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
How does artificial recharge affect groundwater quality?
Recharge can remove contaminants

Non-pristine water can be improved during infiltration
Recharge can remove contaminants

A naturally occurring process that removes nitrate during infiltration
Denitrification during infiltration

Requires an abundance of carbon and a lack of oxygen

Represents a permanent sink for nitrate
Leverage this dataset to model denitrification during infiltration on a landscape scale.
measurements of denitrification during infiltration from **four sites** at three different scales

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

**Laboratory**

**Field**

**MAR Operations**

**MODEL**

Use modeled relationships, in conjunction with spatially mapped soil data to make predictions of denitrification.
measurements of denitrification during infiltration from four sites at three different scales

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

Response Variable

Amount NO₃ Removed = [NO₃]_{initial} - [NO₃]_{final}

Predictors

Soil Residence Time

[NO₃]_{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

Predictors

Soil Residence Time

[NO$_3$]$_{\text{initial}}$ – [NO$_3$]$_{\text{final}}$

Soil Organic Carbon

Soil Nitrogen

Percent Clay

Percent Sand

Percent Silt

Temperature

Soil pH

Amt NO$_3$ Removed = [NO$_3$]$_{\text{initial}}$ – [NO$_3$]$_{\text{final}}$
Model Development

**Model Performance**

- Amt NO$_3$ Removed = $[\text{NO}_3]_{\text{initial}} - [\text{NO}_3]_{\text{final}}$

**Response Variable**

- Predictors:
  - Soil Residence Time
  - $[\text{NO}_3]_{\text{initial}}$
  - $[\text{DOC}]_{\text{initial}}$
  - Soil Organic Carbon
  - Soil Nitrogen
  - Percent Clay
  - Percent Sand
  - Percent Silt
  - Temperature
  - Soil pH

**Predictors**

- Measured Amount NO$_3$ Removed (mg/L)

- Predicted Amount NO$_3$ Removed (mg/L)
measurements of denitrification during infiltration from four sites at three different scales

Soil Properties

Laboratory  Field  MAR Operations

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
1025 topographically delineated hydrological response units (HRUs) 0.1-1.0 km²

Soil properties are averaged with these HRUs

Model is run at the scale of the HRUs

Assuming storm water collection and infiltration using basins
PREDICTORS

Soil Residence Time

Initial Nitrate Concentration

Clay Content

Soil Organic Carbon
What effect will these projects have on groundwater quality?
Drainage areas range from 50-500 acres.

Located in areas of moderate to high suitability.
Groundwater has elevated nitrate concentration in some areas.

**Average Groundwater Nitrate**

<table>
<thead>
<tr>
<th>N-NO₃ (mg/L)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 2.5</td>
</tr>
<tr>
<td>2.5 - 10</td>
</tr>
<tr>
<td>10 - 25</td>
</tr>
<tr>
<td>&gt; 25</td>
</tr>
</tbody>
</table>

* 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

Thank you!
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

Thank you!
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

N-NO₃ mg/L
- > 2.5
- 2.0 - 2.5
- 1.5 - 2.0
- 1.0 - 1.5
- < 1.0

g-N/day\cdot m^{-2}
- > 7.0
- 5.0 - 7.0
- 3.0 - 5.0
- 2.0 - 3.0
- 1.0 - 2.0
- 0.5 - 1.0
- < 0.5