

## Research Article

# Comparison of Two PARAFAC Models of Dissolved Organic Matter Fluorescence for a Mid-Atlantic Forested Watershed in the USA

Shatrughan Singh,<sup>1</sup> Shreeram Inamdar,<sup>2</sup> and Durelle Scott<sup>3</sup>

<sup>1</sup> Department of Geological Sciences, University of Delaware, Newark, DE 19716, USA

<sup>2</sup> Department of Plant and Soil Sciences, University of Delaware, Newark, DE 19716, USA

<sup>3</sup> Biological Systems Engineering, Virginia Tech, Blacksburg, VA 24061, USA

Correspondence should be addressed to Shatrughan Singh; [ssingh.udel@gmail.com](mailto:ssingh.udel@gmail.com)

Received 14 January 2013; Revised 2 June 2013; Accepted 17 June 2013

Academic Editor: Fu-Liu Xu

Copyright © 2013 Shatrughan Singh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The composition of dissolved organic matter (DOM) in a mid-Atlantic forested watershed was evaluated using two fluorescence models—one based on previously validated model (Cory and McKnight, 2005) and the other developed specifically for our study site. DOM samples for the models were collected from multiple watershed sources over a two-year period. The previously validated parallel factor analysis (PARAFAC) model had 13 DOM components whereas our site-specific model yielded six distinct components including two terrestrial humic-like, two microbial-derived humic-like, and two protein-like components. The humic-like components were highest in surficial watershed sources and decreased from soil water to groundwater whereas the protein-like components were highest for groundwater sources. Discriminant analyses indicated that our site-specific model was more sensitive to subtle differences in DOM and the sum of the humic- and protein-like constituents yielded more pronounced differences among watershed sources as opposed to the prevalidated model. Dissolved organic carbon (DOC) and dissolved organic nitrogen (DON) concentrations and selected DOM metrics were also more strongly correlated with the site-specific model components. These results suggest that while the pre-validated model may capture broader trends in DOM composition and allow comparisons with other study sites, a site-specific model will be more sensitive for characterizing within-site differences in DOM.

## 1. Introduction

Dissolved organic matter (typically  $<0.45\ \mu\text{m}$ ) is a heterogeneous mixture of aromatic and aliphatic organic compounds, ranging from proteins, carbohydrates, polysaccharides, and lipids to humic and fulvic acids [1]. In terrestrial and aquatic ecosystems, DOM does not only influence the geochemical and photochemical reactions by participating in carbon (C) and nutrient (N, P, and S) cycles, but also control microbially mediated reactions by serving as potential substrate. It also plays a key role in transport and transformation of major contaminants and/or pollutants and their reactivity with the environment [2, 3]. Additionally, DOM exerts a strong control in the formation of disinfection byproducts (DBPs), for example, trihalomethanes (THMs) and haloacetic acids

(HAAs), during the drinking water supply treatments with disinfectants [4]. The amount and quality of DOM in terrestrial and aquatic environments influence biological processes such as microbial degradation [5, 6]. It is also a key player in altering the depth of the photic zone in aquatic ecosystems by controlling the incident UV radiation [7]. Thus, DOM plays an ecologically important role in various biochemical and physical processes linking terrestrial and aquatic ecosystems.

In the last two decades, the availability of optical measurement techniques such as ultraviolet (UV) absorption [8, 9] and fluorescence spectroscopy [10, 11] has provided important insights into the character and composition of DOM. Some of the optical indices that have been implemented to characterize DOM quality include specific ultraviolet absorbance (SUVA; [9]), absorption coefficient [12, 13], spectral slope

ratio ( $S_R$ ; [14, 15]), humification index (HIX; [16]), fluorescence index (FI; [17]), and %protein-like fluorescence [18] derived from absorption measurements and fluorescence-based excitation emission matrices (EEMs, [11, 17]). While UV absorption and fluorescence spectroscopy have provided important insights into DOM composition, these analyses also yield large amounts of data with high dimensionality and nonlinearity. To process this data and gain meaningful insights, a variety of multivariate statistical tools have been employed [19]. Boehme et al. (2004) examined the DOM fluorescence variability using EEMs coupled with principal component analysis (PCA) [20]. Likewise, parallel factor analysis in conjunction with fluorescence EEMs has been used to identify biogeochemically meaningful components of DOM [19]. One fluorescent model adapted by the water science community is the Cory and McKnight model [17], which was developed using DOM from a wide range of aquatic environments.

Although the benefits of PARAFAC models to extract biogeochemical information from EEMs are indisputable, there is some uncertainty on whether a new, site-specific PARAFAC model should be developed or whether a previously validated model [17] is adequate to characterize the variability of DOM at a given site. Fellman et al. (2009) recently addressed this issue for a study site in Alaska where they compared a new, site-specific PARAFAC model against the Cory and McKnight (2005) model (hereafter referred as CM model) for EEMs derived from soil and stream waters [17, 21]. While they did not find any significant differences between the two PARAFAC models, they did observe that the site-specific model was more sensitive to particular DOM constituents. They found that the Cory and McKnight model [17] was unable to characterize a humic-like component that was specific to soil waters and that the site-specific model did a better job in characterizing the variability of the protein-like region of DOM. Similarly, Larsen et al. (2010) reported a limited resolving power of the Cory and McKnight (2005) model in characterizing protein-like and phenolic DOM fractions [17, 22].

We recently characterized the composition of DOM for multiple watershed sources using UV absorption and fluorescence techniques in a 12 ha forested watershed located in the Piedmont region of Maryland [23, 24]. Watershed DOM sources that were studied included throughfall, litter leachate, soil pore water, wetland soil water, hyporheic water, stream runoff, groundwater seeps, and riparian, shallow, and deep groundwater. EEMs for DOM samples from these watershed sources were fitted to the previously validated Cory and McKnight [17] PARAFAC model. In addition, multiple spectrofluorometric indices were also determined for the DOM samples which included absorption coefficient at 254 nm ( $a_{254}$ ), SUVA at 254 nm ( $SUVA_{254}$ ), HIX, FI, and  $S_R$ . These indices revealed distinct patterns in DOM composition across watershed sources [23]. The surficial watershed sources were rich in humic and aromatic DOM constituents while DOM in groundwater or along deeper flow paths was low in humic-like and high in protein-like DOM [23].

Using the same 714 EEMs collected from our field site, we developed a new, site-specific Fair Hill PARAFAC model

(hereafter referred to as the FH model). Our overall objective here was to investigate how this new site-specific FH model compared with the CM model in characterizing the DOM composition at our site. Specific questions that we addressed were (a) How does the new site-specific FH model characterize the DOM composition? Is the site-specific FH model more sensitive than the CM model for characterizing DOM and does it allow for a greater differentiation among watershed sources? If yes, which model components and watershed sources reveal the largest differences between the FH and CM models? The unique aspect of this study is the availability of strong dataset on multiple and distinct watershed DOM sources to assess the differences between the two PARAFAC models.

## 2. Site Description and Methods

**2.1. Site Description.** A detailed description of the study site and sampling procedures has been previously reported [23, 24]. Briefly, the study watershed (12 ha) is located within the Fair Hill Natural Resources Management Area (NRMA) (39°42'N, 75°50'W) in Cecil County, MD (Figure 1), and is part of the Big Elk Creek drainage basin which lies within the Piedmont physiographic region. Big Elk Creek eventually drains into the Chesapeake Bay.

Cecil County has a humid, continental climate with well-defined seasons and mean annual rainfall of 1221 mm [25]. The study watershed predominantly is forest cover with deciduous canopy dominated by *Fagus grandifolia* (American beech), *Liriodendron tulipifera* (yellow poplar), and *Acer rubrum* (red maple) species [26]. The study watershed is primarily underlain by Wissahickon formation comprised of metamorphosed crystalline, sedimentary, and igneous rocks including mica-rich schist, amphibolites, and gneiss. The soils are coarse loamy, mixed, mesic Lithic Dystrudepts, Oxyaquic Dystrudepts, with subhorizons indicating seasonal water saturation.

**2.2. Watershed Sampling.** Watershed sampling was performed at baseflow (1–3 times a month) as well as during storm events. Manual grab sampling during baseflow over a two-year period (2008–09) was conducted for multiple watershed locations which included stream (ST), seeps (P), hyporheic zone (HY), wetland soil water (WSW), shallow (SGW), and deep (DGW) groundwater, soil pore water (U), and riparian groundwater (RGW). Storm samples were collected using automated ISCO sampler (Teledyne Isco Inc., Lincoln, NE, USA) during the events for stream water (at the 12 ha watershed outlet), forest floor (or litter leachate, LT), throughfall (TF), and rainfall (R) (Figure 1). All water samples were filtered through a 0.45 micron nitrocellulose membrane filter (Millipore Corp., Billerica, MA, USA) within 24 hours of collection and stored at 4°C for further analyses. Dissolved organic carbon analysis for the water samples was conducted at the biogeochemistry laboratory of SUNY-ESF, NY, using the Tekmar-Dohrmann Phoenix 8000 TOC analyzer. Nitrate-N was determined using a Dionex IC,  $NH_4^+$  with an autoanalyzer using the Berthelot Reaction followed by colorimetric analysis, and total dissolved nitrogen (TDN)

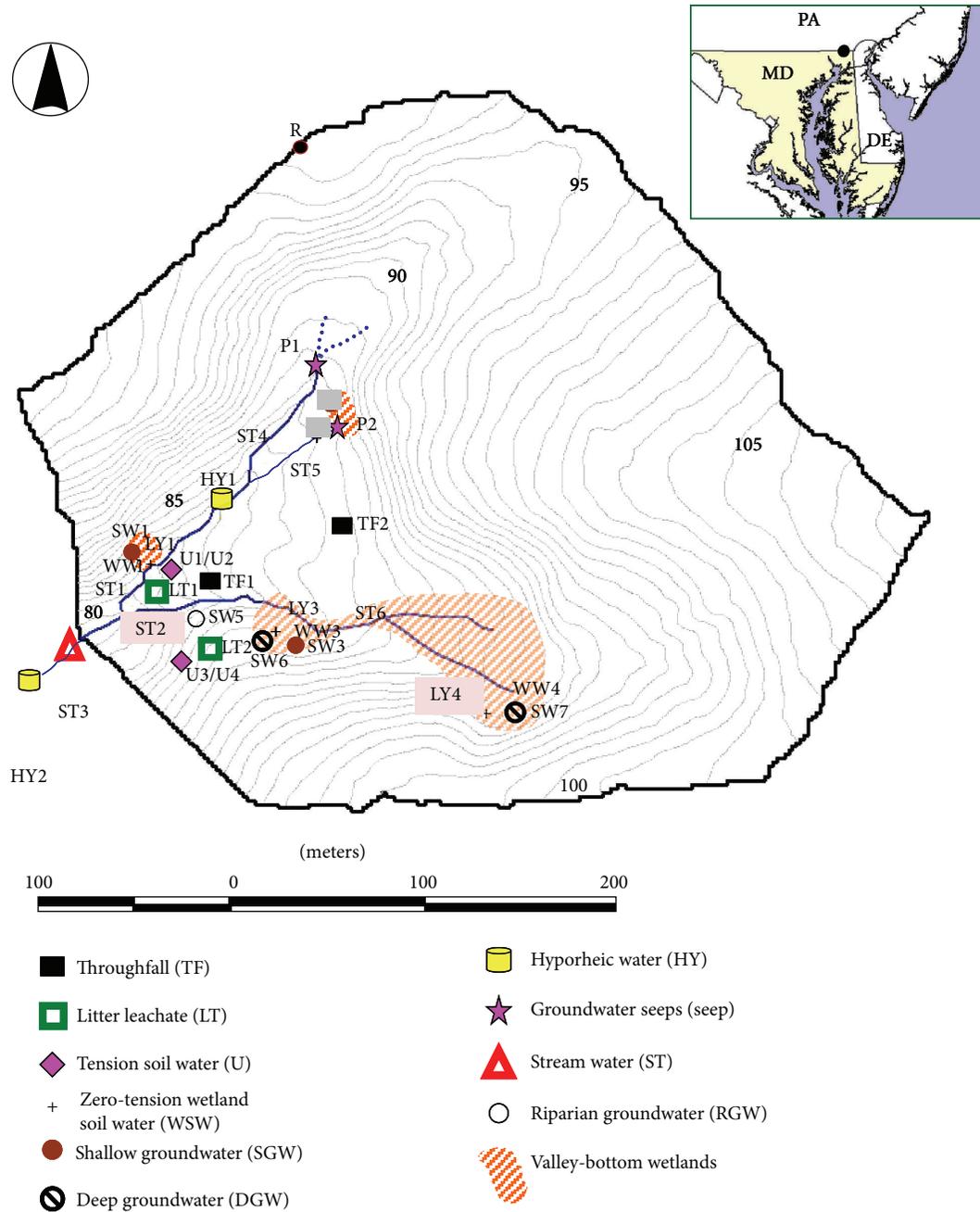


FIGURE 1: Location of the study site in the mid-Atlantic Piedmont region of northeastern Maryland (inset: location of catchment indicated by filled circle) and sampling locations are marked within the 12 ha forested watershed.

using the persulfate oxidation procedure [27] followed by colorimetric analysis on an autoanalyzer. Dissolved organic nitrogen concentrations were computed as the difference between TDN and inorganic N ( $\text{NO}_3^-$ ,  $\text{NH}_4^+$ ).

**2.3. UV-Vis and Fluorescence Spectroscopy.** Absorption spectra (190–1100 nm) were obtained for each sample at room temperature at 1 nm intervals using a UVmini-1240 (Shimadzu Scientific Instruments, Columbia, MD, USA) single-beam spectrophotometer equipped with a 1 cm path-length

quartz cuvette (volume of  $\sim 4$  mL). The instrument was set up and corrected for scattering and baseline fluctuations after running particle-free Nanopure Milli-Q (18.2 M $\Omega$ ) water on daily basis prior to running water samples.

Water samples for fluorescence analysis were treated in a similar manner as for absorption measurements. To account for the inner filter effects (IFE), samples reflecting absorbance greater than 0.2 at 254 nm ( $A_{254} \geq 0.2$ ) were diluted with particle-free Nanopure Milli-Q water (also used as a blank). Excitation-emission matrices (EEMs) were

generated using a Horiba Jobin Yvon Fluoromax-3P (Horiba Scientific, Edison, NJ, USA) spectrofluorometer equipped with a 150 W ozone-free xenon arc lamp. The spectrofluorometer was set to collect the signal in ratio mode (S/R mode) with dark offsets using a 5 nm bandpass on the excitation as well as emission monochromators. Factory-supplied correction factors were applied to the scans to correct for instrument configuration. The EEM spectra were recorded for excitation spectra from 240 to 450 nm at every 10 nm intervals while the emission spectra ranged between 300 and 550 nm, with data saved for every 2 nm over an integration time of 0.25 s. Absorption corrections were applied to account for inner filter effects in “Blank” and sample EEMs. Then, corrected Milli-Q water (Blank) EEMs were subtracted from the sample EEMs to eliminate any influence of Raman peaks. Subsequently, EEMs were normalized to daily determined water Raman integrated area under maximum fluorescence intensity (350 ex/397 em, 5 nm bandpass) as suggested by Lawaetz and Stedmon [28]. Using this approach, EEMs data were normalized and reported in Raman Units (R.U.) which in turn are quantitatively independent from any instrumental parameters provided spectrally corrected data used. Finally, the EEMs were multiplied with dilution factor (if samples were diluted) to obtain the fluorescence intensity for the original undiluted sample [5]. Consequently, the corrected EEMs were exported in MATLAB 7.12 (MathWorks Inc., Natick, MA, USA) for premodel run steps.

**2.4. DOM Characterization Using FH and CM PARAFAC Models.** Following EEMs exported in MATLAB, a premodel step involves a mean centering across the samples to reduce any offsets [29]. Then, each EEM scan was normalized to 1.0 by dividing the whole EEM by the maximum recorded fluorescence intensity value for the sample to ensure that no samples dominated the PARAFAC analysis [30]. Following the procedures of EEM corrections and normalizations, data were fit to a 13-component PARAFAC model developed previously using 379 samples across the wide range of aquatic environments [17]. Residual peaks were less than 10% after fitting EEMs to the 13-component model, confirming that EEMs obtained in this study were well fit to the CM model [31].

The site-specific FH model was developed using the DOMFluor toolbox (ver. 1.7; Feb. 2009) developed for MATLAB by Colin Stedmon (NERI, Aarhus University, Denmark). Based on the Stedmon and Bro [32] study, PARAFAC constraints, such as nonnegativity, and model initialization values derived from singular value decomposition (SVD) were used. The PARAFAC model development was initiated using an EEMs dataset of 747 samples with 121 emission and 22 excitation wavelengths without any assumptions on the number of components (or distinct fluorophores), shape of the resulting spectra, or structure of noise [32]. The number of components (i.e., model validation) was achieved by split-half analysis and by the visual analysis of residuals and corresponding component loadings [19, 32]. Following this methodology, six components were identified for the dataset with some unexplained variability (less than 10%)

remaining in the residuals (Figure 2). A seven- (and eight-) component PARAFAC model was rejected as they could not be validated using split-half and random initialization techniques suggesting that the selection of the six-component model in this study is justified (Figure 2). High scattering Raman and Rayleigh bands were set as missing data and subsequently were removed to avoid any influence by these values in the final dataset according to reported studies [19, 32]. PARAFAC component scores as  $F_{\max}$  are reported for each water sample collected in this EEM dataset. We applied leverage and loading techniques [32] to identify the outliers before a final PARAFAC model. After close inspection of outliers in samples and in excitation-emission loadings, 33 samples, one emission wavelength (300 nm), and the first three excitation wavelengths (240, 250, and 260 nm) were discarded. At the end, the final PARAFAC model was developed on the EEM dataset containing 714 samples, 62 emissions, and 19 excitation wavelengths to derive six validated individual components (Figure 2). While the PARAFAC model was built using 714 samples collected for stormflow and baseflow over a 2-year period, here we present results from 367 samples that include the baseflow sampling and storm-event samples for litter leachate and throughfall.

Similar to the CM model, the % contributions of the individual PARAFAC components for the FH model were determined by dividing each component  $F_{\max}$  score by the sum of the total fluorescence intensity (sum of  $F_{\max}$  scores of all components). To supplement our assessment of the PARAFAC components, we also include the concentrations for DOC and DON and the values for selected DOM metrics such as  $a_{254}$ , HIX, and FI. The  $a_{254}$  values were calculated from the absorbance values using the equation given by Green and Blough [33]. Humification index was calculated according to Ohno [16] and provides a degree of humification in DOM samples (values ranging from 0 to 1). Fluorescence index (FI) was calculated using a ratio of fluorescence emission intensities computed at 470 and 520 nm with excitation intensity at 370 nm [17]. This index has been used to differentiate between terrestrial (FI: 1.2 to 1.5) and microbial sources (FI: 1.6–2.0) [11].

**2.5. Comparison of the Two PARAFAC Models.** The two PARAFAC models were compared using multiple approaches. In the first approach, we compared the sum of the two major fluorescing groups %humic-like and %protein-like DOM from the two models for the watersheds sources. An ANOVA analysis was performed to determine the degree of differentiation among the watershed sources based on data from the PARAFAC models. Additionally, a discriminant function analysis was performed to identify whether watershed sources could be differentiated by their PARAFAC components. Sampling locations were selected as dependent variables whereas humic-like and protein-like DOM compositions were selected as independent variables for the discriminant analysis. Wilks' lambda distribution was used for the forward stepwise selection of independent variables. Finally, we investigated the correlations (Pearson) between the sum of %humic- and %protein-like components against

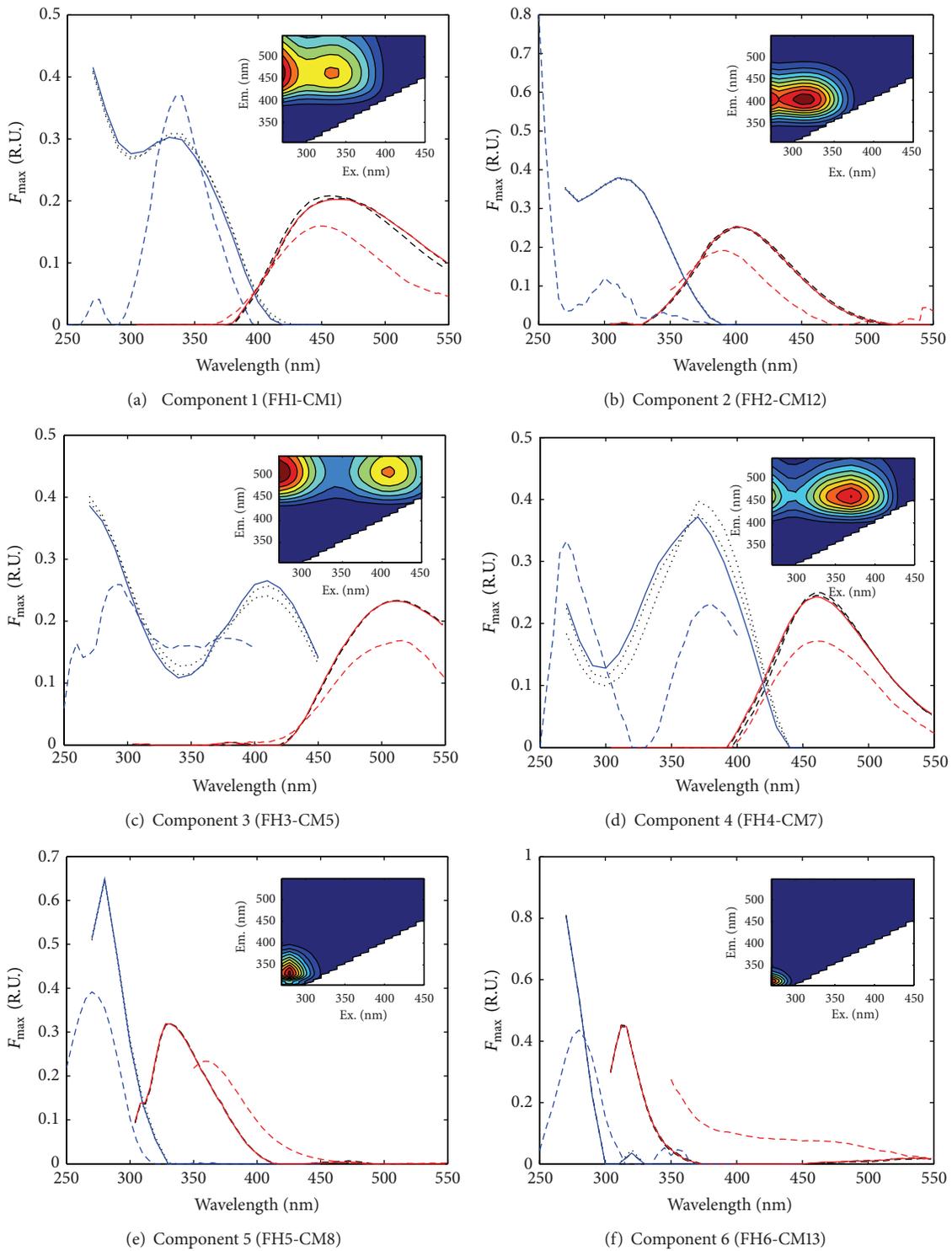


FIGURE 2: Spectral positions of excitation and emission loading maxima derived from the six-component PARAFAC model using split-half validation technique are given in Table 1. Solid blue and red lines represent excitation and emission loadings for site-specific FH model. Dotted and dashed black lines represent excitation and emission loadings, respectively, for two random halves of the complete dataset used for model validation. Broken blue and red lines represent excitation and emission loadings for CM model. EEM contour plots of six different fluorescent components identified by the site-specific PARAFAC (FH) model are shown in corresponding insets.

the DOC and DON concentrations and the DOM metrics of  $a_{254}$ , HIX, and FI. The intent here was to determine if the strength of the correlations differed among the two models and how they varied for watershed sources. All statistical analysis was performed with MATLAB 7.12 and a JMP 9.0 statistical software package (SAS Institute Inc., Cary, NC, USA).

### 3. Results

**3.1. Components for the Fair Hill PARAFAC Model.** A total of six components were identified by the FH PARAFAC model. Component 1 (FH1) was comprised of bimodal excitation maximum peaks at 250 and 330 nm with an emission maximum peak at 460 nm (Figure 2(a) and Table 1). This component is identified as humic-like in character and resembles a mixture of A and C peaks described by Coble [34]. The spectral characteristic of this component is close to the “unknown” component 1 identified by the CM model, despite having a red-shift of 10 nm in the emission maximum obtained in our study. High concentrations of this component (similar to FH1 component) have been reported by Stedmon et al. (2003) in forest streams and wetlands [19]. Likewise, the abundance of this component was reported for wetlands and forest streams by Fellman et al. (2010) [35].

Component 2 (FH2) showed two peaks of excitation maxima at 250 and 310 nm having an emission maximum at 400 nm (Figure 2(b) and Table 1). This component is similar in spectral features to component 12 (Q3) of the CM model, which has a fluorescence comparable to oxidized quinone-like DOM moieties. Previous studies have attributed the origin of this component to the presence of biologically labile organic matter primarily rich in aliphatic carbon content (e.g., Table 1, [3, 35, 36]). Cory and McKnight (2005) suggested that the microbial origin of this component existed typically in oxidized environments whereas Yamashita et al. (2010) attributed the microbial origin to autochthonous production of biologically labile organic matter (e.g., Table 1, [17, 37]).

Component 3 (FH3) is characterized by two excitation maximum peaks at 250 and 410 nm, respectively, with a well-defined emission peak at 512 nm (Figure 2(c) and Table 1). This is similar to the semiquinone (SQ1) component identified in the CM model. Yamashita and Jaffé (2008) attributed the abundance of this component (C2, Table 1) to terrestrial organic matter rich in humic content, thus, indicating its origin from higher vascular plants [3]. Component 4 (FH4; Figure 2(d)) had an emission maximum at 460 nm with a bimodality in excitation wavelengths at 250 and 370 nm. This component is similar to “C” peak in spectral characteristics mainly reported for organic matter of terrestrial origins [34]. Cory and McKnight [17] reported this component as semiquinone (SQ2) and attributed its origin to microbial activity in reducing conditions (Table 1). This DOM fraction has also been reported as DOM of high molecular weight and high aromaticity derived mainly from terrestrial inputs [35].

Component 5 (FH5) and component 6 (FH6) possessed single excitation/emission peaks at 280/328 nm and 270/312 nm, respectively (Figures 2(e) and 2(f); Table 1). These two components are comparable to C8 and C13 as reported in the CM model and are attributed to microbial origin. Components C7 and C8 of the Alaska model of Fellman et al. (2010) are similar to FH5 and FH6, respectively, and have been reported as “free” tryptophan- and tyrosine-like DOM moieties [35]. These protein-like components could together be an indicator of DOM lability and bacterial production in the watershed [39].

**3.2. FH PARAFAC Components for Watershed Sources.** The distribution of FH PARAFAC components for DOM from various watershed sources at our site is illustrated in Figures 3(a)–3(f). Median value of component FH1 was highest (0.42 R.U.) for litter leachate followed by wetland soil water (0.37 R.U.) and throughfall (0.36 R.U.) (Figure 3(a)). Riparian water (0.19 R.U.), seep (0.21 R.U.), and deep groundwater (0.14 R.U.) samples recorded the lowest median values for component FH1. The largest variability in FH1 values was observed for riparian, seep, and deep groundwater sources. Overall, FH1 displayed a decreasing trend from surface to subsurface watershed compartments (Figure 3(a)). Median values for component FH2 were highest in soil pore waters and seep (both, 0.30 R.U.), followed by hyporheic (0.29 R.U.) and stream (0.28 R.U.) water samples. Wetland soil water and shallow and riparian groundwater (all, 0.25 R.U.) recorded intermediate median values for the component FH2 (Figure 3(b)). The lowest median value for FH2 was noted for deep groundwater (0.22 R.U.). Again, similar to FH1, a large variability in FH2 values was observed for riparian and seep water samples.

For the soil-derived humic-like component (FH3), median values of  $F_{\max}$  were highest in the wetland soil water (0.23 R.U.) and litter leachate (0.22 R.U.) and lowest in deep groundwater (0.07 R.U.) (Figure 3(c)). Similar to the trend for FH1, the FH3 values for tension soil pore water (0.16 R.U.) were much lower than the zero-tension wetland soil water. Overall, FH3 displayed a trend similar to FH1 indicating a decrease in humic-like DOM from surficial watershed sources to groundwater DOM sources. Median value for visible humic-like component (FH4) was higher in throughfall (0.14 R.U.) compared to litter leachate (0.11 R.U.; Figure 3(d)). Thereafter, median values for FH4 were highest for wetland soil water and then decreased for groundwater sources.

The two protein-like components (FH5 and FH6) revealed similar trends across the watershed sources (Figures 3(e) and 3(f)). Median  $F_{\max}$  values for FH5 and FH6 were lowest for litter leachate and highest for the groundwater sources including riparian, seep, and deep groundwater. Median values for throughfall for both components were greater than litter indicating that throughfall was enriched in protein-like DOM. Compared to FH6, FH5 displayed a higher variability for groundwater sources that included riparian, seep, and deep groundwater. In contrast, FH6

TABLE 1: Descriptions of the six PARAFAC identified components for the EEM dataset of Fair Hill site and their comparison with previously reported studies.

Components (Ex/Em in nm)	Designated peaks/components in previous studies*	Descriptions	Sources/environments reported in previous studies
FH1 (<250 [330]/460)	C2 <sup>1</sup> , A peak <sup>3</sup> , C1 <sup>4</sup> , C1 <sup>5</sup> , C1 <sup>6</sup> , C1 <sup>7</sup> , C3 <sup>8</sup> , C1 <sup>9</sup>	Humic-like/unknown/UVC-humic-like/fulvic-like	Both terrestrial and marine [34]; both oxidized and reduced [17]; river, estuarine, and near-shore marine ecosystems; common to wetlands/forest streams [35]; terrestrial humic origin [36]; terrestrially derived OM, highest concentration in forest streams and wetlands [19]; observed in bay coastal areas in surface waters, low at the bottom [3]
FH2 (<250 [310]/400)	C6 <sup>1</sup> , Q3 <sup>4</sup> , C4 <sup>5</sup> , C3 <sup>6</sup> , C4 <sup>7</sup> , C4 <sup>8</sup> , C3 <sup>9</sup>	Unknown-nonhumic/quinone-like (lawsone)/humic-like	Microbial origin, aliphatic C; oxidized environment [17]; river, estuarine, and near-shore marine ecosystems; aliphatic C [35]; labile matter; biological production in the water column [36]; combination of N and T peaks, autochthonous and biologically labile component [3]; microbial origin [37]
FH3 (<250 [410]/512)	C5 <sup>1</sup> , D region <sup>2</sup> , SQ1 <sup>4</sup> , C3 <sup>5</sup> , C4 <sup>6</sup> , C2 <sup>8</sup> , C2 <sup>9</sup>	Soil-fulvic-like/semi-quinone-like/humic-like	Black Sea [10]; aromatic C; reducing conditions; higher plant matter [17]; river, estuarine, and near-shore marine ecosystems; higher aromatic content [35]; derived from agricultural catchments and exists in fresh water environments [36]; terrestrial origin, biogeochemical processing of terrestrial POM [3]
FH4 (<250 [370]/460)	C1 <sup>1</sup> , C peak <sup>3</sup> , SQ2 <sup>4</sup> , C2 <sup>5</sup> , C2 <sup>6</sup> , C3 <sup>7</sup> , C5 <sup>8</sup> , C4 <sup>9</sup>	Humic-like/semi-quinone-like/Humic-like	Both terrestrial and marine [34]; microbial origin; reducing conditions [17]; river, estuarine, and near-shore marine ecosystems; high molecular weight/aromatic [35]; terrestrial humic origin, related to terrestrial particulate organic matter [36]; high molecular weight fraction of terrestrially derived humic matter, common to wide range of fresh water environments [19]; terrestrial origin [3]; microbial humic [37]
FH5 (280/328)	C7 <sup>1</sup> , T peak <sup>3</sup> , C8 <sup>4</sup> , C7 <sup>5</sup> , C5 <sup>7</sup> , C7 <sup>8</sup> , C5 <sup>9</sup>	Protein-like (tryptophan)/UVB protein-like (tryptophan)	Both terrestrial and marine [34]; microbial origin [17]; river, estuarine and near-shore marine ecosystems; free-tryptophan-like [35]; combination of N and T peaks, labile matter, autochthonous DOM [19]; tryptophan-like, blue shifted from authentic tryptophan, autochthonous, biologically labile [3]; freshly produced tryptophan-like DOM [37]
FH6 (270/312)	B peak <sup>3</sup> , C13 <sup>4</sup> , C8 <sup>5</sup> , C8 <sup>8</sup>	Protein-like (tyrosine)	Both terrestrial and marine [34]; River, estuarine, and near-shore marine ecosystems; free-tyrosine-like [35]; tyrosine like, degradation processes important for the dynamics of this component [3]

\* [38]<sup>1</sup>, [10]<sup>2</sup>, [34]<sup>3</sup>, [17]<sup>4</sup>, [35]<sup>5</sup>, [36]<sup>6</sup>, [19]<sup>7</sup>, [3]<sup>8</sup>, [37]<sup>9</sup>.

revealed a larger range of variability for stream and shallow groundwater samples.

### 3.3. DOM Concentrations and Metrics for Watershed Sources.

DOC and DON concentrations for the entire sample set and values for selected DOM metrics ( $a_{254}$ , HIX, and FI) are presented in Figure 4. Both DOC and DON concentrations were the highest in surficial sources, especially litter leachate (Figures 4(a) and 4(b), resp.), and declined rapidly from soil water to groundwater sources. The lowest DOC concentration was recorded for groundwater seeps (Figure 4(a)). The values for  $a_{254}$  and HIX, which provide a measure of DOM aromaticity and humic content, respectively, were also

highest for surficial watershed sources and lowest in deep groundwater (Figures 4(c) and 4(d), resp.). In contrast, FI, which has been used to indicate DOM of microbial origins [11], revealed an opposite trend with low values for surficial sources (except throughfall) and high values for groundwater sources (Figure 4(e)).

### 3.4. Comparison of the FH and CM Models for Humic- and Protein-Like Components.

The comparison of the sums of the humic- and protein-like DOM components for the FH and CM models is presented in Figure 5. Summation of the humic-like components for the FH model included components FH1, FH3, and FH4 while for the CM model

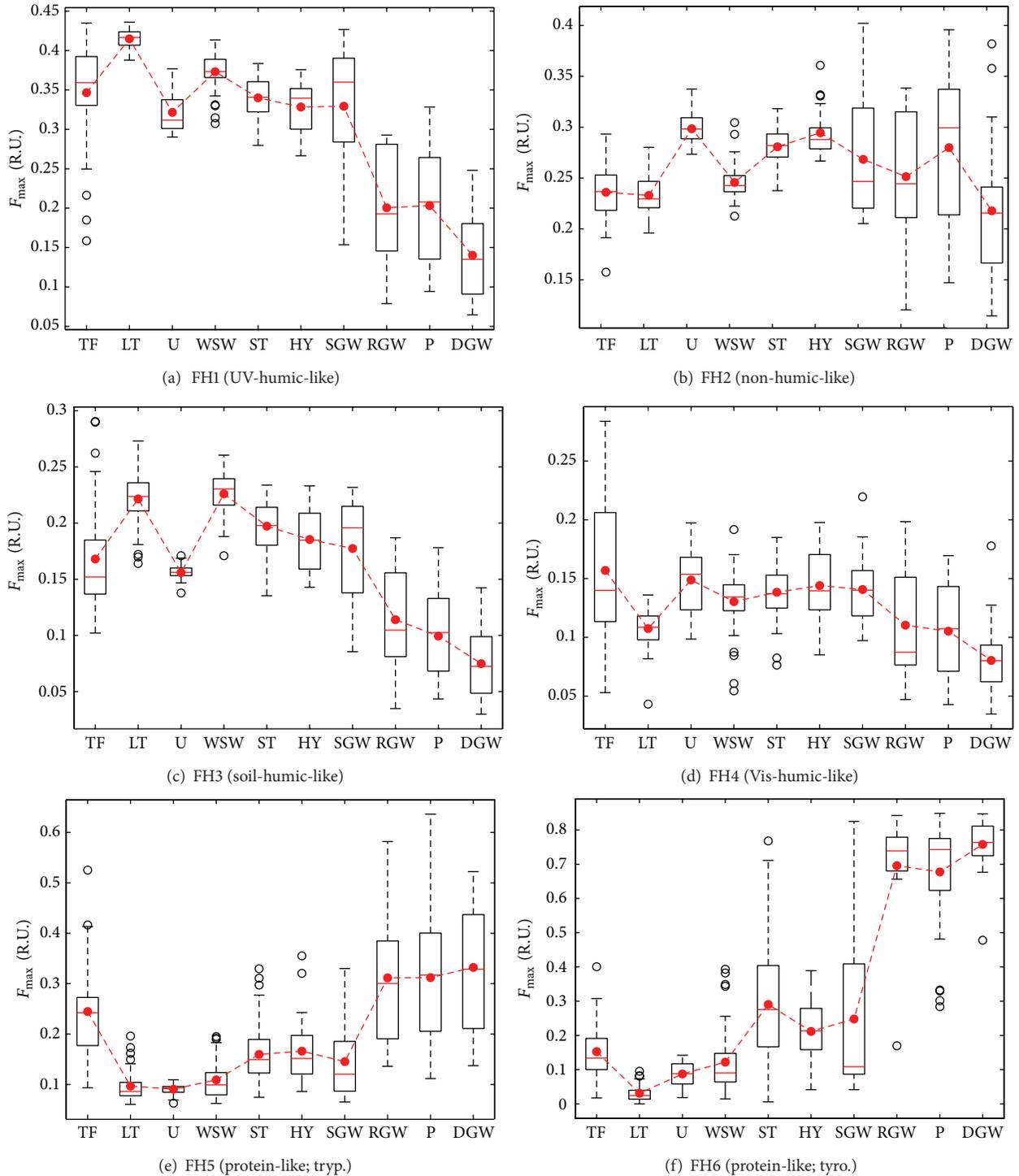


FIGURE 3: Spatial distributions of each PARAFAC component (a) FH1, (b) FH2, (c) FH3, (d) FH4, (e) FH5, and (f) FH6 over the two-year study period (2008-2009) across watershed compartments. Red lines in the box-plot represent median values and red-filled circles represent mean values connected with broken red lines. Empty circles represent outlier data points.

they included components C1, C5, and C7. Similarly, the sum of components FH5 and FH6 represented the protein-like DOM for the FH model while comparable components from the CM model included C8 and C13. Overall, compared to the CM model, the FH model generated a

greater contrast in humic-like and protein-like components among the watershed sources (Figure 5). The FH model also displayed greater variability in humic- and protein-like DOM for individual watershed sources (as indicated by the error bars in Figure 5) as opposed to the CM model. In addition,

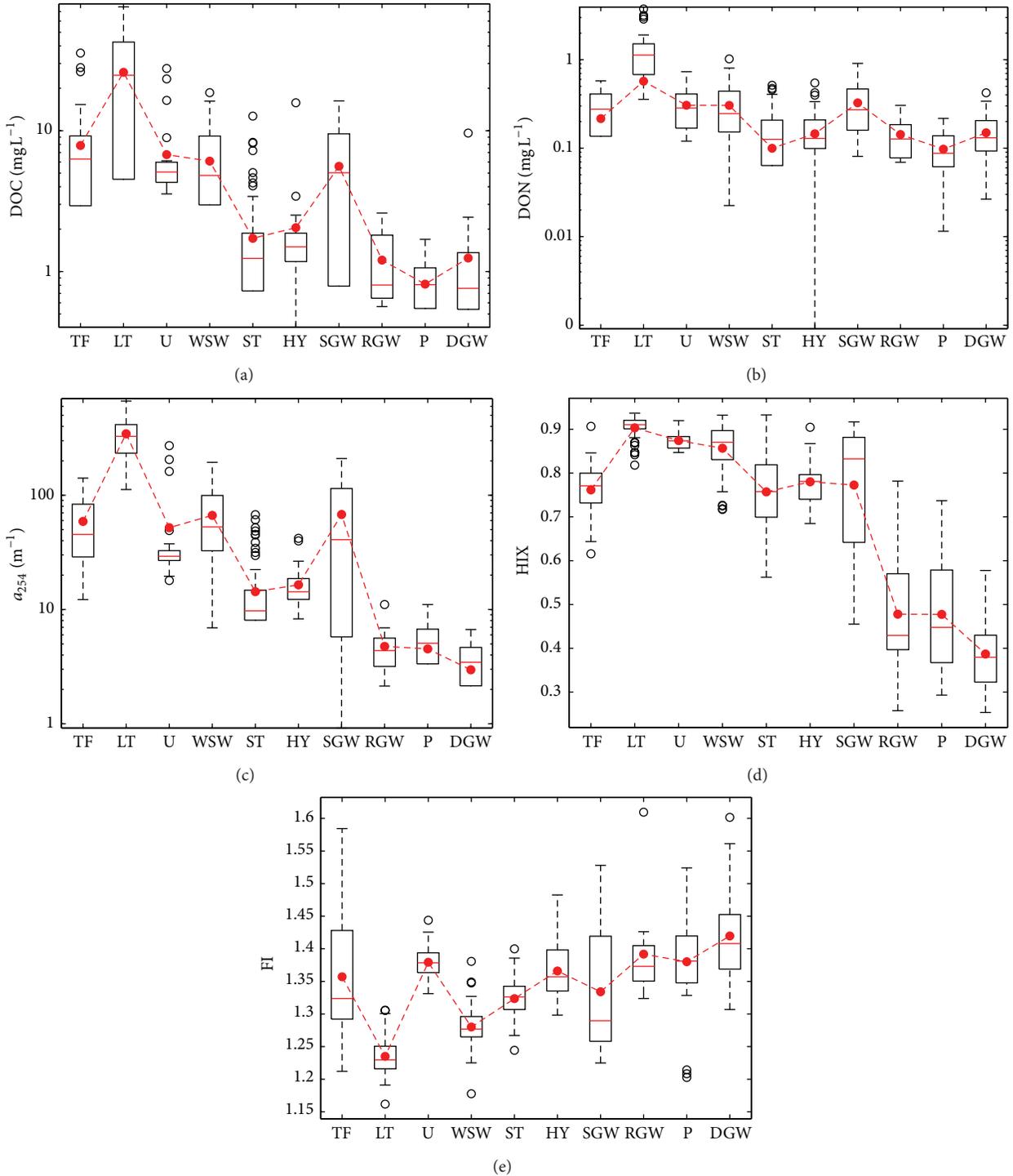


FIGURE 4: Spatial distributions of DOM quantity and quality parameters (a) DOC, (b) DON, (c)  $a_{254}$ , (d) HIX, and (e) FI over the two-year study period (2008-2009) across watershed compartments. Red lines in the box-plot represent median values and red-filled circles represent mean values connected with broken red lines. Empty circles represent outlier data points. (a), (b), and (c) are *log-scaled data* for y-axis to capture distinctiveness in data.

the % values for the humic-like DOM fractions in surficial sources were much greater for the FH model versus the CM model. Similarly, the %protein-like DOM values for groundwater sources were considerably greater for the FH model compared to the CM model. For the FH model, the relative

contribution of %humic-like DOM fluorescence ranged from 18% in deep groundwater to 67% in litter leachate samples, reflecting an almost fourfold increase of humic-like DOM for surficial sources (Figure 5(a)). In contrast, for the CM model, the highest %humic-like contribution was only 26%

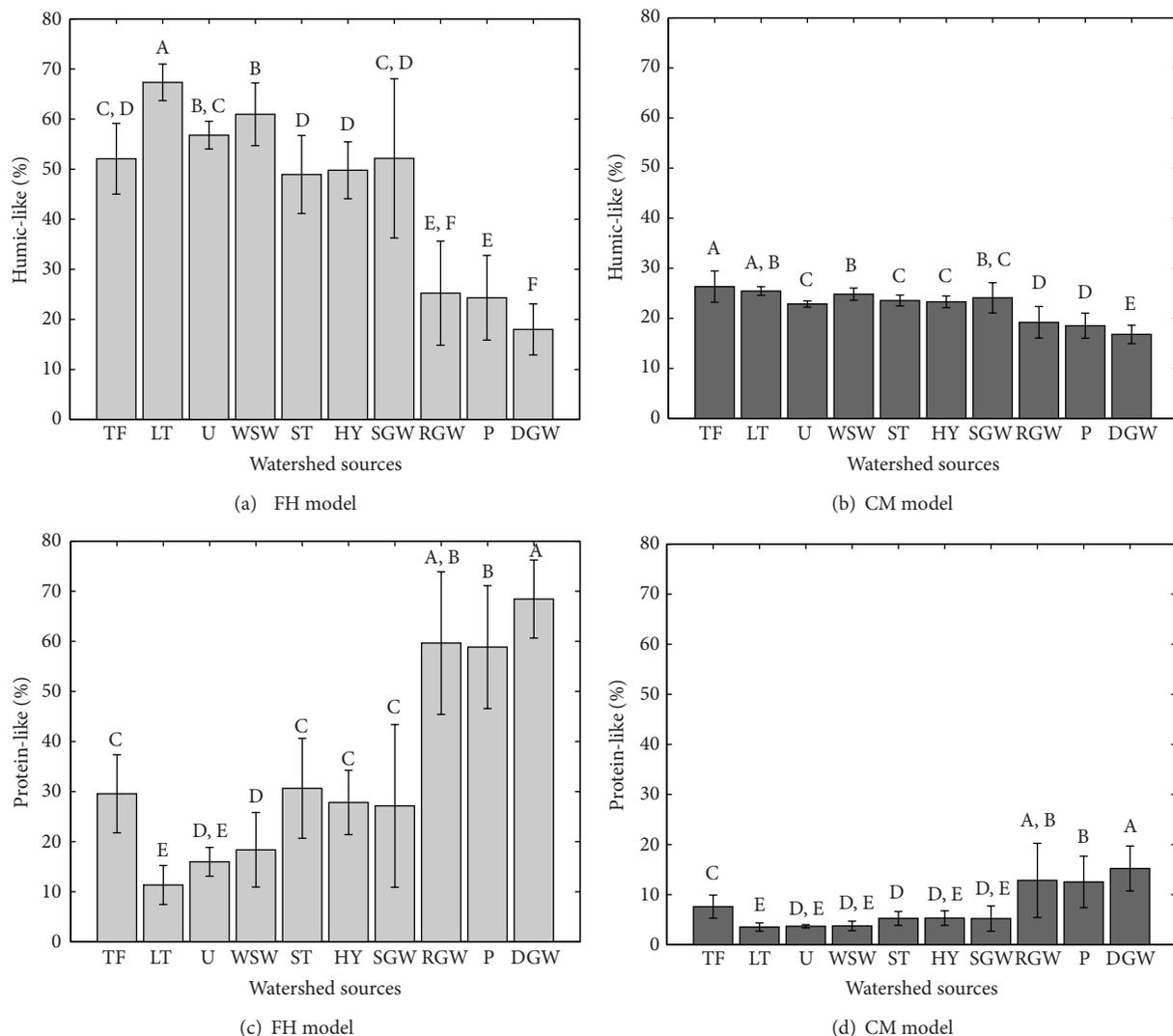


FIGURE 5: Distributions of (a) %humic-like, (c) %protein-like DOM composition derived using site-specific PARAFAC (FH model) across watershed sources are shown. Sources represented with same letters are not significantly different. Results for the same using CM model are shown in (b) and (d), respectively. Standard deviation from means is shown using error bars.

in throughfall, approximately 1.5 times greater than that for deep groundwater (16%, Figure 5(b)). ANOVA on the %humic-like DOM derived from the FH model yielded six distinct classes or categories among the watershed sources (A to F, Figure 5(a)) while the same analyses for the CM model yielded five distinct classes (A to E, Figure 5(b)). The FH model was able to differentiate ( $p < 0.05$ ) between throughfall (TF) and litter leachate (LT) while the same distinction was not possible with the CM model. Similarly, the FH model indicated that litter leachate and wetland soil water were significantly different while the same was not true for the CM model. For groundwater sources, CM model values resulted in a significant difference between the %humic-like DOM for riparian and deep groundwater samples, which was however not replicated by the FH model.

Despite differences, both models produced the same broader trend, that is, a decrease in humic-like DOM from

surficial sources to groundwater sources with a simultaneous increase in %protein-like DOM. With respect to the %protein-like components, the FH model values ranged from 11% for litter leachate to 68% for deep groundwater (Figure 5(c)). In contrast, the range for the CM model was narrower with minimum for litter leachate at 3% to a maximum for deep groundwater at only 15%. However, while the absolute %protein-like DOM values differed substantially among the models, ANOVA analyses indicated that both models classified the watershed sources into five distinct classes. The FH model indicated that throughfall samples were not significantly different from stream, hyporheic, and shallow groundwater for %protein-like DOM (Figure 5(c)). In contrast, the CM model suggested that throughfall was significantly different from all other DOM sources (Figure 5(d)). Similarly, CM model values indicated that %protein-like DOM for the litter samples was not significantly different

from that of wetland soil water, contradicting the results obtained with the FH model.

**3.5. Discriminant Analyses for Watershed Sources Using FH and CM PARAFAC Models.** Forward stepwise discriminant function analysis using FH and CM models revealed distinct differences among DOM for watershed compartments (Figure 6). For each watershed source, the centroid of the data along with the circle representing the 95% confidence region is displayed (Figure 6). Riparian groundwater has a large confidence region, probably due to the small sample size ( $n = 12$ ) whereas stream samples have the smallest confidence region, likely, due to larger sample size ( $n = 82$ ). Biplot rays indicate the direction of variables (humic-like and protein-like DOM compositions) used in space. However, the entire separation (100%,  $p < 0.001$ ) in both the cases (FH and CM models) occurred along the first dimension and humic-like or protein-like DOM characteristic appears to provide good separation. For the FH model, the first discriminant function (Dimension 1) accounted for 81% ( $p < 0.001$ ) of the group (watershed sources) variation, while Dimension 2 accounted for the remaining 19% ( $p < 0.001$ ) (Figure 6(a)). In comparison, the same dimensions (1 and 2) explained 77% ( $p < 0.001$ ) and 23% ( $p < 0.001$ ), of the variability for the CM model (Figure 6(b)). The key observations that come out of these analyses are (a) overall, both models indicate a similar broader trend with seep, riparian, and deep groundwater clustered in the protein-like region while the remaining watershed sources are located in the humic-like region of the plots; (b) in comparison to the CM model, the FH model displays greater separation among watershed sources which is apparent from the reduced overlap of the circular source regions (95% confidence region); and (c) there are slight differences in how the sources are spatially positioned in the discriminant space for the two models. In the CM model space, throughfall is further away from the other humic-like sources; however in the FH model it appears much closer to the other humic-like sources with the exception of tension soil water (U).

Discriminant analysis was used to determine if separation existed between watershed sources based on DOM characteristics (humic- and protein-like DOM composition). This separation was determined using Wilks' lambda multivariate test statistic. Wilks' lambda values for FH and CM model were 0.11 ( $F = 75.88$ ,  $p < 0.001$ ) and 0.17 ( $F = 56.12$ ,  $p < 0.001$ ), respectively. Hence, the greater between-groups variation (separation between watershed sources) as a proportion of the total variation is explained by larger  $F$  statistic observed in case of FH model and the larger the Wilks' lambda statistic, the greater is the within-group variation as a proportion of the total variation as noticed for the CM model (also low  $F$  statistic suggesting lower between-group variability is explained in the case of CM model). Overall, discriminant function analysis revealed a significant association between-watershed sources and DOM characteristics (based on humic- and protein-like DOM composition) for FH model accounting for 59% of between watershed sources variability, thus correctly classifying the groups (watershed

sources) from which samples were collected. In comparison only 51% correct classification (samples predicted belong to the same watershed source) was obtained for the CM model. Biplot rays showed that DOM characteristics were the major factor contributing to the discrimination between watershed sources which is further indicated by the length (or magnitude) of the rays (or vectors) in FH model.

**3.6. Correlation between PARAFAC Model Components and DOM Parameters.** The correlations between the sum of humic-like and protein-like components from the FH and CM models against DOC and DON concentrations and metrics  $a_{254}$ , HIX, and FI are presented in Table 2. One of the most significant results coming out of this comparison was that, overall, the correlations between the FH model components and the DOM parameters were much stronger than the corresponding values for the CM model. This suggested that the FH model characterized DOM more accurately than the CM model. Among watershed sources, the correlations for both FH and CM models were generally weakest for litter and throughfall underscoring the wide variability in DOM composition for these watershed sources. In contrast, correlations were strongest for wetland soil water and shallow groundwater. Across most watershed sources, the humic-like components were positively correlated and the protein-like components were inversely correlated with both DOC and DON (the correlations with DOC being stronger). This would suggest that both C and N fractions of DOM contained a greater fraction of humic DOM. Not surprisingly, in most cases, humic-like components were positively correlated with both  $a_{254}$  and HIX while the protein-like components were positively correlated with FI. This supports the aromatic nature of the humic-like components and the microbial origins of the protein-like DOM.

## 4. Discussion

**4.1. DOM Characterization for Watershed Sources Using the Site-Specific FH Model.** This study revealed a marked variability in humic-like components (FH1, FH3, and FH4) across the watershed sources. Surficial sources (litter leachate and wetland soil water) were especially high in humic-like DOM, whereas groundwater sources, namely, riparian, seeps, and deep groundwater, were not significant contributors of humic-like DOM. These observations are in accordance with previously reported studies [1, 38] as well as our own previous characterization of DOM for this study site based on the CM model [23]. The lack of humic DOM pool for groundwater samples could be attributed to dilution [38] or processes such as sorption and microbial degradation. For instance, Kalbitz and Geyer [40] reported that sorption phenomena with soil mineral surfaces could potentially reduce the humic content of DOM along the hydrologic flow path or as runoff water percolated through the soil profile. Similarly, groundwater sources are prone to microbial decomposition which can result in a significant drop in humic-like DOM in groundwater [38]. The decrease in humic-like components is also supported by DOM concentrations and metrics like  $a_{254}$

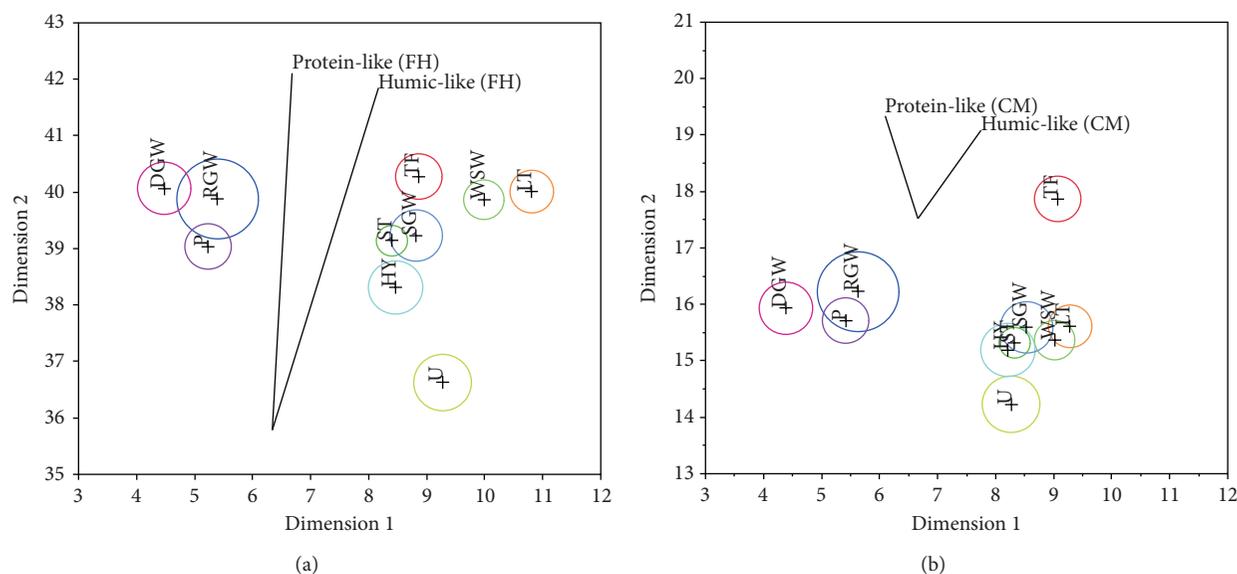


FIGURE 6: Scatter plot of discriminant function analysis showing separation of watershed sources for humic-like and protein-like DOM components based on (a) FH model and (b) CM model results. The extension of biplot rays indicates the strength of correlation between the measured variables.

and HIX (Figures 4(c) and 4(d), resp.). Both  $a_{254}$  and HIX reveal a substantial decrease in values moving from surficial to groundwater sources. Using the UV absorbance values and the equation proposed by Weishaar et al. (2003), we observed lowest aromaticity in subsurface samples (ranged from 14% in deep groundwater to 21% in seep samples) thus corroborating our finding of low aromaticity in subsurface or groundwater samples [9]. Similarly, DOC concentrations in groundwater sources were lower than those for surficial sources suggesting that humic-rich DOC is preferentially sorbed to soil along the hydrologic flow path and/or is degraded by microbial decomposition into lower molecular weight DOM [41–43]. It should be noted that compared to litter leachate, the character of throughfall DOM was much less humic and more protein-rich, indicating a release of less humic and more degradable compounds for the forest canopy.

In contrast to the decrease in humic-like components, relative percent of the protein-like components (FH5 and FH6) increased from surficial to groundwater sources (Figure 3). The increasing trend of protein-like fluorescence is supported by an increase in FI values indicating a greater influence of DOM of microbial origin of DOM (Figure 4). Similar observations have previously been reported by Chen et al. (2010) and Williams et al. (2010) [6, 38]. There was a sharp shift in the protein-like components (FH5 and FH6) between shallow groundwater (SGW) and riparian, seep, and deep groundwater (Figures 3(e) and 3(f), resp.). The shallow groundwater sampled the full soil profile in the wetlands while the other three sources represented water from deeper in the soil profile. Clearly this change had a substantial impact on the protein-like fluorescence of DOM. Furthermore, it should be noted that there were slight differences in the trend for FH5 and FH6 components across watershed sources. FH5

is considered to represent the relatively “fresher” protein-like DOM moieties (tryptophan) whereas FH6 represents more degraded (tyrosine) protein-like fractions [44]. The higher  $F_{max}$  values for FH6 at riparian, seep and deep groundwater (compared to FH5) locations would tend to suggest the existence of more degraded protein-like DOM at these locations. Chen et al. (2010) have also reported more degraded protein-like DOM for groundwater [38].

**4.2. Comparison of Site-Specific versus Prevalidated PARAFAC Models for Characterizing DOM.** The only previous comparison of a site-specific versus a pre-validated model was performed by Fellman et al. (2009) who compared the CM model against a model (referred to as AK) developed from 307 EEMs of soil and stream water samples from southeast Alaska [21]. They compared the two models by (i) examining the EEM residuals obtained from CM model and (ii) determining the correlations between PARAFAC model components and measurements such as DOC and DON concentrations, %bioavailable DOC, and DOM components determined from GC/MS analyses [21]. Based on the EEM residuals, they found that among stream and soil water samples, the CM model did a poor job in fitting the soil water samples from upland and wetland sites. This result may not be surprising considering that the CM model was developed primarily using EEMs from a diverse range of aquatic ecosystems and was not optimized for forested and soil dominated systems. In addition, while comparing the protein-like PARAFAC components against DOM concentrations, Fellman et al. (2009) found that while the two models were not significantly different in predicting DOM concentrations, the CM model explained slightly less variation than the site-specific model [21].

TABLE 2: Correlations between DOM composition (%humic-like or %protein-like, derived from FH and CM model; see text for details) with measured DOM concentrations (DOC and DON) and DOM metrics (derived from UV and fluorescence measurements) across watershed sources.

Watershed sources	DOM composition	DOC	DON	$a_{254}$	HIX	FI
TF	%Humic-like (FH)	0.05	0.34	0.30	0.94	-0.23
	%Humic-like (CM)	0.13	0.49	0.31	0.41	0.46
	%Protein-like (FH)	0.01	-0.27	-0.22	-0.99	0.30
	%Protein-like (CM)	-0.05	-0.27	-0.19	-0.90	0.33
LT	%Humic-like (FH)	-0.05	-0.04	0.09	0.88	-0.52
	%Humic-like (CM)	-0.18	-0.15	-0.23	0.12	-0.20
	%Protein-like (FH)	0.05	-0.15	-0.27	-0.99	0.08
	%Protein-like (CM)	0.04	0.05	-0.07	-0.92	0.40
U	%Humic-like (FH)	0.66	0.42	0.72	0.97	-0.24
	%Humic-like (CM)	-0.12	-0.15	-0.04	0.06	0.40
	%Protein-like (FH)	-0.64	-0.45	-0.68	-0.99	0.22
	%Protein-like (CM)	0.23	0.43	0.12	-0.32	-0.23
WSW	%Humic-like (FH)	0.72	0.38	0.78	0.98	-0.29
	%Humic-like (CM)	0.31	-0.03	0.39	0.39	0.16
	%Protein-like (FH)	-0.73	-0.45	-0.78	-0.99	0.18
	%Protein-like (CM)	-0.55	-0.28	-0.61	-0.93	0.26
ST	%Humic-like (FH)	0.64	0.46	0.65	0.98	-0.23
	%Humic-like (CM)	0.11	0.09	0.23	0.32	-0.15
	%Protein-like (FH)	-0.63	-0.43	-0.60	-0.99	0.11
	%Protein-like (CM)	-0.53	-0.22	-0.52	-0.79	0.15
HY	%Humic-like (FH)	0.42	0.48	0.47	0.96	-0.60
	%Humic-like (CM)	0.05	0.02	0.05	0.41	-0.36
	%Protein-like (FH)	-0.47	-0.45	-0.52	-0.99	0.42
	%Protein-like (CM)	-0.25	-0.18	-0.35	-0.73	0.42
SGW	%Humic-like (FH)	0.94	0.71	0.85	0.99	-0.87
	%Humic-like (CM)	0.89	0.66	0.82	0.90	-0.76
	%Protein-like (FH)	-0.90	-0.67	-0.79	-0.99	0.80
	%Protein-like (CM)	-0.84	-0.61	-0.72	-0.92	0.77
RGW	%Humic-like (FH)	0.46	0.46	0.83	0.99	-0.52
	%Humic-like (CM)	0.20	0.28	0.72	0.88	-0.56
	%Protein-like (FH)	-0.48	-0.44	-0.82	-0.99	0.53
	%Protein-like (CM)	-0.22	-0.04	-0.64	-0.79	0.85
P	%Humic-like (FH)	0.47	0.11	0.30	0.99	0.20
	%Humic-like (CM)	0.27	-0.01	-0.02	0.77	0.30
	%Protein-like (FH)	-0.48	-0.11	-0.30	-0.99	-0.24
	%Protein-like (CM)	-0.25	-0.09	0.08	-0.79	-0.38
DGW	%Humic-like (FH)	0.41	0.17	-0.11	0.99	-0.21
	%Humic-like (CM)	0.47	-0.08	-0.18	0.69	-0.34
	%Protein-like (FH)	-0.31	-0.17	0.12	-0.99	0.09
	%Protein-like (CM)	0.14	0.41	0.17	-0.58	0.42

All correlations significant at  $p < 0.05$ .

We took a slightly different approach than Fellman et al. (2009) in that we directly compared the ability of the two models (CM versus our own FH model) to differentiate DOM from various watershed sources [21]. We did not use the residual EEM approach as used by Fellman et al. (2009) because while fitting the CM model we retained only EEMs that had residuals less than 10% [21]. The direct comparison between

CM and FH models was possible because of the wide variety of watershed sources that were sampled in our study and the large dataset associated with each source. Our results clearly showed that the FH model was much more sensitive to subtle variations in DOM and did a much better job in highlighting the differences in DOM among watershed sources. This was highlighted both by the model differences in humic-like and

TABLE 3: Mean (standard deviation) values for the changes in humic- and protein-like DOM composition between FH and CM models across watershed sources in this study.

Watershed sources	Sample size ( <i>n</i> )	Change in %humic-like	Change in %protein-like
TF	37	0.99 (0.22)	2.94 (0.53)
LT	42	1.65 (0.14)	2.16 (0.51)
U	24	1.49 (0.13)	3.34 (0.73)
WSW	49	1.45 (0.22)	3.74 (0.90)
ST	82	1.07 (0.30)	4.78 (1.24)
HY	27	1.14 (0.20)	4.31 (0.96)
SGW	29	1.12 (0.47)	3.98 (1.16)
RGW	12	0.27 (0.34)	4.30 (1.61)
P	37	0.29 (0.31)	4.17 (1.38)
DGW	28	0.06 (0.22)	3.78 (1.16)

protein-like components as well as the spatial separation of watershed sources in discriminant analysis (Figure 6). Not only did the FH model allow greater differentiation among watershed sources, but the sum of humic- and protein-like model components also yielded stronger relationships with DOC and DON concentrations, thus providing important insights into the ecological nature of the C and N in watershed pools. Among the humic-like and protein-like components, the differences between FH and CM models were higher for the protein-like components (Table 3). This suggests that in our study the CM model tended to underpredict the protein-like moieties for watershed sources. This assessment was also recently alluded to by Larsen et al. (2010) who implemented the CM model to characterize DOM for the ridge and slough landscape in the Florida Everglades [22]. Larsen et al. (2010) found that the highest EEM residuals were associated with the protein-like components which meant that the model had limited capability in resolving these compounds [22]. They further speculated that the inability to correctly resolve these reactive protein-like compounds could hamper our understanding of associated processes such as rapid microbial uptake and photo-degradation.

In light of the results from this study and the observations previously made by Fellman et al. (2009) and Larsen et al. (2010), we suggest that CM model would be appropriate to characterize DOM and should be implemented where (a) EEMs data is not sufficient to develop a PARAFAC model, (b) broad trends in DOM are required, and/or (c) DOM comparisons between disparate watershed sites are performed [21, 22]. However, we suggest that developing a site-specific PARAFAC model should be a preferred approach given that a long-term EEM data is available. Given the sensitivity of our results, a site-specific model could do a slightly better job in characterizing the differences in DOM composition among watershed sources, seasonal or spatial patterns in DOM composition and consequently a better understanding the potential processes (e.g., photodegradation, microbial uptake, and sorption) affecting fate and transport of DOM.

## 5. Conclusions

This study compared a site-specific PARAFAC model (FH model) against a prevalidated PARAFAC model [17] for characterizing DOM composition in a forested, headwater watershed. The PARAFAC models were developed for fluorescence EEMs determined for various watershed sources including throughfall, litter leachate, wetland soil water, stream water, hyporheic zone, shallow and deep groundwater, and groundwater seeps. The site-specific model was more sensitive to subtle differences in DOM and was able to provide a greater level of distinction among watershed sources. The humic- and protein-like constituents derived from the site-specific model displayed more pronounced differences among watershed sources compared to the prevalidated model. DOC and DON concentrations and selected DOM metrics were also more strongly correlated with components of the site-specific PARAFAC model versus the prevalidated model. These results suggest that while a prevalidated PARAFAC model may capture the broad trends in DOM composition and may allow comparisons with other study sites, a site-specific model will do a better job in characterizing within-watershed differences in DOM and consequently provide important insights into processes and mechanisms influencing DOM.

## Conflict of Interests

The authors declare no competing financial interests.

## Acknowledgments

This study was funded through a grant from the National Science Foundation (NSF, Hydrologic Sciences Program, EAR-0809205). The authors would like to thank Captain Wayne Suydam and the Fair Hill NRMA staff for providing access to the study site. They also thank Gurbir Dhillon, Sudarshan Dutta, John Van Stan, and Rachael Vaicunas for sampling and installation of equipment in the watershed and anonymous reviewers for their constructive comments and suggestions to improve this paper.

## References

- [1] E. M. Thurman, *Organic Geochemistry of Natural Waters*, Martinus Nijhoff/Dr W. Junk, Dordrecht, The Netherlands, 1985.
- [2] G. R. Aiken, C. C. Gilmour, D. P. Krabbenhoft, and W. Orem, "Dissolved organic matter in the Florida everglades: implications for ecosystem restoration," *Critical Reviews in Environmental Science and Technology*, vol. 41, no. 1, pp. 217–248, 2011.
- [3] Y. Yamashita and R. Jaffé, "Characterizing the interactions between trace metals and dissolved organic matter using excitation-emission matrix and parallel factor analysis," *Environmental Science and Technology*, vol. 42, no. 19, pp. 7374–7379, 2008.
- [4] T. E. C. Kraus, B. A. Bergamaschi, P. J. Hernes et al., "Assessing the contribution of wetlands and subsided islands to dissolved organic matter and disinfection byproduct precursors in the Sacramento-San Joaquin River Delta: a geochemical approach," *Organic Geochemistry*, vol. 39, no. 9, pp. 1302–1318, 2008.

- [5] M. A. Moran, W. M. Sheldon Jr., and R. G. Zepp, "Carbon loss and optical property changes during long-term photochemical and biological degradation of estuarine dissolved organic matter," *Limnology and Oceanography*, vol. 45, no. 6, pp. 1254–1264, 2000.
- [6] C. J. Williams, Y. Yamashita, H. F. Wilson, R. Jaffe, and M. A. Xenopoulos, "Unraveling the role of land use and microbial activity in shaping dissolved organic matter characteristics in stream ecosystems," *Limnology and Oceanography*, vol. 55, no. 3, pp. 1159–1171, 2010.
- [7] R. G. Wetzel, "Dissolved organic carbon: detrital energetics, metabolic regulators, and drivers of ecosystem stability of aquatic ecosystems," in *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*, S. E. G. Findlay and R. L. Sinsabaugh, Eds., pp. 455–477, Academic Press, San Diego, Calif, USA, 2003.
- [8] A. Bricaud, A. Morel, and L. Prieur, "Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains," *Limnology and Oceanography*, vol. 26, pp. 43–53, 1981.
- [9] J. L. Weishaar, G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper, "Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon," *Environmental Science and Technology*, vol. 37, no. 20, pp. 4702–4708, 2003.
- [10] P. G. Coble, S. A. Green, N. V. Blough, and R. B. Gagosian, "Characterization of dissolved organic matter in the Black Sea by fluorescence spectroscopy," *Nature*, vol. 348, no. 6300, pp. 432–435, 1990.
- [11] D. M. McKnight, E. W. Boyer, P. K. Westerhoff, P. T. Doran, T. Kulbe, and D. T. Andersen, "Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity," *Limnology and Oceanography*, vol. 46, no. 1, pp. 38–48, 2001.
- [12] P. J. Hernes, R. G. M. Spencer, R. Y. Dydá, B. A. Pellerin, P. A. M. Bachand, and B. A. Bergamaschi, "The role of hydrologic regimes on dissolved organic carbon composition in an agricultural watershed," *Geochimica et Cosmochimica Acta*, vol. 72, no. 21, pp. 5266–5277, 2008.
- [13] J. F. Saraceno, B. A. Pellerin, B. D. Downing, E. Boss, P. A. M. Bachand, and B. A. Bergamaschi, "High-frequency in situ optical measurements during a storm event: assessing relationships between dissolved organic matter, sediment concentrations, and hydrologic processes," *Journal of Geophysical Research G*, vol. 114, no. 4, Article ID G00F09, 2009.
- [14] J. R. Helms, A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper, "Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter," *Limnology and Oceanography*, vol. 53, no. 3, pp. 955–969, 2008.
- [15] R. G. M. Spencer, P. J. Hernes, R. Ruf et al., "Temporal controls on dissolved organic matter and lignin biogeochemistry in a pristine tropical river, Democratic Republic of Congo," *Journal of Geophysical Research G*, vol. 115, no. 3, Article ID G03013, 2010.
- [16] T. Ohno, "Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter," *Environmental Science and Technology*, vol. 36, no. 4, pp. 742–746, 2002.
- [17] R. M. Cory and D. M. McKnight, "Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter," *Environmental Science and Technology*, vol. 39, no. 21, pp. 8142–8149, 2005.
- [18] J. B. Fellman, E. Hood, D. V. D'Amore, R. T. Edwards, and D. White, "Seasonal changes in the chemical quality and biodegradability of dissolved organic matter exported from soils to streams in coastal temperate rainforest watersheds," *Biogeochemistry*, vol. 95, no. 2–3, pp. 277–293, 2009.
- [19] C. A. Stedmon, S. Markager, and R. Bro, "Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy," *Marine Chemistry*, vol. 82, no. 3–4, pp. 239–254, 2003.
- [20] J. Boehme, P. Coble, R. Conmy, and A. Stovall-Leonard, "Examining CDOM fluorescence variability using principal component analysis: seasonal and regional modeling of three-dimensional fluorescence in the Gulf of Mexico," *Marine Chemistry*, vol. 89, no. 1–4, pp. 3–14, 2004.
- [21] J. B. Fellman, M. P. Miller, R. M. Cory, D. V. D'Amore, and D. White, "Characterizing dissolved organic matter using PARAFAC modeling of fluorescence spectroscopy: a comparison of two models," *Environmental Science and Technology*, vol. 43, no. 16, pp. 6228–6234, 2009.
- [22] L. G. Larsen, G. R. Aiken, J. W. Harvey, G. B. Noe, and J. P. Crimaldi, "Using fluorescence spectroscopy to trace seasonal DOM dynamics, disturbance effects, and hydrologic transport in the Florida Everglades," *Journal of Geophysical Research G*, vol. 115, no. 3, Article ID G03001, 2010.
- [23] S. Inamdar, N. Finger, S. Singh et al., "Dissolved organic matter (DOM) concentration and quality in a forested mid-Atlantic watershed, USA," *Biogeochemistry*, vol. 108, no. 1–3, pp. 55–76, 2012.
- [24] S. Inamdar, S. Singh, S. Dutta et al., "Fluorescence characteristics and sources of dissolved organic matter for stream water during storm events in a forested mid-Atlantic watershed," *Journal of Geophysical Research G*, vol. 116, no. 3, Article ID G03043, 2011.
- [25] "Maryland State Climatologist Office Data Page," 2008, <http://www.atmos.umd.edu/~climate/weather/marylandnormals.htm>.
- [26] J. T. van Stan II and D. F. Levia Jr., "Inter- and intraspecific variation of stemflow production from *Fagus grandifolia* Ehrh. (American beech) and *Liriodendron tulipifera* L. (yellow poplar) in relation to bark microrelief in the eastern United States," *Ecohydrology*, vol. 3, no. 1, pp. 11–19, 2010.
- [27] J. Ameer, J. R. P. Axler, and C. J. Owen, "Persulfate digestion for determination of total nitrogen and phosphorus in low-nutrient waters," *American Environmental Laboratory*, vol. 10, p. 1, 1993.
- [28] A. J. Lawaetz and C. A. Stedmon, "Fluorescence intensity calibration using the Raman scatter peak of water," *Applied Spectroscopy*, vol. 63, no. 8, pp. 936–940, 2009.
- [29] R. Bro and A. K. Smilde, "Centering and scaling in component analysis," *Journal of Chemometrics*, vol. 17, no. 1, pp. 16–33, 2003.
- [30] P. Westerhoff, W. Chen, and M. Esparza, "Fluorescence analysis of a standard fulvic acid and tertiary treated wastewater," *Journal of Environmental Quality*, vol. 30, no. 6, pp. 2037–2046, 2001.
- [31] N. Mladenov, P. Huntsman-Mapila, P. Wolski, W. R. L. Masamba, and D. M. McKnight, "Dissolved organic matter accumulation, reactivity, and redox state in ground water of a recharge wetland," *Wetlands*, vol. 28, no. 3, pp. 747–759, 2008.
- [32] C. A. Stedmon and R. Bro, "Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial," *Limnology and Oceanography*, vol. 6, pp. 572–579, 2008.
- [33] S. A. Green and N. V. Blough, "Optical absorption and fluorescence properties of chromophoric dissolved organic matter in

- natural waters," *Limnology and Oceanography*, vol. 39, no. 8, pp. 1903–1916, 1994.
- [34] P. G. Coble, "Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy," *Marine Chemistry*, vol. 51, no. 4, pp. 325–346, 1996.
- [35] J. B. Fellman, R. G. M. Spencer, P. J. Hernes, R. T. Edwards, D. V. D'Amore, and E. Hood, "The impact of glacier runoff on the biodegradability and biochemical composition of terrigenous dissolved organic matter in near-shore marine ecosystems," *Marine Chemistry*, vol. 121, no. 1–4, pp. 112–122, 2010.
- [36] S. Singh, E. J. D'Sa, and E. M. Swenson, "Chromophoric dissolved organic matter (CDOM) variability in Barataria Basin using excitation-emission matrix (EEM) fluorescence and parallel factor analysis (PARAFAC)," *Science of the Total Environment*, vol. 408, no. 16, pp. 3211–3222, 2010.
- [37] Y. Yamashita, A. Panton, C. Mahaffey, and R. Jaffé, "Assessing the spatial and temporal variability of dissolved organic matter in Liverpool Bay using excitation-emission matrix fluorescence and parallel factor analysis," *Ocean Dynamics*, vol. 61, no. 5, pp. 569–579, 2010.
- [38] M. Chen, R. M. Price, Y. Yamashita, and R. Jaffé, "Comparative study of dissolved organic matter from groundwater and surface water in the Florida coastal Everglades using multi-dimensional spectrofluorometry combined with multivariate statistics," *Applied Geochemistry*, vol. 25, no. 6, pp. 872–880, 2010.
- [39] N. Hudson, A. Baker, and D. Reynolds, "Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review," *River Research and Applications*, vol. 23, no. 6, pp. 631–649, 2007.
- [40] K. Kalbitz and W. Geyer, "Humification indices of water-soluble fulvic acids derived from synchronous fluorescence spectra—effects of spectrometer type and concentration," *Journal of Plant Nutrition and Soil Science*, vol. 164, no. 3, pp. 259–265, 2001.
- [41] N. S. Bolan, D. C. Adriano, A. Kunhikrishnan, T. James, R. McDowell, and N. Senesi, "Dissolved organic matter: biogeochemistry, dynamics, and environmental significance in soils," in *Advances in Agronomy*, D. Sparks, Ed., pp. 1–75, Elsevier, New York, NY, USA, 1st edition, 2011.
- [42] K. Kalbitz, A. Meyer, R. Yang, and P. Gerstberger, "Response of dissolved organic matter in the forest floor to long-term manipulation of litter and throughfall inputs," *Biogeochemistry*, vol. 86, no. 3, pp. 301–318, 2007.
- [43] H. V.-M. Nguyen, J. Hur, and H.-S. Shin, "Changes in spectroscopic and molecular weight characteristics of dissolved organic matter in a river during a storm event," *Water, Air, and Soil Pollution*, vol. 212, no. 1–4, pp. 395–406, 2010.
- [44] C. Santín, Y. Yamashita, X. L. Otero, M. Á. Álvarez, and R. Jaffé, "Characterizing humic substances from estuarine soils and sediments by excitation-emission matrix spectroscopy and parallel factor analysis," *Biogeochemistry*, vol. 96, no. 1–3, pp. 131–147, 2009.