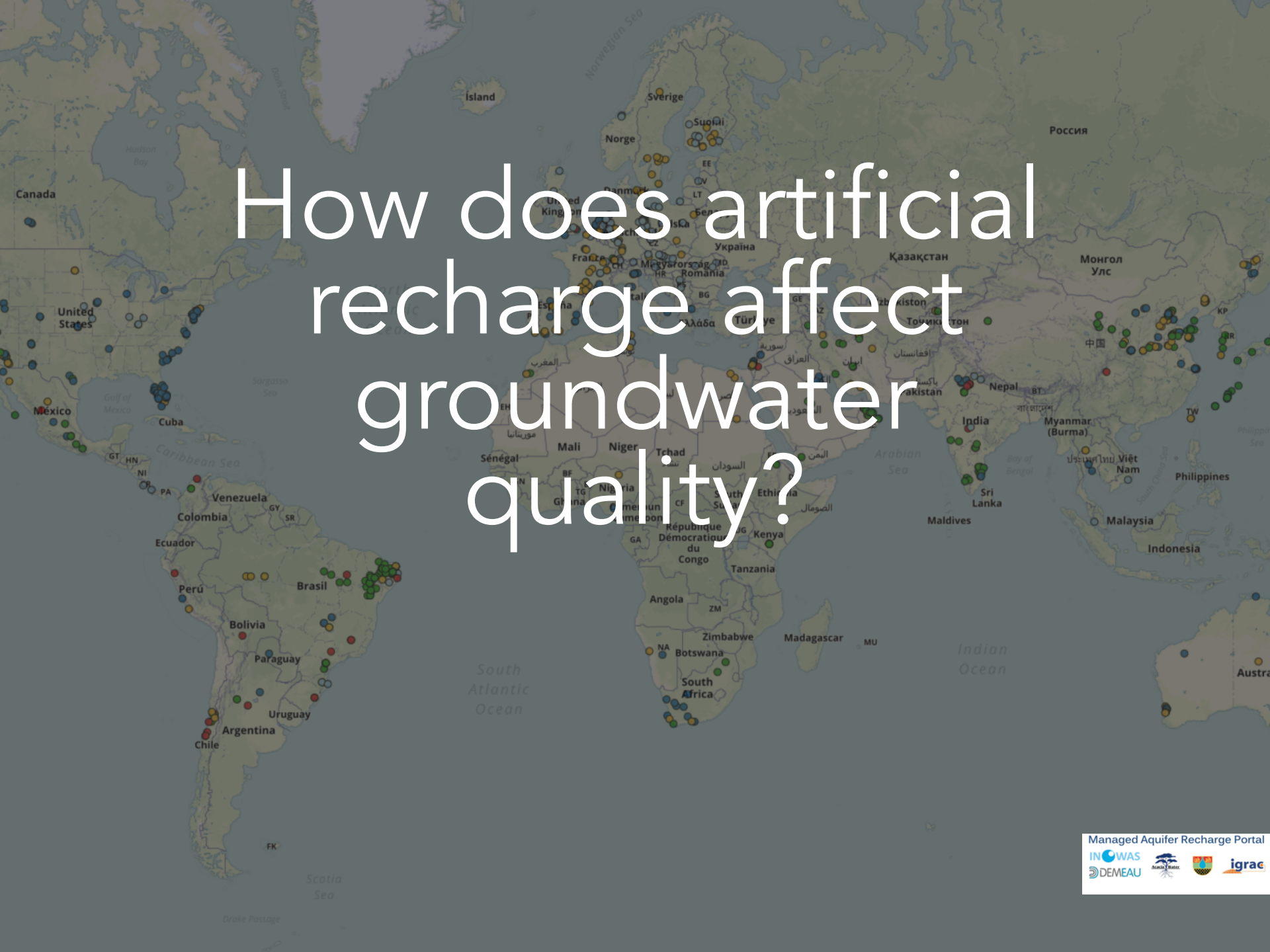
Using machine learning to incorporate water quality improvements for mapping MAR suitability

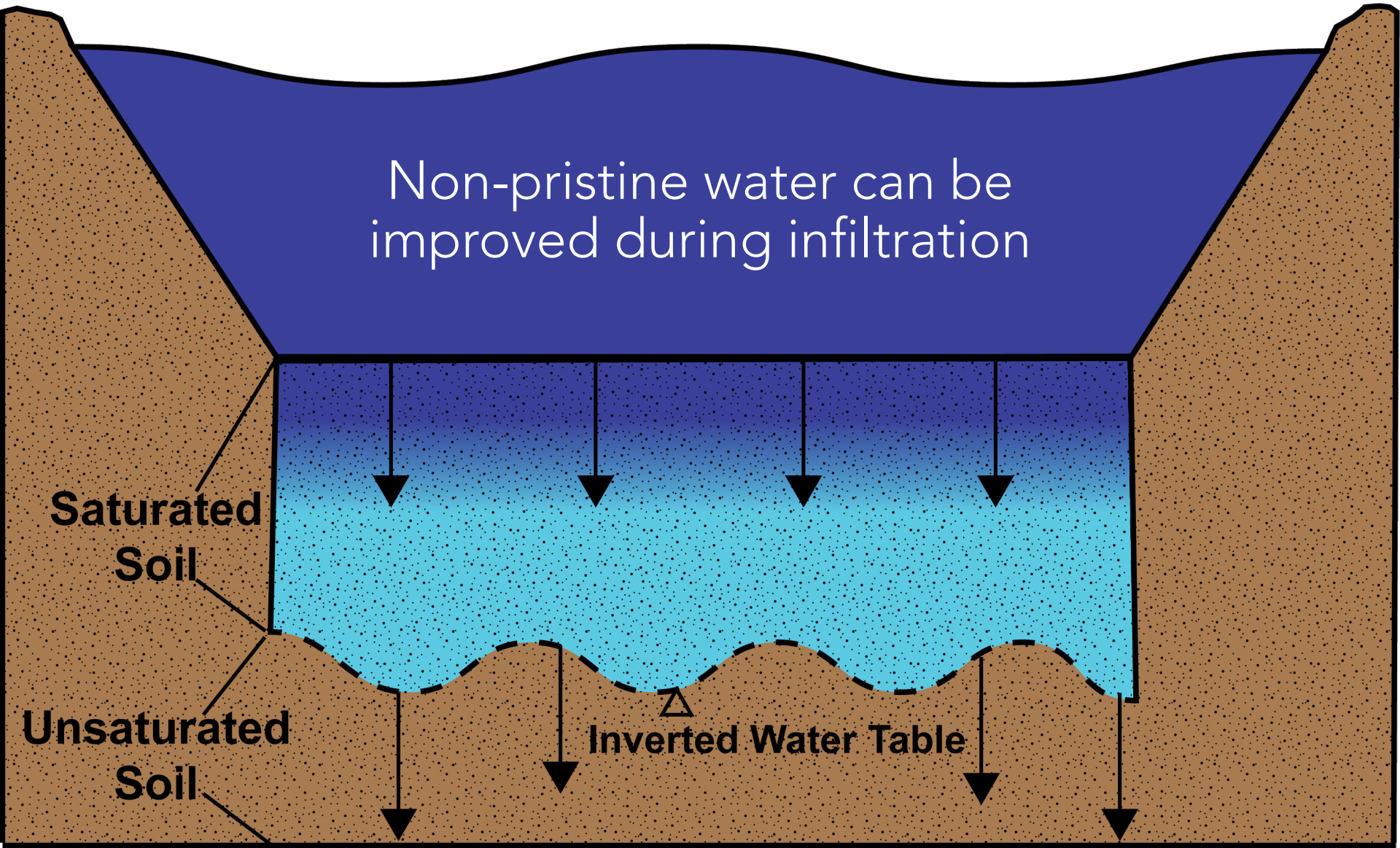
International Symposium for Managed Aquifer Recharge
Madrid, Spain
21 May 2019

Galen Gorski, Andrew Fisher, Sarah Beganskas, Jenny Pensky, Hannah Dailey, and Calla Schmidt

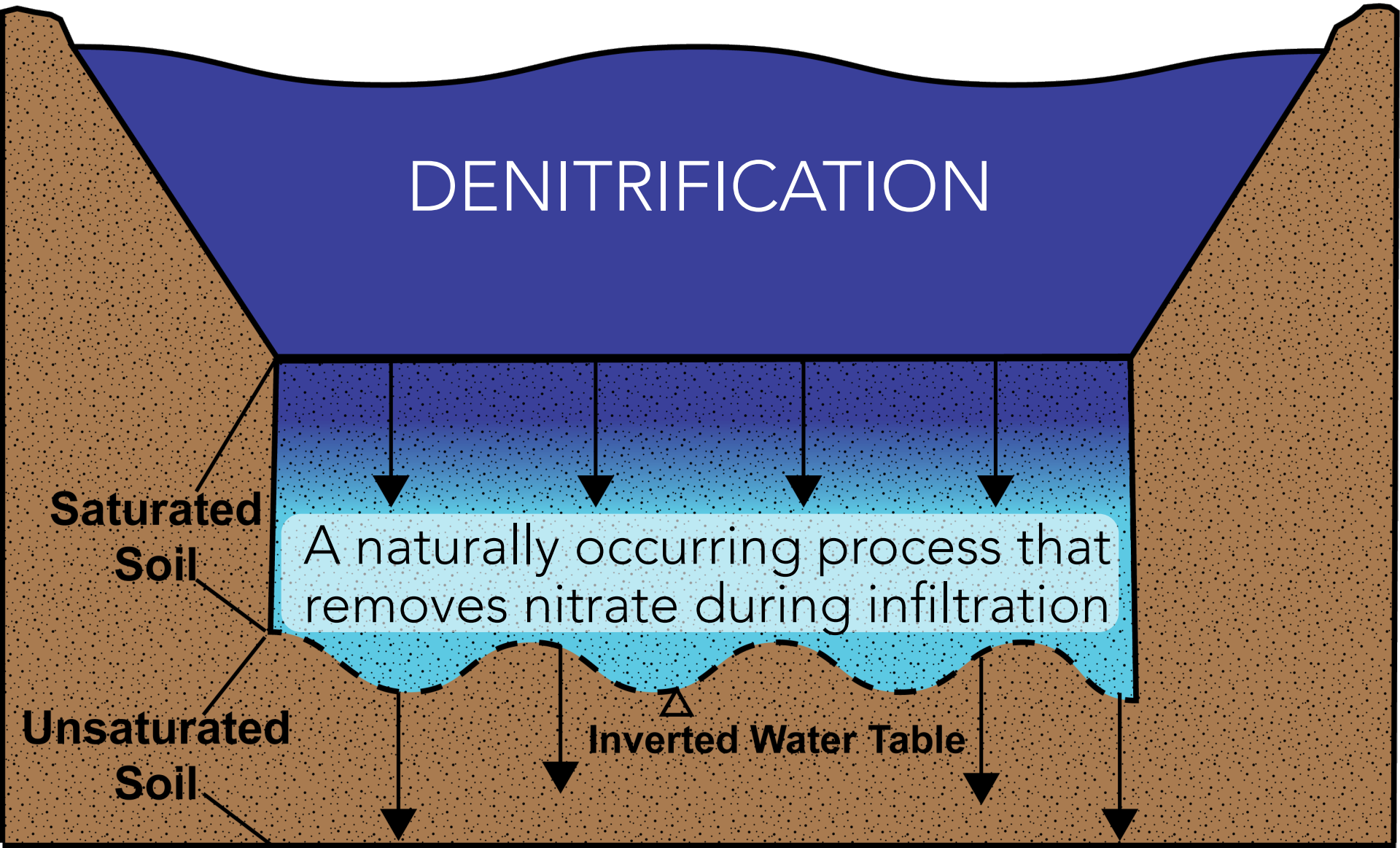
UCSC Hydrogeology


How does artificial recharge affect groundwater quality?
Managed Aquifer Recharge Portal
INOWAS
DEMEAU
igrac

Recharge can remove contaminants



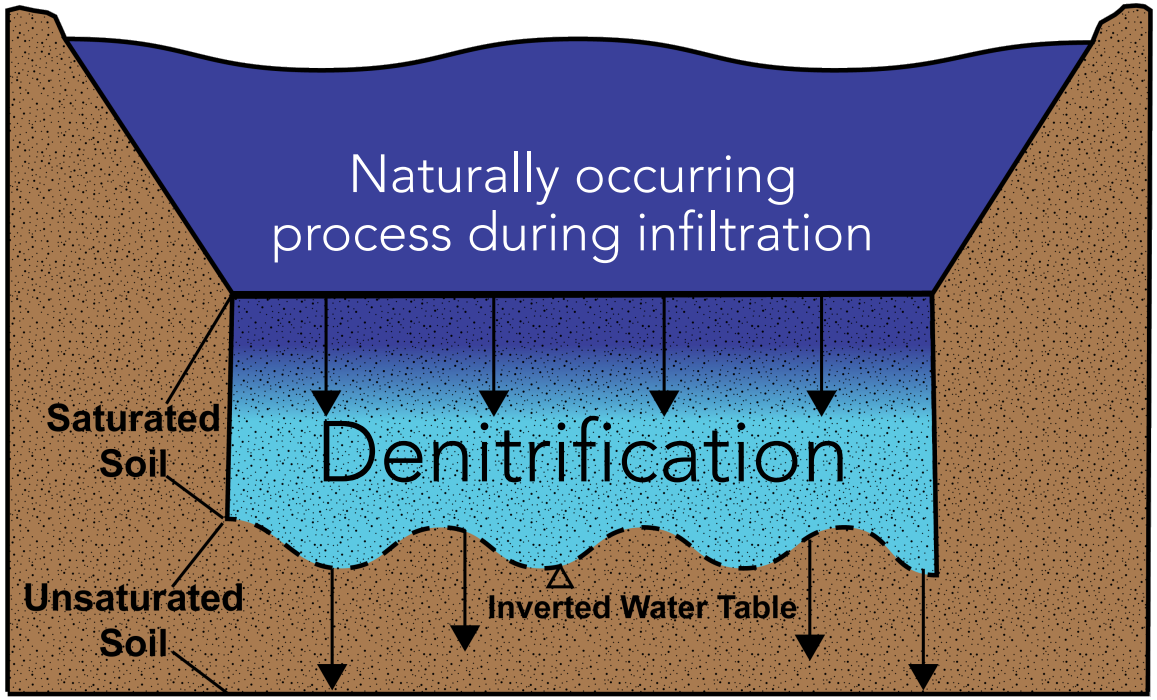
Recharge can remove contaminants

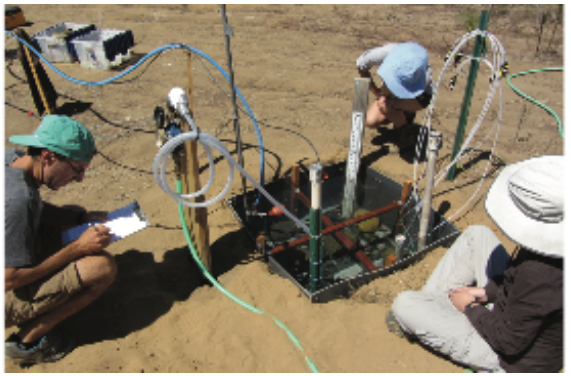


Denitrification during infiltration

Requires an abundance of carbon and a lack of oxygen

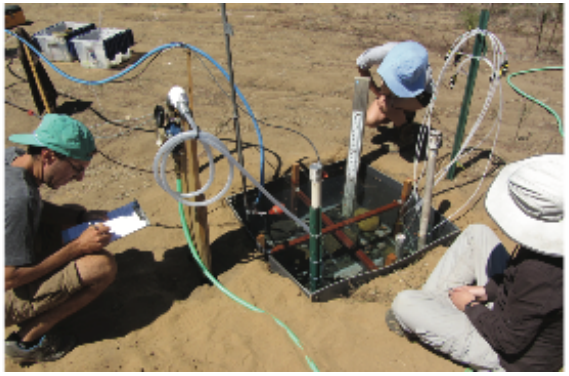
Represents a permanent sink for nitrate

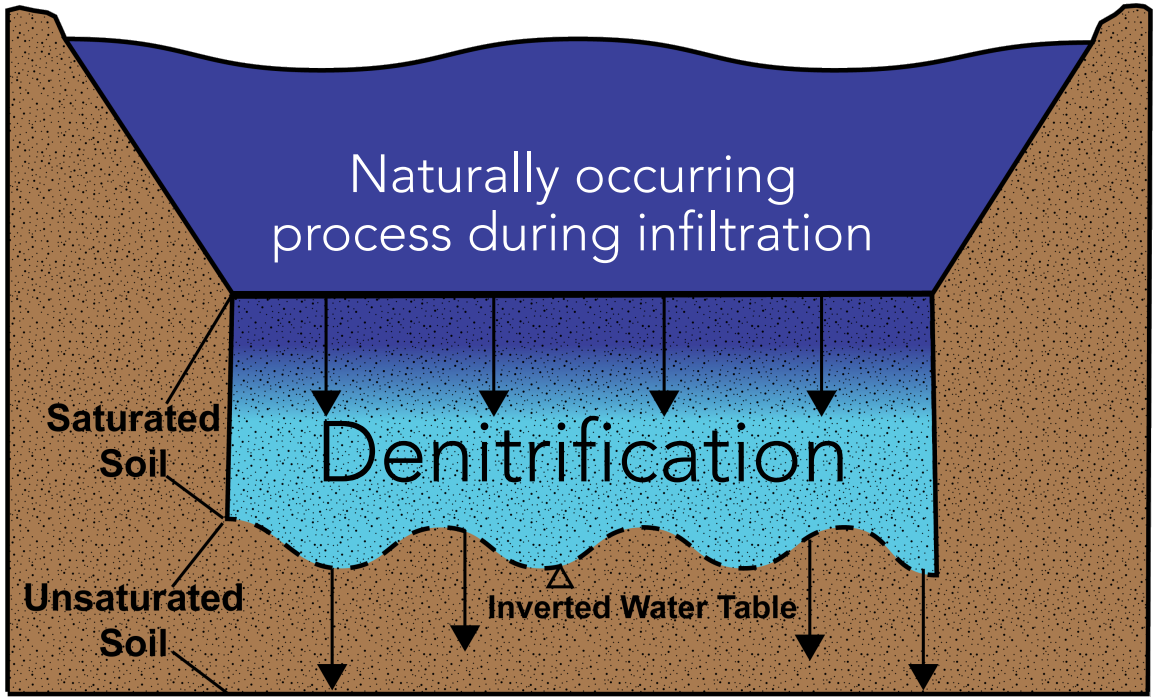




Modeling denitrification during infiltration

Leverage this dataset to model denitrification during infiltration on a landscape scale

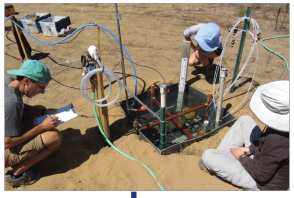




measurements of denitrification during infiltration from **four sites** at **three different scales**

Laboratory

Field



MAR Operations

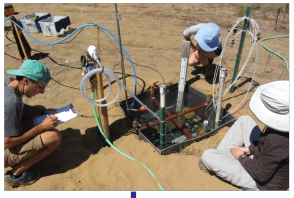
MODEL

Develop, calibrate and validate models of denitrification

measurements of denitrification during infiltration from **four sites** at **three different scales**

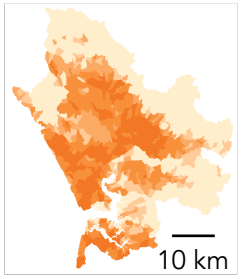
Laboratory

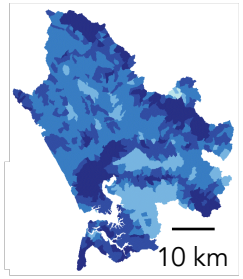
Field

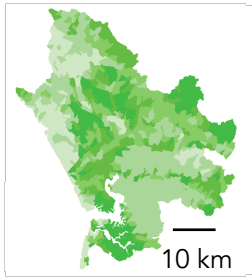


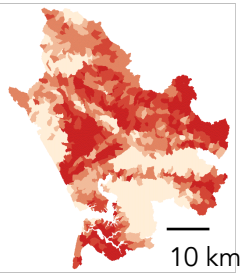
MAR Operations

Soil Properties







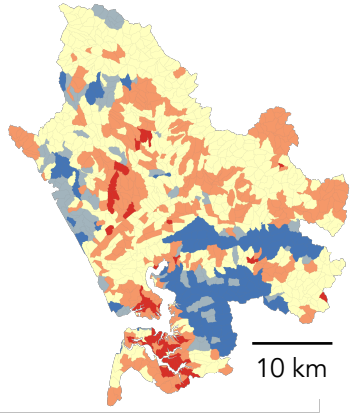


MODEL

Develop, calibrate and validate models of denitrification

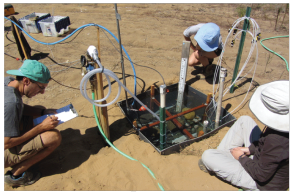
Use modeled relationships, in conjunction with **spatially mapped soil data** to make predictions of denitrification

Amount of nitrate removed

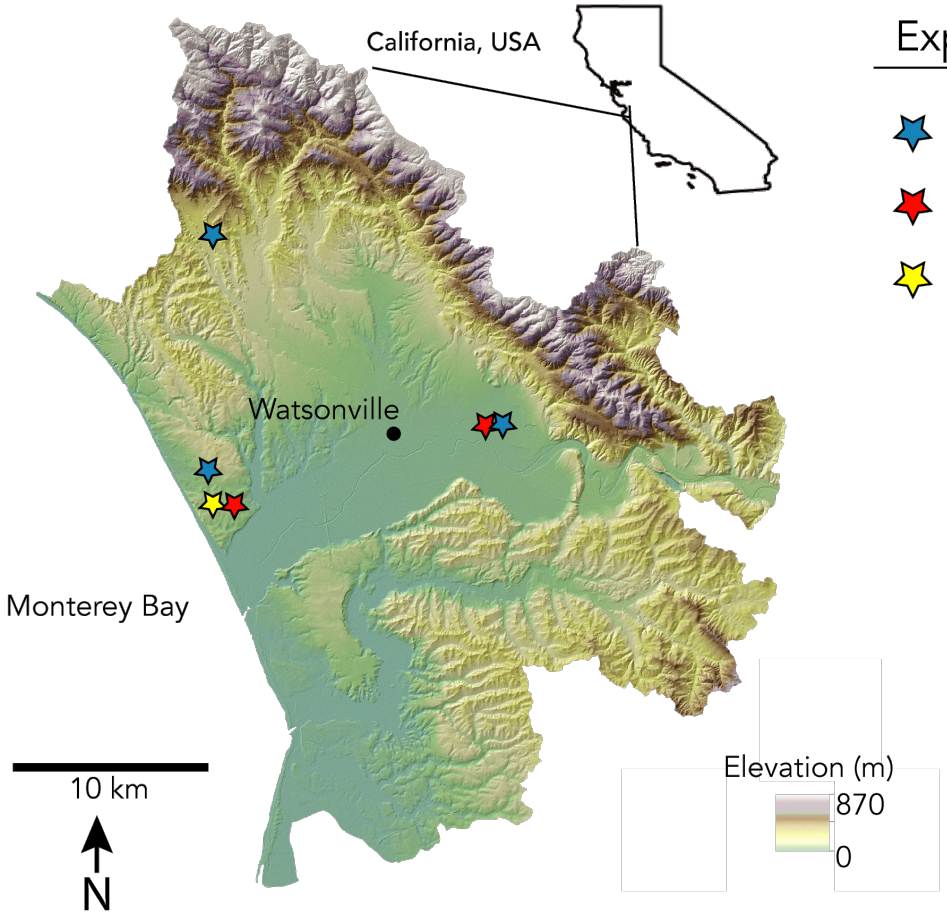


Laboratory

Field



MAR Operations



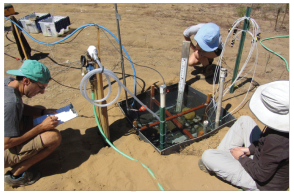
Experimental Scale

- Laboratory Column Studies
- Field Percolation Tests
- MAR Operations

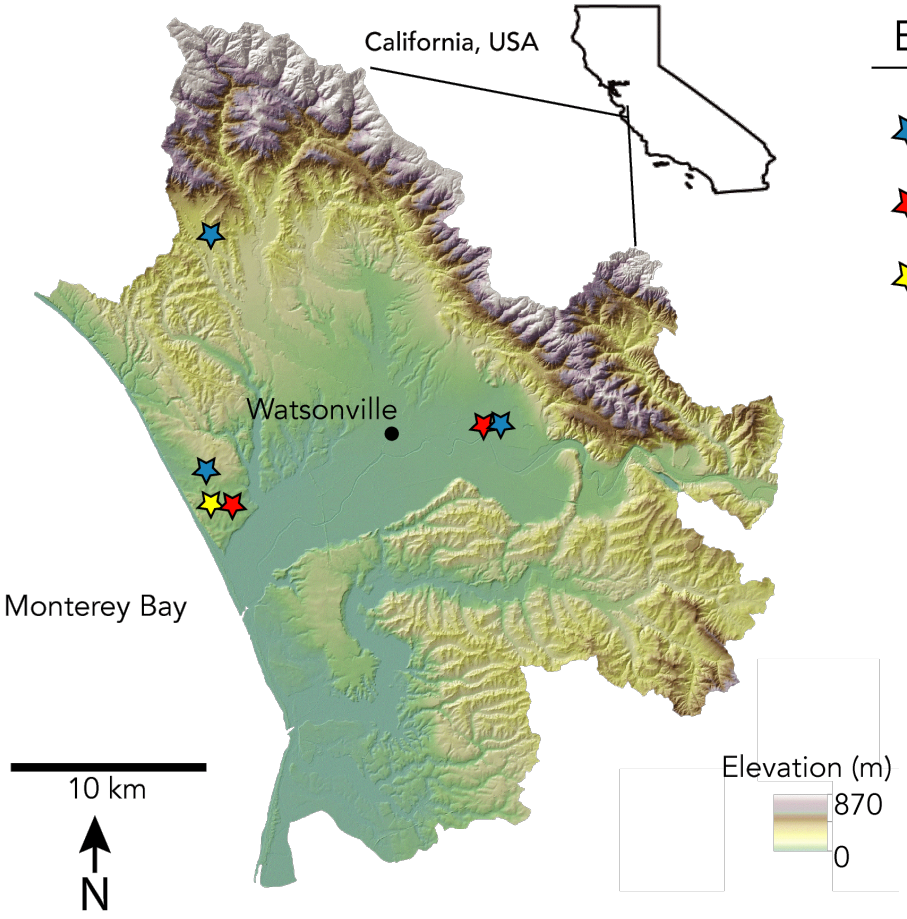
Data split into calibration (75%) and validation (25%) sets to develop models

Laboratory

Field

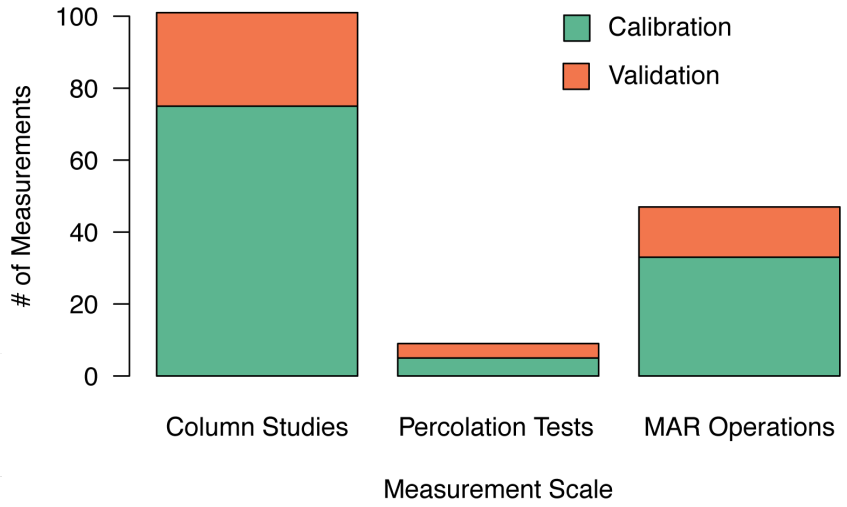


MAR Operations



Experimental Scale

- Laboratory Column Studies
- Field Percolation Tests
- MAR Operations



Model Development

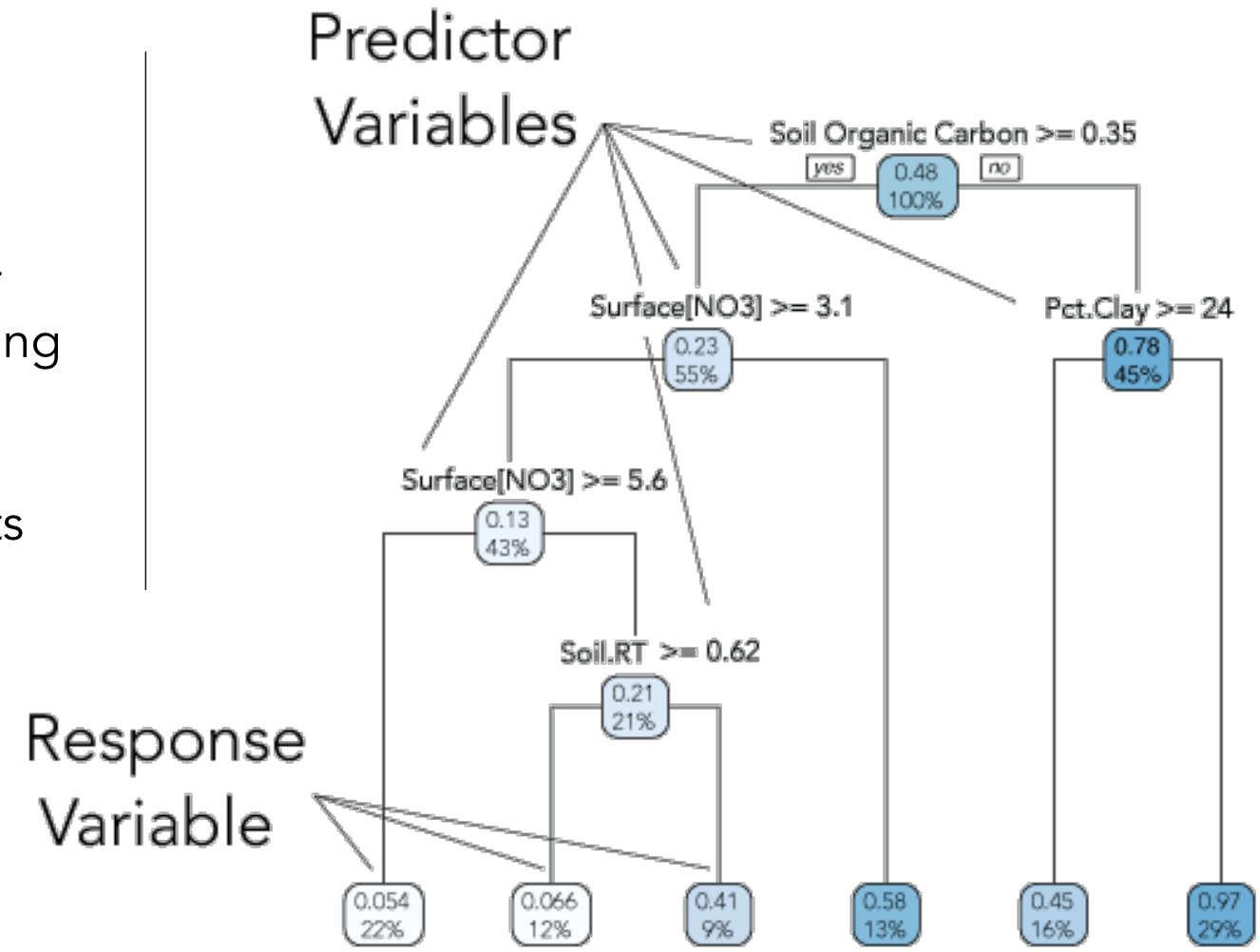
Random Forest

Statistical technique

Robust with non-linear relationships and missing data

Easy to interpret results

Best performer out of several modeling approaches

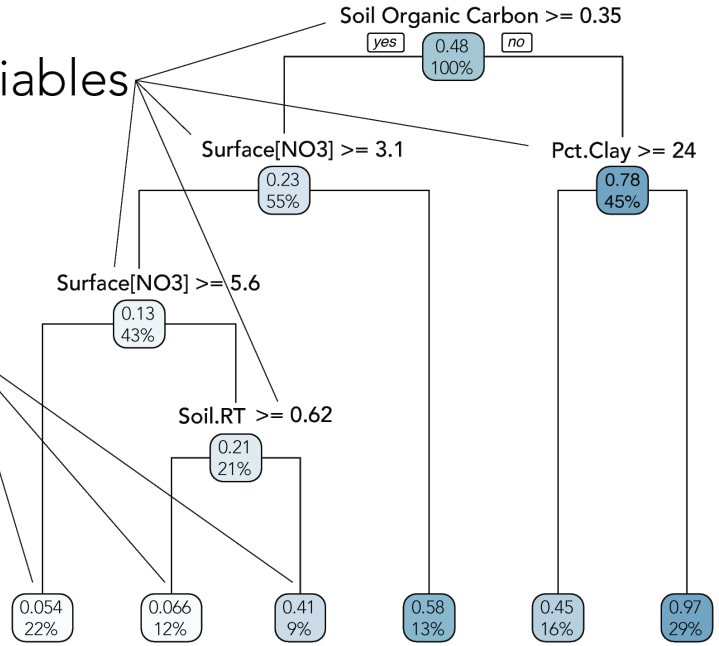


Model Development

Random Forest

Predictor Variables

Response Variable



Response Variable

Amt NO_3 Removed = $[NO_3]_{initial} - [NO_3]_{final}$

Predictors

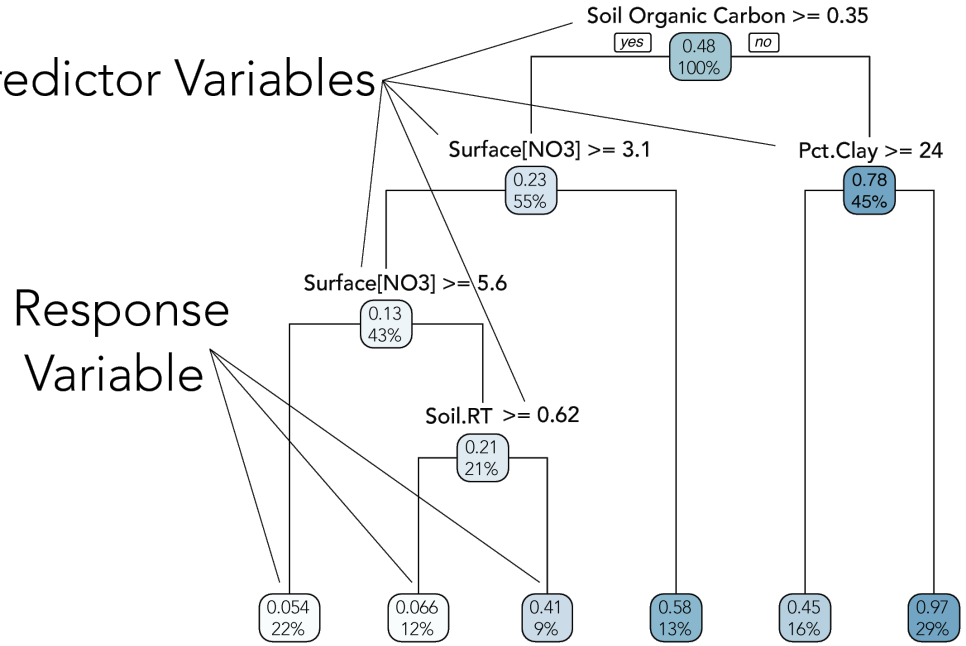
Soil Residence Time
$[NO_3]_{initial}$
$[DOC]_{initial}$
Soil Organic Carbon
Soil Nitrogen
Percent Clay
Percent Sand
Percent Silt
Temperature
Soil pH

Model Development

Random Forest

Predictor Variables

Response Variable



Response Variable

Amt NO_3 Removed = $[NO_3]_{initial} - [NO_3]_{final}$

Predictors

Soil Residence Time

$[NO_3]_{initial}$

$[DOC]_{initial}$

Soil Organic Carbon

Soil Nitrogen

Percent Clay

Percent Sand

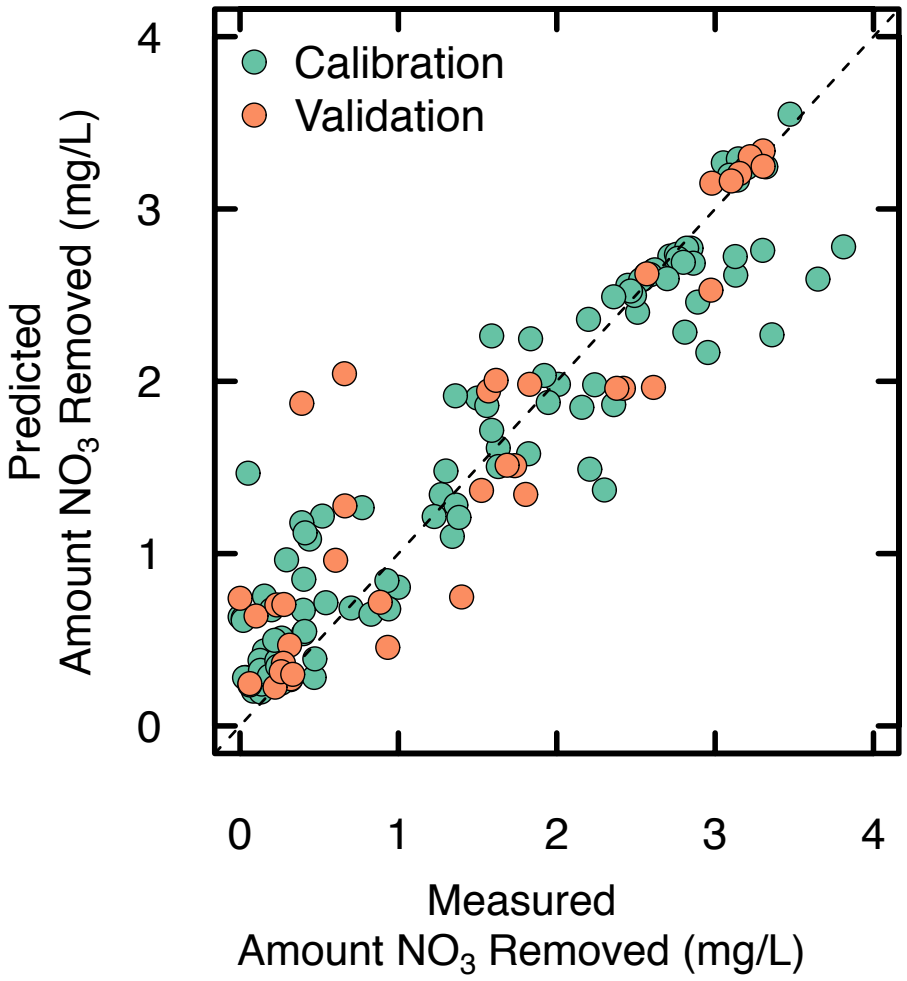
Percent Silt

Temperature

Soil pH

Model Development

Model Performance



Response Variable

Amt NO_3 Removed = $[NO_3]_{initial} - [NO_3]_{final}$

Predictors

Soil Residence Time

$[NO_3]_{initial}$

$[DOC]_{initial}$

Soil Organic Carbon

Soil Nitrogen

Percent Clay

Percent Sand

Percent Silt

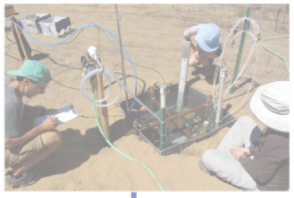
Temperature

Soil pH

measurements of denitrification during infiltration from **four sites** at **three different scales**

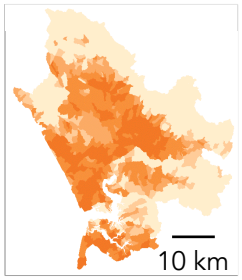
Laboratory

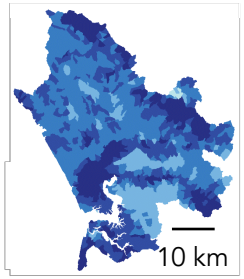
Field

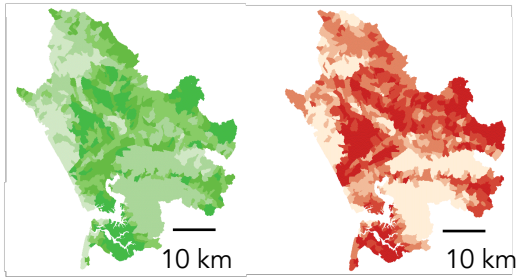


MAR Operations

Soil Properties





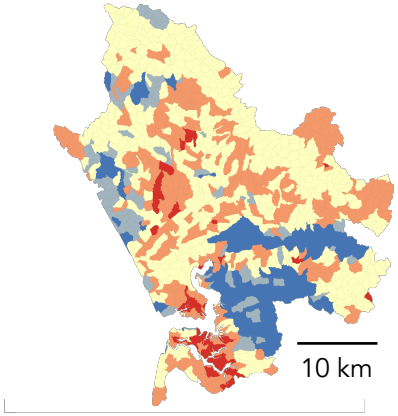


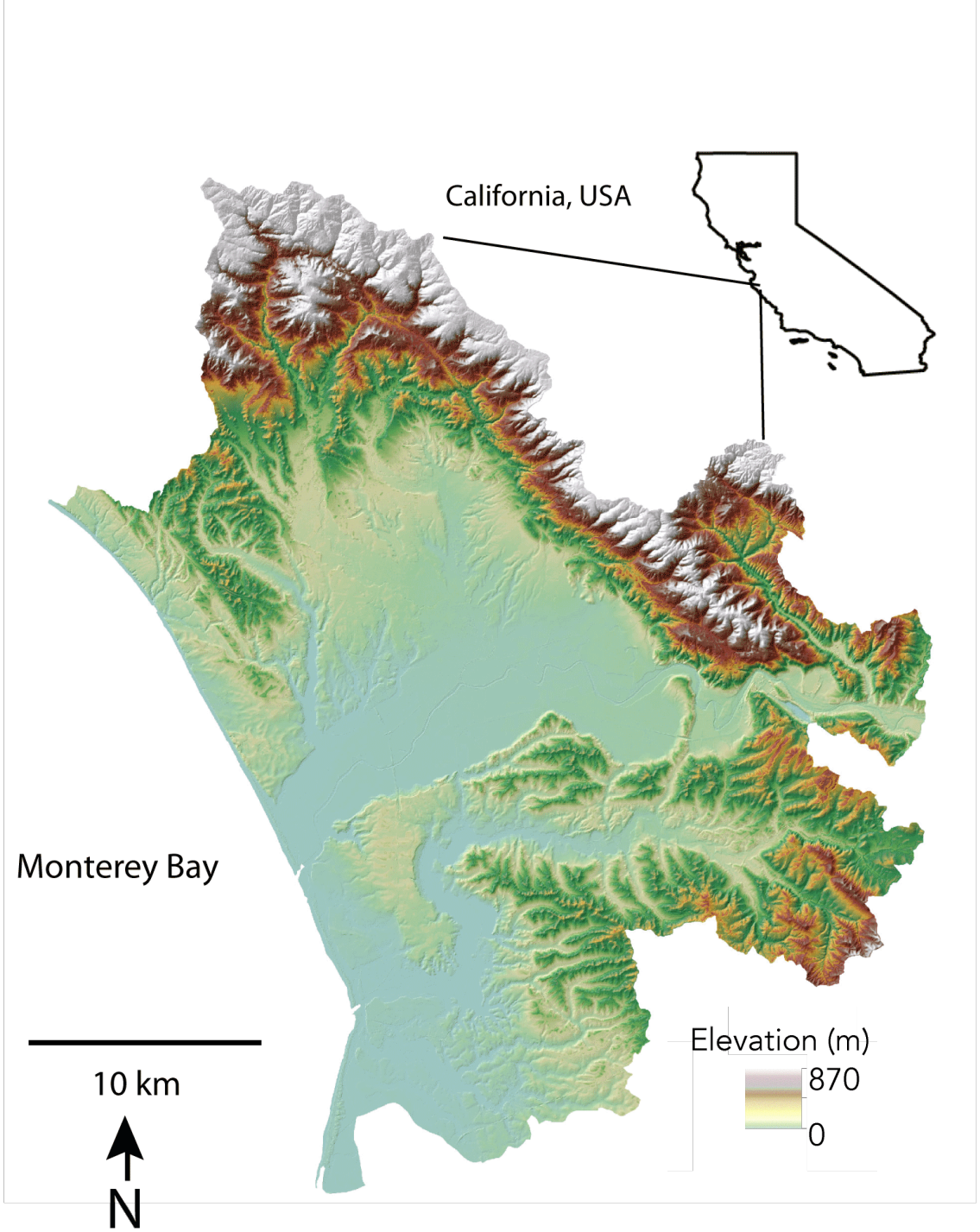
MODEL

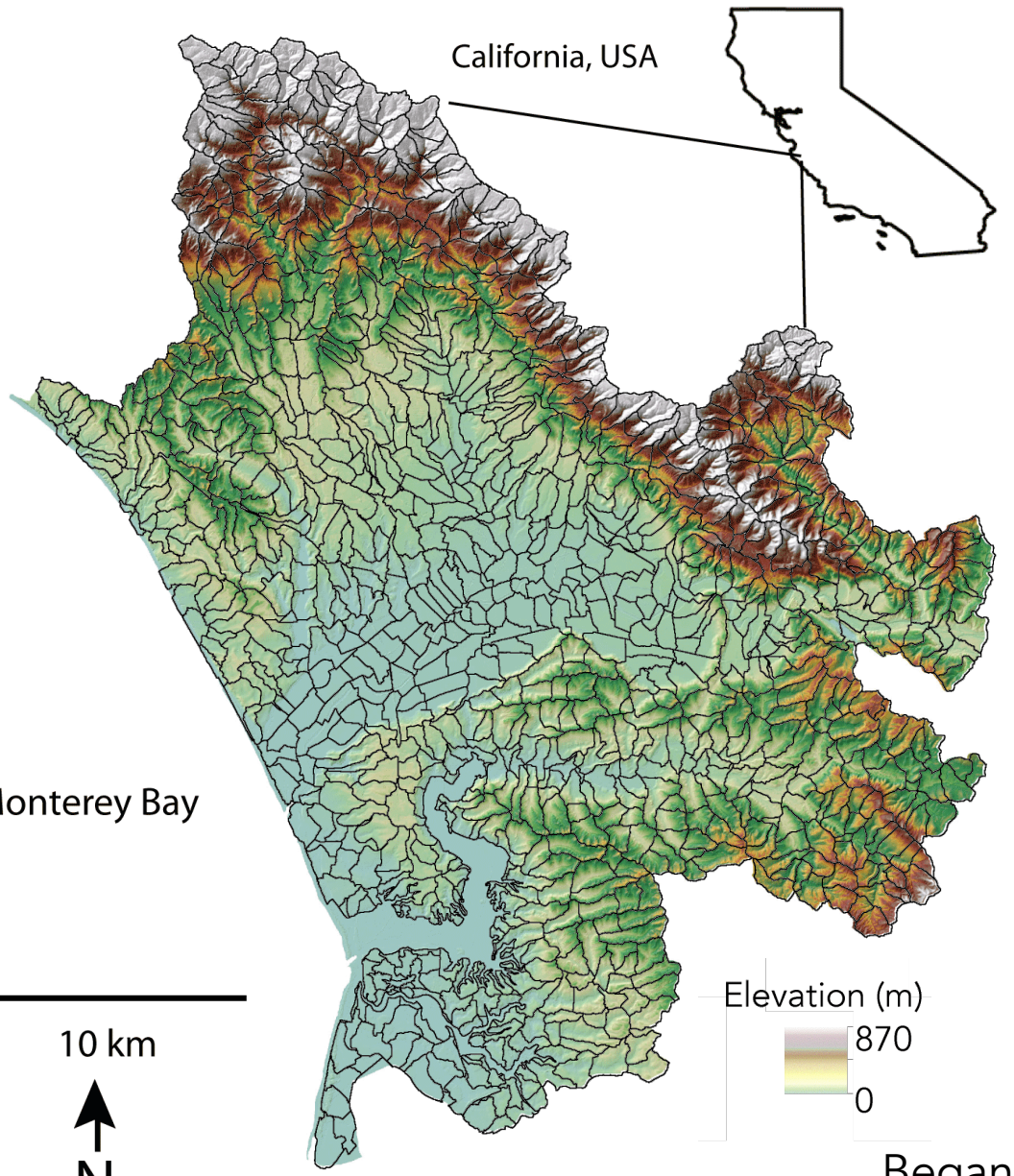
Develop, calibrate and **validate** models of denitrification

Use modeled relationships, in conjunction with **spatially mapped soil data** to make predictions of denitrification

Amount of nitrate removed




California, USA
Monterey Bay
10 km
N
Elevation (m)
870
0



1025 topographically delineated hydrological response units (HRUs) 0.1-1.0 km^2

Soil properties are averaged with these HRUs

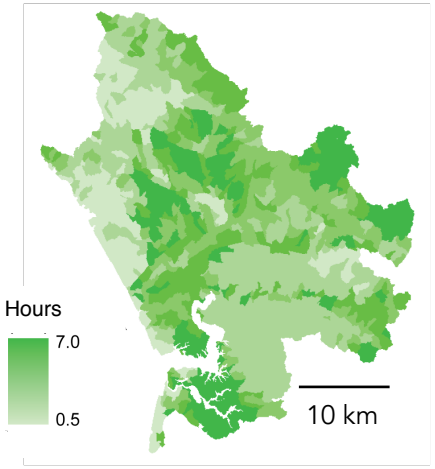
Model is run at the scale of the HRUs

Assuming storm water collection and infiltration using basins

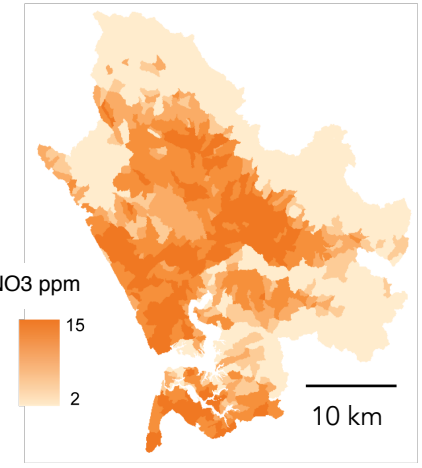
Beganskas et al. 2019

PREDICTORS

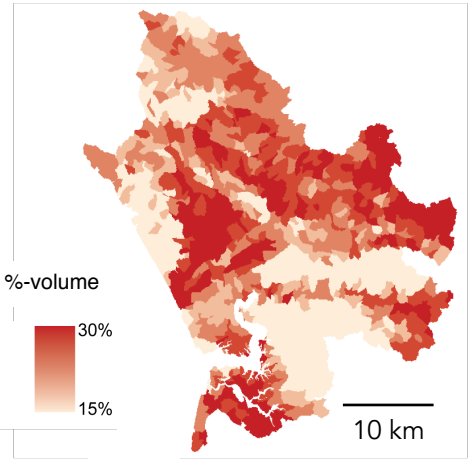
Soil Residence Time



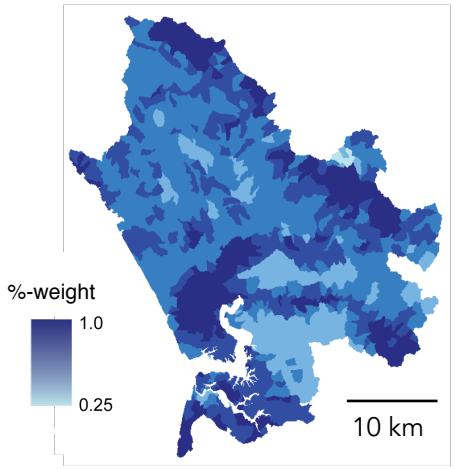
Initial Nitrate Concentration



Clay Content

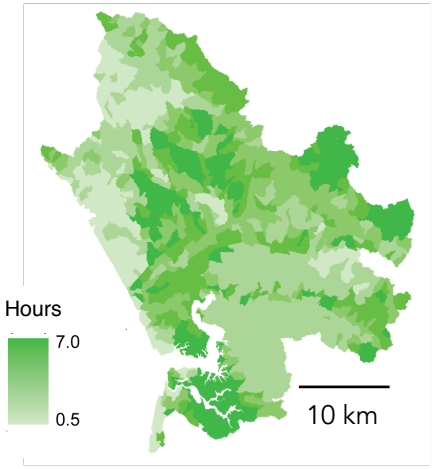


Soil Organic Carbon

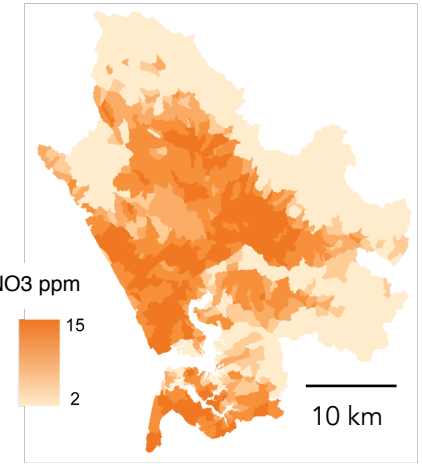


PREDICTORS

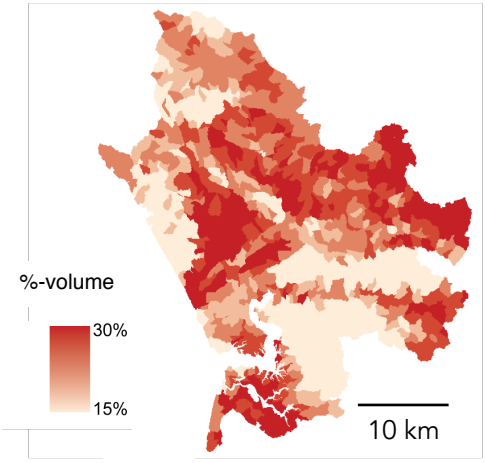
Soil Residence Time



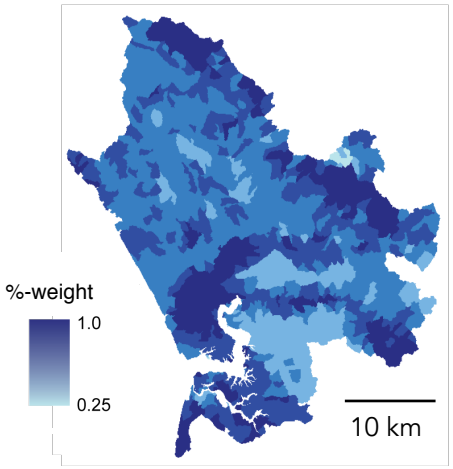
Initial Nitrate Concentration



Clay Content

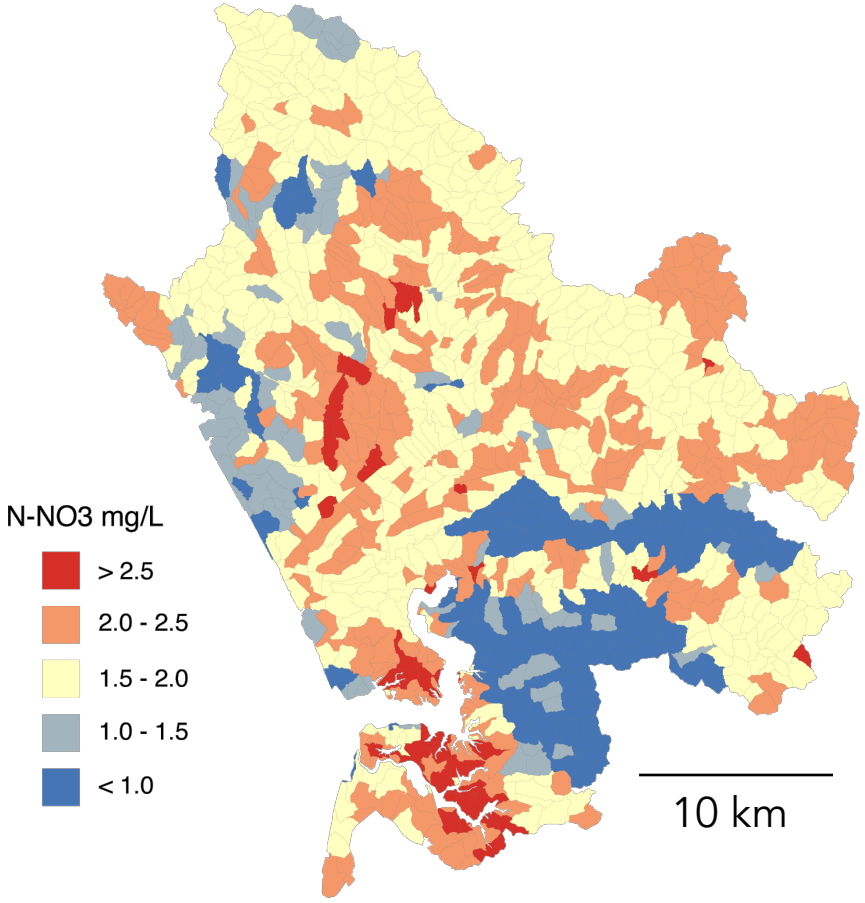


Soil Organic Carbon

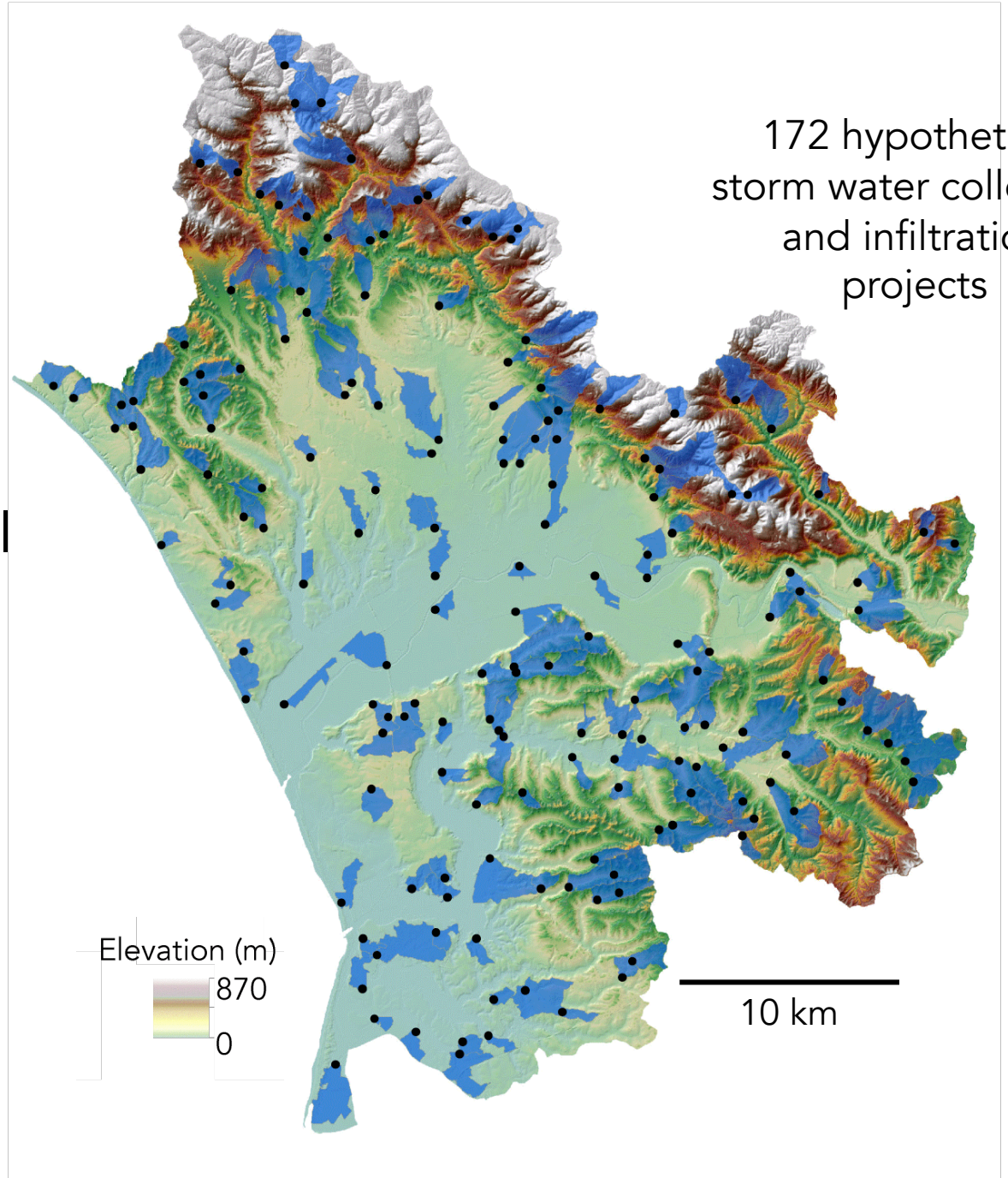


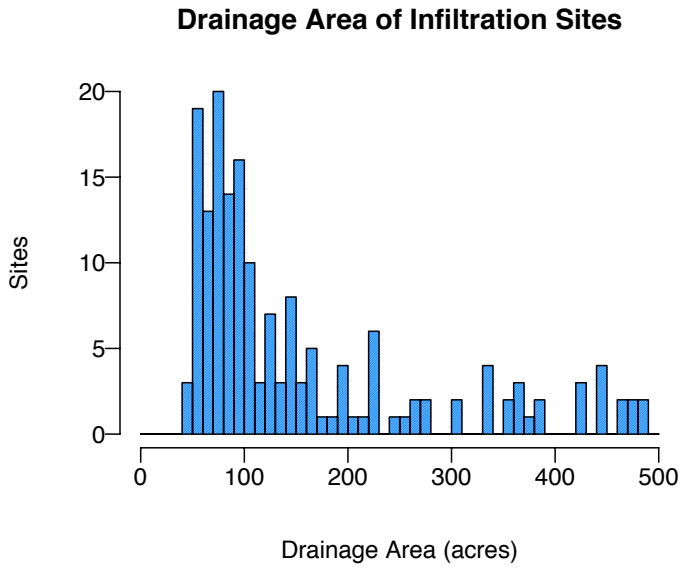
RESPONSE

Amount of nitrate removed

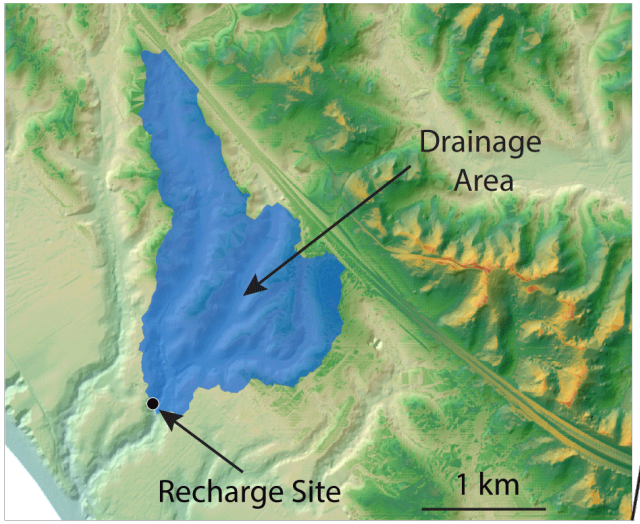


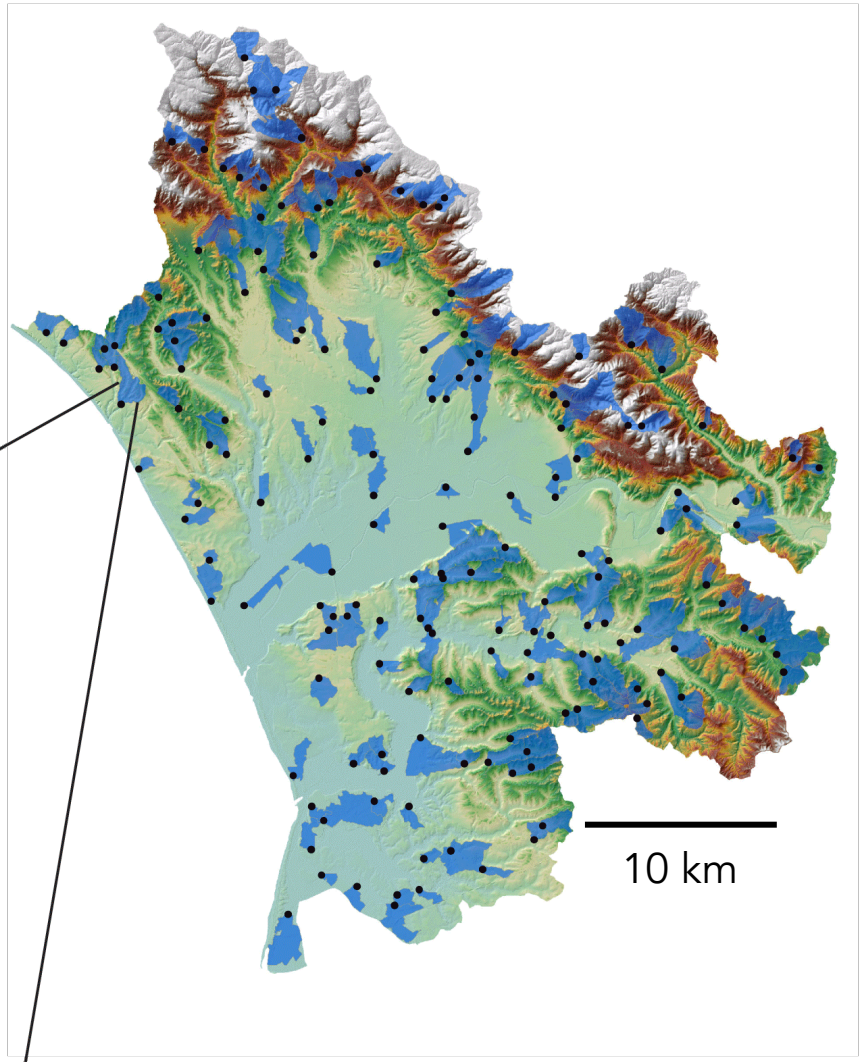
What effect will these projects have on groundwater quality?





Drainage areas range from 50-500 acres

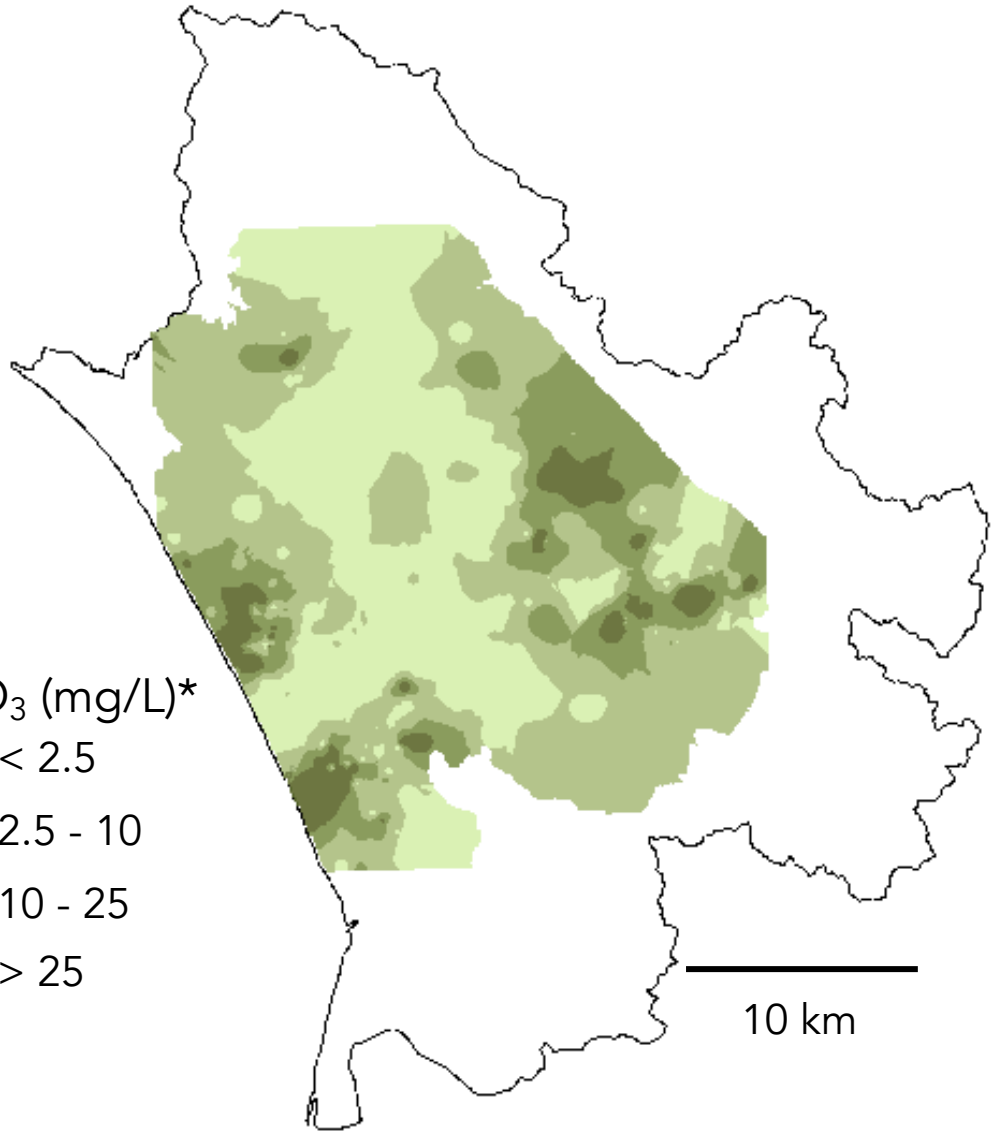




Located in areas of moderate to high suitability

Average Groundwater Nitrate

Groundwater has elevated nitrate concentration in some areas



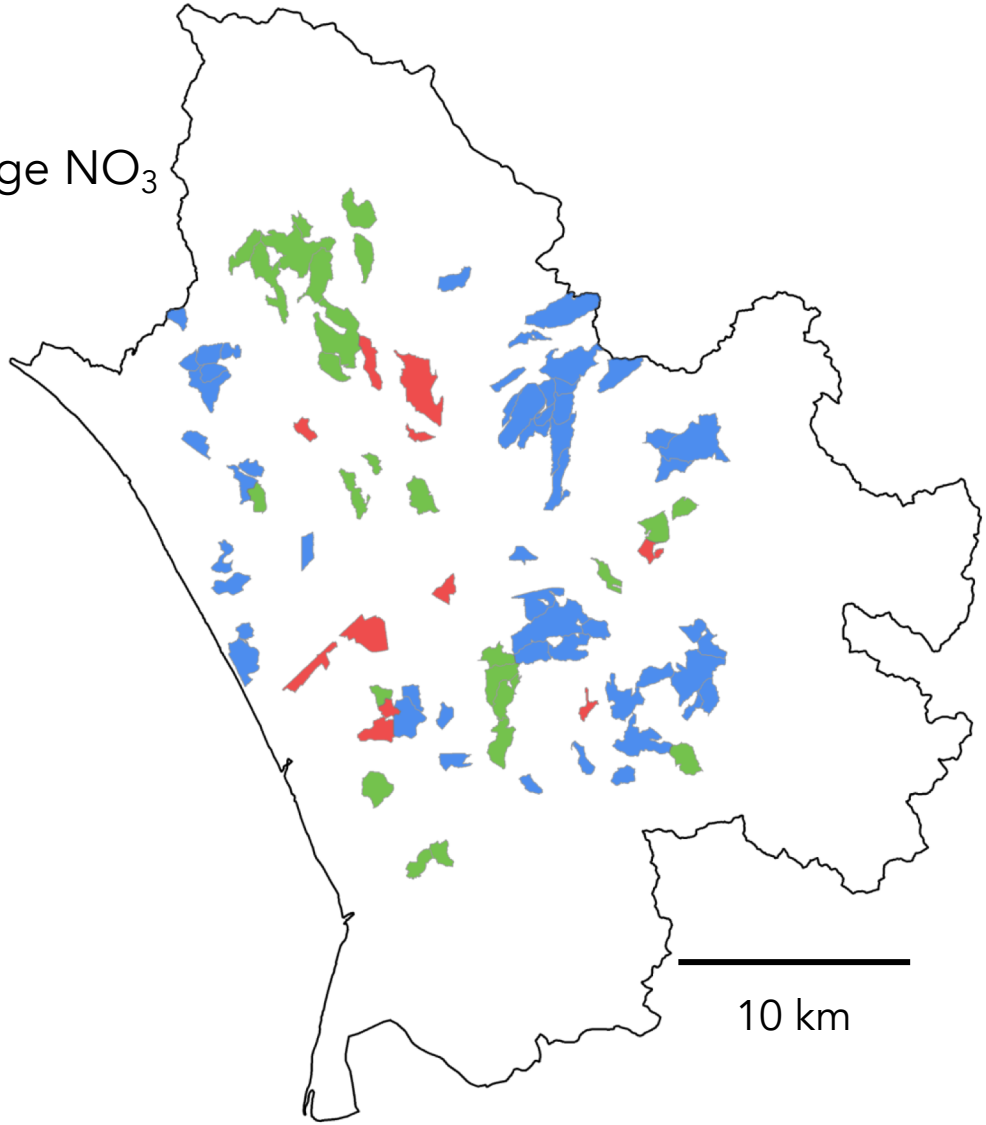
* EPA maximum contaminant level is 10 mg-N/L

How does artificial recharge affect groundwater quality?

Initial NO_3 – Amount NO_3 Removed = Recharge NO_3

Compared to Groundwater NO_3

- Better (> 5 mg/L)
- Marginal (± 5 mg/L)
- Worse (> 5 mg/L)

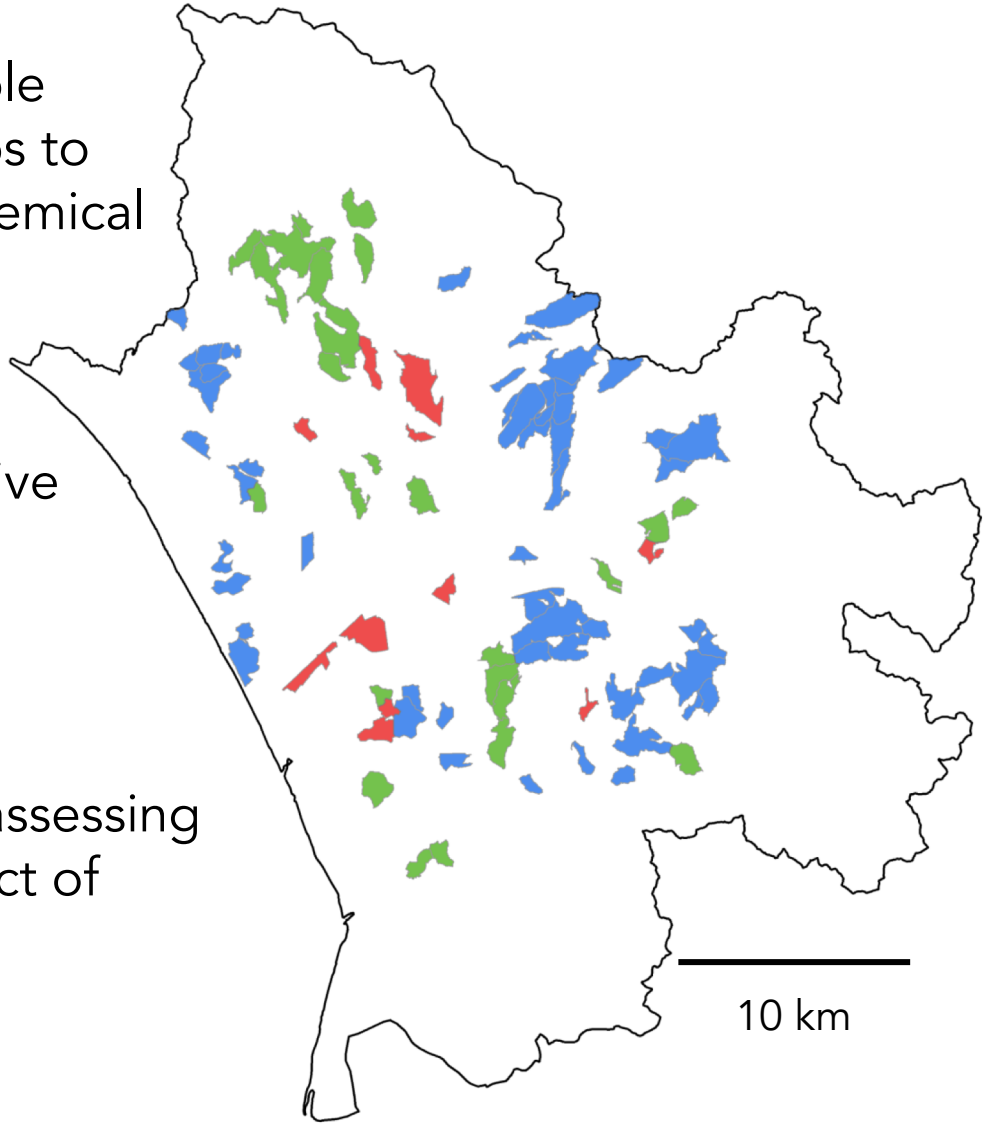


Incorporating water quality in MAR suitability

Incorporated data from multiple scales and experimental setups to develop a model of biogeochemical processing during infiltration

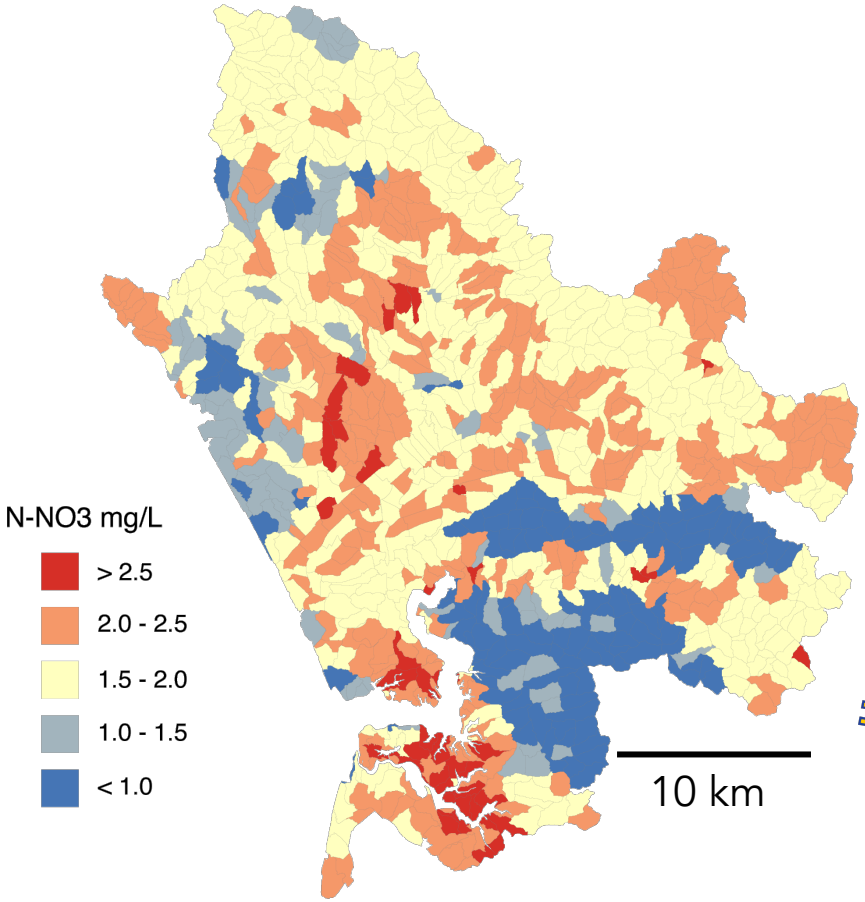
Used the model to make quantitative predictions about the impact of recharge on groundwater quality

Developed a framework for assessing the spatial variability of impact of recharge on water quality



Acknowledgements

Sarah Farola, Paige Borges, Araceli Serrano, Dan Sampson, Stephen and Pamela Storrs, Kelli Camara, Bill Rice, Sarah Beganskas, Mark Burnett, Ryan Nyberg, Walker Weir, Tyler Stewart, Dom van den Dries, Rob Franks



UCWATER
Security and Sustainability Research Initiative

NSF

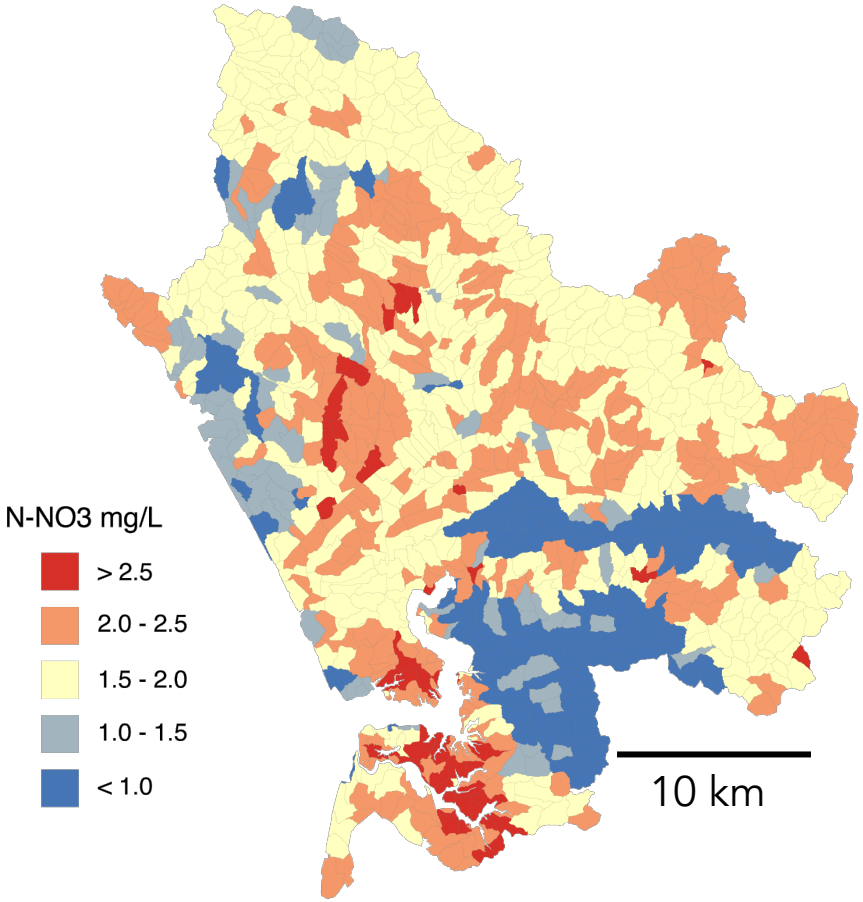
The Recharge Initiative
Replenish • Recover • Restore

UCSC Hydrogeology
THE UNIVERSITY OF CALIFORNIA 1868
"Endeavor to persevere..."

GORDON AND BETTY MOORE FOUNDATION

Acknowledgements

Sarah Farola, Paige Borges, Araceli Serrano, Dan Sampson, Stephen and Pamela Storrs, Kelli Camara, Bill Rice, Sarah Beganskas, Mark Burnett, Ryan Nyberg, Walker Weir, Tyler Stewart, Dom van den Dries, Rob Franks



UC WATER
Security and Sustainability Research Initiative

NSF

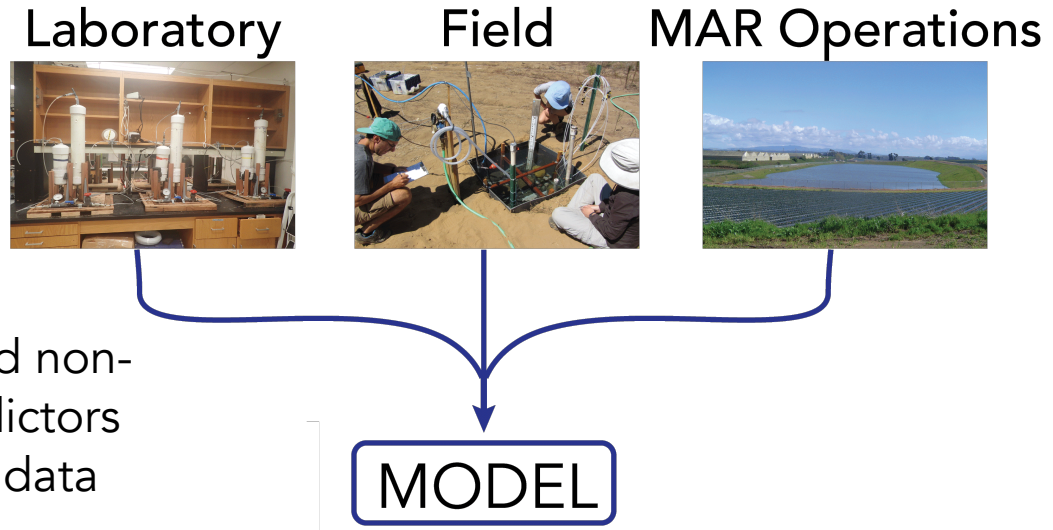
The Recharge Initiative
Replenish • Recover • Restore

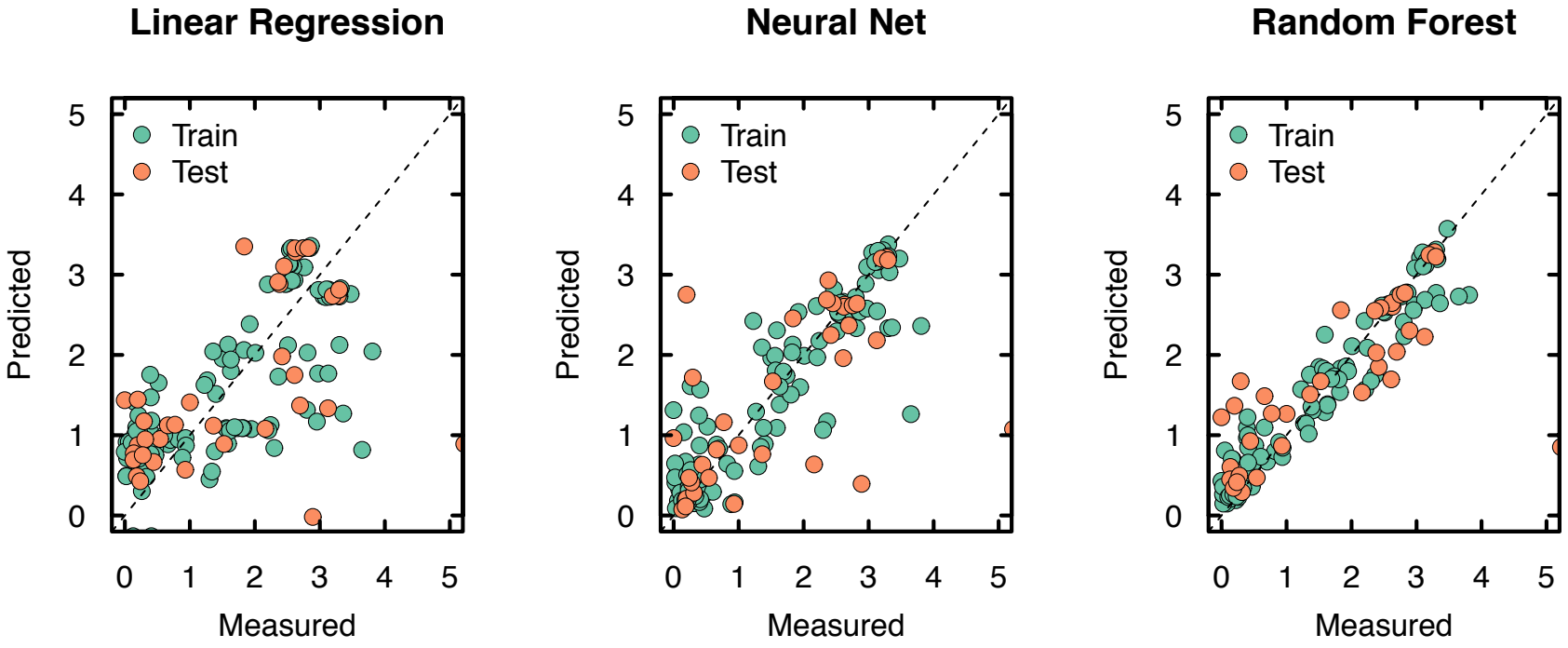
UCSC Hydrogeology
THE UNIVERSITY OF CALIFORNIA
1868
"Endeavor to persevere..."

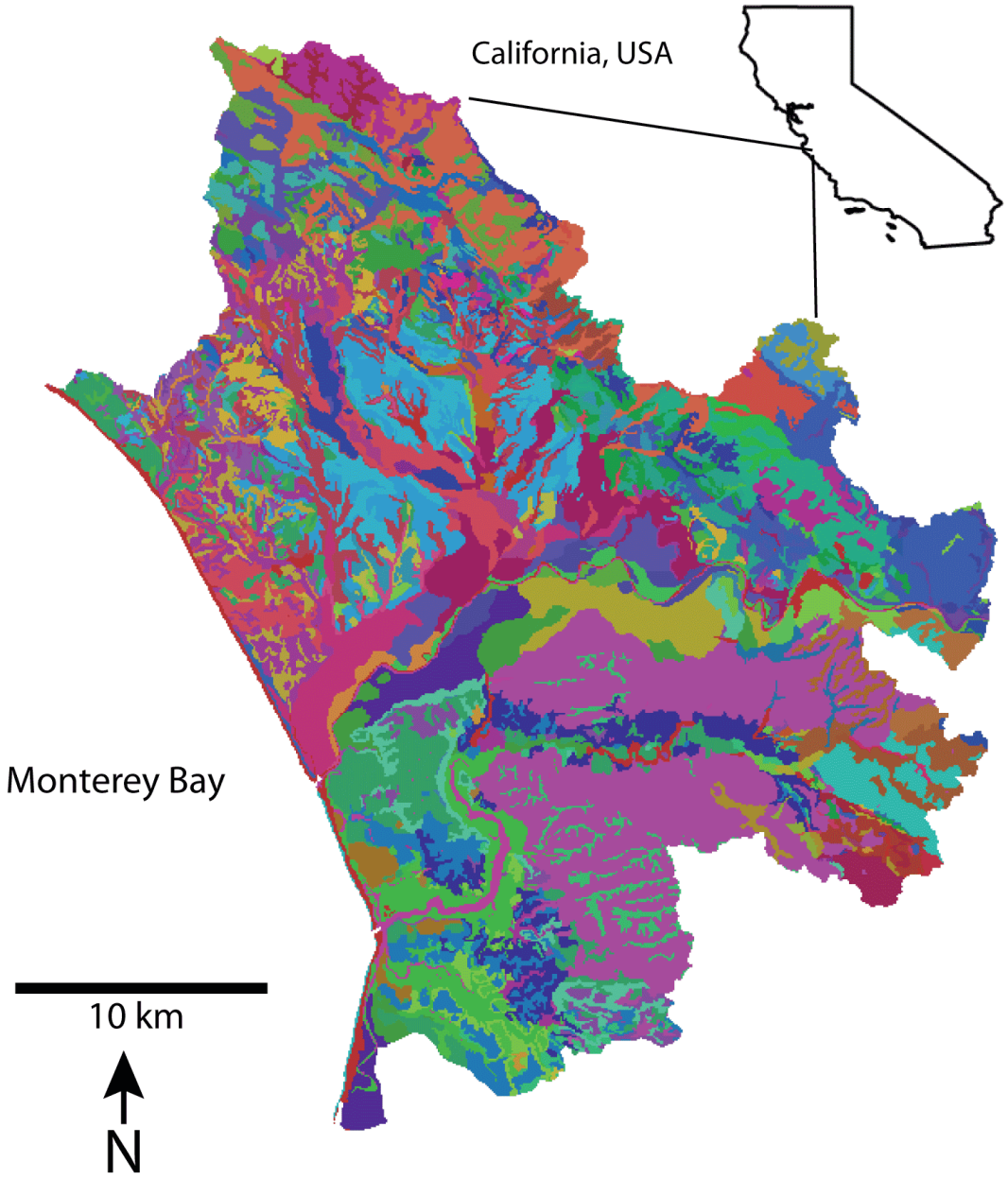
GORDON AND BETTY MOORE FOUNDATION

The neural net and random forest appear to reproduce trends in experimental data better than a linear regression

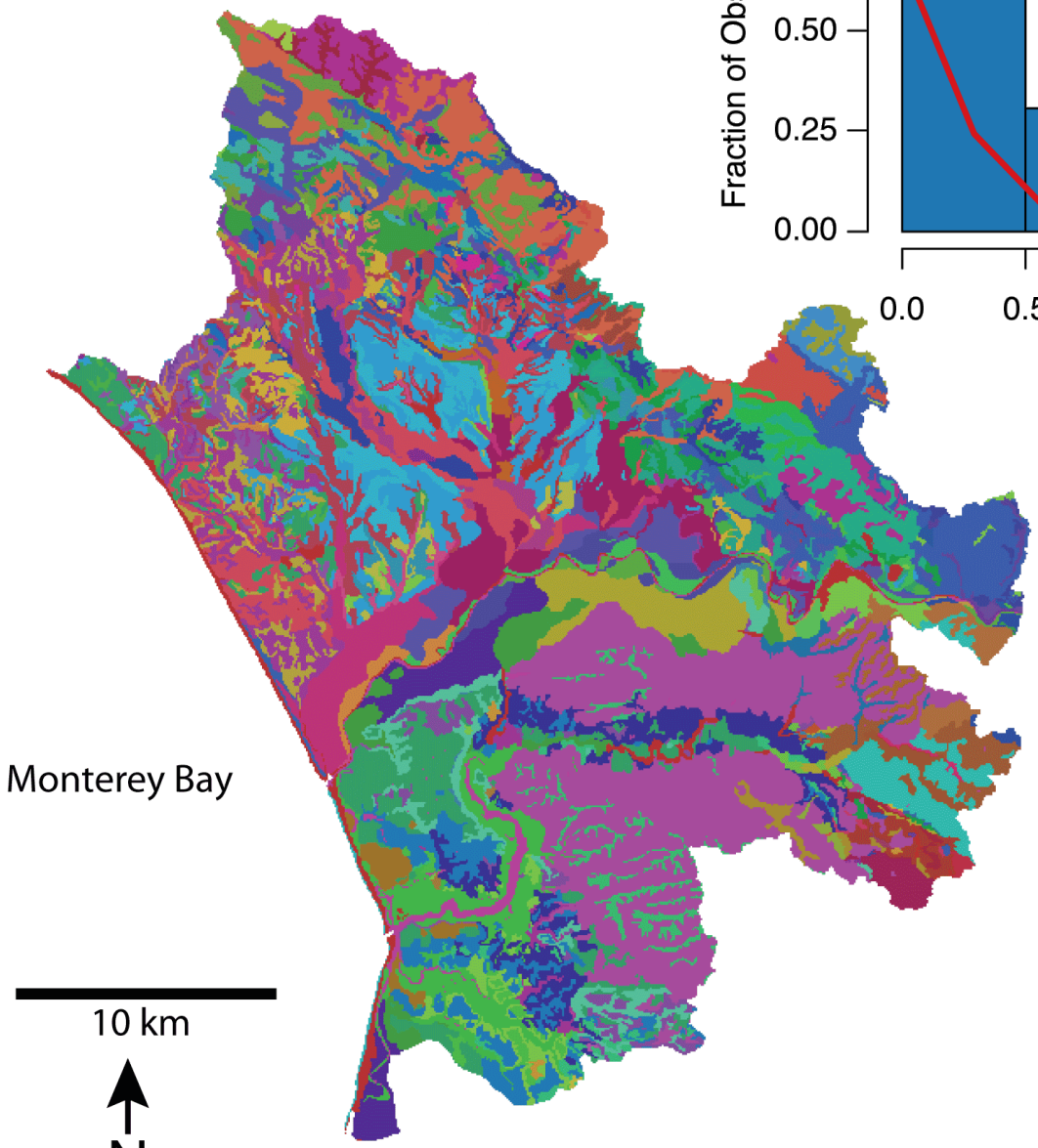
Likely due to interaction terms and non-linear relationships between predictors and response as well as missing data

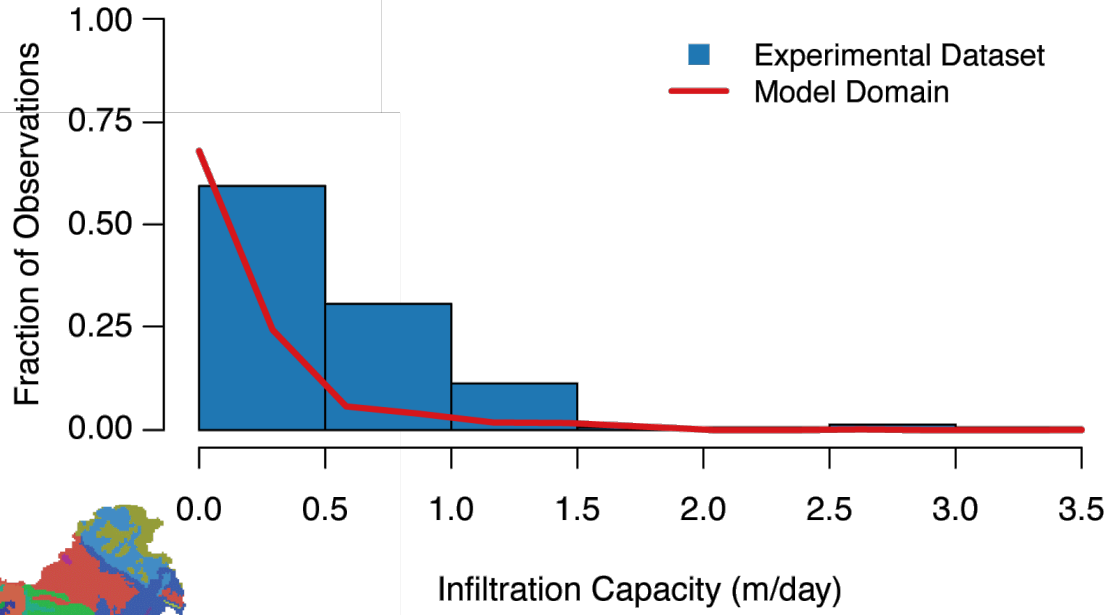






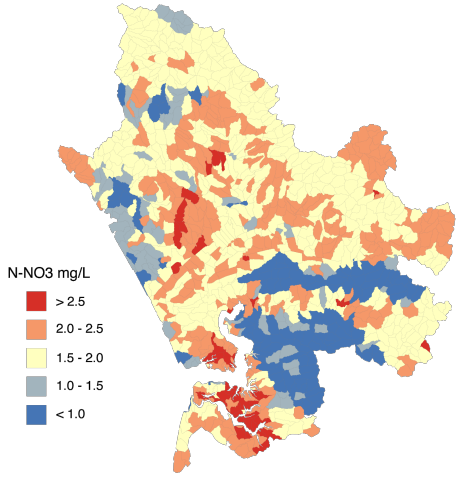
The Pajaro Valley is a complex area with over 100 individual soil units



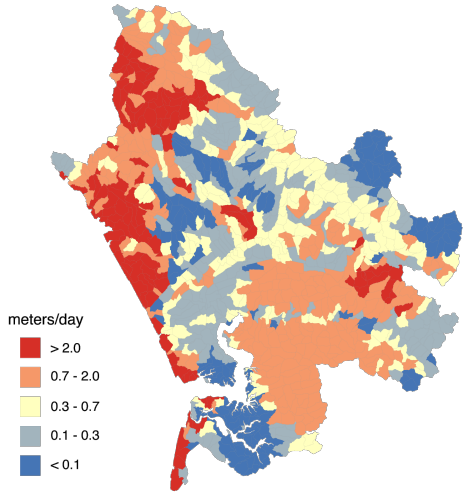


Our experimental data cover much of the range of soil properties in the modeling domain

Potential Nitrate Removal



Soil Infiltration Capacity



Potential Nitrate Load Reduction