east_coast_sparrow

Bryan Milstead

02 August 2017

Introduction

We are interested in using the SPARROW model to estimate how proposed changes to the air rules will affect the trophic states of lakes on the east coast from Maine to Florida. An east coast SPARROW model has been developed that incorporates CMAQ (Community Multi-scale Air Quality Model; see http://tinyurl.com/gv8k9q9) estimates of nitrogen inputs from atmospheric sources. The CMAQ model also provides predicted reductions in atmospheric nitrogen inputs expected to occur under current regulations by 2025. Since the East Coast model is parameterized with the CMAQ nitrogen inputs for 2002 we can look at how reductions will affect nitrogen loading to flow lines all along the east coast. Predictions of 2002 and 2025 nitrogen loads to flowlines can be aggregated to lakes to estimate total loads and in-lake concentrations of nitrogen. The expectation is that these changes in nitrogen inputs will lead to changes in the trophic structure of the lakes. The conventional understanding, however, is that phosphorus is more often the limiting nutrient in lakes and therefore it is questionable whether nitrogen reductions will have a major effect. We cannot deny the importance of phosphorus as a limiting nutrient but increasingly researchers are demonstrating that increased nitrogen loads (especially atmospheric inputs) are strongly impacting lakes leading to increased eutrophication (see Conely et al 2009; http://science.sciencemag.org/content/323/5917/1014.full). Therefore, there is justification for looking at changes in trophic state associated with nitrogen load reductions

Trophic state is a measure of the productivity of a lake. Ideally, we'd like to measure the annual net primary production of a lake to determine its trophic state but this can be costly and technically challenging. In the absence of detailed data on a lake's primary production we often depend on trophic state indices to estimate trophic state. The Carlson (1977) trophic state indices are widely accepted as the best estimates of lake trophic state. These indices use a scale from 0 to 100 (mostly) to define trophic state based on Secchi Depth, Chlorophyll a, or Total Phosphorus. Kratzer and Brezonik (1981) developed a trophic state index that it is compatible with the Carlson indices based on Total Nitrogen. All four indices are nicely summarized on Carlson's secchi dip in website (http://www.secchidipin.org/index.php/monitoring-methods/trophic-state-equations/). The general forumula for the TSI for Nitrogen is:

$TSI(TN) = 54.45 + 14.43 \ln(TN)$

- Trophic State Index for Nitrogen (TSI_N) = 54.45 + (14.43 * ln(total_nitrogen))
  - Note: ln is the natural logrithm; total nitrogen must be in mg/l.

The trophic state index is a continuous linear measure of lake trophic state but it can be used to assign lakes to standard trophic state categories. On the secchi dip in webpage Carlson provides some threshold values for assigning categories using the indices. The thresholds are a little fuzzy with some intermediate states. Below are the threshold values:

- Oligotrophic: $TSI \leq 40$
- Mesotrophic: $TSI > 40$ and $TSI \leq 50$
- Eutrophic: $TSI > 50$ and $TSI \leq 70$
- Hypereutrophic: $TSI > 70$

The TSI for nitrogen appears to be a reasonable predictor of lake tropic state. We validated this claim by comparing trophic state indices for nitrogen, phosphorus, chlorophyll, and secchi depth using the 2007 National Lake Assessment data set (see: https://www.epa.gov/national-aquatic-resource-surveys/nla). The results supported the use of TSI_N (see: https://github.com/willbmisled/east_coast_sparrow/blob/master/r/nlaTrophicStateData2007.Rmd)

data sources

The 2002 base conditions ('data/mrb1mrb2_massbalance_2002base.sas7bdat') and the 2025 predictions ('data/mrb1mrb2_massbalance_2025.sas7bdat') were provided by Anne Hoos in SAS format. See email message below:

From: Hoos, Anne [mailto:abhoos@usgs.gov] Sent: Friday, May 19, 2017 6:16 PM To: Milstead, Bryan Milstead.Bryan@epa.gov Cc: Moore, Richard rmoore@usgs.gov Subject: Re: CMAQ 2025 lake scenario

Good to talk with both of you as well, and thanks for getting this going Rich.

I've placed some files on the ftp depot ftp://ftpext.usgs.gov/pub/er/tn/nashville/Hoos/CMAQScenario_Simulations/ - two separate SAS files, each with the 'mass balance' postprocessing from the SPARROW run (2002 and 2025). They should be visible about an hour from now. Because my postprocessing program does phosphorus at the same time you get that as a bonus in these :) I still need to postprocess the source share calculations for each lake - but you probably don't need that information for your analysis.

NOTE: two updates to these files were received from Anne on 5/23/17

data definitions (see: "data/README_definitionofattributes_massbalancereservoirs_MRB1MRB2.xlsx")

- each file contains the following fields with data for 41,566 lakes.

Attribute	Definition
WBRchCd	Waterbody Reachcode, unique identifier for a reservoir or lake. Most of these corresponds with the attribute 'REACHCODE' in the shapefile NHDwaterbody except where manual fixes were made (documented in the SAS program that populates the SPARROW input data set); in a few cases a new reachcode was assigned by Anne Hoos in 2010
invhload_aggreg	Inverse of the hydraulic load aggregated for lake (i.e. inverse of hload_aggreg, see below for calculation); in days/meter - this is the value reported in DS 820 as "Residence time, d/m"
TN_arealloading_in_perflush	(misnomer, should be named TN_LOAD_in_perflush) Tributary inflow load to lake per flushing rate (i.e. per hydraulic load, or could express as inflow load multiplied by inverse hydraulic load), in kg*d/m, calculated as TN_LOAD_inflow / hload_aggreg - this is the value reported in DS 820 as "Load from watershed per hydraulic flushing rate of receiving waterbody"
TP_arealloading_in_perflush	(misnomer, should be named TP_LOAD_in_perflush) Product of TP_arealloading_inload and invhload_aggreg (see definitions of these attributes elsewhere)
TN_CONC_inload	Tributary inflow Nitrogen Concentration, mg/L - this is value reported in USGS DS 820 as "Concentration of tributary inflow to receiving water", calculated as (TN_LOAD_inflow) / (TOT_CFS * 28.32 * 1/1000000 * 86400 * 365). Would have preferred to use sum of tributary inflow MAFLOWU as the demonimator but this was difficult to compute; thus it's a slightly mixed calculation
TP_CONC_inload	Tributary inflow Phosphorus Concentration, mg/L - this is value reported in USGS DS 820 as "Concentration of tributary inflow to receiving water", calculated as (TP_LOAD_inflow) / (TOT_CFS * 28.32 * 1/1000000 * 86400 * 365). Would have preferred to use sum of tributary inflow MAFLOWU as the demonimator but this was difficult to compute; thus it's a slightly mixed calculation

Attribute	Definition
TNTP_CONC_inload	Ratio of TN_CONC_inload to TP_CONC_inload - this is the value reported in USGS DS 820 as “Ratio of TN:TP”
TN_massdecay	Nitrogen mass removed in the lake, kg/yr, summed from the amount removed in each incremental reservoir segment - this is value reported in USGS DS 820 as “Load assimilated in receiving waterbody”
TP_massdecay	Phosphorus mass removed in the lake, kg/yr, summed from the amount removed in each incremental reservoir segment - this is value reported in USGS DS 820 as “Load assimilated in receiving waterbody”
TN_decayperinflow	Ratio of TN_massdecay to TN_LOAD_inflow
TP_decayperinflow	Ratio of TP_massdecay to TP_LOAD_inflow
Res_NumSegments	Total number of flowline segments associated with reservoir or lake
WB_AREASQKM	Total area of the waterbody polygon (from the NHDwaterbody shapefile). Where multiple waterbody polygons constitute a reservoir this number is smaller than Res_sumofsegmentSA_KM2
WB_FTYPE	Type of waterbody (only Lake/Pond waterbody polygons are included in the SPARROW input data set)
WB_GNIS_ID	Waterbody ID from the Geographic Name Information System, from the shapefile NHDwaterbody (populated for MRB2 only)
WB_Comment	Comment added by Anne Hoos or Jennifer Murphy if changes were made to attributes from NHDwaterbody (e.g. if WBRchCd was changed where two or more WBRchCd values refer to the same reservoir, or if determined that waterbody was coastal/brackish rather than freshwater and therefore SA_KM2 set to 0) (populated for MRB2 only)
TempArcID	Flag (set equal to waterbody comid) indicating that this is a single-polygon reservoir/lake (verified from inspection) that was missing Waterbody Reachcode assignment in the shapefile NHDwaterbody and therefore WBRchCd has been set to waterbody comid (populated for MRB2 only)
outletreach_comid	COMID of outlet reach, i.e. the flowline identified as the most downstream (largest value of cumdrainag) for the reservoir or lake (populated for MRB2 only)
outletreach_LAT	Latitude (decimal degrees) of downstream node of the outlet reach (populated for MRB2 only)
outletreach_LON	Longitude (decimal degrees) of downstream node of the outlet reach (populated for MRB2 only)
hload_aggreg	Hydraulic load of the lake/reservoir (also known as the surface overflow rate, qs), in meters per day, calculated as MAFLOWU * 0.02832 * 86400/ (Res_sumofsegmentSA_KM2 * 1000000)
hload_aggreg_yr	Hydraulic load, in meters per year
RetentCoeff_Nurnberg	Retention Coefficient (Rs) of the reservoir, computed as 15 / (hload + 18) (using Nurnberg, 1998)
FIPS_ST	FIPS code of State in which downstream node of outlet reach is located (populated for MRB2 only)
TN_LOAD_outflow	Total nitrogen load at downstream node of exit reach, in kg/yr - this is the value reported in USGS DS 820 source share table as “Total load”. Wanted to use TN_LOAD_inflow here, to be consistent with use of trib inflow info in the table that follows, but see explanation at top of this file

Attribute	Definition
TP_LOAD_outflow	Total phosphorus load at downstream node of exit reach, in kg/yr - this is the value reported in USGS DS 820 source share table as “Total load” - Wanted to use TN_LOAD_inflow here, to be consistent with use of trib inflow info in the table that follows, but see explanation at top of this file
TNTP_LOAD_outflow	Ratio of TN_LOAD_outflow to TP_LOAD_outflow
TN_LOAD_inflow	Total nitrogen load entering lake from all tributaries and adjacent areas, calculated from mass balance (load at downstream node of exit reach plus load removed in lake)
TP_LOAD_inflow	Total phosphorus load entering lake from all tributaries and adjacent areas, calculated from mass balance (load at downstream node of exit reach plus load removed in lake)
TNTP_LOAD_inflow	Ratio of TN_LOAD_inflow to TP_LOAD_inflow
TN_arealloading_inload	Net specific or areal nitrogen loading (L) of the reservoir (kg/km2-yr), computed as ratio of TN_LOAD_inflow to AreaSqKm
TN_arealloading_outload	Net specific or areal nitrogen loading (L) of the reservoir (kg/km2-yr), computed as ratio of TN_LOAD_outflow to AreaSqKm
TN_CONC_inlake	Estimated In-Lake Nitrogen Concentration of reservoir (mg/L), calculated as TN_arealloading_outload / hload * (1 - RetentCoeff_Nurnberg) (equation corrected 10/3/2014)
TP_arealloading_inload	Net specific or areal phosphorus loading (L) of the reservoir (kg/km2-yr), computed as ratio of TP_LOAD_inflow to AreaSqKm
TP_arealloading_outload	Net specific or areal phosphorus loading (L) of the reservoir (kg/km2-yr), computed as ratio of TP_LOAD_outflow to AreaSqKm
TP_CONC_inlake	Estimated In-Lake Phosphorus Concentration of reservoir (mg/L), calculated as TP_arealloading_outload / hload * (1 - RetentCoeff_Nurnberg) (equation corrected 10/3/2014)
TNTP_CONC_inlake	Ratio of TN_CONC_inlake to TP_CONC_inlake
MRB	1 = MRB1, 2 = MRB2
huc8	Numeric version of 8-digit hydrologic unit code for the reach segment at the downstream end of the watebody polygon
Group_Generic	Grouping number so that lakes can be sorted in hydrologic groups, equal to HUC4 in MRB1, and to a nominal ordering number in MRB2
WB_NAME	Waterbody name from the Geographic Name Information System (with corrections/revisions from Jennifer Murphy and Anne Hoos, 2010
UPSTRM_AREA	Cumulative drainage area (km2) for the downstream node of the flowline reach segment at the downstream end of the reservoir (same as CUMDRAINAG in MRB2 and DEMTAREA in MRB1)
FLOW_OUTLET_CFS	Mean annual streamflow (cfs) for the reach segment at the downstream end of the reservoir (TOT_CFS for MRB1, MAFLOWU for MRB2)

data steps

- read the sas datafiles and convert to r data.frames

- write a function to convert "TN_CONC_inlake" (mg/l) to trophic state
- join datasets, assign trophic state, and calc trophic state change
- write results to 'output/ec_sparrow_lake_tropic_state.csv'
- "ec_sparrow_lake_tropic_state.csv" is avaiable here: https://github.com/willbmisled/east_coast_sparrow/blob/master/output/ec_sparrow_lake_tropic_state.csv
- The file contains estimates of total nitrogen and trophic state designations for the two time periods (2002 & 2025) for 41,566 lakes.
- The dataset also indicates whether trophic state is estimated to change between the years.
- The lakes can be joined to NHDplus with the field "WBRchCd".

data definitions for "ec_sparrow_lake_tropic_state.csv"

- each file contains tye following fields with data for 41,566 lakes.

Attribute	*Definition**
WBRchCd	Waterbody Reachcode, unique identifier for a reservoir or lake. Most of these corresponds with the attribute 'REACHCODE' in the shapefile NHDwaterbody except where manual fixes were made (documented in the SAS program that populates the SPARROW input data set); in a few cases a new reachcode was assigned by Anne Hoos in 2010. Note this was used to join the original two datasets.
tn02	Estimated In-Lake Nitrogen Concentration of reservoir (mg/L), calculated as TN_arealloading_outload / hload /=* (1 - RetentCoeff_Nurnberg) (equation corrected 10/3/2014) for year=2002. The name of this field in the source dataset ("mrb1mrb2_massbalance_2002base.sas7bdat") was: "**TN_CONC_inlake**"
tn25	Estimated In-Lake Nitrogen Concentration of reservoir (mg/L), calculated as TN_arealloading_outload / hload /=* (1 - RetentCoeff_Nurnberg) (equation corrected 10/3/2014) for year=2025. The name of this field in the source dataset ("mrb1mrb2_massbalance_2025base.sas7bdat") was: "**TN_CONC_inlake**"
ts02	Predicted trophic state for the 2002 lakes based on "tn02";"oligo"="oligotrophic", "meso"="mesotrophic", "eu"="eutrophic", and, "hyper"="hypertrophic".
ts25	Predicted trophic state for the 2025 lakes based on "tn25";"oligo"="oligotrophic", "meso"="mesotrophic", "eu"="eutrophic", and, "hyper"="hypertrophic".
ts_chg	Indicated whether the predicted lake trophic state changed between 2002 and 2025; ts_chg=0 indicates the lakes were assigned the same trophic state in both years; ts_chg=1 indicates the predicted trophic state for 2025 was less eutrophic by 1 trophic state category compared to 2002. No lakes became more eutrophic or changed more than one trophic state category.

Results

- The 2002 and 2025 in-lake concentations of nitrogen (mg/l) were estimated for 41,566 lakes along the east coast.
- The estimated 2025 in-lake nitrogen concentrations were slightly lower than the estimated values for

2002.
- The results are summarized below.

```
## [1] "Summary of the 2002 TN data"

##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
##   0.00077   0.09780   0.17820   0.35320   0.36140 111.20000

## [1] "Summary of the 2025 TN data"

##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
##   0.00068   0.08115   0.15270   0.31280   0.30780 111.10000
```

- The field "ts_chg" shows which lakes changed in predicted trophic state between 2002 and 2025. Some lakes stayed the same (ts_chg=0) and some became less eutrophic by one trophic state level (ts_chg=0; e.g., a lake that was eutrophic in 2002 was mesotrophic in 2025). No lakes became more eutrophic and none changed by more than 1 trophic state class. A total of 2385 lakes "improved" in 2025.

- It is more interesting, however, to look at the details of the tropic state changes. The table below shows all possible changes (or not) and the number of lakes in each category:

```
##    ts02  ts25     n
## 1 oligo oligo 31341
## 2  meso oligo  1587
## 3  meso  meso  4104
## 4    eu  meso   754
## 5    eu    eu  3400
## 6 hyper    eu    44
## 7 hyper hyper   336
```