Spatial modeling of elevated groundwater nitrate concentrations using random forests and regression kriging

Andy Canion
Dean Dobberfuhl
Lori McCloud

How do we find vulnerable areas to best focus projects?

Screening Approach:
Where do nitrogen sources and geologic vulnerability overlap?

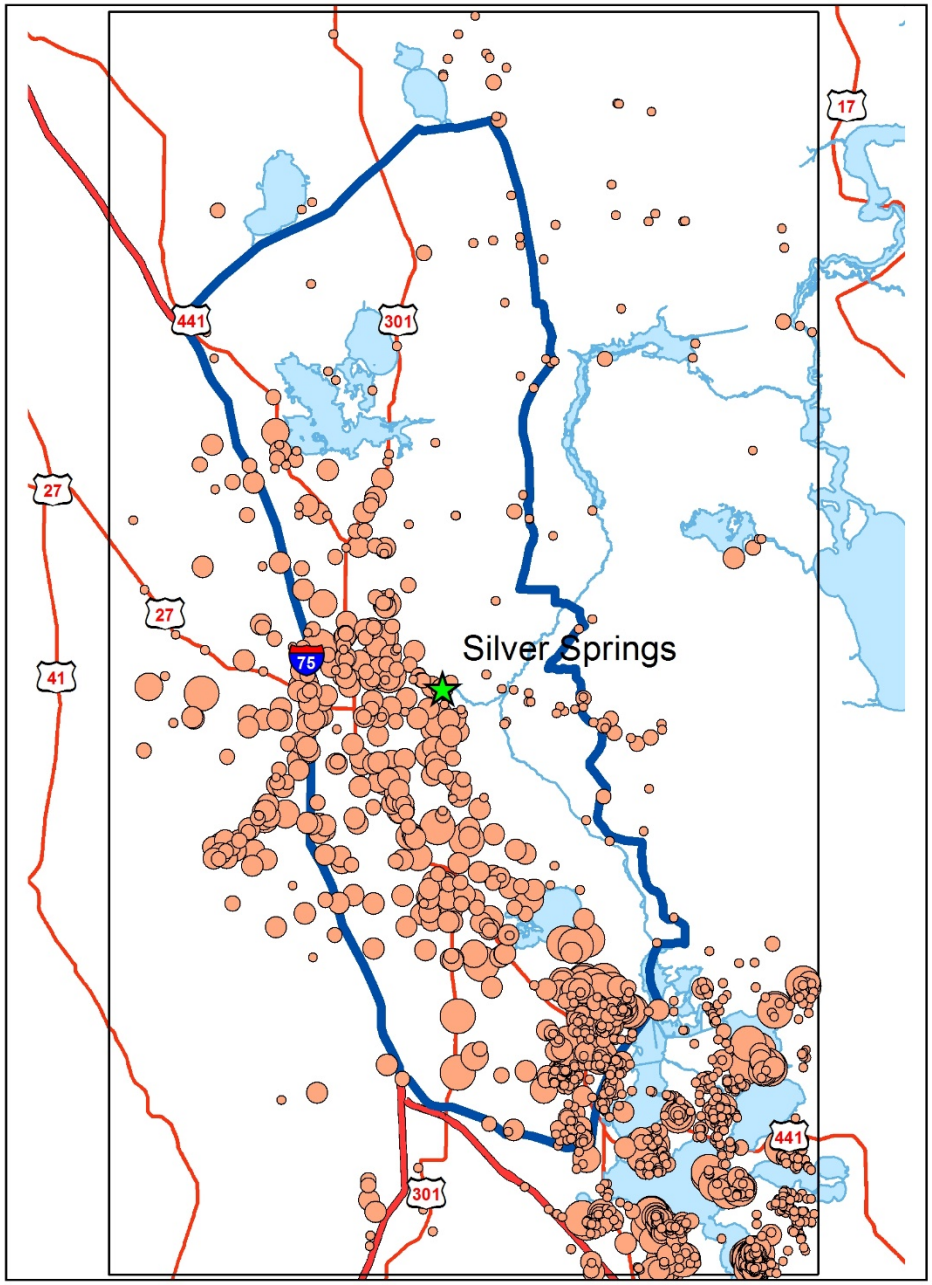
Can we use the wealth of nitrate data from well monitoring to focus efforts?

Mean Nitrate

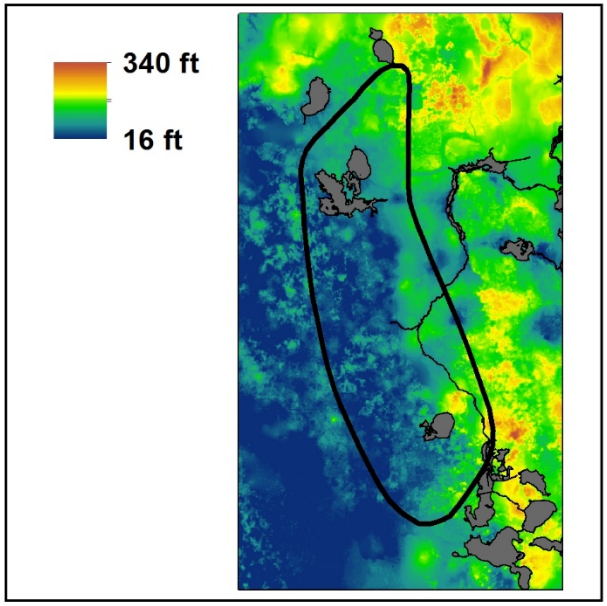
- 0.0 - 0.35 mg/L
- 0.35 - 1.0 mg/L
- 1.0 - 3.0 mg/L
- 3.0 - 5.0 mg/L
- 5.0 - 17.0 mg/L

Silver Springs BMAP Boundary

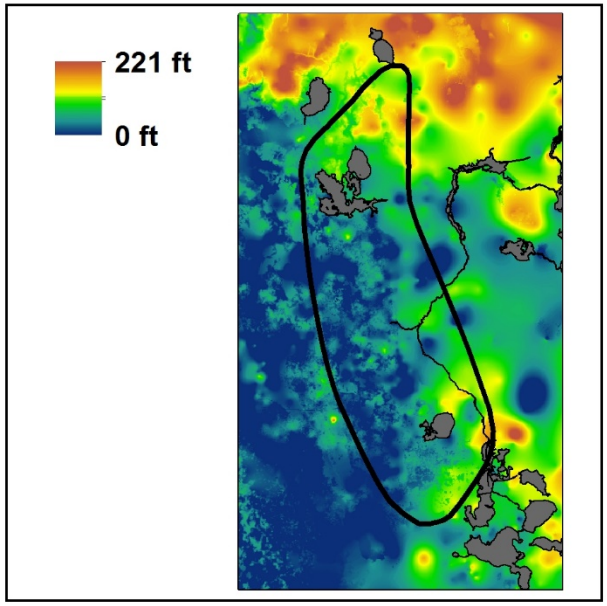
Dataset	# Wells	POR
FDEP Public Water System (PWS)	499	2009 – 2013
FDEP Water Supply Restoration (WSRP)	968	2000 – 2015
SJRWMD	44	2000 – 2013
SWFWMD	7	2000 – 2013
USGS (Phelps 2004)	36	2001 – 2002



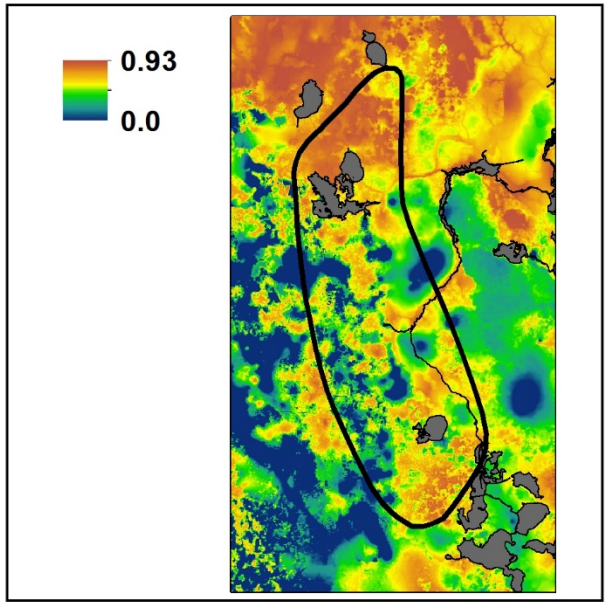
Depth to UFA



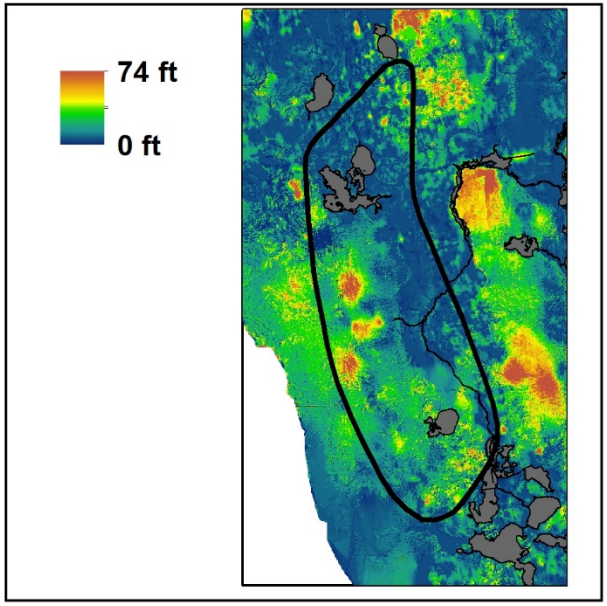
ICU Thickness



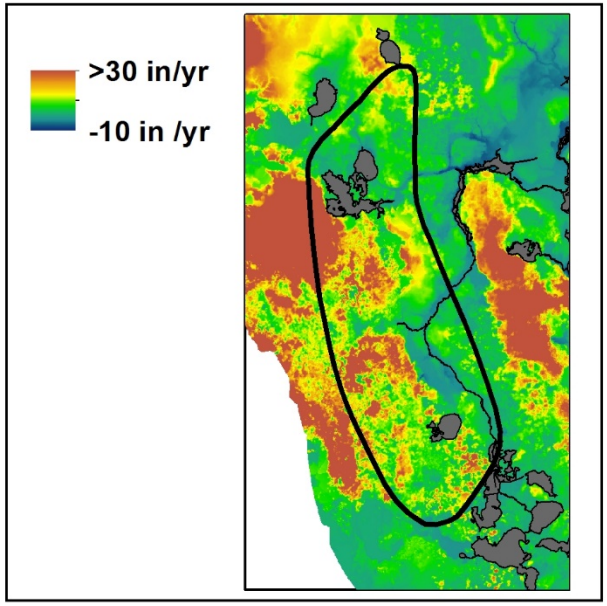
ICU to OB ratio



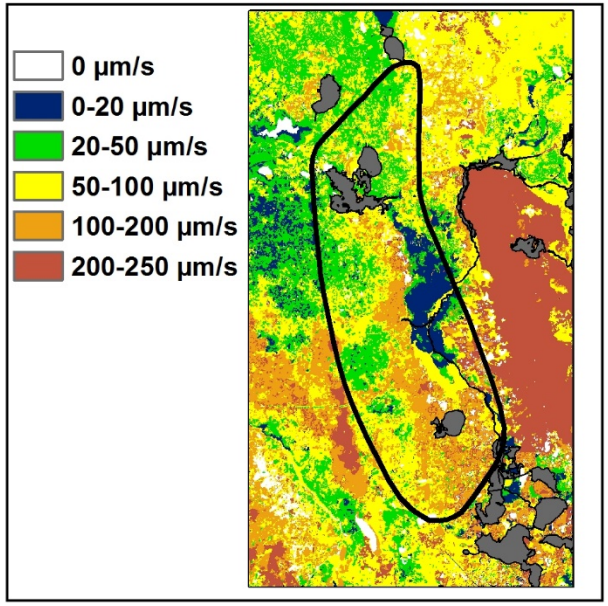
Water Table Depth



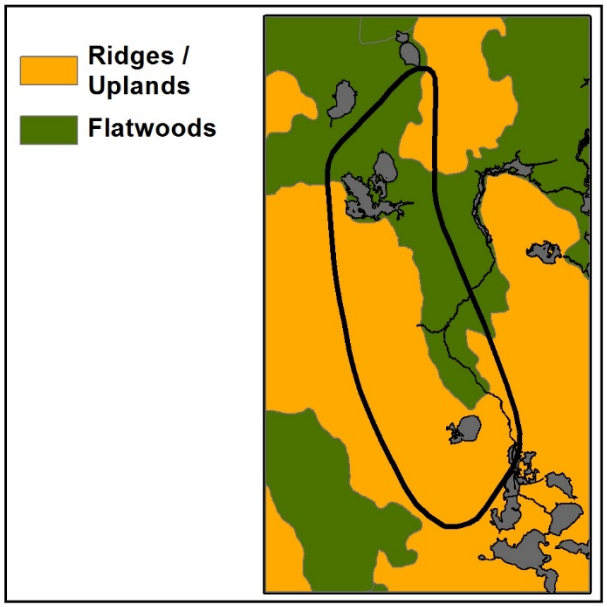
UFA Recharge



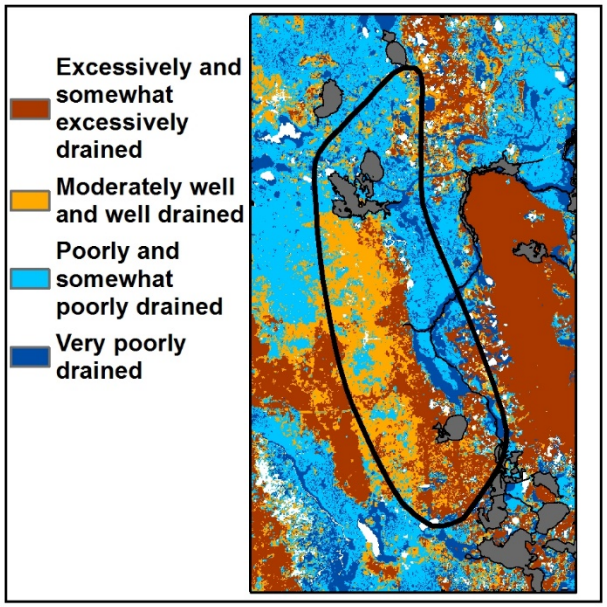
Soil Hydraulic Conductivity (Ksat)



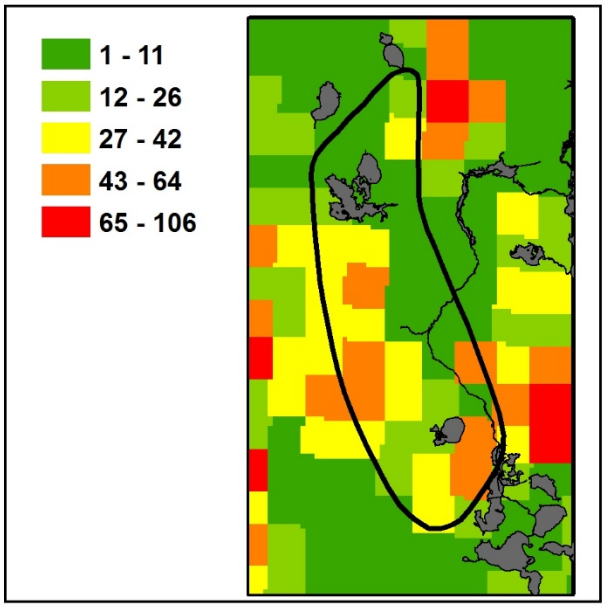
Ecoregions



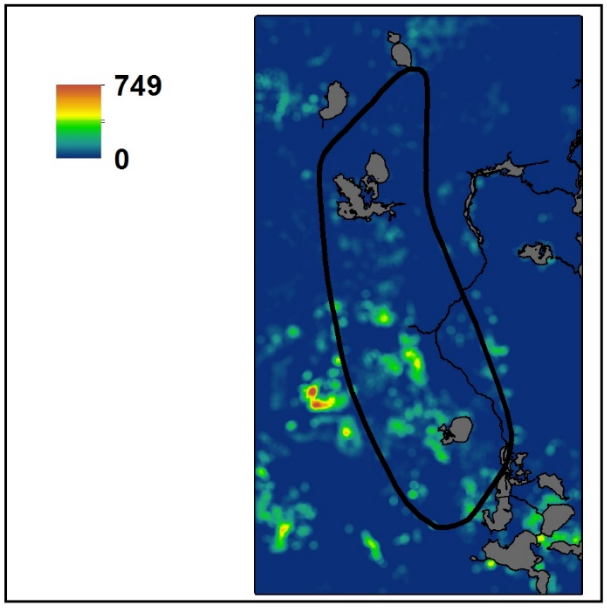
Soil Drainage Class



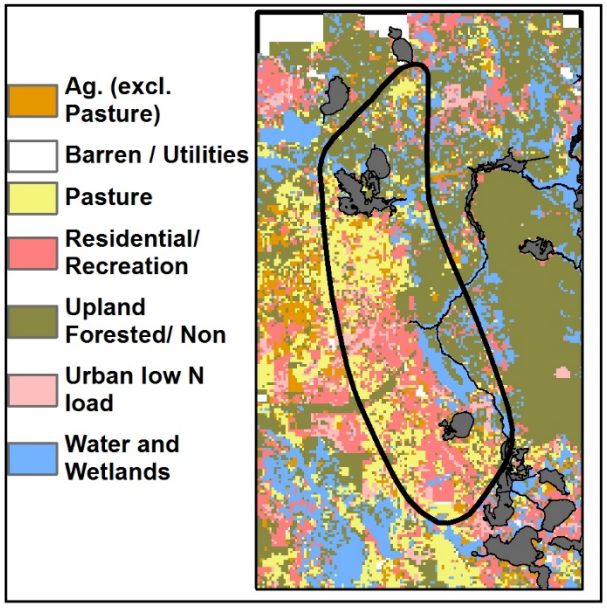
Sinkholes - Number per 25 sq mi



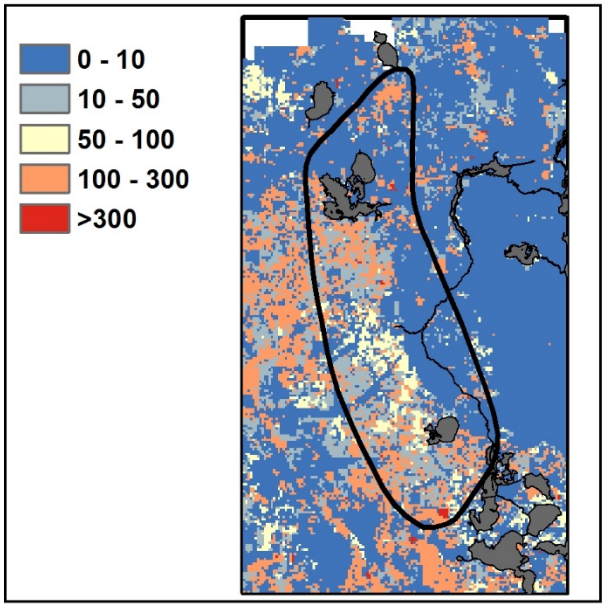
Septic Tanks- Number per sq. km



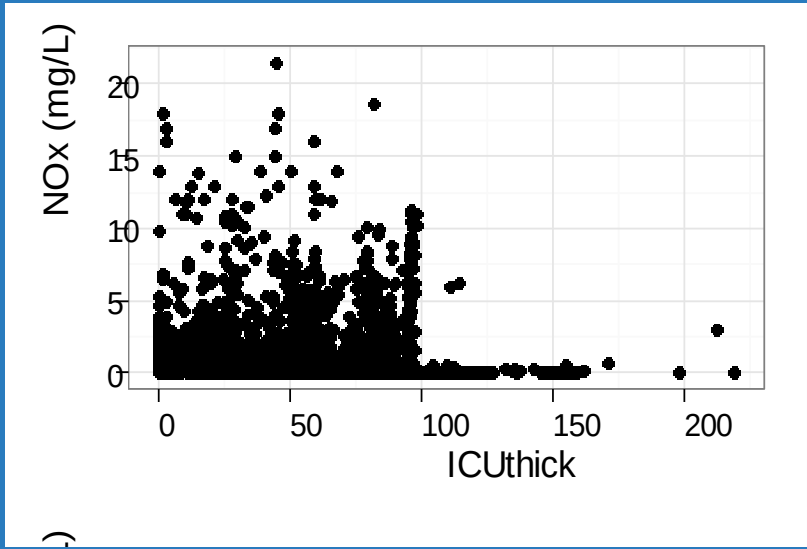
Dominant Land Use (25 ha Grid)

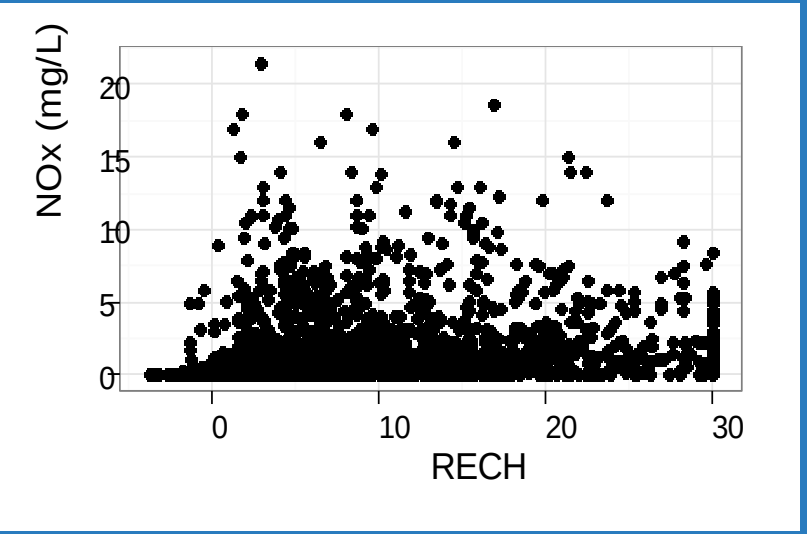


Land Surface N Load (Kg/Ha)

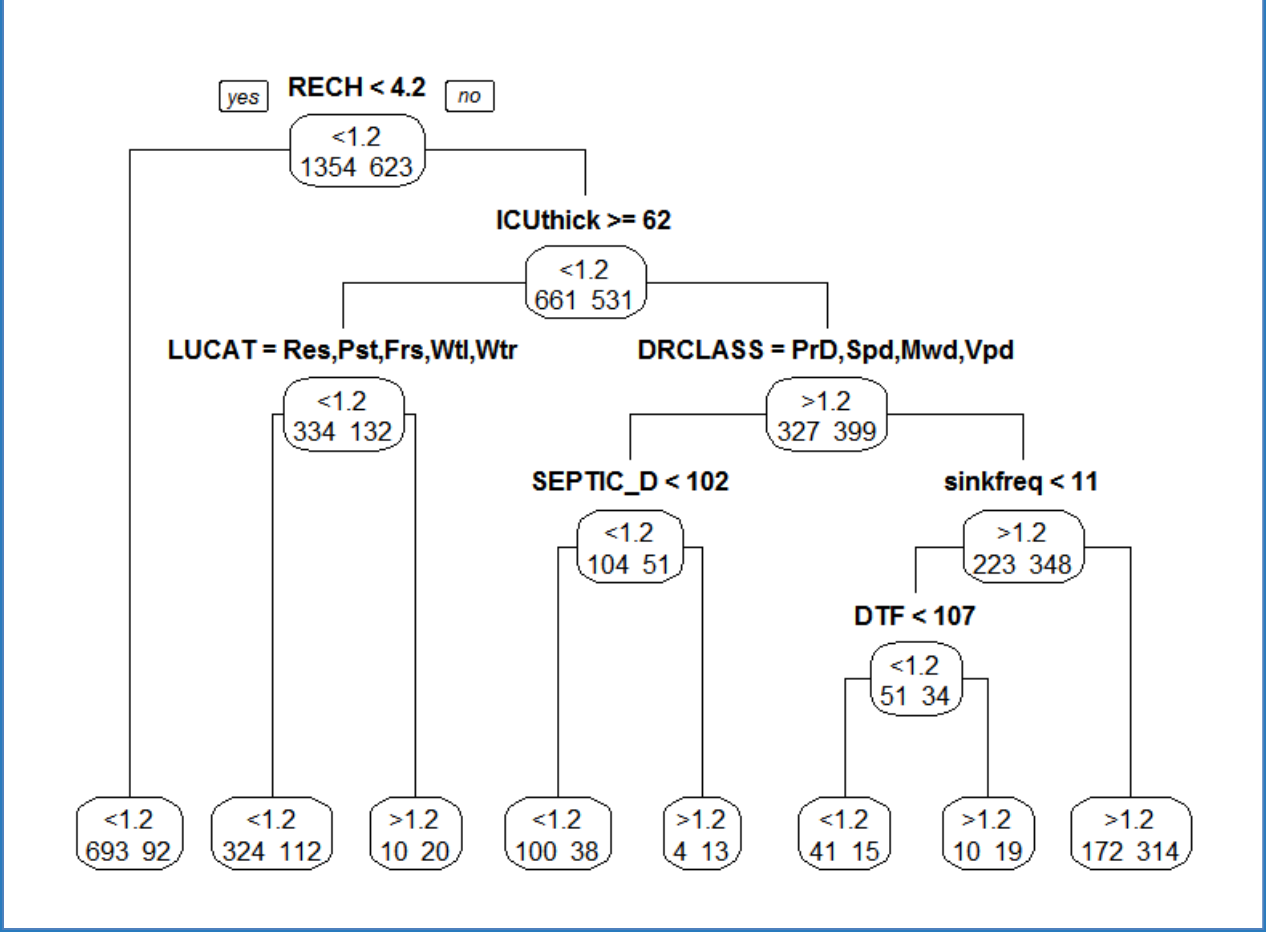


Is there information in the data?

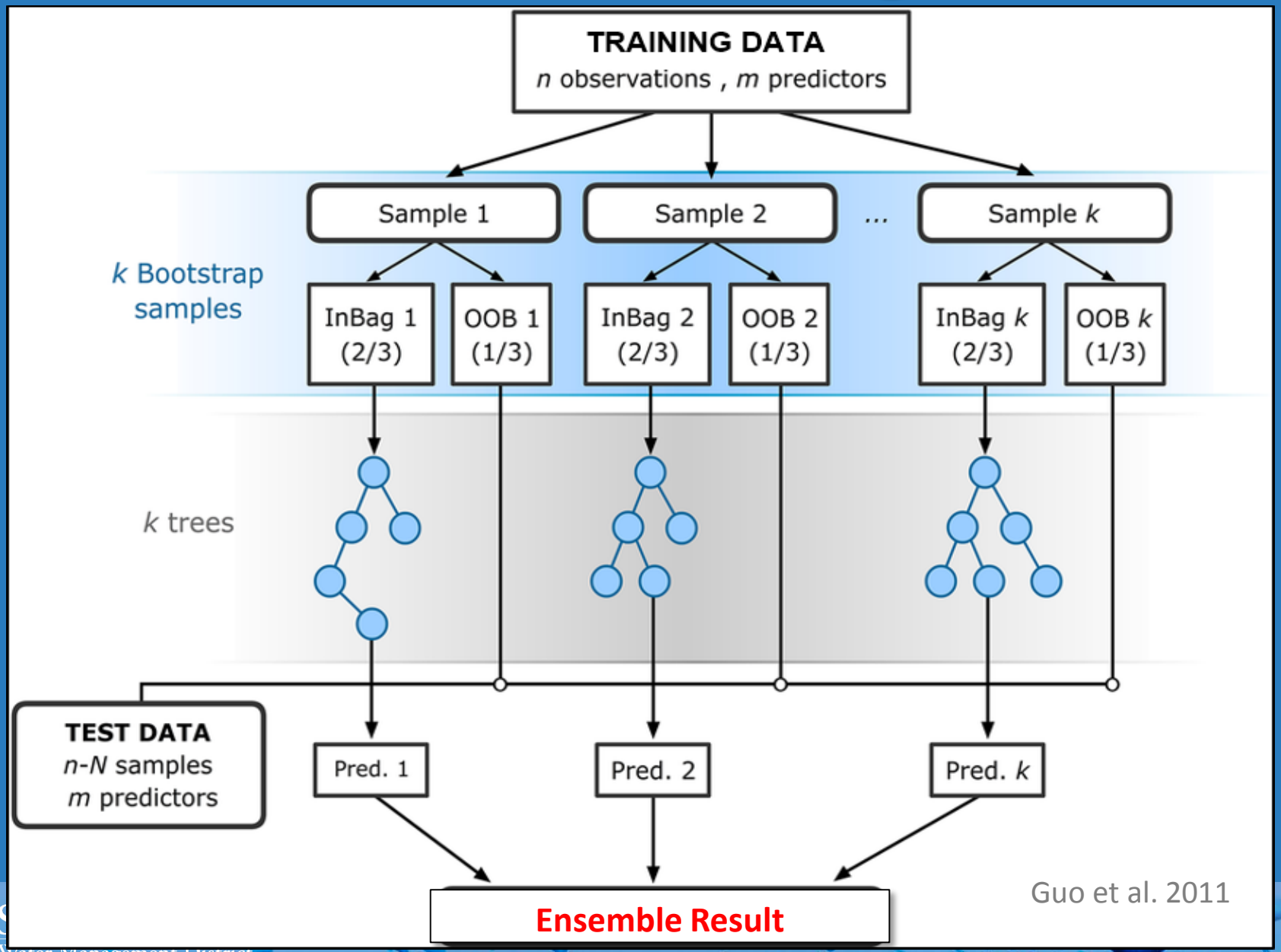




Regression Tree Approach



Random Forest Algorithm

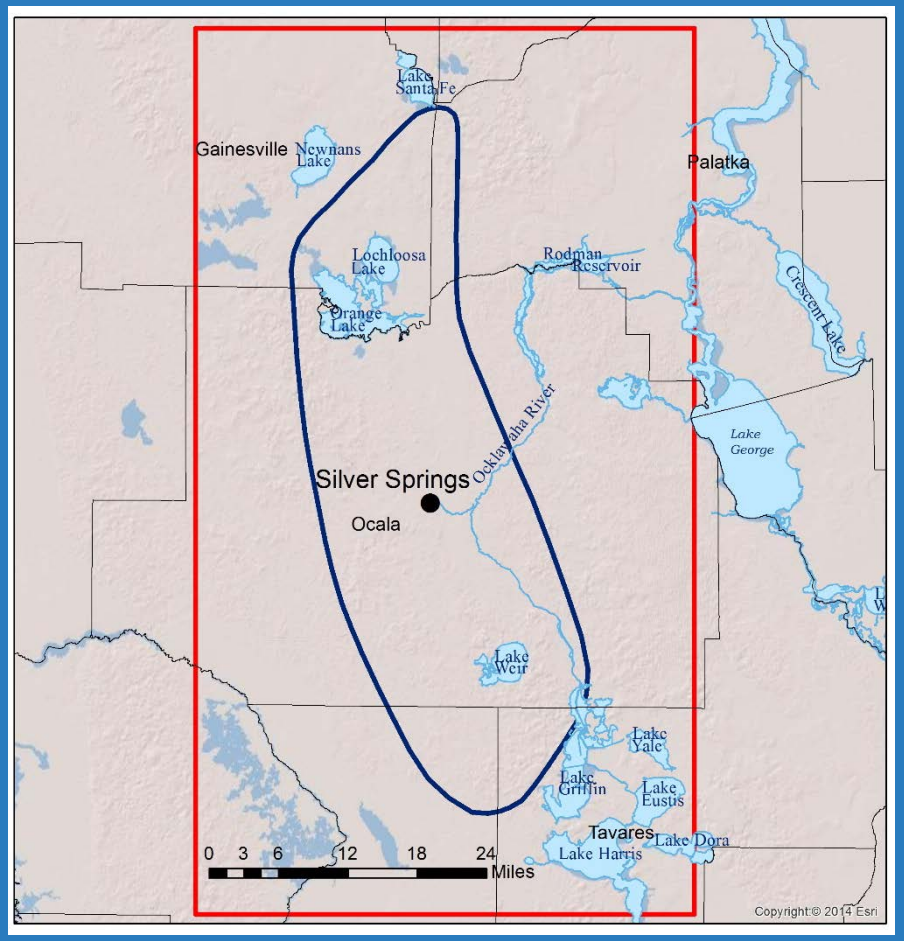


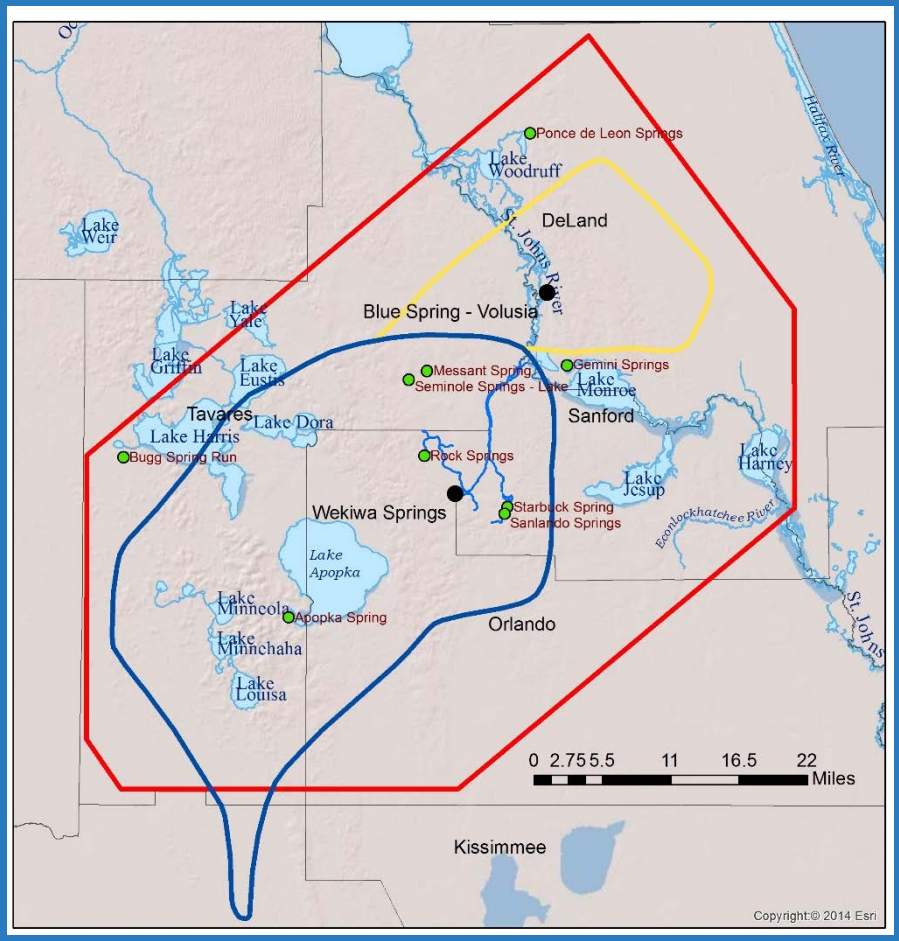
Guo et al. 2011

Random Forest Classification Model

Probability of nitrate exceeding a chosen threshold is predicted by:

- Well Depth
- Confinement
- Aquifer Depth
- Recharge
- Water Table Depth
- Soil Hydraulic Conductivity
- Soil Drainage Class
- Ecoregion
- Sinkhole Frequency
- Land Use
- Nitrogen Load
- Septic Density





Legend

- Silver Springs Main Vents
- County Boundaries
- Silver Springs Springshed
- Model boundary

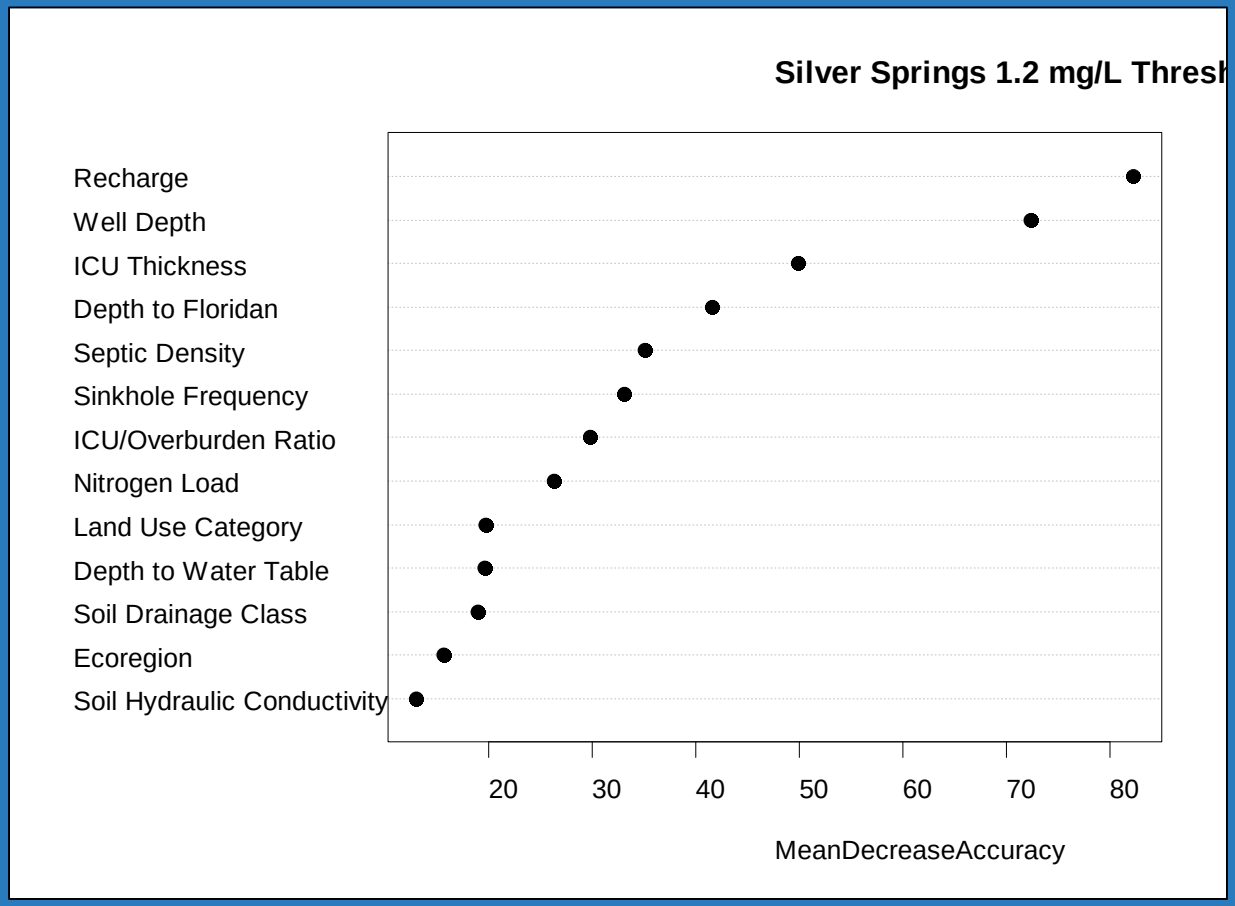
Legend

- Wekiwa and Blue Springs
- Other 2nd magnitude springs
- Wekiva River
- Wekiwa Springs Springshed
- Blue Springs (Volusia) Springshed
- Model Boundary
- County Boundaries

Model Diagnostics

Model Area	Nitrate Threshold (mg/L)	Predictor Variable Subset (m)	Number of Trees	Out-of-Bag Error	Area Under ROC Curve
Silver	0.35	2	1,000	17.31%	0.89
Silver	1.2	7	1,000	21.24%	0.84
Wekiwa/ Blue	0.35	2	1,000	21.25%	0.86
Wekiwa/ Blue	0.8	2	1,000	20.75%	0.86

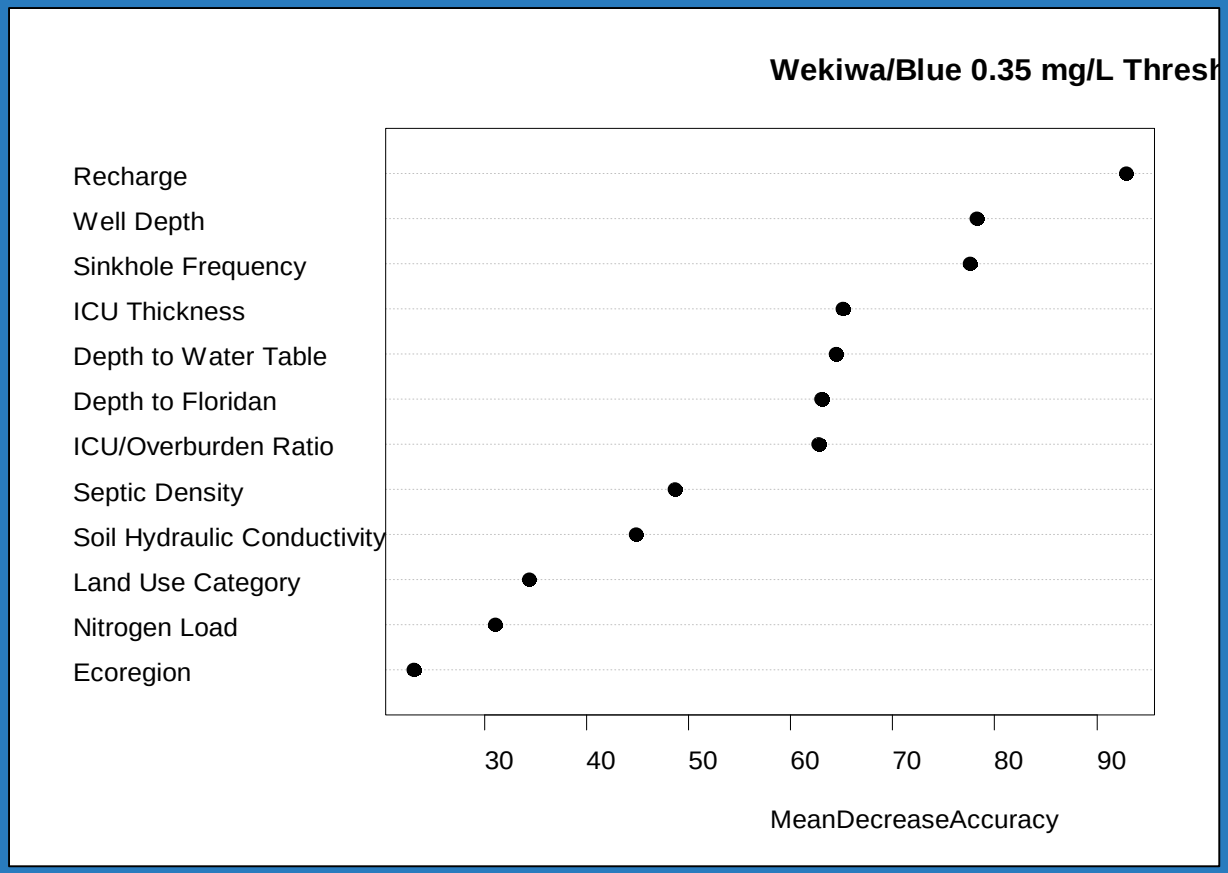
Model Diagnostics - Silver



Geology / Hydrogeology

Soils, Loading

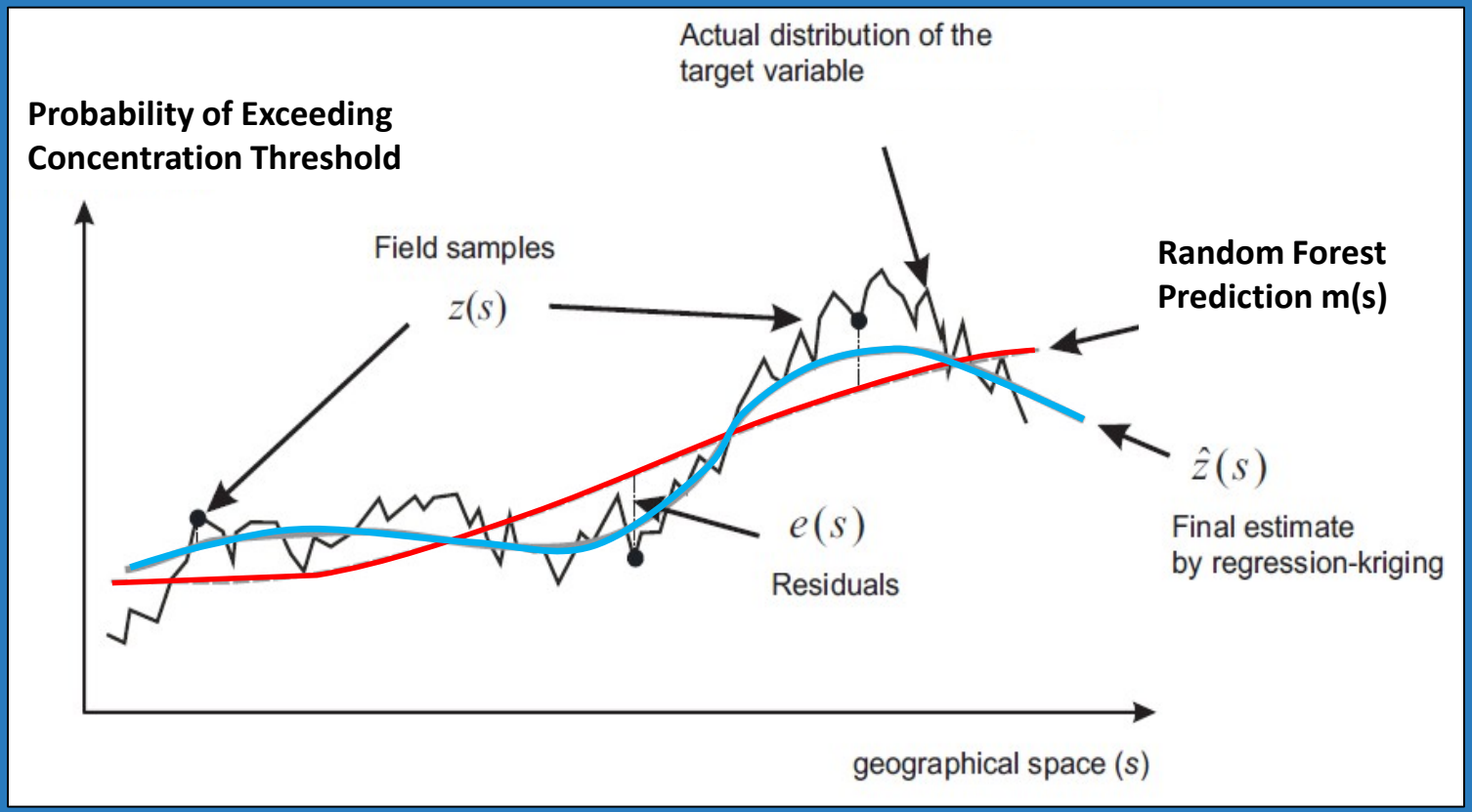
Model Diagnostics – Wekiwa/Blue

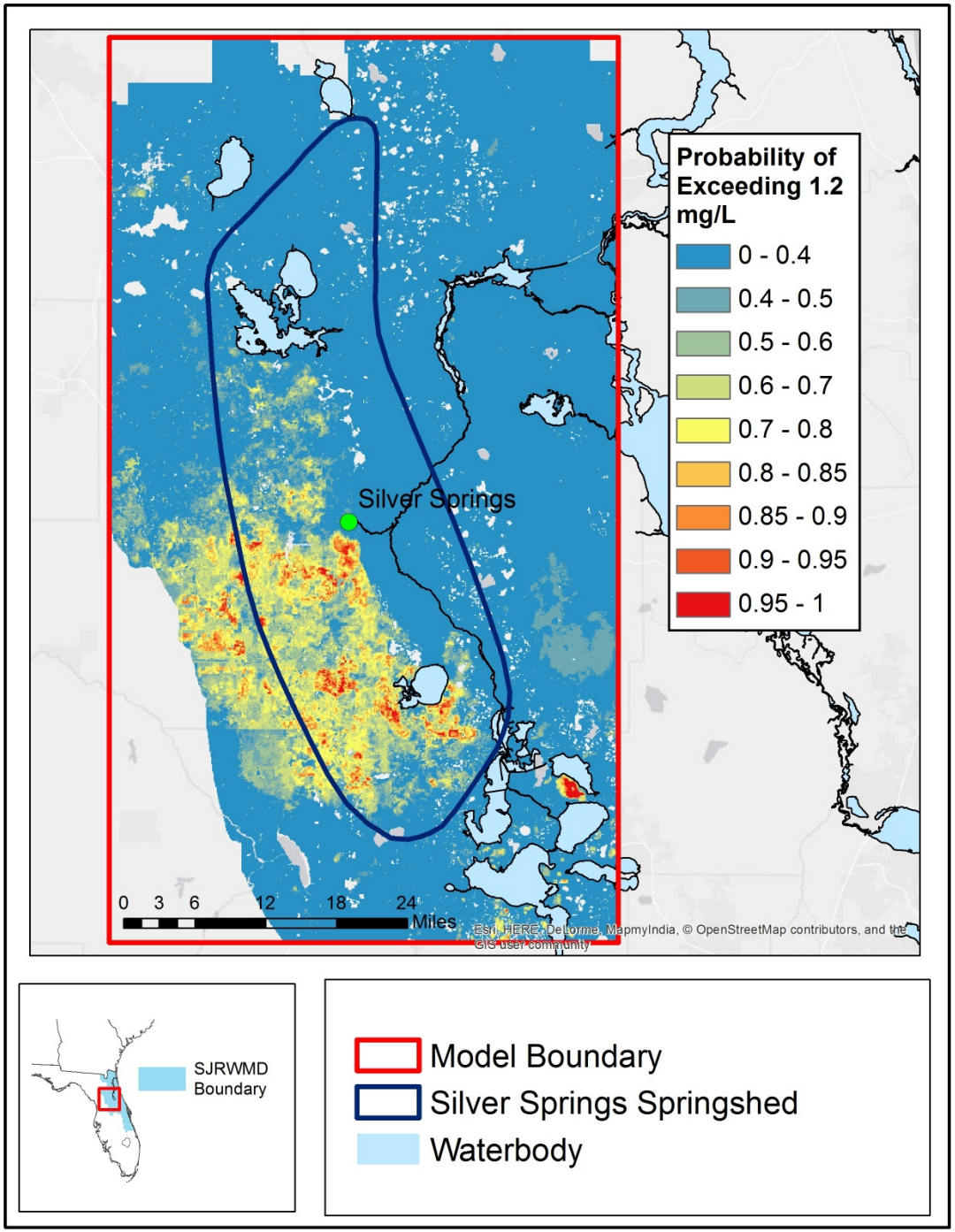


Geology / Hydrogeology

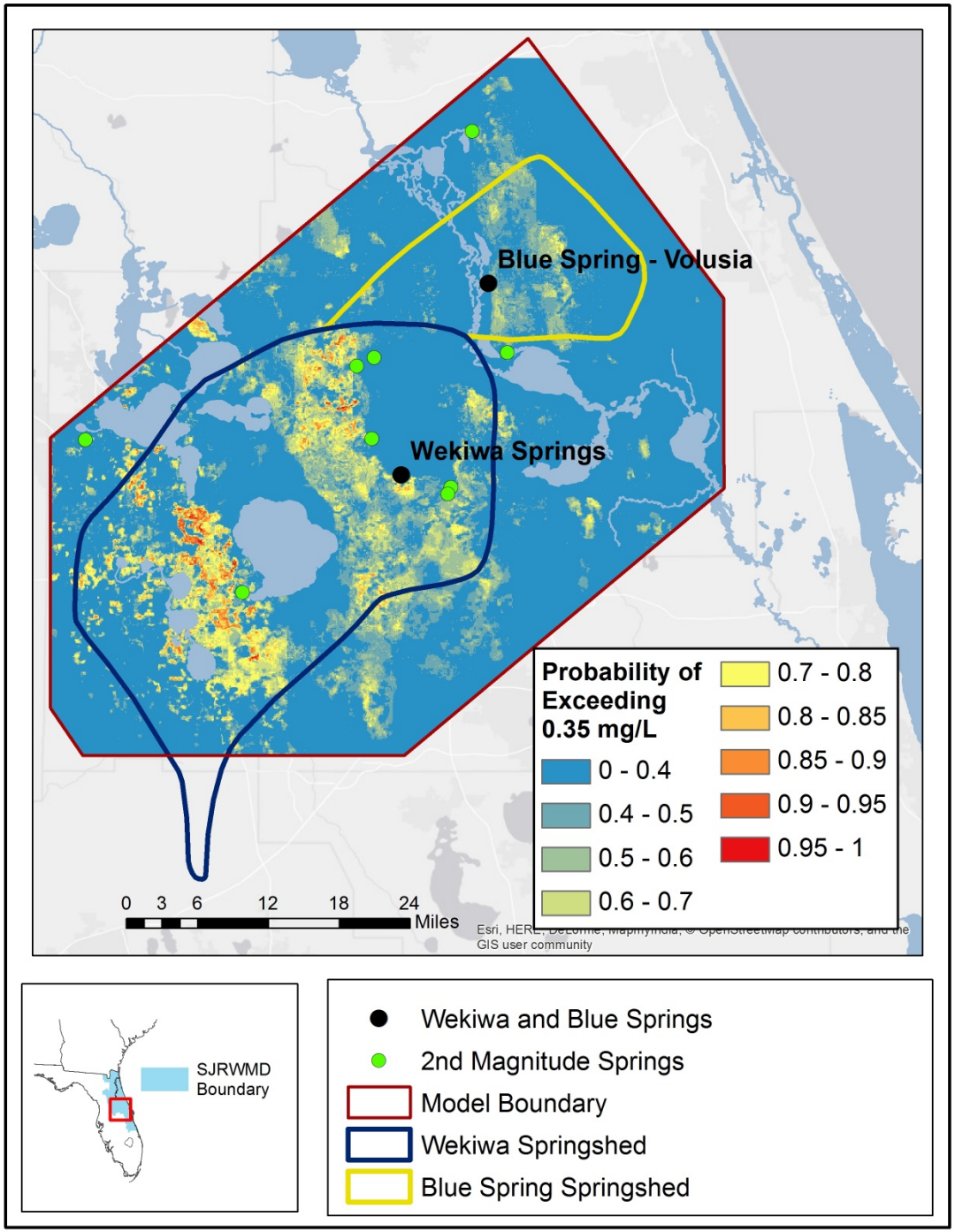
Soils, Loading

Kriging of Random Forest Residuals





Includes kriged residuals



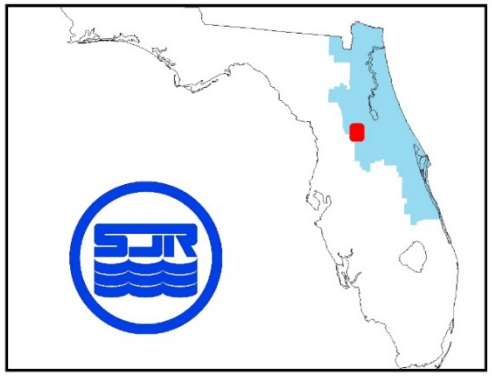
No spatial correlation in residuals – not kriged

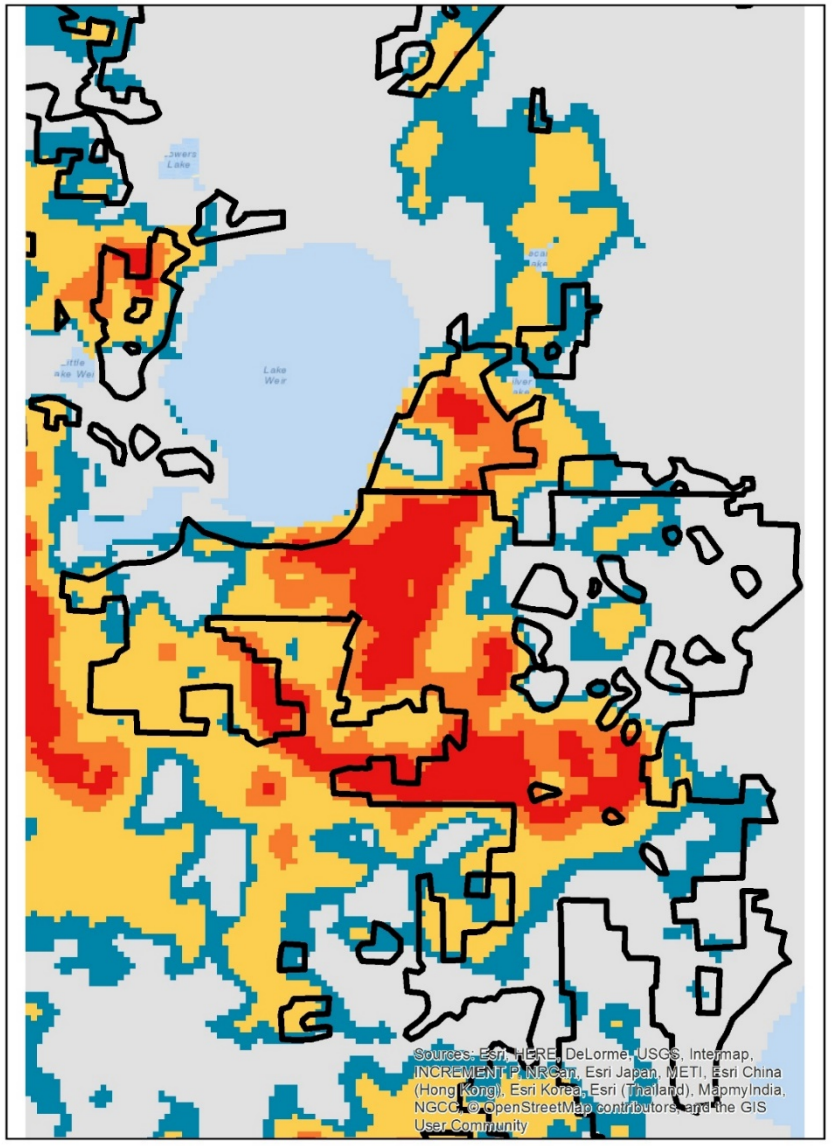
Legacy Nitrate

1970's Land Use - Orchards and Groves (USGS)

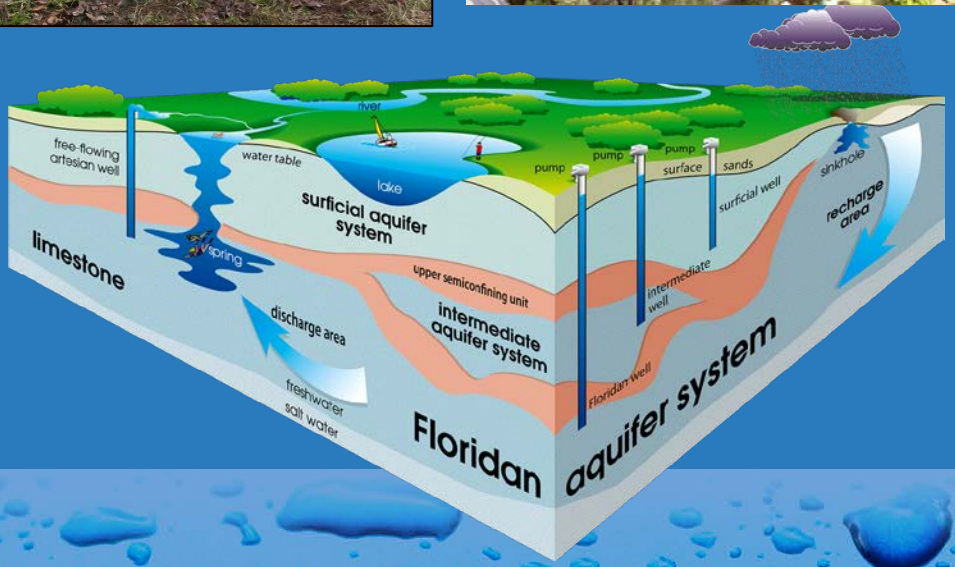
Probability of Exceeding 1.2 mg L^{-1} Nitrate

- < 40%
- 40 -50 %
- 50 - 70 %
- 70 - 80 %
- 80 - 100 %





Thank You



St. Johns River
Water Management District