Predicting ungauged stream flow and nutrient loading through hierarchical modeling of the Falls Lake watershed (1982-2015)

WRRI Conference
March 18, 2016
Justin Davenport
Daniel Obenour Ph.D

OVERVIEW OF FALLS LAKE RESERVOIR

• Falls Lake impoundment was completed in December of 1983
• Total lake area at normal pool elevation is 50 km²
• Problems with eutrophication in the reservoir

WRAL January 5, 2010
OVERVIEW OF FALLS LAKE RESERVOIR

- Added to NC Impaired 303d list in 2008
- State legislature passed Falls Lake Rules in 2011
- Algae in the lake affects drinking water treatment

Indy Week May 4, 2011

OVERVIEW OF FALLS LAKE BASIN

NC STATE UNIVERSITY
WHY A MULTI-LEVEL MODEL?

• Based upon linear regressions
• Accounts and allows variation at two levels
  • Observation level (flow, time)
  • Watershed level (land-use, reservoir presence)
• Allows model and prediction uncertainty to be characterized
• Shown to be better than modelling all the observations together.

WHY A MULTI-LEVEL MODEL?
AN EXAMPLE WITH A SIMULATED DATASET
WHY A MULTI-LEVEL MODEL?
AN EXAMPLE WITH A SIMULATED DATASET

[Graphs showing varying intercepts and slopes for different groups.]

NC STATE UNIVERSITY
WHY A MULTI-LEVEL MODEL? 
AN EXAMPLE WITH A SIMULATED DATASET

1. Description and Source of Data 
   a. Overview 
   b. Land Cover Data 
   c. Flow Data 
   d. Concentration Data 

2. Methods 

3. Flow Model 
   a. Parameter Specification 
   b. Calibration 
   c. Validation 

4. Nutrient Model 
   a. Parameter Specification 
   b. Calibration 

5. Conclusions
DATA

• Flow data came from USGS Gage Stations

• Nutrient data came from the EPA STORET database and USGS Gage Stations

• Land cover data came from the 1992, 2001, 2006, and 2011 datasets in the National Land Cover Database.

• Exchange data came from NC Division of Water Resources
• Used WRTDS for daily river concentration values (Hirsch et al. 2010)
MODEL CREATION

- Removal of correlated variables prior to calibration
- Backwards variable selection at 95% confidence levels
- Box-Cox transformation was used for response variables and some predictor variables

VALIDATION

- Validation was performed by removing one river at a time, calibrating the model, and then using the removed river for validation.
- Evaluated the model validation by examining predicted vs. observed plots, Nash-Sutcliffe Efficiency, and the bias in predictions.
FLOW MODEL PREDICTORS

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Min. Value</th>
<th>Avg. Value</th>
<th>Max. Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eno Flow ($Q_E$)</td>
<td>0.007</td>
<td>3.39</td>
<td>54.53</td>
<td>(cms)</td>
</tr>
<tr>
<td>Area Ratio*</td>
<td>0.008</td>
<td>0.41</td>
<td>1.19</td>
<td></td>
</tr>
<tr>
<td>Hard Area Ratio*</td>
<td>0.0006</td>
<td>0.24</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Month</td>
<td>1</td>
<td>-</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>1982</td>
<td>2000</td>
<td>2015</td>
<td></td>
</tr>
</tbody>
</table>

* Ratio Predictors are multiplied by the Eno Flow in the model

CALIBRATED FLOW MODEL

$$Response = \beta_{fixed}X_{fixed} + \gamma_{random,j}X_{random,j} + \epsilon$$

<table>
<thead>
<tr>
<th>Model</th>
<th>$Q_{eno}$</th>
<th>$Q_{eno} \times \text{Ratio}_\text{Area}$</th>
<th>Varying Slopes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow</td>
<td>0.03 ± 0.06</td>
<td>1.00 ± 0.06</td>
<td>$Q_{eno}$</td>
</tr>
</tbody>
</table>
Negative Nash-Sutcliffe Efficiencies are set to zero.
NUTRIENT MODEL

• Three different models, one for each nutrient
  • Total Organic Nitrogen (Organic N)
  • Total Nitrogen (Total N)
  • Total Phosphorous (Total P)
• Have only performed calibration currently
• These are preliminary results!

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Min. Value</th>
<th>Avg. Value</th>
<th>Max. Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eno Conc. (C_E) Organic N</td>
<td>0.17</td>
<td>0.34</td>
<td>0.76</td>
<td>(mg/L)</td>
</tr>
<tr>
<td>Eno Conc. (C_E) Total N</td>
<td>0.39</td>
<td>1.21</td>
<td>11.4</td>
<td>(mg/L)</td>
</tr>
<tr>
<td>Eno Conc. (C_E) Total P</td>
<td>0.02</td>
<td>0.12</td>
<td>0.84</td>
<td>(mg/L)</td>
</tr>
<tr>
<td>River Flow (Q_R)</td>
<td>0.002</td>
<td>1.41</td>
<td>43.97</td>
<td>(cms)</td>
</tr>
<tr>
<td>% Agriculture*</td>
<td>2.9</td>
<td>14.7</td>
<td>27.8</td>
<td>%</td>
</tr>
<tr>
<td>% Urban*</td>
<td>5.8</td>
<td>28.9</td>
<td>72.2</td>
<td>%</td>
</tr>
<tr>
<td>Month</td>
<td>1</td>
<td>-</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>1982</td>
<td>2000</td>
<td>2015</td>
<td></td>
</tr>
</tbody>
</table>

* Average value for the entire period of record
NUTRIENT MODEL PREDICTORS

Total Organic Nitrogen Concentrations (mg/L)

CALIBRATED NUTRIENT MODEL

\[\text{Response} = \beta_{\text{fixed}} X_{\text{fixed}} + \gamma_{\text{random}, j} X_{\text{random}, j} + \epsilon\]

<table>
<thead>
<tr>
<th>Model</th>
<th>(Q_R)</th>
<th>(C_E)</th>
<th>% Ag</th>
<th>% Urban</th>
<th>Year</th>
<th>Varying Slopes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organic Nitrogen</td>
<td>-0.11 ± 0.05</td>
<td>0.92 ± 0.08</td>
<td>0.59 ± 0.04</td>
<td>0.044 ± 0.017</td>
<td>-0.07 ± 0.001</td>
<td>(Q_{\text{river}})</td>
</tr>
<tr>
<td>Total Nitrogen</td>
<td>-0.42 ± 0.26</td>
<td>0.48 ± 0.02</td>
<td>-0.18 ± 0.02</td>
<td>-0.04 ± 0.002</td>
<td>(Q_{\text{river}})</td>
<td></td>
</tr>
<tr>
<td>Total Phosphorous</td>
<td>-0.16 ± 0.06</td>
<td>5.87 ± 0.13</td>
<td></td>
<td></td>
<td>-0.04 ± 0.001</td>
<td>(Q_{\text{river}})</td>
</tr>
</tbody>
</table>
Negative Nash-Sutcliffe Efficiencies are set to zero.
FLOW MODEL CONCLUSIONS

- The total area ratio multiplied by the flow in the Eno was positive and significant
- No land cover term was significant
- Model still performs better than simple area ratio method model
- Issues with predicting flows in smaller watersheds such as the Flat River Tributary
NUTRIENT MODEL CONCLUSIONS

• River Flow and Year were significant negative factors in all three nutrient models
• Need to improve modelling of larger watersheds
• Need to take into account nutrient loading into watersheds with wastewater treatment plants

TEMPORAL CONCLUSIONS

• Year was only significant in the nutrient models
• Month was not significant in any model
• Flow model does capture some temporal variation via fluctuations in the Eno River watershed.
QUESTIONS?