Excitation Emission Matrix Fluorescence Spectroscopy based Sensing of Combustion Generated Particulate Matter

Gaurav Mahamuni

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2020

Reading Committee:
Igor Novosselov, Chair
Gregory Korshin
John Kramlich

Program Authorized to Offer Degree:
Mechanical Engineering

University of Washington

Excitation Emission Matrix Fluorescence Spectroscopy based Sensing of Combustion Generated Particulate Matter

Gaurav Mahamuni

Abstract

Human exposure to particulate matter (PM) is associated with adverse health effects, including cardiopulmonary diseases, neurological diseases, and lung cancer. PM can originate from mobile and industrial combustion sources, forest fires, domestic fuel burning, and other natural (soil dust, sea salt, etc.) and anthropogenic sources. Within the human respiratory tract, PM size determines the region of deposition, residence time, solubility, and tissue uptake. The chemical composition of the particle determines the potential for biochemical reaction with tissues and cells. The effect of PM exposure on health outcomes is a critically important research area. Miniaturized exposure monitoring devices and flexible orthogonal *in-situ* analysis of samples would greatly enhance the ability to clarify the relationships between PM exposure and health. Real time knowledge of these relationships can lead to improved health risk assessment due to PM exposure, guidance for the management of diseases, and targeted intervention strategies to reduce health risk. This work addresses the need for improved exposure assessment, quantification, and characterization of ultrafine particles in the environment. The anticipated outcome of this research is the development of a real-time analysis tool that will collect airborne nanoparticles and provide *in-situ* analysis of PM chemical composition—which will determine their toxic potential using spectroscopic techniques.

Chemical characterization of PM is critical for the assessment of the hazard associated with human exposure to potentially harmful agents, notably combustion-generated PM. Specifically, polycyclic aromatic hydrocarbons (PAHs) found in combustion PM have been associated with carcinogenic and mutagenic effects. Chapter 1 introduces the key concepts useful for development of the PM analysis tool. In Chapter 2, we analyse PM chemical composition from various sources for (1) organic carbon content, (2) the presence and concentrations of PAHs with low molecular weight (LMW, $126<MW<202$), which may exhibit acute toxicity, and (3) higher molecular weight PAHs (HMW, $226<MW<302$) which may have carcinogenic impacts. We use combustion-generated PM from both laboratory sources as well as real-world sources (e.g., woodsmoke and diesel exhaust PM). The colder flames result in lower PM yields; however, fraction of PAHs constituting the PM increases significantly. For hotter flames, PM yields are higher with very low fractions of PAHs constituting the PM.

In Chapter 3, we applied the Principal Component Regression – Excitation Emission Matrix (PCR-EEM) method to predict LMW and HMW PAH concentrations as well as individual PAH concentrations in combustion generated PM from 16 reference PAHs[1]. The reference PAH concentration are obtained using Gas Chromatography Mass Spectroscopy (GCMS). The models are trained on samples from the UW inverted gravity flame reactor (IGFR)[2–5] and tested on PM emissions from biomass cookstove combustion and diesel exhaust samples accurately predicts HMW PAH concentrations with $R^2 = 0.976$ and overestimates LMW PAHs.

In Chapter 4, we present the design and evaluation of am electrostatic precipitation based capillary-based sensor that collects fine and ultrafine particulate matter onto the outer surface of a capillary for *in-situ* spectroscopic analysis. At optimized operating conditions, the capillary collection efficiency is relatively constant at 65 % for all particles in the 0.2 µm - 2 µm range while it drops below 50 % for the larger particles, likely due to a lower charge-to-mass ratio. Fluorescence spectroscopy is used to demonstrate the integration of the capillary collector with *in-situ* spectroscopic analysis techniques.

Automation of fluorescence analysis is required for in-situ spectroscopic analysis of PM. However, solvent extraction as a preprocessing step and the need for benchtop instrumentation hinders development of a miniaturized sensor. In Chapter 5, we present a methodology that eliminates labor-intensive sample preparation and allows to miniaturize the detection platform. We describe the development of a solid phase excitation emission matrix (SP-EEM) analysis method that eliminates pre-processing steps and reduces the cost of the standard liquid phase excitation emission matrix (LP-EEM) analysis (the experimental procedure followed in Chapter 3 to obtain EEMs of combustion generated PM). We evaluated external and internal excitation schemes for their ability to produce the best SP-EEM signatures with lowest scattering and reflection interference. Analysis of woodsmoke and cigarette smoke samples showed good agreement with LP-EEMs. In internal excitation, SP-EEM spectra shows fluorescent interference from PDMS for wavelengths < 250nm. The external excitation EEM spectra are dependent on the incident angle; ranges of 30-40º and 55-65º showed the best results. SP-EEM technique can be used for development of miniaturized sensors for chemical analysis of combustion generated PM.

In summary, this thesis: (i) provides a statistical model for determining concentrations of potentially carcinogenic PAHs in combustion generated PM using EEM fluorescence spectroscopy; (ii) includes design and evaluation of a miniaturized collection and fluorescence analysis setup for PM2.5; and (iii) shows the feasibility of SP-EEM analysis for the development of miniaturized sensors for *in-situ* chemical analysis of combustion generated PM.

Acknowledgements

I would like to thank my advisor Dr. Igor Novosselov for constant guidance and support. Thank you for providing the freedom to follow my interests and providing the necessary help whenever needed. I also want to thank Dr. Gregory Korshin, Dr. John Kramlich, Dr. Jonathan Posner, Dr. Edmond Seto and Dr. Arka Majumdar for their valuable feedback on my research.

This work would not have been possible without the constant support of Jay Rutherford, Justin Davis, Jiayang He and Byron Ockerman. Thank you for being the best people to discuss ideas with and work in collaboration on projects with throughout graduate school. I would also like to thank Eric Molnar and Saurabh Gupta for working with me.

I have had a great work environment in the Novosselov Research Group and I would like to thank all my fellow lab mates. Lastly, I would like to thank my parents, sisters, and friends for providing constant support throughout the PhD program.

Table of Contents

List of Figures

List of Abbreviations

PM	particulate matter
UPM	ultrafine particulate matter
PAHs	polycyclic aromatic hydrocarbons
LMW	low molecular weight
HMW	high molecular weight
IGFR	inverted gravity flame reactor
PCR	principal component analysis
EEM	excitation emission matrix
SP-EEM	solid phase EEM
LP-EEM	liquid phase EEM
TOC	total organic carbon
GCMS	gas chromatography mass spectroscopy
NP	nanoparticles
PM2.5	particulate matter having an aerodynamic diameter less than 2.5 μm
PM10	particulate matter having an aerodynamic diameter less than 10 μm
PM0.1	particulate matter having an aerodynamic diameter less than 100 nm
MW	molecular weight
WHO	World Health Organization
SVOCs	semi volatile organic compounds
VOCs	volatile organic compounds
VVOCs	very volatile organic compounds

EPA	The Environmental Protection Agency
IARC	International Agency for Research on Cancer
BC	black carbon
OC	organic carbon
BrC	brown carbon
UV-LIF	ultra violet laser-induced fluorescence
OBG	optical band gap
PCA	principal component analysis
TEOM	tapered element oscillating microbalances
BAM	beta attenuation monitors
APS	aerodynamic particle sizers
ESP	electrostatic precipitator
PAS	photoacoustic spectroscopy
CRDS	cavity ringdown spectroscopy
IBBCEAS	incoherent broadband cavity-enhanced absorption spectroscopy
SP2	single-particle soot photometer
WIBS	wideband integrated bioaerosol sensor
PDMS	Polydimethylsiloxane
RH	relative humidity
NIOSH	National Institute for Occupational Safety and Health
NMAM	NIOSH Manual of Analytical Methods
EC	elemental carbon
TC	total carbon

PMRM	pseudo multiple reaction monitoring
Ar	argon
PCs	principal components
MLR	multiple linear regression
MSE	mean squared error
LOOCV	leave one out cross-validation
BFL	Benzofluoranthenes
INP	Indeno-123cd-Pyrene
TC	Total Collection
CC	Capillary Collection
PSL	polystyrene latex
FCC	fluorescence-based capillary collection
I	bulk fluorescence intensity

List of Symbols

T_{max}	maximum flame temperature
T^*	highest soot luminosity region temperature
$T_{max,c}$	maximum flame temperature – critical
T^*_c	highest soot luminosity region temperature – critical
τ_{res}	flame residence time
φ	global equivalence ratio
T_{TC}	thermocouple temperature
Nu	Nusselt number
k	thermal conductivity
T	flame temperature
ε	emissivity
Pt	platinum
Rh	rhodium
ε_{TC}	ε of the thermocouple
T_{cor}	corrected flame temperature
T_{carb}	carbonization temperature
d	number of PCs
C	data matrix
E	residual matrix
s	scores matrix
v	loadings matrix

n	total number of samples
C_{test}	test data matrix
R^2	coefficient of determination
m	number of test samples
λ_{ex}	emission wavelength
λ_{em}	excitation wavelength
φ_C	corona voltage
φ_R	repelling voltage
$TC_{N,M}$	TC at fixed $\varphi_C = N$ (kV) and $\varphi_R = M$ (kV)
$APS_{N,M}$	concentration of particles measured by the APS at $\varphi_C = N$ (kV) and $\varphi_R = M$ (kV)
$CC_{N,M}$	CC at $\varphi_C = N$ (kV) and $\varphi_R = M$ (kV)
θ	angle between the incident excitation light and the substrate

Chapter 1 Introduction

1.1 Background

Particulate matter (PM) is a pollutant that is emitted from agricultural processes, mobile and industrial combustion sources, forest fires, and other natural and anthropogenic sources. Environmental and occupational exposure to PM is linked to adverse health impacts. PM2.5 (particles having an aerodynamic diameter less than 2.5 μm) associated with combustion products cause significant lung inflammation and other adverse respiratory effects.[6–12] Cardiovascular diseases are also suspected to be related to environmental exposure.[13,14] PM pollution exposure was shown to be responsible for 4.2 million deaths in the year 2016.[15] Around 45%-60% of urban PM2.5 exposure and 40%-58% of urban PM10 exposure around the world is attributed to combustion generated PM.[16] The toxic potential of inhaled particles depends on particle size and chemical composition. Current exposure monitoring approaches do not, however, provide chemical analysis of collected samples, especially in field settings. Robust and low-cost strategies for sampling and analysis are needed to measure individual exposure to PM2.5. Ultimately, such improved assessment methods can enhance our understanding of the health effects of air pollutant mixtures, leading to improved human health risk assessment and guidance for the managing of intervention strategies to reduce disease.

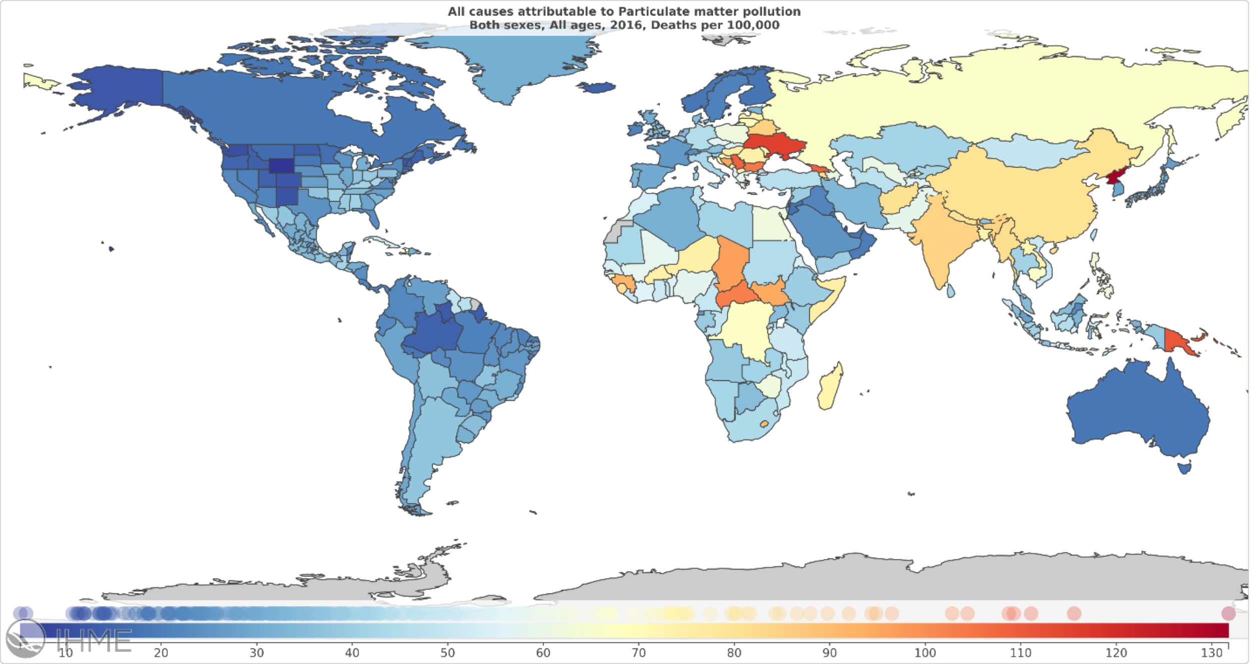
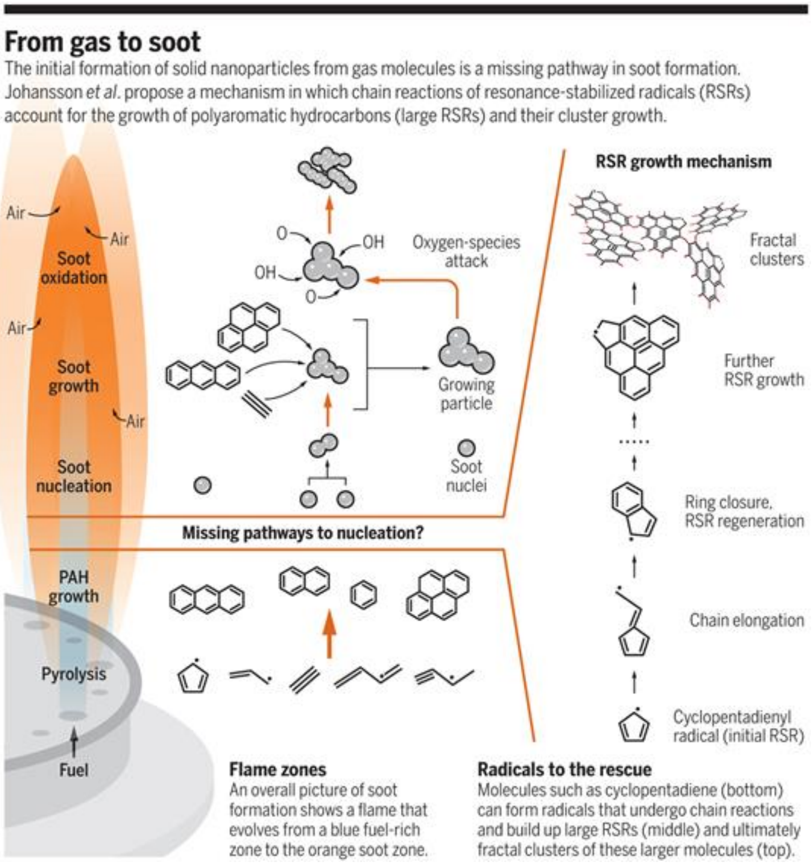


Figure 1.1: Deaths per 100,000 due to particulate matter pollution around the world. In the year 2016, ambient air pollution was responsible for 4.2 million deaths, from ref [17]

1.2 PAHs in Combustion Aerosols

Particulate matter with aerodynamic diameter less than 2.5 μm is referred to as PM2.5 and ultrafine particles (also called Nanoparticles - NP) which have aerodynamic diameter less than 100 nm are referred to as PM0.1. These PM can have varying chemical composition and structure depending on the emission source, and can, thus, have differing impacts on air quality and health effects.[18] The PM structure from different types of combustion scenarios have similarities based on the structures and chemical composition. Several mechanisms to describe the formation and growth of the particles have been proposed.[19,20] Soot formation during combustion can be broken down into several steps. First, fuel pyrolysis and entropic-driven chemical reactions form PAHs, which are stable in the flame environment. PAHs grow in mass with the addition of fuel. As PAHs grow in size, attractive Van der Waals forces overcome the dissociation forces in the flame environment and PAHs start forming clusters.[21] Previous numerical modeling studies use irreversible dimerization of pyrene for soot inception for condensed phase.[22,23]. However, other computational studies show that a larger PAH size is necessary for soot inception through dimerization.[24–27] Experimental evidence suggests PAH coalescence (formation of dimers, trimers, and clusters) and that incipient soot is formed due to the stacking of PAHs.[28,29]



From gas to soot

The initial formation of solid nanoparticles from gas molecules is a missing pathway in soot formation. Johansson *et al.* propose a mechanism in which chain reactions of resonance-stabilized radicals (RSRs) account for the growth of polyaromatic hydrocarbons (large RSRs) and their cluster growth.

Flame zones
An overall picture of soot formation shows a flame that evolves from a blue fuel-rich zone to the orange soot zone.

Radicals to the rescue
Molecules such as cyclopentadiene (bottom) can form radicals that undergo chain reactions and build up large RSRs (middle) and ultimately fractal clusters of these larger molecules (top).

Figure 1.2: Soot formation process and a proposed pathway for soot nucleation, from ref [30]

Polycyclic aromatic hydrocarbons (PAHs) are organic compounds having 2-7 fused aromatic rings. PAHs are released by both natural and anthropogenic sources such as forest fires, fuel and biomass burning, cigarette smoking, etc. PAHs were one of the first atmospheric pollutants linked with carcinogenic and mutagenic activity.[31–33] Physical properties of PAHs vary with molecular weight (MW) and structure. PAHs in the environment can exist in the vapor phase or adsorb into airborne PM depending on the atmospheric conditions (ambient temperature, relative humidity, etc.), the nature of the aerosol (i.e., origin and properties), and the properties of the individual PAH. [34,35]

The World Health Organization (WHO) classifies semi volatile organic compounds (SVOCs) as indoor organic pollutants with a boiling point range between 240/260 and 380/400 °C. They differ from volatile organic compounds (VOCs) and very volatile organic compounds (VVOCs) that have boiling point ranges of 50/100 to 240/260 °C and <0 to 50/100 °C, respectively.[36] In the United States, emissions of VOCs to the outdoors are regulated by Environmental Protection Agency (EPA) mostly to prevent the formation of ozone. However, among them, only some VOCs are considered "reactive" enough to be of concern. For indoor air quality, all organic chemical compounds whose compositions give them the potential to evaporate under normal atmospheric conditions are considered VOCs. Semivolatile organic compounds (SVOCs) are a subgroup of VOCs that tend to have a higher molecular weight than other VOCs. PAHs can be generally divided into two categories: low molecular weight (LMW) and high molecular weight (HMW) PAHs. Usually, LMW PAHs (two, three or four ring hydrocarbons) include PAHs which exist mainly in the gas phase and can be classified as VOCs while HMW PAHs exist primarily in the particulate phase and are considered SVOCs.

In combustion aerosol formation, the planar nature and inherent stability of aromatic compounds are linked to the formation of soot particles.[2] More than 100 PAHs cluster together to form liquid-like particles in flames with temperatures greater than 1350 K.[37] The size of PAHs in young soot has been approximated by molecular dynamics simulations to be between 4 rings and 19 rings.[26,38] Thus, in this research, we consider PAHs having a molecular weight less than or equal to the molecular weight of pyrene (MW 128-202 g/mol, ≤ four rings) as LMW PAHs. PAHs with a molecular weight greater than that of pyrene are HMW PAHs (MW 226-302 g/mol, ≥ four rings). HMW PAHs have significantly lower vapor pressures compared to LMW PAHs.

The Environmental Protection Agency (EPA) has established a panel of 16 PAH compounds as priority pollutants that represent a range of molecular structures with 128<MW<278 g/mol.[39] More recent reports[40] show that among these, HMW PAHs have high cancerogenic potential. The majority of these HMW PAHs are classified as probably or possibly carcinogenic (groups 2A and 2B, respectively), and benzo[a]pyrene is classified as carcinogenic (group 1) by the International Agency for Research on Cancer (IARC).[41] The lower molecular weight (LMW) PAHs are considered non-genotoxic, except naphthalene;[41,42] Although the LMW PAHs are less toxic compared to HMW PAHs, they are able to react with other pollutants (such as ozone, nitrogen oxides, and sulfur dioxide) to form diones, nitro- and dinitro-PAHs, and sulfuric acids, which can increase their toxicity.[35]

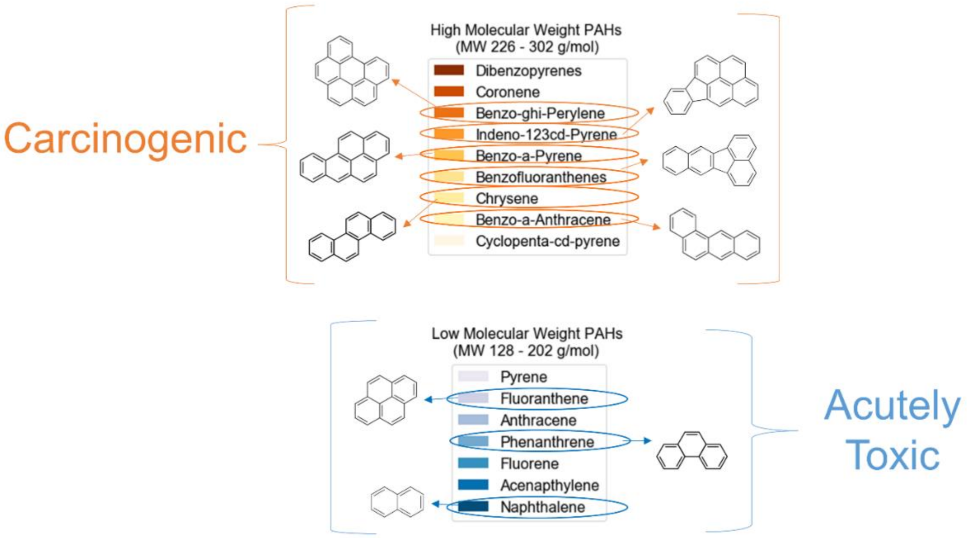


Figure 1.3: HMW PAHs are carcinogenic and found in higher percentage in particle phase compared to LMW PAHs.[43]

The two principal components of PM from combustion sources are elemental carbon (EC) and organic carbon (OC). PM with high concentration of EC is often called black carbon (BC), or those synthesized for industrial purposes, these particles are also called Carbon Black. PM with a high fraction of OC is also termed 'brown carbon' (BrC). The chemical species generated in flames span hundreds to thousands of organic components that make up a relatively small mass fraction of OC compared to BC. PAHs, their oxidized and nitrogenated byproducts as well as hydrocarbons with more aliphatic-like structures and higher oxygen content may constitute a fraction of these chemical species.[44] Though organic fraction in combustion PM contains a variety of complex hydrocarbon compounds[45], among them, PAHs have been reported to be a major cause of oxidative damage.[46–48]

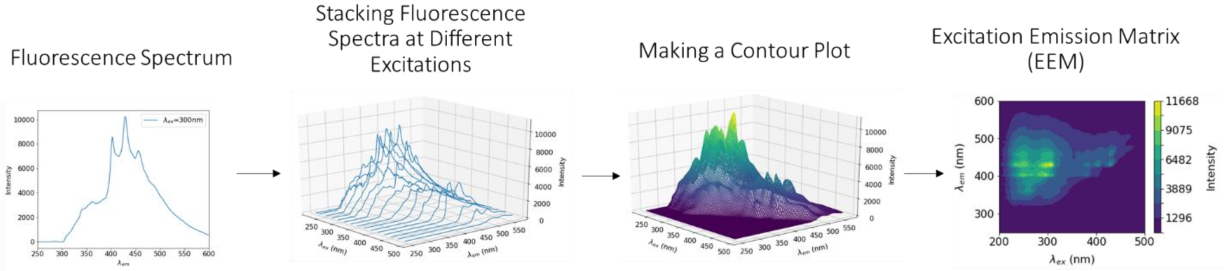
1.3 Excitation Emission Matrix Analysis of Combustion Generated Aerosols to Predict Carcinogenic and Non-carcinogenic Compounds

Chemical analysis of PM is routinely performed to determine total organic carbon (TOC) fraction in the sample[49]; TOC analysis can be complemented by more specific GCMS[50] or LCMS[51,52] to provide more detailed information about the organic fraction. The determination of chemical composition using these laboratory techniques involves PM collection on filters, followed by solvent extraction as pre-processing steps. This offline laboratory analysis is cumbersome. Furthermore, these techniques require expensive equipment and several hours for analysis even after the solvent extraction of target compounds from PM (~$100/sample). Hence, new technologies are being developed for PM analysis to complement standard analysis techniques.[53–57] Among these, spectroscopic and microscopic methods have a number of advantages such as measurement of chemical species of individual particles, ambient pressure operation, nondestructive analysis, and real-time analysis. Spectroscopic analysis such as UV laser-induced fluorescence (UV-LIF) and Raman spectroscopy have been used for classification of the aerosol particles.[58–61] These spectroscopic techniques are particularly useful for the analysis of bioaerosols and PM containing species with high fluorescent yield.

PAHs are good candidates for detection by spectroscopic techniques as they have high absorption coefficients and quantum yields.[62] The UV–visible electronic transitions in sp^2 carbon systems such as PAHs rely on $\pi*$–π transitions.[63] Several studies show that there is an inverse power law relationship between the optical band gap (OBG) and the number of benzene rings in PAHs[64,65]. An increase in MW of PAHs broadens their absorption bands towards longer wavelengths and red-shifts their emission bands due to the decreasing OBG.[66,67] This variation in spectral properties as a function of MW can be used to distinguish PAH content in combustion aerosols.

In general, fluorescence spectroscopy is highly sensitive (~1ng/mL) for PAH detection in ultrafine PM[68–70]; however, it is not specific. For analysis of multicomponent mixtures, scanning single-wavelength-excitations and stacking fluorescence emissions at each excitation wavelength provides a three-dimensional spectral fingerprint of the sample known as Excitation Emission Matrix (EEM). Among other applications, EEM spectroscopy for PM PAHs can be used to investigate soot formation routes in various combustion scenarios, for source apportionment studies[71], and also provide an estimate of PM toxic potential.[72]

In general, fluorescence spectroscopy is highly sensitive (~1ng/mL) for PAH detection in ultrafine PM[68–70]; however, it is not specific. PAH content determination in liquid samples, and qualitative analysis of diesel spray.[72–76] However, a convolution of individual compound signatures in a multi-component mixture, such as environmental samples, generates complex EEM spectra that are difficult to interpret. Principal Component Analysis (PCA) is commonly used for multivariate statistical analysis of analytical chemistry data, including EEM.[77,78] Several studies have successfully used the principal component regression (PCR) method, a combination of PCA and linear regression, to characterize and quantify PAHs from atmospheric samples and suspended sediments.[79,80] Therefore, advanced multivariate statistical techniques, such as PCR can be used to decompose the EEM complex fluorescence signal into the underlying component (or class of components) fluorescence spectra. We employ the PCR-EEM method to determine concentrations of PAH fractions in combustion generated PM.


Fluorescence Spectrum

Stacking Fluorescence Spectra at Different Excitations

Making a Contour Plot

Excitation Emission Matrix (EEM)

Figure 1.4: EEM is emission spectra of single excitation wavelengths stacked together.

1.4 Electrostatic Collection for *In-situ* Spectroscopic Analysis of Aerosols

Current methods for PM exposure monitoring, such as using gravimetric filters, tapered element oscillating microbalances (TEOM), beta attenuation monitors (BAM) and aerodynamic particle sizers (APS) are accurate and precise. These are, however, expensive and cumbersome for personal monitoring purposes. For example, filter collection and analysis is limited by high elution volumes, the form factor of the sampling setup, and high-power requirements. Hence, applications such as personal monitoring need the development of inexpensive and compact instruments that do not sacrifice sizing accuracy and reliability. Due to developments in miniaturization of sampling pumps and electronics, new exposure methods are becoming more popular in epidemiological studies and air quality monitoring applications. Recent examples of miniature PM samplers are presented in[81–85], including biological aerosol collectors[86–89], real-time time PM monitoring devices[90–93], and low-cost distributed sensor networks[94–96] for pollution exposure estimates.

Of particular interest is the collection and analysis of PM2.5. Their smaller particle size presents a challenge for real-time optical scattering instruments due to their low optical cross-section and their low Stokes number, which presents difficulties for inertial focusing[97–99] and inertial collection[100–103]. Filter collection of PM2.5 is currently a standard method used in air quality assessment. Electrostatic precipitation is another advantageous method for PM2.5 collection because the electrostatic force acting on a charged particle in the electric field is significantly higher than gravitational, inertial, and thermal forces, allowing for the manipulation of particles. Electrostatic precipitator (ESP) devices have been reported in the mid-20th century[104,105]. Although electrostatic precipitators are widely used in the industry, miniaturized electrostatic precipitator devices have not been widely used for exposure monitoring. ESPs remove suspended PM from gas by charging the PM in a corona discharge and separating it from the gas using an electric field. In a single-stage ESP, the electric field generates a corona discharge to charge the PM and to attract them to the grounded wall, where they are collected. In a two-stage ESP, charging and removal of PM occurs in two separate electric fields.[106] Collectors using electrostatic precipitation having novel designs and high PM2.5 collection efficiencies are being developed[86,87,107]. More recent work includes the design and evaluation of collectors integrated with analysis capabilities such as microscale optical spectroscopies[108] and fluidic assays[109–112]

In field settings, spectroscopic and microscopic methods have a number of advantages such as measurement of chemical species of individual particles, ambient pressure operation and nondestructive analysis. Other field-capable instruments that utilize spectroscopic approaches include photoacoustic spectroscopy (PAS), cavity ringdown spectroscopy (CRDS), incoherent broadband cavity-enhanced absorption spectroscopy (IBBCEAS), single-particle soot photometer (SP2) and wideband integrated bioaerosol sensor (WIBS).[113] All these techniques incorporate single and broadband wavelengths covering the UV, visible, and infrared, and give information about key climate relevant parameters like absorption and scattering, and about chemical composition from emission signals such as fluorescence, incandescence, etc. We present the design and evaluation of a novel two-stage ESP-based capillary sensor which can be integrated for subsequent *in-situ* spectroscopic analysis of PM2.5.

1.5 Solid-Phase Extraction and Fluorescence Analysis of Combustion Generated Aerosols

Analysis of the PM chemical composition is often a significant barrier in large-scale studies (especially in developing countries) due to cost, limited facilities, and personnel qualification required to perform pre-processing steps and analysis. EEM analysis requires sample mass measurements, extraction of organics into non-polar solvents, and dilution to appropriate concentrations levels as pre-processing steps.[71,114] To simplify sample prep and miniaturize the detection platform, solid-phase solvents can be used. Among solid-phase solvents, polydimethylsiloxane (PDMS) provides an attractive option as it is inert, non-toxic, non-flammable, and optically transparent. PDMS has been previously used for extraction and sensing of PAHs.[115,116] When PM is deposited onto PDMS, the organic molecules partition into solvent. Given that PDMS has a relatively low fluorescent background, it is well suited for EEM analysis. We report the development of a novel SP-EEM technique for chemical analysis of combustion generated PM.

Chapter 2 Generation and Chemical Composition of Combustion Aerosols

2.1 Summary

In this chapter, we examine the presence and concentrations of PAHs in combustion generated PM. Combustion PM samples were generated in an inverted gravity flame reactor (IGFR) operated on either ethylene or ethane fuel. The fuel was diluted by Ar in over a range of 0% to 90% to control the flame temperature. PM samples were collected and analysed from additional sources, including diesel exhaust and woodsmoke from UW clean cookstove. Analysis was performed via conventional laboratory techniques such as TOC and GCMS. The colder flames result in lower PM yields; however, the PM PAH fraction increases significantly. Temperature thresholds for soot transition from low to high organic carbon content were characterized based on the maximum flame temperature ($T_{max,c}$ ~1791-1857K) and the highest soot luminosity region temperature (T^*_c~1600-1650K). These results are published.[117]

2.2 Experimental Methods

2.2.1 PM Collection in Inverted Gravity Flame Reactor

Soot formation has been studied in many combustion configurations, including opposing flow burners, co-flow, and flat flame burners, e.g., refs [118–122]. Among them, inverted co-flow diffusion flames have some advantages due to their stability, which allows one to achieve a broad range of temperatures and residence times.[123–125] An IGFR generates ultrafine PM over a range of combustion temperatures due to the stable flame and high residence time compared to an upright flame. The combustor allows for control of flame residence times (τ_{res}) by altering the balance of the convective and buoyant terms.[126] The increased residence time and flame stability in the IGFR have been utilized to synthesize materials such as carbon and TiO_2 aerosol gels[127,128] and for testing of several hypotheses related to PM formation and growth.[4] Here, the IGFR is used to produce combustion PM samples with controlled organic carbon fraction by varying the flame temperature with fuel dilution by Ar. The IGFR consists of a 150 mm long quartz tube with an inner diameter of 32 mm. A mixture of fuel and Ar enters the combustor in a "+g" (downward) direction. The air is introduced as a co-flow into the reactor via concentric inlets having diameters 9.5 mm and 32 mm. A flow straightener at the combustor inlet prevents the propagation of instabilities from the air manifold. The combustion conditions, such as temperature, dilution, fuel, composition, and flowrates determine the PM formation rates and routes, thus controlling PAH fractions in the sample. The dilution ratio is defined as the volumetric ratio of Ar versus the sum of fuel plus Ar in the fuel stream. The dilution ratios for each fuel were chosen to obtain flame structure stability at each dilution ratio as well as flame structure variability among different dilution ratios for the same fuel. Stable flame structures produce PM with repeatable chemical composition. High variation in flame structures generates PM with variable chemical composition.

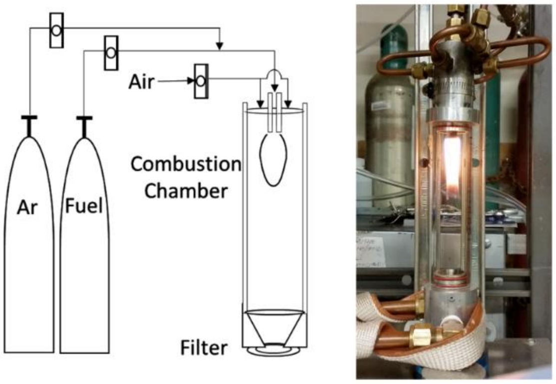
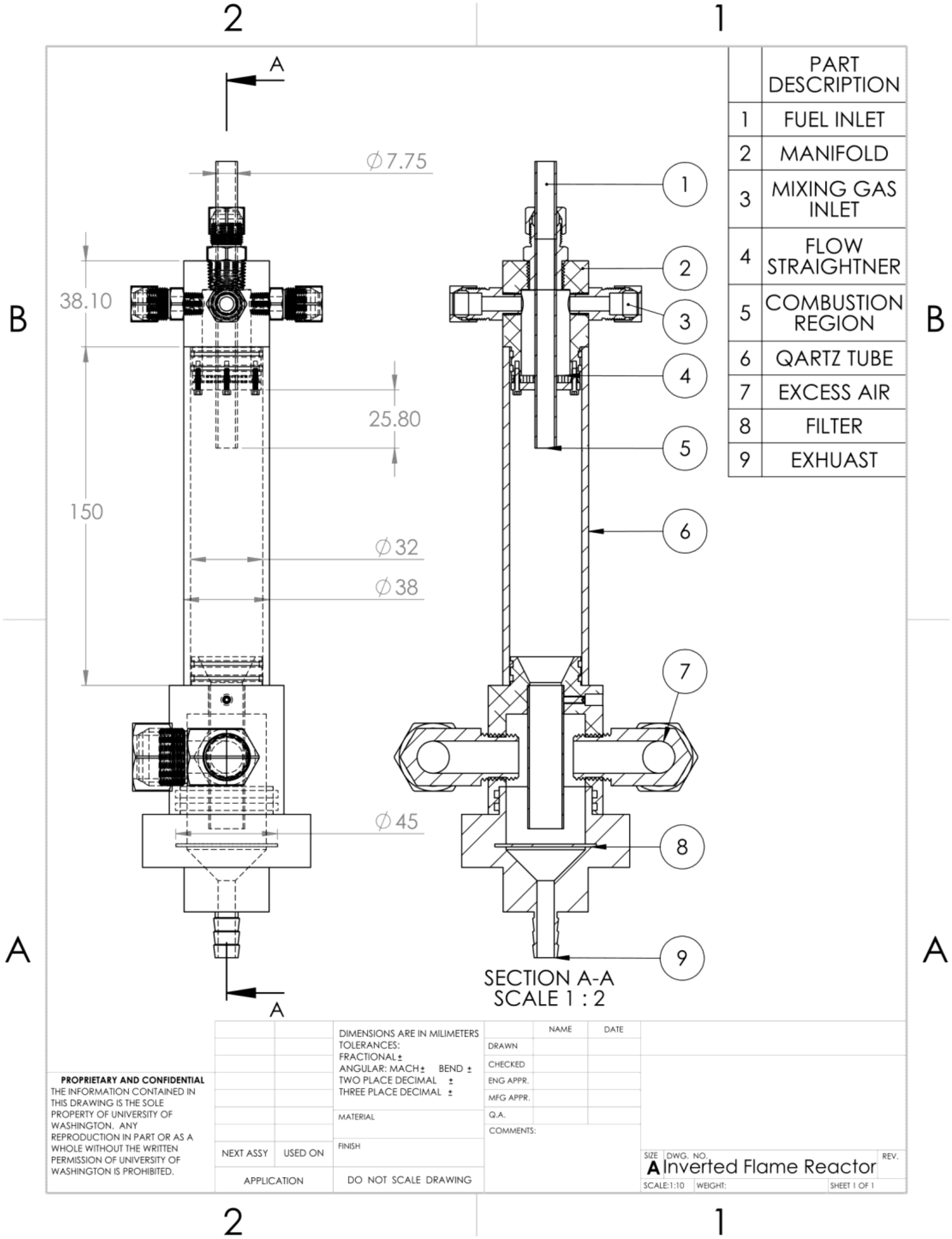


Figure 2.1: IGFR operating principle (left) and IGFR image (right). Fuel is diluted with argon with different dilution ratios. Soot produced in the flame is collected at the bottom of the combustion chamber.



	PART DESCRIPTION
1	FUEL INLET
2	MANIFOLD
3	MIXING GAS INLET
4	FLOW STRAIGHTNER
5	COMBUSTION REGION
6	QARTZ TUBE
7	EXCESS AIR
8	FILTER
9	EXHUAST

Figure 2.2: Design of the IGFR combustion chamber.

Introduction of Ar reduces the maximum flame temperature (T_{max}). For each fuel, the ratio of fuel flow rate to air flow rate defined as the global equivalence ratio (φ) was constant. Figure 2.1 shows the IGFR schematic, and Figure 2.2 shows engineering drawings of the reactor. Table 2.1 shows the experimental conditions matrix used in this study. The soot emission factor, defined as mg of PM per gram of fuel, was calculated by gravimetric analysis. A weighed filter was placed in an open-face filter holder downstream of the flame. The sampling flow rate was set at 4 slpm, which was greater than the inlet flow, to ensure that all PM from the flame was collected. Make-up air compensated for the difference in inlet and sampling flow rates. The sample collection time was varied according to PM yields to gain an adequate mass for TOC, GCMS, and EEM analyses. Table 2.1 shows the number of replicates for each condition.

Table 2.1: Experimental conditions for different samples and # samples used for different analysis techniques

Fuel	Dilution Ratio	Duration of Sampling (min)	Equivalence Ratio (φ)	# samples for OC analysis	# samples for GCMS analysis	# samples for EEM analysis
Ethylene	0.9	120	0.65	1	1	1
Ethylene	0.85	15	0.65	1	1	1
Ethylene	0.8	15	0.65		3	3
Ethylene	0.67	2	0.65	1	2	2
Ethylene	0.5	0.5	0.65	1	1	1
Ethylene	0	0.33	0.65	1	3	3
Ethane	0.85	120	0.8	1	1	1
Ethane	0.8	30	0.8	1	1	1
Ethane	0.67	5	0.8	1	1	1
Ethane	0.6	3	0.8		1	1
Ethane	0.5	2	0.8	1	1	1
Ethane	0.25	1	0.8		2	2
Ethane	0	0.5	0.8	1	2	2

2.2.2 Maximum Temperature Measurements (T_{max})

The experimental measurements and calculation shown in this section are performed in collaboration Mr. Eric Molnar and Dr. Justin Davis. Some parts of the IGFR measurement and the temperature correction have been published in ref[2,3,5,114], these measurements and methodology are presented here for completeness. Flame temperature affects the condensation, coalescence, transport, carbonization, and aggregation rates during soot formation. Flame temperatures were measured using an uncoated R-type thermocouple (wire diameter 0.075 mm, bead head diameter (d) 0.19 mm via optical microscopy). Due to competing radiative and conductive heat transfer, the thermocouple reached an equilibrium temperature lower than the flame temperature. Therefore, the thermocouple measurements (T_{TC}) were corrected following previously used radiation balance methods.[129,130] Heat loss from conduction down the thermocouple wire was neglected. The Nusselt number (Nu) was approximated as that of a sphere. For diluted flames, the relationship between thermal conductivity (k) and flame temperature (T) is linear as given by (2.1).

$$k = \frac{55.4T_{cor} + 1228.9}{10^6} \tag{2.1}$$

Furthermore, the emissivity (ε) of the thermocouple is dependent on its temperature and material. Since the R-type thermocouple is a platinum (Pt) wire and a 13% rhodium (Rh)/ Pt wire, the ε values for an S-type (Pt -10% Rh/Pt) thermocouple are assumed to be representative of the R-type thermocouple and are given by (2.2) and (2.3). The arithmetic average ε value of the two wires was used to obtain ε of the thermocouple (ε_{TC}) (2.4).[131] (2.5) was used to calculate corrected flame temperature (T_{cor}). An iterative function solved for k alternatively based on T_{TC} measurements.

$$\varepsilon_{Pt} = 0.136\ln(T_{TC}) - 0.8 \tag{2.2}$$

$$\varepsilon_{Pt,10\%\,Rh} = 0.135\ln(T_{TC}) - 0.78 \tag{2.3}$$

$$\varepsilon_{TC} = \frac{\varepsilon_{Pt,10\%\,Rh} + \varepsilon_{Pt}}{2} \tag{2.4}$$

$$T_{cor} = T_{TC} + \frac{\sigma\varepsilon_{TC}(T_{Tc}^4 - T_{amb}^4)d}{(k)(Nu)} \tag{2.5}$$

As the thermocouple is placed inside a sooting flame, thermophoresis drives soot particles to deposit on the thermocouple, increasing ε and d. However, it was assumed that soot deposition is negligible because the environment was oxidative.[132] Additionally, a rapid thermocouple insertion technique was used to reduce temperature correction error.[131]

For measuring flame temperature, thermocouples have been an effective and low-cost tool, but probing the flame with a thermocouple introduces errors, e.g., flame temperature and flow field perturbation.[124,127] Semi and non-intrusive techniques such as pyrometry, laser light incandescence, spectral soot emission, and laser extinction have been proposed [133–136]; however, optical distortions in the IGFR quartz tube makes their implementation challenging. In the current work, the temperature was measured using an uncoated R-type thermocouple. The temperature measurement method is based on two assumptions of flame structure: (i) the flame is axisymmetric and (ii) the maximum temperature on the wing is located at the flame front.[137] For each flame condition, the flame front temperatures are recorded with a 5 mm resolution in the axial direction. The maximum value among these measurements determines T_{max}.

The flame temperature in the IGFR is measured with a bare wire thermocouple attached to a stepper motor (0.9-degree rotation per step) on a vertical stage. A microprocessor is programmed to periodically move the exposed junction of the thermocouple from the sheath air region into the flame. Thermocouple junction exposure to the flame is 2.5 seconds. The thermal response time-step, τ_{TC}, of the P13R-003 thermocouple is less than 200 ms. The duration of measurement (2.5 s) allows for thermal equilibrium while mitigating possible soot deposition. Soot deposition is minimal in the high-temperature oxidative region and at axial positions near reactor inlet. For low dilution flames and downstream positions in the flame, soot escapes the flame front boundary and soot deposition is more likely.

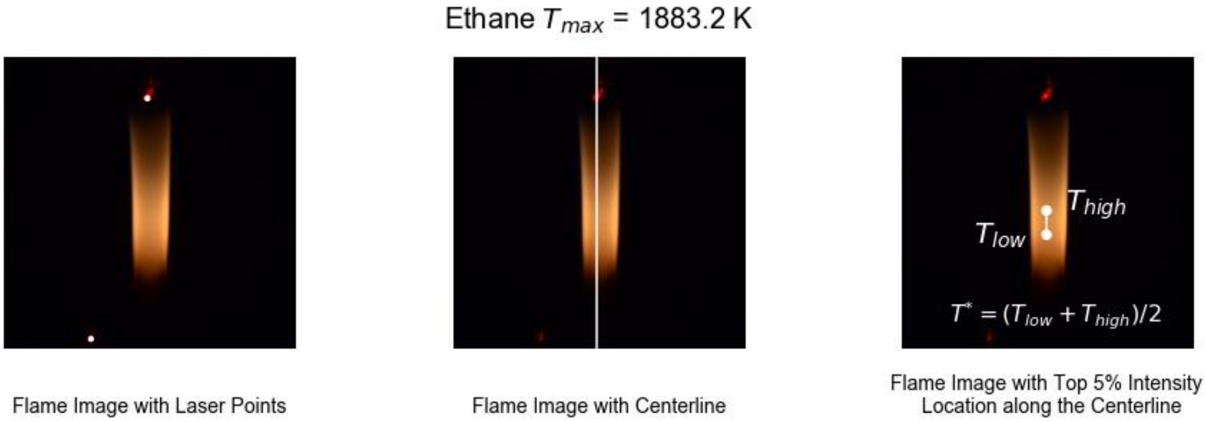
The frequency of measurement is limited by the acquisition module at 200 ms (*OMEGA* TC-08, Omega, Norwalk, CT, USA). The recorded value for each axial measurement along the flame front is determined from the maximum of ~10 measurements. The axial position is varied by adjusting the linear stage at 5 mm increments from the end of the fuel inlet tube to the end of the visible flame. Thermocouple measurements are recorded by acquisition software (*PicoLog Beta 6,* Pico Tech, Cambridgeshire, UK) in Celsius. The R-type thermocouple has a tolerance of 0.25% of measured value (e.g. 3.75 C for 1500 C). The acquisition module has an additional tolerance of 0.5 C plus 0.2% of the measured value. Measurement values are converted to Kelvin before radiation correction.

2.2.3 Temperature at Maximum Soot Radiation Location (T^*)

Note that T_{max} generally does not correspond to the region of the maximum soot production.[5] The soot formation region temperature (T^*) can be a more relevant measure for particle exposure to temperature.[3,5] The estimation of T^* is based on the flame luminosity in the red spectrum and the flame front temperature profiles.[5] The locations with 95% percentile intensity in the red spectrum were determined as the soot formation region, with the measured temperature assumed to be soot formation temperature (T*).

To determine temperature at the maximum soot radiation location (T^*), the temperature measurements followed a similar procedure as described previously. Flame temperatures were recorded using an uncoated R-type thermocouple (wire diameter 0.075 mm, bead head diameter (d) 0.17 mm via optical microscopy). However, instead of measuring temperature around one axial position, the temperature was recorded for varying axial positions along the flame axis with a 5mm resolution. The thermocouple thermal response time was around 200 ms, as was the data collection frequency of the acquisition module and software. For each axial location, 10 measurements were taken within 2 minutes and the maximum of the 10 measurements was considered as the recorded data point. Temperature measurements along the flame axis were used to estimate the axial temperature profile for each flame condition using linear interpolation. This temperature profile was used to estimate the temperature at the maximum soot radiation location (T^*) as described in the next section.

The axial flame temperature profiles were used to identify the gas temperature of the maximum soot radiation region, that corresponds to the highest soot temperature. Temperature in the vicinity of the maximum soot radiation was determined based on flame images for each flame condition. Digital images of flames were captured with a Nikon DSLR 5100, a widely available digital camera using a standard CMOS sensor operated on manual mode. Aperture, focal length, and sensitivity were kept constant and shutter speed was adjusted to reduce overexposure for flame conditions with lower Ar dilution. Two laser points were used as reference points to denote the inlet nozzle location and a location 8 mm apart from the inlet nozzle along the flame axis. A python script was used to extract the red layer of the RGB image and evaluate the location of pixels with the top 5% intensity along the centreline of the flame to approximate the region of high luminosity. Temperatures at the endpoints of the top 5% intensity line were estimated based on the temperature profile obtained from the temperature measurements taken along the flame axis. T^* is reported as the average of the temperatures estimated at the two endpoints as shown in Figure 2.3.


Ethane T_{max} = 1883.2 K

Flame Image with Laser Points

Flame Image with Centerline

Flame Image with Top 5% Intensity Location along the Centerline

Figure 2.3: The python script used to estimate T^* detects the two reference laser points in the flame image (distance between the laser points is 8 mm) (left), evaluates the location of pixels with top 5% intensity in the red layer of the image along the centerline (shown in the middle), interpolates the thermocouple temperature measurements to estimate T_{low} and T_{high} based on the location of the endpoints of the top 5% intensity line (right), and reports T^* as the average of T_{low} and T_{high}.

Figure 2.4 (left) shows variation of T^* with dilution for both ethylene and ethane. T^* decreases with an increase in Ar dilution except for an Ar dilution increase from 0.67 to 0.8 for ethylene where T^* increases. This indicates that the temperature experienced by the soot in the flame decreases with an increase in Ar dilution. Figure 2.4 (right) shows variation of T^* with respect to T_{max} for both ethylene and ethane. The variation can be approximated as a linear relationship. Although T^* can be used for characterization of each flame, there is a lot of uncertainty in terms of estimating the range of highest temperatures experienced by the soot in the flame. The radial temperature variation for each axial location is rapid and can cause error in maximum temperature measurement for each axial location. Based on the previously reported results for laminar diffusion flames, the radial location of the maximum soot concentration is shifted 1-2 mm towards the axis of symmetry from the stoichiometric flame temperature and temperature correction associated with this radial shift can be in the range of 10 K - 650 K.[130,138–141] Hence, temperatures estimated in the region of high luminosity are only an approximation of the highest temperatures experienced by the soot. Here we provide an estimated range of temperatures between T_{low} and T_{high} that the soot experiences as highest temperatures in the flame and T^* as the average of that range.

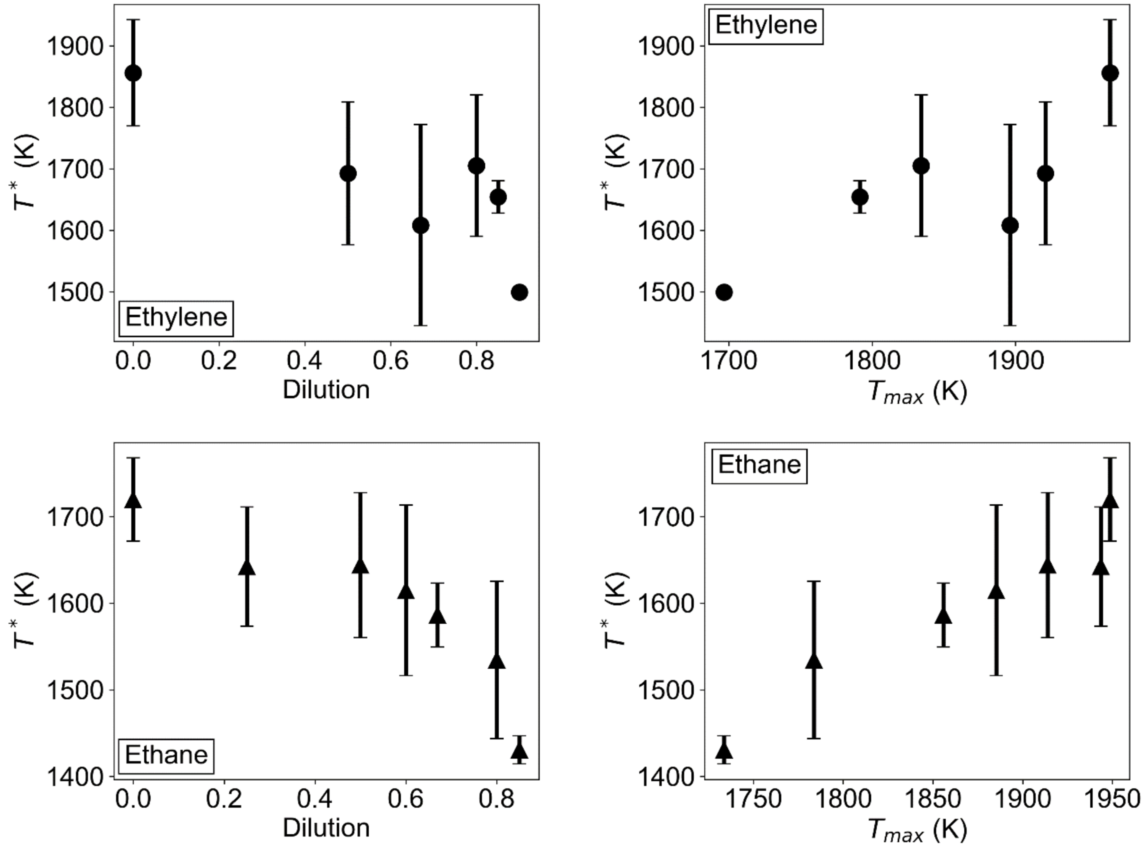


Figure 2.4: The temperature at the maximum soot radiation location (T^*) is denoted by the midpoint of the temperature range bar. The endpoints of the temperature range bar are T_{low} and T_{high}. T^* decreases with increasing Ar dilution, in general, for both ethylene and ethane except for an increase in T^* when the Ar dilution increases from 0.67 to 0.8 for ethylene (left). T^* increases linearly with T_{max} for both fuels (right).

Due to the high uncertainties in estimating the range and location of T^*, we use T_{max} for flame characterization, T^* is reported for reference; the typical values are lower than T_{max} by 150-300 K.[3,138,142,143]

2.2.4 Woodsmoke and Diesel PM Collection

Woodsmoke samples were collected by burning 1½ by ¾ inch Douglas fir sticks cut from dimensional lumber in a prototype side-feed, natural-draft cookstove.[144] The sampling filters were placed in a sealed chamber connected to the exhaust duct at the sampling point. Diesel PM samples were collected in the UW diesel exposure facility; the samples were generated from a 435 cc direct-injection single-cylinder diesel engine (Yanmar LW Series, Yanmar, Adairsville GA, USA) operated on ultra-low-sulfur diesel.[145] PM was collected from the exhaust duct connected to the engine using 2.0 µm pore PTFE membrane filters (Pall Zefluor®, Pall Cat. # P5PJ037, Pall Corporation, Port Washington, New York, USA) housed in the Harvard impactor (Cat. # HP2518, BGI, Butler, NJ, USA) sampling cassettes placed in a sealed chamber. Sampling flowrate was set to 1.8 lpm using either portable or stationary vacuum pumps (AirChek XR5000 pump, SKC Inc., Eighty Four, PA, USA or VP0625-V1014-D2-0511, Medo USA, Roselle, IL, USA) with a custom manifold of nine VFB-65-BV roto-meters (Dwyer Instruments, Michigan City, IN, USA). Flowrates were verified using a flow calibrator (Gilian Gilibrator PN# 800268, Sensidyne, St. Petersburg, FL, USA).

2.2.5 Gravimetric Analysis

PM samples weights collected from IGFR, cookstove and diesel burning are typically of the order of micrograms. The U.S. EPA requirements for the filter conditioning and weighing environment are a mean temperature of 20 - 23 ± 2 °C and a mean RH of 30 – 40 ± 5 % over 24 hrs.[146] Filter contamination, balance vibration, electrostatic charges, temperature and relative humidity (RH) variations in the filter-conditioning environment all contribute to uncertainties in filter handling and weighing. These variations must be controlled to achieve reliable weight measurements. Hence we use an inexpensive weighing chamber designed for use in PM exposure assessment studies at the University of Washington in Seattle for gravimetric analysis.[147] Zeflour PTFE filters (Pall Zefluor®, Pall Cat. # P5PJ037, Pall Corp., Port Washington, NY, USA) were conditioned inside the chamber for atleast 24 hours before weighing. Electrostatic charges are minimized before weighing by passing the filters between two Po/sup 210/ sources (500 microcuries each), and during weighing by two Po/sup 210/ strips placed on top of the filter-weighing pan. The Po/sup 210/ strips emit alpha particles neutralizing static charges in filters. The filters were weighed before and after PM collection using a microbalance with 0.5 µg resolution (Mettler-Toledo UMT-2, Greifensee, Switzerland).[147]

2.2.6 Total Organic Carbon (TOC) Analysis

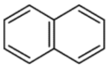
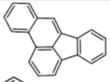
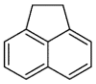
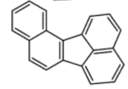
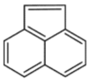
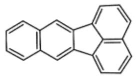
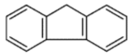
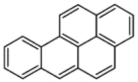
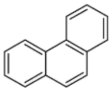
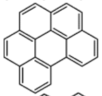
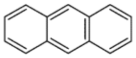
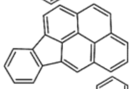
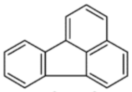
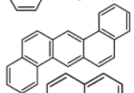
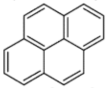
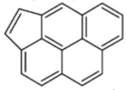
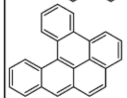
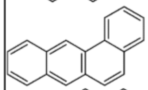
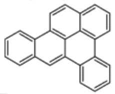
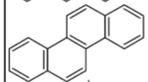
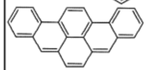
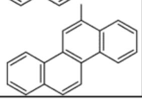
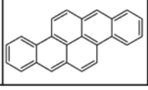
TOC analysis was performed by Sunset Laboratories Inc (Tigard, OR-97223, USA) using a thermal/optical method based on National Institute for Occupational Safety and Health (NIOSH) Method 5040 in the NIOSH Manual of Analytical Methods (NMAM).[148] The thermal optical analyzer[136,149] is equipped with a pulsed diode laser and photodetector that permit continuous monitoring of the sample filter transmittance. PM samples were collected on fiberglass filters (934-AH 47mm, SterliTech, Kent, WA, USA) instead of PTFE filters. In the thermal-optical analysis, a portion of a sample is placed in the sample oven, which is then tightly sealed. The analysis proceeds in inert and oxidizing atmospheres. First, OC is removed in helium as the temperature is increased to a preset maximum (about 870°C in NMAM 5040). After removal of OC, an oxygen-helium mix is introduced to effect combustion of the remaining material. An increase in filter transmittance is seen as light-absorbing carbon is removed. The split between the organic and elemental carbon (EC) is designated at the point where the filter transmittance reaches its initial baseline value. The carbon removed prior to the OC-EC split is considered organic, and that removed after the split is considered elemental. The total carbon (TC) in the sample is the sum of OC and EC.

2.2.7 Extraction of PM organic fraction

After performing the gravimetric analysis, the samples were extracted in cyclohexane due to its non-polar characteristics and common use for PM PAH extraction.[150] (Cat. #1.02822.2500, Uvasol® Cyclohexane for Spectroscopy, MilliporeSigma, Burlington, Massachusetts, USA). The filters were cut into four equal parts and placed in 20 mL glass vials (Cat # 89096-774, VWR, Edison, NJ, USA). The vials were filled with ~3mL of cyclohexane and sonicated for 30 min in a water bath sonicator (42 kHz, 2510R-MT Branson, Ultrasonic Corp., Danbury, CT, USA). After sonication, the samples were soaked for 24 hours and filtered with 0.2 µm PTFE syringe filters (Cat. #28145-491, VWR, Edison, NJ, USA) into 4 mL vials (Cat # 66009-876, VWR, Edison, NJ, USA). Blank filters, used for reference, were extracted in the same manner. Aliquots (~200 µL) of filtered extracts were transferred into 2 mL GCMS vials having 250 µL inserts and PTFE screw tops (Cat# 5182-0715, 5181-8872 and 5182-0717, Agilent, Santa Clara, CA, USA) for GCMS analysis. The remaining extracts were diluted further into 20 mL glass vials. The IGFR, diesel exhaust, and woodsmoke samples were diluted to a final concentration of 40, 20 and ~3.5-4 µg(soot)/mL(cyclohexane) respectively for EEM analysis.

2.2.8 Gas Chromatography Mass Spectroscopy (GCMS) Analysis

The GCMS analysis was performed in collaboration with Dr. Jay Rutherford and is published in a joint publication.[114] Here the methodology and the results are presented for completeness. PM extracts were analyzed for the PAHs listed in Figure 2.5, these include the 16 PAHs recommended as primary pollutants by the EPA and eight additional species. The extracts were analyzed using an Agilent 7000 GC/MS Triple Quad Mass Spectrometer using two 15 m columns (Part #: HP-5MS UI, Agilent, Santa Clara, CA, USA) equipped with a backflush. 24 PAHs (EPA 16 PAHs[39] and eight additional compounds with MW up to 302 g/mol) were included in the analysis. Calibration curves for PAH species with seven concentration levels in the range of 1-1000 ng/mL were obtained. The calibration standards for the 24 PAHs were a mixture of 23 compounds (Wellington Laboratories Cat. # PAH-STK-A, Guelph, ON, Canada) and one additional compound, coronene, a standard PAH used in several studies on mechanisms of soot formation[28] (AccuStandard Cat. # H-116, New Haven, CT, USA). Each calibrant included 16 deuterium-labeled PAH internal standards (Wellington Laboratories Cat. # PAH-LCS-A, Guelph, ON, Canada). PAH internal standards were at a concentration of 100 ng/mL in the calibrants, and an equivalent concentration of the same internal standard was spiked into each sample for use in quantification. The instrument was operated in pseudo multiple reaction monitoring (PMRM) mode[151].

PAH		Structure	High/Low MW	PAH		Structure	High/Low MW
Naphthalene (NA)	$C_{10}H_8$ 128		Low	**Benzo[b] fluoranthene** (BbF)	$C_{20}H_{12}$ 252		High[#]
Acenaphthene (AC)	$C_{12}H_{10}$ 155		Low *	Benzo[j] fluoranthene (BbF)	$C_{20}H_{12}$ 252		High[#]
Acenapthylene (ACN)	$C_{12}H_8$ 152		Low	**Benzo[k] fluoranthene** (BkF)	$C_{20}H_{12}$ 252		High[#]
Fluorene (FL)	$C_{13}H_{10}$ 166		Low	**Benzo[a]pyrene (BaP)**	$C_{20}H_{12}$ 252		High
Phenanthrene (PHE)	$C_{14}H_{10}$ 178		Low	**Benzo[ghi]- perylene (BgP)**	$C_{22}H_{12}$ 276		High
Anthracene (AN)	$C_{14}H_{10}$ 178		Low	**Indeno[1,2,3-cd]- pyrene** (IP)	$C_{22}H_{12}$ 276		High
Fluoranthene (FA)	$C_{16}H_{10}$ 202		Low*	**Dibenz[a,h]- anthracene** (DaA)	$C_{22}H_{14}$ 278		High*
Pyrene (PY)	$C_{16}H_{10}$ 202		Low	Coronene (COR)	$C_{24}H_{12}$ 300		High
Cyclopenta(CD) Pyrene (CpPY)	$C_{18}H_{10}$ 226		High	Dibenzo[a,l] pyrene (DBP-al)	$C_{24}H_{14}$ 302		High[†]
Benz[a]anthracene (BaA)	$C_{18}H_{12}$ 228		High	Dibenzo[a,e] pyrene (DBP-ae)	$C_{24}H_{14}$ 302		High[†]
Chrysene (CHR)	$C_{18}H_{12}$ 228		High	**Dibenzo[a,i] pyrene (DBP-ai)**	$C_{24}H_{14}$ 302		High[†]
5-Methyl Chrysene (5M-CHR)	$C_{19}H_{14}$ 242		High*	Dibenzo[a,l] pyrene (DBP-al)	$C_{24}H_{14}$ 302		High[†]

*Excluded from PCA-MLR analysis due to infrequent detection in GCMS

[#] Isomers co-eluted so quantified as a group in GCMS

[†] Isomers lumped together for analysis, quantified by GCMS separately

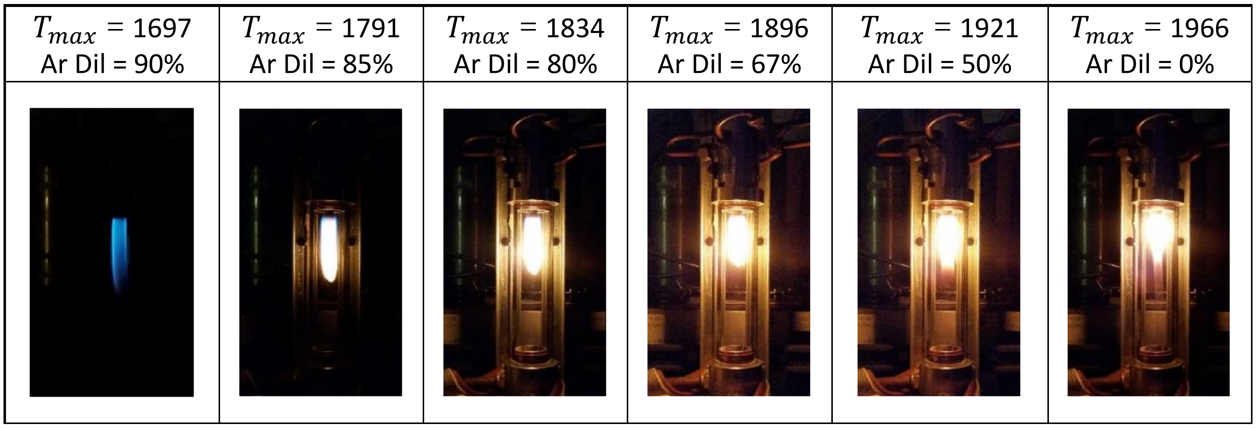
Figure 2.5: PAHs used for GCMS analysis. PAHs having infrequent detection were excluded and a few PAHs were lumped together for PCA-MLR analysis in which 16 PAHs are considered. Deuterium labeled analogs of the compounds shown in bold were used as internal standards.

Calibration standards were prepared for the 24 compounds listed in Figure 2.5. Twenty-three of the native compounds were provided as a mixture (Wellington Laboratories Cat. # PAH-STK-A, Guelph, ON, Canada) and Coronene was added as an additional analyte (AccuStandard Cat. # H-116, New Haven, CT, USA). Sixteen deuterium-labeled internal standard compounds (shown in bold in Figure 2.5) were provided in a mixture (Wellington Laboratories Cat. # PAH-LCS-A, Guelph, ON, Canada). These internal standard compounds were included in each calibrant and spiked into each sample at a constant concentration of 100 ng/mL. For compounds without a deuterium labelled analogue, we used the deuterium labelled compound with the closest molecular weight as an internal standard. Our calibration curve consisted of the following concentration levels: 1, 4, 10, 40, 100, 400, and 1000 ng/mL.

2.3 Results and Discussion

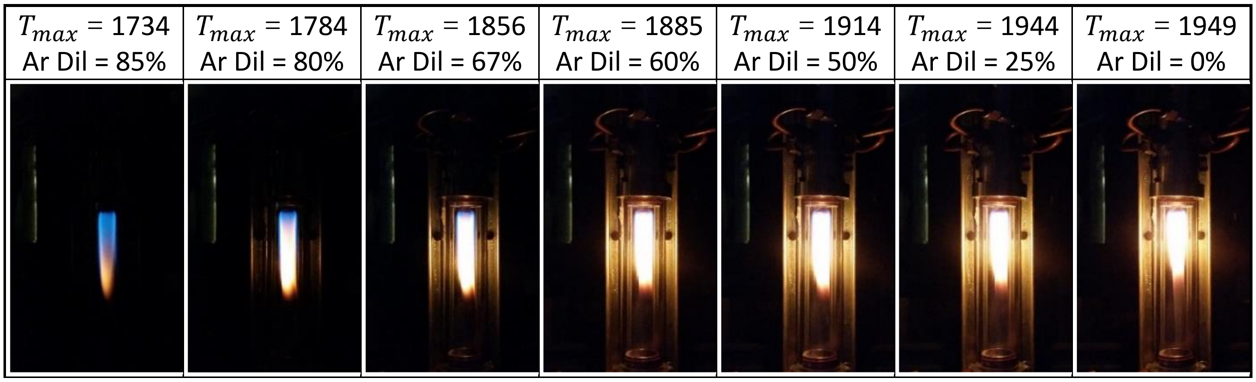
2.3.1 Effect of Argon Dilution

Effects of fuel dilution by Ar on the maximum flame temperatures, soot yields, and TOC are presented. Figure 2.8 A shows that T_{max} decreases with fuel dilution due to heat transfer from the combustion reaction to Ar. As the dilution ratio changes, the difference between the adiabatic temperature and the measured flame temperature remains similar ~250-300K.[5] There are multiple processes responsible in the changes of the flame temperature, e.g., (i) reduction of soot radiation in the diluted flames, (ii) downstream shift of the T_{max} location (longer residence time) in the diluted flames based on flame speed (iii) changes in the flame shape from open-tip to close-tip at higher dilutions.[5] The flame brightness and color are shown in Figure 2.6 and Figure 2.7. Figure 2.8 B shows that the emission factor increases with T_{max}. The amount of soot collected for the flame condition with the highest T_{max} is two orders of magnitude greater than that of the lowest temperature. Although the emission factor is measured *ex-situ* gravimetrically, visual observation confirms that flames with greater T_{max} have significantly higher soot radiation. At the higher Ar dilutions, the flame is mostly blue with a visible orange region at the tip of the flame, which agrees with the previous reports of suppression of soot formation due to inert additives.[133] The contributing factors include (i) dilution (lower concentrations) of species participating in the reaction and (ii) thermal effects – lower collision energies. Dilution by inert additives has been shown to have a greater impact on the sooting tendency than thermal effects alone.[133,152,153]



T_{max} = 1697 Ar Dil = 90%
T_{max} = 1791 Ar Dil = 85%
T_{max} = 1834 Ar Dil = 80%
T_{max} = 1896 Ar Dil = 67%
T_{max} = 1921 Ar Dil = 50%
T_{max} = 1966 Ar Dil = 0%

Figure 2.6: Ethylene flame brightness variation with T_{max} and argon dilution (Ar Dil), image exposure is varied.



T_{max} = 1734 Ar Dil = 85%
T_{max} = 1784 Ar Dil = 80%
T_{max} = 1856 Ar Dil = 67%
T_{max} = 1885 Ar Dil = 60%
T_{max} = 1914 Ar Dil = 50%
T_{max} = 1944 Ar Dil = 25%
T_{max} = 1949 Ar Dil = 0%

Figure 2.7: Ethane flame brightness variation with T_{max} and argon dilution (Ar Dil), image exposure is varied.

Figure 2.8 C shows that TOC decreases with the increase in flame temperature. The higher TOC in the soot from lower temperature flames is possible due to a reduction in hydrogen abstraction, hindrance of carbonization of young soot particles and reduction in PAH oxidation rates.[28] Detailed analysis of the effects of flame conditions on soot formation mechanisms[154,155] is beyond the scope of this work. However, the PAH types and concentration measured in the study can be used to gain insight into PM maturation as a function of temperature or position in the flame.

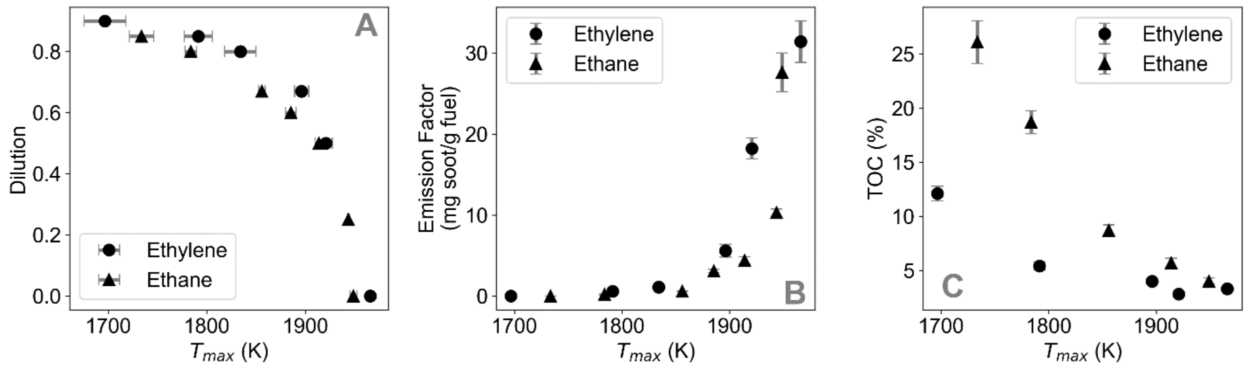


Figure 2.8: Maximum flame temperature (T_{max}) varies with argon dilution of fuel affecting the soot emission factor and TOC from ethylene and ethane flames. (A): The maximum flame temperature decreases with increasing Ar dilution for both fuels. (B): The emission factor (soot yields) increases with flame temperature. (C): TOC fraction decreases with increasing flame temperature.

2.3.2 Change in PAH Composition with Flame Temperature

Figure 2.9 shows PM-PAH content (by GCMS) from IGFR operated on ethylene and ethane as a function of flame temperatures and PAH content in PM from woodsmoke and diesel exhaust. A total of 24 PAHs were analyzed, but only 16 groups are shown, as the PAHs detected at low concentrations were excluded, and structural isomers were grouped together. For both ethylene and ethane, HMW PAHs are observed at lower T_{max}. The LMW PAHs are mostly found at intermediate and high T_{max}. Curiously, in ethylene's case, lowest T_{max} conditions yield both LMW and some HMW PAHs associated with PM. We do not have an explanation for this observation though this data point was repeated multiple times. Figure 2.10 and Figure 2.11 show the complete GCMS dataset. For woodsmoke, HMW PAHs are dominant while for diesel exhaust, LMW PAHs are dominant. Note that the concentration of PAHs in woodsmoke is significantly high compared to diesel exhaust if normalized by the sample mass.

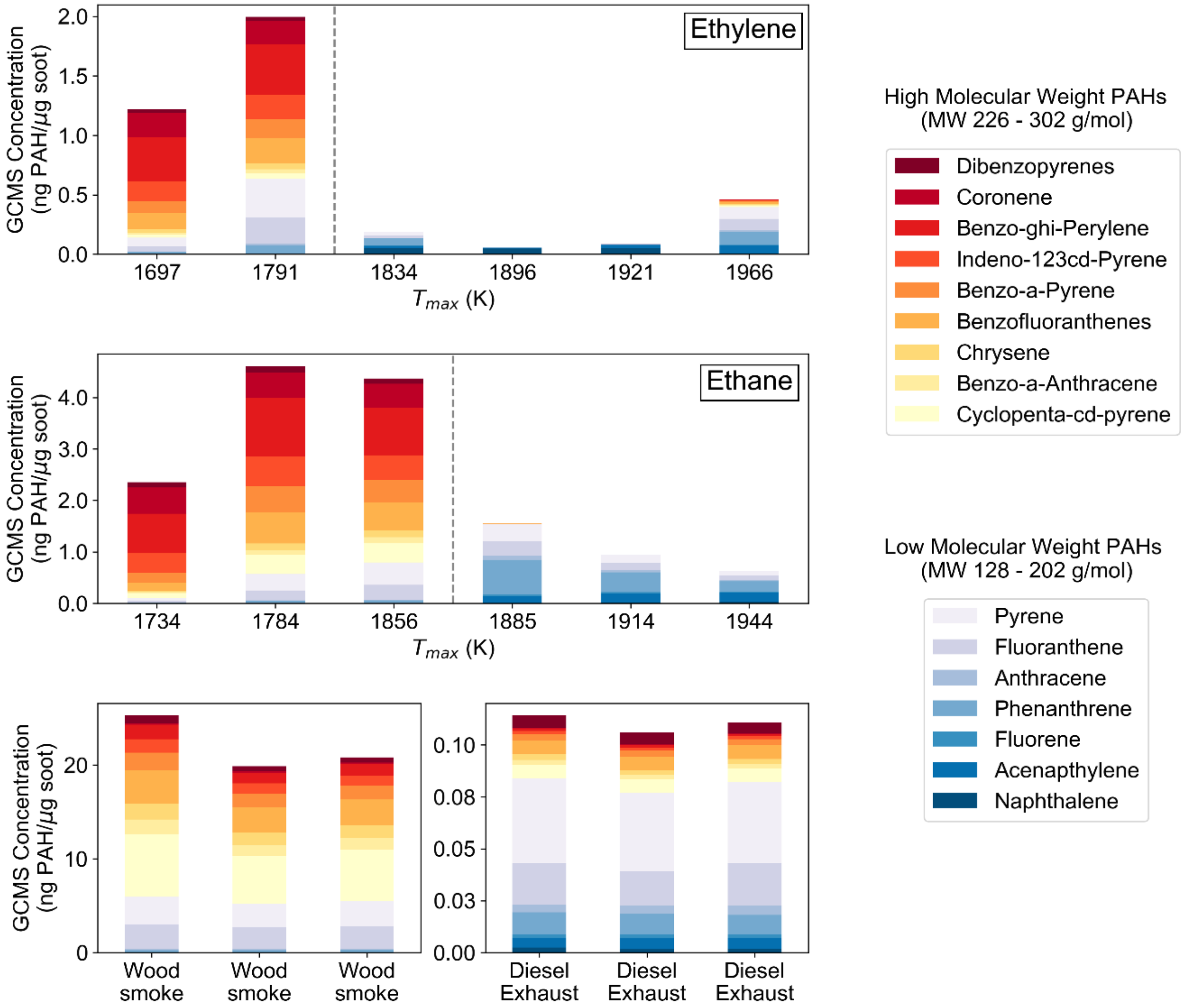


Figure 2.9: PM PAH fraction as a function of T_{max} for IGFR flames, woodsmoke, and diesel exhaust. GCMS concentration of 16 PAHs is divided into LMW and HMW PAH. In IGFR samples, for higher T_{max}, LMW PAHs are prevalent; for lower T_{max}, HMW PAHs are the dominant group. The vertical dashed gray line indicates the carbonization threshold in terms of maximum flame temperature. The woodsmoke contains mostly HMW PAHs, diesel exhaust - LMW PAHs. Note the very low concentration of the diesel sample.

For both ethylene and ethane, we observe two temperature ranges that correspond to the presence of HMW and LMW PAHs in the PM sample. The transition is associated with soot maturation, i.e., a transition of incipient soot with liquid-like properties, high organic fraction and unstructured morphology to particles with graphitic core-shell carbon structures.[3] This transition is typically defined by carbonization temperature (T_{carb})[156,157]. From a phenomenological perspective, the temperature in the highest soot concentration region is a more relevant metric to evaluate this transition than the maximum flame temperature but due to high uncertainty in determining T^* (as shown in Figure 2.4), T_{max} is used. For high-temperature combustion sources ($T^* > T_{carb}$), the HMW PAHs transition to graphitic structures forming BC[3] which is also observed in other co-flow diffusion flames[158]. BrC, characterized by high TOC fraction, is often associated with low-temperature combustion.[159] These conditions are typically found in biomass combustion due to the lower heating value of the fuel (high moisture content) and complex (endothermic) surface chemistry reaction. The samples in our lower temperature experiments also exhibit a brown color. It is consistent with other reports showing HMW PAHs found in biomass burning.[160] In terms of maximum flame temperature, the threshold for transition from high PAH PM to highly graphitic PM is in the range 1791K-1857K, denoted by $T_{max,c}$, the subscript 'c' stands for critical. The transition is shown as a dotted grey line in Figure 2.9. These flame temperatures correspond to the estimated local temperature condition for soot exposure of T^*_c ~ 1600K-1650K, see Figure 2.4.

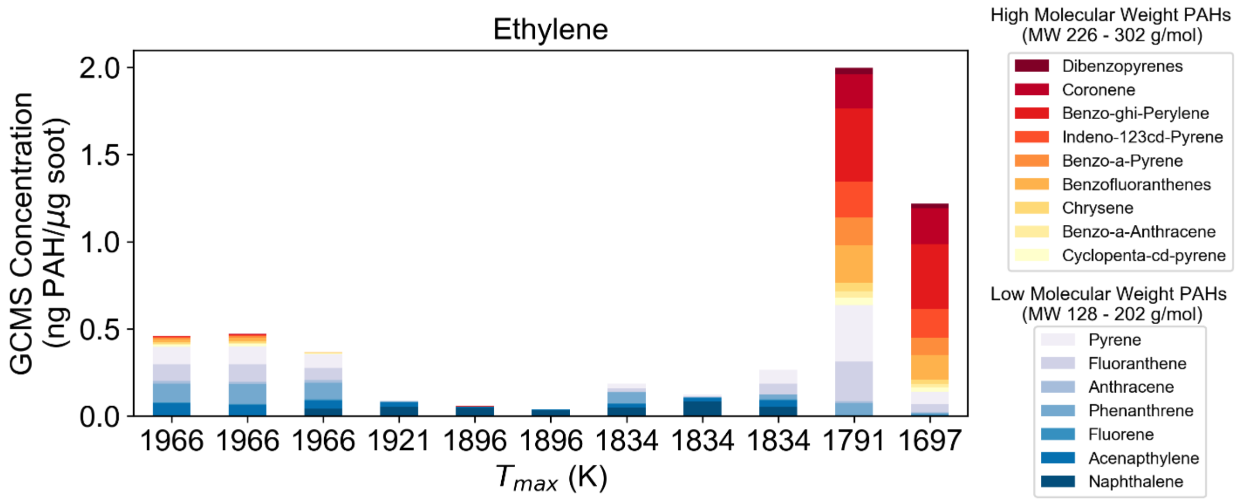


Figure 2.10: GCMS concentration of PAHs of three different ethylene soot samples having the same T_{max} are similar.

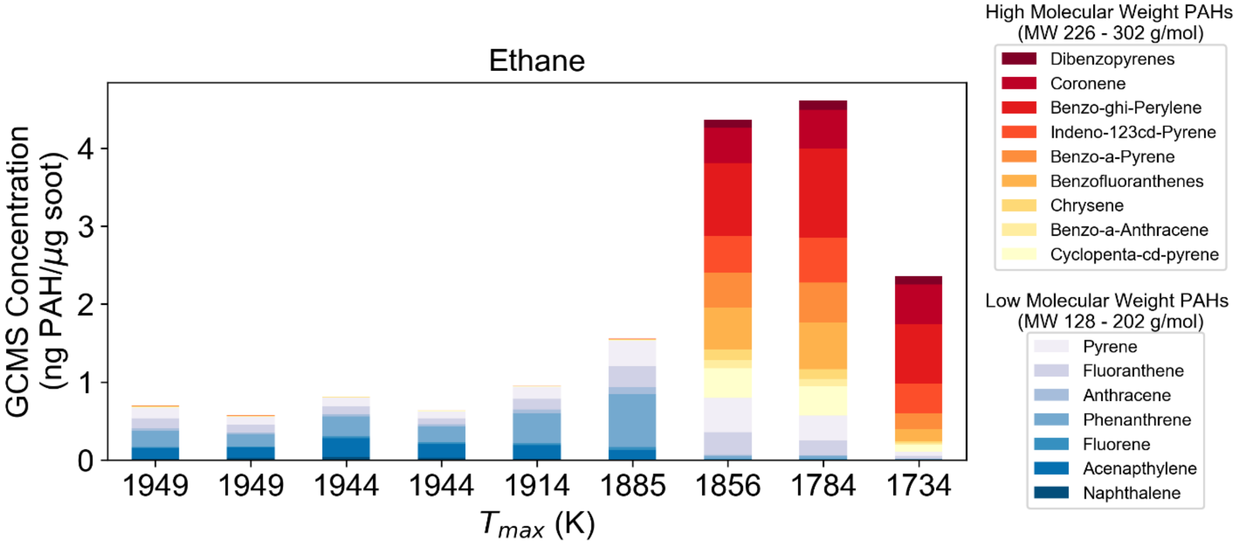


Figure 2.11: GCMS concentration of PAHs of three different ethane soot samples having the same T_{max} are similar.

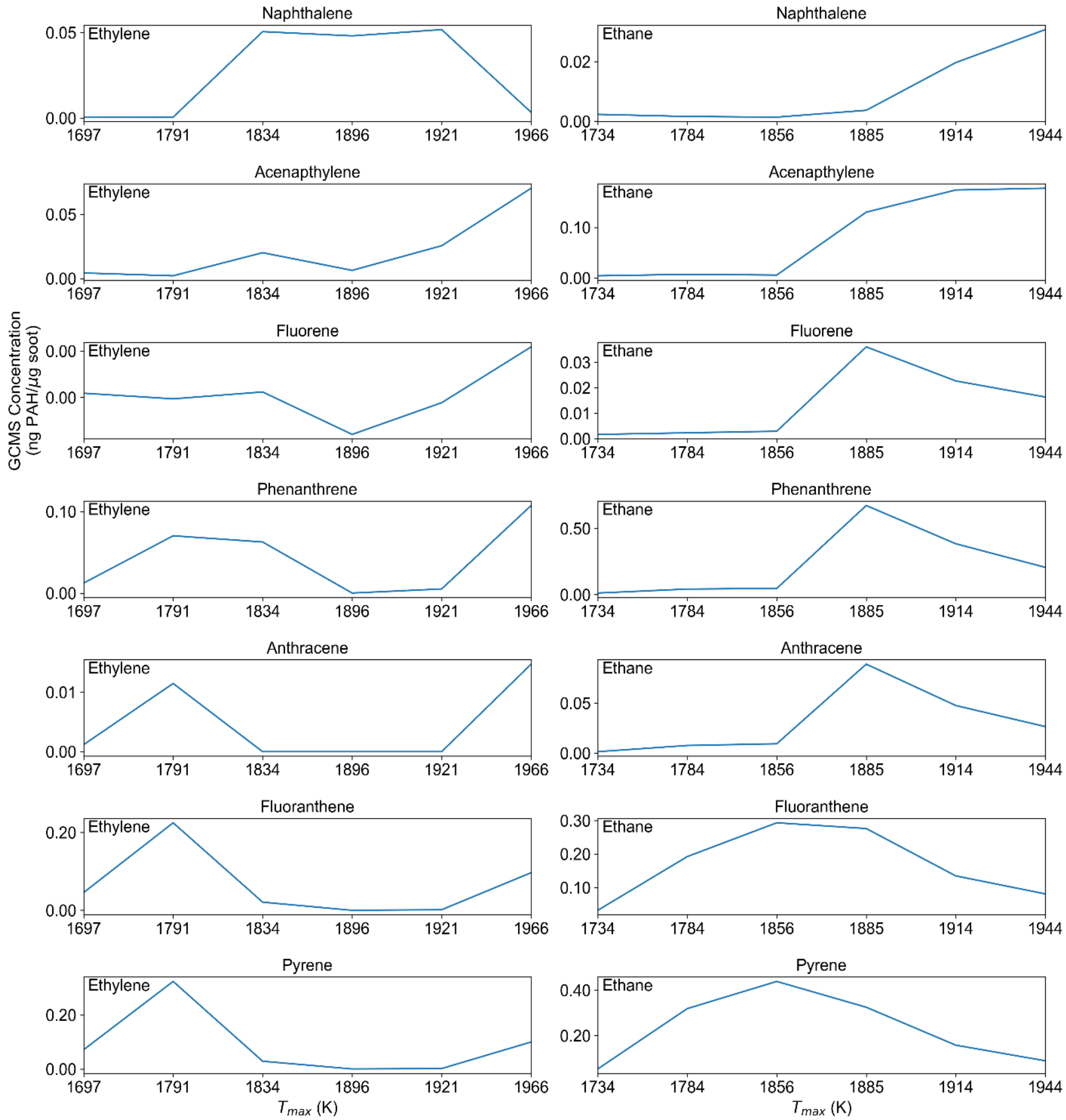


Figure 2.12: Variation of GCMS concentration of each of the LMW PAHs vs. T_{max} for ethylene and ethane fuels.

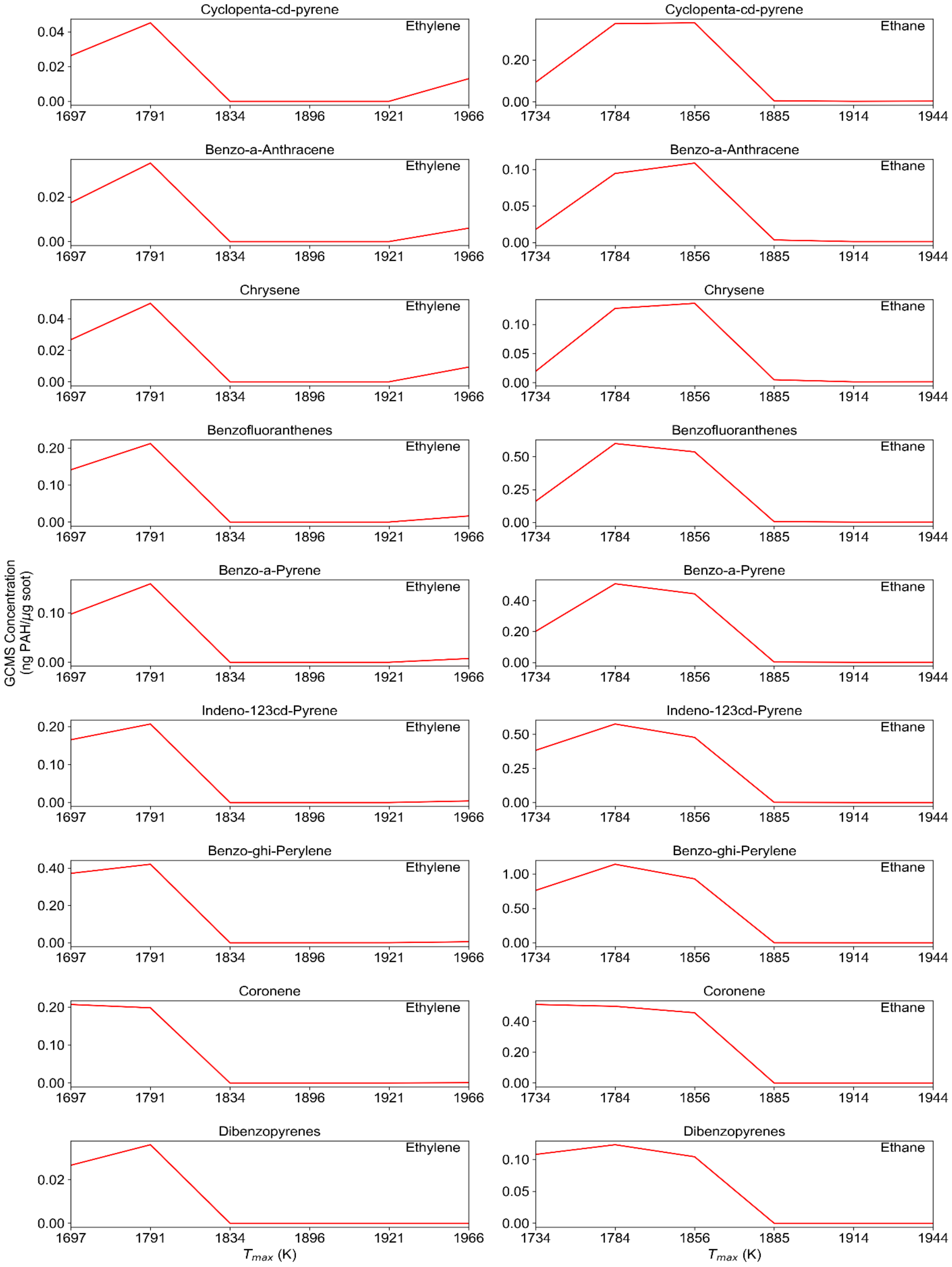


Figure 2.13: Variation in GCMS concentration of each of the HMW PAHs with varying T_{max} for both fuels. The HMW PAH concentration increases for low temperatures.

2.4 Chapter 2 Conclusions

The colder flames produce lower PM yield; however, the PM TOC content increased from 3-5% up to 13-27% for ethylene and ethane respectively. The increase in organic carbon is associated with the presence of HMW PAHs. This PAH group is associated with BrC emissions. Temperature thresholds for PM transition from low to high TOC PM were determined based on the maximum flame temperature $T_{max,c}$ ~1791K-1857K, which corresponds to the temperature of the highest soot luminosity (concentration) region of T^*_c~1600-1650K (see Figure 2.4).

Pyrene is four-ring aromatic PAH and common precursor to formation of graphitic structures and is used in the condensation models of soot surface interaction with SVOC species.[38,161] Pyrene has also been used as a model compound in particle inception models. Davis et al.[4] investigated the for mature soot (with core shell structure) formed in IGFR operating on ethane. The authors detemined that the average fringe lengths in the particle core 0.67- 0.71 nm range. Assuming D_{2h} symmetry[162], the cores are mostly composed of PAHs between pyrene and ovalene ($C_{32}H_{14}$); these PAHs are anticipated to be involved in particle nucleation[26]. The structured shell contains larger average fringe length of 0.77- 0.86 nm. Again, assuming D_{2h} symmetry, the shell is likely to be mostly composed of a graphitic structure formed from compound equivalent in size to ovalene and circumpyrene ($C_{42}H_{16}$). Thus, pyrene present a reasonable threshold for PAH sizes that are involved in the particle carbonization. PAH smaller than pyrene are present in the fuel-rich side of the flame at much higher concentration than pyrene[4] and are likely to participated in the condensation kinetics but not form the particle bound PAH.

Chapter 3 Excitation Emission Matrix Analysis to Evaluate Chemical Composition in Combustion Generated Particulate Matter

3.1 Summary

In this chapter, we use EEM to examine the presence and concentrations of PAHs with low molecular weight (LMW) -- lower than pyrene (MW<202) and higher molecular weight (HMW) -- higher than pyrene (226<MW<302) in combustion-generated PM. PAH analysis of PM performed via conventional laboratory techniques such as gas GCMS and LCMS is labor-intensive, require expensive equipment and require several hours for analysis. PAHs are good candidates for detection by spectroscopic techniques as they have high absorption coefficients and quantum yields. In general, fluorescence spectroscopy is highly sensitive (~1ng/mL) for PAH detection in ultrafine PM; however, it is not specific as overlap of fluorescence signatures of several fluorescing compounds in a multi-component mixture such as PM extracts is difficult to interpret. Principal component regression (PCR) analysis of the EEM spectra of IGFR samples correlates to GCMS data with $R^2 = 0.988$ for LMW and 0.998 for HMW PAHs. PCR-EEM analysis trained on the IGFR samples was applied to PM samples from woodsmoke and diesel exhaust, the model accurately predicts HMW PAH concentrations with $R^2 = 0.976$ and overestimates LMW PAHs. These results are published.[117]

3.2 Experimental Methods

3.2.1 Excitation Emission Matrix (EEM) Analysis

The EEM spectra of PM extract samples were obtained using a spectrofluorometer (Aqualog-880-C, HORIBA Instruments Inc. Edison, NJ, USA). The spectra were recorded in the range of excitation wavelength λ_{ex} = 200-600 nm with a 2 nm resolution. For each excitation wavelength, the instrument records emissions using CCD array in the range of λ_{em} = 246-826 nm with a 0.58 nm resolution. The spectra were normalized to Raman units (R.U.).[163] EEM spectra of blank cyclohexane were subtracted from each EEM spectra to account for Rayleigh scattering peaks. The remaining Rayleigh and Raman scattering peaks were removed computationally.[164] Absorption of PM extracts in the range λ_{ex} = 200-600 nm was recorded to correct the EEM measurements for inner filter effect.[165] The processed EEMs were then passed through a Gaussian filter (sigma=2) and negative values were removed numerically (MATLAB, MathWorks Inc., Natick Massachusetts, USA) to smooth the data (see Figure 3.1). The EEM spectra were normalized by the mass of PM in the solution.

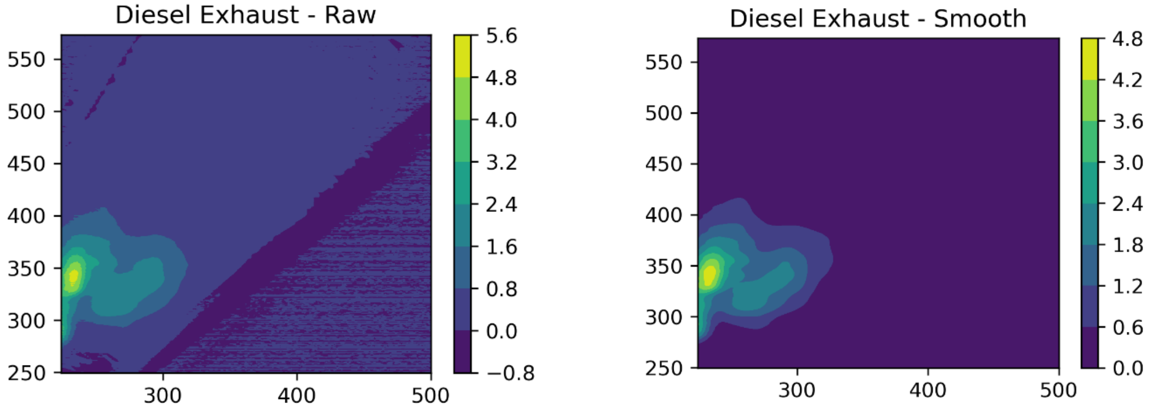


Figure 3.1: Comparison of raw vs. smooth EEM of diesel exhaust sample after passing through a Gaussian filter and removing negative values.

3.2.2 Principal Component Regression to estimate concentration of LMW and HMW PAHs in PM

PCR was performed as a two-step process: (1) transforming the EEM data onto its principal components (PCs) and (2) fitting regression models on the transformed EEM data and PAH concentrations obtained from GCMS analysis. EEM is recorded as a 2D matrix with a fluorescence intensity value for each pixel, which corresponds to one of 1000 row positions (one for each emission wavelength) and 201 column positions (one for each excitation wavelength).

In the first step of PCR, each EEM 2D matrix (1000 x 201) was unfolded into a 1D row vector (1 x 20100). The EEM row vectors of IGFR samples were stacked to create a data matrix C (n x 201000), where n is the total number of samples. The data matrix C was reduced with PCA using the scikit-learn library in Python 3.6.0.[166] The number of PCs (d), required to represent the data in the data matrix C was varied from 1 to 15. The data matrix C can be represented as the matrix product of s and v and a residual matrix E, shown in (3.1)

$$C = sv + E, \tag{3.1}$$

where s (n x d) is the scores matrix and v (d x 201000) is the loadings matrix having d PC vectors. The scores matrix s is the reduced form of data matrix C; the operation reduces the number of data points required to represent each EEM from 20100 to d. The value of d is chosen to minimize the number of PCs that account for the maximum variance in the EEM dataset. The variance in the EEM data accounted for by the PCs for d=1 to 15 is shown in Figure 3.3.

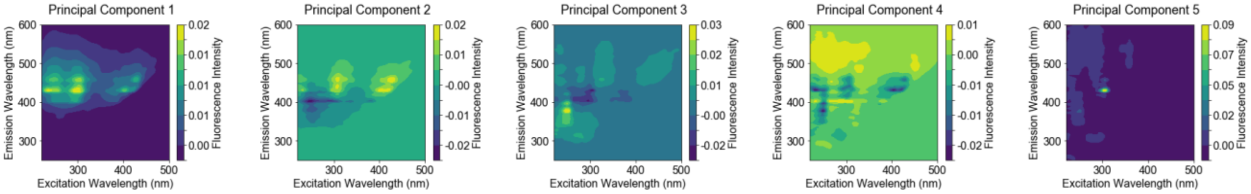


Figure 3.2: Principal Components of the IGFR EEM data are non-interpretable. The first five principal components account for 99.99% variance in data.

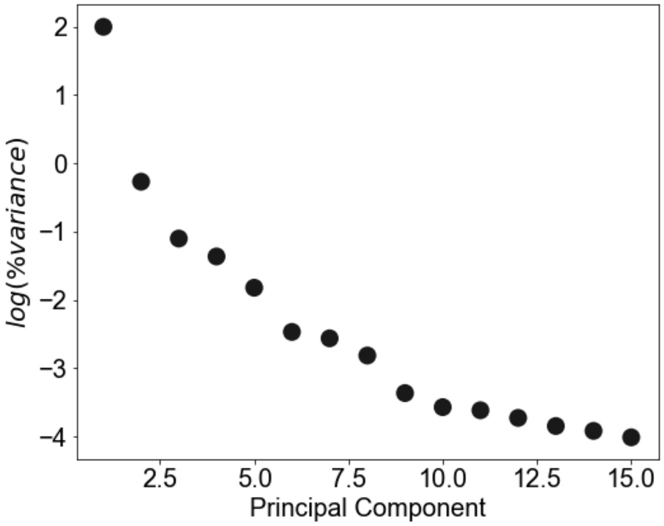


Figure 3.3: Contribution of each PC to the total variance of the EEM dataset is represented in terms of $log(\%variance)$.

For the second step of PCR, the scores matrix s (n x d) is used to fit linear models for estimating concentrations of PAH fractions. We used the ordinary least squares method in the scikit-learn library in Python 3.6.0 to perform multiple linear regression (MLR) to fit two linear models between: (i) EEM scores matrix and the LMW PAH concentrations from GCMS analysis and (ii) EEM scores matrix and HMW PAH concentrations from GCMS analysis. The mean squared error (MSE), representing the error in correlation for each model, was calculated using the "leave one out cross-validation" method (LOOCV).[166] The MSEs for range d=1-15 are shown in Figure 3.4 for LMW and HMW PAH groups. Based on the variance in the EEM data for d=1-15 PCs and the MSE, $d = 5$ and $d = 4$ were chosen for fitting the linear models for LMW and HMW PAHs, respectively, as these PCs accounted for 99.99% variance. Finally, PCA and MLR, referred together as principal component regression (PCR), was performed for LMW (d=5) and HMW PAHs (d=4) separately for all IGFR samples ($n = 20$).

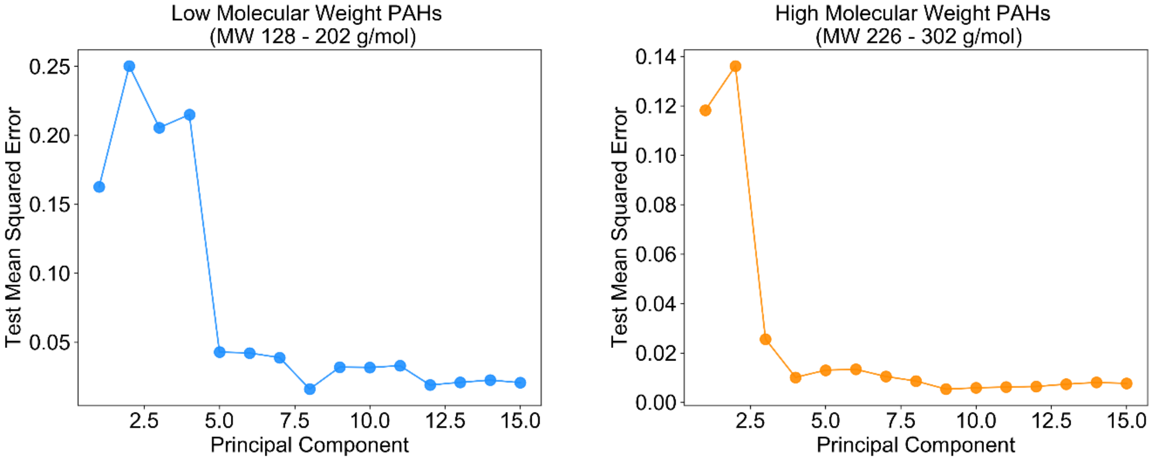


Figure 3.4: Mean Squared Error (MSE) calculated using leave one out cross-validation (LOOCV) for different number of PCs for LMW and HMW PAHs.

The models for estimating LMW and HMW PAH concentrations trained on IGFR samples were tested using the woodsmoke and diesel exhaust samples. The EEM row vectors of the test samples were stacked to create a test data matrix C_{test} (6 x 201000), where 6 is the total number of test samples. The test data matrix C_{test} was reduced along the PCs obtained for IGFR samples to obtain a test scores matrix (6 x d). The coefficient of determination (R^2) for both models are reported.

3.2.1 Matrix Effects on EEM analysis

In this section, we explore the performance of the PCR-EEM method for PM matrix effects. We consider two situations: 1. PM is from a mixture of real-world sources 2. Aging of the sample in the environment.

3.2.1.1 PM Mixture Samples Generation and Analysis

To evaluate the PCR-EEM method for PM mixtures, we perform computational mixing of EEM data and corresponding GCMS concentration of PAHs to generate 16 new samples from the six diesel-exhaust and woodsmoke samples used for testing the PCR-EEM method. For 16 mathematically mixed samples, three samples were picked randomly from the six original test samples, their EEM data and PAH concentrations were multiplied by a constant value between 2/15 and 8/15 (1/3 ± 1/5) and then added to generate a new EEM and its corresponding PAH concentration profile. The EEM generated by computational mixing is similar to the EEM recorded from physically mixed PM samples.[71] GCMS concentration of PAHs can be computationally mixed to generate new PAH profiles as PAH fluorescence intensity vary linearly with their concentration.[167]

3.2.1.2 PM aging effect on EEM spectra

Although PM aging effects on PAH composition are not part of the study, in a pilot experiment, we compare EEMs of young and aged diesel exhaust. Diesel PM was generated by a 435 cc direct-injection single-cylinder diesel engine (Yanmar LW Series, Adairsville GA, USA) operated on ultra-low-sulfur diesel. Young diesel exhaust PM was collected from an exhaust duct connecting the diesel engine exhaust to the exposure room at the University of Washington’s controlled inhalation diesel exhaust exposure facility by sampling devices placed in a sealed chamber.[145] Aged diesel exhaust PM was collected in the exposure room.[71] Note that the aged exhaust was not exposed to long term solar radiation or temperature change, and was not mixed with other sources. Although PM aging occurs for only a few minutes, there is a change in PM concentration and environmental exposure between young and aged PM.

3.2.2 Estimating Concentrations for Individual PAHs Using PCR-EEM Method

3.2.2.1 Creating EEMs

The training dataset consists of IGFR samples, which has 20 EEM matrices of size 1000 X 201 and their corresponding HMW PAH concentrations. The test dataset consists of woodsmoke and diesel exhaust samples which have 6 EEM image matrices of size 1000 X 201 and their corresponding HMW PAH concentrations. We reduce the size of the EEMs to 100 X 33 using 2D MaxPool in Python with filter size (10,6).

To generate more data for training and testing of the PCR model, we compute linear combinations of the existing EEMs and their known HMW PAH concentrations. 580 EEMs were generated from the training set and 144 from the test set. Each generated EEM is a linear combination of three randomly chosen EEMs from its respective set with each chosen EEM matrix is multiplied by a random number between 2/15 and 8/15 (1/3 ± 1/5). The EEM spectra generated by computational mixing is similar to the EEM recorded from physically mixed PM samples.[71] GCMS concentration of PAHs can be computationally mixed to generate new PAH profiles as PAH concentrations vary linearly with fluorescence intensity of PAHs.[167]

3.2.2.2 Principal Component Regression to Estimate Concentration single HMW PAHs.

Each EEM in the training set is unfolded as a row array and stacked to create an n x p matrix where n is the number of training samples (n=600) and p is the number of pixels (p=3300). The resultant data matrix is demeaned column-wise. The 'PCA' function from class 'sklearn.decomposition' is used to decompose the data matrix. It uses the LAPACK implementation of the full SVD or a randomized truncated SVD, depending on the shape of the input data and the number of components to extract. We plot the percentage variance explained by each component vs. the number of components in Figure 3.5. A score matrix obtained after PCA of EEMs is representative of each EEM. The dimension of the scores for each EEM is 600 which is reduced to d based on Figure 3.5. The reduced data matrix n x d represents the training data.

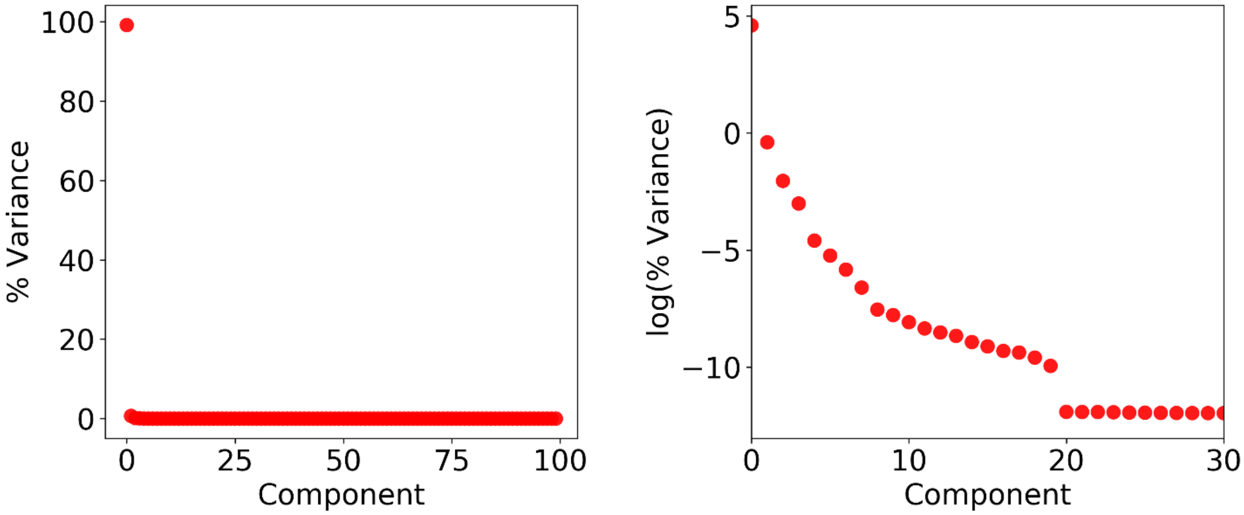


Figure 3.5: % Variance for each component (left). Since the first component accounts for most of the variance, a log plot gives more information on the relative contribution of other components to the variance (right).

The 'linear_model' function from class 'sklearn' is used for multiple linear regression analysis. 'LinearRegression' fits a linear model with coefficients to minimize the residual sum of squares between the observed responses in the dataset. We use 'LinearRegression' to compute a linear model between reduced data matrix of size n x d and training labels of size n x 9 (Concentration of 9 HMW PAHs). The training error is calculated using the linear model. Each EEM in the test set is unfolded and demeaned the same way as EEMs in training set to create a matrix of size m x p, where m (m=150) is the number of test samples. The function 'pca.transform()' is used to obtain scores of test samples based on dimensionality reduction of training samples. First, d scores out of 150 scores are chosen. Test PAH concentrations are predicted based on the d scores for each test sample using the linear model developed using the training set. The test error is calculated.

3.3 Results and Discussion

3.3.1 Spectral Information from EEM Analysis

Figure 3.6 shows the EEM spectra of PM generated in IGFR operated on ethylene and ethane for three T_{max} conditions and EEM spectra of PM from the woodsmoke and diesel exhaust. For IGFR samples, the location and intensity of fluorescence emission peaks varies with T_{max}. At T_{max}=1966K for ethylene and T_{max}=1949K for ethane, the emission peaks are present at shorter excitation wavelengths and have lower intensities. At T_{max}=1697K for ethylene and T_{max}=1734K for ethane, the range of excitation wavelengths where fluorescence peaks are observed, increases, indicating a broadening of PAH absorption bands, and the emission peaks shift towards longer wavelengths. This suggests that colder flames produce PM with higher MW PAHs than hotter (less diluted) flames, which agrees with the GCMS analysis, as shown in Figure 2.9, and is also consistent for woodsmoke and diesel exhaust samples. The diesel cycle generates particle BC particles at high temperature; the EEMs show shorter excitation and emission wavelengths and have lower intensities. On the other hand, woodsmoke generates BrC at a lower temperature; the emission peaks have in a broader excitation wavelength range and longer emission wavelengths.

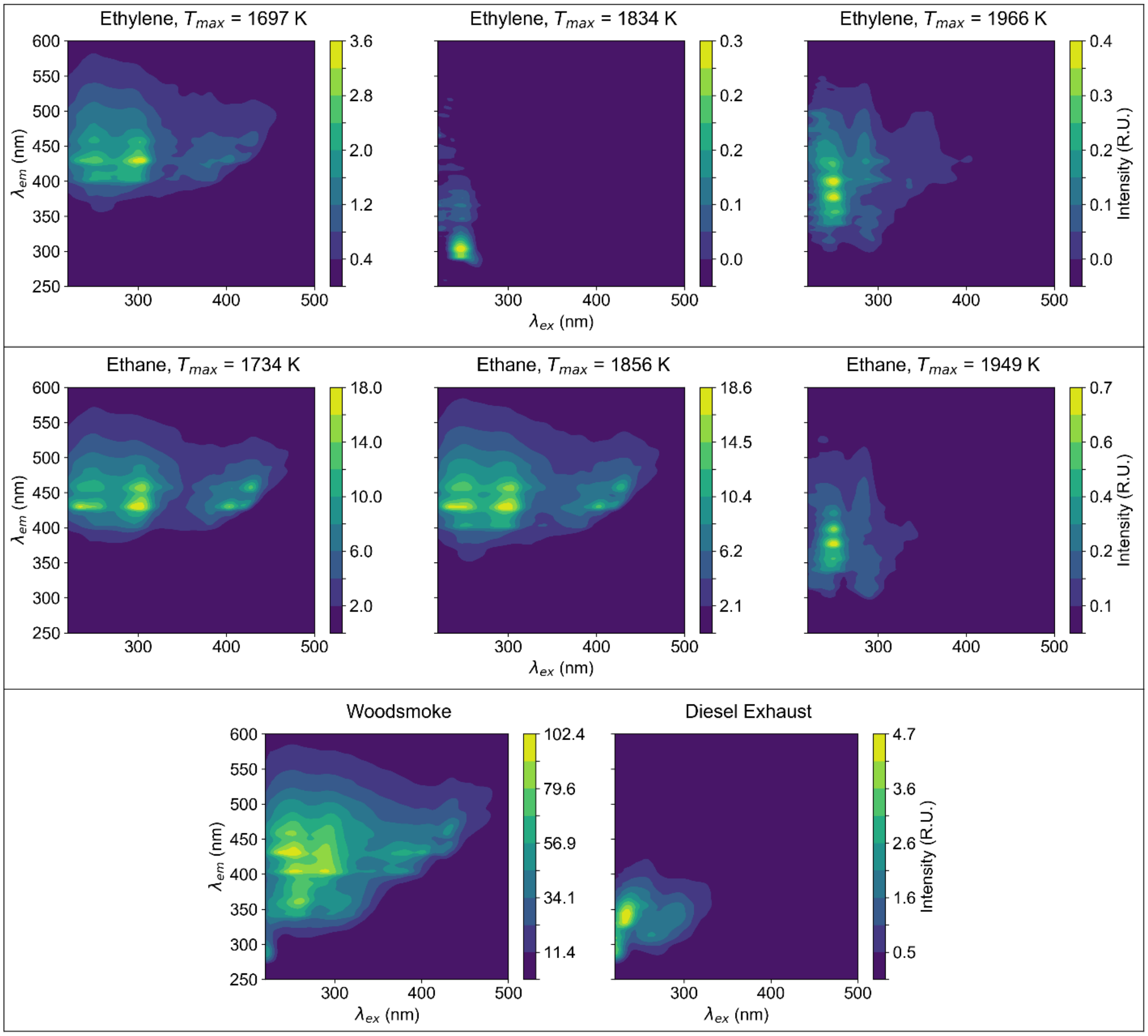


Figure 3.6: EEM spectra for IGFR flames, woodsmoke, and diesel exhaust. Emission intensities are in Raman Units (R.U.). The location and intensity of fluorescence peaks vary with temperature. Diesel PM sample has emission peaks at shorter emission wavelengths and in narrow excitation bands. PM samples from woodsmoke have emission peaks at longer emission wavelengths and broader excitation bands.

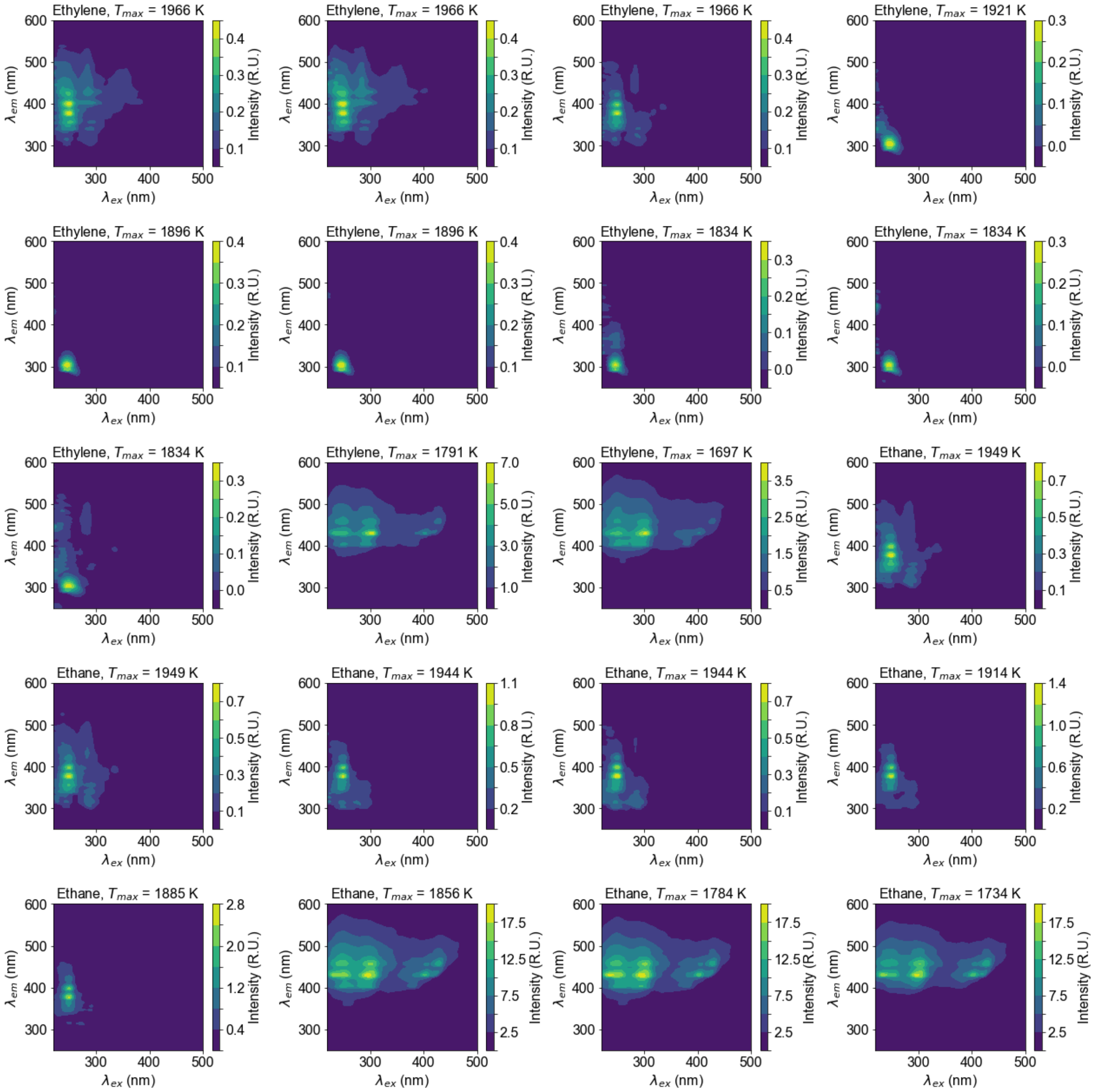


Figure 3.7: EEM spectra of PM from the combustion of ethylene and ethane fuels in the entire range of T_{max} considered in this study.

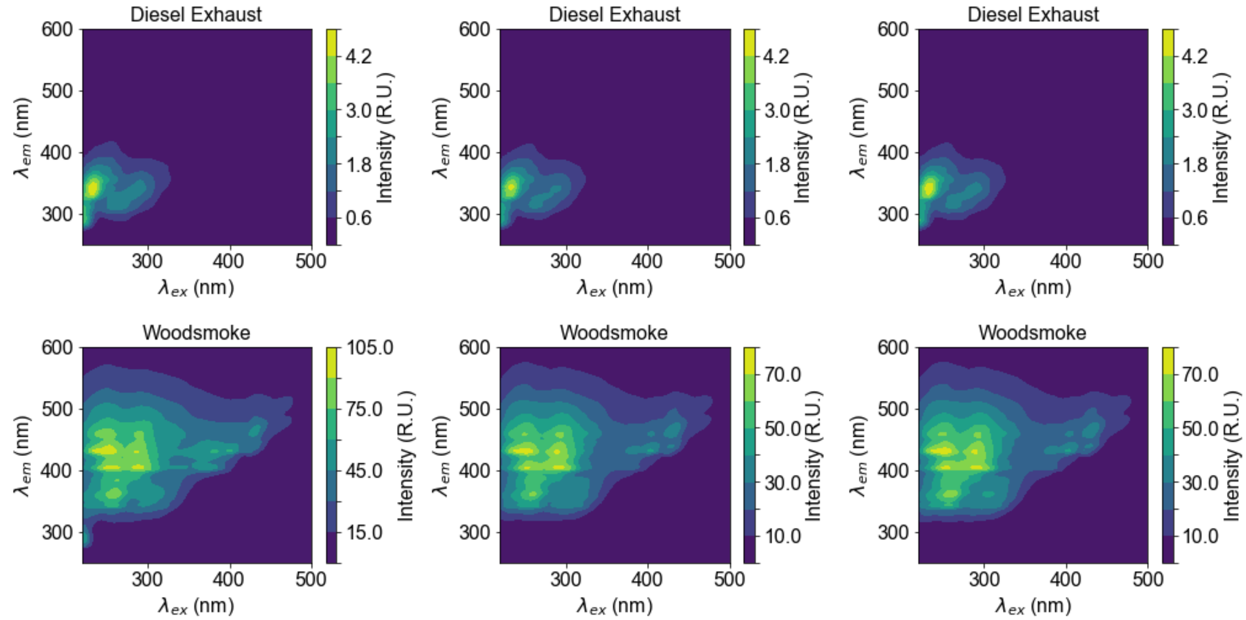
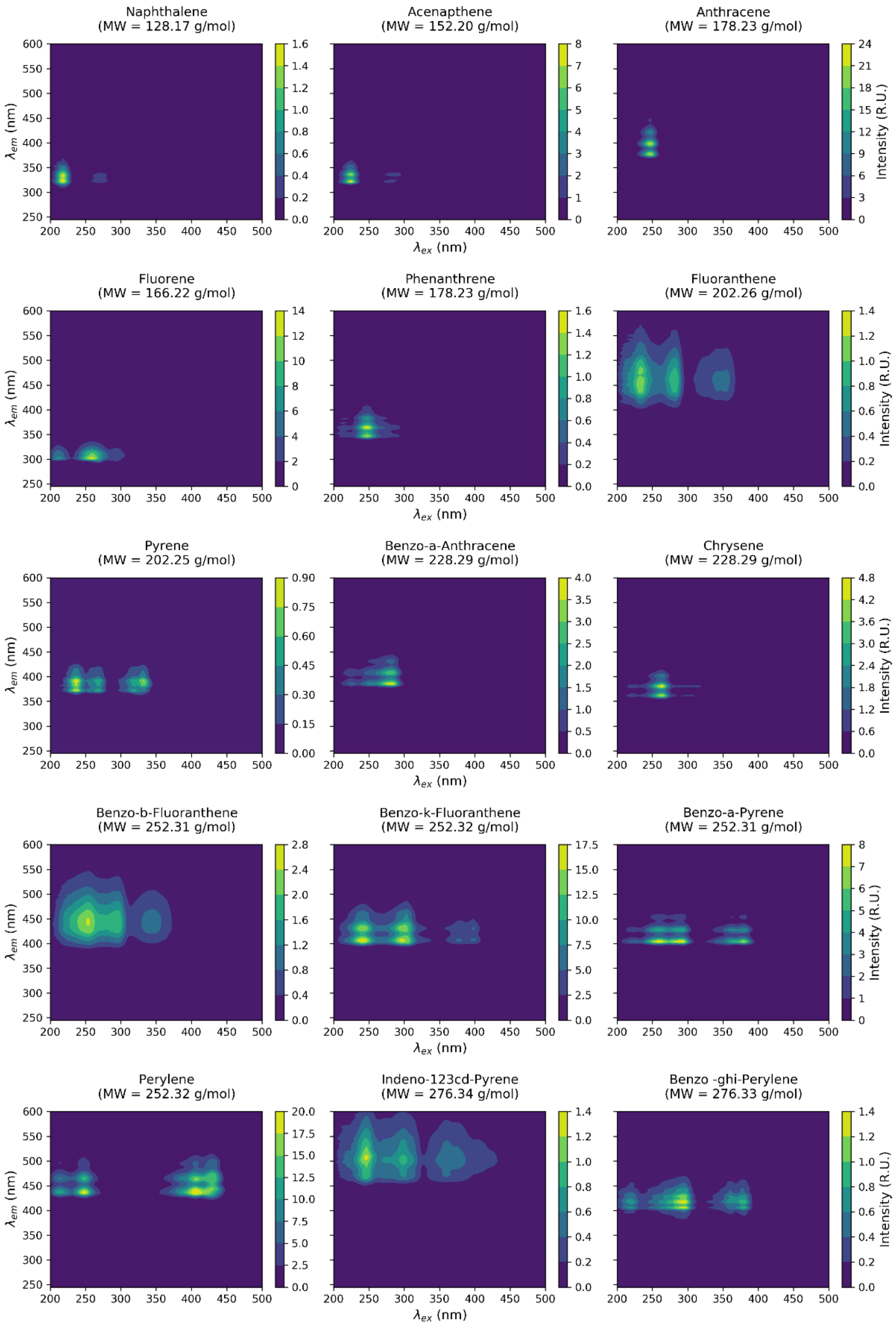


Figure 3.8: EEMs of three different samples of the same source have similar peaks for diesel exhaust and woodsmoke.

The EEM spectra of IGFR samples for the entire range of flame conditions are shown in Figure 3.7, and EEM spectra of all 6 test samples are shown in Figure 3.8. Figure 3.9 shows the emission spectra for pure PAHs in the target molecular weight range. In general, as the MW of PAHs increase, their emission shifts towards longer emission wavelengths and broader excitation wavelength ranges.[72] Similar to the GCMS data, the temperature threshold correlates to the change in EEM fluorescent peaks. Note that the EEM signal intensities are significantly greater for lower temperature samples due to the high fluorescence quantum yield of HMW PAHs (sample mass normalized).[168] The odd result of having HMW PAHs for the hottest ethylene condition, T_{max}=1966K, as pointed out in the GCMS analysis, is also observed in EEM analysis via the fluorescence peaks associated with HMW PAHs at λ_{ex} = 300 – 400 nm and λ_{em} = 400 – 500 nm.



Naphthalene (MW = 128.17 g/mol)

Acenapthene (MW = 152.20 g/mol)

Anthracene (MW = 178.23 g/mol)

Fluorene (MW = 166.22 g/mol)

Phenanthrene (MW = 178.23 g/mol)

Fluoranthene (MW = 202.26 g/mol)

Pyrene (MW = 202.25 g/mol)

Benzo-a-Anthracene (MW = 228.29 g/mol)

Chrysene (MW = 228.29 g/mol)

Benzo-b-Fluoranthene (MW = 252.31 g/mol)

Benzo-k-Fluoranthene (MW = 252.32 g/mol)

Benzo-a-Pyrene (MW = 252.31 g/mol)

Perylene (MW = 252.32 g/mol)

Indeno-123cd-Pyrene (MW = 276.34 g/mol)

Benzo -ghi-Perylene (MW = 276.33 g/mol)

Figure 3.9: EEM spectra of 15 of the 24 PAHs analyzed using GCMS.

3.3.2 Total Fluorescence Analysis

Figure 3.10 compares the total PAH concentrations in PM samples from GCMS analysis against the total fluorescence intensity, defined as the sum of intensities at each excitation-emission pair in the EEM. IGFR samples with T_{max} > $T_{max,c}$ and diesel exhaust have predominantly LMW PAHs (see Figure 2.9), thus variation in the total fluorescence intensity can be attributed to the variation in predominantly LMW PAH concentrations. LMW PAH fluorescence peaks are located in the lower-left quadrant of the EEM spectra at shorter excitation and emission wavelengths ($\lambda_{ex} < 300nm; \lambda_{em} < 450nm$), as shown in Figure 3.6. Similarly, variation in the total fluorescence intensity in PM samples for T_{max} < $T_{max,c}$ and woodsmoke can be attributed to the variation in concentrations of HMW PAHs; their fluorescence peaks located at higher wavelengths ($\lambda_{ex} > 300nm;\ \lambda_{em} > 400nm$). There is a clear distinction between emission wavelength and width of excitation bands in the EEMs of PM samples with dominant LMW PAHs vs. HMW PAHs. The integrated fluorescence intensity follows the total PAH concentration levels for all samples. Compounds other than the 24 PAHs from GCMS panel contribute to the fluorescence; however, their peak location in the EEM spectra and relative intensity are not known.

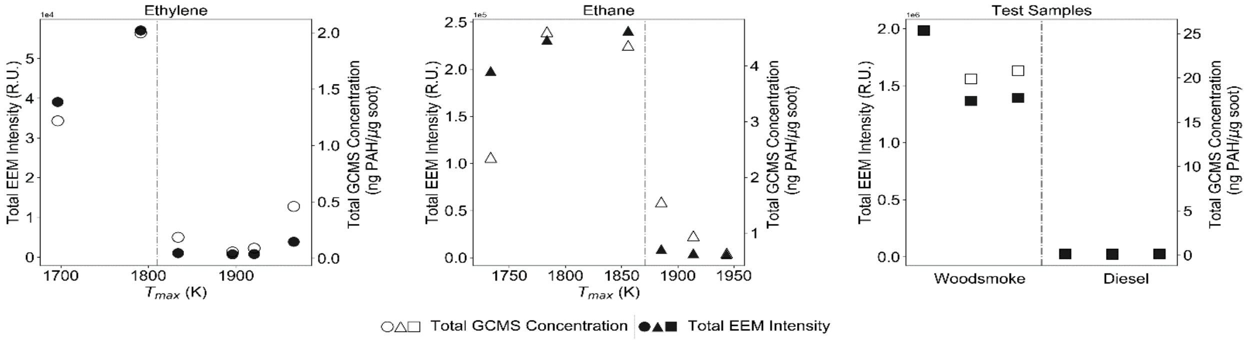


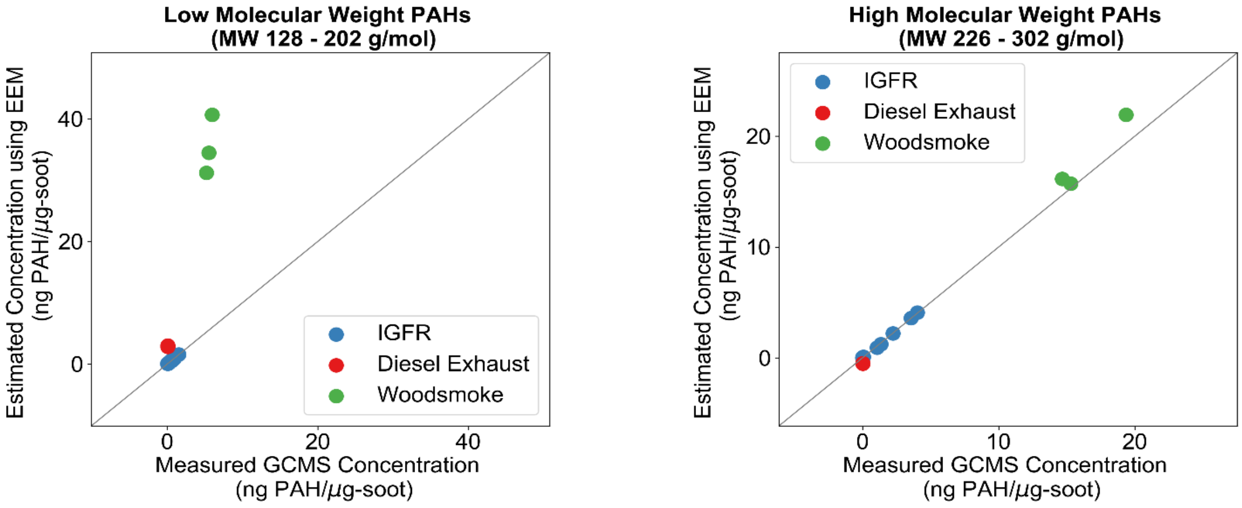
Figure 3.10: Comparison of total EEM fluorescence to PAH concentration from GCMS analysis of 24 PAHs panel for PM from diluted IGFR ethylene and ethane flames, the woodsmoke, and diesel exhaust. The dotted line represents transition threshold from black carbon (low TOC, low HMW PAHs) to brown carbon (high TOC, high HMW PAHs).

3.3.3 Application of IGFR-trained EEM Model to Real-world Combustion Sources

This section describes the application of PCR-EEM model training on IGFR samples for two different fuels, over a range of Ar dilution ratios, and application of the model to real-world combustion sources, i.e., fresh cookstove PM samples and the fresh diesel exhaust PM emission. Figure 3.11 shows the estimated LMW and HMW PAH concentrations based on the PCR-EEM analysis. When considering only the IGFR samples, the correlations between PCR model and GCMS data have R^2 values of 0.988 and 0.998 for the LMW and HMW PAHs, respectively. Although the intensity of the EEM fluorescence peaks for high-temperature PM samples is lower than that for colder flames (see Figure 3.10), the unique attributes of the spectra, i.e., peak locations, widths, and other overall shape parameters, allow for accurate estimation of LMW PAH concentrations in IGFR samples. The fluorescence intensity in the specific EEM region, indicative of the concentration of LMW and HMW PAHs, can be used as a fluorescent fingerprint for organic compounds in combustion PM. For samples where both LMW and HMW PAH groups are present, their emission wavelengths and excitation bands overlap, and it is necessary to determine the fraction of each group.

Here, the model trained on the IGFR samples is tested on woodsmoke and diesel exhaust samples. The model predicts HMW PAH concentrations well (R^2=0.976), and it overestimates LMW PAH concentrations in the woodsmoke and diesel exhaust samples. As the samples from these combustion sources were extracted and analysed by GCMS and EEM in the same manner as IGFR samples, we cannot attribute the difference to PM aging or other matrix effects. However, it is likely that the formation and destruction of the LMW PM-PAHs follow different routes in different combustion systems. These routes are likely to be affected by the fuel type (gaseous, liquid, solid), combustion conditions (smoldering vs flaming combustion), soot oxidation process, etc. However, the similarity in HMW PAHs between different combustion sources suggests the soot maturation process (transition from BrC to BC) may follow similar routes. More detailed investigations of these phenomena are required.

Although the PCR-EEM method can be used to analyse PAHs in PM having concentrations as low as 0.1 ng PAH/µg soot, it is not specific in terms of determining the concentrations of individual PAHs due to the overlap in fluorescence peaks and presence of other fluorescent species which were not analysed in the present study. Multivariate analysis techniques can improve the accuracy and specificity of this method if the data used to develop similar models includes a greater variation in concentration of the specific PAH target.



R^2_{IGFR} = 0.988 R^2_{D+WS} = -59.410

R^2_{IGFR} = 0.998 R^2_{D+WS} = 0.976

Figure 3.11: Parity plots of PAH concentrations estimated by PCR-EEM vs. concentration measured by GCMS for LMW PAHs (left) and HMW PAHs (right) in PM samples from IGFR, diesel exhaust and woodsmoke (D+WS).

3.3.4 PCR-EEM Method with PM Matrix Effects

We performed computational mixing of EEM data and corresponding GCMS concentration of PAHs to generate 16 new samples from the six diesel-exhaust and woodsmoke samples used for testing the PCR-EEM method. Computationally mixed 16 new samples were added to the original six test samples to create a new test dataset having 22 samples. The PCR-EEM models for LMW and HMW PAH concentration estimation, developed using IGFR samples, were tested using the new test dataset. The results in Figure 3.12 show that the R^2 for estimating HMW PAH concentrations is 0.973 while the LMW PAH concentrations are overestimated. These correlations are similar to the ones when the original test dataset with 6 samples was used. Hence, PM mixing does not affect the PCR-EEM estimates.

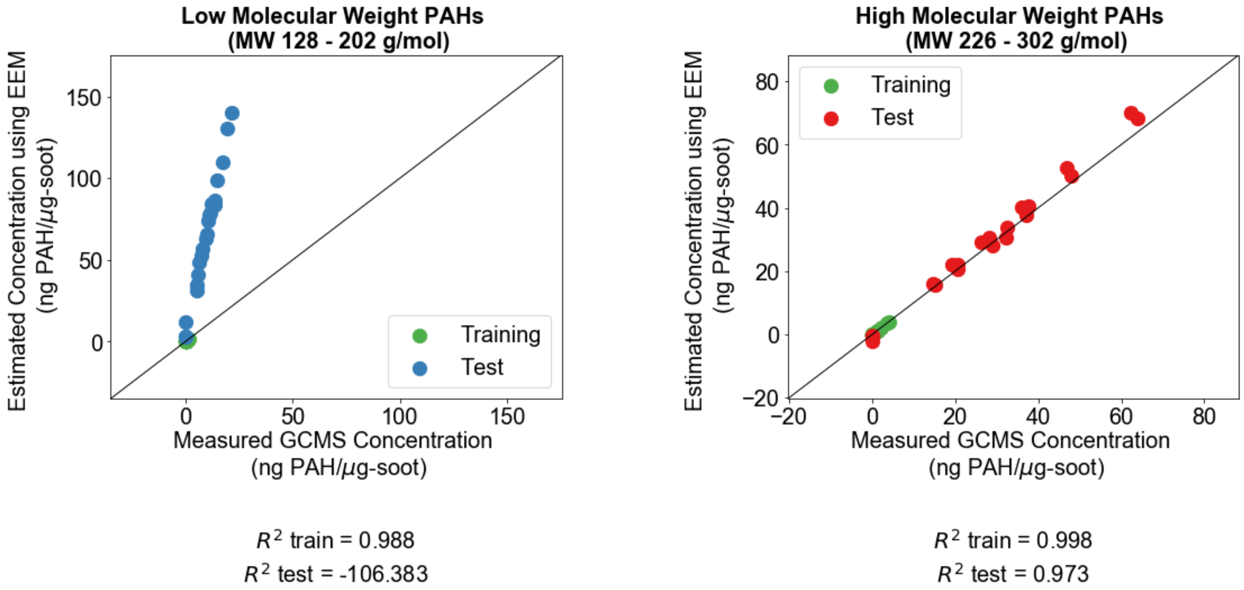


Figure 3.12: Parity plots of PAH concentrations estimated by PCR-EEM vs. concentration measured by GCMS for LMW (left) and HMW (right) PAHs in PM samples from ethylene and ethane flames (training), an original and computationally created woodsmoke and diesel exhaust (test). PM mixing does not affect PCR-EEM method performance.

Lower molecular weight compounds are overpredicted if training data includes only the IGFR samples. Though we no do have full insight explaining the discrepancy, the more volatile compounds are significantly affected during sample collection, and photodegradation also changes PAH composition in young vs. aged soot.[169] Perhaps the greatest uncertainty is the presence of the other fluorescent unsubstituted PAHs, alkylated PAHs, nitrated PAHs (N-PAHs), oxygenated PAHs (oxy-PAHs) and non-PAH species in these combustion sources. The route to PM formation for gaseous, liquid and solid fuel is likely to influence the composition of organic carbon fraction. While simple gaseous fuels were used in the training set, diesel exhaust and the biomass combustion are significantly more complex. The initial fuel pyrolysis step produces a great variety of the C-H compounds that are not present in simple flames.[170] We also found that high molecular weight compounds in PM are more stable (fluorescent signature of the extract does not change) during sample collection. The prediction of the HMW compounds over combustion source tested is in excellent agreement with the original training set, see Figure 3.12.

Limitations in the analysis of LMW PAH compounds require further investigations, possible reasons for the discrepancies between the training and test sets can be related to sampling procedures, analysis of spectra, and differences in fuel type. Some of these issues may be resolved by retraining the PCR algorithm using different training data. (i) Related to sampling procedure, PAHs<Pyrene considered in the analysis are more volatile, their boiling point range is in 240-400 °C, while PAHs>Pyrene have boiling point > 400°C.109,110 The high volatility hydrocarbons are more likely to be influenced by the difference in the temperature and flow rates during the sample collection from laboratory flame and other environmental sources. (ii) Related to spectra analysis, LMW compounds fluoresce in a narrow UV wavelength range resulting in overlapping signatures which are harder to deconvolute than signatures of HMW compounds, which occupy a larger EEM region. (iii) Fuel type and combustion regime may influence the abundance of LMW hydrocarbons. Species associated with fuel pyrolysis, particle formation and maturation can be significantly different for liquid fuels or solid fuels (used in the testing set) from the simple gaseous fuels (used in the training set). PAHs>Pyrene have been shown to be fixed in the particle structure and the carbonization chemistry is likely to be similar for different types of fuels. Thus, the comparison between different combustion sources yields good agreement for HMW PAHs.

Figure 3.13 shows EEMs of young and aged diesel exhaust. The fluorescent peak with the highest intensity is in the same location for both samples; however, other fluorescent regions have slightly different features. Also, maximum fluorescence intensity for the aged sample is less than the young sample (The samples are concentration normalized). This might be reflective of loss in volatile PAHs during aging. This suggests that EEM measurements might be capable of showing aging effects and that the PCR-EEM method can provide an accurate estimation of HMW PAH concentrations for young as well as aged PM samples.

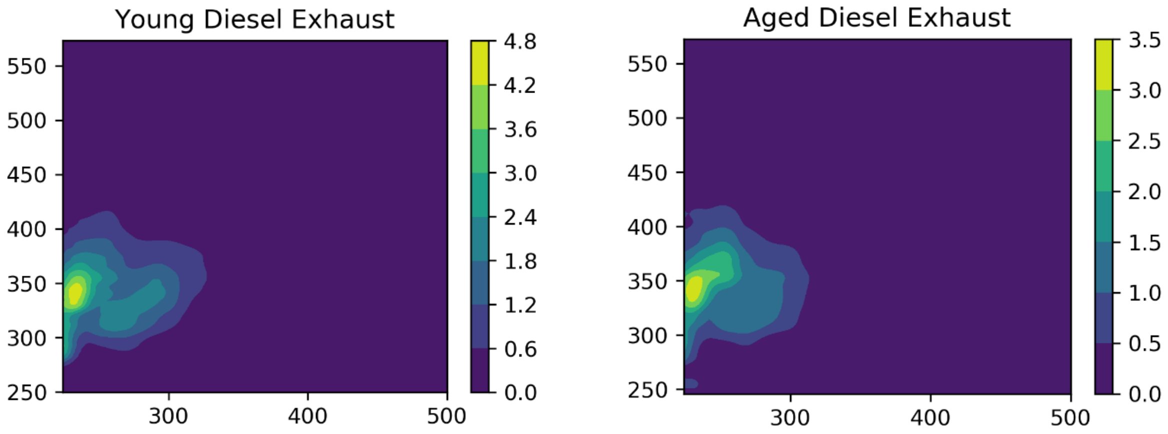


Figure 3.13: EEMs of young vs. aged PM diesel exhaust have slightly different fluorescent peaks and intensities which are reflective of PM chemistry evolution during aging.

3.3.5 Performance of Fitted PCR-EEM Model to Predict 9 HMW PAH concentrations

Figure 3.5 (right) shows that % variance elbows beyond 19 principal components, hence we choose d=19, in this case the EEM matrix dimension is reduced from 3300 to 19 using PCA. A linear model obtained for scores and HMW PAH concentrations of samples in the training dataset is used to estimate HMW PAH concentrations of samples in the test dataset. The error for estimating HMW PAH concentrations for samples in the test dataset is shown in Figure 3.14. The %error is low for Benzofluoranthenes (BFL) and Indeno-123cd-Pyrene (INP) while estimates of other HMW PAH concentrations have higher error.

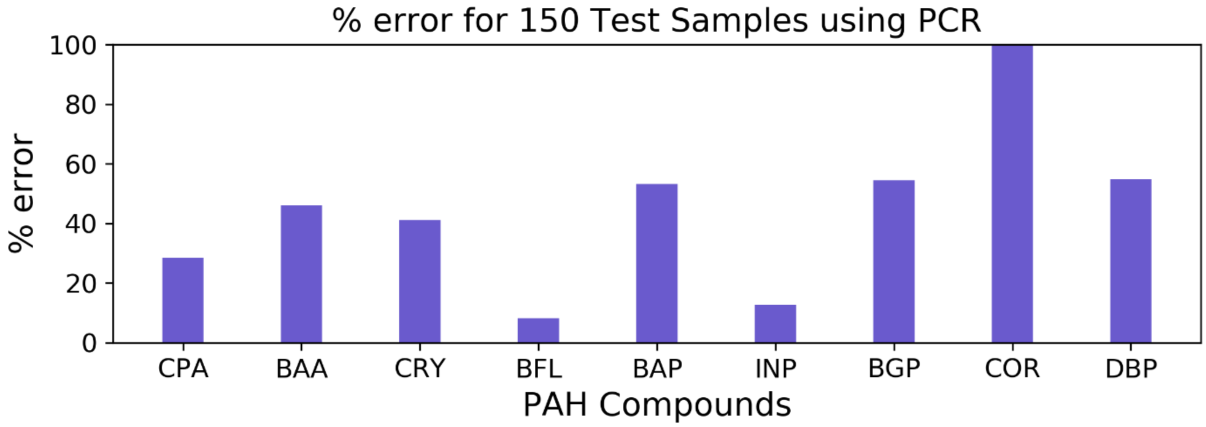


Figure 3.14: The error for all 9 HMW PAH compounds. The average %error is 36.6%.

The PCR-EEM method estimates HMW PAH concentrations for single PAHs with reasonable accuracy for only a couple of individual PAHs. Although the relationship between PAH fluorescence intensity and PAH concentration in PM is linear[167], there are hundreds other fluorescing species in combustion generated PM. The fluorescence spectra of these species overlap. This makes predicting concentration of single PAHs difficult, especially when the number of fluorescent species is high as expected in woodsmoke and diesel exhaust samples due to the complex nature of combustion mechanisms and kinetics for these fuels. The HMW PAH group seems to represent the dominant fluorophores fluorescing at longer wavelengths and broad excitation wavelength ranges for samples in the training and test set as the estimates of HMW PAH concentrations are accurate with the PCR model, hence PCR-EEM method is appropriate for estimating concentrations for a group of PAHs but may not be accurate for estimating the concentration of single PAH compounds in the complex mixtures compounded by the possible aging effects.

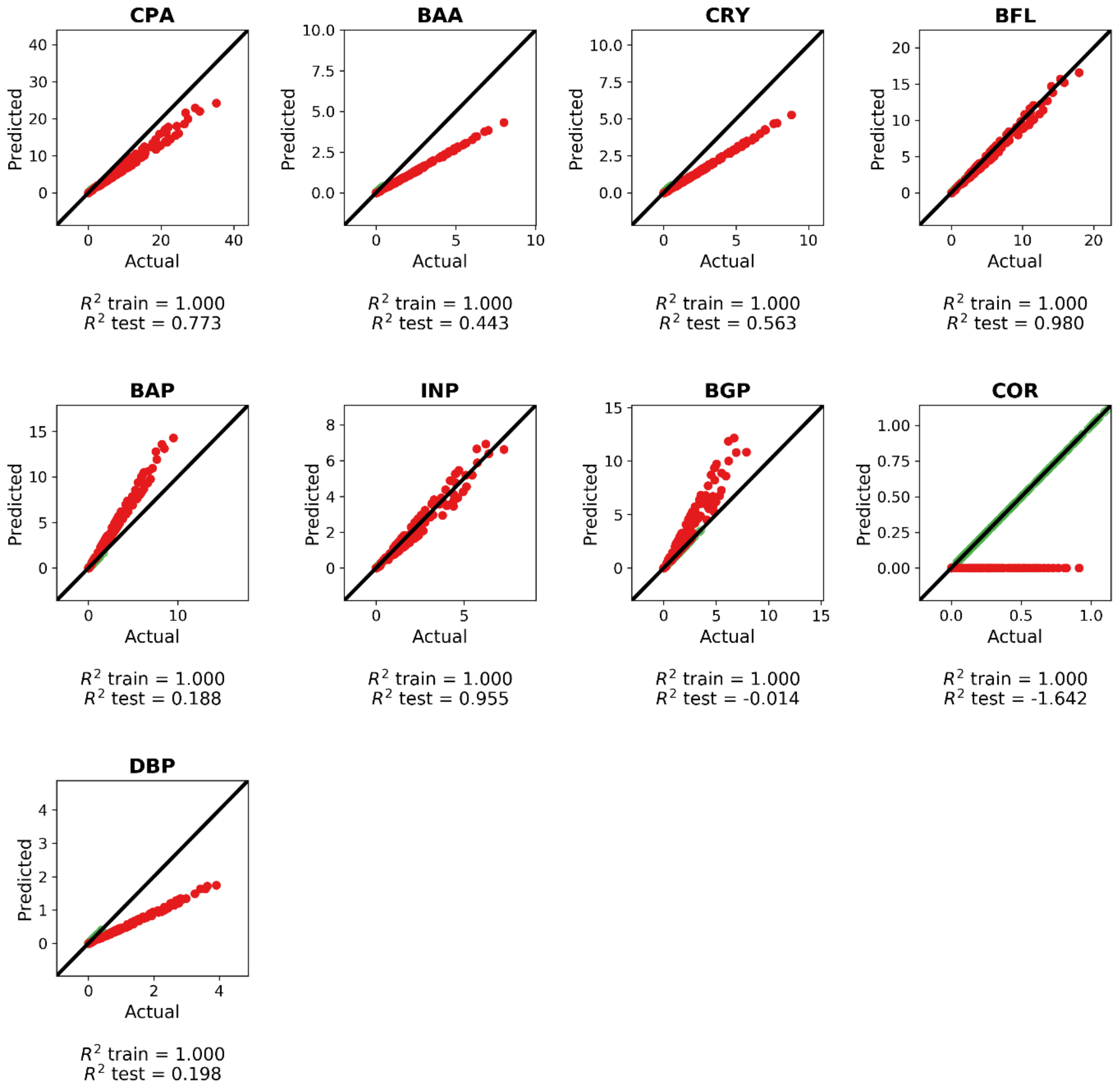


Figure 3.15: R^2 values for training (green) and test (red) for estimating HMW PAH concentration for the individual PAH using the PCR-EEM method. Benzofluoranthenes (BFL) and Indeno-123cd-Pyrene (INP) have test R^2 close to 1.

3.4 Chapter 3 Conclusions

The PCR-EEM method based on fluorescent fingerprint attributes, such as location, width, overall shape, and intensity of the three-dimensional spectra, can be used to estimate HMW PAHs in combustion generated aerosols. Although the relationship between PAH fluorescence intensity and PAH concentration is linear[167], the presence of multiple fluorescent species with overlapping spectra complicates the analysis of individual compounds. Our attempt to determine the concentration of individual PAH from the EEM spectra by comparing it with GCMS analysis did not yield a good agreement. Only BFL and INP could be estimated within 20% accuracy, see Figure 3.14 and Figure 3.15. Overall, the HMW PAHs having strong and unique EEM fingerprints (at longer wavelengths and with broad excitation wavelength) show the best correlation with GCMS data for laboratory flames and real-world combustion sources; thus the PCR-EEM model for this group of compounds shows the best predictive power, $R^2=0.976$. Though the PCR-EEM method is not as specific as GCMS, EEM is (i) more sensitive (>10x, as it analyzes thousands of OC species instead of a single target)[71], (ii) significantly less-expensive (<$10/sample), (iii) time-efficient (as it take only a few minutes for analysis after solvent extraction of target compounds from PM), and (iv) can be miniaturized into a small, low-cost sensor[171].

Chapter 4 Electrostatic Capillary Collector for *In-situ* Spectroscopic Analysis of Aerosols

4.1 Summary

This chapter presents the design and evaluation of a capillary sensor that collects PM2.5 onto the outer surface of a capillary for *in-situ* spectroscopic analysis. The capillary collector uses a needle-to-ring corona particle charger; the electric field between the cylinder and wire inserted into the bore of the capillary is used to collect the charged particles. Corona and repelling voltages are optimized for maximum capillary collection of PM2.5 particles for ambient aerosol and fluorescent polystyrene microspheres. Capillary collection efficiency for PM2.5 in ambient air at 3 slpm operating flow rate and optimal operating voltages is 63%. Fluorescence spectroscopy is used to quantify particle collection using polystyrene latex microspheres with high fluorescent yield. The collection and analysis methodology can be used to develop a compact, low-cost sensor for *in-situ* spectroscopic analysis of aerosols to determine their chemical compositions for source apportionment. These results are published.[172]

4.2 Experimental Methods

4.2.1 Capillary Collector Design and Operation

The capillary collector is divided into two regions: the ionization region and the collection region, see Figure 4.1 (top). The individual elements of the collector are held in the 3D-printed enclosure to control the spacing and alignment. The ionization region consists of a high voltage corona needle and a grounded ring electrode. The anode needle is a 0.5 mm thick tungsten needle with a radius of curvature of 1 μm at the tip (measured using optical microscopy). The ground electrode is a

1.58 mm thick solid solder with an inner radius of 10 mm. The needle is held at the axis of the ring at 5 mm from the center of the ring. A variable high voltage power supply (Glassman High Voltage Inc. EH Series, High Bridge, NJ, USA) is used to create the electric potential between the electrodes, referred to as corona voltage (φ_C) in this paper. In the positive corona mode, discharge occurs at a corona voltage > 2.5 kV, indicated by a small positive anode current of around 0.1µA.

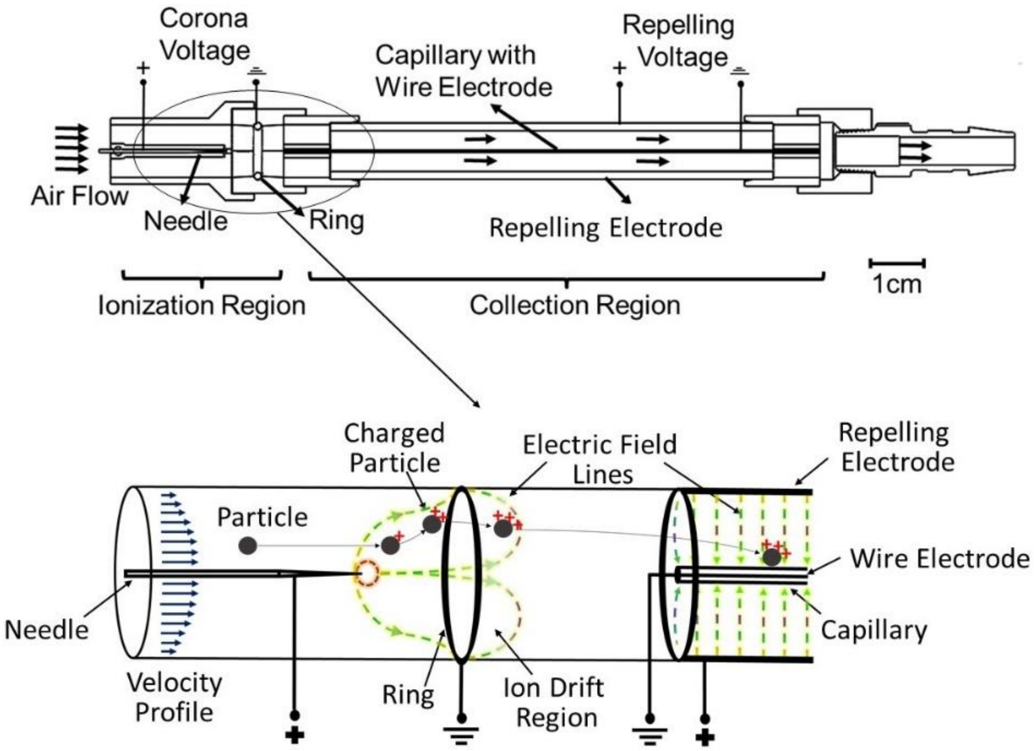


Figure 4.1: Capillary collector design: The needle ring assembly is called the ionization region and wire tube assembly is called the collection region (top). Particles entering the device are charged due to the corona discharge in the ionization region and are repelled by the electric field in the collection region leading to the collection of particles on the capillary surface (bottom).

Particles are charged as they enter the capillary collector. The ionization region is fluidically connected to the collection region, consisting of a high voltage aluminum tube and a grounded thin metal wire placed inside the fused silica capillary. The positively charged aluminum tube (100 mm long, 12.7 mm OD) acts as a repelling electrode. The ground electrode wire (100 µm in diameter) is inserted into the bore of a fused silica capillary (250 µm ID, 350 µm OD, Length: 100 mm). The capillary with the grounded wire is placed concentric to the aluminum tube.

Charged particles enter the electrical field formed between the aluminum tube (repelling electrode) and the capillary (with the ground electrode) and are collected on the outer capillary wall. Hence, this region is called the collection region. A variable high voltage power supply (Glassman High Voltage Inc. EH Series, High Bridge, NJ, USA) is used to create the electric potential between the wire and tube, referred to as repelling voltage (φ_R) in this paper. The collection region is connected to a pump through a barb connection. From the Lagrangian perspective, the particles passing through the corona discharge become positively charged, as shown in Figure 4.1 (bottom). Their trajectories in the ionization region depend on their mass/charge ratio, the strength of the electrical field, their residence time, their initial position and properties of the fluid (e.g. drag force on particles). The time-dependent particle charging in the strong, rapidly changing electric field and the high ion concentration is a very complex process and is a subject of future research. As the ring in the ionization region is grounded, a fraction of the charged particles can be collected on the ring and the rest of the particles enter the collection region. The electric field between the wire and tube in the collection region repels the positively charged particles towards the ground wire. Since the wire is concentric to the capillary, particles which change their trajectories significantly due to the force exerted by the electric field in the ionization region collect on the outer surface of the capillary wall. During this time, a fraction of the PM can exit the device. The optimization of the collector operating conditions is performed.

4.2.2 Ambient Particle Collection Experiments

The experimental study to assess the collection efficiency was performed for several corona voltages (φ_C) and repelling voltages (φ_R) in a temperature range of 22-25 ^oC, a relative humidity of 23%-25% and at ambient pressure. The experimental setup was placed on a benchtop in a laboratory setting having a PM2.5 mass concentration of 8µg/m^3 measured by an aerodynamic particle sizer (APS 3310, TSI, Shoreview, MN). In the ionization region, corona discharge was stable for a range of corona voltages, from 2.5 kV - 7 kV. The onset of ion generation was at $\varphi_C >$ 2.5 kV. The current measured on the cathode ring, based on the voltage drop across a 1 MΩ resistor, increased with increasing corona voltage, having values of 0.8 µA for φ_C = 4 kV and 7.5 µA for φ_C = 7 kV. For φ_C > 7 kV, the cathode current became unstable and not reproducible over a range of atmospheric and operating conditions (RH, T, flow rate), eventually leading to a spark discharge. Hence, φ_C was varied from 3 kV to 7 kV for collection efficiency experiments and φ_R was varied from 0 kV to 2 kV. Increasing φ_R above 2 kV resulted in arcing between the collection and repelling electrodes which damaged the capillary substrate. The capillary collector requires lower pump power to operate compared to commercially available personal aerosol samplers because of the ionic wind generation, which adds to the flow in the device, and because of the absence of flow restricting collection filters. In this paper, we present the collection results for a flow rate of 3 slpm to operate the device; however, the device can be optimized for higher flow rates.

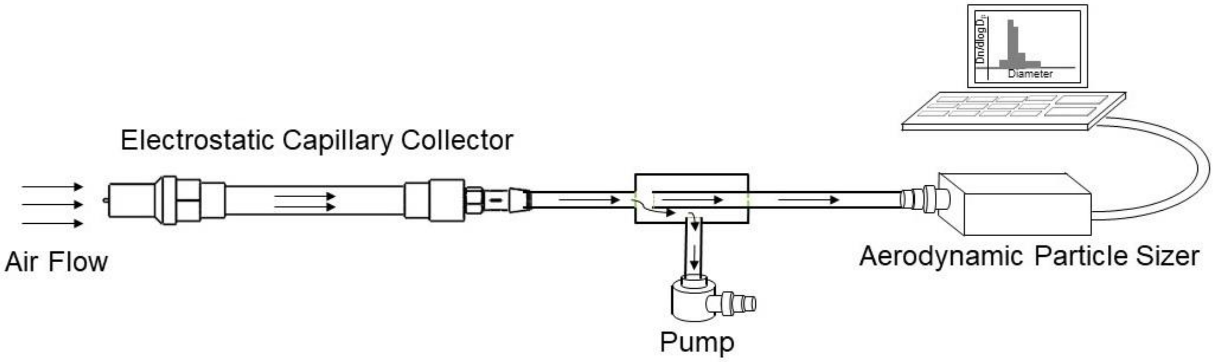


Figure 4.2: Schematic diagram of the experimental setup used to optimize φ_C and φ_R for maximum capillary collection efficiency. The particle concentration leaving the device was measured by the APS, and the CE of the device was calculated.

In ambient particle collection experiments, efficiency was determined based on the total number of ambient air particles collected in the capillary collector with respect to the number of particles entering the device, referred to as Total Collection (TC). The outlet of the device was attached to a T-connection, as shown in Figure 4.2. One end of the T-connection was connected to a pump, and the other end was connected to the APS. The sampling flow rate of the APS was set to 1 slpm. The flow rate of the pump was set to 2 slpm via a digital rotameter. The dimension of the sampling tube and the inner dimension of the T-connector were chosen, such that the velocities were approximately the same for a total flow rate of 3 slpm. For each sampling experiment, φ_C and φ_R were fixed. The sampling time was 60 seconds. Five reference measurements of particle concentrations were recorded by the APS with $\varphi_{C,}$ and φ_R turned off. Five sample measurements were recorded with $\varphi_{c,}$ and φ_R turned on. The average of these five measurements was used to calculate the TC, as given by (4.1). The measurements were repeated for each sampling condition,

$$TC_{N.M} = (1 - \frac{APS_{N,M}}{APS_{0,0}}) \times 100 \tag{4.1}$$

where $TC_{N,M}$ (%) is the TC at fixed φ_C = N (kV) and φ_R = M (kV), $APS_{N,M}$ (μg/m^3) is the concentration of particles measured by the APS at φ_C = N (kV) and φ_R = M (kV), and $APS_{0,0}$ is the concentration recorded for reference measurement. When $\varphi_C > 0$ kV and $\varphi_R > 0$ kV, TC measurement includes particles collected on the capillary and particles collected elsewhere in the device. The fraction of particles entering the device which are collected on the capillary substrate can be quantified using Capillary Collection (CC). While TC is the measurement of all particles collected in the device, CC is a fraction of TC for particles which are collected on the capillary. It is necessary since only the particles on the substrate will be used for subsequent spectroscopic analysis; the particles collected on the grounded ring electrode or other parts of the geometry cannot be analyzed. When $\varphi_C > 0$ kV and $\varphi_R = 0$ kV, TC represents the fraction of particles lost on the ring and other parts of the ionization region. The difference in TC between $\varphi_R > 0$ kV and $\varphi_R = 0$ kV for the same φ_C can be used to calculate CC given by (4.2),

$$CC_{N,M} = TC_{N,M} - TC_{N,0} \tag{4.2}$$

where $CC_{N,M}$ (%) is the CC at φ_C = N (kV) and φ_R = M (kV). The N and M values for which $CC_{N,M}$ is the highest, are set as the optimum operating conditions for a given flow rate. The ozone concentration was measured using an ozone analyzer (Model 450, Teledyne Instruments, San Diego, CA, USA) downstream of the capillary collector at optimum operating conditions and sampling flow rate of 3 slpm.

4.2.3 Capillary Collection Efficiency Using Bulk Fluorescence

While the collection of polydisperse ambient PM provides the most relevant data for exposure monitoring applications and allows for the collection of large amounts of data over a range of particle sizes, it may have uncertainties with calculations of particle loss in the ionization zone and losses associated with particle transmission in the tubing connected to the particle sizer. To validate the ambient PM results, additional experiments with monodisperse model particles were performed. The capillary collector was placed in an aerosol chamber, which is a sealed container having an inlet port to press fit the nebulizer and output ports for connecting devices, filters, and the APS. Two variable high voltage power supplies (Glassman High Voltage Inc. EH Series, High Bridge, NJ, USA), which provide φ_C and φ_R, were placed outside the aerosol chamber and were electrically connected to the collector. The operating voltages were set at the optimum values obtained from the experiments with ambient particles. A reference filter (Spectrum Poretics® polycarbonate membrane filters, 47 mm, 0.1 µm or 0.6 µm pore sizes) was placed in an open face aerosol filter holder (EDM Millipore, Billerica, MA, model XX5004710). The collector and the filter were operated at a sampling flow rate of 3 slpm. The experiments used polystyrene latex (PSL) particles (Fluoresbrite® YG Microspheres, Polysciences, Inc., Warrington, PA, USA) having sizes 0.2 µm, 0.5 µm, 1 µm, and 2 µm. A stock particle solution from the manufacturer (typically 2 % - 3 % solid fraction) was diluted 100X in distilled water. Figure 4.3 (a) shows the schematic of the experimental setup.

During the experiment, about 5 ml of prepared solution was nebulized from the liquid suspension using the MADA Up-Mist™ Medication nebulizer (MADA Products, Carlstadt, NJ). The aerodynamic particle sizer (APS 3321, TSI, Shoreview, MN) was used to verify the particle size and monitor the particle concentration in the chamber. Typical particle size distribution in the aerosol chamber for each particle size is shown in Figure 4.3 (b). APS measurements confirmed that there was no significant aerosol agglomeration during the experiments. Particles were collected for 20 minutes in each experiment. Three collection experiments were performed for each particle size.

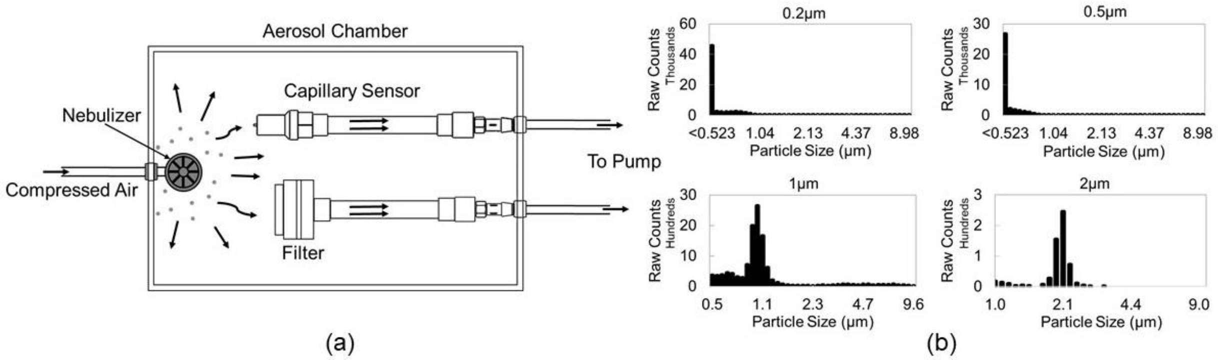


Figure 4.3: (a) Fluorescent PSL particles in the PM2.5 size range were sampled using the capillary collector and a reference filter in an aerosol chamber. The capillary collection was determined by measuring the bulk fluorescence signal of particles collected on the capillary and filter. (b) Typical size distribution of PSL particles during the collection, recorded by the APS.

After collection, the capillary was removed from the collector for analysis. The collection efficiency was calculated by comparing the bulk fluorescence emission intensity of particles captured on the capillary surface to the bulk fluorescence emission intensity of particles captured on the reference filters. Particles on the capillary substrate and the reference filters were first eluted with 10 mL of 0.05 % Tween 20 solution using 50 mL centrifuge tubes. The centrifuge tubes containing the capillary and the reference filters were placed in an ultrasonic bath for 5 minutes which was followed by 5 minutes of shaking in a vortexer. Additional elution steps or increased times did not produce any significant increase in the fluorescent signal. A 2 mL aliquot of total 10 mL volume was then transferred into a 12 mm × 75 mm borosilicate glass cuvette, and the solution fluorescence was measured with a spectrofluorometer (Sequoia-Turner model 450). The calibration of the spectrofluorometer and the flow meters was confirmed before and after each experiment. Fluorescence-based capillary collection (FCC) is defined as the efficiency of particle collection on the capillary surface based on the fluorescence intensity reading from the capillary solution and reference filter solution:

$$FCC = \frac{I(Capillary)}{I(Filter)} \times 100 \tag{4.3}$$

where I is the bulk fluorescence intensity from the filter or the capillary eluent.

4.2.4 Fluorescence Spectroscopy Using Capillary Substrate

We demonstrate the integration of capillary substrate used in the capillary collector for *in-situ* analysis through fluorescence spectroscopy of polystyrene latex (PSL) particles. PSL particles having sizes 0.1 µm, 0.2 µm, 0.5 µm, 1 µm, and 2 µm were used for fluorescence analysis. Particle collection experiments were performed at 3 slpm sampling flow rate and optimum operating voltages. Three collectors were used in the aerosol chamber for collection. Five collection experiments were performed for each particle size. After the collection, the capillary was removed and inserted into the optical analysis module. We tested two different excitation configurations: 1. internal excitation, where the light was passed from the capillary end, and 2. standard external excitation used in epifluorescence microscopy in which the light source was fixed perpendicular to the capillary. The lens used for the collection of emitted light was always placed perpendicular to the capillary. Figure 4.4 shows the schematic of the experimental setup. Fluorescence emission with respect to excitation light captured by the detector was investigated for both internal and external excitation methods.

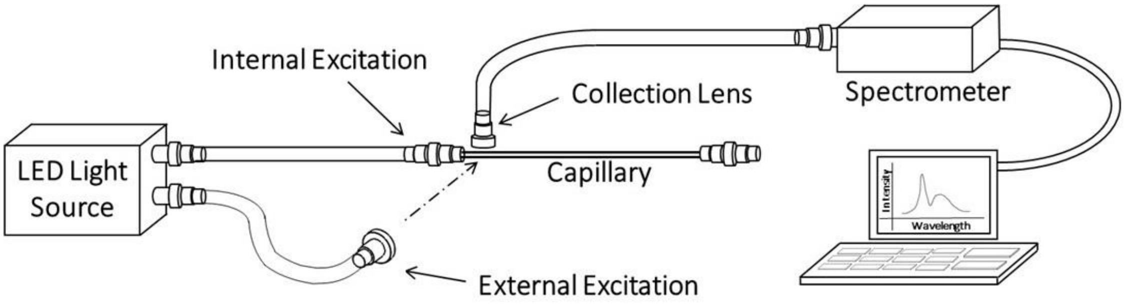


Figure 4.4: Fluorescence spectra of PSL particles collected on capillary were recorded using internal excitation and external excitation as shown in the schematic.

4.3 Results and Discussion

4.3.1 Optimization of Operating Voltages

Operating voltages were optimized in the ambient particle collection experiments at a collector flow rate of 3 slpm. The ambient particle concentration in each bin did not change significantly during the experiments. An average of 5 measurements were used for the calculation of TC. Figure 4.6 (a) shows the TC for the PM2.5 range. When $\varphi_C > 0$ kV, particles are charged in the ionization zone. The particle charging mechanism and their exact trajectories in the ionization region, which has a strong electrical field gradient and a high ion concentration, are beyond the scope of this manuscript. However, three possibilities with respect to particle fate are apparent and are also observed experimentally: 1. the particle is charged and captured in the ionization zone before entering the collection region, 2. the particle travels through the ionization zone and is captured on the collection substrate, and 3. the particle is not captured anywhere in the device, exits the collection zone, and is counted by the APS. In the first scenario, the charged particle deviates in the electrical field of the ionization region and eventually lands on the wall or the ground electrode. The fraction of particles exhibiting this behavior can be measured by operating the device with the corona voltage “on” ($\varphi_C > 0$ kV) and the repelling voltage “off” ($\varphi_R = 0$ kV).

For zero repelling voltage, as corona voltage is increased, the number of particles charged and the amount of charge on the particles increases. Hence, changing φ_C from 3 kV to 7 kV increases particle capture from 10 % to 35 %. When the repelling voltage is turned on, TC increases significantly due to particle capture in the collection region. An increase in the repelling voltage results in a stronger electrical field and a greater force acting on the charged particles. Hence, TC increases with increasing repelling voltage. For φ_C > 4 kV and φ_R = 2 kV, more than 90 % of all particles entering the device are collected. CC is found to be maximum for φ_C ~ 4 – 5 kV; additional increase in φ_C results in an efficiency drop due to particle loss in the ionization zone. At lower values of φ_C, a significant number of particles escape the collector (35 % at φ_C = 3 kV) even when φ_R is set to the highest value, suggesting that these particles are not charged in the ionization zone. A maximum CC of 63 % is observed for φ_C = 4 kV and φ_R = 2 kV; this condition is chosen as optimal for the collector operation. φ_R is restricted to 2 kV to prevent arcing and damage to the capillary. It was verified that 3 slpm is the optimum flow rate for maximum CC as shown in Figure 4.5. Ozone concentration measured at φ_C = 4 kV and φ_R = 2 kV and a sampling flow rate of 3 slpm is 28 ppb.

Maximum CC is obtained at operating optimum conditions of φ_C = 4 kV, φ_R = 2 kV, and flow rate = 3 slpm. At optimum operating voltages, if the flow rate is increased or decreased by 1 slpm, the CC reduces from 63 % to a little over 54 %. Hence, a lower flow rate can be used with little reduction in CC. If φ_C = 3 kV, φ_R = 2 kV, and flow rate = 2 slpm, the CC is 57.45 %. Hence, power can be reduced using lower φ_C and flow rate with only 5.55 % less CC.

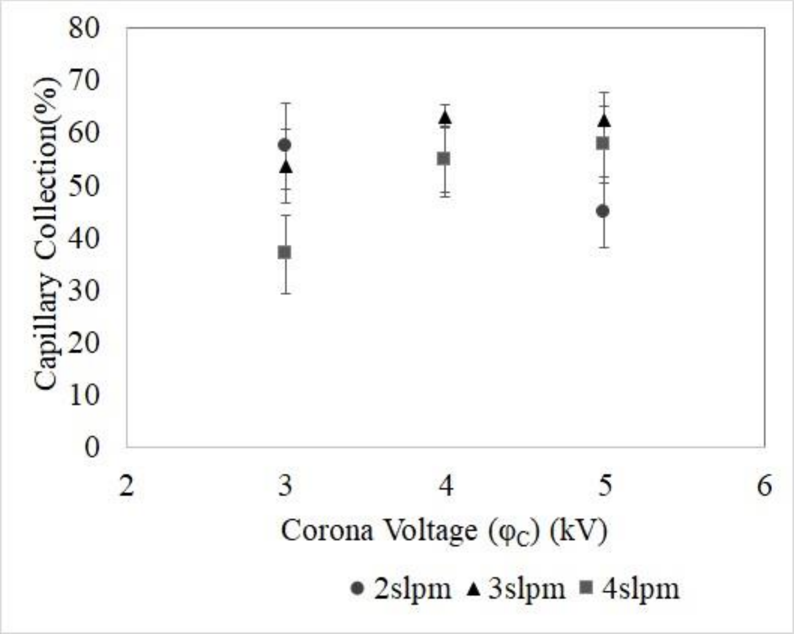
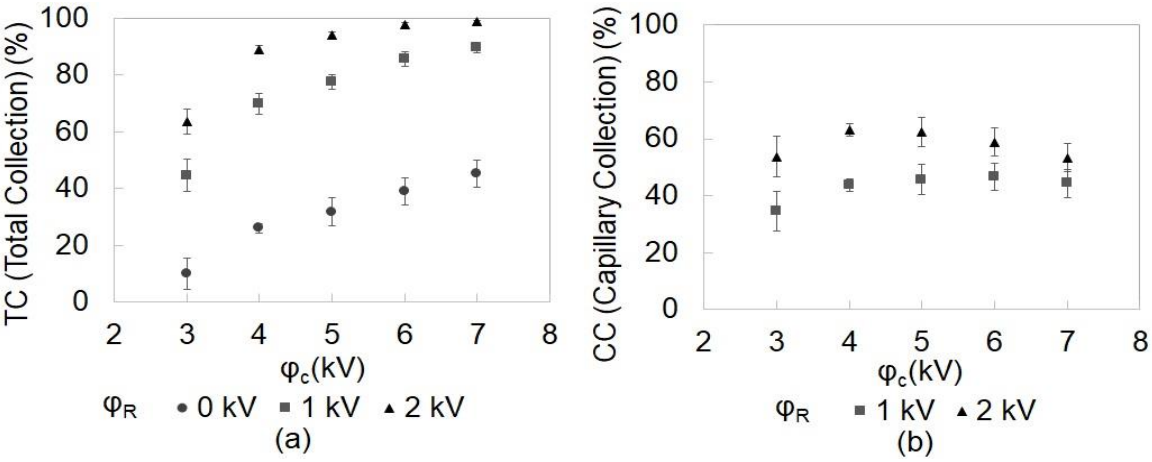


Figure 4.5: Capillary Collection (CC) of particles in the ambient air are obtained at repelling voltage (φ_R) = 2 kV, and varying corona voltage (φ_C) and flow rate.



(a)

(b)

Figure 4.6: (a) The TC of the capillary sensor for PM2.5 increases with increasing corona voltage (φ_C). (b) The CC of the capillary sensor for PM2.5 is maximum at corona voltage (φ_C) = 4 kV and repelling voltage (φ_R) = 2 kV.

4.3.2 Collection Efficiency Using Polystyrene Latex (PSL) Particles

Experiments with monodisperse fluorescent PSL particles were performed to validate the ambient particle collection efficiency measurements. Here, the fluorescence from the capillary collection is directly compared with the reference filters operated in parallel in the aerosol chamber. Figure 4.7 compares capillary collection as a function of particle size for ambient air particles (APS sizing) and PSL particles. In ambient air experiments, APS measurements show that particle concentration is highest in the 0.5 µm to 1 µm size range and decreases as particle size increases. Capillary collection obtained from bulk fluorescence measurements is the highest for 0.5 µm and 1 µm particle sizes and decreases for 2 µm particles, which compares well with the APS data. The average FCC for PM2.5 particles at optimum voltage conditions is 61 % (based on equal particle binning), which is close to the 63 % CC, which validates the measurements taken with the ambient particles.

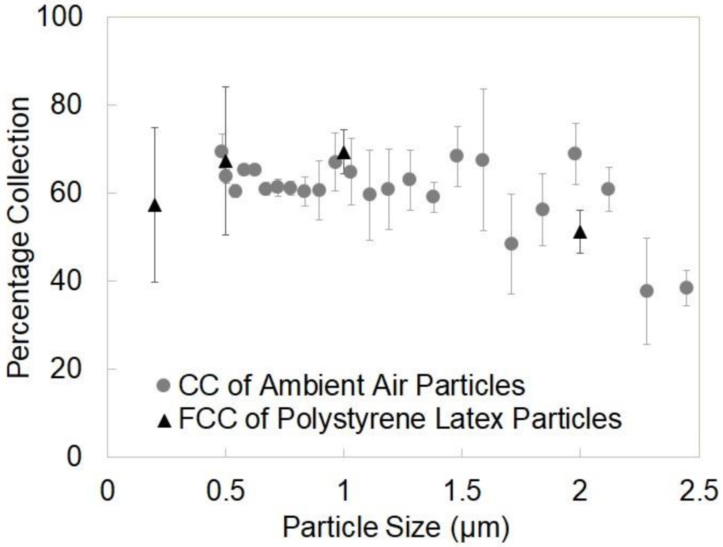


Figure 4.7: Comparison of PM2.5 collection on the capillary obtained by two methods: 1. bulk fluorescent measurements of PSL particles (FCC), and 2. particle concentration measurement of ambient air (CC)

4.3.3 Fluorescence Emission of PSL Particles

Emission spectra recorded by internal and external excitation consist of fluorescence peaks from PSL particles and an excitation peak due to the detection of scattered excitation light, as shown in Figure 4.8. For both internal and external excitation, the maximum intensity of the excitation peak is greater than the maximum intensity of the fluorescence peaks in the emission spectra. If fluorescence peaks have very low intensity compared to the excitation peak, it will be difficult to discern fluorescence peaks in emission spectra since they will be covered by the broad excitation peak. Therefore, for the same intensity of the excitation peak, spectra which have higher fluorescence intensity will be easy to identify and interpret. The emission spectra of collected samples are normalized by their maximum value. The normalized spectrum of the LED light source is subtracted from the emission spectra to get the fluorescence spectra for each method. The maximum fluorescence intensity in the resultant spectra for internal excitation are compared to those for external excitation in Figure 4.9.

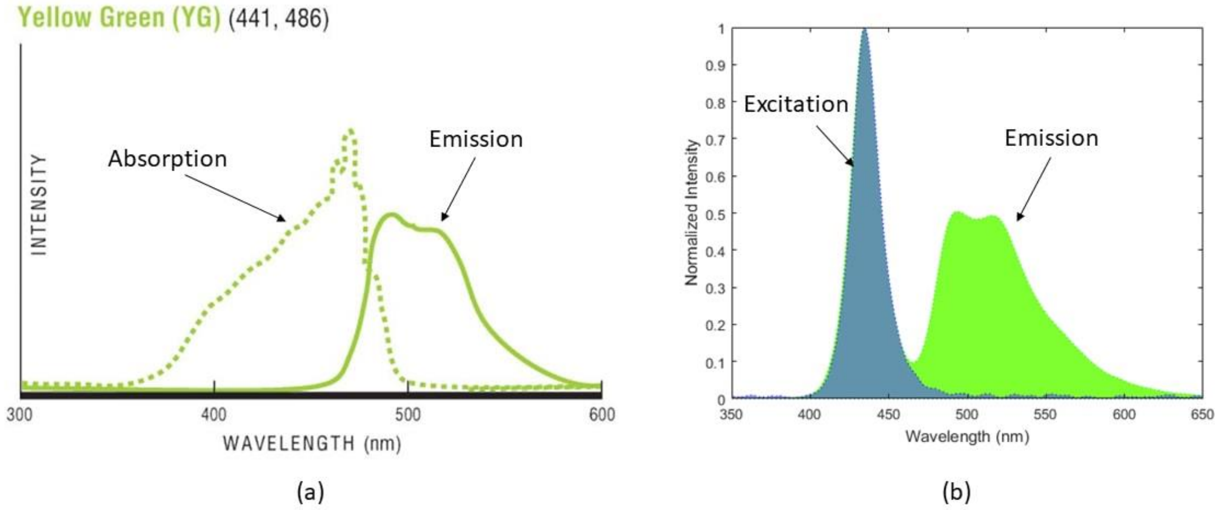


Figure 4.8: Manufacturer's data on the emission spectrum of PSL particles (a) compares well with the emission spectrum of the same particles (0.1 µm) collected on the capillary (b).[173]

For all particle sizes, the peak fluorescence intensity for the internal excitation method is always greater than that for the external excitation. This means that fluorescence peaks can be distinguished easily from emission spectra obtained using internal excitation than the ones obtained using external excitation. Internal excitation technique with a capillary substrate will work better in compact, low-cost devices to analyze fluorescence of collected particles when broad excitation LEDs are used. The result does not imply that the internal excitation technique has higher sensitivity than the external excitation technique. Additional work is needed for sensitivity analysis.

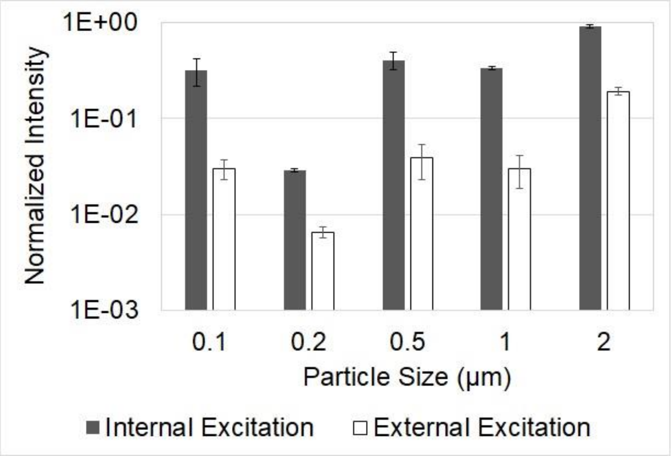


Figure 4.9: The normalized maximum fluorescence intensity for the internal excitation method is greater than the external excitation method; plotted with standard error (n = 3).

4.4 Chapter 4 Conclusions

We designed and evaluated a capillary collector for the collection of PM2.5 which can be integrated with *in-situ* spectroscopic analysis techniques. The device combines particle charging and electrostatic collection to collect PM2.5. Two methods are used to determine the capillary collection: real-time particle concentration measurements of ambient air using particle sizers and bulk fluorescent measurements based on the collection of fluorescent PSL particles. The average capillary collection of PM2.5 from APS measurements of ambient air particles is 63 % and is 61% from bulk fluorescence measurements of PSL particles. At optimized operating conditions, the collection efficiency is relatively constant (~ 65 %) for all particles in the 0.2 µm - 2 µm range and drops below 50 % for particles larger than 2 µm, likely due to a lower charge-to-mass ratio. Total collection measurements for ambient PM2.5 also show that most of the lost particles are in the particle charging region due to their collection on the walls and on the ground electrode in the ionization region. Fluorescence spectroscopy is used to demonstrate the integration of the capillary collector with *in-situ* spectroscopic analysis techniques.

Chapter 5 Solid Phase Excitation-Emission Matrix Spectroscopy for Combustion Generated Particulate Matter Analysis

5.1 Summary

Chapter 3 describes the EEM analysis as a cost-effective and time-efficient technique for chemical composition evaluation and source apportionment of combustion generated PM compared to conventional lab-based techniques. However, real-time monitoring of PM chemical composition using EEM analysis requires development of a PM collection and analysis substrates which can extract PM organic compounds and aid EEM measurements *in-situ*. In this chapter, we report the development of a SP-EEM technique for chemical analysis of combustion generated PM. The UPM samples are collected onto transparent substrate electrostatically and analysed *in-situ*. The SP-EEM spectra for wood smoke and cigarette smoke is compared with more conventional LP-EEM analysis. A preliminary analysis on internal and external excitation of sample shows feasibility for development of a low-cost sensor for chemical composition analysis of combustion generated PM. These results are submitted for publication.[174]

5.2 Experimental Methods

5.2.1 Collection Substrate

PDMS solution was prepared from the platinum-based Dow Corning Sylgard 184 silicone elastomer kit by mixing the base with the curing agent in a 10:1 mass ratio. We coated glass and quartz slides (75mm X 25mm X 1mm) with the PDMS using a spin coater (Laurell WS-650, Laurell Technologies, North Wales, PA, USA) with 2000 rpm for 1 minute. We cured the coating in a heating oven at 70^{o}C for 4 hours. The cured PDMS layer thickness was measured at ~150 µm.

5.2.2 Particulate Matter Collection

A miniature single-stage ESP collected PM directly onto the PDMS layer, see Figure 5.1. The device did not employ a charging stage; instead, we relied on PM native charge from the combustion process. For aerosols aged in the atmosphere the Boltzmann distribution can be expected. Woodsmoke was generated by burning 1 ½ by ¾ inch Douglas fir sticks in a side-feed natural-draft cookstove.[144] Cigarette smoke was generated by lighting cigarettes in a sealed chamber and allowed the cigarettes to smoulder; flow was aspirated by the pump and the smoke was circulated back into the sealed chamber.[71,175] Typical size distribution peak for woodsmoke is ~30 nm (Figure 5.2) and for cigarette smoke is ~60 nm (Figure 5.3).

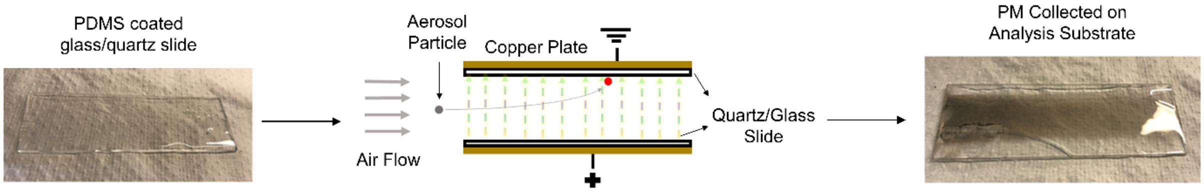


Figure 5.1: Quartz slide coated with PDMS (left) was placed in a parallel plate electrostatic collector. Combustion generated PM was collected directly onto the PDMS coated surface of the substrate. The substrate was used for SP-EEM analysis after collection.

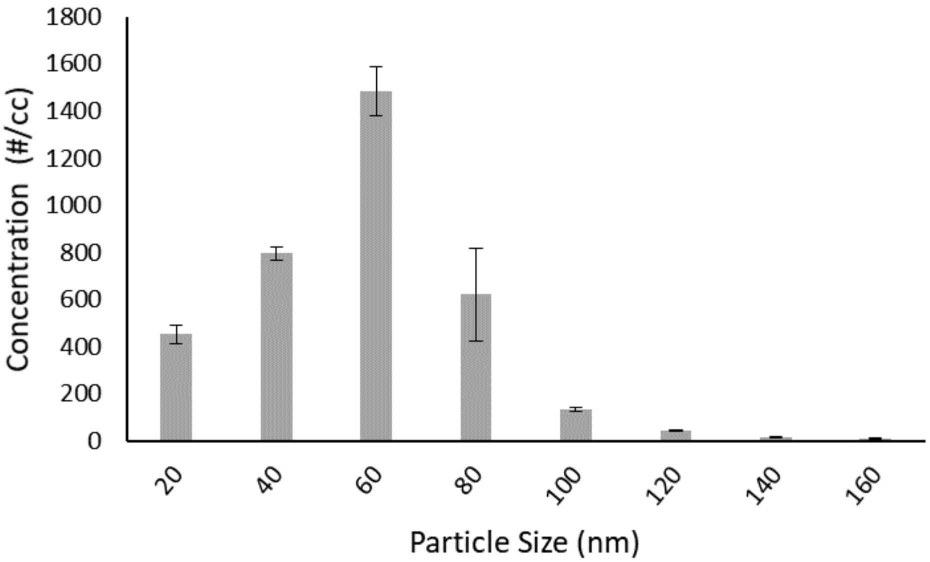


Figure 5.2: Size distribution of wood smoke particulate matter sampled at UW Clean Cookstoves Lab measured with TSI SMPS NanoScan (10-200 nm).

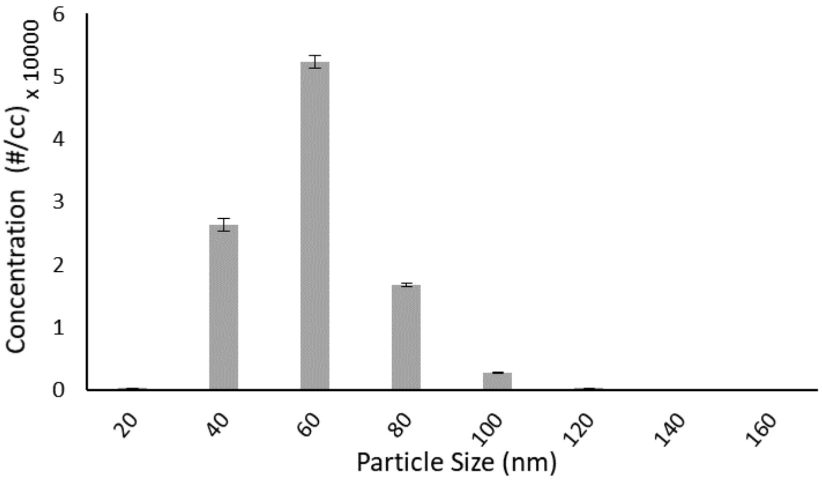


Figure 5.3: Size distribution of cigarette smoke particulate matter measured with TSI SMPS NanoScan (10-160 nm).

Two collection substrates were placed in a Teflon holder, 5 mm apart, forming a rectangular flow channel, with PDMS layers facing the flow. Particle-laden flow was aspirated between the collection substrates at the flow rates of 1.8 slpm. Two copper electrodes positioned on either side were connected to a benchtop HV power supply (Glassman High Voltage Inc. EH Series, High Bridge, NJ, USA). An electric potential φ = 3 kV was applied in the experiment resulting in electric field strength between the electrodes E=0.42 kV/mm. Charged particles were collected on the PDMS coated surfaces; the tests were run until visible particle deposition was observed (90 min for woodsmoke and 10min for cigarette smoke). We did not characterize aerosol collection efficiency in these experiments; in our previous work, a single-stage collector with similar field strength, gap size, and residence time yielded ~30-40% collection efficiencies for ambient ultrafine PM (Figure 5.4) and ~55% collection efficiency for woodsmoke PM (Figure 5.5). The substrates were removed from the collector and isolated for at least 24 hours to allow adequate time for the extraction of organic compounds into PDMS.

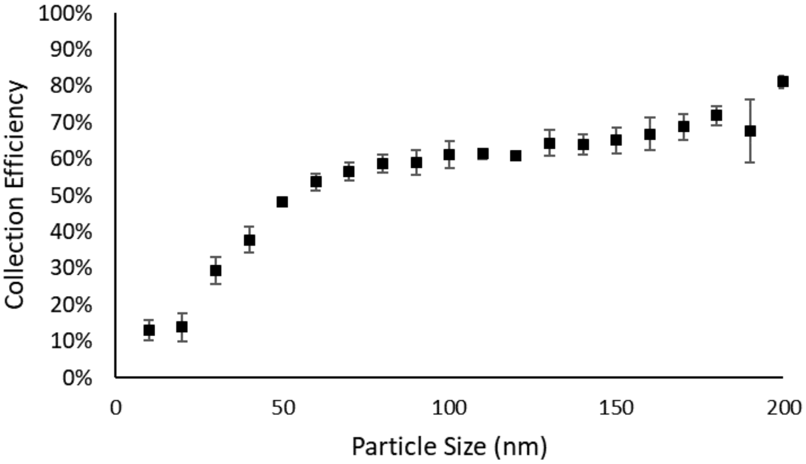


Figure 5.4: Collection efficiency of parallel plate collector for ambient PM as a function of particle size at 0.75 slpm. Each data point corresponds to three or more measurements of particle count by a TSI SMPS NanoScan particle counter (10-200 nm) with error bars representing one standard deviation.

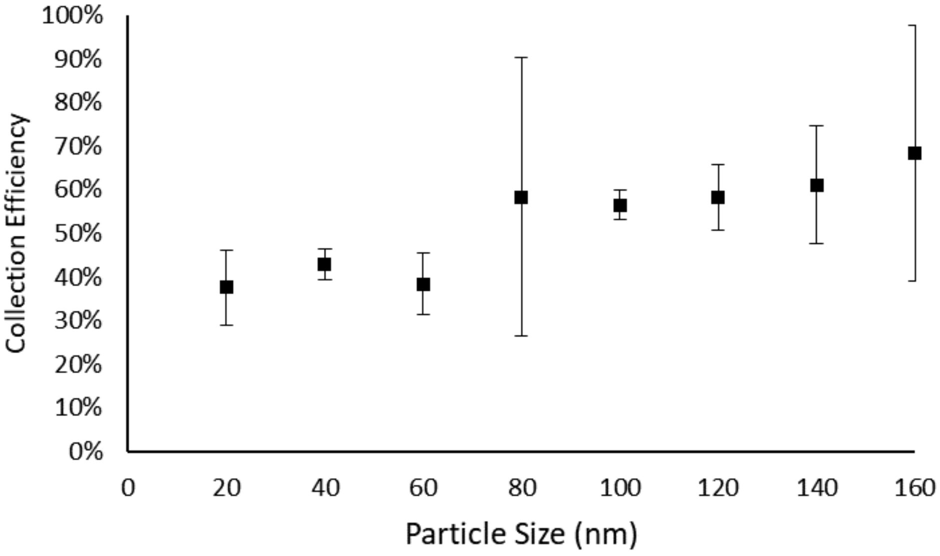


Figure 5.5: Parallel plate collection efficiency for woodsmoke at 0.75 slpm, 5 kV, dimensions - 75 x 25 x 5 mm

The reference sample was collected for LP-EEM analysis.[71,114] PM samples were collected on 2.0 µm pore PTFE membrane filters (Pall Zefluor®, Pall Cat. # P5PJ037, Pall Corporation, Port Washington, New York, USA) housed in the Harvard impactor sampling cassettes (Cat. # HP2518, BGI, Butler, NJ, USA). After gravimetric analysis, the samples were extracted in cyclohexane due to its non-polar characteristics and common use for PM PAH extraction[150] (Cat. #1.02822.2500, Uvasol® Cyclohexane for Spectroscopy, MilliporeSigma, Burlington, Massachusetts, USA). The sample extracts were filtered with 0.2 µm PTFE syringe filters (Cat. #28145-491, VWR, Edison, NJ, USA), diluted to achieve ~3.5-5 µg/mL concentration, and processed EEMs of filtered extracts were taken.

5.2.3 Solid Phase Excitation Emission Matrix (SP-EEM) Analysis

The EEM spectra of PM collected on the analysis substrates were obtained using a spectrofluorometer (Aqualog-880-C, HORIBA Instruments Inc. Edison, NJ, USA). The spectra were recorded in the range of excitation wavelength λ_{ex} = 200-600 nm with a 2 nm resolution. For each excitation wavelength, the instrument records emissions using CCD array in the range of λ_{em} = 246-826 nm with a 0.58 nm resolution. The angle between the incident excitation light and the substrate (θ) was varied from 5° to 85° to optimize for maximum fluorescence signal and minimum Rayleigh and Raman scattering interference, see Figure 5.6 B and Figure 5.6 C. Elastic scattering and reflections are significantly higher in SP-EEM measurements, compared to LP-EEM, due to presence of PM on the surface and PDMS surface non-uniformity (in LP-EEM, PM is filtered after solvent extraction). SP-EEM measurements for woodsmoke and cigarette smoke PM samples were obtained at 5°<θ<85°, this is referred to as external excitation. Glass substrates data is presented in the manuscript, the results for quartz at optimum θ are similar, see Figure 5.7.

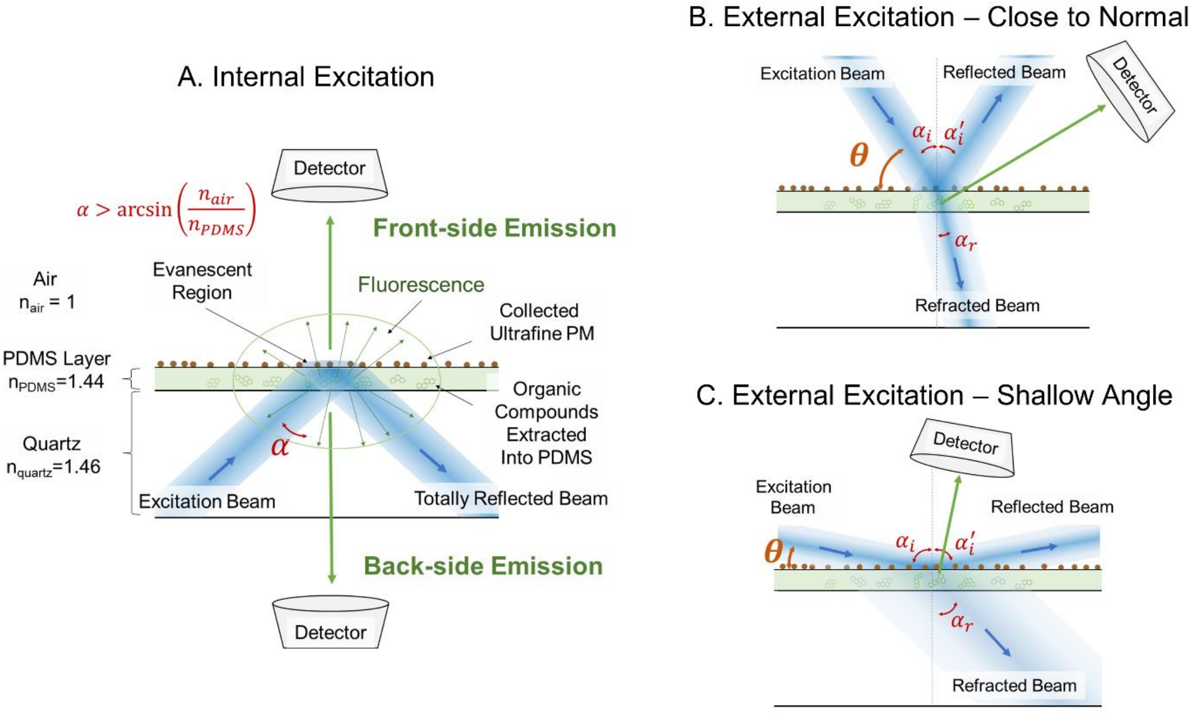


Figure 5.6: (A) In internal excitation, light passes through the analysis substrate and the fluorescence signal is recorded from the PDMS side of collection substrate (front-side emission, θ = 0°) or the uncoated side of the collection substrate (back-side emission, θ = 180°). (B) SP-EEM analysis using external excitation for excitation angle (θ) closer to normal and (C) SP-EEM analysis using external excitation for shallow excitation angle between the collection surface and excitation light, the angle between the incident light and collection optics was held constant at 90°.

Internal excitation was explored at $\theta = 0°$ and $\theta = 180°$, where the quartz substrate acts as a waveguide for excitation light. Since transmission of glass is negligible for $\lambda<300$nm (Figure 5.8), quartz substrates were used in internal excitation experiments. Excitation light couples to PDMS layer and deposited PM as refracted light or by an evanescent wave mechanism[176,177]. The excitation light interacts with organic compounds extracted into the PDMS layer. The excited compounds emit fluorescence which is collected at an 90° angle to the excitation beam, see Figure 5.6 A. It has been shown previously that in internal excitation, interference of scattering signal with the fluorescence is lower compared to external excitation.[172] We refer to $\theta = 0°$ configuration as 'front-side emission' as fluorescence emission is collected from the PM deposition side of the substrate while $\theta = 180°$ configuration is referred to as 'back-side emission' in which fluorescence emission is collected from the side opposite to the PM deposition. SP-EEM of a blank substrate was subtracted from each PM sample. Rayleigh and Raman scattering peaks were removed computationally.[164] The processed EEMs were passed through a Gaussian filter (sigma=2) and the negative values were removed numerically (MATLAB, MathWorks Inc., Natick Massachusetts, USA) to smooth the data. The resulting EEMs were normalized to Raman units (R.U.).[163]

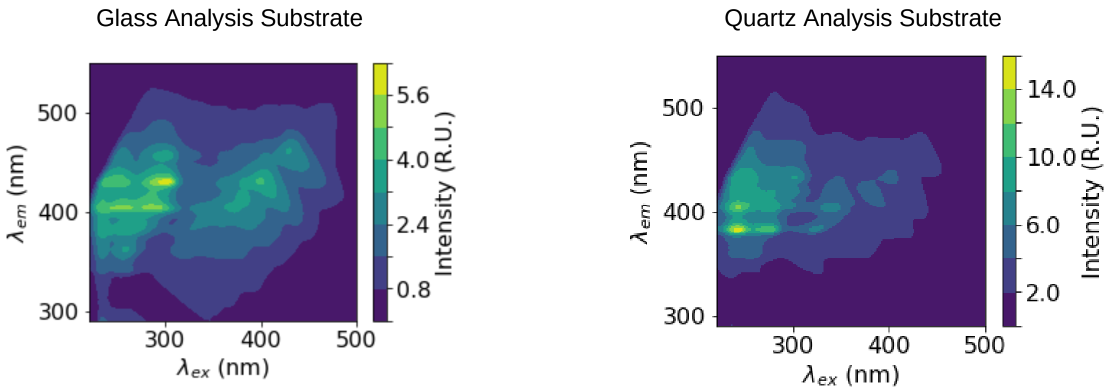


Figure 5.7: External excitation SP-EEMs at θ = 60° match well for glass and quartz analysis substrates. The number and relative intensity of peaks defining woodsmoke PM fluorescence signature is similar for both SP-EEMs.

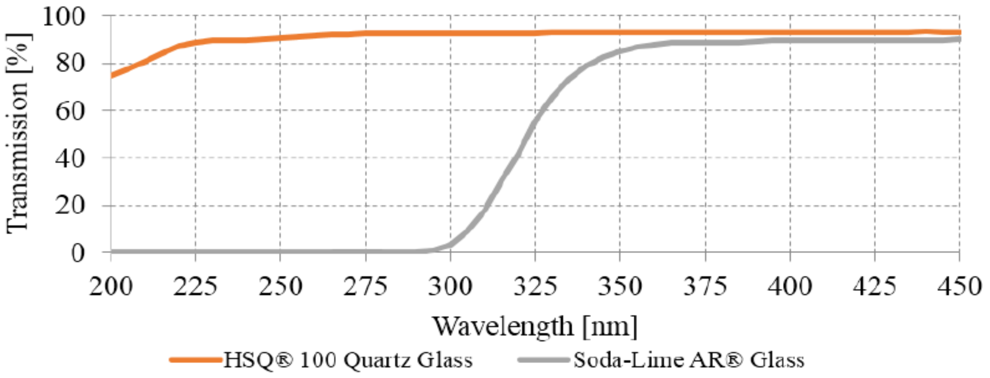


Figure 5.8: Transmission of HSQ® 100 quartz glass (92% at 280 nm) and soda-lime glass (2% at 280 nm) with a thickness of 1 mm between 200 and 450 nm measured with a Specord 250 Plus spectrophotometer from Analytik Jena (Jena, Germany); reference: air.[178]

5.3 Results and Discussion

5.3.1 External Excitation SP-EEM Measurements

External excitation of woodsmoke samples were used to optimize the excitation angle. The fluorescent peak increases up to about θ = 50° and then decreases, see in Figure 5.9. The combined reflection and scattering signals are also at their maximum at θ = 40-55° masking the fluorescent signal. Low excitation angles θ = 5-20° do not result in very strong fluorescent signal likely due to broadening and diffusion of gaussian excitation beam in the substrate resulting in lower excitation intensity as shown in Figure 5.6 C. At angles θ >70° the fluorescent signal drops significantly likely due to decrease in the excitation path length and reduction of the view factor of the collection optics as shown in Figure 5.6 B. Additional ray optics modelling is required to gain better understanding of the system; however, it is beyond the scope of this communication. To balance the intensity of PM fluorescence peaks detected and number of peaks masked due to reflection/scattering, two angle ranges were found optimal θ = 30-40° and θ = 55-65°. For the comparison with other samples and methods, θ = 60° is used here.

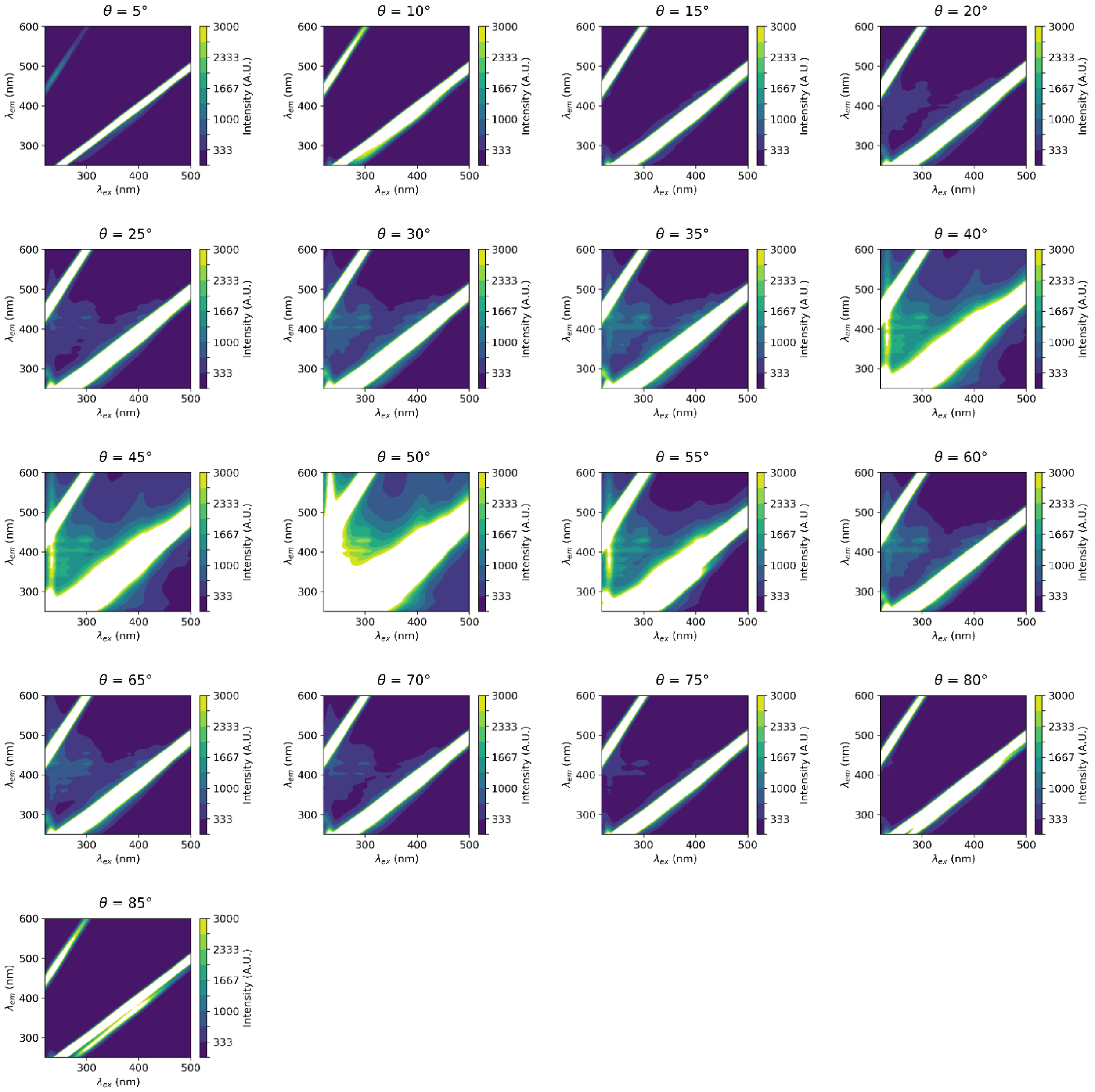
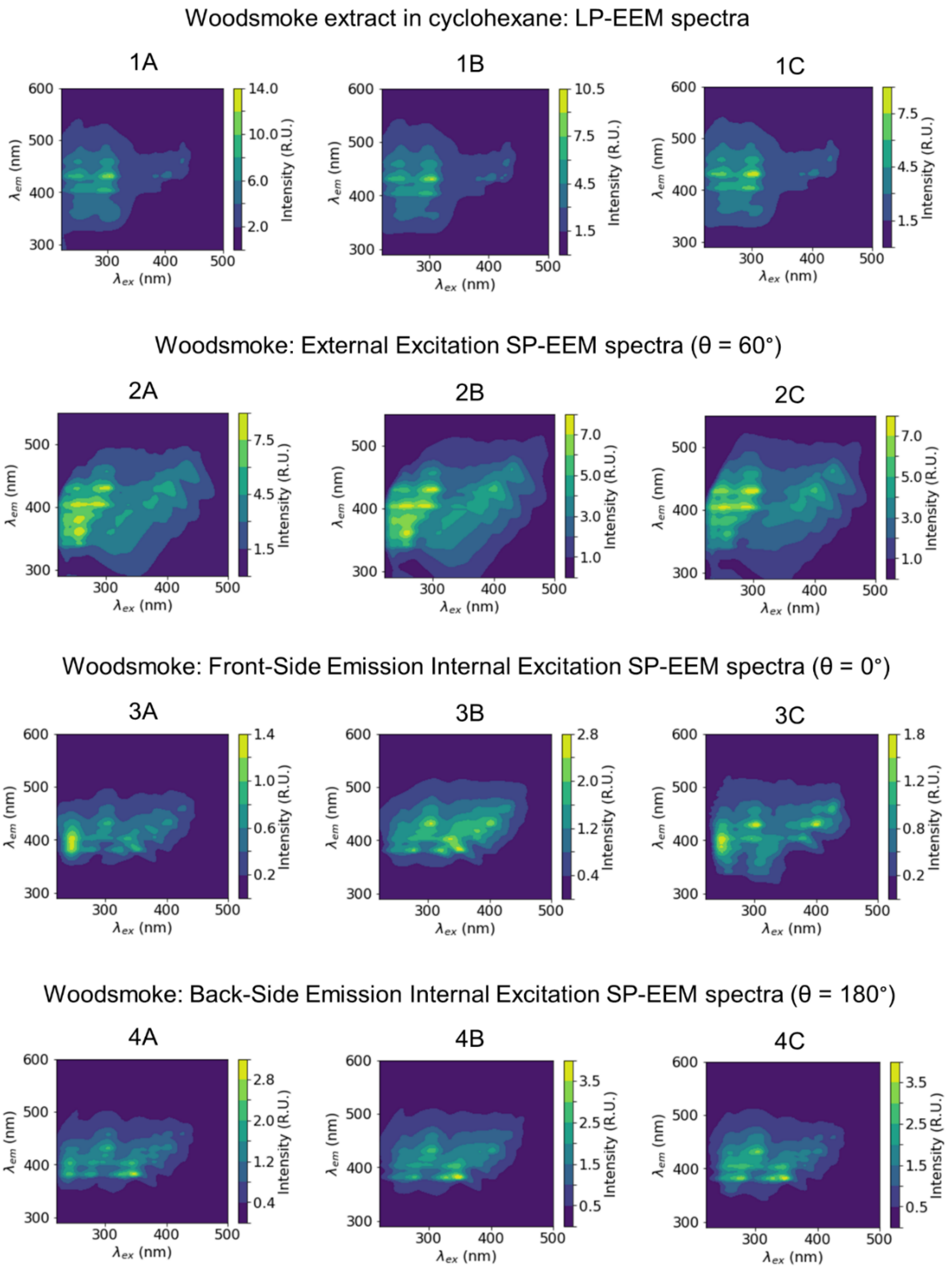


Figure 5.9: External excitation SP-EEMs of woodsmoke at varying θ. θ = 60° was chosen as optimum angle for external excitation SP-EEM measurements.

5.3.2 Comparison of SP-EEMs and LP-EEMs

The internal excitation SP-EEMs at θ = 0°, θ = 180°and external excitation θ = 60° for three different woodsmoke samples are compared to LP-EEMs in Figure 5.10. In internal excitation SP-

EEM, PDMS layer acts as waveguide, so the optical path length is significantly greater than in the external excitation leading to the increase in the PDMS fluorescent background, see Figure 5.12. Blank signal subtraction results in either (i) masking of sample fluorescence peaks at $\lambda_{ex} < 250$nm as visible in Figure 5.10 (3A and 3C) or (ii) the suppression or complete removal of woodsmoke fluorescence peaks at $\lambda_{ex} < 250$nm as visible in Figure 5.10 (3B, 4A-C). In internal excitation SP-EEMs the intensity of scattering peaks is lower compared to external excitation. After the subtraction of the background scatter and interpolation of fluorescence[71,114], the internal excitation yields better features near scattering lines than the external excitation arrangement. This is beneficial for analysis of the samples with high molecular weight (more complex) compounds as their EEM fingerprints are typically shifted to the upper-right corner of the spectra.[114]


Woodsmoke extract in cyclohexane: LP-EEM spectra

Woodsmoke: External Excitation SP-EEM spectra (θ = 60°)

Woodsmoke: Front-Side Emission Internal Excitation SP-EEM spectra (θ = 0°)

Woodsmoke: Back-Side Emission Internal Excitation SP-EEM spectra (θ = 180°)

Figure 5.10: Comparison of SP-EEMs obtained using three different excitation configurations with LP-EEMs for woodsmoke PM. The external excitation SP-EEMs match well with LP-EEMs. Both front-side emission and back-side emission internal excitation SP-EEMs lose spectral information for $\lambda_{ex} < 250$nm due to interference of PDMS fluorescence (Figure 5.12).

To demonstrate the applicability of the SP-EEM to other pollutants, in addition to wood smoke, external excitation SP-EEM analysis (at θ = 60°) was performed for cigarette smoke samples. Figure 5.11 shows all three SP-EEM sample show good repeatability and correlation with LP-EEM. Cyclohexane extraction yielded spectra with higher relative peak intensity for $\lambda_{ex} < 250$nm, which is possible due to difference in extraction efficiencies, sample aging in the PDMS, or uncertainties in the background subtraction in the SP-EEM analysis.

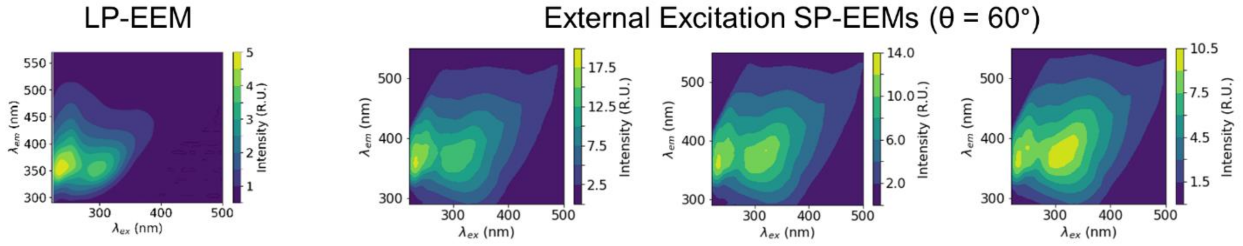


Figure 5.11: External excitation SP-EEMs show most of the peaks defining cigarette smoke PM fluorescence signature when compared to LP-EEM. The removal and interpolation of Rayleigh and Raman scattering lines leads to broadening or removal of fluorescence in SP-EEMs compare to LP-EEMs as seen for woodsmoke PM (Figure 5.10).

5.3.3 Interference of PDMS Fluorescence in Internal vs External Excitation SP-EEMs

Figure 5.12 shows external excitation SP-EEM at optimum θ (θ = 60°), front-side emission and back-side emission internal excitation SP-EEMs of woodsmoke PM samples without subtracting SP-EEM of a blank analysis substrate (shown in Figure 5.12) during pre-processing of EEM data. No blank subtraction results in overlap of PDMS fluorescence ($\lambda_{ex} < 250$nm, $350\text{nm} < \lambda_{em} < 450$nm) with woodsmoke PM fluorescence. For external excitation SP-EEM, the PDMS fluorescence is negligible compared to woodsmoke PM fluorescence however for internal excitation SP-EEMs, PDMS fluorescence is comparable to woodsmoke PM fluorescence which masks much of the PM fluorescence signal at $\lambda_{ex} < 250$nm. Hence internal excitation SP-EEM fluorescence at $\lambda_{ex} < 250$nm is either masked by PDMS fluorescence or completely removed during blank subtraction.

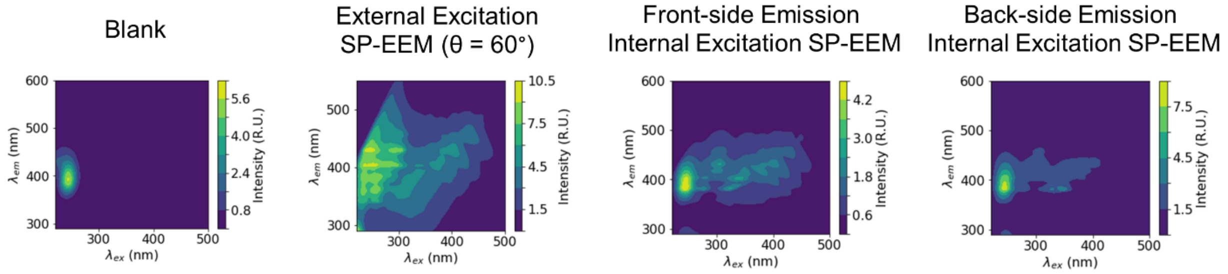


Figure 5.12: SP-EEMs of woodsmoke PM at θ = 60° without blank subtraction using three different excitation emission optics. Intensity of PDMS fluorescence ($\lambda_{ex} < 250$nm, 350nm < $\lambda_{em} < 450$nm) compared to woodsmoke fluorescence peaks ($\lambda_{ex} > 250$nm) is least in external excitation SP-EEM.

5.4 Chapter 5 Conclusions

This chapter demonstrates novel chemical analysis technique of combustion generated ultrafine PM, which eliminates the need for extraction into the liquid phase and can be performed *in-situ*. The method can be used in the development of sensors for measurements of chemical composition of combustion PM. Optical configuration and the choice of the solid phase solvent should be considered in optimization. In external excitation, for shallow angles, excitation intensity is lower; for angles close to normal, one needs to consider optical path length, solid angle of the detection optics and losses due to inner filter effect.[165] Internal excitation yielded stronger interference from PDMS fluorescence for $\lambda_{ex} < 250$nm. For compounds that fluoresce at longer wavelengths, such as higher molecular weight hydrocarbons[114], the internal excitation can provide significant advantages in sensor design and lower elastic scatter from excitation light source.[172] To reduce the fluorescent background other (than PDMS) solid phase solvents with lower fluorescent signatures can be used.

Chapter 6 Thesis Conclusions and Future Work

- Organic fraction of PM obtained by controlling the ethane and ethylene flame temperatures by Ar dilution in an IGFR reveal that the colder flames produce lower PM yield; however, the PM TOC content increased significantly. This increase in organic carbon, also observed in PM from low temperature cookstove burning vs high temperature diesel combustion, is associated with the presence of potentially carcinogenic HMW PAHs.
- Multivariate statistical analysis of overlapping fluorescence signatures of PM solvent extracts, (PCR-EEM method) can be used to estimate the combustion PM organic fraction, in particular, the concentration of the HMW PAH group.
- A novel two-stage ESP-based capillary sensor, designed and evaluated for collection and analysis of fine and ultrafine PM, has a collection efficiency of ~65% and can be integrated with *in-situ* spectroscopic analysis techniques such as fluorescence spectroscopy.
- A novel *in-situ* spectroscopic analysis technique (SP-EEM analysis) for combustion generated ultrafine PM eliminates the need for extraction into the liquid phase and can be used in the development of sensors for measurements of chemical composition of combustion PM.
- This research provides an analysis tool for collection of airborne nanoparticles and *in-situ* analysis of PM chemical composition to determine their toxic potential using a spectroscopic technique.

- This research direction can be expanded to demonstrate the health impact value of the EEM analysis approach. Research on investigating association between EEM derived concentrations and lung function changes using metrics such as PM2.5 exposure vs. exhaled NO; woodsmoke exposure vs. peak expiratory flow (PEF) and mid expiratory flow (MEF), etc.

- A current barrier in development of a UPM sensors for continuous PM toxicity monitoring using SP-EEM is quantifying PM concentration *in-situ*. Preliminary experiments on time-resolved measurements of UPM concentration via absorbance spectroscopy show promising results to solve this problem. Further, optimization of the developed UPM sensor cartridge for absorption, UV-vis and SP-EEM measurements is essential. The cartridge can further be developed into a UPM wearable monitor, sensor integration for absorption, and EEM measurements.

- A fully developed PM chemical sensor can be integrated with other sensors for monitoring individual PM exposure and demonstrating the public health relevance and the broader applications of the technology in large, long-term epidemiological studies. Analysis using the UPM sensors can help evaluate associations of PM-PAH exposure biomarkers for altered lung function. Future applications of this technology can be focused on established and emerging areas of health research, personalized medicine, global health, and regulatory and security areas.

Funding Acknowledgements

This work was supported by the National Institute of Health grants U01 EB021923 and NIEHS R42ES026532 subcontract to the University of Washington.

Chapter 7 Appendix

7.1 Introduction

Attached are the Python codes used for analysis of data generated in this thesis. The codes in the Jupyter Notebook (.ipynb) and the data associated with the codes can be found here: https://github.com/gauravsm31/Gaurav_Mahamuni_PhD_Thesis_Code

7.2 Python Code for LMW and HMW PAH Concentration Estimation

```
import numpy as np
import numpy.matlib as M
from numpy.matlib import rand,zeros,ones,empty,eye
import matplotlib as mpl
from scipy import linalg
from scipy import misc
from sklearn import mixture
import matplotlib.pyplot as plt
import matplotlib as mpl
from numpy import genfromtxt
from sklearn import mixture
from pylab import *
from sklearn.decomposition import PCA
from pandas import DataFrame
from sklearn import linear_model
import statsmodels.api as sm
import pandas as pd

# Load 100 samples and their labels
X = np.zeros((20,201000))
Y1 = np.zeros((20,1))
Y2 = np.zeros((20,1))
Y = np.zeros((20,16))

for i in range(20):
   j=i+1
   s = np.genfromtxt("../20samples Original/Sample%s.csv" %j, dtype=float, delimiter=',')
   X[i,:] = np.reshape(s,(1,201000))

Ynorm = pd.read_csv("../20samples Original/Normalized PAH Conc ng_ug 5_13_2019.csv",
delimiter=',')
```

```
# Y1[:20,0] = np.genfromtxt("./20labels Original/LMW20samples.csv", dtype=float,
delimiter=',')
# Y2[:20,0] = np.genfromtxt("./20labels Original/HMW20samples.csv", dtype=float,
delimiter=',')
# Y[:20,:] = np.genfromtxt("./20labels Original/GCMS20samples.csv", dtype=float,
delimiter=',')

Y1[:20,0] = np.array(Ynorm.iloc[:,2:9].sum(axis=1))
Y2[:20,0] = np.array(Ynorm.iloc[:,9:].sum(axis=1))
Y[:20,:] = np.array(Ynorm.iloc[:,2:])

SamplesInfo = pd.read_csv("../20samples Original/SampleInfo.csv", delimiter=',')
TempRanges = pd.read_csv("../20samples Original/TempRanges.csv", delimiter=',')
TEF = pd.read_csv("../20samples Original/TEF.csv", delimiter=',')

T_low_list = []
T_high_list = []
T_star_list = []
for i in range(SamplesInfo.shape[0]):
   fuel = SamplesInfo.iloc[i][1]
   dil = SamplesInfo.iloc[i][2]
   T_low_list = T_low_list + [round(TempRanges[(TempRanges['fuel']==fuel) &
(TempRanges['dil']==dil)].iloc[0][2], 2)]
   T_high_list = T_high_list + [round(TempRanges[(TempRanges['fuel']==fuel) &
(TempRanges['dil']==dil)].iloc[0][3], 2)]
   T_star_list = T_star_list + [round(TempRanges[(TempRanges['fuel']==fuel) &
(TempRanges['dil']==dil)].iloc[0][4], 2)]
SamplesInfo.insert(loc=6, column='T_low (K)', value=T_low_list)
SamplesInfo.insert(loc=7, column='T_high (K)', value=T_high_list)
SamplesInfo.insert(loc=8, column='T_star (K)', value=T_star_list)

PAH_Conc_22 = np.matrix(SamplesInfo.iloc[:,12:])
TEF_22PAHs = np.matrix(TEF.iloc[:,1])
TEF_Samples = np.matmul(TEF_22PAHs,PAH_Conc_22.T).T
SamplesInfo['TEF'] = TEF_Samples

PAH_Conc_8 = PAH_Conc_22[:,:8]
PAH_Conc_14 = PAH_Conc_22[:,8:]
TEF_8PAHs = np.matrix(TEF.iloc[:8,1])
TEF_14PAHs = np.matrix(TEF.iloc[8:,1])
TEF_LMW_Samples = np.matmul(TEF_8PAHs,PAH_Conc_8.T).T
TEF_HMW_Samples = np.matmul(TEF_14PAHs,PAH_Conc_14.T).T
SamplesInfo['TEF LMW'] = TEF_LMW_Samples
SamplesInfo['TEF HMW'] = TEF_HMW_Samples
```

```
Ex = np.genfromtxt('../20samples Original/Ex.csv', dtype=float, delimiter=',')
Em = np.genfromtxt('../20samples Original/Em.csv', dtype=float, delimiter=',')
import pandas
df = pandas.read_csv('../20labels Original/PAHs.csv')
index_col='PAHs'
leg = df[index_col][:16]
lst = list(leg)


# Leave One Out Cross Validation
from sklearn.decomposition import PCA
from sklearn.metrics import r2_score

pcdim = 15


Rsq1_train = np.zeros((pcdim,1))
Rsq2_train = np.zeros((pcdim,1))
Rsq1_test = np.zeros((pcdim,1))
Rsq2_test = np.zeros((pcdim,1))

err1_train = np.zeros((pcdim,1))
err2_train = np.zeros((pcdim,1))
err1_test = np.zeros((pcdim,1))
err2_test = np.zeros((pcdim,1))

for j in range(pcdim):

   Ypred1_train = np.zeros(Y1.shape)
   Ypred2_train = np.zeros(Y2.shape)

   Ypred1_test = np.zeros(Y1.shape)
   Ypred2_test = np.zeros(Y2.shape)

   M = X.mean(0)
   X_dmean = X - np.array([M,]*20)

   pca = PCA()
   scores = pca.fit_transform(X_dmean)
   comps = pca.components_
   D = pca.explained_variance_

   pca_dim = pcdim

   regr = linear_model.LinearRegression()
   x_train = scores[:,:pca_dim]
```

```
y1 = Y1
y2 = Y2

# Y1
lm1 = linear_model.LinearRegression()
model1 = lm1.fit(x_train,y1)
Ypred1_train = lm1.predict(x_train)

# Y2
lm2 = linear_model.LinearRegression()
model2 = lm2.fit(x_train,y2)
Ypred2_train = lm2.predict(x_train)

for i in range(Y1[:,0].size):

    # Split x into training set and validation set
    X_train = np.delete(X, i, 0)
    X_val = X[i]
    X_val = np.reshape(X_val,(1,X_val.size))

    # Demean X_train and X_val
    M = X_train.mean(0)
    X_train_dmean = X_train - np.array([M,]*19)
    X_val_dmean = X_val - M

    # Transform X_train_dmean to PC space
    pca = PCA()
    scores = pca.fit_transform(X_train_dmean)
    comps = pca.components_
    D = pca.explained_variance_

    # Transform X_test_dmean to PC space
    scores_test = pca.transform(X_val_dmean)

    # Defining training and validation x and y
    regr = linear_model.LinearRegression()
    xtrain = scores[:,:j+1]
    ytrain1 = np.delete(Y1, i, 0)
    ytrain2 = np.delete(Y2, i, 0)
    xtest = scores_test[:,:j+1]
    ytest1 = Y1[i]
    ytest2 = Y2[i]

    # Fit linear model on training data
    lm1 = linear_model.LinearRegression()
    model1 = lm1.fit(xtrain,ytrain1)
```

```
        lm2 = linear_model.LinearRegression()
        model2 = lm2.fit(xtrain,ytrain2)

        # Predict validation data using fitted model
        Ypred1_test[i] = lm1.predict(xtest)
        Ypred2_test[i] = lm2.predict(xtest)

    Rsq1_train[j,0] = r2_score(Y1, Ypred1_train)
    Rsq1_test[j,0] = r2_score(Y1, Ypred1_test)
    Rsq2_train[j,0] = r2_score(Y2, Ypred2_train)
    Rsq2_test[j,0] = r2_score(Y2, Ypred2_test)

    err1_train[j,0] = np.sum(np.square(Ypred1_train - Y1))/20
    err1_test[j,0] = np.sum(np.square(Ypred1_test - Y1))/20
    err2_train[j,0] = np.sum(np.square(Ypred2_train - Y2))/20
    err2_test[j,0] = np.sum(np.square(Ypred2_test - Y2))/20

from matplotlib import rcParams
import matplotlib.pyplot as plt
rcParams['font.family'] = 'arial'

f12 = plt.figure(figsize=(15,6))
ax1 = f12.add_subplot(1,2,1)
#
ax1.plot(range(1,16),err1_train[:,0],markersize=10,marker="o",c='dodgerblue',alpha=0.9)
ax1.plot(range(1,16),err1_test[:,0],markersize=10,marker="o",c='dodgerblue',alpha=0.9)
ax1.set_xlabel('Principal Component',fontsize=20)
ax1.set_ylabel('Test Mean Squared Error',fontsize=20)
ax1.tick_params(axis='x', labelsize=20)
ax1.tick_params(axis='y', labelsize=20)
# ax1.set_ylim([0.9, 1.01])
# ax1.legend(['Test'],fontsize=12,loc=1)
ax1.set_title('Low Molecular Weight PAHs \n (MW 128 - 202 g/mol)',fontsize=20)
plt.tight_layout()

ax1 = f12.add_subplot(1,2,2)
#
ax1.plot(range(1,16),err2_train[:,0],markersize=10,marker="o",c='dodgerblue',alpha=0.9)
ax1.plot(range(1,16),err2_test[:,0],markersize=10,marker="o",c='darkorange',alpha=0.9)
ax1.set_xlabel('Principal Component',fontsize=20)
ax1.set_ylabel('Test Mean Squared Error',fontsize=20)
ax1.tick_params(axis='x', labelsize=20)
ax1.tick_params(axis='y', labelsize=20)
# ax1.set_ylim([0.9, 1.01])
# ax1.legend(['Test'],fontsize=12,loc=1)
ax1.set_title('High Molecular Weight PAHs \n (MW 226 - 302 g/mol)',fontsize=20)
```

```
plt.tight_layout(pad=0.4, w_pad=5, h_pad=5)


pca_dim_LMW = 5
pca_dim_HMW = 4

Ypred1_train = np.zeros(Y1.shape)
Ypred2_train = np.zeros(Y2.shape)

Ypred1_test = np.zeros(Y1.shape)
Ypred2_test = np.zeros(Y2.shape)

M = X.mean(0)
X_dmean = X - np.array([M,]*20)

pca = PCA()
scores = pca.fit_transform(X_dmean)
comps = pca.components_
D = pca.explained_variance_

print(np.cumsum(pca.explained_variance_ratio_))

plt.plot(np.cumsum(pca.explained_variance_ratio_))
plt.xlabel('number of components')
plt.ylabel('cumulative explained variance');

regr = linear_model.LinearRegression()
x1 = scores[:,:pca_dim_LMW]
x2 = scores[:,:pca_dim_HMW]
y1 = Y1
y2 = Y2

# Y1
lm1 = linear_model.LinearRegression()
model1 = lm1.fit(x1,y1)
Ypred1_train = lm1.predict(x1)

# Y2
lm2 = linear_model.LinearRegression()
model2 = lm2.fit(x2,y2)
Ypred2_train = lm2.predict(x2)

for i in range(Y1[:,0].size):

  # Split x into training set and validation set
  X_train = np.delete(X, i, 0)
```

```
    X_val = X[i]
    X_val = np.reshape(X_val,(1,X_val.size))

    # Demean X_train and X_val
    M = X_train.mean(0)
    X_train_dmean = X_train - np.array([M,]*19)
    X_val_dmean = X_val - M

    # Transform X_train_dmean to PC space
    pca = PCA()
    scores = pca.fit_transform(X_train_dmean)
    comps = pca.components_
    D = pca.explained_variance_

    # Transform X_test_dmean to PC space
    scores_test = pca.transform(X_val_dmean)

    # Defining training and validation x and y
    regr = linear_model.LinearRegression()
    xtrain1 = scores[:,:pca_dim_LMW]
    xtrain2 = scores[:,:pca_dim_HMW]
    ytrain1 = np.delete(Y1, i, 0)
    ytrain2 = np.delete(Y2, i, 0)
    xtest1 = scores_test[:,:pca_dim_LMW]
    xtest2 = scores_test[:,:pca_dim_HMW]
    ytest1 = Y1[i]
    ytest2 = Y2[i]

    # Fit linear model on training data
    lm1 = linear_model.LinearRegression()
    model1 = lm1.fit(xtrain1,ytrain1)
    lm2 = linear_model.LinearRegression()
    model2 = lm2.fit(xtrain2,ytrain2)

    # Predict validation data using fitted model
    Ypred1_test[i] = lm1.predict(xtest1)
    Ypred2_test[i] = lm2.predict(xtest2)

# Plot parity plots for prediction of each of the 16 compounds

def parity_plots(y_test, pred_test,
                 y_train=None, pred_train=None,
                titles = None):
    """Make parity plots for  16 regression outputs
    **In the future make it so any number of parity plots can be accomidate
    Args:
```

```
    y_test (np.array): expected values for test dataset
        (shape = samples X values, currently expected as samples X 16)
    pred_test (np.array): predicted values from test dataset
    y_train (np.array, optional): expected values from test dataset
    pred_test (np.array, optional): predicted values from test dataset
    titles (list of strings, optional) : list of strings to use as tiles on plots

Returns:
    fig: (matplotlib.figure.Figure) : Parity Plots
"""

import numpy as np
import matplotlib.pyplot as plt
from matplotlib import gridspec
from sklearn.metrics import r2_score


# Define colors for test and train
# http://colorbrewer2.org/#type=qualitative&scheme=Set1&n=3
c_low = [55 / 255, 126 / 255, 184 / 255]
c_high = [228 / 255, 26 / 255, 28 / 255]
c_train = [77 / 255, 175 / 255, 74 / 255]

fig = plt.figure(figsize=(18, 9))

outer = gridspec.GridSpec(1, 2, wspace=0.5, hspace=0.2)


# Generate a parity plot for each class using a loop
for i in range(0, pred_test.shape[1]):
    inner = gridspec.GridSpecFromSubplotSpec(2, 1,
                    subplot_spec=outer[i], wspace=0.1, hspace=0.4,
                                        height_ratios=[3,1])

    #first axis - parity plot
    ax = plt.Subplot(fig, inner[0])

    # calculate appropriate scale for x and y axis
    maximums = [np.max(pred_test[:, i]), np.max(y_test[:, i])]
    minimums = [np.min(pred_test[:, i]), np.min(y_test[:, i])]
    ax_max = np.max(maximums)
    ax_min = np.min(minimums)
    ax_range = ax_max - ax_min
    #add some padding so no data are on the edges
    ax_max = ax_max + ax_range*.25
    ax_min = ax_min - ax_range*.25
```

```
    # add the training data to the plot
    if y_train is None:
        pass
    else:
        ax.scatter(y_train[:, i], pred_train[:, i], color=c_train, s=100)

    # plot the test data
    # select blue color for low MW, select red color for high MW
    if i<1:
        color = c_low
    else:
        color = c_high
    ax.scatter(y_test[:, i], pred_test[:, i], color=color, s=100)
    # add perfect prediction line
    ax.plot([ax_min, ax_max], [ax_min, ax_max], color='k',
            label='Perfect Prediction', linewidth=1)

    ax.set_ylim(ax_min, ax_max)
    ax.set_xlim(ax_min, ax_max)
    ax.set_xlabel('Measured GCMS Concentration \n (ng PAH/$\mu$g-soot)', fontsize=20)
    ax.set_ylabel('Estimated Concentration using EEM \n (ng PAH/$\mu$g-soot)',
fontsize=20)
    ax.tick_params(labelsize=20)
    if titles is None:
        pass
    else:
        ax.set_title(titles[i], fontsize=20, weight='bold')
    # add an option for putting a letter on each plot later
    #       # add letters to each plot
    #       letter = ['A', 'B', 'C']
    #       bbox_props = dict(boxstyle="round, pad=0.1", fc="white", ec="white", lw=2)
    #       plt.text(-0.5, 6.5, letter[i], ha="center", va="center",
    #           size=20,
    #           bbox=bbox_props)
    fig.add_subplot(ax)

    ######## Put R^2 values in the second panel below the plot #########
    ax = plt.Subplot(fig, inner[1])

    # Calulate and report R^2 train if training data provided
    if y_train is None:
        pass
    else:
        r2_train = r2_score(y_train[:, i], pred_train[:, i])
        t = ax.text(0.5,0.5, '$R^2$ train = %.3f' % r2_train, size=20)
```

```
        t.set_ha('center')

      # Calulate and report R^2 test
      r2_test = r2_score(y_test[:, i], pred_test[:, i])
      t = ax.text(0.5, 0.2, '$R^2$ test = %.3f' % r2_test, size=20)
      t.set_ha('center')
      fig.add_subplot(ax)

      #hide the axis
      ax.get_xaxis().set_visible(False)
      ax.get_yaxis().set_visible(False)
      ax.axis('off')

  return fig


titles = ['Low Molecular Weight PAHs \n (MW 128 - 202 g/mol)','High Molecular Weight
PAHs \n (MW 226 - 302 g/mol)']

ytest = np.zeros((20,2))
ytest[:,0] = Y1[:,0]
ytest[:,1] = Y2[:,0]

ypred_test = np.zeros((20,2))
ypred_test[:,0] = Ypred1_test[:,0]
ypred_test[:,1] = Ypred2_test[:,0]

ytrain = np.zeros((20,2))
ytrain[:,0] = Y1[:,0]
ytrain[:,1] = Y2[:,0]

ypred_train = np.zeros((20,2))
ypred_train[:,0] = Ypred1_train[:,0]
ypred_train[:,1] = Ypred2_train[:,0]

fig = parity_plots( \
            y_test=ytest, \
            pred_test=ypred_test, \
            y_train=ytrain, \
            pred_train = ypred_train,
            titles=titles)



# Importing Woodsmoke and Diesel Samples
```

```
test_files = pd.read_excel('../2019-11-09 Diesel Woodsmoke GCMS Diluted2
Samples/TestSamplesData.xlsx')
test_eem_data = np.zeros((test_files.shape[0],1000,201))

j=0
for file in test_files['filename'].to_list():
    test_eem_data[j,:,:] = np.genfromtxt("../2019-11-09 Diesel Woodsmoke GCMS Diluted2
Samples/" + file, dtype=float, delimiter=',')
    j=j+1

test_ex = np.genfromtxt("../2019-11-09 Diesel Woodsmoke GCMS Diluted2
Samples/ex_woodsmoke.csv", dtype=float, delimiter=',')
test_em = np.genfromtxt("../2019-11-09 Diesel Woodsmoke GCMS Diluted2
Samples/em_woodsmoke.csv", dtype=float, delimiter=',')

def em_interpolation(eem):
    intp_eem = np.zeros(eem.shape)
    intp_eem[8:,:] = eem[:-8,:]
    return intp_eem

for j in range(test_eem_data.shape[0]):
    test_eem_data[j,:,:] = em_interpolation(test_eem_data[j,:,:])

def view_eems(ex,em,eems):
    import matplotlib.ticker as ticker

    def myfmt(x, pos):
        return '{0:.1f}'.format(x)

    import math
    fig = plt.figure(figsize=(4*3.5,math.ceil((eems.shape[0])/3)*3.5))
    sams = [0,1,2,3,7,8,4,5,6]
    titles = ['Diesel','Diesel','Diesel',\
             'Woodsmoke\nForced Draft','Woodsmoke\nNatural Draft','Woodsmoke\nNatural
Draft',\
            'Woodsmoke\nNatural Draft','Woodsmoke\nForced Draft','Woodsmoke\nForced
Draft']
    j=0
    for i in sams:
        ax = fig.add_subplot(math.ceil((eems.shape[0])/3),3,j+1)
        fl = eems[i]
        cont = ax.contourf(ex, em, fl)
        ax.set_xlim([220,500])
        ax.set_ylim([250,600])
        ttl = ax.set_title(titles[i],fontsize=16)
        ax.set_ylabel('$\lambda_{em}$ (nm)',fontsize=16)
```

```
        ax.set_xlabel('$\lambda_{ex}$ (nm)',fontsize=16)
        ax.tick_params(axis='x', labelsize=16)
        ax.tick_params(axis='y', labelsize=16)
        cbar = fig.colorbar(cont, spacing='proportional',format=ticker.FuncFormatter(myfmt))
        cbar.ax.set_ylabel('Intensity (R.U.)',fontsize=16)
        cbar.ax.tick_params(axis='y',labelsize=16)
        for label in cbar.ax.yaxis.get_ticklabels()[::2]:
          label.set_visible(False)
        j=j+1
    plt.tight_layout(pad=0.4, w_pad=1, h_pad=2)

    return fig



# Importing GCMS data for test samples

test_GCMS_diesel_data = pd.read_csv("../GCMS Data/GCMS_diesel.csv", delimiter=',')
test_GCMS_woodsmoke_data = pd.read_csv("../GCMS Data/GCMS_woodsmoke.csv",
delimiter=',')
test_GCMS_data = pd.concat([test_GCMS_diesel_data, test_GCMS_woodsmoke_data])


# Left Join GCMS data on samples data
test_GCMS_data_merged = test_files.join(test_GCMS_data.set_index('Sample ID'),
            on='Sample ID',
            how='left')


Y1_test = np.zeros((test_GCMS_data_merged.shape[0],1))
Y2_test = np.zeros((test_GCMS_data_merged.shape[0],1))
Y_test = np.zeros((test_GCMS_data_merged.shape[0],16))


Y1_test[:,0] = np.array(test_GCMS_data_merged.iloc[:,4:11].sum(axis=1))
Y2_test[:,0] = np.array(test_GCMS_data_merged.iloc[:,11:].sum(axis=1))
Y_test[:,:] = np.array(test_GCMS_data_merged.iloc[:,4:])


# Load 20 samples and their labels
X = np.zeros((20,201000))
Y1 = np.zeros((20,1))
Y2 = np.zeros((20,1))
Y = np.zeros((20,16))

for i in range(20):
```

```
   j=i+1
   s = np.genfromtxt("../20samples Original/Sample%s.csv" %j, dtype=float, delimiter=',')
   X[i,:] = np.reshape(s,(1,201000))

Ynorm = pd.read_csv("../20samples Original/Normalized PAH Conc ng_ug 5_13_2019.csv",
delimiter=',')

# Y1[:20,0] = np.genfromtxt("./20labels Original/LMW20samples.csv", dtype=float,
delimiter=',')
# Y2[:20,0] = np.genfromtxt("./20labels Original/HMW20samples.csv", dtype=float,
delimiter=',')
# Y[:20,:] = np.genfromtxt("./20labels Original/GCMS20samples.csv", dtype=float,
delimiter=',')

Y1[:20,0] = np.array(Ynorm.iloc[:,2:9].sum(axis=1))
Y2[:20,0] = np.array(Ynorm.iloc[:,9:].sum(axis=1))
Y[:20,:] = np.array(Ynorm.iloc[:,2:])

# Demean data

X_test = np.reshape(test_eem_data,(test_eem_data.shape[0],201000))
M = X.mean(0)
X_dmean = X - np.array([M,]*20)
X_test_dmean = X_test - np.array([M,]*test_eem_data.shape[0])

from sklearn.decomposition import PCA

pca = PCA()
scores = pca.fit_transform(X_dmean)
comps = pca.components_
D = pca.explained_variance_

pca_dim_LMW = 5
pca_dim_HMW = 4

# Transform X_test_dmean to PC space
scores_test = pca.transform(X_test_dmean)

# Defining training and validation x and y
xtrain1 = scores[:,:pca_dim_LMW]
xtrain2 = scores[:,:pca_dim_HMW]
ytrain1 = Y1
ytrain2 = Y2
xtest1 = scores_test[:,:pca_dim_LMW]
xtest2 = scores_test[:,:pca_dim_HMW]
ytest1 = Y1_test
```

```
ytest2 = Y2_test

# Fit linear model on training data
lm1 = linear_model.LinearRegression()
model1 = lm1.fit(xtrain1,ytrain1)
lm2 = linear_model.LinearRegression()
model2 = lm2.fit(xtrain2,ytrain2)


# Predict validation data using fitted model
Ypred1_test = lm1.predict(xtest1)
Ypred2_test = lm2.predict(xtest2)

# Predict validation data using fitted model
Ypred1_train = lm1.predict(xtrain1)
Ypred2_train = lm2.predict(xtrain2)

titles = ['Low Molecular Weight PAHs \n (MW 128 - 202 g/mol)','High Molecular Weight
PAHs \n (MW 226 - 302 g/mol)']

ytest = np.zeros((9,2))
ytest[:,0] = Y1_test[:,0]
ytest[:,1] = Y2_test[:,0]

ypred_test = np.zeros((9,2))
ypred_test[:,0] = Ypred1_test[:,0]
ypred_test[:,1] = Ypred2_test[:,0]

ytrain = np.zeros((20,2))
ytrain[:,0] = Y1[:,0]
ytrain[:,1] = Y2[:,0]

ypred_train = np.zeros((20,2))
ypred_train[:,0] = Ypred1_train[:,0]
ypred_train[:,1] = Ypred2_train[:,0]

fig = parity_plots(y_test=ytest, pred_test=ypred_test, y_train=ytrain, pred_train =
ypred_train, titles=titles)
```

7.3 Python Code for Computational Mixing of EEMs

```
import numpy as np
import torch
import torch.optim as optim
import torch.nn as nn
import torch.nn.functional as F
import matplotlib.pyplot as plt
import matplotlib as mpl

import numpy as np
import numpy.matlib as M
from numpy.matlib import rand,zeros,ones,empty,eye
import matplotlib as mpl
from scipy import linalg
from scipy import misc
from sklearn import mixture
import matplotlib.pyplot as plt
import matplotlib as mpl
from numpy import genfromtxt
from sklearn import mixture
from pylab import *
from sklearn.decomposition import PCA
from pandas import DataFrame
from sklearn import linear_model
import statsmodels.api as sm
import pandas as pd

from itertools import cycle, islice
import pandas
np.random.seed(2)


# X has 20 original samples
X_orig = []
for i in range(20):
   Orig_data = np.genfromtxt('../20samples Original/Sample%s.csv' %(i+1), delimiter=',')
   #Orig_data = np.reshape(Orig_data,(1,201000))
   X_orig.append(Orig_data)
X_orig = np.array(X_orig)
Ex = np.genfromtxt('../20samples Original/Ex.csv', delimiter=',')
Em = np.genfromtxt('../20samples Original/Em.csv', delimiter=',')
Orig_labels = np.genfromtxt('../20labels Original/GCMS20samples.csv', delimiter=',')

# Importing Woodsmoke and Diesel Samples
```

```
test_files = pd.read_excel('../../Manuscript 2/6labels Original/TestSamplesData.xlsx')
test_eem_data = np.zeros((test_files.shape[0],1000,201))

j=0
for file in test_files['filename'].to_list():
    test_eem_data[j,:,:] = np.genfromtxt("../../Manuscript 2/6samples Original/" + file,
dtype=float, delimiter=',')
    j=j+1

test_ex = np.genfromtxt("../../Manuscript 2/6samples Original/ex_woodsmoke.csv",
dtype=float, delimiter=',')
test_em = np.genfromtxt("../../Manuscript 2/6samples Original/em_woodsmoke.csv",
dtype=float, delimiter=',')

def em_interpolation(eem):
    intp_eem = np.zeros(eem.shape)
    intp_eem[8:,:] = eem[:-8,:]
    return intp_eem

for j in range(test_eem_data.shape[0]):
    test_eem_data[j,:,:] = em_interpolation(test_eem_data[j,:,:])

# Importing GCMS data for test samples

test_GCMS_diesel_data = pd.read_csv("../GCMS Data/GCMS_diesel.csv", delimiter=',')
test_GCMS_woodsmoke_data = pd.read_csv("../GCMS Data/GCMS_woodsmoke.csv",
delimiter=',')
test_GCMS_data = pd.concat([test_GCMS_diesel_data, test_GCMS_woodsmoke_data])

# Left Join GCMS data on samples data
test_GCMS_data_merged = test_files.join(test_GCMS_data.set_index('Sample ID'),
                 on='Sample ID',
                 how='left')

Y_test = np.zeros((test_GCMS_data_merged.shape[0],16))
Y_test[:,:] = np.array(test_GCMS_data_merged.iloc[:,4:])

def makspool(data, size_x, size_y, stride_x, stride_y):
    data_torch = torch.from_numpy(data)
    max_pooled = nn.MaxPool2d((size_x,size_y),stride=(stride_x, stride_y))
    output = max_pooled(data_torch)
    return output.numpy()

X_test_down = makspool(test_eem_data, 10,6,10,6)
X_orig_downsampled = makspool(X_orig, 10,6,10,6)
```

```
np.savetxt("x_orig_downsampled.csv", np.reshape(X_orig_downsampled, (20, 3300)),
delimiter=',')
X_train_down = X_orig_downsampled
Y_train = Orig_labels


X_train_new = []
Y_train_new = []
for i in range(580):
  rand_im = np.random.randint(0,19,3)
  rand = np.random.uniform(0.133,2.33,3)
  new_image = np.zeros(np.shape(X_train_down[0]))
  new_label = np.zeros(np.shape(Y_train[0]))
  for j in range(3):
    new_image = new_image + X_train_down[rand_im[j]] * rand[j]
    new_label = new_label + Y_train[rand_im[j]] * rand[j]
  new_image = new_image + np.random.normal(0, 0.01, np.shape(new_image))
  X_train_new.append(new_image)
  Y_train_new.append(new_label)
X_train_new = np.array(X_train_new)
Y_train_new = np.array(Y_train_new)


X_test_new = []
Y_test_new = []
for i in range(144):
  rand_im = np.random.randint(0,6,3)
  rand = np.random.uniform(0.133,2.33,3)
  new_image = np.zeros(np.shape(X_test_down[0]))
  new_label = np.zeros(np.shape(Y_test[0]))
  for j in range(3):
    new_image = new_image + X_test_down[rand_im[j]] * rand[j]
    new_label = new_label + Y_test[rand_im[j]] * rand[j]
  new_image = new_image + np.random.normal(0, 0.01, np.shape(new_image))
  X_test_new.append(new_image)
  Y_test_new.append(new_label)
X_test_new = np.array(X_test_new)
Y_test_new = np.array(Y_test_new)



Ex = np.genfromtxt('../20samples Original/Ex.csv', dtype=float, delimiter=',')
Em = np.genfromtxt('../20samples Original/Em.csv', dtype=float, delimiter=',')
import pandas
df = pandas.read_csv('../20labels Original/PAHs.csv')
index_col='Acronyms'
```

```
leg = df[index_col][:16]
lst = list(leg)

fsize = 33
n = int(np.around(Ex.size/fsize))
Ex_new = zeros(fsize)
for i in range(fsize):
   Ex_new[i] = np.mean(Ex[i*n+1:(i+1)*n])

fsize = 100
n = int(np.around(Em.size/fsize))
Em_new = zeros(fsize)
for i in range(fsize):
   Em_new[i] = np.mean(Em[i*n+1:(i+1)*n])


X_train_new = np.append(X_train_new, X_train_down, axis=0)
X_test_new = np.append(X_test_new, X_test_down, axis=0)
Y_train_new = np.append(Y_train_new, Y_train, axis=0)
Y_test_new = np.append(Y_test_new, Y_test, axis=0)

X_train_reshape = [np.shape(X_train_new)[0], 1, np.shape(X_train_new)[1],
np.shape(X_train_new)[2]]
X_test_reshape = [np.shape(X_test_new)[0], 1, np.shape(X_test_new)[1],
np.shape(X_test_new)[2]]
X_train_new = X_train_new.reshape(X_train_reshape)
X_test_new = X_test_new.reshape(X_test_reshape)

np.savetxt("xtrain.csv", np.reshape(X_train_new, (600, 3300)), delimiter=',')
np.savetxt("xtest.csv", np.reshape(X_test_new, (150, 3300)), delimiter=',')
np.savetxt("ytrain.csv", Y_train_new, delimiter=',')
np.savetxt("ytest.csv", Y_test_new, delimiter=',')
```

7.4 Python Code for Estimating Concentration of Individual PAHs from the 9 EPA Priority HMW PAHs

```
# Make all numpy available via shorter 'num' prefix

import numpy as np
import numpy.matlib as M
from numpy.matlib import rand,zeros,ones,empty,eye
import matplotlib as mpl
from scipy import linalg
from scipy import misc
from sklearn import mixture
import matplotlib.pyplot as plt
import matplotlib as mpl
from numpy import genfromtxt
from sklearn import mixture
from pylab import *
from sklearn.decomposition import PCA
from pandas import DataFrame
from sklearn import linear_model
import statsmodels.api as sm

from itertools import cycle, islice
import pandas

np.random.seed(3)

# Load samples and their labels

X_train = np.genfromtxt("./xtrain.csv", dtype=float, delimiter=',')
X_test = np.genfromtxt("./xtest.csv", dtype=float, delimiter=',')
Y_train = np.genfromtxt("./ytrain.csv", dtype=float, delimiter=',')
Y_test = np.genfromtxt("./ytest.csv", dtype=float, delimiter=',')
X_orig_ds = np.genfromtxt("./x_orig_downsampled.csv", dtype=float, delimiter=',')

# Load excitation wavelength and emission wavelength data

Ex = np.genfromtxt('../20samples Original/Ex.csv', dtype=float, delimiter=',')
Em = np.genfromtxt('../20samples Original/Em.csv', dtype=float, delimiter=',')

# Import names of toxic compounds (PAHs)

df = pandas.read_csv('../20labels Original/PAHs.csv')
```

```
index_col='Acronyms'
leg = df[index_col][:16]
lst = list(leg)

# Reduce the size of Ex and Em to match that of the EEMs for plotting

fsize = 33
n = int(np.around(Ex.size/fsize))
Ex_new = zeros(fsize)
for i in range(fsize):
   Ex_new[i] = np.mean(Ex[i*n+1:(i+1)*n])

fsize = 100
n = int(np.around(Em.size/fsize))
Em_new = zeros(fsize)
for i in range(fsize):
   Em_new[i] = np.mean(Em[i*n+1:(i+1)*n])


# Demean X data

M = X_train.mean(0)
X_train = X_train - np.array([M,]*600)

X_test = X_test - np.array([M,]*150)


# Perform PCA

from sklearn.decomposition import PCA

pca = PCA()
scores = pca.fit_transform(X_train)
comps = pca.components_

# Variance plot for PCA

D = pca.explained_variance_
f2 = plt.figure(figsize=(12,5))
ax = f2.add_subplot(1,2,1)
ax.scatter(range(100),100*np.divide(D[:100],sum(D[:100])),s=100,marker="o",c='r',alpha
=0.9,edgecolors='none')
ax.set_ylabel('% Variance',fontsize=20)
ax.set_xlabel('Component',fontsize=20)
ax.tick_params(axis='x', labelsize=20)
ax.tick_params(axis='y', labelsize=20)
```

```
ax2 = f2.add_subplot(1,2,2)
ax2.scatter(range(100),log(100*np.divide(D[:100],sum(D[:100]))),s=100,marker="o",c='r',
alpha=0.9,edgecolors='none')
ax2.set_ylabel('log(% Variance)',fontsize=20)
ax2.set_xlabel('Component',fontsize=20)
ax2.tick_params(axis='x', labelsize=20)
ax2.tick_params(axis='y', labelsize=20)
ax2.set_xlim([0,30])
plt.tight_layout(pad=0.4, w_pad=4, h_pad=1.0)

plt.show()

f2.savefig('../../Manuscript 2/v5 Figures/SI Figures/IndiPAHs_variance.png',dpi=300)

# Plotting first four principal components

f2 = plt.figure(figsize=(15,15))
for i in range(4):
  levels=linspace(min(comps[i,:]),max(comps[i,:]),10)
  ax = f2.add_subplot(2,2,i+1)
  eem = np.reshape(comps[i,:],(100,33))
  cont = ax.contourf(Ex_new,Em_new,eem,levels)
  ax.set_xlim([210,500])
  ax.set_ylim([250,600])
  ax.set_title('Principal Component %s' %(i+1))
  ax.set_ylabel('Em')
  ax.set_xlabel('Ex')
  cbar = f2.colorbar(cont)
  cbar.ax.set_ylabel('Fluorescence Intensity')

plt.show()


# Fitting a linear model and validating using test data

scores_test = pca.transform(X_test)
scores_orig_ds = pca.transform(X_orig_ds)

pca_dim = 19

regr = linear_model.LinearRegression()
xtrain = scores[:,:pca_dim]
ytrain = Y_train[:,7:]
xtest = scores_test[:,:pca_dim]
ytest = Y_test[:,7:]
```

```
lm = linear_model.LinearRegression()
model = lm.fit(xtrain,ytrain)

ypred_train = lm.predict(xtrain)

err = np.zeros((ytrain[:,0].size,ytrain[0,:].size))
for i in range(ytrain[:,0].size):
    for j in range(ytrain[0,:].size):
        err[i,j] = np.absolute(ytrain[i,j]-ypred_train[i,j])
err_train = 100*np.divide(np.sum(err,axis=0),np.sum(ytrain,axis=0))



ypred_test = lm.predict(xtest)

err = np.zeros((ytest[:,0].size,ytest[0,:].size))
for i in range(ytest[:,0].size):
    for j in range(ytest[0,:].size):
        err[i,j] = np.absolute(ytest[i,j]-ypred_test[i,j])
err_test = 100*np.divide(np.sum(err,axis=0),np.sum(ytest,axis=0))


x_orig_ds = scores_orig_ds[:,:pca_dim]
ypred_orig_ds = lm.predict(x_orig_ds)


# Checking residual error to check linearity of the model

plt.scatter(ypred_train,ypred_train-ytrain,c='b',s=20,alpha=0.5)
plt.scatter(ypred_test,ypred_test-ytest,c='g',s=20,alpha=0.5)
plt.hlines(y=0,xmin=-0.1,xmax=0.5)
plt.xlabel('Predicted Concentrations')
plt.ylabel('Residuals')
plt.legend(['training data','test data'])


# zero negative predictions
for i in range(ypred_train[:,0].size):
    for j in range(ypred_train[0,:].size):
        if ypred_train[i,j] < 0:
            ypred_train[i,j]=0

for i in range(ypred_test[:,0].size):
    for j in range(ypred_test[0,:].size):
        if ypred_test[i,j] < 0:
            ypred_test[i,j]=0
```

```
# Plotting error for prediction of 16 compound concentration levels using PCR

f5 = plt.figure(figsize=(8,3))
ax = f5.add_subplot(1,1,1)

df1 = ax.bar(np.arange(9),err_test[:],0.35,color='slateblue',label='Test Samples Original')
ax.set_ylabel('% error',fontsize=15)
ax.set_xlabel('PAH Compounds',fontsize=15)
ax.tick_params(axis='x', labelsize=12)
ax.tick_params(axis='y', labelsize=12)
ax.set_xticks(np.arange(9))
ax.set_xticklabels([lst[i] for i in [7,8,9,10,11,12,13,14,15]])
ax.set_title('% error for 150 Test Samples using PCR',fontsize=15)
ax.set_ylim([0, 100])
plt.tight_layout()

f5.savefig('../../Manuscript 2/v5 Figures/SI Figures/IndiPAH_ERROR.png',dpi=300)


# Plot parity plots for prediction of each of the 16 compounds

def parity_plots(y_test, pred_test,
                y_train=None, pred_train=None,
               titles = None):
    """Make parity plots for  16 regression outputs
    **In the future make it so any number of parity plots can be accomidate
    Args:
        y_test (np.array): expected values for test dataset
            (shape = samples X values, currently expected as samples X 16)
        pred_test (np.array): predicted values from test dataset
        y_train (np.array, optional): expected values from test dataset
        pred_test (np.array, optional): predicted values from test dataset
        titles (list of strings, optional) : list of strings to use as tiles on plots

    Returns:
        fig: (matplotlib.figure.Figure) : Parity Plots
    """

    import numpy as np
    import matplotlib.pyplot as plt
    from matplotlib import gridspec
    from sklearn.metrics import r2_score
```

```
    # Define colors for test and train
    # http://colorbrewer2.org/#type=qualitative&scheme=Set1&n=3
    c_low = [55 / 255, 126 / 255, 184 / 255]
    c_high = [228 / 255, 26 / 255, 28 / 255]
    c_train = [77 / 255, 175 / 255, 74 / 255, 0.4]

    fig = plt.figure(figsize=(15, 21))

    outer = gridspec.GridSpec(4, 4, wspace=0.5, hspace=0.2)


    # Generate a parity plot for each class using a loop
    for i in range(0, pred_test.shape[1]):
        inner = gridspec.GridSpecFromSubplotSpec(2, 1,
                        subplot_spec=outer[i], wspace=0.1, hspace=0.4,
                                    height_ratios=[3,1])

        #first axis - parity plot
        ax = plt.Subplot(fig, inner[0])

        # calculate appropriate scale for x and y axis
        maximums = [np.max(pred_test[:, i]), np.max(y_test[:, i])]
        minimums = [np.min(pred_test[:, i]), np.min(y_test[:, i])]
        ax_max = np.max(maximums)
        ax_min = np.min(minimums)
#       ax_min = 0.0
#       ax_max = 20
        ax_range = ax_max - ax_min
        #add some padding so no data are on the edges
        ax_max = ax_max + ax_range*.25
        ax_min = ax_min - ax_range*.25

        # add the training data to the plot
        if y_train is None:
            pass
        else:
            ax.scatter(y_train[:, i], pred_train[:, i], color=c_train)

        # plot the test data
        # select blue color for low MW, select red color for high MW
#       if i<7:
#           color = c_low
#       else:
        color = c_high
        ax.scatter(y_test[:, i], pred_test[:, i], color=color)
        # add perfect prediction line
```

```
        ax.plot([ax_min, ax_max], [ax_min, ax_max], color='k',
                label='Perfect Prediction', linewidth=3)

        ax.set_ylim(ax_min, ax_max)
        ax.set_xlim(ax_min, ax_max)
        ax.set_xlabel('Actual', fontsize=14)
        ax.set_ylabel('Predicted', fontsize=14)
        ax.tick_params(labelsize=12)
        if titles is None:
            pass
        else:
            ax.set_title(titles[i], fontsize=16, weight='bold')
        # add an option for putting a letter on each plot later
        #       # add letters to each plot
        #       letter = ['A', 'B', 'C']
        #       bbox_props = dict(boxstyle="round, pad=0.1", fc="white", ec="white", lw=2)
        #       plt.text(-0.5, 6.5, letter[i], ha="center", va="center",
        #           size=20,
        #           bbox=bbox_props)
        fig.add_subplot(ax)

        ######## Put R^2 values in the second panel below the plot #########
        ax = plt.Subplot(fig, inner[1])

        # Calulate and report R^2 train if training data provided
        if y_train is None:
            pass
        else:
            r2_train = r2_score(y_train[:, i], pred_train[:, i])
            t = ax.text(0.5,0.5, '$R^2$ train = %.3f' % r2_train, size=14)
            t.set_ha('center')

        # Calulate and report R^2 test
        r2_test = r2_score(y_test[:, i], pred_test[:, i])
        t = ax.text(0.5, 0.2, '$R^2$ test = %.3f' % r2_test, size=14)
        t.set_ha('center')
        fig.add_subplot(ax)

        #hide the axis
        ax.get_xaxis().set_visible(False)
        ax.get_yaxis().set_visible(False)
        ax.axis('off')

    return fig
```

```
titles = [
#     'NAP','ACE','FLU','PHE', 'ANT', 'FLA', 'PYR',
         'CPA','BAA','CRY','BFL','BAP','INP',
        'BGP','COR','DBP']

fig = parity_plots(ytest, ypred_test, ytrain, ypred_train, titles=titles)
```

Chapter 8 References

(1) *Office of the Federal Registration (OFR) Appendix A: Priority Pollutants. Fed Reg. 1982;47:52309. Https://Www.Epa.Gov/Sites/Production/Files/2015-09/Documents/Priority-Pollutant-List-Epa.Pdf*.

(2) Davis, J. Characterization of Combustion Generated Particulates Produced in an Inverted Gravity Flame Reactor, University of Washington, Seattle, 2019, Vol. Ph.D.

(3) Davis, J.; Molnar, E.; Novosselov, I. Nanostructure Transition of Young Soot Aggregates to Mature Soot Aggregates in Diluted Diffusion Flames. *Carbon* **2020**, *159*, 255–265. https://doi.org/10.1016/j.carbon.2019.12.043.

(4) Davis, J.; Tiwari, K.; Novosselov, I. Soot Morphology and Nanostructure in Complex Flame Flow Patterns via Secondary Particle Surface Growth. *Fuel* **2019**, *245*, 447–457. https://doi.org/10.1016/j.fuel.2019.02.058.

(5) Molnar, E. Effects of Temperature and Fuel Dilution on Soot Yields in an Inverted Gravity Flame Reactor. Master of Science in Mechanical Engineering, Univiersity of Washington, Seattle,WA, 2019.

(6) Gilmour, M. I.; McGee, J.; Duvall, R. M.; Dailey, L.; Daniels, M.; Boykin, E.; Cho, S.-H.; Doerfler, D.; Gordon, T.; Devlin, R. B. Comparative Toxicity of Size-Fractionated Airborne Particulate Matter Obtained from Different Cities in the United States. *Inhal. Toxicol.* **2007**, *19* (sup1), 7–16. https://doi.org/10.1080/08958370701490379.

(7) Gilmour, M. I.; O'Connor, S.; Dick, C. A. J.; Miller, C. A.; Linak, W. P. Differential Pulmonary Inflammation and in Vitro Cytotoxicity of Size-Fractionated Fly Ash Particles from Pulverized Coal Combustion. *J. Air Waste Manag. Assoc. 1995* **2004**, *54* (3), 286–295.

(8) Happo, M. S.; Salonen, R. O.; Hälinen, A. I.; Jalava, P. I.; Pennanen, A. S.; Kosma, V. M.; Sillanpää, M.; Hillamo, R.; Brunekreef, B.; Katsouyanni, K.; Sunyer, J.; Hirvonen, M.-R. Dose and Time Dependency of Inflammatory Responses in the Mouse Lung to Urban Air Coarse, Fine, and Ultrafine Particles From Six European Cities. *Inhal. Toxicol.* **2007**, *19* (3), 227–246. https://doi.org/10.1080/08958370601067897.

(9) Dick, C. A. J.; Singh, P.; Daniels, M.; Evansky, P.; Becker, S.; Gilmour, M. I. Murine Pulmonary Inflammatory Responses Following Instillation of Size-Fractionated Ambient Particulate Matter. *J. Toxicol. Environ. Health A* **2003**, *66* (23), 2193–2207. https://doi.org/10.1080/716100636.

(10) Seagrave, J.; McDonald, J. D.; Reed, M. D.; Seilkop, S. K.; Mauderly, J. L. Responses to Subchronic Inhalation of Low Concentrations of Diesel Exhaust and Hardwood Smoke Measured in Rat Bronchoalveolar Lavage Fluid. *Inhal. Toxicol.* **2005**, *17* (12), 657–670. https://doi.org/10.1080/08958370500189529.

(11) Singh, P.; DeMarini, D. M.; Dick, C. A. J.; Tabor, D. G.; Ryan, J. V.; Linak, W. P.; Kobayashi, T.; Gilmour, M. I. Sample Characterization of Automobile and Forklift Diesel Exhaust Particles and Comparative Pulmonary Toxicity in Mice. *Environ. Health Perspect.* **2004**, *112* (8), 820–825. http://dx.doi.org/10.1289/ehp.6579.

(12) Becker, S.; Dailey, L. A.; Soukup, J. M.; Grambow, S. C.; al, et. Seasonal Variations in Air Pollution Particle-Induced Inflammatory Mediator Release and Oxidative Stress. *Environ. Health Perspect. Res. Triangle Park* **2005**, *113* (8), 1032–1038.

(13) Dominici, F.; Peng, R. D.; Bell, M. L.; Pham, L.; McDermott, A.; Zeger, S. L.; Samet, J. M. Fine Particulate Air Pollution and Hospital Admission for Cardiovascular and

Respiratory Diseases. *JAMA* **2006**, *295* (10), 1127–1134. https://doi.org/10.1001/jama.295.10.1127.

(14) Pope C. Arden; Burnett Richard T.; Thurston George D.; Thun Michael J.; Calle Eugenia E.; Krewski Daniel; Godleski John J. Cardiovascular Mortality and Long-Term Exposure to Particulate Air Pollution. *Circulation* **2004**, *109* (1), 71–77. https://doi.org/10.1161/01.CIR.0000108927.80044.7F.

(15) WHO | Mortality and burden of disease from ambient air pollution https://www.who.int/gho/phe/outdoor_air_pollution/burden_text/en/ (accessed Mar 10, 2019).

(16) Karagulian, F.; Belis, C. A.; Dora, C. F. C.; Prüss-Ustün, A. M.; Bonjour, S.; Adair-Rohani, H.; Amann, M. Contributions to Cities' Ambient Particulate Matter (PM): A Systematic Review of Local Source Contributions at Global Level. *Atmos. Environ.* **2015**, *120*, 475–483. https://doi.org/10.1016/j.atmosenv.2015.08.087.

(17) *Institute for Health Metrics and Evaluation (IHME). GBD Compare Data Visualization. Seattle, WA: IHME, University of Washington, 2018. Available from Http://Vizhub.Healthdata.Org/Gbd-Compare. (Accessed 2/28/2019)*.

(18) Yang, Y.; Ruan, Z.; Wang, X.; Yang, Y.; Mason, T. G.; Lin, H.; Tian, L. Short-Term and Long-Term Exposures to Fine Particulate Matter Constituents and Health: A Systematic Review and Meta-Analysis. *Environ. Pollut.* **2019**, *247*, 874–882. https://doi.org/10.1016/j.envpol.2018.12.060.

(19) Richter, H.; Howard, J. B. Formation of Polycyclic Aromatic Hydrocarbons and Their Growth to Soot—a Review of Chemical Reaction Pathways. *Prog. Energy Combust. Sci.* **2000**, *26* (4), 565–608. https://doi.org/10.1016/S0360-1285(00)00009-5.

(20) McEnally, C. S.; Pfefferle, L. D.; Atakan, B.; Kohse-Höinghaus, K. Studies of Aromatic Hydrocarbon Formation Mechanisms in Flames: Progress towards Closing the Fuel Gap. *Prog. Energy Combust. Sci.* **2006**, *32* (3), 247–294. https://doi.org/10.1016/j.pecs.2005.11.003.

(21) Frenklach, M. Reaction Mechanism of Soot Formation in Flames. *Phys. Chem. Chem. Phys.* **2002**, *4* (11), 2028–2037. https://doi.org/10.1039/B110045A.

(22) Bhatt, J. S.; Lindstedt, R. P. Analysis of the Impact of Agglomeration and Surface Chemistry Models on Soot Formation and Oxidation. *Proc. Combust. Inst.* **2009**, *32* (1), 713–720. https://doi.org/10.1016/j.proci.2008.06.201.

(23) Nakamura, H.; Tanimoto, R.; Tezuka, T.; Hasegawa, S.; Maruta, K. Soot Formation Characteristics and PAH Formation Process in a Micro Flow Reactor with a Controlled Temperature Profile. *Combust. Flame* **2014**, *161* (2), 582–591. https://doi.org/10.1016/j.combustflame.2013.09.004.

(24) Elvati, P.; Violi, A. Thermodynamics of Poly-Aromatic Hydrocarbon Clustering and the Effects of Substituted Aliphatic Chains. *Proc. Combust. Inst.* **2013**, *34* (1), 1837–1843. https://doi.org/10.1016/j.proci.2012.07.030.

(25) Totton, T. S.; Misquitta, A. J.; Kraft, M. A Quantitative Study of the Clustering of Polycyclic Aromatic Hydrocarbons at High Temperatures. *Phys. Chem. Chem. Phys.* **2012**, *14* (12), 4081–4094. https://doi.org/10.1039/C2CP23008A.

(26) Chen, D.; Akroyd, J.; Mosbach, S.; Opalka, D.; Kraft, M. Solid–Liquid Transitions in Homogenous Ovalene, Hexabenzocoronene and Circumcoronene Clusters: A Molecular Dynamics Study. *Combust. Flame* **2015**, *162* (2), 486–495. https://doi.org/10.1016/j.combustflame.2014.07.025.

(27) Botero, M. L.; Adkins, E. M.; González-Calera, S.; Miller, H.; Kraft, M. PAH Structure Analysis of Soot in a Non-Premixed Flame Using High-Resolution Transmission Electron Microscopy and Optical Band Gap Analysis. *Combust. Flame* **2016**, *164*, 250–258. https://doi.org/10.1016/j.combustflame.2015.11.022.
(28) Wang, H. Formation of Nascent Soot and Other Condensed-Phase Materials in Flames. *Proc. Combust. Inst.* **2011**, *33* (1), 41–67. https://doi.org/10.1016/j.proci.2010.09.009.
(29) Happold, J.; Grotheer, H.-H.; Aigner, M. Distinction of Gaseous Soot Precursor Molecules and Soot Precursor Particles through Photoionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21* (7), 1247–1254. https://doi.org/10.1002/rcm.2955.
(30) Thomson, M.; Mitra, T. A Radical Approach to Soot Formation. *Science* **2018**, *361* (6406), 978–979. https://doi.org/10.1126/science.aau5941.
(31) Baird, W. M.; Hooven, L. A.; Mahadevan, B. Carcinogenic Polycyclic Aromatic Hydrocarbon-DNA Adducts and Mechanism of Action. *Environ. Mol. Mutagen.* **2005**, *45* (2–3), 106–114. https://doi.org/10.1002/em.20095.
(32) Kim Oanh, N. T.; Nghiem, L. H.; Phyu, Y. L. Emission of Polycyclic Aromatic Hydrocarbons, Toxicity, and Mutagenicity from Domestic Cooking Using Sawdust Briquettes, Wood, and Kerosene. *Environ. Sci. Technol.* **2002**, *36* (5), 833–839. https://doi.org/10.1021/es011060n.
(33) de Kok, T. M. C. M.; Driece, H. A. L.; Hogervorst, J. G. F.; Briedé, J. J. Toxicological Assessment of Ambient and Traffic-Related Particulate Matter: A Review of Recent Studies. *Mutat. Res. Mutat. Res.* **2006**, *613* (2), 103–122. https://doi.org/10.1016/j.mrrev.2006.07.001.
(34) Becker, S.; Dailey, L. A.; Soukup, J. M.; Grambow, S. C.; Devlin, R. B.; Huang, Y.-C. T. Seasonal Variations in Air Pollution Particle-Induced Inflammatory Mediator Release and Oxidative Stress. *Environ. Health Perspect.* **2005**, *113* (8), 1032.
(35) Kim, K.-H.; Jahan, S. A.; Kabir, E.; Brown, R. J. C. A Review of Airborne Polycyclic Aromatic Hydrocarbons (PAHs) and Their Human Health Effects. *Environ. Int.* **2013**, *60*, 71–80. https://doi.org/10.1016/j.envint.2013.07.019.
(36) US EPA, O. Text Version of the Indoor Air Quality House Tour http://www.epa.gov/indoor-air-quality-iaq/text-version-indoor-air-quality-house-tour (accessed Mar 30, 2020).
(37) Dobbins, R. A. Soot Inception Temperature and the Carbonization Rate of Precursor Particles. *Combust. Flame* **2002**, *130* (3), 204–214. https://doi.org/10.1016/S0010-2180(02)00374-7.
(38) Aubagnac-Karkar, D.; El Bakali, A.; Desgroux, P. Soot Particles Inception and PAH Condensation Modelling Applied in a Soot Model Utilizing a Sectional Method. *Combust. Flame* **2018**, *189*, 190–206. https://doi.org/10.1016/j.combustflame.2017.10.027.
(39) *Office of the Federal Registration (OFR) Appendix A: Priority Pollutants. Fed Reg. 1982;47:52309.*; https://www.epa.gov/sites/production/files/2015-09/documents/priority-pollutant-list-epa.pdf.
(40) Samburova, V.; Zielinska, B.; Khlystov, A. Do 16 Polycyclic Aromatic Hydrocarbons Represent PAH Air Toxicity? *Toxics* **2017**, *5* (3), 17. https://doi.org/10.3390/toxics5030017.
(41) *Some Non-Heterocyclic Polycyclic Aromatic Hydrocarbons and Some Related Occupational Exposures*; IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, International Agency for Research on Cancer, Eds.; IARC monographs on the

evaluation of carcinogenic risks to humans; IARC Press ; Distributed by World Health Organization: Lyon, France : Geneva, 2010.

(42) National Toxicology Program: 14th Report on Carcinogens https://ntp.niehs.nih.gov/go/roc14 (accessed Mar 12, 2019).

(43) Kim, K.-H.; Jahan, S. A.; Kabir, E.; Brown, R. J. C. A Review of Airborne Polycyclic Aromatic Hydrocarbons (PAHs) and Their Human Health Effects. *Environ. Int.* **2013**, *60*, 71–80. https://doi.org/10.1016/j.envint.2013.07.019.

(44) West, C.; Hettiyadura, A.; Mahamuni, G.; Davis, J.; Novosselov, I.; Laskin, A. *Molecular Composition of Brown Carbon Aerosol Generated in the Diluted Ethane Flame*; 2019.

(45) Lin, P.; Fleming, L. T.; Nizkorodov, S. A.; Laskin, J.; Laskin, A. Comprehensive Molecular Characterization of Atmospheric Brown Carbon by High Resolution Mass Spectrometry with Electrospray and Atmospheric Pressure Photoionization. *Anal. Chem.* **2018**, *90* (21), 12493–12502. https://doi.org/10.1021/acs.analchem.8b02177.

(46) Wei, Y.; Han, I.-K.; Hu, M.; Shao, M.; Zhang, J.; Tang, X. Personal Exposure to Particulate PAHs and Anthraquinone and Oxidative DNA Damages in Humans. *Chemosphere* **2010**, *81* (10), 1280–1285. https://doi.org/10.1016/j.chemosphere.2010.08.055.

(47) Bonetta, Sa.; Gianotti, V.; Bonetta, Si.; Gosetti, F.; Oddone, M.; Gennaro, M. C.; Carraro, E. DNA Damage in A549 Cells Exposed to Different Extracts of PM2.5 from Industrial, Urban and Highway Sites. *Chemosphere* **2009**, *77* (7), 1030–1034. https://doi.org/10.1016/j.chemosphere.2009.07.076.

(48) Wei, Y.; Han, I.-K.; Shao, M.; Hu, M.; Zhang, J. (Jim); Tang, X. PM2.5 Constituents and Oxidative DNA Damage in Humans. *Environ. Sci. Technol.* **2009**, *43* (13), 4757–4762. https://doi.org/10.1021/es803337c.

(49) Buzcu-Guven, B.; Brown, S. G.; Frankel, A.; Hafner, H. R.; Roberts, P. T. Analysis and Apportionment of Organic Carbon and Fine Paniculate Matter Sources at Multiple Sites in the Midwestern United States. *J. Air Waste Manag. Assoc.* **2007**, *57* (5).

(50) Bergauff, M.; Ward, T.; Noonan, C.; Palmer, C. P. Determination and Evaluation of Selected Organic Chemical Tracers for Wood Smoke in Airborne Particulate Matter. *Int. J. Environ. Anal. Chem.* **2008**, *88* (7), 473–486. https://doi.org/10.1080/03067310701809110.

(51) Lin, P.; Fleming, L. T.; Nizkorodov, S. A.; Laskin, J.; Laskin, A. Comprehensive Molecular Characterization of Atmospheric Brown Carbon by High Resolution Mass Spectrometry with Electrospray and Atmospheric Pressure Photoionization. *Anal. Chem.* **2018**, *90* (21), 12493–12502. https://doi.org/10.1021/acs.analchem.8b02177.

(52) Samburova, V.; Connolly, J.; Gyawali, M.; Yatavelli, R. L. N.; Watts, A. C.; Chakrabarty, R. K.; Zielinska, B.; Moosmüller, H.; Khlystov, A. Polycyclic Aromatic Hydrocarbons in Biomass-Burning Emissions and Their Contribution to Light Absorption and Aerosol Toxicity. *Sci. Total Environ.* **2016**, *568*, 391–401. https://doi.org/10.1016/j.scitotenv.2016.06.026.

(53) Diwakar, P. K.; Loper, K. H.; Matiaske, A.-M.; Hahn, D. W. Laser-Induced Breakdown Spectroscopy for Analysis of Micro and Nanoparticles. *J. Anal. At. Spectrom.* **2012**, *27* (7), 1110–1119. https://doi.org/10.1039/C2JA30012E.

(54) Sivaprakasam, V.; Lin, H.-B.; Huston, A. L.; Eversole, J. D. Spectral Characterization of Biological Aerosol Particles Using Two-Wavelength Excited Laser-Induced Fluorescence

and Elastic Scattering Measurements. *Opt. Express* **2011**, *19* (7), 6191–6208. https://doi.org/10.1364/OE.19.006191.

(55) Ji, Z.; Dai, R.; Zhang, Z. Characterization of Fine Particulate Matter in Ambient Air by Combining TEM and Multiple Spectroscopic Techniques – NMR, FTIR and Raman Spectroscopy. *Environ. Sci. Process. Impacts* **2015**, *17* (3), 552–560. https://doi.org/10.1039/C4EM00678J.

(56) Sheesley, R. J.; Mieritz, M.; DeMinter, J. T.; Shelton, B. R.; Schauer, J. J. Development of an in Situ Derivatization Technique for Rapid Analysis of Levoglucosan and Polar Compounds in Atmospheric Organic Aerosol. *Atmos. Environ.* **2015**, *123*, 251–255. https://doi.org/10.1016/j.atmosenv.2015.10.047.

(57) Yu, X.; Song, W.; Yu, Q.; Li, S.; Zhu, M.; Zhang, Y.; Deng, W.; Yang, W.; Huang, Z.; Bi, X.; Wang, X. Fast Screening Compositions of PM2.5 by ATR-FTIR: Comparison with Results from IC AndOC/EC Analyzers. *J. Environ. Sci.* **2018**, *71*, 76–88. https://doi.org/10.1016/j.jes.2017.11.021.

(58) Dippel, B.; Heintzenberg, J. Soot Characterization in Atmospheric Particles from Different Sources by NIR FT Raman Spectroscopy. *J. Aerosol Sci.* **1999**, *30*, S907–S908. https://doi.org/10.1016/S0021-8502(99)80464-9.

(59) Huffman, J. A. Online Techniques for Quantification and Characterization of Biological Aerosols. In *Microbiology of Aerosols*.

(60) Sengupta, A.; Brar, N.; Davis, E. J. Bioaerosol Detection and Characterization by Surface-Enhanced Raman Spectroscopy. *J. Colloid Interface Sci.* **2007**, *309* (1), 36–43. https://doi.org/10.1016/j.jcis.2007.02.015.

(61) Pan, Y.-L. Detection and Characterization of Biological and Other Organic-Carbon Aerosol Particles in Atmosphere Using Fluorescence. *J. Quant. Spectrosc. Radiat. Transf.* **2015**, *150*, 12–35. https://doi.org/10.1016/j.jqsrt.2014.06.007.

(62) Berlman, I. *Handbook of Florescence Spectra of Aromatic Molecules*; Elsevier, 2012.

(63) Robertson, J.; O'Reilly, E. P. Electronic and Atomic Structure of Amorphous Carbon. *Phys. Rev. B* **1987**, *35* (6), 2946–2957. https://doi.org/10.1103/PhysRevB.35.2946.

(64) Menon, A.; Dreyer, J. A. H.; Martin, J. W.; Akroyd, J.; Robertson, J.; Kraft, M. Optical Band Gap of Cross-Linked, Curved, and Radical Polyaromatic Hydrocarbons. *Phys. Chem. Chem. Phys.* **2019**, *21* (29), 16240–16251. https://doi.org/10.1039/C9CP02363A.

(65) Martin, J. W.; Bowal, K.; Menon, A.; Slavchov, R. I.; Akroyd, J.; Mosbach, S.; Kraft, M. Polar Curved Polycyclic Aromatic Hydrocarbons in Soot Formation. *Proc. Combust. Inst.* **2019**, *37* (1), 1117–1123. https://doi.org/10.1016/j.proci.2018.05.046.

(66) Birks, J. B. *Photophysics of Aromatic Molecules*; WILEY - INTERSCIENCE a division of John Wiley & Sons Ltd, 1970.

(67) Rieger, R.; Müllen, K. Forever Young: Polycyclic Aromatic Hydrocarbons as Model Cases for Structural and Optical Studies. *J. Phys. Org. Chem.* **2010**, *23* (4), 315–325. https://doi.org/10.1002/poc.1644.

(68) Moerner, W. E.; Fromm, D. P. Methods of Single-Molecule Fluorescence Spectroscopy and Microscopy. *Rev. Sci. Instrum.* **2003**, *74* (8), 3597–3619. https://doi.org/10.1063/1.1589587.

(69) Elcoroaristizabal, S.; de Juan, A.; García, J. A.; Durana, N.; Alonso, L. Comparison of Second-Order Multivariate Methods for Screening and Determination of PAHs by Total Fluorescence Spectroscopy. *Chemom. Intell. Lab. Syst.* **2014**, *132*, 63–74. https://doi.org/10.1016/j.chemolab.2014.01.005.

(70) L. Nahorniak, M.; S. Booksh, K. Excitation-Emission Matrix Fluorescence Spectroscopy in Conjunction with Multiway Analysis for PAH Detection in Complex Matrices. *Analyst* **2006**, *131* (12), 1308–1315. https://doi.org/10.1039/B609875D.

(71) Rutherford, J. W.; Dawson-Elli, N.; Manicone, A. M.; Korshin, G. V.; Novosselov, I. V.; Seto, E.; Posner, J. D. Excitation Emission Matrix Fluorescence Spectroscopy for Combustion Generated Particulate Matter Source Identification. *Atmos. Environ.* **2020**, *220*, 117065. https://doi.org/10.1016/j.atmosenv.2019.117065.

(72) Aizawa, T.; Kosaka, H. Investigation of Early Soot Formation Process in a Diesel Spray Flame via Excitation—Emission Matrix Using a Multi-Wavelength Laser Source. *Int. J. Engine Res.* **2008**, *9* (1), 79–97. https://doi.org/10.1243/14680874JER01407.

(73) JiJi, R. D.; Cooper, G. A.; Booksh, K. S. Excitation-Emission Matrix Fluorescence Based Determination of Carbamate Pesticides and Polycyclic Aromatic Hydrocarbons. *Anal. Chim. Acta* **1999**, *397* (1), 61–72. https://doi.org/10.1016/S0003-2670(99)00392-X.

(74) Muroski, A. R.; Booksh, K. S.; Myrick, M. L. Single-Measurement Excitation/Emission Matrix Spectrofluorometer for Determination of Hydrocarbons in Ocean Water. 1. Instrumentation and Background Correction. *Anal. Chem.* **1996**, *68* (20), 3534–3538. https://doi.org/10.1021/ac960252b.

(75) Booksh, K. S.; Muroski, A. R.; Myrick, M. L. Single-Measurement Excitation/Emission Matrix Spectrofluorometer for Determination of Hydrocarbons in Ocean Water. 2. Calibration and Quantitation of Naphthalene and Styrene. *Anal. Chem.* **1996**, *68* (20), 3539–3544. https://doi.org/10.1021/ac9602534.

(76) Beltrán, J. L.; Ferrer, R.; Guiteras, J. Multivariate Calibration of Polycyclic Aromatic Hydrocarbon Mixtures from Excitation–Emission Fluorescence Spectra. *Anal. Chim. Acta* **1998**, *373* (2), 311–319. https://doi.org/10.1016/S0003-2670(98)00420-6.

(77) Chen, H.; Liao, Z.; Gu, X.; Xie, J.; Li, H.; Zhang, J. Anthropogenic Influences of Paved Runoff and Sanitary Sewage on the Dissolved Organic Matter Quality of Wet Weather Overflows: An Excitation–Emission Matrix Parallel Factor Analysis Assessment. *Environ. Sci. Technol.* **2017**, *51* (3), 1157–1167. https://doi.org/10.1021/acs.est.6b03727.

(78) Osburn, C. L.; Handsel, L. T.; Peierls, B. L.; Paerl, H. W. Predicting Sources of Dissolved Organic Nitrogen to an Estuary from an Agro-Urban Coastal Watershed. *Environ. Sci. Technol.* **2016**, *50* (16), 8473–8484. https://doi.org/10.1021/acs.est.6b00053.

(79) Harrison, R. M.; Smith, D. J. T.; Luhana, L. Source Apportionment of Atmospheric Polycyclic Aromatic Hydrocarbons Collected from an Urban Location in Birmingham, U.K. *Environ. Sci. Technol.* **1996**, *30* (3), 825–832. https://doi.org/10.1021/es950252d.

(80) Sofowote, U. M.; McCarry, B. E.; Marvin, C. H. Source Apportionment of PAH in Hamilton Harbour Suspended Sediments: Comparison of Two Factor Analysis Methods. *Environ. Sci. Technol.* **2008**, *42* (16), 6007–6014. https://doi.org/10.1021/es800219z.

(81) Cai, J.; Yan, B. Z.; Ross, J.; Zhang, D. N.; Kinney, P. L.; Perzanowski, M. S.; Jung, K.; Miller, R.; Chillrud, S. N. Validation of MicroAeth (R) as a Black Carbon Monitor for Fixed-Site Measurement and Optimization for Personal Exposure Characterization. *Aerosol and Air Quality Research*, 2014, *14*, 1–9. https://doi.org/10.4209/aaqr.2013.03.0088.

(82) Chartier, R.; Phillips, M.; Mosquin, P.; Elledge, M.; Bronstein, K.; Nandasena, S.; Thornburg, V.; Thornburg, J.; Rodes, C. A Comparative Study of Human Exposures to Household Air Pollution from Commonly Used Cookstoves in Sri Lanka. *Indoor Air*, 2017, *27*, 147–159. https://doi.org/10.1111/ina.12281.

(83) Koehler, K. A.; Peters, T. M. New Methods for Personal Exposure Monitoring for Airborne Particles. *Curr Environ Health Rep*, 2015, *2*, 399–411. https://doi.org/10.1007/s40572-015-0070-z.

(84) Nakayama, T.; Matsumi, Y.; Kawahito, K.; Watabe, Y. Development and Evaluation of a Palm-Sized Optical PM2.5 Sensor. *Aerosol Science and Technology*, 2018, *52*, 2–12. https://doi.org/10.1080/02786826.2017.1375078.

(85) Volckens, J.; Quinn, C.; Leith, D.; Mehaffy, J.; Henry, C. S.; Miller-Lionberg, D. Development and Evaluation of an Ultrasonic Personal Aerosol Sampler. *Indoor Air*, 2017, *27*, 409–416. https://doi.org/10.1111/ina.12318.

(86) Foat, T. G.; Sellors, W. J.; Walker, M. D.; Rachwal, P. A.; Jones, J. W.; Despeyroux, D. D.; Coudron, L.; Munro, I.; McCluskey, D. K.; Tan, C. K. L.; Tracey, M. C. A Prototype Personal Aerosol Sampler Based on Electrostatic Precipitation and Electrowetting-on-Dielectric Actuation of Droplets. *J. Aerosol Sci.* **2016**, *95*, 43–53. https://doi.org/10.1016/j.jaerosci.2016.01.007.

(87) Han, T. T.; Thomas, N. M.; Mainelis, G. Design and Development of a Self-Contained Personal Electrostatic Bioaerosol Sampler (PEBS) with a Wire-to-Wire Charger. *Aerosol Sci. Technol.* **2017**, *51* (8), 903–915. https://doi.org/10.1080/02786826.2017.1329516.

(88) Ariessohn, P. C.; Novosselov, I. V. Aerosol Collection and Microdroplet Delivery for Analysis, 2010.

(89) Roux, J. M.; Sarda-Esteve, R.; Delapierre, G.; Nadal, M. H.; Bossuet, C.; Olmedo, L. Development of a New Portable Air Sampler Based on Electrostatic Precipitation. *Environmental Science and Pollution Research*, 2016, *23*, 8175–8183. https://doi.org/10.1007/s11356-015-5522-3.

(90) Austin, E.; Novosselov, I.; Seto, E.; Yost, M. G. Laboratory Evaluation of the Shinyei PPD42NS Low-Cost Particulate Matter Sensor (Vol 10, E0137789, 2015). *Plos One*, 2015, *10*, 2. https://doi.org/10.1371/journal.pone.0141928.

(91) Dacunto, P. J.; Klepeis, N. E.; Cheng, K.-C.; Acevedo-Bolton, V.; Jiang, R.-T.; Repace, J. L.; Ott, W. R.; Hildemann, L. M. Determining PM2.5 Calibration Curves for a Low-Cost Particle Monitor: Common Indoor Residential Aerosols. *Environmental Science: Processes & Impacts*, 2015, *17*, 1959–1966. https://doi.org/10.1039/C5EM00365B.

(92) Dons, E.; Laeremans, M.; Orjuela, J. P.; Avila-Palencia, I.; Carrasco-Turigas, G.; Cole-Hunter, T.; Anaya-Boig, E.; Standaert, A.; De Boever, P.; Nawrot, T.; Gotschi, T.; de Nazelle, A.; Nieuwenhuijsen, M.; Panis, L. I. Wearable Sensors for Personal Monitoring and Estimation of Inhaled Traffic-Related Air Pollution: Evaluation of Methods. *Environmental Science & Technology*, 2017, *51*, 1859–1867. https://doi.org/10.1021/acs.est.6b05782.

(93) Njalsson, T.; Novosselov, I. Design and Optimization of a Compact Low-Cost Optical Particle Sizer. *Journal of Aerosol Science*, 2018, *119*, 1–12. https://doi.org/10.1016/j.jaerosci.2018.01.003.

(94) Gao, M. L.; Cao, J. J.; Seto, E. A Distributed Network of Low-Cost Continuous Reading Sensors to Measure Spatiotemporal Variations of PM2.5 in Xi'an, China. *Environmental Pollution*, 2015, *199*, 56–65. https://doi.org/10.1016/j.envpol.2015.01.013.

(95) Johnson, K. K.; Bergin, M. H.; Russell, A. G.; Hagler, G. S. W. Using Low Cost Sensors to Measure Ambient Particulate Matter Concentrations and On-Road Emissions Factors. *Atmospheric Measurement Techniques*, 2016, 1–22. https://doi.org/10.5194/amt-2015-331.

(96) Seto, E.; Austin, E.; Novosselov, I.; Yost, M. G. Use of Low-Cost Particle Monitors to Calibrate Traffic-Related Air Pollutant Models in Urban Areas; 2014.
(97) Ariessohn, P. C.; Novosselov, I. V. Skimmer for Concentrating an Aerosol, 2011.
(98) Marple, V. A.; Chien, C. M. Virtual Impactors: A Theoretical Study. *Environmental Science and Technology*, 1980. https://doi.org/10.1021/es60168a019.
(99) Novosselov, I. V.; Ariessohn, P. C. Rectangular Slit Atmospheric Pressure Aerodynamic Lens Aerosol Concentrator. *Aerosol Science and Technology*, 2014, *48*, 163–172. https://doi.org/10.1080/02786826.2013.865832.
(100) He, J.; Beck, N. K.; Kossik, A. L.; Zhang, J.; Seto, E.; Meschke, J. S.; Novosselov, I. Evaluation of Micro-Well Collector for Capture and Analysis of Aerosolized Bacillus Subtilis Spores. *PloS One* **2018**, *13* (5), e0197783.
(101) Marple, V. A.; Rubow, K. L.; Behm, S. M. A Microorifice Uniform Deposit Impactor (MOUDI): Description, Calibration, and Use. *Aerosol Science and Technology*, 1991, *14*, 434–446. https://doi.org/10.1080/02786829108959504.
(102) Marple, V. A.; Rubow, K. L.; Olson, B. A. Diesel Exhaust/Mine Dust Virtual Impactor Personal Aerosol Sampler: Design, Calibration and Field Evaluation. *Aerosol Science and Technology*, 1995, *22*, 140–150. https://doi.org/10.1080/02786829408959735.
(103) Novosselov, I. V.; Gorder, R. A.; Van Amberg, J. A.; Ariessohn, P. C. Design and Performance of a Low-Cost Micro-Channel Aerosol Collector. *Aerosol Sci. Technol.* **2014**, *48* (8), 822–830. https://doi.org/10.1080/02786826.2014.932895.
(104) Morrow, P. E.; Mercer, T. T. A Point-to-Plane Electrostatic Precipitator for Particle Size Sampling. *American Industrial Hygiene Association Journal, Vol. 25, No. 1, January-February 1964: pp. 8–14*, 2007. https://doi.org/10.1080/00028896409342547.
(105) Trichel, G. W. The Mechanism of the Negative Point to Plane Corona near Onset. *Physical Review*, 1938, *54*, 1078–1084. https://doi.org/10.1103/PhysRev.54.1078.
(106) Bedmutha, R. J.; Ferrante, L.; Briens, C.; Berruti, F.; Inculet, I. Single and Two-Stage Electrostatic Demisters for Biomass Pyrolysis Application. *Chem. Eng. Process. Process Intensif.* **2009**, *48* (6), 1112–1120. https://doi.org/10.1016/j.cep.2009.02.007.
(107) Lee, M. H.; Kim, J. H.; Biswas, P.; Kim, S. S.; Suh, Y. J.; Jang, H. D.; Bhang, S. H.; Yu, T.; Cho, K. Enhanced Collection Efficiency of Nanoparticles by Electrostatic Precipitator with Needle-Cylinder Configuration. *Journal of Nanoscience and Nanotechnology*, 2016, *16*, 6884–6888. https://doi.org/10.1166/jnn.2016.11322.
(108) Zheng, L.; Kulkarni, P.; Zavvos, K.; Liang, H.; Birch, M. E.; Dionysiou, D. D. Characterization of an Aerosol Microconcentrator for Analysis Using Microscale Optical Spectroscopies. *J. Aerosol Sci.* **2017**, *104*, 66–78. https://doi.org/10.1016/j.jaerosci.2016.11.007.
(109) Han, T.; An, H. R.; Mainelis, G. Performance of an Electrostatic Precipitator with Superhydrophobic Surface When Collecting Airborne Bacteria. *Aerosol Sci. Technol.* **2010**, *44* (5), 339–348. https://doi.org/10.1080/02786821003649352.
(110) Han, T.; Nazarenko, Y.; Lioy, P. J.; Mainelis, G. Collection Efficiencies of an Electrostatic Sampler with Superhydrophobic Surface for Fungal Bioaerosols. *Indoor Air* **2011**, *21* (2), 110–120. https://doi.org/10.1111/j.1600-0668.2010.00685.x.
(111) Ma, Z.; Zheng, Y.; Cheng, Y.; Xie, S.; Ye, X.; Yao, M. Development of an Integrated Microfluidic Electrostatic Sampler for Bioaerosol. *J. Aerosol Sci.* **2016**, *95*, 84–94. https://doi.org/10.1016/j.jaerosci.2016.01.003.

(112) Pardon, G.; Ladhani, L.; Sandström, N.; Ettori, M.; Lobov, G.; van der Wijngaart, W. Aerosol Sampling Using an Electrostatic Precipitator Integrated with a Microfluidic Interface. *Sens. Actuators B Chem.* **2015**, *212*, 344–352. https://doi.org/10.1016/j.snb.2015.02.008.

(113) Ault, A. P.; Axson, J. L. Atmospheric Aerosol Chemistry: Spectroscopic and Microscopic Advances. *Anal. Chem.* **2017**, *89* (1), 430–452. https://doi.org/10.1021/acs.analchem.6b04670.

(114) Gaurav, M.; Jay, R.; Justin, D.; Eric, M.; Jonathan D., P.; Edmund, S.; Gregory V., K.; Igor V., N. *Excitation-Emission Matrix Spectroscopy for Analysis of Chemical Composition of Combustion Generated Particulate Matter*; 2019. https://doi.org/10.26434/chemrxiv.9702242.v1.

(115) Hawthorne, S. B.; Grabanski, C. B.; Miller, D. J.; Kreitinger, J. P. Solid-Phase Microextraction Measurement of Parent and Alkyl Polycyclic Aromatic Hydrocarbons in Milliliter Sediment Pore Water Samples and Determination of KDOC Values. *Environ. Sci. Technol.* **2005**, *39* (8), 2795–2803. https://doi.org/10.1021/es0405171.

(116) Hawthorne, S. B.; St Germain, R. W.; Azzolina, N. A. Laser-Induced Fluorescence Coupled with Solid-Phase Microextraction for In Situ Determination of PAHs in Sediment Pore Water. *Environ. Sci. Technol.* **2008**, *42* (21), 8021–8026. https://doi.org/10.1021/es8011673.

(117) Mahamuni, G.; Rutherford, J.; Davis, J.; Molnar, E.; Posner, J. D.; Seto, E.; Korshin, G. V.; Novosselov, I. V. Excitation-Emission Matrix Spectroscopy for Analysis of Chemical Composition of Combustion Generated Particulate Matter. *Environ. Sci. Technol.* **2020**. https://doi.org/10.1021/acs.est.0c01110.

(118) Hwang, J. Y.; Lee, W.; Kang, H. G.; Chung, S. H. Synergistic Effect of Ethylene–Propane Mixture on Soot Formation in Laminar Diffusion Flames. *Combust. Flame* **1998**, *114* (3), 370–380. https://doi.org/10.1016/S0010-2180(97)00295-2.

(119) Bladh, H.; Johnsson, J.; Olofsson, N.-E.; Bohlin, A.; Bengtsson, P.-E. Optical Soot Characterization Using Two-Color Laser-Induced Incandescence (2C-LII) in the Soot Growth Region of a Premixed Flat Flame. *Proc. Combust. Inst.* **2011**, *33* (1), 641–648. https://doi.org/10.1016/j.proci.2010.06.166.

(120) Violi, A.; D’Anna, A.; D’Alessio, A.; Sarofim, A. F. Modeling Aerosol Formation in Opposed-Flow Diffusion Flames. *Chemosphere* **2003**, *51* (10), 1047–1054. https://doi.org/10.1016/S0045-6535(03)00125-5.

(121) Cuoci, A.; Frassoldati, A.; Faravelli, T.; Jin, H.; Wang, Y.; Zhang, K.; Glarborg, P.; Qi, F. Experimental and Detailed Kinetic Modeling Study of PAH Formation in Laminar Co-Flow Methane Diffusion Flames. *Proc. Combust. Inst.* **2013**, *34* (1), 1811–1818. https://doi.org/10.1016/j.proci.2012.05.085.

(122) Botero, M. L.; Chen, D.; González-Calera, S.; Jefferson, D.; Kraft, M. HRTEM Evaluation of Soot Particles Produced by the Non-Premixed Combustion of Liquid Fuels. *Carbon* **2016**, *96*, 459–473. https://doi.org/10.1016/j.carbon.2015.09.077.

(123) Stipe, C. B.; Higgins, B. S.; Lucas, D.; Koshland, C. P.; Sawyer, R. F. Inverted Co-Flow Diffusion Flame for Producing Soot. *Rev. Sci. Instrum.* **2005**, *76* (2), 023908. https://doi.org/10.1063/1.1851492.

(124) Kazemimanesh, M.; Moallemi, A.; Thomson, K.; Smallwood, G.; Lobo, P.; Olfert, J. S. A Novel Miniature Inverted-Flame Burner for the Generation of Soot Nanoparticles. *Aerosol Sci. Technol.* **2019**, *53* (2), 184–195. https://doi.org/10.1080/02786826.2018.1556774.

(125) Yan, S.; Jiang, Y.-J.; Marsh, N. D.; Eddings, E. G.; Sarofim, A. F.; Pugmire, R. J. Study of the Evolution of Soot from Various Fuels. *Energy Fuels* **2005**, *19* (5), 1804–1811. https://doi.org/10.1021/ef049742u.

(126) Ghazi, R.; Tjong, H.; Soewono, A.; Rogak, S. N.; Olfert, J. S. Mass, Mobility, Volatility, and Morphology of Soot Particles Generated by a McKenna and Inverted Burner. *Aerosol Sci. Technol.* **2013**, *47* (4), 395–405. https://doi.org/10.1080/02786826.2012.755259.

(127) Chakrabarty, R. K.; Novosselov, I. V.; Beres, N. D.; Moosmüller, H.; Sorensen, C. M.; Stipe, C. B. Trapping and Aerogelation of Nanoparticles in Negative Gravity Hydrocarbon Flames. *Appl. Phys. Lett.* **2014**, *104* (24), 243103. https://doi.org/10.1063/1.4884057.

(128) Liu, P.; Arnold, I.; Wang, Y.; Yu, Y.; Fang, J.; Biswas, P.; K. Chakrabarty, R. Synthesis of Titanium Dioxide Aerosol Gels in a Buoyancy-Opposed Flame Reactor. *Aerosol Sci. Technol.* **2015**, *49*, 00–00. https://doi.org/10.1080/02786826.2015.1112356.

(129) Shaddix, C. R. *Correcting Thermocouple Measurements for Radiation Loss: A Critical Review*; CONF-990805-; American Society of Mechanical Engineers, New York, NY (US), 1999.

(130) Kholghy, M. R.; Afarin, Y.; Sediako, A. D.; Barba, J.; Lapuerta, M.; Chu, C.; Weingarten, J.; Borshanpour, B.; Chernov, V.; Thomson, M. J. Comparison of Multiple Diagnostic Techniques to Study Soot Formation and Morphology in a Diffusion Flame. *Combust. Flame* **2017**, *176*, 567–583. https://doi.org/10.1016/j.combustflame.2016.11.012.

(131) Hindasageri, V.; Vedula, R. P.; Prabhu, S. V. Thermocouple Error Correction for Measuring the Flame Temperature with Determination of Emissivity and Heat Transfer Coefficient. *Rev. Sci. Instrum.* **2013**, *84* (2), 024902. https://doi.org/10.1063/1.4790471.

(132) Kholghy, M. R.; Veshkini, A.; Thomson, M. J. The Core–Shell Internal Nanostructure of Soot – A Criterion to Model Soot Maturity. *Carbon* **2016**, *100*, 508–536. https://doi.org/10.1016/j.carbon.2016.01.022.

(133) Glassman, I. Sooting Laminar Diffusion Flames: Effect of Dilution, Additives, Pressure, and Microgravity. *Symp. Int. Combust.* **1998**, *27* (1), 1589–1596. https://doi.org/10.1016/S0082-0784(98)80568-7.

(134) Jerez, A.; Cruz Villanueva, J. J.; Figueira da Silva, L. F.; Demarco, R.; Fuentes, A. Measurements and Modeling of PAH Soot Precursors in Coflow Ethylene/Air Laminar Diffusion Flames. *Fuel* **2019**, *236*, 452–460. https://doi.org/10.1016/j.fuel.2018.09.047.

(135) Karataş, A. E.; Gülder, Ö. L. Dependence of Sooting Characteristics and Temperature Field of Co-Flow Laminar Pure and Nitrogen-Diluted Ethylene–Air Diffusion Flames on Pressure. *Combust. Flame* **2015**, *162* (4), 1566–1574. https://doi.org/10.1016/j.combustflame.2014.11.020.

(136) Birch, M. E.; Cary, R. A. Elemental Carbon-Based Method for Monitoring Occupational Exposures to Particulate Diesel Exhaust. *Aerosol Sci. Technol.* **1996**, *25* (3), 221–241. https://doi.org/10.1080/02786829608965393.

(137) Glassman, I.; Yetter, R. A.; Glumac, N. G. *Combustion*; Academic press, 2014.

(138) Smooke, M. D.; McEnally, C. S.; Pfefferle, L. D.; Hall, R. J.; Colket, M. B. Computational and Experimental Study of Soot Formation in a Coflow, Laminar Diffusion Flame. *Combust. Flame* **1999**, *117* (1), 117–139. https://doi.org/10.1016/S0010-2180(98)00096-0.

(139) Saffaripour, M.; Veshkini, A.; Kholghy, M.; Thomson, M. J. Experimental Investigation and Detailed Modeling of Soot Aggregate Formation and Size Distribution in Laminar

Coflow Diffusion Flames of Jet A-1, a Synthetic Kerosene, and n-Decane. *Combust. Flame* **2014**, *161* (3), 848–863. https://doi.org/10.1016/j.combustflame.2013.10.016.

(140) Yan, W.; Lou, C. Two-Dimensional Distributions of Temperature and Soot Volume Fraction Inversed from Visible Flame Images. *Exp. Therm. Fluid Sci.* **2013**, *50*, 229–233. https://doi.org/10.1016/j.expthermflusci.2013.05.013.

(141) Liu, F.; Thomson, K. A.; Smallwood, G. J. Soot Temperature and Volume Fraction Retrieval from Spectrally Resolved Flame Emission Measurement in Laminar Axisymmetric Coflow Diffusion Flames: Effect of Self-Absorption. *Combust. Flame* **2013**, *160* (9), 1693–1705. https://doi.org/10.1016/j.combustflame.2013.02.007.

(142) Kholghy, M. R.; Afarin, Y.; Sediako, A. D.; Barba, J.; Lapuerta, M.; Chu, C.; Weingarten, J.; Borshanpour, B.; Chernov, V.; Thomson, M. J. Comparison of Multiple Diagnostic Techniques to Study Soot Formation and Morphology in a Diffusion Flame. *Combust. Flame* **2017**, *176*, 567–583. https://doi.org/10.1016/j.combustflame.2016.11.012.

(143) Liu, F.; Guo, H.; Smallwood, G. J. Effects of Radiation Model on the Modeling of a Laminar Coflow Methane/Air Diffusion Flame. *Combust. Flame* **2004**, *138* (1), 136–154. https://doi.org/10.1016/j.combustflame.2004.04.007.

(144) Sullivan, B.; Allawatt, G.; Emery, A.; Means, P.; Kramlich, J.; Posner, J. Time-Resolved Particulate Emissions Monitoring of Cookstove Biomass Combustion Using a Tapered Element Oscillating Microbalance. *Combust. Sci. Technol.* **2017**, *189* (6), 923–936. https://doi.org/10.1080/00102202.2016.1253564.

(145) Gould, T.; Larson, T.; Stewart, J.; Kaufman, J. D.; Slater, D.; McEwen, N. A Controlled Inhalation Diesel Exhaust Exposure Facility with Dynamic Feedback Control of PM Concentration. *Inhal. Toxicol.* **2008**, *20* (1), 49–52. https://doi.org/10.1080/08958370701758478.

(146) 40 CFR Appendix L to Part 50 - Reference Method for the Determination of Fine Particulate Matter as PM2.5 in the Atmosphere https://www.law.cornell.edu/cfr/text/40/appendix-L_to_part_50 (accessed Apr 1, 2020).

(147) Allen, R.; Box, M.; Liu, L.-J. S.; Larson, T. V. A Cost-Effective Weighing Chamber for Particulate Matter Filters. *J. Air Waste Manag. Assoc.* **2001**, *51* (12), 1650–1653. https://doi.org/10.1080/10473289.2001.10464392.

(148) Birch, M. E. Occupational Monitoring of Particulate Diesel Exhaust by NIOSH Method 5040. *Appl. Occup. Environ. Hyg.* **2002**, *17* (6), 400–405. https://doi.org/10.1080/10473220290035390.

(149) Cassinelli, M. E.; O`Connor, P. F. *NIOSH Manual of Analytical Methods (NMAM{trademark}). Fourth Edition, Second Supplement*; PB-99-128621/XAB; NIOSH-98-119; National Inst. for Occupational Safety and Health, Cincinnati, OH (United States), 1998.

(150) Szulejko, J. E.; Kim, K.-H.; Brown, R. J. C.; Bae, M.-S. Review of Progress in Solvent-Extraction Techniques for the Determination of Polyaromatic Hydrocarbons as Airborne Pollutants. *TrAC Trends Anal. Chem.* **2014**, *61*, 40–48. https://doi.org/10.1016/j.trac.2014.07.001.

(151) Shang, D.; Kim, M.; Haberl, M. Rapid and Sensitive Method for the Determination of Polycyclic Aromatic Hydrocarbons in Soils Using Pseudo Multiple Reaction Monitoring Gas Chromatography/Tandem Mass Spectrometry. *J. Chromatogr. A* **2014**, *1334*, 118–125. https://doi.org/10.1016/j.chroma.2014.01.074.

(152) Du, D. X.; Axelbaum, R. L.; Law, C. K. The Influence of Carbon Dioxide and Oxygen as Additives on Soot Formation in Diffusion Flames. *Symp. Int. Combust.* **1991**, *23* (1), 1501–1507. https://doi.org/10.1016/S0082-0784(06)80419-4.

(153) Gülder, Ö. L.; Snelling, D. R. Influence of Nitrogen Dilution and Flame Temperature on Soot Formation in Diffusion Flames. *Combust. Flame* **1993**, *92* (1), 115–124. https://doi.org/10.1016/0010-2180(93)90202-E.

(154) Kailasanathan, R. K. A.; Yelverton, T. L. B.; Fang, T.; Roberts, W. L. Effect of Diluents on Soot Precursor Formation and Temperature in Ethylene Laminar Diffusion Flames. *Combust. Flame* **2013**, *160* (3), 656–670. https://doi.org/10.1016/j.combustflame.2012.11.004.

(155) Zhang, Y.; Wang, L.; Liu, P.; Guan, B.; Ni, H.; Huang, Z.; Lin, H. Experimental and Kinetic Study of the Effects of CO2 and H2O Addition on PAH Formation in Laminar Premixed C2H4/O2/Ar Flames. *Combust. Flame* **2018**, *192*, 439–451. https://doi.org/10.1016/j.combustflame.2018.01.050.

(156) Kholghy, M.; Saffaripour, M.; Yip, C.; Thomson, M. J. The Evolution of Soot Morphology in a Laminar Coflow Diffusion Flame of a Surrogate for Jet A-1. *Combust. Flame* **2013**, *160* (10), 2119–2130. https://doi.org/10.1016/j.combustflame.2013.04.008.

(157) Vander Wal, R. L.; Tomasek, A. J. Soot Nanostructure: Dependence upon Synthesis Conditions. *Combust. Flame* **2004**, *136* (1), 129–140. https://doi.org/10.1016/j.combustflame.2003.09.008.

(158) Mitra, T.; Zhang, T.; Sediako, A. D.; Thomson, M. J. Understanding the Formation and Growth of Polycyclic Aromatic Hydrocarbons (PAHs) and Young Soot from n-Dodecane in a Sooting Laminar Coflow Diffusion Flame. *Combust. Flame* **2019**, *202*, 33–42. https://doi.org/10.1016/j.combustflame.2018.12.010.

(159) Andreae, M. O.; Gelencsér, A. Black Carbon or Brown Carbon? The Nature of Light-Absorbing Carbonaceous Aerosols. *Atmospheric Chem. Phys.* **2006**, *6* (10), 3131–3148. https://doi.org/10.5194/acp-6-3131-2006.

(160) Samburova, V.; Connolly, J.; Gyawali, M.; Yatavelli, R. L. N.; Watts, A. C.; Chakrabarty, R. K.; Zielinska, B.; Moosmüller, H.; Khlystov, A. Polycyclic Aromatic Hydrocarbons in Biomass-Burning Emissions and Their Contribution to Light Absorption and Aerosol Toxicity. *Sci. Total Environ.* **2016**, *568*, 391–401. https://doi.org/10.1016/j.scitotenv.2016.06.026.

(161) Guilloteau, A.; Bedjanian, Y.; Nguyen, M. L.; Tomas, A. Desorption of Polycyclic Aromatic Hydrocarbons from a Soot Surface: Three- to Five-Ring PAHs. *J. Phys. Chem. A* **2010**, *114* (2), 942–948. https://doi.org/10.1021/jp908862c.

(162) Adkins, E. M.; Miller, J. H. Extinction Measurements for Optical Band Gap Determination of Soot in a Series of Nitrogen-Diluted Ethylene/Air Non-Premixed Flames. *Phys. Chem. Chem. Phys.* **2014**, *17* (4), 2686–2695. https://doi.org/10.1039/C4CP04452E.

(163) Murphy, K. R.; Butler, K. D.; Spencer, R. G. M.; Stedmon, C. A.; Boehme, J. R.; Aiken, G. R. Measurement of Dissolved Organic Matter Fluorescence in Aquatic Environments: An Interlaboratory Comparison. *Environ. Sci. Technol.* **2010**, *44* (24), 9405–9412. https://doi.org/10.1021/es102362t.

(164) Zepp, R. G.; Sheldon, W. M.; Moran, M. A. Dissolved Organic Fluorophores in Southeastern US Coastal Waters: Correction Method for Eliminating Rayleigh and Raman

Scattering Peaks in Excitation–Emission Matrices. *Mar. Chem.* **2004**, *89* (1), 15–36. https://doi.org/10.1016/j.marchem.2004.02.006.

(165) Larsson, T.; Wedborg, M.; Turner, D. Correction of Inner-Filter Effect in Fluorescence Excitation-Emission Matrix Spectrometry Using Raman Scatter. *Anal. Chim. Acta* **2007**, *583* (2), 357–363. https://doi.org/10.1016/j.aca.2006.09.067.

(166) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-Learn: Machine Learning in Python. *J Mach Learn Res* **2011**, *12*, 2825–2830. http://dl.acm.org/citation.cfm?id=1953048.2078195.

(167) Sun, R.; Zobel, N.; Neubauer, Y.; Cardenas Chavez, C.; Behrendt, F. Analysis of Gas-Phase Polycyclic Aromatic Hydrocarbon Mixtures by Laser-Induced Fluorescence. *Opt. Lasers Eng.* **2010**, *48* (12), 1231–1237. https://doi.org/10.1016/j.optlaseng.2010.06.009.

(168) Serio, N.; Prignano, L.; Peters, S.; Levine, M. Detection of Medium-Sized Polycyclic Aromatic Hydrocarbons via Fluorescence Energy Transfer. *Polycycl. Aromat. Compd.* **2014**, *34* (5), 561–572. https://doi.org/10.1080/10406638.2014.918889.

(169) Kim, D.; Kumfer, B. M.; Anastasio, C.; Kennedy, I. M.; Young, T. M. Environmental Aging of Polycyclic Aromatic Hydrocarbons on Soot and Its Effect on Source Identification. *Chemosphere* **2009**, *76* (8), 1075–1081. https://doi.org/10.1016/j.chemosphere.2009.04.031.

(170) Wang, H. Formation of Nascent Soot and Other Condensed-Phase Materials in Flames. *Proc. Combust. Inst.* **2011**, *33* (1), 41–67. https://doi.org/10.1016/j.proci.2010.09.009.

(171) Vaddi, R. S.; Mahamuni, G.; Novosselov, I. Development of an EHD Induced Wind Driven Personal Exposure Monitor and In-Situ Analysis for Characterization of Exposure; St. Petersburg, Russia, 2019.

(172) Mahamuni, G.; Ockerman, B.; Novosselov, I. Electrostatic Capillary Collector for In-Situ Spectroscopic Analysis of Aerosols. *Aerosol Sci. Technol.* **2019**, 1–40. https://doi.org/10.1080/02786826.2019.1600653.

(173) Http://Www.Polysciences.Com/Default/Catalog-Products/Microspheres-Particles/Polymer-Microspheres/Fluoresbrite-Sup-r-Sup-Fluorescent-Microspheres/Fluoresbrite-Sup-r-Sup-Microspheres/Fluoresbrite-Yellow-Green-Microspheres.

(174) Mahamuni, G.; He, J.; Rutherford, J.; Ockerman, B.; Seto, E.; Korshin, G. V.; Novosselov, I. V. Solid Phase Excitation-Emission Matrix Spectroscopy for In-Situ Chemical Analysis of Combustion Aerosols. **2020**. https://doi.org/10.26434/chemrxiv.12167865.v1.

(175) He, J. Y.; Novosselov, I. V. Design and Evaluation of an Aerodynamic Focusing Micro-Well Aerosol Collector. *Aerosol Science and Technology*, 2017, *51*, 1016–1026. https://doi.org/10.1080/02786826.2017.1329515.

(176) Agnarsson, B.; Ingthorsson, S.; Gudjonsson, T.; Leosson, K. Evanescent-Wave Fluorescence Microscopy Using Symmetric Planar Waveguides. *Opt. Express* **2009**, *17* (7), 5075–5082. https://doi.org/10.1364/OE.17.005075.

(177) Tao, S. Q.; Gong, S. D.; Fanguy, J. C.; Hu, X. M. The Application of a Light Guiding Flexible Tubular Waveguide in Evanescent Wave Absorption Optical Sensing. *Sens. Actuators B-Chem.* **2007**, *120* (2), 724–731. https://doi.org/10.1016/j.snb.2006.03.039.

(178) Gross, A.; Stangl, F.; Hoenes, K.; Sift, M.; Hessling, M. Improved Drinking Water Disinfection with UVC-LEDs for Escherichia Coli and Bacillus Subtilis Utilizing Quartz Tubes as Light Guide. *Water* **2015**, *7* (9), 4605–4621. https://doi.org/10.3390/w7094605.