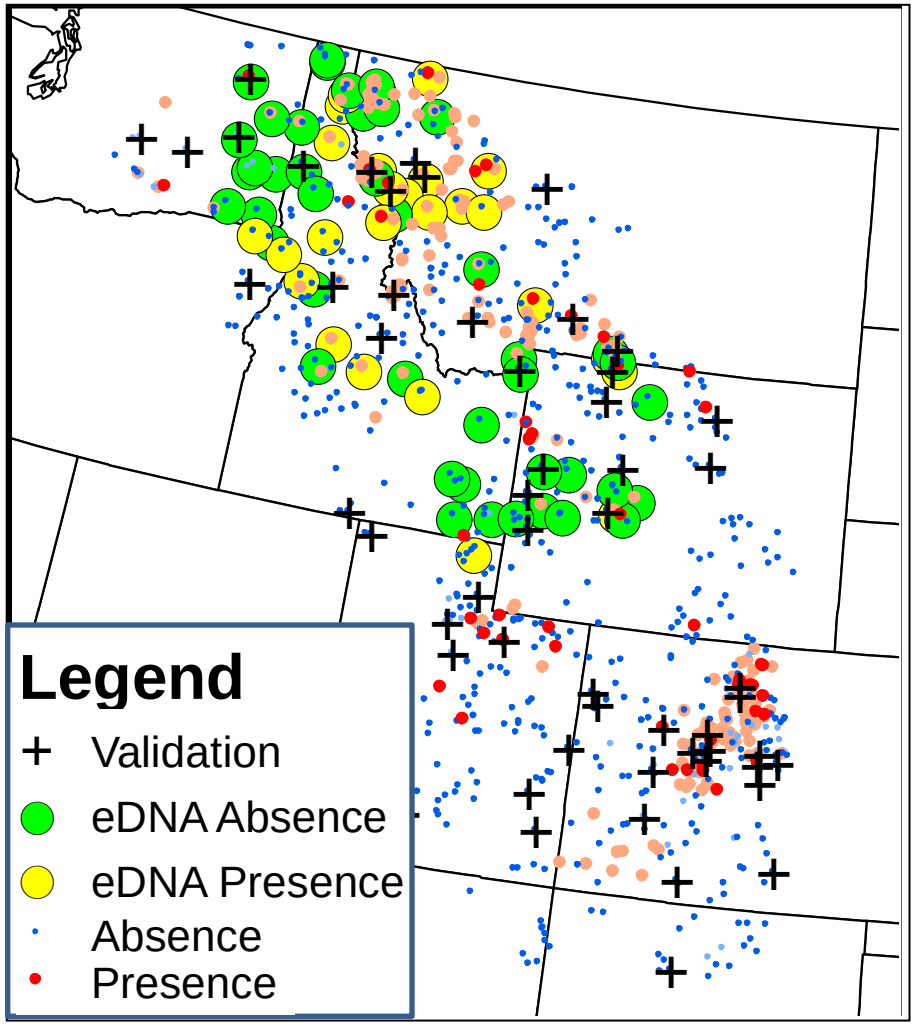


NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)



NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

Data	Absence	Presence
eDNA only	42	28
All	684	292

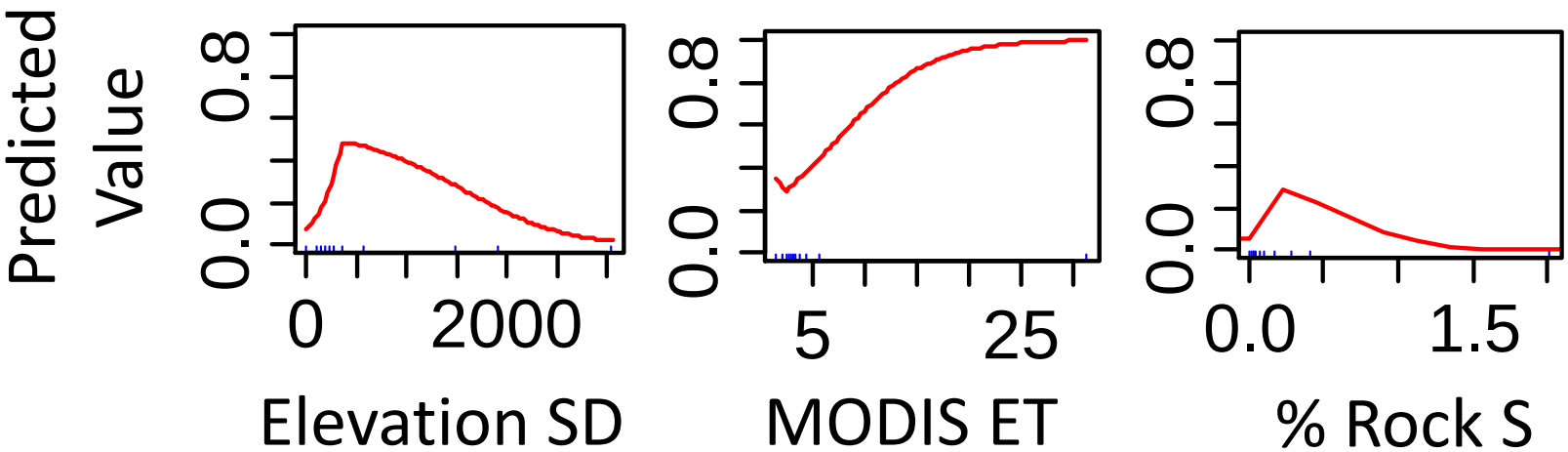


NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

MaxEnt Model	AUC	Sensitivity	Specificity	Correct Class Rate	True Skill Statistic	# of predictors
eDNA only	0.77	0.93	0.62	0.70	0.55	8
All Data Static	0.94	0.86	0.95	0.93	0.81	15
All Data Temporal	1.00	1.00	1.00	1.00	1.00	14

NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

Top 3 Predictors (of 14) in All Data Temporal Model



NASA Ecological Forecasting - System for Mapping and Predicting Species of Concern (SMAP-SOC)

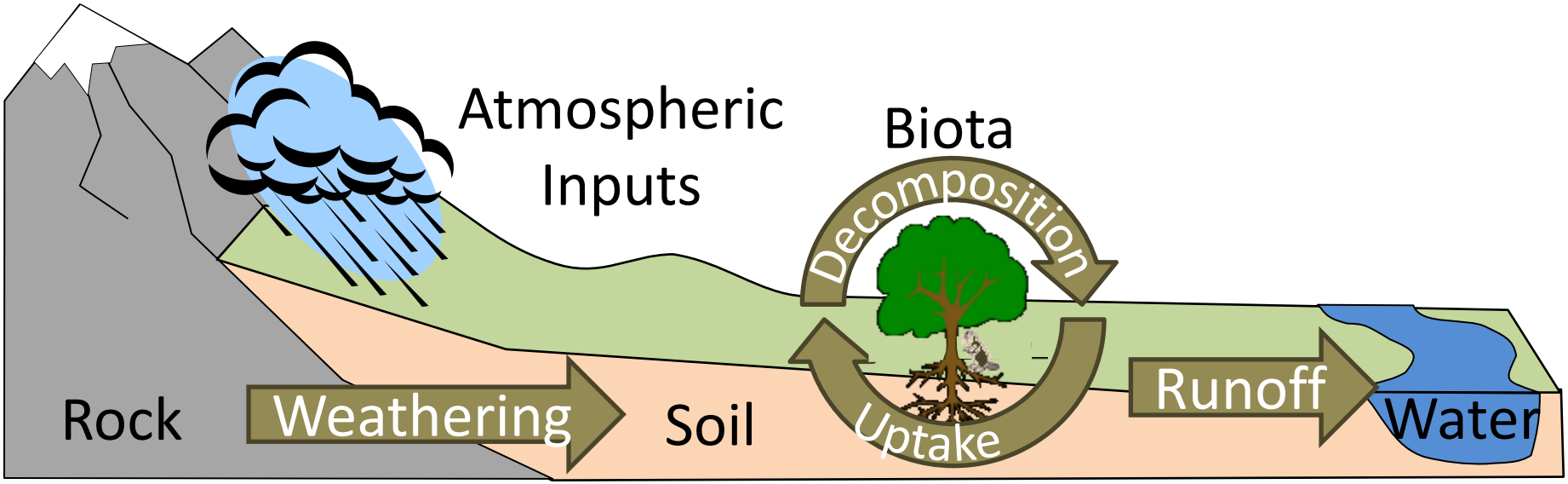
Maps of known locations and probability of occurrence

Current

Future

Conclusions

1. By taking advantage of new data and modeling techniques we can effectively model:
   A. Natural water chemistry
   B. Distributions of species of concern
2. These model give us both predictive ability and a better understanding of environmental influences



Acknowledgments:

Chuck Hawkins
Ryan Hill
Jake Vander Laan
Scott Miller
Michelle Baker
Pete Kolesar
Helga Van Miegroet
Matt Baker

This research is funded by
U.S. EPA - Science To Achieve Results (STAR) Program
Grant # R-82863701

Questions