Use of Dynamic SPARROW Modeling in Characterizing Time-Lags in Nitrogen Transport in the Potomac River Basin

Presented By
Richard A. Smith
US Geological Survey, Reston, VA

Workshop:
“Management Effects on Water Quality Trends”

Scientific and Technical Advisory Committee, Chesapeake Bay Program
March 26, 2014
Contributors

Scott Ator
Joel Blomquist
John Brakebill
Paul Capel
Anne Hoos
Andrew Lamotte
Molly Macauley (RFF)
Rich Moore
Anne Nolin (Oregon State U.)
Ward Sanford
Andrew Sekellick
Greg Schwarz
Jhih-Shyang Shih (RFF)
• Project objectives
• Dynamic SPARROW modeling
  • Role of MODIS data
• Model calibrations and results
  • South Carolina estuaries
  • Potomac River/Chesapeake Bay
  • Long Island Sound
• Simulating the effects of source reductions
• Exploring the role of groundwater
• Conclusions
SPARROW models are now widely used for investigating spatial aspects of water quality (conditions, processes, and management); however…..

*Limitations of the steady state “assumption” in SPARROW modeling:*
1. **In calibration:** requires long-term averaging and load adjustments for changes in flow and sources.
2. **In interpretation:** contaminant storage and residence time in the watershed are unknown.
3. **In application:** can’t predict the response of contaminant flux to changes in precipitation, temperature, sources, etc (seasonal or longer term). Also, requires “space for time” assumption.

*Research Questions:*
1. Can dynamic SPARROW models be constructed by calibration with time series data?
2. Can dynamic SPARROW models be constructed with “memory” so that they account for storage and time lags in contaminant flux using a recursive form?
3* . Does use of earth observation data facilitate dynamic SPARROW modeling?

• NASA supported research
New seasonal data sets derived from MODIS 8-day 500-m surface reflectance (2001-2009)

Seasonal snow cover frequency

\[
snow \text{ freq} = \frac{N_{snow}}{N_{observations}}
\]

Median Enhanced Vegetation Index (EVI)

\[
EVI = G \frac{NIR \cdot \text{red}}{NIR + (C_1 \cdot \text{red} - C_2 \cdot \text{blue}) + L}
\]

Also: Gross Primary Productivity and Land Surface Water Index
SPARROW’s Reach-Scale Mass Balance
Reach network relates watershed data to monitored loads

\[ LOAD_i = \left\{ \sum_{j \in J(i)} \left[ \sum_{n=1}^{N} S_{n,j} \beta_n \exp(-\alpha'Z_j) \right] \prod_m \exp(-\delta_m^z T_{i,j,m}) \prod_l 1/(1 + \lambda' q_{i,l}^{-1}) \right\} \exp(\varepsilon_i) \]

Required Modification of SPARROW Equation for dynamic modeling
1. Addition of runoff, and lag-1 runoff, to Land-to-water transport term
2. Addition of lag-1 source term(s) based on predicted “concentration” in previous time step.
3. Operates recursively
The procedure for dynamic calibration requires specifying the “source” and “land-to-water” terms for individual reach-level catchments as follows:

\[ E_t = \sum_i f_i S_{t,i} + f_0 E_0 \]

where

- \( E_t \) is export (m/t) of \( N_R \) from the catchment during the current time step;
- \( E_0 \) is export (m/t) of \( N_R \) from the catchment during the previous time step;
- \( \sum f_i S_{t,i} \) is the sum of sources of new \( N_R \) to the catchment during the current time step;
- \( f_t \) and \( f_0 \) are functions of catchment characteristics (e.g. precipitation, temp, runoff, etc).

When sources of \( N_R \) are zero, catchment export falls exponentially according to the recursive relation: \( E_t / E_0 = f_0 \). Assuming the export rate of material from transient storage at a given point in time is proportional to the quantity in storage, the exponential decay rate can be used to estimate the mean residence time of stored material: \( E_t / E_0 = \exp[-(r+k) \Delta t] \) and mean residence time is \( 1/\ln(r+k) \Delta t \)

Where \( r \) and \( k \) are rate coefficients for stream export and watershed loss, and \( \Delta t \) is time step length.

The contribution to export of transient storage relative to “quick” runoff (CTS) is estimated as \( f_0 E_0 / (\sum_i f_i S_{t,i} + f_0 E_0) \)
Basin locations for three dynamic SPARROW models.

- Long Island Sound Drainage
- Potomac River Basin
- South Carolina Coastal Drainage
Independent Variables

N sources (4 or 5)
Precipitation
Temperature
Runoff
Delta runoff
Enhanced vegetation index*
Snow/ice % cover (LIS only)*
In-stream decay
Other (slope, base flow index, depth to gw, carbonate %)

* MODIS data
<table>
<thead>
<tr>
<th>Drainage Basin</th>
<th>Area (km²)</th>
<th>Storage Processes</th>
<th>Model $r^2$ (RMSE)</th>
<th>Estimated Mean Residence Time (yr)</th>
<th>Contribution of Storage to Export (CTS in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Carolina</td>
<td>134,300</td>
<td>Plant/Soil Uptake</td>
<td>0.90 (0.64)</td>
<td>1.1</td>
<td>72</td>
</tr>
<tr>
<td>Potomac River</td>
<td>30,100</td>
<td>Plant/Soil Uptake</td>
<td>0.90 (0.69)</td>
<td>1.5</td>
<td>79</td>
</tr>
<tr>
<td>Long Island Sound</td>
<td>22,059</td>
<td>Plant/Soil Uptake</td>
<td>0.94 (0.44)</td>
<td>0.27</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Snow/Ice Cover</td>
<td></td>
<td></td>
<td>22</td>
</tr>
</tbody>
</table>
Total Nitrogen Yield (kg km$^{-2}$ day$^{-1}$)  
Winter (J, F, M) 2006
Total Nitrogen Yield (kg km$^{-2}$ day$^{-1}$)
Spring 2006
Total Nitrogen Yield (kg km\(^{-2}\) day\(^{-1}\))
Summer 2006
Total Nitrogen Yield \( (\text{kg km}^{-2} \text{ day}^{-1}) \)
Fall 2006
Total Nitrogen Yield (kg km\(^{-2}\) day\(^{-1}\))
Winter 2008
Prediction Errors?

Seasonal Biases?
Seasonal Accuracy

Winter residuals

- Residuals range from -3 to 4
Seasonal Accuracy

Spring residuals

- Standardized residuals range from -2 to 3
Seasonal Accuracy
Seasonal Accuracy

Fall residuals


Standardized Residual

-7.5 to 5.0
Application: Response Time Following Management Action
Simulation: TN Flux, Potomac River at Chain Bridge
All Nonpoint Sources Set to Zero After First Time Step (Winter 2002)
Simulation: TN Flux in Christians Creek, VA
All Nonpoint Sources Set to Zero After First Time Step (Winter 2002)
Predicted Incremental TN Load - Spring 2003

kg/ha per 90 days
High : 175,000,000,000
Low : 0
Predicted Incremental TN Load - Spring 2006

kg/ha per 90 days

High: 175,000,000,000
Low: 0
Predicted Incremental TN Load - Spring 2008

kg/ha per 90 days
High: 175,000,000,000
Low: 0
Remaining Questions?

1. Are there ways to verify the model predictions of response times?
   There may be some special opportunities for this.

2. Response times in some parts of the Chesapeake Bay watershed (esp. Coastal Plain) may be very long (e.g. decades) due to long groundwater residence times. Would a dynamic SPARROW model calibration capture such behavior. Not as currently constructed. Perhaps with additional, longer lag terms.

3. Could a dynamic SPARROW model be constructed to specifically address long groundwater residence times (and the long-term history of nitrogen sources)? Yes, hopefully, as suggested below........
USGS Focus on Integrated Modeling of Storage Processes and “Lags” in Effects of Nutrient and Sediment Controls in the Chesapeake Bay Watershed

• Many participants:
  - MD Water Science Center
  - National Water Quality Assessment Program
  - National Research Program

• Combined statistical and deterministic modeling of nutrients and sediment.

• Initial emphasis on developing time series of groundwater N inputs to SPARROW models based on historical source data.

• Will explore multiple methods for estimating groundwater age distributions, discharge rates, and nitrate concentrations to combine with historical source data.
Approximate Groundwater and Stream Water Age Distributions for Potomac Basin Based on Preliminary Modflow Modeling (Ward Sanford, personal communication)

<table>
<thead>
<tr>
<th>Time Interval</th>
<th>% of Groundwater</th>
<th>% of Stream Water (runoff + groundwater)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Year (2014)</td>
<td>33</td>
<td>66</td>
</tr>
<tr>
<td>2010 - 2013</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2000 - 2009</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>1990 – 1999</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1980 - 1989</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>1970 - 1979</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>1960 - 1969</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>Older than 5 yr</td>
<td>48</td>
<td>32</td>
</tr>
<tr>
<td>“ “ 10 yr</td>
<td>37</td>
<td>25</td>
</tr>
<tr>
<td>“ “ 25 yr</td>
<td>24</td>
<td>16</td>
</tr>
<tr>
<td>“ “ 50 yr</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>“ “ 100 yr</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>
Historical Data Development

- **Atmospheric Deposition estimates 1940-1980:**

- **Atmospheric Deposition estimates 1985-2010:**

- **Housing density estimates for 1940-2030:**

- **Fertilizer estimates for 1945-1985:**

- **Fertilizer estimates for 1982-2001:**

- **Animal numbers for 1950-2002:**

- **Manure estimates 1950-2002:**

- **Recharge, Bedrock Geology, and Surficial Geology**

- Data is averaged in years where gaps exist (for example, between Agricultural Census years and atmospheric deposition data from 1981-1984).

- Atmospheric deposition data from 1940-1980 is only available as a regional multiplier that was applied to 1985 NADP grid data.

- Manure estimated from Agricultural Census animal counts and manure averages per animal from USDA NRCS. High and low ranges estimated.

- Data organized by NHDPlus V1 COMID. Since agriculture census data is only available by county, these data were processed with a simple spatially weighted average.
Specifying a Groundwater Source in SPARROW: Multiple Approaches

1. Groundwater characteristics (NO3 conc., recharge, BF load, rock type, etc) included in either “source” or “LTW” terms in steady state models.

2. Time series of estimated groundwater inputs included as “sources” in dynamic models.
   - based on historical source data, groundwater age distributions, BF load data.
   - based on historical source data, groundwater age distributions, stream recharge rate, and groundwater concentration data.
Summary

- The results of initial calibrations of dynamic SPARROW TN models based on seasonal time series of water quality and basin attribute data were highly encouraging.
- MODIS EVI and Snow/Ice Cover were especially strong predictor, appearing to account for seasonal retention of nitrogen in basin vegetation and snowpack.
- Model predictions for the stream network show moderately accurate (and seemingly realistic) seasonal and year-to-year variations in yield. Model coefficient estimates were very precise due to many observations.
- Predicted response times are “reasonable” but difficult to verify.
- Efforts to include specific terms for groundwater residence times (and the long-term history of nitrogen sources) in SPARROW models are underway.
Remaining Tasks

1. Refine calibrations (e.g. try a few additional predictors and specifications)

2. Test serial correlation of residuals and develop variance correction method

3. Write up and publish overall method and results
Additional Projects

1. Apply method to additional basins.
   a. Upper Klamath lake (project initiated)
   b. *Entire Chesapeake Bay watershed

2. *Examine/test reasonableness of model predictions and interpretations with independent data.

* Possible collaboration with IWS
Economic Value of MODIS Enhancements: Cost of Nutrient Control vs Model Error

Cost of 90% Certain Nutrient Reduction

RMSE of SPARROW Model

With MODIS Data

Without MODIS Data
Premise: Cost of management actions could be lowered if spring TN loads were more accurately known.
Percent of the year with frozen ground (snow) has a significant effect on increasing the delivery of phosphorus from animal manure to streams throughout the Mississippi River Basin.
Preliminary Calibration of Dynamic SPARROW Model of Total Nitrogen in Potomac Basin

- Chesapeake system: Need for spatially-detailed, seasonal TN loads, and information on response times (i.e. lags) of controls
- Based on NHD stream network (16,000+ reaches/catchments)
- 81 water-quality monitoring stations for “observed” flux
- TN sources: point, urban runoff, atmosphere, fertilizer, farm animal waste, catchment “storage”*
- Land-to-water drivers: runoff, lag-1 runoff, MODIS vegetation index
- Seasonal time series of all data for fall 2001 through fall 2008

* New mechanism required for dynamic modeling
Improving Water Quality Management: Use of Earth Observations in SPARROW

Presented by
Richard A. Smith
US Geological Survey

Sixth International Nitrogen Conference
Kampala, Uganda
November 18 – 22, 2013
Improving Water Quality Management: Use of Earth Observations in SPARROW

Research Team

John Brakebill\textsuperscript{2}, Anne Hoos\textsuperscript{2}, Molly Macauley\textsuperscript{1}, Richard Moore\textsuperscript{2}, Anne Nolin\textsuperscript{3}, Dale Robertson\textsuperscript{2}, Gregory Schwarz\textsuperscript{2}, Jhih-Shyang Shih\textsuperscript{1}, and Richard Smith\textsuperscript{2}

Thank you:

Resources for the Future

USGS

Oregon State University

NASA
• **Mechanistic Features**
  - contaminant sources and landscape attributes linked to stream/river network
  - nonlinear contaminant processes
  - non-conservative transport
  - Steady-state mass balance form
  - dynamic version (with MODIS input) under development

• **Statistical Features**
  - “data-driven” (from large, long-term, monitoring network (1000+ sites)
  - statistical calibration (nonlinear regression)
  - coefficients estimated from the data, not literature
  - promotes hypothesis testing of mechanistic interpretation
  - provides error quantification
Water quality parameters modeled with SPARROW to date

1. Reactive Nitrogen
2. Total Phosphorus
3. Total Organic Carbon
4. Total Dissolved Solids
5. Suspended Sediment
6. Ammonium
7. Fecal Coliform Bacteria
8. Cryptosporidium
9. Atrazine

* This study
Example Application: Quantifying the Sources of Nutrients Delivered to the Gulf of Mexico

Mississippi/Atchafalaya River Basin

![Map of the Mississippi/Atchafalaya River Basin](image)

![Bar chart showing nutrient sources](image)

# Calibration Results

<table>
<thead>
<tr>
<th>N Source</th>
<th>Units</th>
<th>Coeff.</th>
<th>Significance (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point sources</td>
<td>kg/yr</td>
<td>0.66</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>Urban</td>
<td>sq km</td>
<td>427</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>Atmos.</td>
<td>kg/yr</td>
<td>0.11</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>Fertilizer</td>
<td>kg/yr</td>
<td>0.034</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>Animal waste</td>
<td>kg/yr</td>
<td>0.060</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>“Storage”</td>
<td>kg/yr</td>
<td>0.85</td>
<td>&lt; 10^{-4}</td>
</tr>
<tr>
<td>ln EVI</td>
<td>-</td>
<td>-0.90</td>
<td>&lt; 10^{-4}</td>
</tr>
</tbody>
</table>

No. of observations 2268

R^2 0.90

Yield R^2 0.68

RMSE 0.69
Partners and User Community

• USGS (4 science Centers + Reston)
• USEPA
  ➢ Chesapeake Bay Program
  ➢ Narragansett Laboratory
  ➢ Gulf of Mexico Program
• State of South Carolina
• NEIWPCC
• Four NE State DEPs
Low concentrations of dissolved oxygen (hypoxia), as a result of nitrogen enrichment, often occur during the summer in the western part of LIS.

**Partners / User Community**

- New England Interstate Water Pollution Control Commission (NEIWPCC)
- Four New England States especially Connecticut Dept. Env. Protection
- New York Dept. Env. Conservation
- U.S. Environmental Protection Agency (USEPA)
Nitrogen transport from the watershed to LIS varies seasonally. Much of the nitrogen transport occurs during the spring freshet.

**Modeling Approach**

Dynamic – Seasonal SPARROW
Winter 2001 - Summer 2009

Seasonal loads at monitoring sites are the dependent variable

Standard suite of SPARROW predictors plus NASA predictors compiled seasonally:
- The Enhanced Vegetation Index (EVI)
- Percent snow cover
Long Island Sound (LIS)

**Anticipated outcome**

Improved understanding of the source and transport of nitrogen to LIS and how it varies seasonally.

Intended to aid in targeting nutrient controls.

Modeling “in progress”.
Algae and Cyanobacteria blooms are significant problem.

Existing SPARROW models are used to predict mean annual N and P conditions.

State HEC officials would prefer seasonally-specific (spring/summer) predictions.

Preliminary Seasonal SPARROW TN model with MODIS EVI input has been (highly) successfully calibrated.

Especially close communication with (and high anticipation from!) state and local officials.
## Study Basins and Model Details

<table>
<thead>
<tr>
<th>Drainage Basin</th>
<th>Area (km²)</th>
<th>Stream Segments</th>
<th>In-Stream Monitoring Stations</th>
<th>No. Time Steps</th>
<th>Nitrogen Sources</th>
<th>Storage Processes (MODIS DATA)</th>
<th>Total Est. Coeff</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Carolina Coastal</td>
<td>134,300</td>
<td>76,742</td>
<td>34</td>
<td>15</td>
<td>Point, Urban, Atmosphere, Fert., Manure</td>
<td>Plant/Soil Uptake</td>
<td>14</td>
</tr>
<tr>
<td>Potomac River (Ches. Bay)</td>
<td>30,100</td>
<td>16,500</td>
<td>81</td>
<td>28</td>
<td>Point, Urban, Atmosphere, Fert., Manure</td>
<td>Plant/Soil Uptake</td>
<td>13</td>
</tr>
<tr>
<td>Long Island Sound</td>
<td>41,600</td>
<td>22,059</td>
<td>38</td>
<td>35</td>
<td>Point, Urban, Atmosphere, Fert. Rotation, Fert. Other, Manure</td>
<td>Plant/Soil Uptake, Snow/Ice Cover</td>
<td>15</td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th>Drainage Basin</th>
<th>Area (km²)</th>
<th>Storage Processes</th>
<th>Model r² (RMSE)</th>
<th>Est. Mean Residence Time (yr)</th>
<th>Δ N Flux / Δ GPP</th>
<th>Δ Flux / Δ Spring Snow Cover (% / %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Carolina</td>
<td>134,300</td>
<td>Plant/Soil Uptake</td>
<td>0.90 (0.64)</td>
<td>1.1</td>
<td></td>
<td>-4.2</td>
</tr>
<tr>
<td>Potomac River</td>
<td>30,100</td>
<td>Plant/Soil Uptake</td>
<td>0.90 (0.69)</td>
<td>1.5</td>
<td></td>
<td>-2.2</td>
</tr>
<tr>
<td>Long Island Sound</td>
<td>22,059</td>
<td>Plant/Soil Uptake</td>
<td>0.94 (0.44)</td>
<td>0.27</td>
<td></td>
<td>-4.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Snow/Ice Cover</td>
<td></td>
<td></td>
<td></td>
<td>0.5 - 1.0</td>
</tr>
</tbody>
</table>
**Approach**: estimate time-variable baseflow nitrate (BF-NO₃) and groundwater discharged nitrate (GWD-NO₃) loads at selected stream sites within the Chesapeake Bay watershed that have continuous [NO₃], [SC], and discharge data.

**Objectives**:
1. Demonstrate the difference between BF-NO₃ loads and GWD-NO₃ loads. Specifically, by investigation of the ratio of BF-NO₃ and GWD-NO₃ loads to STR-NO₃ loads, we will be able to assess the relative importance of in-stream processing over time and space.
2. Provide “gold standard” estimates of BF-NO₃ and GWD-NO₃ loads using continuous nitrate data against which similar estimates obtained using discrete [NO₃] data will be compared.
Groundwater TN Loads: Potomac River at Chain Bridge

Options…
1. Baseflow discharge x LOADEST Predicted [TN]
2. Baseflow discharge x Average Winter [TN] at BFI=1.0
3. Baseflow discharge x Time Variable Winter [TN] at BFI=1.0

From “big picture” perspective there is little difference in loads among groundwater [TN] options
SPARROW and Groundwater

Traditional SPARROW models do not include GW sources of N to the stream.

ASSUMPTION: GW, as a source of N to the stream, can be added to SPARROW

1) by the addition of “land-to-water variables” for subsurface characteristics, N loads, …that the model could choose
   → No changes in equations
   → Multiple descriptors of the subsurface could be useful

2) by adding a N in GW as a “source term”

Sources of N to the stream: N processed “recently” through landscape