Predicting Sources of Dissolved Organic Nitrogen to an Estuary from an Agro-Urban Coastal Watershed

Christopher L. Osburn,*† Lauren T. Handsel,†§ Benjamin L. Peierls,‡ and Hans W. Paerl‡

*Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina 27695 United States
†Institute of Marine Sciences, University of North Carolina at Chapel Hill, Morehead City, North Carolina 28557 United States

Supporting Information

ABSTRACT: Dissolved organic nitrogen (DON) is the nitrogen (N)-containing component of dissolved organic matter (DOM) and in aquatic ecosystems is part of the biologically reactive nitrogen pool that can degrade water quality in N-sensitive waters. Unlike inorganic N (nitrate and ammonium) DON is comprised of many different molecules of variable reactivity. Few methods exist to track the sources of DON in watersheds. In this study, DOM excitation−emission matrix (EEM) fluorescence of eight discrete DON sources was measured and modeled with parallel factor analysis (PARAFAC) and the resulting model (“FluorMod”) was fit to 516 EEMs measured in surface waters from the main stem of the Neuse River and its tributaries, located in eastern North Carolina. PARAFAC components were positively correlated to DON concentration. Principle components analysis (PCA) was used to confirm separation of the eight sources and model validation was achieved by measurement of source samples not included in the model development with an error of <10%. Application of FluorMod to surface waters of streams within the Neuse River Basin showed that while >70% of DON was attributed to natural sources, nonpoint sources, such as soil and poultry litter leachates and street runoff, accounted for the remaining 30%. This result was consistent with changes in land use from urbanized Raleigh metropolitan area to the largely agricultural Southeastern coastal plain. Overall, the predicted fraction of nonpoint DON sources was consistent with previous reports of increased organic N inputs in this river basin, which are suspected of impacting the water quality of its estuary.

INTRODUCTION

Eutrophication, the increase in organic matter supply to aquatic ecosystems, is a widespread problem often linked to anthropogenic nutrient enrichment in estuaries, especially nitrogen (N), since it is the primary nutrient limiting algal production.1−4 Unlike inorganic N (nitrate, nitrite, ammonium), dissolved organic N (DON) in aquatic systems is likely made up of diverse compounds (e.g., amino acids, urea, humic substances) with varying reactivity, bioavailability, and concentration, which, along with inorganic N, supports the growth of phytoplankton and bacteria.5−10 and may differentially favor certain phytoplankton taxa, including harmful algal bloom (HAB) species.11−13 Prior work suggests that riverine fluxes of DON will increase in the future due to anthropogenic and climactic factors.14−17 In rivers, DON can comprise a substantial fraction of the total dissolved N (TDN, or DON + DIN; the latter being the sum of nitrate, nitrite, and ammonium) load from forested catchments, yet in urban and agricultural catchments, this ratio typically is smaller.18 Tracking these sources of N (and other organic and inorganic nutrients) is of key importance for computing load estimates from streams and rivers to receiving waters such as lakes, reservoirs and estuaries. Sources of organic matter (OM) have been tracked by using their fluorescence properties (e.g., chromophoric dissolved organic matter, CDOM) (e.g., refs 19−22). For example, the indole moiety directly imparts characteristic protein-like fluorescence properties to organic

Received: January 7, 2016
Revised: June 4, 2016
Accepted: July 12, 2016
Published: July 12, 2016

DOI: 10.1021/acs.est.6b00053
matter, including N bound to aromatic macromolecular material (so-called "humic substances"), making this technique relevant for measurements of DON.\textsuperscript{23,24} Fluorescence of filtered water samples is relatively easy to measure and is highly informational, and can be evaluated using multivariate analysis techniques, especially parallel factor analysis (PARAFAC).\textsuperscript{25,26} EEM-PARAFAC models have been developed for a number of natural and artificial systems ranging from rivers, lakes, and estuaries to recycled water and water treatment systems and a searchable database (OpenFluor, \texttt{http://www.openfluor.org}) exists where new models can be compared to existing models.\textsuperscript{27,28}

Substantial challenges exist to utilizing EEM-PARAFAC models as water quality proxies. First, the signals (components) from a PARAFAC model are not unique to specific compounds but rather are thought to be specific to classes of molecules. Thus, PARAFAC models of organic matter do not identify discrete fluorophores.\textsuperscript{23,29} This means empirical relationships between fluorescent components and water quality parameters must be developed. Second, any water quality model developed using organic matter fluorescence potentially is system specific. In other words, a model developed for one watershed may not be directly applicable to other watersheds. Rivers impacted by different land uses exhibit distinct fluorescence patterns.\textsuperscript{30} EEM fluorescence of farm wastes (swine and cattle slurry, sheep barn waste) and rivers influenced by sewage and by wastewater treatment facility (WWTF) effluences are enriched in protein-like fluorescence resembling the amino acids tryptophan and tyrosine.\textsuperscript{31} These differences are distinct from natural organic matter (NOM) in streams and have important implications for organic matter metabolism in and health of aquatic ecosystems.\textsuperscript{32}

**Figure 1.** Representative excitation–emission matrices (EEMs) of eight probable DON sources to the Neuse River and its tributaries. Sources measured were (a) Reference (Deep Creek); (b) WWTF Effluent (Kinston); (c) WWTF Influent (Kinston); (d) Poultry (litter leachate); (e) Swine lagoon; (f) Septic; (g) Soil (topsoil leachate); and (h) Street (runoff in Pigeon House Creek after rain). Note QSU scaling differences on the color bar for each EEM.
The aim of this study was to use organic matter fluorescence as a means to track potential sources of DON in the Neuse River Basin (NRB), a coastal river watershed in eastern North Carolina, whose estuary (the Neuse River Estuary, NRE) is a major tributary to the Pamlico Sound, the nation’s second largest estuarine complex. Fluorescence was measured on a series of eight potential sources of DON to this river system and modeled with PARAFAC. We hypothesized that PARAFAC model components determined for the eight sources would be correlated to source DON concentrations. Further, we expected that when applied to streams and the Neuse River proper, fluorescence would be dominated by natural background but reflect contributions from urban and agricultural sources of organic matter in response to land cover and land use. The output from PARAFAC was used in a mixing model, termed FluorMod, to estimate the relative amounts of...
DON originating from the eight sources at three locations along the Neuse River proper and three tributaries to the river. Each site was located along a land use gradient from the urbanized Raleigh-Durham metropolitan area in the Piedmont to the rural and agricultural-Atlantic coastal plain of the U.S. We discuss the development, calibration, and validation of this modeling approach and its utility for monitoring nutrient water quality.

### MATERIALS AND METHODS

#### Sample Collection. The Neuse River Basin is approximately 16,000 km² in size and the river itself is about 320 km in length, originating as outflow from Falls Lake, a reservoir near the Raleigh metropolitan area, and flowing southeasterly from the northern Piedmont through the Atlantic coastal plain through an increasing rural and agriculturally dominated landscape. Nutrient loading is primarily from nonpoint sources to tributaries and the Neuse River proper. Concentrated animal feeding operations (swine, poultry) are heaviest in the Middle Neuse near the city of Kinston, NC. The lower Neuse River and upper estuary have experienced nutrient loadings from the northern Piedmont through the Atlantic coastal plain (see Supporting Information (SI)). Monthly samplings were conducted in the NRB at 10 sites (SI Table S1). Samples were collected from just below surface using water samplers lowered into streams from bridge overpasses approximately midstream, kept on ice or cold, in the dark, until returned to the laboratory, generally within 24 h. In the laboratory, samples were filtered through 0.7 μm porosity glass filters (precombusted at 450 °C for 6 h) to collect particulates (POM), the filtrate (DOM) was collected in detergent-washed and ultrapure water-rinsed polycarbonate or glass vials and kept refrigerated and in the dark for up to 1 week until analysis.

#### Optical Analyses. Absorbance (200–800 nm) was measured on filtered samples using a Varian Cary 300UV spectrophotometer in 1, 5, or 10 cm quartz cells and diluted if absorbance was greater than 0.4 at 240 nm (1 cm cells only). Ultrahigh purity laboratory water (18.2 MΩ resistivity) was used as a blank, and the blank-corrected absorbance values were converted to Napierian absorption coefficients (aλ) using the following equation:

$$ a_\lambda = 2.303 \times \frac{A_{\lambda,\text{meas}} - A_{\lambda,\text{blank}}}{L} $$  \hspace{1cm} (1)

where $A_{\lambda,\text{meas}}$ is the measured Absorbance (aka optical density “OD”); unitless of a sample at wavelength, $\lambda$; $A_{\lambda,\text{blank}}$ is the measured Absorbance of the Milli-Q water blank and $L$ is the path length in meters.

Fluorescence was measured on samples, diluted to match the absorbance measurements, if necessary, on a Varian Eclipse spectrofluorometer. Excitation (Ex) was measured from 240 to 450 nm, in 5 nm increments, and emission (Em) was measured from 300 to 600 at 2 nm intervals. Slit widths of 5 nm were used in both Ex and Em modes and scanning speed was set to 9600 nm/min with an integration time of 0.0125 s. Corrections for lamp intensity, detector response, inner-filtering effects, and dilution were applied (see SI). Final values were calibrated in quinine sulfate units (QSU, where 1 QSU = 1 ppb quinine sulfate).

#### FluorMod Design. We took a forward modeling approach wherein probable sources of organic N were measured and modeled and then fit to streamwater samples. Fluorescence was measured on discrete point and nonpoint sources of DOM to the NRB (Figure 1). Details on source acquisition and preparation are in the SI. Two “Reference” streams represented the natural background source of ON in unimpacted streams within the NRB. These streams are classified as “outstanding resource waters (ORW)” by the North Carolina Department of Environment and Natural Resources (NCDENR). “Influent” was raw sewage inflow to WWTF that are permitted discharges into the Neuse River and several of its tributaries. “Effluent” was treated water discharged from those WWTF. “Swine” was surface samples of swine lagoons collected by farmers and sent to the North Carolina Department of Agriculture (NCDA) laboratory for nitrogen analysis. “Poultry” was water-soluble extracts (“leachates”) leached from turkey, hen, and broiler litters. “Street” samples were runoff collected roadside near storm drains in Raleigh, NC, during and after rain events. “Septic” was samples collected from residential community septic discharge ditches around the City of Durham, NC. “Soil” samples were water-soluble extracts prepared in similar fashion to Poultry leachates.

All source samples were scanned for absorbance and fluorescence and processed as described for the stream and river samples. Some soil leachates and all of the swine lagoon and poultry leachates were highly concentrated and required substantial dilution with Milli-Q water (up to 1:1000). These dilution values were recorded and used to correct final values of absorbance and fluorescence. Street runoff and WWTF Effluent contained residual natural DOM signals that may obscure modeling. Mean normalized fluorescence intensity (I) of the Reference stream measurements was thus subtracted from mean-normalized Effluent and Street fluorescence. The residual was rescaled, negative values zeroed, and used as the discrete fluorescence fingerprint of each source (SI Figure S1). A total of 180 source EEMs were modeled with PARAFAC to create FluorMod and resulted in a nine component model that was validated through analysis of leverages (using a criteria < 0.25), analyzing residuals, and conducting split-half analysis and random initialization tests (Figure 2 and Table 1). Split-half validation results are shown in SI Figure S2. The EEMs from each source category were resampled to produce a normal distribution. All EEMs were normalized to total integrated fluorescence prior to PARAFAC modeling.

#### Using PARAFAC in a Mixing Model to Compute Proportional DON Sources. We applied the PARAFAC model to 516 unknown sample EEMs taken from tributaries and the main stem of the Neuse River so that FMax values for the model’s nine components were determined for each sample. Next, a least-squares approach was taken, whereby the 9 components were used in a mixing model, formulated with a multivariate normal distribution following Voli et al.:
Table 1. Excitation (Ex.) and Emission (Em.) Maxima of the Nine Components Validated From a PARAFAC Model fit to 181 EEMs of DON Sources to the Neuse River Basin

<table>
<thead>
<tr>
<th>component</th>
<th>max ex. (nm)</th>
<th>max em. (nm)</th>
<th>probable origin</th>
<th>number of fluoronor matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>245</td>
<td>372</td>
<td>leaf material?</td>
<td>none</td>
</tr>
<tr>
<td>C2</td>
<td>240</td>
<td>430</td>
<td>natural DOM in</td>
<td>2</td>
</tr>
<tr>
<td>C3</td>
<td>275</td>
<td>350</td>
<td>protein (tryptophan)</td>
<td>10</td>
</tr>
<tr>
<td>C4</td>
<td>250</td>
<td>525</td>
<td>soil leachate</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>275</td>
<td>304</td>
<td>protein (tyrosine)</td>
<td>4</td>
</tr>
<tr>
<td>C6</td>
<td>240 (300)</td>
<td>354</td>
<td>urban runoff</td>
<td>none</td>
</tr>
<tr>
<td>C7</td>
<td>350</td>
<td>440</td>
<td>effluent</td>
<td>6</td>
</tr>
<tr>
<td>C8</td>
<td>305</td>
<td>400</td>
<td>microbial activity</td>
<td>1</td>
</tr>
<tr>
<td>C9</td>
<td>250</td>
<td>412</td>
<td>urban runoff</td>
<td>None</td>
</tr>
</tbody>
</table>

*Value in parentheses is secondary maximum. See text for discussion of probable origins.*


multivariate normal distributions of the eight DON sources. Each set of random values were used to solve the following mass balance equation:

\[
z^T = \sum_k (x_k^T \times P_k), 0 \leq P_k \leq 1, \sum_k P_k = 1
\]  

where \( z \) is the stream or river FMax value of the nine components, \( x \) is the DON source FMax values, and \( P \) is the fraction of each \( k \) DON source in an unknown sample. A non-negative least-squares solution to the mass balance eq 3 provided the estimate of the relative proportion of each source to the total DON. We used this approach to limit the uncertainty in the data due infrequent and random sampling inherent to routine monitoring programs, whereas the Monte Carlo approach incorporated uncertainty into FluorMod by random sampling of derived probability distributions of the nine components for each source (cf. ref 37). Once it was determined that the mean percent contributions of each source between each group of trials varied by less than 5%, the number of trials was downscaled from 10001 to 5001 to 51 in the final version of FluorMod to decrease computing time.

FluorMod was validated by through its application to a random subset of the source EEMs not used to build the model (Table 2). Effluent and Swine sources were estimated to be 93% Effluent and 83% Swine, respectively. Street sources were estimated to be 83% Street. The less discrete sources had lower actual percentages predicted by FluorMod. Septic source samples had a mean prediction of 30% Septic, the lowest of all sources, possibly a result of mixing of septic discharges with leachate from soil. Poultry leachate was modeled as 71% Poultry, 10% Reference, the remaining 19% distributed among other categories. This result likely was influenced by materials mixed in with poultry waste. Our model compared well to a PCA-based model of fluorescence used to determine the contribution of wastewater organic matter in streams near Portland, OR, which was able to separate out Effluent from Reference sources with a confidence of 80% (ref 19).

**Nitrogen Analyses.** Particulate nitrogen (PN), which is assumed to be primarily organic nitrogen (PON), was measured on seston collected on GF/F filters using a 2400 Series II PerkinElmer elemental analyzer36 or a CE Elantech 1112 Flash elemental analyzer. Acetanilide (71% C, 5.9% N) was used as a calibration standard for both instruments. Known masses of acetanilide were analyzed to compute a mass-based response for each instrument. The mass of PON for a given seston sample divided by the volume of water filtered for that sample produced the concentration of PON (in micrograms-N per liter). Blank GF/F filters were measured and typically contained no detectable amounts of N.

Dissolved organic nitrogen (DON) was measured on samples by difference using two methods (see SI). One method was based on the difference between total Kjeldahl N (TKN) and the sum of ammonium and PN. The other method was based on the difference between TDN and dissolved inorganic nitrogen (DIN: sum of nitrate, nitrate, and ammonium). TKN and DIN concentration data for samples run by an external laboratory were obtained from the EPA STORET repository; methodologies are also available at the STORET Web site. TKN, TDN, and DIN measurements were made on the same samples that were used for fluorescence.

**Land Cover and Statistical Methods.** Total watershed areas were delineated for each sampling site using discharge accumulation and direction information from the National Hydrography Data Set (http://www.horizon-systems.com/NHDPlus/NHDPlusV1_home.php). Once watershed areas were calculated, the spatial analysis toolbox in ArcGIS software was used to compute percentages of each land cover (LC) type within each sampling site’s watershed, using data obtained from the 2011 National Land Cover Data set (NLCD) (http://www.mrlc.gov/nlcd2011.php). We simplified the NLCD classifications; urban land use percentage was defined as the sum total area of developed land. Agricultural land use was estimated as the total cultivated cropland area.

Table 2. Mean Percent of Source Contributions Predicted by FluorMod on a Random Subset of the Eight DON Sources, N = Number of Samples

<table>
<thead>
<tr>
<th>source</th>
<th>N</th>
<th>reference</th>
<th>effluent</th>
<th>influent</th>
<th>poultry</th>
<th>swine</th>
<th>septic</th>
<th>street</th>
<th>soil</th>
</tr>
</thead>
<tbody>
<tr>
<td>reference</td>
<td>20</td>
<td>75%</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>0%</td>
<td>1%</td>
<td>1%</td>
<td>18%</td>
</tr>
<tr>
<td>effluent</td>
<td>24</td>
<td>0%</td>
<td>93%</td>
<td>1%</td>
<td>0%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>3%</td>
</tr>
<tr>
<td>influent</td>
<td>24</td>
<td>3%</td>
<td>1%</td>
<td>57%</td>
<td>9%</td>
<td>8%</td>
<td>18%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>poultry</td>
<td>15</td>
<td>10%</td>
<td>3%</td>
<td>4%</td>
<td>71%</td>
<td>6%</td>
<td>2%</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>swine</td>
<td>24</td>
<td>1%</td>
<td>0%</td>
<td>5%</td>
<td>8%</td>
<td>83%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>septic</td>
<td>17</td>
<td>10%</td>
<td>3%</td>
<td>22%</td>
<td>3%</td>
<td>24%</td>
<td>30%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>street</td>
<td>20</td>
<td>2%</td>
<td>1%</td>
<td>8%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
<td>83%</td>
<td>2%</td>
</tr>
<tr>
<td>soil</td>
<td>18</td>
<td>14%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>81%</td>
<td></td>
</tr>
</tbody>
</table>
RESULTS AND DISCUSSION

Fluorescence Characteristics of Potential DON Sources in the Neuse River Basin. Fluorescence characteristics were similar among the eight potential sources of DON to the NRB (Figure 1). All EEMs were similar in that most fluorescence occurred at the shorter excitation wavelengths (generally below 350 nm) yet over a wide range of emission (ca. 280–500 nm). Most organic fluorophores share these properties. Effluent and runoff sources were fluvial—not specific leachates such as the Soil and Poultry—and appeared to be mixtures with NOM likely leached from leaves, soils, dust, etc. These materials accumulate on road surfaces and would make their way into storm drains and other stormwater conveyance systems such as Pigeon House Creek, a tributary of Crabtree Creek in Raleigh, NC, where we observed accumulation of leaf material during near-stagnant base flow. In overland flow during rain events, this material could be leached and mobilized as might loose soil on impervious surfaces, essentially creating fluorescence similar to Reference prior to entering a stream as runoff.

Effluent discharged from treatment plants also exhibited a similarity to the Reference stream, possibly as a result of natural streamwater commingling with Influent entering a WWTF. However, the Influent signal was quite distinct from the Reference signal (Figure 1). Therefore, much like the Street runoff source, the Effluent source had the Reference signal removed to establish its unique signal (SI Figure S1).

Human and animal waste sources are enriched in proteinaceous material and the EEMs of these sources were dominated by protein fluorescence, which is comprised mostly of tryptophan (excitation/emission (ex/em) maxima of 280 and 340 nm, respectively) and tyrosine (ex = 275, em = 304 nm). Influent, poultry leachate, swine lagoon, and septic samples all had prominent fluorescence in the region of tryptophan and tyrosine as has been observed for farm wastes and sewage (Figure 1). These signals were distinctive enough to separate sources dominated by them from the Reference. While the Poultry and Soil sources were leachates that shared features similar to the other EEMs, each represented direct leachable inputs of DON to streams during percolation of rainwater through piles of poultry litter or surface soil, respectively. In general, rainwater fluorescence is very low compared to the fluorescence of these extracts.

PARAFAC modeling of these eight sources produced a nine component model (Figure 2). Split-half validation results are in the SI (Figure S2). PARAFAC model results were tested against the OpenFluor database (http://www.openfluor.org) to determine which of FluorMod’s components matched to components of models in the published literature (Table 2). Components 1 (C1), C6, and C9 did not match to any published models in the OpenFluor database. This result was expected for some components because the FluorMod model was built of fluorescence measurements on septic and animal waste sources of DOM and few OpenFluor models were fit to those sources. C2 and C4 resembled aromatic, conjugated macromolecular substances of terrestrial origin. C7 matched with 6 models, all from recycled water studies, which included samples of wastewater, treated water, and gray water. C3 and C5 were similar to tryptophan and tyrosine, respectively. C8 resembled aromatic, conjugated macromolecular substances of microbial origin. Thus, most of the components we modeled were widely distributed in a variety of natural and anthropogenic systems.

Separation of DON Sources Using EEM Fluorescence and PARAFAC Analysis. DON concentrations for sources and stream waters were significantly correlated to corresponding rescaled FMax values for each of the 9 PARAFAC components (SI Table S4). While this demonstrated that fluorescence can be used to assess DON sources, principal components analysis (PCA) conducted on the rescaled FMax values from the PARAFAC model demonstrated our ability to separate the DON sources by their fluorescence properties (Figure 3). Autoscaling was used on FMax values from the original samples, and cross-validation of the model was achieved with a Venetian blinds procedure with 10 data blocks. PC1 and PC2 explained roughly 70% of the variability in the data with most variance explained on PC1, which separated urban runoff sources of DON (Street) from natural sources (Reference, Soil). PC2 explained 27% of the variation and could be attributed to protein content as most waste sources were positive along PC2.

Scores for the “waste” sources (influent, poultry leachate, swine lagoon, and septic) formed a distinct cluster which was clearly separated from WWTF Effluent, as well as from natural reference, soil, and runoff sources. Some septic sample scores fell very near the origin, indicating these samples were the least distinctive of all sources and thus most difficult to separate. The proximity of Poultry sample scores to reference and soil scores likely results from the mixture of soil or wood chips, or straw in...
the bedding material, with poultry waste, which when leached may contribute some signal resembling "natural" DON sources such as Reference or Soil.

Loadings for the DON sources showed enrichment of protein- and microbial fluorescence components C3, C5, and C8 in swine, influent, septic, and to a lesser degree Poultry (Figure 3). The loading for C4 corresponded to the cluster of soil and reference scores and matched with several sources described as soil fulvic acid but also with several models of recycled water plants.27 Loadings for C1, C6, and C9 all clustered together and best match the cluster of Street runoff scores. No variable loading aligned closely with Effluent scores. These results indicated that no one component served as a clear marker for any single DON source.

However, using the nine PARAFAC components found for the eight sources in the mixing model allowed us to estimate the relative amounts of each source in a stream or river sample. The mixing model assumed a stream or river sample's fluorescence was some combination of the eight sources. We anticipated that the major fraction of DON in the stream and river samples would be natural organic matter and thus the results would be dominated by reference and soil sources. Our results clearly showed this was the case and also elucidated the smaller but meaningful contributions of urban and agricultural DON to the river and its estuary (Table 3 and SI Table S5).

### Linking DON Sources to Land Use

Normalization of EEM fluorescence intensities used in the model eliminated its ability to predict DON concentrations from sources ab initio. Furthermore, fluorescence is not solely based on organic N. Despite strong correlation between PARAFAC FMax values and DON concentration (SI Table S4), relative amounts of DON sources were estimated in the NRB by multiplying the fraction of each source modeled by FluorMod by its bulk DON concentration. Multiplying the fractional concentration of each DON source by streamflow estimates its load in a stream. As we anticipated, natural sources of DON were dominant in the NRB because Reference and Soil were nearly always major fractions. Street runoff and Poultry sources represented the dominant anthropogenic fractions of DON in the Neuse proper and its tributaries.

In both tributaries and in the main stem of the Neuse River, reference sources generally were >70% (SI Table S5). Maximum Reference contribution was 91% at Contentnea Creek and minimum Reference contribution was <30% both in Bear Creek and at Ft. Barnwell. No significant correlation between % land use and Reference was found in the data set. The contribution of Effluent signal was generally <5%, though a maximum of 15% was estimated at Middle Creek. Influent also contributed <5% of the DON signal in most water samples modeled with FluorMod. However, maxima of up to 15–30% Influent were measured in the Neuse Basin in tributaries and main-stem locations. Similar to Reference, no significant correlations were found between land use and Effluent or Effluent sources.

Poultry was a more dominant source of agricultural DON than Swine in both the main stem of the Neuse River and its tributaries (Table 3 and SI Table S5). Poultry ranged from 0 to 8% in terms of mean contribution, but the maximum was 22% in Bear Creek, whose watershed was 55% agriculture (SI Table S1 and S5). Virtually no Swine contribution was measured in the streams of the NRB. Mean Septic percentages were <10%, while maximum values were similar to other anthropogenic point and nonpoint sources. Street runoff showed a surprisingly small overall contribution, with a mean <6% and maxima not much higher. However, a clear positive correlation between Watershed Developed cover and Street DON was found (R = 0.79; P = 0.02; N = 8), corresponding to a negative correlation between Watershed Cropland cover and Street DON (R = −0.84; P < 0.001; N = 8). Septic, Influent, and Effluent sources of DON also were correlated with Watershed Developed cover (Table 4). Soil DON also was substantial, with means ranging from 7% in Crabtree Creek to 14% in Bear Creek (SI Table S5). The Neuse River at Kinston and Ft. Barnwell had mean Soil DON of 11%, not quite double the mean contribution at Clayton (7%). At most sites, Soil DON had maxima of at least 20%, making this source second to Reference in terms of being major contributions of DON to the NRB.

Urban sources of DON were at times very large in the Neuse River at Clayton, Kinston, and Ft. Barnwell (SI Table S5). In the tributaries, it was surprising that DON in Bear Creek was as much as 45% urban but that percentage was dominated by a 24% Septic DON signal measured on July 2, 2012. The maxima were very episodic in nature; in other words, no sustained dominant "non-natural" source of DON was apparent.

### Table 3. Percentages of DON in the Neuse River Basin Tributaries and Main-Stem Sites Predicted by FluorMod from August 2011 to May 2013**

<table>
<thead>
<tr>
<th>stream</th>
<th>percentage of DON in grouped sources</th>
<th>yield of DON in grouped sources</th>
<th>total DON load</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>urban mean</td>
<td>max</td>
<td>min</td>
</tr>
<tr>
<td>Crabtree Creek</td>
<td>16%</td>
<td>35%</td>
<td>7%</td>
</tr>
<tr>
<td>Middle Creek</td>
<td>9%</td>
<td>30%</td>
<td>0%</td>
</tr>
<tr>
<td>Little River</td>
<td>4%</td>
<td>10%</td>
<td>1%</td>
</tr>
<tr>
<td>Bear Creek</td>
<td>5%</td>
<td>45%</td>
<td>1%</td>
</tr>
<tr>
<td>Contentnea Creek</td>
<td>3%</td>
<td>20%</td>
<td>0%</td>
</tr>
<tr>
<td>Neuse-Clayton</td>
<td>22%</td>
<td>45%</td>
<td>6%</td>
</tr>
<tr>
<td>Neuse-Kinston</td>
<td>6%</td>
<td>48%</td>
<td>1%</td>
</tr>
<tr>
<td>Neuse-Ft. Barnwell</td>
<td>4%</td>
<td>67%</td>
<td>0%</td>
</tr>
<tr>
<td>Trent River</td>
<td>2%</td>
<td>34%</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Urban = urban sources (sum of EFF + INF + SEP + STR); Agriculture = animal waste sources (sum of PLT + SWI). Gray shaded values indicate model runs in which a particular source was excluded due to land use analysis. Also included are DON yields (kg N/ha/yr) for select tributaries in the NRB as well as the yield at Ft. Barnwell. The Trent River at Trenton is listed last because it flows in to the Neuse River Estuary downstream of Ft. Barnwell. Also included is the total DON load from each watershed. "nd" = not determined. 6Little River near Goldsboro, NC, a different stream than that sampled for the reference source signal.
Land use was a key driver of the DON quality in the NRB as modeled by FluorMod (Figure 4). DON source results estimated for the main stem stations along the Neuse River were separated using PCA, whereby PC1 appeared to separate urban from agricultural land use (Figure 4). Upper Coastal Plain stations at Raleigh and Clayton were influenced mostly by street, septic, influent, and effluent sources, while lower Coastal Plain stations of Kinston, and Ft. Barnwell, all show increased influence of soil and poultry sources. Land use in the lower Coastal Plain changes from a strong urban influence to more agricultural influence (SI Table S1).

Tributary scores also showed separations based on land use (Figure 4). Urbanized Crabtree Creek near Raleigh, NC, had scores along PC1 that was influenced by loadings for Septic, Influent (probably sewage), and Street runoff sources of DON. Variability in Bear Creek was strongly influenced by Soil and Poultry. Contentnea Creek DON was primarily influenced by Reference and Soil sources with episodic influences from Street and Poultry sources. Middle Creek was mainly influenced by the urban sources of DON, similar to Crabtree Creek, but with some episodic influences from Poultry, indicating the importance of the rainfall- and runoff-based events. Middle Creek has Watershed Cropland cover of 20% compared to 2% for Crabtree Creek, which likely explained this difference. Little River showed separation of DON sources between samples collected downstream of a WWTF (LittleDNS), where scores were close to the loading for Effluent DON, in contrast to samples collected upstream of the WWTF (LittleUPS). Thus, land use and land cover also strongly influenced the DON quality modeled by FluorMod in NRB tributaries.

Seasonal and Annual Trends in DON Loads to the Neuse River Estuary. Observations of increased relative importance of organic nitrogen to the total N load in the NRB over the past 20 years were placed in context of DON sources modeled in this study. Loads of DON estimated by source showed that proportionally more urban, agricultural, and soil-derived sources of DON were present in the river at Ft. Barnwell during the study period (Figure 5). From August 2011 to April 2013, increases in the load of urban, agricultural, and soil DON generally increased with discharge. The largest flow during the period was late August 2011, just after Hurricane Irene crossed eastern NC at which time loads of DON were greatest. While the largest overall source of DON to the estuary was Reference (Table 3), it was clear that high flows in the river also translated into higher loads of anthropogenic sources. The increase in the source of DON attributed to soil with discharge suggested some effect of erosion.

Table 4. Pearson’s Correlation Coefficients Calculated between Percent of DON Source Predicted by FluorMod and Watershed Characterization

<table>
<thead>
<tr>
<th></th>
<th>developed</th>
<th>forest</th>
<th>cropland</th>
<th>wetland</th>
</tr>
</thead>
<tbody>
<tr>
<td>reference</td>
<td>0.00</td>
<td>−0.24</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>effluent</td>
<td>0.40</td>
<td>0.48</td>
<td>−0.64</td>
<td>−0.61</td>
</tr>
<tr>
<td>influent</td>
<td>0.40</td>
<td>0.50</td>
<td>−0.59</td>
<td>−0.61</td>
</tr>
<tr>
<td>poultry</td>
<td>−0.62</td>
<td>−0.26</td>
<td>0.61</td>
<td>0.57</td>
</tr>
<tr>
<td>septic</td>
<td>0.57</td>
<td>0.22</td>
<td>−0.61</td>
<td>−0.75</td>
</tr>
<tr>
<td>street</td>
<td>0.80</td>
<td>0.24</td>
<td>−0.84</td>
<td>−0.92</td>
</tr>
<tr>
<td>soil</td>
<td>−0.73</td>
<td>−0.45</td>
<td>0.83</td>
<td>0.76</td>
</tr>
</tbody>
</table>

P < 0.05
P < 0.10

*Data from SI Tables S1 and S5 were used in this analysis. Significant relationships are indicated in italics and bold and are based on significant levels <0.05 and <0.10, respectively.*
Annual DON load estimates for three sites on the Neuse Proper (Clayton, Kinston, and Ft. Barnwell) were computed using the LOADEST hydrologic model. Mean source percentages were then multiplied by the yearly mass flux (kg DON per year) computed by LOADEST to estimate the kg N/yr of each DON source for the Neuse River at Clayton, Kinston, and Ft. Barnwell (Figure 6). As the river’s discharge increased downstream, variations in the mass flux of DON sources were observed. For example, septic mass flux was 35 000 kg DON/yr at Clayton, yet decreased to 8000 kg DON/yr at Ft. Barnwell. Land use transitioned from urban to rural downstream, with a corresponding increase in the Cropland land cover class (SI Table S1). This transition increased the loading of Poultry DON load in the Neuse River from 0 kg N/yr at Clayton, to 28 000 kg N/yr at Kinston, and then to 64 000 kg N/yr at Ft. Barnwell. To put the latter number in perspective, total N load (sum of organic N, nitrate-N, and ammonium-N) in 2012 from operations permitted to discharge into the Neuse River was 245 345 kg N, roughly 25% of which is DON (61 336 kg N). Thus, the poultry DON load at Ft. Barnwell was estimated to be about the same as the permitted point-source load of total ON discharged to the NRB in 2013.

Annual yields of DON sources were computed to enable comparisons across watersheds of various sizes (Table 3 and SI Table S8). Bear Creek had the largest yield of Poultry DON (0.47 kg N/ha/yr) and Watershed Cropland Cover of 55% (SI Tables S1 and S8). As the percentage of agricultural land use (Cropland cover) increased in the NRB, so did the yield of Poultry DON (R = 0.80; P = 0.08; N = 5), though the lack of
significance may be due to the limited number of observations. Yields of urban DON (sum of effluent, influent, septic, and street) were highest in Crabtree Creek watershed (0.61 kg N/ha/yr). Contention Creek had equal yields of urban and agricultural (poultry plus swine) DON (0.04 kg N/ha/yr each).

**Implications for DON Loads to the Neuse River Estuary.** In terms of the Neuse River Basin above Ft. Barnwell, the highest yield of DON by source from this watershed was Reference—about an order of magnitude greater than the next largest source, Soil. This may be considered “natural” sources (but see “Management Implications” below). As far as anthropogenic sources, Poultry was the highest yield (0.06 kg N/ha/yr) from Ft. Barnwell, nearly 3-fold higher than Septic or Street runoff (both 0.02 kg N/ha/yr; SI Table S8).

The Trent River at Trenton drains a watershed having LULC of 4% urban, 39% forest, 25% agriculture, and 32% wetlands based on GIS analysis. The Trent enters the NRE at New Bern, NC, downstream of where this system is gauged. At the gauged site, this system yields about 70% of the Poultry DON yield from Contention Creek and about half the yield at Ft. Barnwell. The total DON load from Trent River at Trenton was similar to Crabtree Creek near Raleigh.

Lebo et al. observed that ON was an increasingly large fraction of the total N load in the NRB and suggested an increase in TKN to the basin on the order of 15% at Streets Ferry Bridge (downstream of Ft. Barnwell). TKN in the NRB was ca. 90% organic-N and 10% NH₄–N. Other work in this region has identified substantial inputs of “new” N from atmosphere, partially driven by ammonium volatilization from animal operations. Our results indicated that the increasing ON is dominated by a natural background fluorescence signature (Reference DON) that matched with DOM produced in forested catchments. The PCA showed this most clearly as sample scores clustered near loadings for Reference and Soil (Figure 4). However, on average, 8% of the DON at Ft. Barnwell was urban plus agricultural sources (Table 3). Base-extracted particulate organic matter fluorescence was measured using a recently developed fluorescence technique and we estimated that 5% of particulate ON was urban plus agricultural sources. Thus, the fractional input of anthropogenic (urban + agriculture) ON at Ft. Barnwell was 13%, a value remarkably similar to an estimate of a relative organic-N increase (13.5%), which we based on 15% relative increase in TKN made by Lebo et al. Thus, our results are in agreement with earlier analysis of organic-N dynamics in the NRB.

**Management Implications.** We hypothesize that the increase in anthropogenic DON could be due to land use practices, making sub-basins of the NRB more conducive to “leaking” organic nitrogen. Adding fertilizer (DIN) to the urbanizing and agriculturalizing NRB over time likely has increased the export of DON through conversions of DIN to DON. Substantial amounts of DOM can be mobilized by rainfall splash during precipitation. In that study, large soil OM particles between 0.05 and 0.5 mm contributed the most to organic matter enrichment. The predominant poultry signal we modeled in this river basin could result from rain events saturating piles of poultry litter near streams. Such a dynamic set of events might also explain the large soil signals found as contributing substantially to nonpoint DON sources. Soil particles exposed by clearing of vegetative cover would similarly be mobilized by rainwater splashes. The effect of humification—the so-called Malliard reaction—which quenches protein and microbial fluorescence signals specific to the animal waste sources could transform a poultry (or swine) signal into a soil signal. Resolving the modification of swine or poultry fluorescence signatures after land application against reference and/or soil fluorescence signatures is a critical information gap in our present model that warrants further study.

Soil DON yield at Ft. Barnwell was three times greater than the Poultry DON yield (SI Table S8), suggesting that soil is potentially an anthropogenically influenced source important for the increasing N load to the Neuse River Estuary. Soil organic matter is leached from soils during rain events and constitutes a large fraction of the organic load in streams. Soil DON in streams could also originate from erosion of soil when land cover (grasses, shrubs, trees) is removed during development in urbanized watersheds. In agricultural watersheds, tiling (especially surface ditching) facilitates soil ON mobilization. Forestry practices too can mobilize soil particles and contribute to water quality impairments. Moreover, ditching during silviculture may lead to new hydrologic connectivity between nonriparian and riparian wetlands, facilitating organic N transport into streams. These possibilities in the NRB and other coastal watersheds subject to similar land use pressures (e.g., the Chesapeake Bay) may be investigated with fluorescence-based models of organic matter quality such as FluorMod.

Application of fluorescence as organic nitrogen (or organic carbon) source markers in other river basins may require local calibrations of fluorescence-based models of organic matter. Thus, FluorMod might not be directly applicable to other basins without some modification. Despite this possible limitation, we demonstrated that predicting sources of DON can be useful in evaluating and comparing nonpoint sources of N to point sources (e.g., from permitted dischargers). Use of fluorescence to model organic matter quality in watersheds is a growing field and studies such as ours have presented new ways in which this technique can be applied to a variety of water quality concerns.

**ASSOCIATED CONTENT**

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.6b00053.

Details of the nitrogen reduction policy for the NRB, sampling location and land cover characteristics for the sites in NRB (Table S1), fluorescence analytical methodology, source sample preparation, isolation of the Street and Effluent source signals (Figure S1), the mixing model (Tables S2–S5), measurement of DON, and nitrogen load calculations (Table S8) (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**
Phone: +1-919-515-0382; closburn@ncsu.edu.

**Present Address**
Cardno, Inc., Raleigh, North Carolina 27612, United States.

**Notes**
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

This project was funded under an EPA Section 319 Grant administered by the North Carolina Department of Environment and Natural Resources (NC DENR Contract #4443), the
NC Dept. of Environment and Natural Resources and Lower Neuse Basin Association (ModMon Project), and the National Science Foundation, Project 1119704. Mr. Haywood Phthisic and the Lower Neuse Basin Association are thanked for providing logistical support and a forum to communicate project results. We appreciate the suggestions of Mr. Jay Sauber, Professor Deanna Osmond, and Ms. Brenda Cleveland for providing locations of reference streams and for providing blind samples of swine lagoons and poultry litter. Technical assistance with water collection, sample processing, and sample analysis was provided by Brittnie Adams, Ruchi Bhattacharya, Adeline Brym, Jennifer Dixon, and Daniel Rojas-Jimenez. Professor Karl Wegmann provided thoughtful and useful discussions and logistical support for soil sample collections. Professors Youhei Yamashita and Rudolf Jaffe and Dr. Kaelin Cawley are thanked for sharing Matlab code we adapted for FluorMod development. Professor Del Bohnenstiehl is acknowledged for assistance with Monte Carlo model development and coding in Matlab. Professor Erin Hestir and three anonymous reviewers provided helpful suggestions that improved this manuscript.

**REFERENCES**


(32) Graeber, D.; Gelbrecht, J.; Pushe, M. T.; Anlanger, C.; von Schiller, D. Agriculture has changed the amount and composition of
(36) Volf, M. T.; Wegmann, K. W.; Bohnenstiehl, D. R.; Leithold, E.; Osburn, C. L.; Polyakov, V. Fingerprinting the sources of suspended sediment delivery to a large municipal drinking water reservoir: Falls Lake, Neuse River, North Carolina, USA. J. Soils Sediments 2013, 13 (10), 1692−1707.