Spring 1-1-2012

Advances in Understanding the Contributions of Microbial Dissolved Organic Matter to the Fluorescence Signature of Natural Waters Using Parallel Factor Analysis

Jessica L. Ebert
University of Colorado at Boulder, jessica.l.ebert@colorado.edu

ADVANCES IN UNDERSTANDING THE CONTRIBUTIONS OF MICROBIAL DISSOLVED ORGANIC MATTER TO THE FLUORESCENCE SIGNATURE OF NATURAL WATERS USING PARALLEL FACTOR ANALYSIS

by
JESSICA L. EBERT
B.S. Syracuse University, 2010

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Masters of Science
Department of Civil, Environmental and Architectural Engineering
2012

This thesis entitled:
Advances in understanding the contributions of microbial dissolved organic matter to the fluorescence signature of natural waters using parallel factor analysis

written by Jessica L. Ebert
has been approved for the Department of Civil, Environmental and Architectural Engineering

_____
(Diane McKnight)

_____
(Fernando Rosario-Ortiz)

_____
(Natalie Mladenov)

Date_____

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

Abstract

Ebert, Jessica L. (M.S., Environmental Engineering)

Advances in understanding the contributions of microbial dissolved organic matter to the fluorescence signature of natural waters using parallel factor analysis

Thesis directed by Professor Dr. Diane McKnight

Dissolved organic matter (DOM) is present in all aquatic ecosystems and plays and important role in the global carbon cycle. DOM can be characterized using fluorescence spectroscopy. The measurements that are collected from a DOM fluorescence scan are commonly referred to as three dimensional excitation-emission matrices (EEMs) (Coble et al, 1990; McKnight et al 2001). In this study EEMs were run through an already established model (Cory and McKnight), but we were left with areas of high residuals, e.g. areas of the EEM that were not modeled as well as one would like. Therefore, these results identified the need for the creation of a new model; one specialized towards fluorescence data for DOM enriched in microbial products. The new model used samples taken from 3 low humic sites and 2 microbial incubations. Using parallel factor analysis (PARAFAC) a 5-component model was validated, with component C1 being the microbial input.

Table of Contents

List of Tables

Tables

List of Figures

1. Introduction

Dissolved organic matter (DOM) is present in all aquatic ecosystems and plays an important role in the global carbon cycle. DOM is comprised of a complex, heterogeneous mixture of organic compounds derived from both microbial and terrestrial sources. Most DOM in freshwater contains a large humic fraction predominantly derived from the degradation of plant litter organic matter, leaching of soils and other terrestrial sources. Degradation of microbial biomass in lakes and streams can also contribute to the humic fraction of DOM. The humic fraction has the ability to affect the quality of life for in-stream microorganisms and is responsible for controlling the physical habitat by absorbing visible light.

It has also been shown that humic DOM can act as an electron shuttle to stimulate the reduction of Fe (III) oxides (Nevin and Lovely 2002) under anoxic conditions. Humic substances are also able to act as electron acceptors during anaerobic microbial respiration (Scott et al 1998, Lovely et al 1999). It has also been shown that electron transfer by humic substances uses multiple redox functional groups, including quinones and sulfur-containing groups (Finmen et al. 2007). This leads one to believe that these redox processes are occurring by several mechanisms other than just the formation of quinones and dihydroquinones. However the microbial electron transfer is less energetic than that of catalytic reduction techniques because the microbes are unable to use non-quinone sites for electron transfer (Ratasuk and Nanny 2007).

DOM can be characterized using fluorescence spectroscopy. Fluorescence is a process that occurs when a molecule is excited to a higher energy level by absorbing energy and then returns to its ground state. The energy that is lost as the molecule returns to its ground state is emitted as light and is called fluorescence. The measurements that are collected from a DOM fluorescence scan are commonly referred to as three dimensional excitation-emission matrices (EEMs) (Coble et al, 1990; McKnight et al 2001). An example EEM is shown in the following figure.

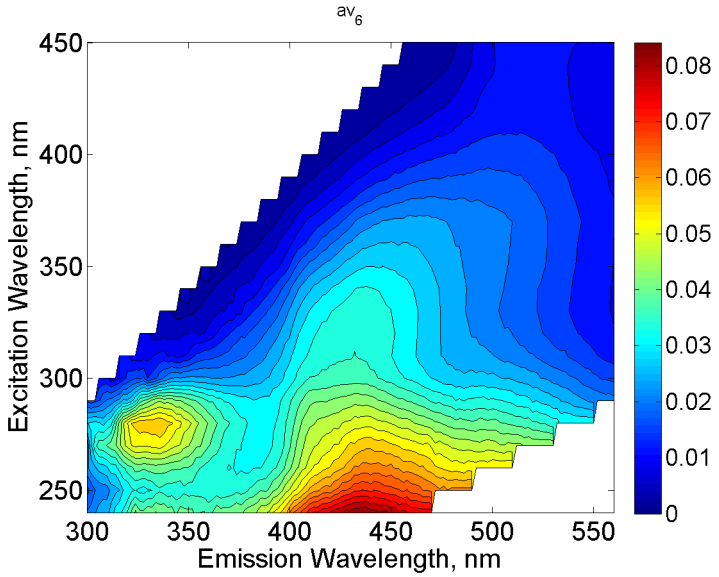

Figure 1: Example excitation-emission matrix (EEM) from the Sierra Nevada Mountain range in Spain.

EEMs can provide a large amount of information about the chemical character of the molecules in the DOM pool that fluoresce, mostly based on where peaks are occurring. These various peaks are commonly referred to as fluorophores. This information can be used to calculate a variety of indices. The fluorescence index (FI) is calculated as the ratio of emission wavelengths at 470nm and 520nm at an excitation of 370 nm (Cory and McKnight 2005). Having a high FI (1.8) indicates that the DOM in the sample is derived mostly from microbial organic matter sources while having a low FI (1.2) indicates that the DOM in the sample is derived mostly from terrestrial organic matter sources. Because there are large amounts of data contained in each individual EEM, a statistical method called parallel factor analysis (PARAFAC) was developed to separate the overall signal into those of its fluorophores (Stedmon and Bro 2008).

The Cory and McKnight PARAFAC model was created for the purpose of exploring the overall role of DOM in redox processes from the perspective of humics as electron shuttles (Klapper et al). The Cory and McKnight PARAFAC model is broad in biogeochemical scope. EEMs were obtained for samples collected from all over the globe, including many reducing environments. Because of this approach, the Cory and McKnight model a great starting point when looking at large sets of data (Cory and McKnight 2005). Using this model also allows one to calculate the redox index (RI) for a sample. The RI is calculated by the following equation. In this equation the amount (or loading, Q_{red}) of reduced quinone-like fluorophores is divided by the total number of quinone-like fluorophores ($Q_{red} + Q_{ox}$)(Miller et al 2006).

$$RI = \frac{Q_{red}}{Q_{red} + Q_{ox}}$$

In this study I examined the EEMs of a specialized set of samples, coming from microbial incubation experiments and field sites with very high FI values. The samples that have undergone anoxic microbial incubation resulted in changes in the amount of humic DOM over the course of the experiment, showing that there were changes occurring during that process. The post incubation data showed a decrease in the relative proportion of the humic fraction and an increase in the protein fraction. In this study EEMs were run through an already established model (Cory andMcKnight), which breaks down the data into 13 separate fluorophores. When samples were run through the Cory and McKnight model, we were left with areas of high residuals, e.g. areas of the EEM that were not modeled as well as one would like. When looking at a residual obtained in applying a PARAFAC model to an EEM the overall appearance should be noisy (e.g. a scratchy appearance with no noticeable peaks) with no peak of greater than 10% in the residual. Therefore, these results identified the need for the creation of a new model; one specialized towards fluorescence data for DOM enriched in microbial products.

The following figure shows the EEM of one of the incubation samples as modeled using Cory and McKnight. The top left hand image represents the original EEM, the top right hand image represents the modeled version, the bottom left hand represents the overall residual, and the bottom right hand shows the residuals as a percentage. A large residual (>10%) at em 525mn and ex 375nm is apparent. The

new model aims to address that residual by accurately representing the associated fluorophore or fluorophores in an expanded PARAFAC model. This model will provide greater ability to model fluorescence data from microbial sources. In turn, this model will allow for EEMs from new and old microbial sample sets to be reanalyzed in hope that they will be better modeled in the future.

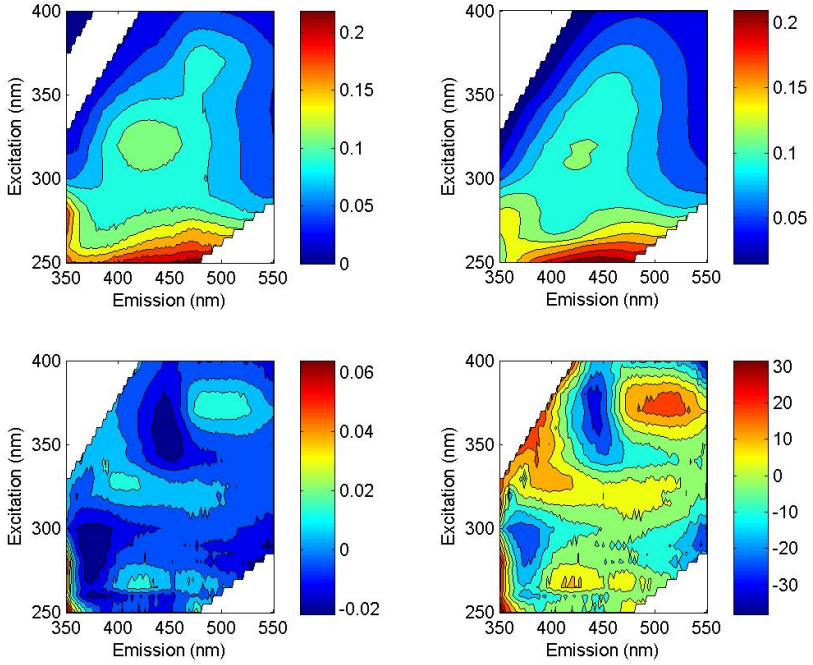

Figure 1: Sample modeled using Cory and McKnight. The top 2 EEMs represent the original (left) and modeled (right) EEM. The bottom 2 EEMs are the overall; residual (left) and the residual as a percentage of the total (right).

2. Site Description

Sample EEMS were collected from 5 separate locations including Bangladesh, 2 sites in Antarctica (Lake Bonney and the Cotton Glacier), Boulder Creek, and the Sierra Nevada mountain range in Spain. For two of these sites (Bangladesh and Boulder Creek) the data came from anaerobic microbial incubation experiments. The other three sites the data came from locations that have little to no inputs from terrestrial organic matter, this includes Lake Bonney, Cotton Glacier, and the high alpine lakes in the Sierra Nevada mountain range.

2. 1 Microbial Incubations

2.1.1 Bangladesh

Water and sediment samples were obtained from the Araihazar region of Bangladesh (Figure 3), which is located about 30 km northeast of Dhaka. Samples were collected from groundwater wells and were used in a study focusing on the arsenic contamination of groundwater supplies. The Araihazar region has monsoon climate with most precipitation occurring between June and September; this causes groundwater levels to vary seasonally (Immerzeel 2008, Stute et al 2007). Samples used in the incubations were collected from both site B and site K. Site K is located in a rural region within the floodplain of the Old Brahmaputra River, which is now considered a small stream due to sedimentation. Site B is located in the Bayalakandi village and consists of 1 well with various depths. The following pictures (Figure 3) show the people in this region of Bangladesh and some sample collection activities.

Figure 3: Photo collage of the Bangladesh site showing the people of the region and the sampling effort.

2.1.2 Boulder Creek

Boulder Creek is a tributary of the South Platte River with headwaters that flow from the east side of the Continental Divide. The upper region of Boulder Creek contains mostly high alpine vegetation with mix of forest and shrubland. The lower region of Boulder Creek flows through a predominantly urban setting to plains before it reaches the South Platte River. All water and sediment samples were collected from the same site on Boulder Creek, upstream of the wastewater treatment plant (40^0 0' N, 105^0 14' W). The site is located in an urban transition with vegetated banks (width less than 20 meters) on either side and is near a pedestrian path. Boulder Creek EEMs were collected for an emission range of 300-600nm. The following picture shows Boulder Creek.

Figure 4: Photograph of Boulder Creek

2.2 Low humic and high microbial activity sites

2.2.1 Lake Bonney

Lake Bonney is located in McMurdo Dry Valleys, Antarctica and has two lobes (Spigel and Priscu 1998). Each lobe has a maximum depth of 40 m and is separated by a sill at a depth of ~13 m. Lake Bonney has a perennial ice-cover (between 3-5 m thick) that prevents wind driven mixing, creating a vertical stratification. The chemocline in the west lobe of Lake Bonney (WLB) occurs at ~15m (Spigel and Priscu 1998). The bottom water of WLB is derived from the evapoconcentration of marine water and inputs of subglacial outflow from Taylor Glacier, whereas the

surface water of WLB is derived from glacial runoff into seasonal streams. Thus, the biological assemblages in these lobes reflect the effects of landscape position and ecological legacy (Lyons et al. 2001). Samples were collected for fluorescence spectroscopy and biogeochemistry during the austral summer, November 2003. Lake Bonney EEMs were collected for an emission range of 300-600nm. The following picture shows a researcher looking over Lake Bonney the day of sampling.

Figure 5: Photograph of Lake Bonney taken near Blood Falls feature at the west end of the West Lobe.

2.2.2 Cotton Glacier

The Cotton Glacier is located in the Transantarctic Mountains in the region of Victoria Land, Antarctica. Samples of a supraglacial stream were collected from the glacier north of Cape Roberts and on the south side of the Clare Range (Foreman et al, in review). This section of the glacier is a region where high winds commonly occur, which deposit dark sediment on the surface. These sediments are then heated up by solar radiation, producing meltwater. The following photo is of sampling Cotton Glacier.

Figure 6: Photograph of Cotton Glacier supraglacial stream being sampled

2.2.3 Sierra Nevada Lakes

The Sierra Nevada Mountain Range is located in Andalusia provinces of Spain near the cities of Granada and Almeria and is known for containing the highest elevation peak in continental Spain. Several Sierra Nevada lakes were sampled during a field study to determine the effects of dust inputs on dissolved organic matter in high alpine lakes (Mladenov et al 2011). Samples from various lakes were included in this model.

Figure 7: Photograph of Lago Aguas Verdes in Sierra Nevada, Spain

3. Methods

3.1 Microbial Incubations

3.1.1 Bangladesh

Groundwater samples for the incubations were collected anaerobically in serum vials with butyl stoppers until the incubations were prepared. Sediments for the incubations were preserved on ice in mylar bags with oxygen-absorbing packets until the incubations were prepared. Incubations were prepared under an N2 atmosphere in a temporary glove bag, with an average of 3.5g of sediment added to each 20ml amber serum incubation vial. For control incubations 18ml of groundwater was added to the vial and then anaerobically capped with a butyl stopper and an aluminum crimp seal. For the treatment incubations the treatment carbon source was mixed with 18ml of groundwater, and then added to the incubation vial, which was capped in the temporary glove box. The vials were then incubated for 90 days. After the 90 days, the groundwater and sediment were sediment in an anaerobic glove box with an N2/CO2/H2 atmosphere. Aliquots of groundwater were then filtered and acidified to a pH of 2 in the glove box and stored at 4°C until analysis. Anaerobic conditions were maintained during fluorescence spectroscopy and specific UV absorbance measurements.

3.1.2 Boulder Creek

Two separate laboratory incubations of distinct dissolved organic matter fractions were performed For the first experiment, 90 mL of creek water was passed through a 0.22 µm polycarbonate filter using a vacuum pump in order to establish a microbial community for each replicate microcosm. For the second experiment approximately 5 L of creek water was pre-filtered through 1.2 µm glass fiber filters using a vacuum pump. The less than 1.2 µm size fraction is devoid of nearly all protozoan grazers, this sample was then used to establish a microbial community in each replicate microcosm by filtering 90 mL through a 0.22 µm polycarbonate filter using a vacuum pump and then placing one filter into each microcosm. Bottles were immediately sealed and crimped with butyl stoppers and then flushed with purified helium for 10 minutes, followed by an addition of 5 mL of acetylene (C_2H_2). Acetylene was produced in anaerobic bottles by reacting Calcium Carbide and DI water. Replicates were then spiked with approximately 100 µM NO_3^- (as Sodium Nitrate) in the first incubation and approximately 40 µM NO_3^- in the second incubation (lower concentrations in the second experiment were due to experimenter error), and placed on a rotator. All microcosms were shielded from light with aluminum foil and were incubated at approximately 23° C for approximately 30 days.

3.2 EEM Collection

All EEMs were collected using the Jobin-Yvon Horiba Fluoromax series fluorometers over an excitation range of 240-450nm and an emission range of 300-550nm unless otherwise noted. All EEMs were collected using a 1cm quartz cuvette and if necessary were diluted to an absorbance of <0.2 cm^{-1} at λ= 254 nm to ensure that the sample data is in the range for optimal correction of the inner filter effect. The EEMs were blank subtracted using a MilliQ EEM, corrected for instrument optics using factors from the manufacturer, Raman area normalized, and inner filter corrected before modeling began (Ohno 2002). This level of processing allows inter-comparison of the EEMs.

3.3 Model Creation

The model was created in Matlab using the tutorial written by Colin Stedmon and Rasmus Bro. Data from 94 samples were used; the number of EEMS from each site varied according to the following table.

Site Location	Number of samples
Bangladesh	43
Cotton Glacier	8
Lake Bonney	10
Boulder Creek	23
Sierra Nevada	10

Table 1: Site locations and number of samples used from each to create the model

The first step in creating the model was combining all the data from the various locations; this step included trimming the various EEMS to the same excitation-emission range and removing the Rayleigh scattering. The excitation range used was from 240nm to 450nm at every 5mn and the emission range was from 300nm to 550nm at every 2 nm. After compiling all the samples the data was put into Matlab and plotted to make sure each EEM appeared correctly. Samples were then normalized to the same maximum intensity because of differences of greater than an order of magnitude in maximum intensity among the EEMs. This approach is the same as described by Cawley et al. (2012).

The data were then analyzed to determine if there were any outliers. Outlier tests were performed step-wise on all wavelengths of data constrained by non-negativity. Leverages and loadings of each sample emission wavelength and excitation wavelength were used to determine if there were any outliers. The Loadings show how much of each component is in each sample. When looking at these all of them should be positive and not go over 100%. Leverages are a measure used in regression statistics; samples that have high leverage will greatly effect model results if removed. When looking at these they should all be similar almost in a straight line. The following figure (8) shows the results of plotting the leverage and loadings for 5 components obtained through development of the new PARAFAC model.

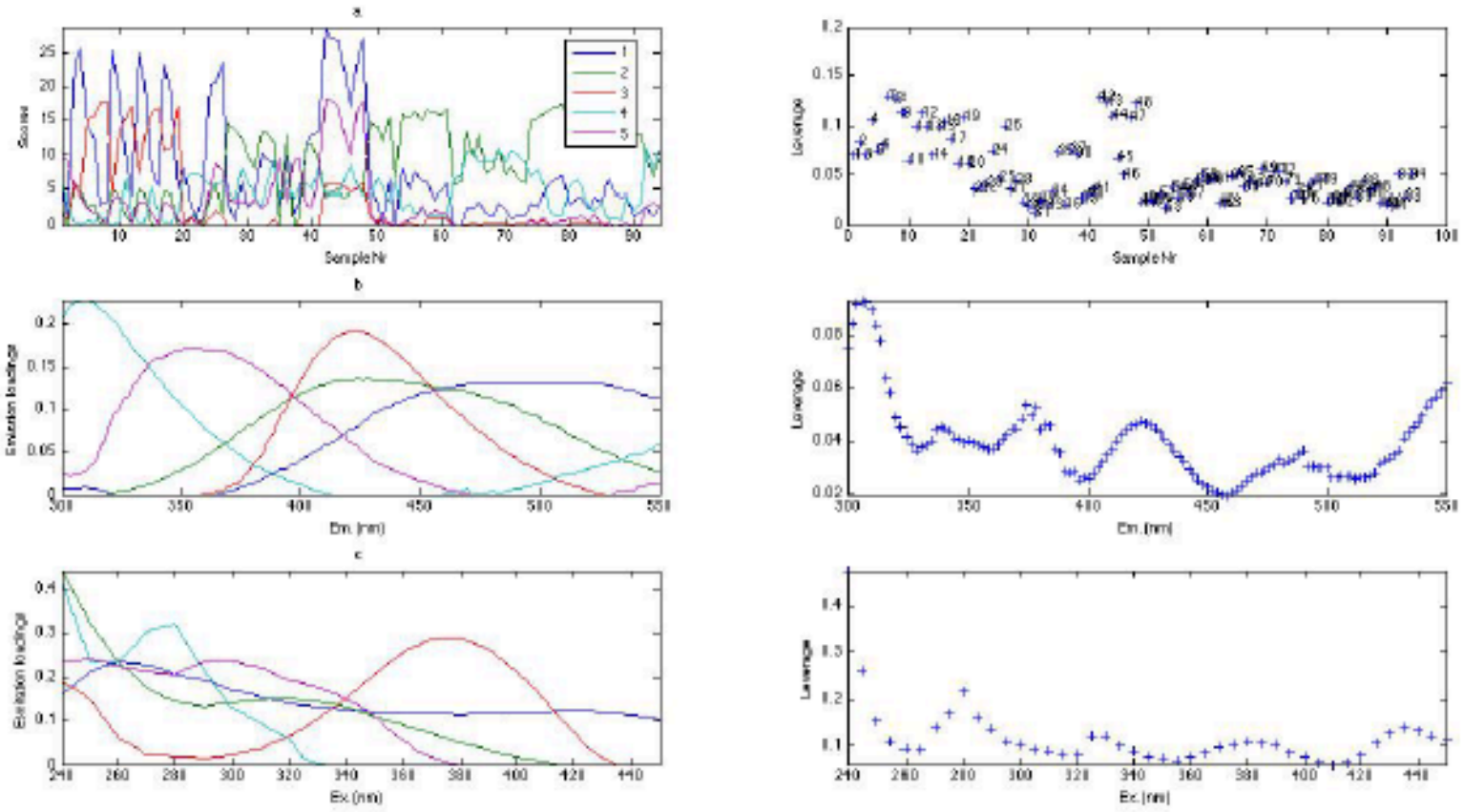

Figure 8: The left hand column represents the loadings for 5 components based on sample number, excitation, and emission. The right hand column shows the leverages for all the samples based on sample number, emission, and excitation.

Examining this figure shows that there are no samples that are clear outliers. The leverages for all the samples are in the same range, with no samples being significantly higher than another.

The model was evaluated by comparing plots of the modeled data and the original data and then examining a plot of the residuals. This approach allows for comparison between various samples and number of modeled components. In this approach it is possible to see which number of components might be needed to have the least amount of residual available.

The model was then evaluated using split half analysis; the data was split into 4 unique sets and then recombined into 4 halves that will compare all of the data. The following figure (9) from the tutorial shows an example of how the split half works; both the split half sections contain all of the data so if either validates it means that your model is valid.

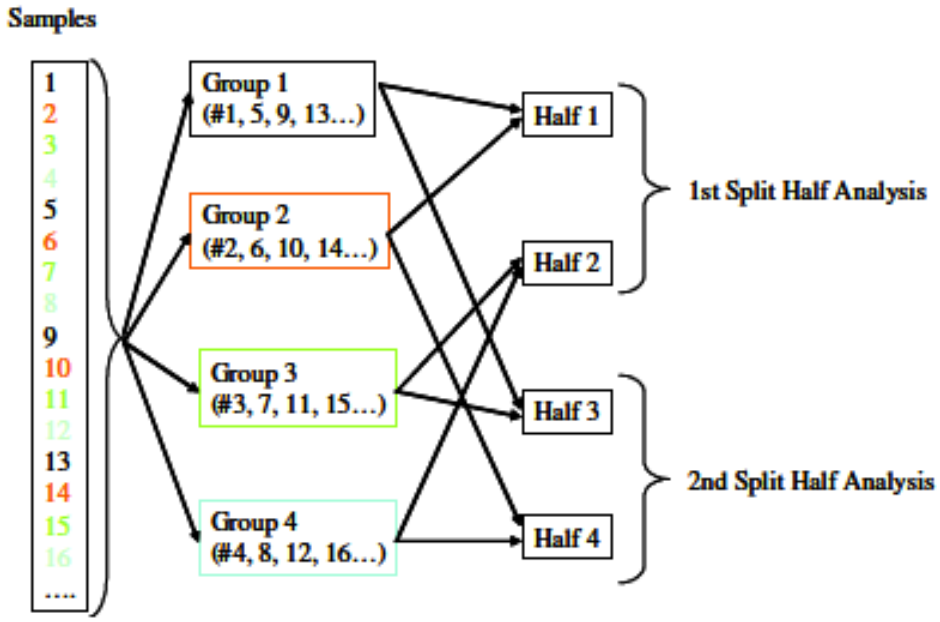

Figure 9: Visual display of how the split half validation works. It shows that each of the split halves includes all of the data for the model (Stedmon and Bro tutorial)

The following figure (10) shows the actual split of the data.

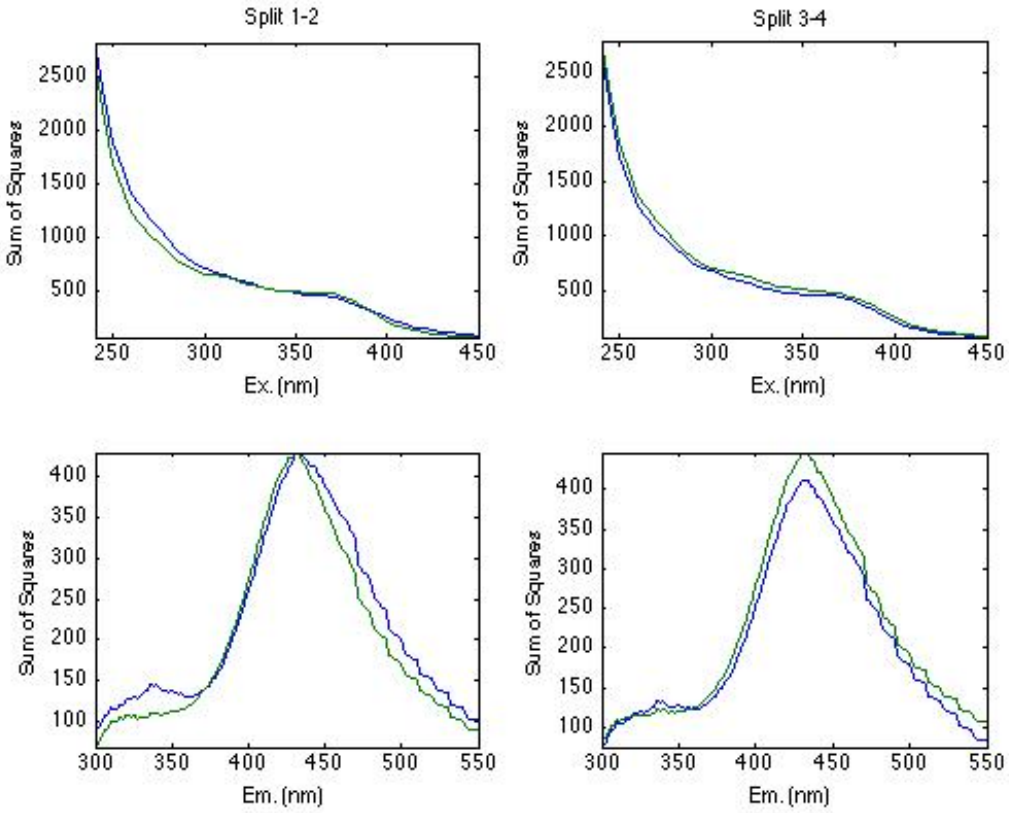

Figure 10: Shows the sum of squares for excitation and emission for the data used in both split half 1-2 and split half 3-4

The final step was to fit a series of models to the data using random initialization to make sure that the final version of the model is the least squares result. Ten different models were fit to determine which one had the least squares fit.

The next step in the modeling process was to add the identified microbial component back into the Cory and McKnight PARAFAC model. This step was done using the loading factors and adjusting the original Cory and McKnight matlab code to run a 14-component model. The data from all of the samples needed to be cut to

fit in the Cory and McKnight range because the new model's excitation and emission ranges are larger.

4. Results and Discussion

4.1 EEM Description

The following figure (Figure 11) shows an example EEM for each of the 5 samples sites. Fig. 11(a) is a sample from Boulder Creek incubations; it has a tail that stretches from em 500-600 nm and ex 380-450 nm. Fig. 11(b) is a sample from the Bangladesh incubations; it has high fluorescence in the amino acid range along with a peak located at em 480 nm and ex 360 nm that is associated with microbial activity. Fig. 11(c) is a sample from Lake Bonney in Antartica; it has very little humic fluorescence and a peak located at em 480nm and ex 410 nm that is similar to ones seen in the Boulder Creek samples. Fig. 11(d) is a sample from the Sierra Nevadas in Spain; it has high amino acid like fluorescence and a peak extending outward starting at em 440 nm and ex 380-450 similar to that of the Bangladesh sample. Fig. 11(e) is a sample from Cotton Glacier in Antarctica; it has very little humic fluorescence and the same peak extending outward from em 450 nm and ex 380-450 as both the Bangladesh and Sierra Nevada Lake samples.

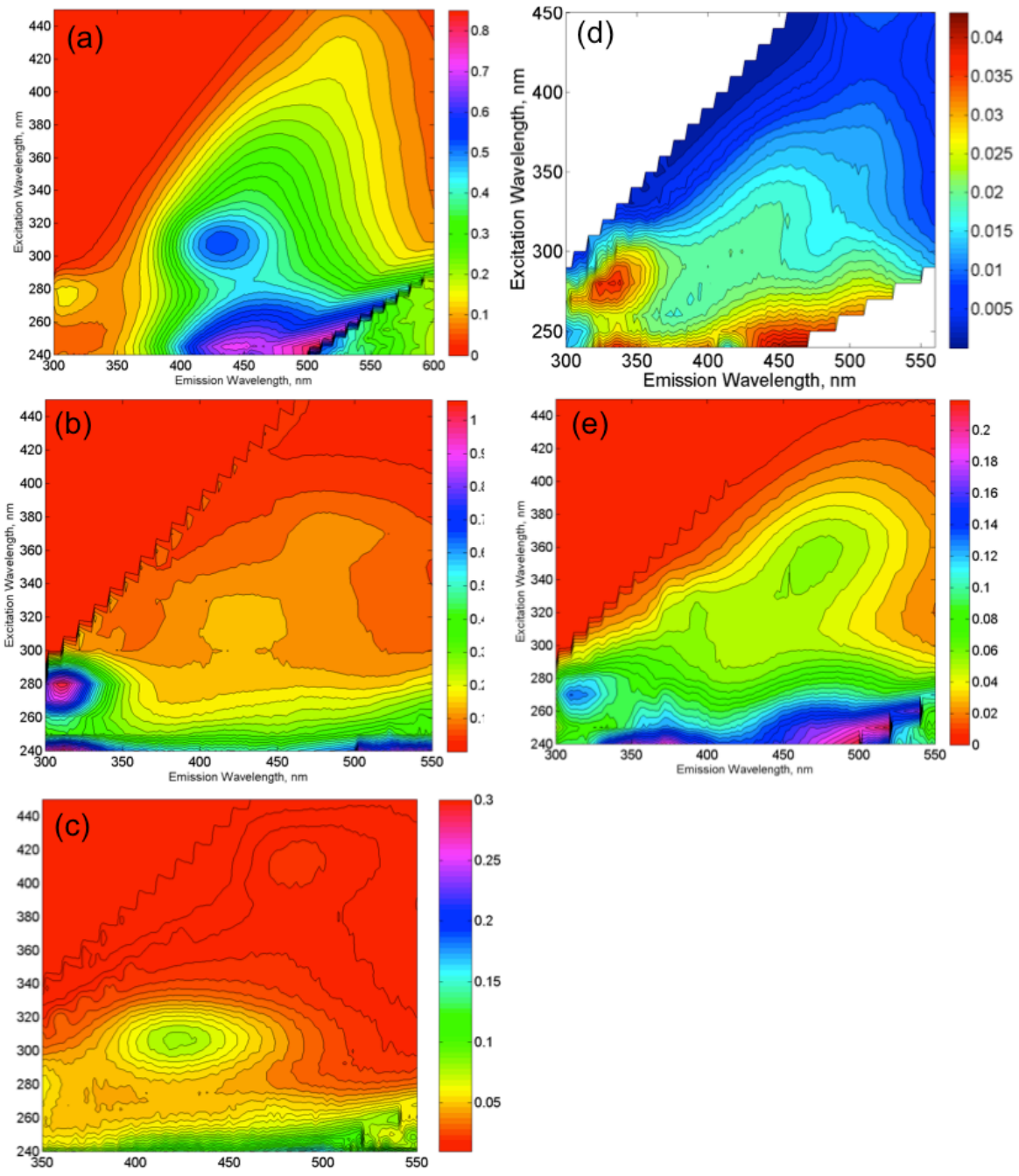

Figure 11: Sample EEMs from all 5 sites. (a) Boulder Creek, (b) Bangladesh, (c) Lake Bonney, (d) Sierra Nevada, (e) Cotton Glacier

4.2 Model Results

The application of the modeling criteria of Stedmon and Bro resulted in a model with 5 validated components. Figure 12 shows the results of evaluating the model using both a 2-component approach and with the final validated 5-component model. With the 2-component approach there are still large residuals. The 5-component results (Figure 13) show an overall good fit to the sample with an overall noisy residual that has no distinct peaks.

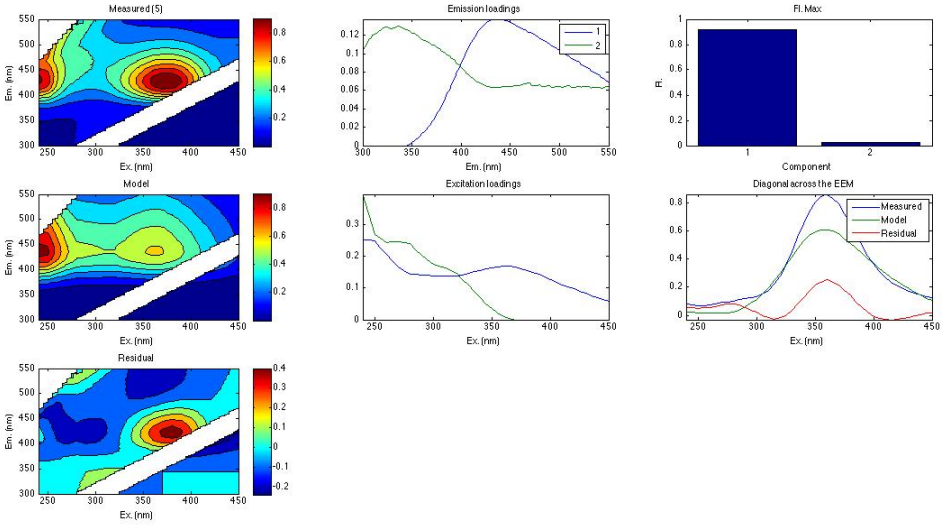

Figure 12: Modeled sample using a 2-component approach. There is a large residual at ex375 em425 showing that only having 2 components is not an effective model.

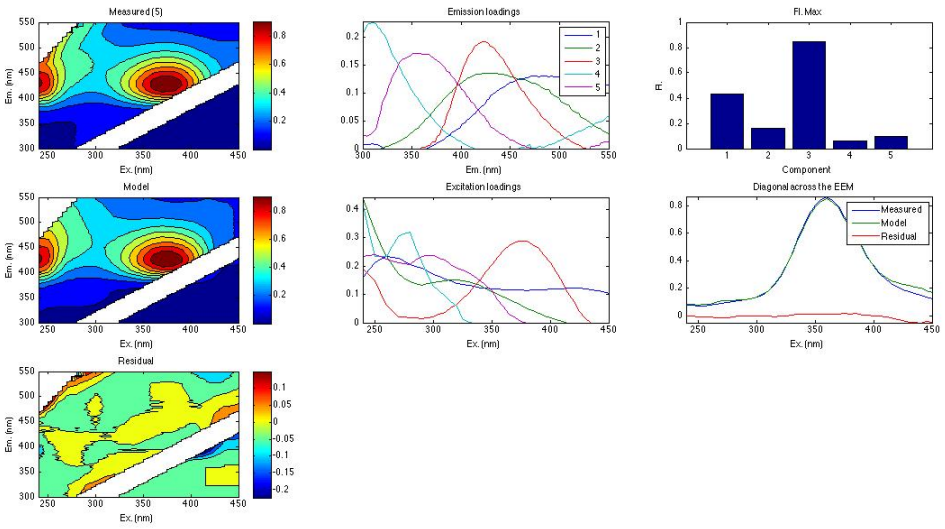

Figure 13: Modeled sample using 5 components. There are overall no peaks in the residual; it has a noisy appearance showing that the sample is being modeled effectively.

The following figure (Figure 14) shows the results of fitting 10 different models to the data to determine the least squares fit. The 1st model had the least squares fit and that is the one that is used.

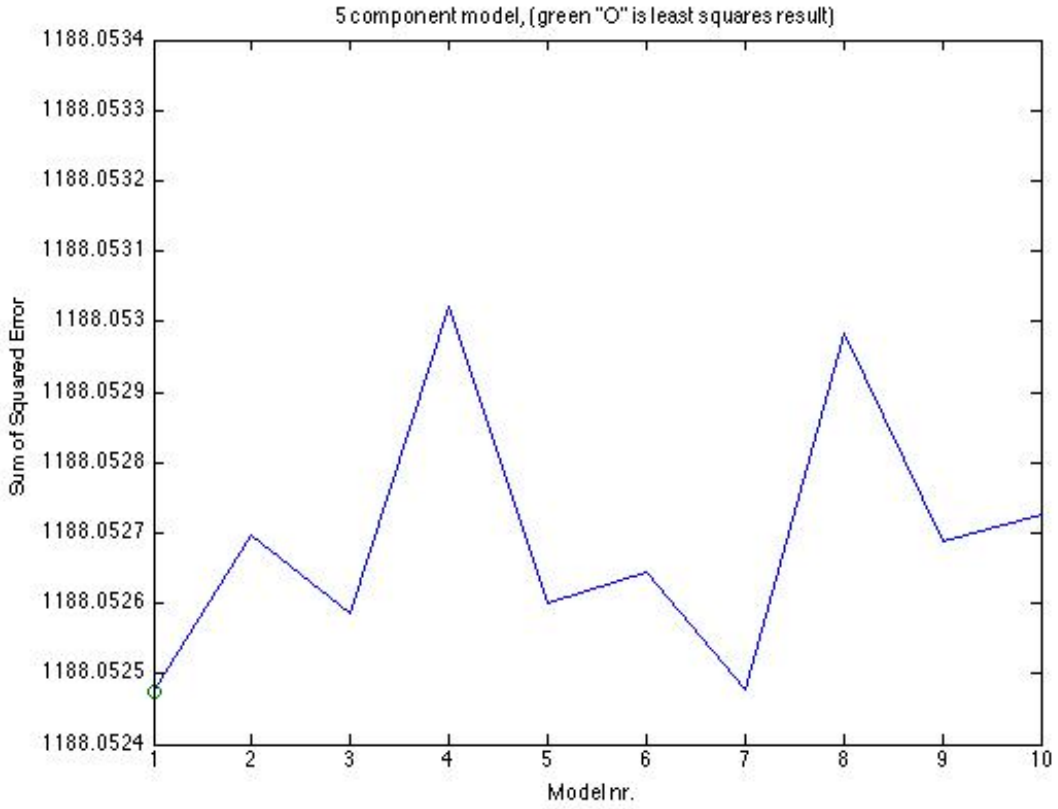

Figure 14: This figure shows the random initialization of the 5-component model. The 1st model had the least squares fit so that is the model that is used.

The resulting 5 components of the model are shown in the figure below (Figure 15). Component one (C1) represents one of the microbial components that the study was designed to describe. C1 has one emission maximum at 500 nm and two excitation maxima located at 250 nm and 410 nm. C1 is not similar to any of the thirteen components in the Cory and McKnight model.

Component two (C2) is the peak associated with the microbial humic substances. C2 also has one emission maximum and two excitation maxima occurring at em 440 nm and ex 240 nm and 310 nm. Component three (C3) is most likely also from some microbial sources and similar to the reduced quinone-like

components SQ2 and SQ3 from the Cory and McKnight model. C3 has two excitation maximum and one emission maximum. The emission maximum occurs at 420 nm and the most prominent excitation maximum occurs at 380 nm, while the other excitation maximum may be truncated at the lowest excitation of the measured EEM.

Component four (C4) is one that mostly occurred in the incubation samples from the Bangladesh incubation experiments, and is also beyond the range of the Cory and McKnight model. This component is most likely a product from the microbial growth incubation process. C4 has one excitation and emission maximum that occurs at em 310 and ex 275 nm.

The 5th and final component (C5) that the model identified is the amino acid-like fluorescence present in many of the samples. This component is a combination of the fluorescence that comes from both tyrosine-like and tryptophan-like fluorophores. C5 has one emission maximum and two excitation maxima. The emission maximum occurs at 360 nm and the excitation maxima occur at 245 nm and 295 nm.

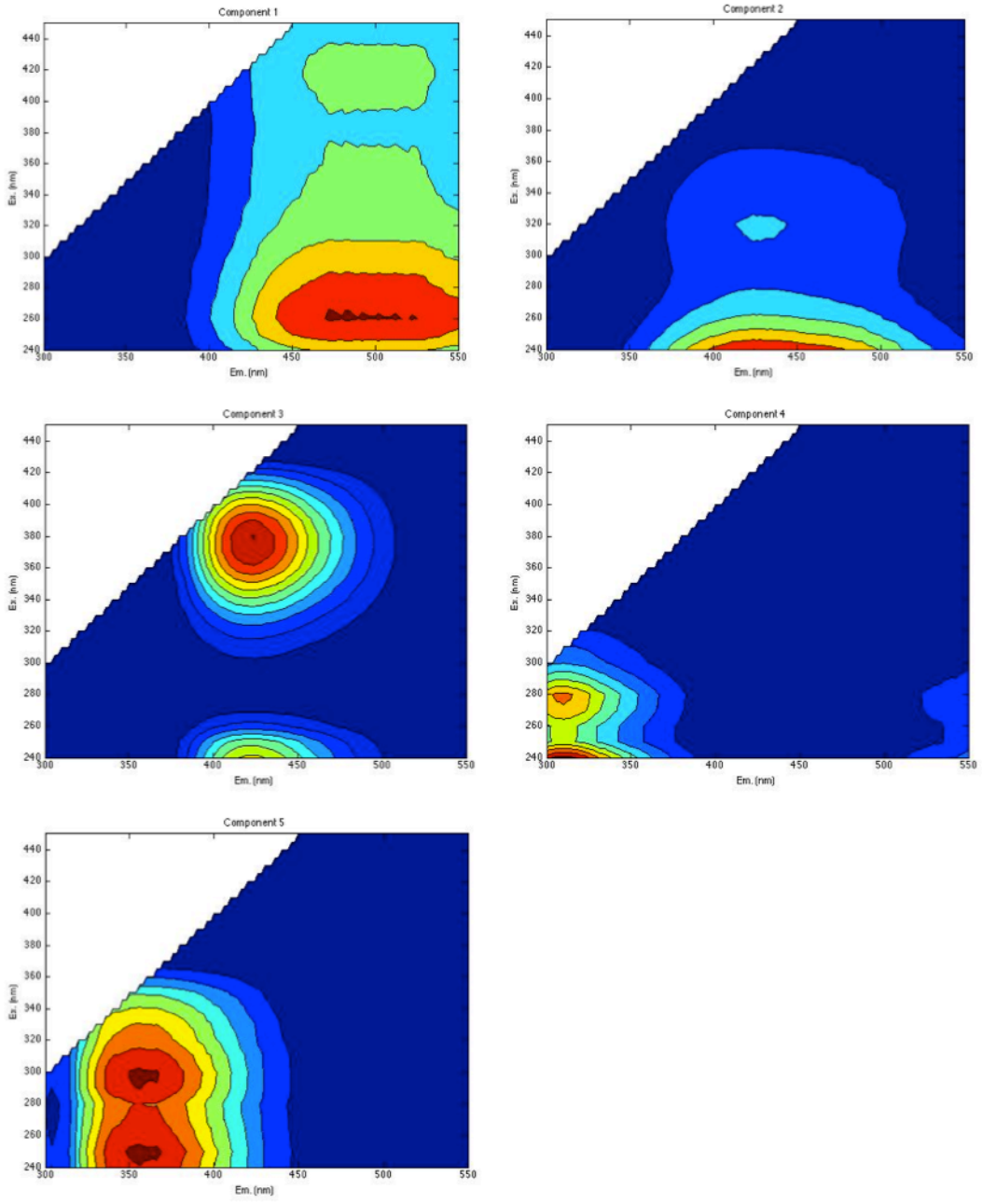

Figure 15: The resulting 5 components from the new microbial model

4.3 Modeling the EEMS

All samples used to create the model were then analyzed using the new 5-component model, the original Cory and McKnight (CM) model, and then the new Cory and McKnight plus (CM+) model. This analysis was done to compare how the residuals appear and change through the various models.

4.3.1 5 Component Model

The following two figures shows the same 5 samples from above being modeled by the new 5-component model; figure 16 presents the incubation samples and figure 17 presents the low humic and high microbial activity sites. The incubation experiment EEMs in Fig. 16(a) is a sample from Boulder Creek and Fig. 16(b) is from Bangladesh. Fig. 17(c) is from Lake Bonney, Fig. 17(d) is from Cotton Glacier, and Fig. 17(e) is from the Sierra Nevada mountains. The top row represents the original EEM, the 2nd row is the modeled EEM, the 3rd row is the residual plot, and the bottom row is a pie chart of the component make up for that sample.

For the Boulder Creek incubation experiments, the modeled version (row 2) looks very similar to the original EEM (row 1) with no noticeable peaks of high residual. The component with the greatest loading was C1, which is the microbial component not represented well in the Cory and McKnight model. For Bangladesh the modeled version also looks very similar to the original EEM. There is a slightly high residual located at em 350 nm and ex 275 nm. The main component for this EEM was C4, which was a product from the incubation process.

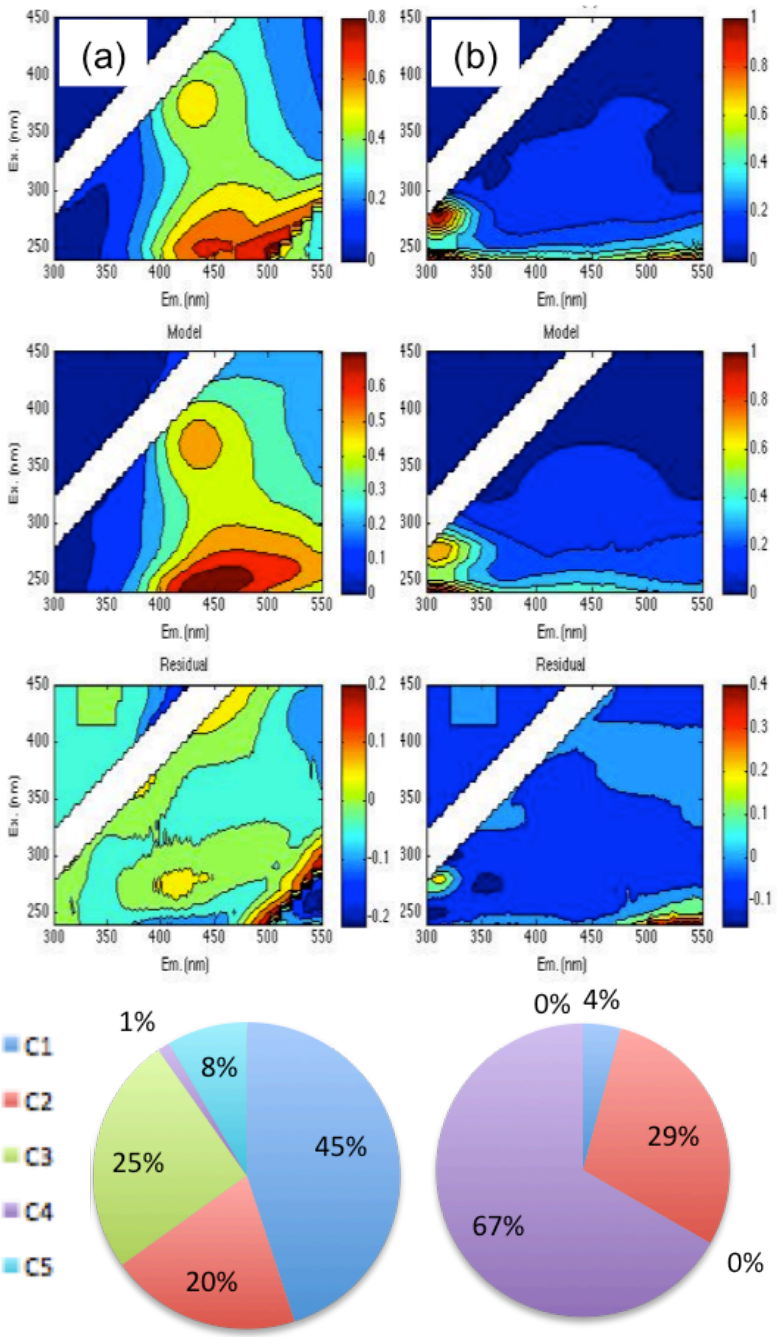

Figure 16: 5-Component model results for microbial incubation sites (a) Boulder Creek and (b) Bangladesh. The top row represents the original EEM, the 2nd row is the modeled EEM, the 3rd row is the residual plot, and the bottom row is a pie chart of the component make up for that sample

For Lake Bonney the modeled EEM is similar to that of the original EEM (Fig. 17 (c). There is more variation in the modeled version, but this is due to different scaling. This variation is more apparent when looking at the residual plot which is noisy with no noticeable peaks. The main component in that sample is C1, the microbial component.

For Cotton Glacier (Fig. 17 (d)) the modeled EEM is similar to the original EEM with an overall noisy residual. The high points near em 500 nm and ex 240-250 nm are due to scattering. The main component in the Cotton Glacier EEM is C4 which is also found in the Bangladesh incubation samples, but is out of the range of the Cory and McKnight 2005 model.

Lastly is the Sierra Nevada sample (Fig. 17 (e). Although it looks like the model is not working as well as for the other samples, there are different scales on the two images and the overall value of the residual is relatively low. The main component in this sample is also C4.

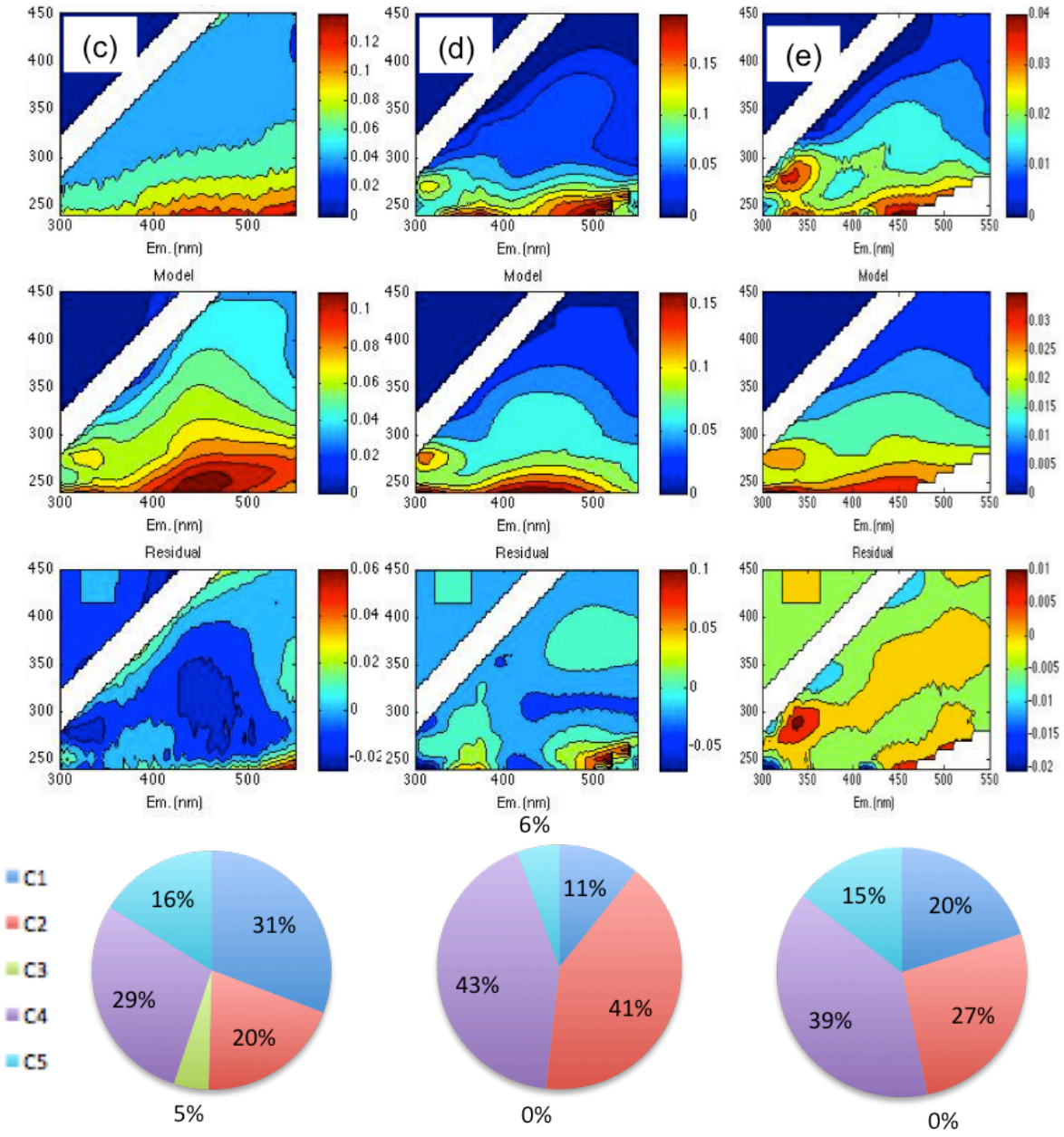

Figure 17: 5-Component model results for microbial incubation sites (c) Lake Bonney, (d) Cotton Glacier, (e) Sierra Nevada. The top row represents the original EEM, the 2nd row is the modeled EEM, the 3rd row is the residual plot, and the bottom row is a pie chart of the component make up for that sampl

4.3.2 Cory and McKnight Plus (CM+)

Component 1 from the 5-component model was then added back into the original Cory and McKnight model to create a new 14-component model. Component 1 was chosen because it was a new microbial component that occurred in all of the samples; it was thought that this would alleviate some of the high residuals seen in the results from Cory and McKnight. This method was done following the guidelines of Cawley et al 2012. The sample set from Cawley et al was from the Gulf of Maine, Penobscot Bay, Penobscot River, and the Androscoggin River in Maine. When modeled with Cory and McKnight a high residual was noted at em 436 nm ex 275 nm, so a new model was created to identify this component. The component was then added back into the original Cory and McKnight model and drastically reduced the observed high residual region. The following figure shows a comparison of the data found in Cawley et al 2012. In the case of Cawley et al there was a definite source to this new component; the residual was caused by pulp mill effluent. The fluorescence signature of the pulp mill effluent was equivalent to that of the new component.

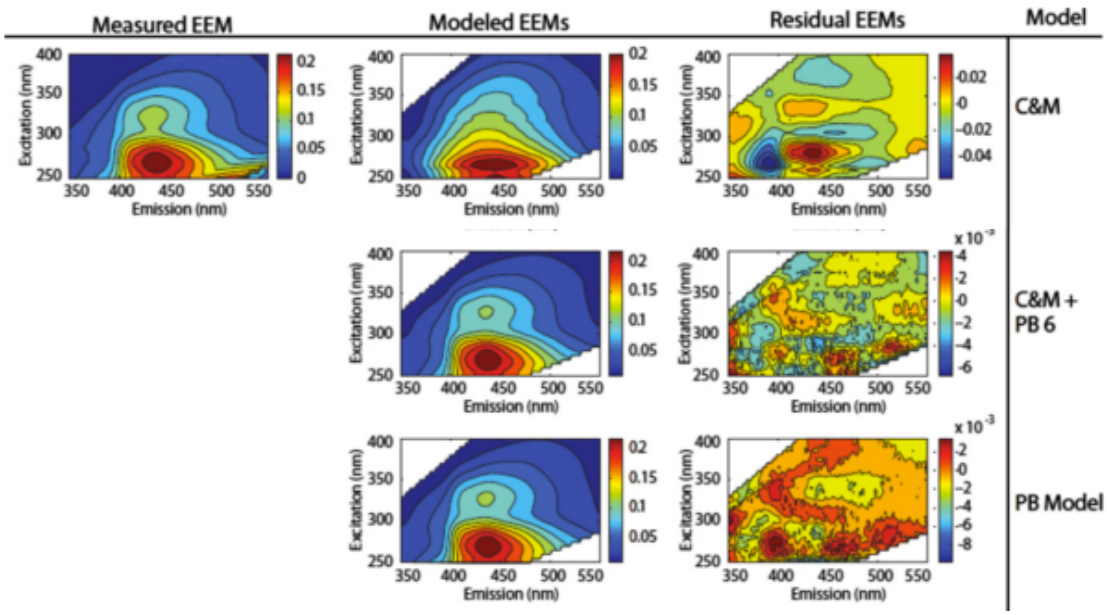

Figure 18: From Cawley et al 2012, shows the differences in Modeled and Residual EEMs using the original Cory and McKnight (C&M) model, the Cory and McKnight model plus the high residual component, 6 (C&M+PB6), and the Penobscot Bay model (PB Model).

The following figure (19) shows the results from C1 being added into Cory and McKnight. In Fig. 19 the left column presents the new model results, the center column presents the Cory and McKnight (CM) results, and the right column represents the results for the Cory and McKnight (CM+) with the C1 component added. Looking at the 5-component model one can notice that the modeled EEM looks very much like the original EEM and that the residuals are low and noisy. Looking at the CM modeled version it also does a good job of modeling the EEM. It places the peak in the right place but the peak is more rounded than in the original EEM, the residual plot is overall low and noisy as well. Looking at the CM+ version of the modeled EEM it seems to capture the intensity of the peak better but otherwise does not quite look right. There is an indent around ex 270 nm that should not be

there; it disconnects the two peaks making it look unnatural. This leads one to believe that there is something different going on in this system vs. the Penobscot Bay system example from Cawley et al. The added component, C1, could also be interfering with a component that is already in the Cory and McKnight model (e.g. SQ1), which would cause this strange appearance. This means that an accurate redox index is not going to be able to be determined for these microbial samples as of right now.

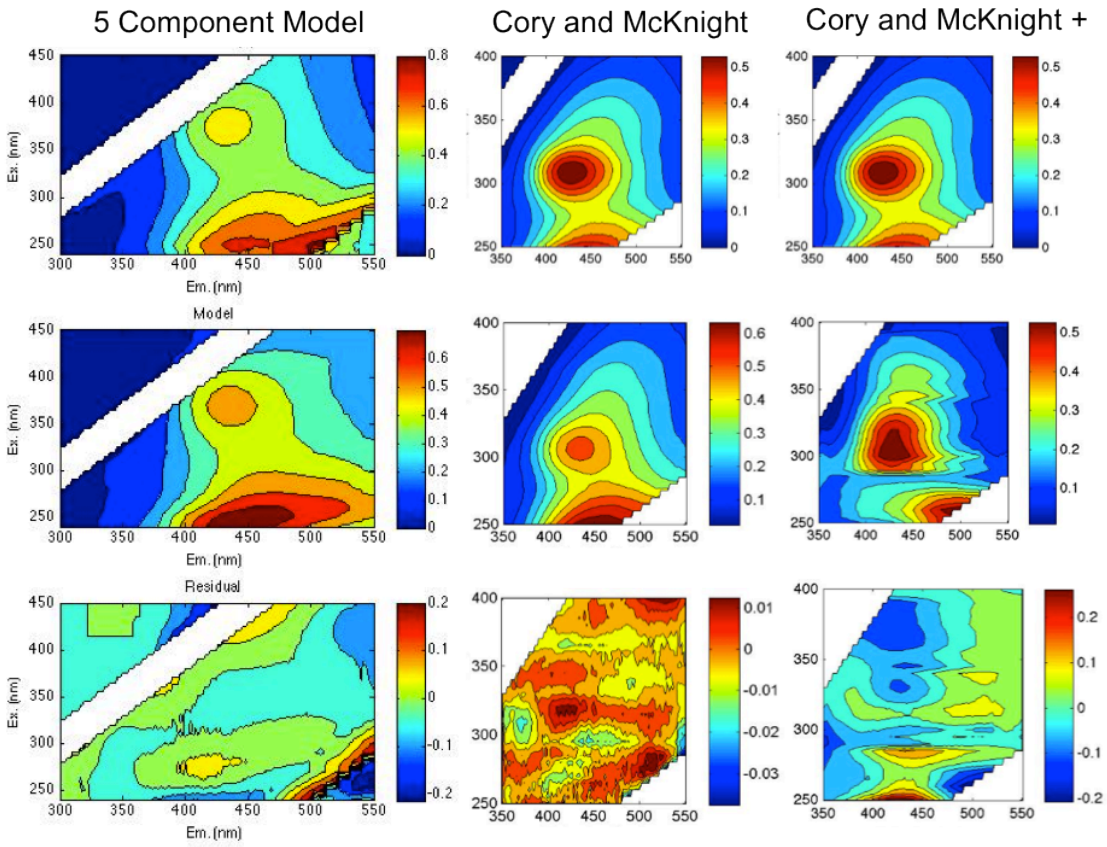

Figure 19: Comparison of all model results for a Boulder Creek sample 0384. The right hand column represents the new 5-component model, the center column is the Cory and McKnight model, and the right hand column is the Cory and McKnight plus model. The top row represents the original EEM, the middle row is the modeled data, and the bottom row is the model residuals.

5. Conclusions

In this study I was able to create a new model for DOM fluorescence based upon EEMs from samples that had high microbial inputs, whether those are from microbial incubation experiments or from sites that have little to no terrestrial inputs. This new model identified several new components that are most likely microbial in origin. This result shows that while Cory and McKnight is an excellent starting point to large and varied data sets there are benefits to creating a system specific model. System specific models are able to catch nuances that other, broader, models would not be able to identify. When attempting to add a microbial component back into Cory and McKnight the data did not model correctly, this means that there has to be something more complicated occurring here than the example study of Cawley et al 2012. It was also shown that it is possible to study microbial activity from a fluorescence standpoint and large amounts of information can be gathered from this, from identifying sites that have high microbial activity to even identifying what types of microbes might be in that region.

6. References

Coble, P., Green, S.A., Blough, N.V., Gagosian, R.B. (1990) Characterization of DOM in the Black Sea by fluorescence spectroscopy. *Nature.* 348, 432-435

Cory, R.M., McKnight, D.M. (2005) Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in DOM. *Environmental Science and Technology.* 39, 8142-8149

Fellman, Jason B., Hood, Eran, Spencer, Robert G.M. (2010) Fluorescence spectroscopy opens new windows into dissolved organic matter dynamics in freshwater ecosystems: A review. *Limnology and Oceanography.* 55, 2452-2462

Foreman, C. M., R. M. Cory, C. E. Morris, M. D. SanClements, H. J. Smith, J. T. Lisle, P. L. Miller, Y.-P. Chin, and D. M. McKnight. Microbial growth under humic-free conditions in a supraglacial stream system on the Cotton Glacier, Antarctica. *Geomicrobiology.* (In Review)

Immerzeel, W. (2008). Historical trends and future predictions of climate variability in the Brahmaputra basin. *International Journal of Climatology.* 28, 243-254.

Klapper, L., McKnight, D.M., Fulton, J.R., Blunt Harris, E.L., Nevin, K.P., Lovely, D.R., Hatcher, P.G. (2002) Fulvic acid oxidation state detection using fluorescence spectroscopy. *Environmental Science and Technology.* 36(14), 3170-3175

Lovely, D.R., Frage, J.L., Coates, J.D., Blunt-Harris, E.L. (1999) Humics as an electron donor for anaerobic respiration. *Environmental Microbiology.* 1, 89-98

McKnight, D.M., Boyer, E.W., Westerhoff, P.K., Doran, P.T., Kulbe, T., Andersen, D.T. (2001) Spectrofluorometric characterization of DOM for indication of precursor material and aromaticity. *Limnology and Oceanography.* 46, 38-48

Miller, M.P., McKnight, D.M., Cory, R.M., Williams, M.W., Runkel, R.L. (2006) Hyporheic exchange and fluvic acid redox reactions in an alpine stream/wetland ecosystem, Colorado Front Range. *Environmental Science and Technology.* 40, 5943-5949

Nevin, K., Lovely, D. (2002) Mechanisms for Fe(III) oxide reduction in sedimentary environments. *Geomicrobiologia*. J. 19, 141–159.

Lyons, W. B., Welch, K. A., Snyder, G., Olesik, J., Graham, E. Y., Marion, G. M., Poreda, R. J. (2005) Halogen geochemistry of the McMurdo dry valleys lakes, Antarctica: Clues to the origin of solutes and lake evolution. *Geochimica et Cosmochimica Acta* 69: 305-323.

Mladenov N., Sommaruga, R., Morales-Baquero, R., Laurion, I., Camarero, L., Dieguez, M.C., Camacho, A., Delgado, A., Torres, O., Chen, Z., Felip, M., Reche, I. (2011) Dust inputs and bacteria influence dissolved organic matter in clear alpine lakes. *Nature Communications.*

Ohno, T. 2002 Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. *Environmental Science and Technology.* 36, 742-746.

Ratasuk, Nopawan, Nanny, Mark A. (2007) Characterization and quantification of reversible redox sites in humic substances. *Environmental Science and Technology.* 41, 7844-7850

Scott, Durelle T., McKnight, Diane M., Blunt-Harris, Elizabeth L., Kolesar, Sarah E., Lovley, Derek R. (1998) Quinone moieties act as electron acceptors in the reduction of humic substances by humics-reducing microorganisms. *Environmental Science and Technology.* 32, 2984-2989.

Spigel, R.H., and Priscu, J.C. (1996) Evolution of temperature and salt structure of Lake Bonney, a chemically stratified Antarctic lake. *Hydrobiologia*, 321, 177-190.

Stedmon, Colin A., Bro, Rasmus. (2008) Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. *Limnology and Oceanography Methods.* 6, 572-579.

Stute, M., Zheng, Y., Schlosser, P., Horneman, A., Dhar, R., Datta, S., Hoque, M., Seddique, A. A., Shamsudduha, M., Ahmed, K. M., van Geen, A. (2007) Hydrological control of As concentrations in Bangladesh groundwater. *Water Resources Research.* 43, W09417