

**Quantifying the contribution of distant, regional, and  
local sources to Particulate Matter (PM) loadings in  
north-west India using a combination of statistical  
tools and low-cost PM sensors**

**HARSHITA PAWAR**

*A thesis submitted for the partial fulfillment of  
the degree of  
**Doctor of Philosophy***



Department of Earth and Environmental Sciences  
Indian Institute of Science Education and Research Mohali  
Knowledge City, Sector 81, SAS Nagar, Manauli PO 140306, Punjab, India.

**December 2020**

## *Dedication*

*This thesis is dedicated to my beloved grandparents.*

*They rose above the status quo and shaped our destinies through an unrelenting commitment to hard work and education.*

*Lt. Smt. Sona Devi and Sh. Netram Pawar*

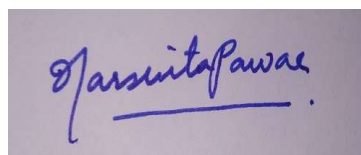
*&*

*Smt. Krishna Sachdeva and Lt. Sh. Ram Lal Sachdeva*

## Declaration

---

The work presented in this thesis has been carried out by me under the guidance of Dr. Baerbel Sinha at the Indian Institute of Science Education and Research Mohali. This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me, and all sources listed within have been detailed in the bibliography.

A photograph of a handwritten signature in blue ink on a light-colored surface. The signature reads "Harshita Pawar" and is underlined with a single horizontal line.

Harshita Pawar

May 2021

In my capacity as the supervisor of the candidate's thesis work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Baerbel Sinha

# Acknowledgments

---

I wish to extend my heartfelt thanks to my supervisor, Dr. Baerbel Sinha, for her valuable guidance and constant support. Her unmatched dedication to her work and resilience have always inspired me.

I am thankful to members of my monitoring committee, Dr. Vinayak Sinha and Dr. Arijit Kumar De, for their insightful suggestions and motivation. Dr. Vinayak Sinha's unique ability to visualize the bigger picture, attention to minute details, and exceptional communication skills have benefitted me immensely. I also want to thank Dr. Arijit Kumar De for providing helpful comments on my research work and being extremely kind and encouraging during all interactions.

I want to thank the Founding Director of IISER Mohali, Dr. N. Sathyamurthy, for his monumental efforts in establishing this institute from ground zero. I am also thankful to all the forthcoming Directors, Dr. Debi Prasad Sarkar, Dr. Arvind, and Dr. J. Gowrishankar, for their sustained endeavors that led to IISER scaling new heights. I am grateful to the entire IISER Mohali administration for providing state-of-the-art infrastructure and a peaceful environment conducive to research. I would like to thank the institute for providing me Junior and Senior Research Fellowships (JRF, SRF) in a very timely manner.

On a personal note, I am incredibly grateful for the love, warmth, and blessings bestowed upon me by Mrs. Suguna Sathyamurthy and Dr. N. Sathyamurthy.

I acknowledge the Department of Earth and Environmental Sciences at IISER Mohali for providing all the necessary resources for research work. It has been a privilege to witness the phenomenal progress of this nascent Department right from its inception. I thank the Atmospheric Chemistry Facility of IISER Mohali for its state-of-the-art measurement facility where I could get hands-on training.

I duly acknowledge the efforts of Mr. Inderjit Singh and Mr. Vivek Singh for training and assisting me in working with the Scanning Electron Microscope. Additionally, I acknowledge Mr. Piyush Aggarwal for providing the Laser Egg low-cost particulate matter sensors and assisting in their installation at IISER Mohali. I also want to thank the entire staff and students at Government

Senior Secondary School, Sangrur, Punjab, and in particular, Mr. Parmal Singh Teja, for their cooperative efforts in the installation and maintenance of the Airveda low-cost sensor.

I would also like to thank the editors and the anonymous reviewers whose helpful suggestions and critical comments have uplifted the quality and clarified the results of my peer-reviewed publications.

I am thankful to IISER Mohali, the International Centre for Integrated Mountain Development (ICIMOD), and the American Geophysical Union (AGU) for providing travel grants and enabling my participation in international scientific conferences.

I want to thank all the past and present colleagues, friends, and members of the Aerosol Research Group and the Atmospheric Chemistry and Emissions Research Group- Saryu, Vinod, Praphulla, Chinmoy, Savita, Gaurav, Haseeb, Ashish, Pallavi, Tess, Abhishek, Vaishali, Nimya, Ebin, Shabin, Bharti, Pooja, Astha, Adarsh and Preeti for always being helpful in innumerable ways.

I am grateful for the steadfast support and kindness I received from the entire #AcademicChatter community on Twitter. Additionally, I also want to thank the magical coffee beans that have kept me going through the highs and lows of my academic career.

I am thankful to my friends Divya, Monika, and Manpreet for being by my side through every thick and thin. Most of the familiar faces I had known from my first five years at IISER left after 2015. I want to acknowledge those who filled that void with happiness and joy: Manisha, Shikha, Nimya, Bharti, Shweta, Sudha, Sandita, Tess, and Astha. I am thankful to my incredibly caring brother Abhishek Verma for always looking out for me in more ways than I can imagine. A special dedication to both Shweta Mishra and Abhishek Mishra for supporting me at my worst and being with me every step of the way. I also want to acknowledge my best friend Dilraj for reminding me of my strengths, for picking me up every time I felt lost, and for always being there.

I want to thank my mother, Mrs. Poonam Pawar, my father, Mr. Balraj Pawar, my brother, Apoorav, for being my pillars of strength, for having unquestioned faith in my abilities, and for patiently supporting me in my endeavors all these years. I am thankful to the Universe for being a silent guardian and a watchful protector on my life's journey.

Harshita Pawar

# Synopsis

## Chapter 1: Introduction

Aerosol particles, ranging from 0.001  $\mu\text{m}$  to 100  $\mu\text{m}$  in size, are ubiquitous in the atmosphere. They interact with and influence all elements of the Climate system. Atmospheric aerosols are defined as an ensemble of solid and liquid particles suspended in the air with lifetimes ranging from a few hours to weeks. Unlike the well-mixed greenhouse gases, aerosol particles exhibit vast spatial and temporal heterogeneity. The first chapter of my thesis reviews the literature and discusses the theoretical foundation of atmospheric aerosols. The chapter discusses the methods to categorize aerosol particles based on sources (natural and anthropogenic) and formation mechanisms (primary and secondary). An aerosol particle can be characterized based on its physical (shape, size, density) and chemical properties. It is tricky to define the size (diameter) of aerosol particles because they often assume complicated non-spherical morphologies. Therefore, aerosol particles are assigned an “equivalent diameter.” Broadly, aerosol particles can be split into *nucleation mode* (freshly nucleated particles <10 nm), *Aitken mode* (10-100 nm), *accumulation mode* (particles that have a long atmospheric residence time because they have neither high mobility nor high inertia 0.1-2.5  $\mu\text{m}$ ), and *coarse mode* (>2.5  $\mu\text{m}$ ) particles based on their size. This chapter describes the physical properties of aerosol particles along with a formulation of their size distribution. Aerosol particles are also chemically diverse. They can be categorized into secondary inorganic (sulfate, nitrate, and ammonium) aerosol particles, primary and secondary organic aerosol particles, elemental carbon (soot) particles, sea-salt, dust made of crustal minerals, biological material including viruses, bacteria, spores, pollen, and plant debris, ash, anthropogenic particles generated during the wear and tear of various materials and trace metals. Aerosol particles are also subject to intricate aging processes in the atmosphere, which alter their physical, chemical, and radiative properties. Aerosols evolve under the influence of changes in meteorological conditions, particularly in response to changes in relative humidity and temperature, resulting in a reversible mass transfer from the carrier gas, air, to the liquid or solid particles therein. In summary, this chapter attempts to rationalize the vast spatial and temporal heterogeneity characteristic of atmospheric aerosol particles.

Aerosol particles primarily scatter and occasionally absorb (black and brown carbon) sunlight. They also serve as cloud condensation nuclei on which water condenses to form clouds. Therefore, aerosol particles impact both the radiative budget and the Earth’s hydrological cycle. While we have a robust understanding of the impacts of greenhouse gases on the climate, that of aerosol particles remains riddled with uncertainties. Further, aerosol particles enter the human body via

the nasal passages and eventually lead to respiratory and cardiovascular ailments. Therefore, given the implications of these tiny particles on human health and climate, it is essential to investigate their sources, sinks, and transformation mechanisms.

Globally, several resolutions aim to mitigate the levels of aerosol particles in the atmosphere. Most countries have designated the mass concentration of aerosol particles below the aerodynamic diameters of 10  $\mu\text{m}$  and 2.5  $\mu\text{m}$  as a marker of air quality and human exposure. However, the efforts to combat particulate pollution are limited by an inadequate understanding of the relative contribution of different sources to the overall mass loading and lack of robust measurements, particularly in India. In such a scenario, mitigation policies are often guided by popular opinion rather than scientific facts. Policies implemented elsewhere worldwide with totally different sources have been blindly put into action with little output. In my thesis, I have tried to evaluate the relative contribution of local, regional, and distant sources to the overall pollutant loading in north-west India using a combination of statistical tools and low-cost sensors. Additionally, the results in my thesis pave the way for more accurate low-cost sensors that can ramp up spatial and temporal data availability.

Chapter 1 concludes by stating the motivation for writing this thesis and listing the questions that I have aimed to address:

- 1) What is the general climatology of the air masses arriving at Mohali?
- 2) What is the contribution of long-range transport to the particulate matter loadings at Mohali, a regionally representative site in north-west India?
- 3) How does long-range transport impact the percentage of days when the receptor site, Mohali, is not in compliance with the national ambient air quality standards of  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$ ?
- 4) In a complex atmospheric environment with varied meteorology, what factors impact the accuracy of low-cost optical sensors when used for continuous ambient air quality monitoring?
- 5) How to correct the inaccurate particulate matter measurements from the low-cost sensors?
- 6) What is the contribution of paddy residue burning and domestic biofuel burning to the particulate matter loadings in north-west India?

## **Chapter 2: Quantitative measurement of ambient particulate matter and approaches to link particulate matter measurements with air-mass history**

This chapter is split into two parts: quantitative measurement of ambient particulate matter (PM) and approaches to link the measurements of PM with air-mass history. The first part of this chapter highlights the general properties of aerosol particles that can be quantified and their associated

challenges. Globally, the air quality standards of PM are based on the mass concentration of ambient particulate matter below the aerodynamic diameters of 10  $\mu\text{m}$  and 2.5  $\mu\text{m}$ . Therefore, measurements of the mass concentration are considered critical. Gravimetry-based analyzers with a size-selective inlet, designated as the Federal Reference Methods (FRMs) by the US Environmental Protection Agency (EPA), are generally considered the gold standard for PM mass concentration measurement. To foster research and development in new techniques to measure the PM mass concentration, EPA designates another category of analyzers called the Federal Equivalent Method (FEM). In comparison to FRMs, automated FEMs provide data at a higher temporal resolution. This chapter briefly discusses the measurement principle of a few of the most widely used FRMs.

The reference  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  measurements used in this thesis were acquired using the Thermo Fischer Scientific 5014i  $\beta$ -continuous ambient particulate monitors installed at the IISER Mohali Atmospheric Chemistry Facility (30.667 °N, 76.729 °E). This chapter discusses the theory, functioning, and troubleshooting of these analyzers in detail.

Aerosol particles are also efficient scatterers of light. The intensity and the angular distribution of the scattered light can provide information about the number, size, and nature of particles. This chapter discusses the principle behind particle characterization using light scattering in detail. Recently, the availability of cheap lasers, along with an increase in demand for open-source air quality data, has led to the development of several low-cost optical instruments. I briefly discuss some of the popular low-cost optical sensors. A description of the Laser Egg and the Airveda low-cost sensors that have been used in this thesis (Chapters 4 and 5) to monitor ambient particulate matter has also been provided.

The particulate matter concentration at any place is a complex function of sources, sinks, and meteorology. Winds can transport pollutants emitted at the source to sites downwind. The relative contribution of transported PM to the overall PM loading is a valuable piece of information. Linking air mass history with the observed aerosol composition at the receptor site can reveal necessary information about the spatial location and the contribution of the pollution source. Several approaches to do so have been discussed in the second part of this chapter. Combining the measurements of locally measured pollutants with wind speed and wind direction can help visualize the source location with respect to the receptor. I discuss the conditional probability and polar bivariate plots in this chapter. Air mass back-trajectories are a powerful yet convenient tool to assess the local impact of distant pollution sources. I explain the methodology for computing back-trajectories, their validation, and the factors that affect their accuracy. This chapter also consists of an in-depth description of the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPPLIT) model, one of the most popular trajectory computation tools. A brief discussion on



more sophisticated tools, Particle Dispersion Model and Chemistry Transport Model, has been included.

Out of the several approaches mentioned above to explain the observed pollutant concentrations using air-mass history, I have used back-trajectories in my thesis. The concluding part of the chapter builds further on extracting information from trajectories using clustering methods (hierarchical and non-hierarchical) and geographical sector classification (concentration field and potential source contribution function).

### **Chapter 3: Quantifying the contribution of long-range transport to particulate matter mass loadings at a suburban site in the north-western Indo-Gangetic Plain (NW-IGP)**

Many sites in the densely populated Indo-Gangetic Plain (IGP) frequently exceed the national ambient air quality standard (NAAQS) of  $100 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{10}$  and  $60 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{2.5}$ . Particulate matter mass loadings at any site are a complex function of the sources, sinks, and meteorological conditions. Wind transports pollutants emitted at the source to downwind sites. Therefore, it is essential to quantify the relative contribution of long-range transport facilitated by high wind speeds to the overall regional PM levels for formulating effective mitigation policies.

In this chapter, I quantify the contribution of long-range transport and local pollution episodes to the elevated PM levels and the number of exceedance events through a back-trajectory climatology analysis of the air masses arriving at the IISER Mohali Atmospheric Chemistry Facility for the period Aug 2011-June 2013. First, I compute 72-hour back-trajectories using the National Oceanic and Atmospheric Administration's HYSPLIT model and subject them to a non-hierarchical clustering algorithm (k-means clustering). The ensemble of trajectories is split into six distinct clusters, each representing unique synoptic-scale air-mass transport patterns.

Then, I evaluate the season-wise average coarse ( $\text{PM}_{10-2.5}$ ) and fine mode PM ( $\text{PM}_{2.5}$ ) mass loading associated with each air-mass cluster and assess if the intra-cluster differences are statistically significant. I find that the air-mass transport from the westerly clusters belonging to the arid regions of Arabia, Thar desert, Middle East, and Afghanistan leads to significant enhancements in the average coarse PM mass loading (varying between 9 % and 57 % of the total coarse PM mass) at Mohali during all seasons. The contribution of long-range transport from the westerly clusters to the elevated fine mode PM was most pronounced during the summer season (varying between 4 % to 31 %). On the other hand, south-easterly air masses (source: eastern Indo-Gangetic Plain) were found to bring cleaner air masses with significantly lower levels of coarse and fine mode PM to the receptor site during all seasons. During the winter season, local pollution episodes (wind speed  $< 1 \text{ m s}^{-1}$ ) contributed  $\sim 14$  % and 22 % to the overall coarse and fine mode PM mass loading,

respectively. Additionally, I investigate the possible identities of the sources of the fine mode PM by observing its correlation with specific gas-phase tracers (CO, NO<sub>2</sub>, benzene, and acetonitrile) measured at the IISER Mohali Atmospheric Chemistry Facility. Excellent correlation ( $r^2$  between 0.39 and 0.70) was found between fine mode PM and CO, an incomplete combustion tracer, during all seasons. This suggests that such inefficient combustion sources (municipal waste burning, biomass burning) should be targeted for mitigating the particulate matter pollution in this region. This chapter also provides evidence of the rampant formation of coarse mode PM from the aqueous-phase processing of gas-phase precursors during winter and monsoon seasons.

Long-range transport was found to have a minimal impact on the percentage of exceedance events in all seasons (days not in compliance with the NAAQS of PM<sub>10</sub> and PM<sub>2.5</sub>). For the local cluster, which represents regional air masses (source region: NW-IGP), the fraction of days during which the 24h average NAAQS of PM<sub>2.5</sub> was exceeded varied between 36 % of the days associated with this synoptic-scale transport during the monsoon and 95 % during post-monsoon and winter seasons; the fraction of days during which the NAAQS of the 24h average PM<sub>10</sub> was exceeded, varied between 48 % during the monsoon and 98 % during the post-monsoon season. Whenever long-range transport enhanced the fraction of exceedance days, the increase varied between a few percent and at most 20 %. This confirms that the residents in this region suffer from chronic exposure to high levels of coarse and fine mode PM due to the local sources rather than the long-range transport.

In conclusion, the results of this chapter establish a general climatology of the air masses at Mohali using back-trajectory analysis. Further, the link between air-mass history and locally measured PM loadings is probed. Additionally, I ascertain a baseline of the season-wise regional backgrounds of coarse and fine mode PM against which the contribution from local sources can be gauged.

#### **Chapter 4: Humidity, density, and inlet aspiration efficiency corrections improve the accuracy of a low-cost particulate matter sensor during field calibration at a suburban site in the north-western Indo-Gangetic plain (NW-IGP)**

To meet the rising demand for open-source air quality data, several low-cost sensors have permeated the consumer market with little to no field validation. Cost-effectiveness and compact size make them desirable candidates to integrate with the existing sparse measurement facilities and achieve a denser network of real-time data. However, only a few studies have assessed the long-term performance of these sensors in complex Indian environments. In this chapter, I evaluate the efficacy of Laser Egg, a low-cost optical sensor, for monitoring ambient particulate matter in Mohali during the summer and monsoon seasons in the year 2016. Further, I highlight the factors

which affect the accuracy of Laser Egg, including hygroscopic growth, aerosol density, aspiration losses of particles at high wind speeds, and the implications of assuming optical equivalent diameter to be the same as aerodynamic diameter and discuss site-specific corrections to address each issue.

Laser Egg sensors are mainly intended for indoor air monitoring. However, this study used modified sensors with an additional fan at the inlet to increase the airflow inside the chamber. Finally, two identical sensors were placed next to the inlets of the US EPA-approved  $\beta$ -attenuation analyzers, and their performance was tracked from April 2016 to July 2016.

The Laser Egg sensors reported precise measurements of  $PM_{10}$  and  $PM_{2.5}$  with the coefficient of variance  $< 12\%$  and Pearson's  $r > 0.9$ . A negative mean bias error (MBE) of  $-38.2 \mu\text{g m}^{-3}$  (Sensor 1) and  $-44.9 \mu\text{g m}^{-3}$  (Sensor 2) indicated that the Laser Egg underestimated  $PM_{10}$  as it used a sub-optimal value of aerosol density and suffered from aspiration losses at high wind speeds. Hygroscopic growth of aerosol particles and the lack of an RH regulation system resulted in a positive MBE in Laser Egg  $PM_{2.5}$  measurements ( $19.1$  and  $28.3 \mu\text{g m}^{-3}$  for Sensor 1 and Sensor 2, respectively) during the monsoon season (characterized by  $RH > 60\%$ ). After applying site-specific corrections, the accuracy of the LE monitors increased. The MBE in Airveda  $PM_{10}$  measurements during the summer season (27 April - 15 June 2016) reduced from  $\sim -90 \mu\text{g m}^{-3}$  to  $-30.9 \mu\text{g m}^{-3}$  (Sensor 1) and  $-23.2 \mu\text{g m}^{-3}$  (Sensor 2) after correcting for particle density and aspiration losses. During the monsoon period (16 June - 25 July 2016), the MBE of the Laser Egg  $PM_{2.5}$  measurements decreased from  $19.1 \mu\text{g m}^{-3}$  to  $8.7 \mu\text{g m}^{-3}$  and from  $28.3 \mu\text{g m}^{-3}$  to  $16.5 \mu\text{g m}^{-3}$  for Sensor 1 and Sensor 2, respectively after correcting for particle density and hygroscopic growth. The change in the magnitude of slopes (obtained after performing reduced major axis regression with the reference PM) from  $0.61$ - $1.23$  to  $0.83$ - $1.03$  indicated improvement in sensor accuracy. Likewise, the value of Pearson's  $r$  increased from the range of  $0.41$  -  $0.53$  to  $0.51$  -  $0.67$ . The corrections reduced the overall MBE to  $< -20 \mu\text{g m}^{-3}$  for  $PM_{10}$  and  $< 3 \mu\text{g m}^{-3}$  for  $PM_{2.5}$ , indicating that the modified version of the LE could be used for ambient PM monitoring with appropriate correction and meteorological observations.

As discussed in this chapter, accurate measurements of relative humidity and temperature from co-located measurement devices can be useful for correcting the PM measurements from low-cost sensors. Several commercial low-cost sensors already come with in-built meteorological detectors. In this chapter, I show that the incorrect placement of these sensors can adversely impact data accuracy. Absorption of solar radiation by the casing of the LE caused an overestimation of temperature (by  $\sim 5$  °C to  $10$  °C) during the daytime, while evaporative cooling at night resulted in temperature readings that were lower than the reference ambient temperature by  $0.5$  °C to  $2$  °C. The sensors also underestimated RH by almost  $50\%$  throughout the calibration period.

Further, I report that Wi-Fi is not an optimum choice for data transmission, especially if low-cost sensors are deployed at remote locations. Even though the Laser Egg monitors were continuously charged, and there was no power shortage, the data availability was only 48% and 28% for Sensors 1 and 2, respectively. Those looking to use low-cost sensors for extensive regulatory monitoring must either invest in long-term cloud storage or a SIM card with an active internet connection rather than relying on Wi-Fi alone.

In conclusion, the results presented in this chapter provide critical insights that can be used for designing more accurate low-cost sensors.

### **Chapter 5: Particulate matter measurements from rural sites in north-west India reveal that local fuel choices and residue management have a disproportionate impact on ambient air quality.**

This chapter showcases the potential applications of low-cost sensors for the source apportionment of ambient particulate matter in north-west India, a region with inadequate air quality monitoring. November onwards, the poor air quality over India's National Capital Region is blamed on the large-scale agricultural residue burning in Punjab and Haryana. However, the emission strength remains poorly constrained due to the lack of ground-based measurements over rural areas within the source region and issues in the satellite detection of small-scale paddy residue fires. In this chapter, variations in the PM<sub>10</sub> and PM<sub>2.5</sub> levels were investigated across four sites in Punjab (Nadampur, Kalal Majra, Mohali, and Patiala) with varying socioeconomic levels and land-use categories from 1 October 2019 to 19 December 2019. Nadampur and Kalal Majra are rural sites, while Mohali and Patiala are urban areas.

PM mass concentration at Nadampur and Mohali was measured using the Airveda low-cost sensors. A multi-season field calibration with respect to US EPA approved  $\beta$ -attenuation analyzers at IISER Mohali revealed that although the Airveda sensors were more accurate (Pearson's  $r$  between 0.6 and 0.8) in comparison to the Laser Egg sensors (Pearson's  $r$  between 0.4 and 0.5), high relative humidity and high PM mass loading affected their performance. Random Forest machine learning algorithm was used to correct the biases in the Airveda sensor's PM measurements. Post-correction, the sensors provided sufficient accuracy to substitute the bulkier and costlier  $\beta$ -attenuation instrument in field deployment and sensor networks, as indicated by a Pearson's  $R > 0.9$  and a slope of 0.8 to 1.0 with respect to the reference analyzers.

I used the Airveda sensors to quantify the contribution of paddy residue fires and heating demand induced domestic biofuel burning to the overall PM mass loading in north-west India. The daily average PM<sub>10</sub> and PM<sub>2.5</sub> mass concentration in late October and early November at Nadampur

correlated well (Pearson's  $R > 0.7$ ) with the daily sum of VIIRS (Visible Infrared Imaging Radiometer Suite Emissions) fire counts in a  $50 \text{ km} \times 50 \text{ km}$  area surrounding the rural site. I calculated two independent proxies to constrain the satellite detection of paddy residue fires over this region. Correlating the PM emissions with the VIIRS detected Fire Radiative Energy (FRE), I estimated the coefficient of emissions ( $C_e$ ), which directly relates the FRE to the mass of aerosol emitted. The estimated  $C_e$  for  $\text{PM}_{10}$  ( $0.038 \text{ kg MJ}^{-1}$ ) was comparable to the satellite-derived  $C_e$  of  $0.04 \text{ kg MJ}^{-1}$  for  $\text{PM}_{10}$  over this region. A top-down value of the  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  emission factor (EF) was estimated by correlating the PM emissions to the paddy residue burnt, derived using Fire Radiative Power detected by VIIRS. The estimated  $\text{PM}_{2.5}$  EF of  $351.6 \text{ g per kg}$  of paddy residue burnt was  $>17$  times larger than the corresponding EFs reported in laboratory studies. Agreement of the  $C_e$  estimated in this study with the reported value and a disagreement of the top-down estimate of PM emission factors with the laboratory reported values indicates an underdetection of paddy residue fires. It is likely that only a small fraction of the total fires that burn each day are detected during the short satellite overpass time (13:30 LT) of VIIRS, explaining the overestimation of estimated EFs. Overall, I find that Punjab and northern Haryana were responsible for more than 80% of the total  $\text{PM}_{10}$  ( $61.7 \text{ Gg}$ ) and  $\text{PM}_{2.5}$  ( $38.4 \text{ Gg}$ ) emissions over the NW-IGP from paddy residue fires between 1 October and 19 December.

From 20 November to 19 December 2019, the daily average temperature dipped from  $17.0 \text{ }^\circ\text{C}$  to  $9.9 \text{ }^\circ\text{C}$ , leading to the enhanced use of solid biofuels, such as fuelwood and cow dung cakes, for cooking and space heating. I find that every  $1 \text{ }^\circ\text{C}$  increase in the heating demand led to the emission of  $5.7$  tonnes of  $\text{PM}_{10}$  ( $r = 0.53$ ) and  $5.3$  tonnes ( $r = 0.61$ ) of  $\text{PM}_{2.5}$  in Nadampur. However, increased heating demand had no statistically significant effect on the  $\text{PM}_{10}$  ( $-0.1 \text{ t }^\circ\text{C}^{-1}$ ,  $r = -0.02$ ) and  $\text{PM}_{2.5}$  ( $1.2 \text{ t }^\circ\text{C}^{-1}$ ,  $r = 0.36$ ) emissions in Mohali.

Finally, source apportionment of the PM mass concentration was performed at the four sites using Multiple Linear Regression (MLR) and Positive Matrix Factorization (PMF). In the descending order, paddy harvest enhanced the  $\text{PM}_{10}$  ( $\text{PM}_{2.5}$ ) mass concentration in Nadampur and Kalal Majra by  $44.8 \pm 1.7 \text{ } \mu\text{g m}^{-3}$  ( $20.1 \pm 5.2 \text{ } \mu\text{g m}^{-3}$ ),  $41.1 \pm 7.0 \text{ } \mu\text{g m}^{-3}$  ( $12.0 \pm 4.4 \text{ } \mu\text{g m}^{-3}$ ), respectively. In comparison, local and regional paddy residue fires enhanced the  $\text{PM}_{10}$  ( $\text{PM}_{2.5}$ ) mass concentration in Patiala, Nadampur, Mohali, and Kalal Majra by  $103.1 \pm 13.7 \text{ } \mu\text{g m}^{-3}$  ( $50.1 \pm 7.6 \text{ } \mu\text{g m}^{-3}$ ),  $97.0 \pm 36.6 \text{ } \mu\text{g m}^{-3}$  ( $53.4 \pm 16.8 \text{ } \mu\text{g m}^{-3}$ ), and  $58.1 \pm 6.6 \text{ } \mu\text{g m}^{-3}$  ( $34.1 \pm 8.7 \text{ } \mu\text{g m}^{-3}$ ),  $55.4 \pm 12.8 \text{ } \mu\text{g m}^{-3}$  ( $22.1 \pm 5.6 \text{ } \mu\text{g m}^{-3}$ ) respectively, from 20 October to 19 November 2019. Even though the paddy residue fires had subsided after 20 November 2019, the air quality continued to be poor. The dip in the daily average temperature from  $17.0 \text{ }^\circ\text{C}$  to  $9.9 \text{ }^\circ\text{C}$  increased the heating demand, enhancing the residential burning of solid biofuels for space and water heating. In the descending order, the

increased heating demand enhanced the PM<sub>10</sub> (PM<sub>2.5</sub>) mass concentrations in Nadampur, Kalal Majra, Patiala, and Mohali by  $151.2 \pm 47.2 \mu\text{g m}^{-3}$  ( $120.1 \pm 8.8 \mu\text{g m}^{-3}$ ),  $84.2 \pm 24.6 \mu\text{g m}^{-3}$  ( $34.1 \pm 7.3 \mu\text{g m}^{-3}$ ),  $38.3 \pm 7.6 \mu\text{g m}^{-3}$  ( $22.5 \pm 0.3 \mu\text{g m}^{-3}$ ), and  $30.7 \pm 0.4 \mu\text{g m}^{-3}$  ( $22.3 \pm 17.3 \mu\text{g m}^{-3}$ ), respectively, from 20 November to 19 December 2019. Unlike paddy residue burning, which typically affects the air quality for over a month, heating-related emissions profoundly impact the air quality for multiple months. Thus, promoting cleaner technology for water and space heating can play a key role in improving air quality during the cold season.

## **Chapter 6: Conclusions: Major findings and outlook**

This concluding chapter highlights the major results of this thesis and their implications. The results in Chapter 3 establish a regional background of particulate matter above which contribution from the local and synoptic scale sources can be quantified. The air masses originating from the arid regions of the Thar desert, Arabia, the Middle East, and Afghanistan lead to significant enhancements in the coarse mode loading. On the contrary, those from the eastern Indo-Gangetic Plain bring substantially cleaner air masses to the receptor site. One of the most noteworthy findings is that the residents of the north-western Indo-Gangetic Plain suffer from long-term chronic exposure to high levels of particulate matter solely from the regional sources and not long-range transport. PM<sub>2.5</sub> depicted an excellent correlation with CO across all seasons, which suggests the co-emission of fine mode particles from inefficient combustion sources. Additionally, aqueous-phase processing of gas-phase precursors during periods of high humidity in the winter and monsoon season was found to contribute significantly to the coarse-mode loading in this region. Therefore, standalone mitigation policies aimed at reducing just the levels of particulate matter will not be successful unless the gas phase precursors are also targeted simultaneously.

Globally, several people, independent organizations, and firms have started using low-cost particulate sensors like a black box without giving much thought to the data quality. In this thesis, I report the performance of two low-cost sensors, the Laser Egg sensor (Chapter 4) and the Airveda sensor (Chapter 5). The Airveda sensor, which has been calibrated for Indian conditions, outperformed the Laser Egg, a sensor calibrated in China, in terms of accuracy. This thesis presents two approaches to correct the PM measurements of low-cost sensors. PM measurements of the Laser Egg sensors were corrected using empirically derived corrections (Chapter 4), while that of the Airveda sensors were corrected using a machine learning algorithm (Chapter 5). Deriving season-specific and issue-specific empirical corrections were more challenging than the machine learning-based approach. However, the Random Forest algorithm could not extrapolate beyond the limits of the training dataset. Further efforts are needed to evaluate the efficacy of a machine learning model at predicting accurate PM measurements for periods beyond the training datasets.

My thesis highlights the pitfalls of using factory-calibrated low-cost sensors for ambient air-quality monitoring in India. Apart from the well-known artifacts arising due to water absorption at high relative humidity, the effect of inlet losses due to wind speeds on the sensor accuracy has been documented for the first time. The results of Chapter 4 can be applied to design more accurate optical sensors, which would eliminate the need for complicated post-processing of the raw sensor data. Further, the high spatial and temporal heterogeneity in PM measurements needs to be accounted for while implementing sensor calibration. A sensor calibrated for optimum performance in one region may not report accurate measurements for another region with distinct sources. Likewise, the same sensor calibration may not be valid across all seasons. Accurate measurement of ambient relative humidity is one of the essential parameters to rectify sensor measurements. Although RH sensors are relatively cheap, their proper placement is crucial to ensure data correctness.

Using ground-based PM measurements from the Airveda low-cost sensor, I show that the polar-orbiting satellites underestimate the fire counts over north-west India due to their limited overpass times. Moreover, paddy residue burning cannot be blamed for the poor air quality beyond the one month of the peak burning. Increased residential fuel usage during cold periods for space heating and water heating has a profound and currently neglected impact on ambient air.

This thesis evaluates the contribution of distant, regional, and local PM sources over the NW-IGP to air quality deterioration and finds that mitigation of local sources has the greatest potential to improve the air quality at all receptor sites studied. Different mitigation policies are needed to tackle urban and rural PM pollution. Promoting cleaner technology for water and space heating can play a crucial role in improving air quality during the cold season. A low-cost PM sensor monitoring network employing spatial interpolation techniques that account for dispersion could potentially be used to locate PM sources with a disproportionate impact on air quality in urban environments and industrial areas in the future.

# Table of Contents

Synopsis .....	vi
List of Publications .....	xix
<b>1. Introduction .....</b>	<b>1</b>
1.1. Atmospheric composition in the Anthropocene .....	1
1.2. What are aerosols? What is particulate matter? .....	2
1.3. Types of atmospheric aerosols .....	3
1.4. Physicochemical properties of atmospheric aerosols .....	6
1.5. Atmospheric processing and sinks of aerosols .....	7
1.6. Health impacts of atmospheric aerosols .....	9
1.7. Climate impacts of atmospheric aerosols .....	12
1.8. Research motivation and thesis outline .....	13
<b>2. Quantitative measurement of ambient particulate matter (PM) and approaches to link PM measurements with air-mass history .....</b>	<b>15</b>
2.1. Analytical techniques to quantify ambient Particulate Matter .....	15
2.2. Federal Reference and Federal Equivalent Methods to measure the mass concentration of particulate matter .....	18
2.3. PM monitoring at IISER Mohali using 5014i Beta Continuous Ambient Particulate Monitor .....	20
2.4. Optical methods to characterize atmospheric aerosols .....	24
2.4.1. Theoretical background .....	25
2.4.2. The mathematical formulation of light scattering .....	27
2.4.3. PM monitoring at IISER Mohali using low-cost optical sensors .....	30
2.5. Approaches to link air mass history with atmospheric composition at a receptor site .....	33
2.5.1. Wind speed and wind direction measurements .....	34
2.5.2. Trajectory based models .....	36
2.5.3. Lagrangian particle dispersion models .....	40
2.5.4. Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) Model .....	40
2.5.5. Chemistry transport models .....	43
2.6. Approaches to link air mass trajectory with atmospheric composition .....	43
2.6.1. Geographical sector classification .....	43
2.6.1.1. Potential Source contribution function (PSCF) .....	43
2.6.1.2. Concentration Field (CF) .....	44
2.6.2. Clustering methods .....	45
2.6.2.1. Hierarchical clustering methods .....	45
2.6.2.2. Non-hierarchical clustering methods .....	46



<b>3. Quantifying the contribution of long-range transport to particulate matter (PM) mass loadings at a suburban site in the north-western Indo-Gangetic Plain (NW-IGP).....</b>	<b>47</b>
3.1. Abstract.....	47
3.2. Introduction.....	48
3.3. Materials and Methods .....	50
3.3.1. Study location, air quality data, and general meteorology .....	50
3.3.2. Back trajectory modelling .....	58
3.3.3. Back trajectory cluster analysis .....	59
3.4. Results and discussion .....	60
3.4.1. Optimisation of the number of clusters .....	60
3.4.2. Spatial and dynamic patterns of the airflow associated with the clusters .....	61
3.4.3. Impact of air-mass transport on particulate matter (PM) mass loadings.....	66
3.4.3.1. Winter season.....	67
3.4.3.2. Summer season .....	72
3.4.3.3. Monsoon season.....	75
3.4.3.4. Post-monsoon season.....	76
3.4.4. Impact of air-mass transport on particulate matter (PM) exceedance events.....	77
3.4.4.1. Winter season.....	77
3.4.4.2. Summer season .....	78
3.4.4.3. Monsoon season.....	79
3.4.4.4. Post-monsoon season.....	79
3.5. Conclusions.....	79
<b>4. Humidity, density, and inlet aspiration efficiency correction improve accuracy of a low-cost particulate matter sensor during field calibration at a suburban site in the north-western Indo-Gangetic plain (NW-IGP).....</b>	<b>82</b>
4.1. Abstract.....	82
4.2. Introduction.....	83
4.3. Materials and methods .....	86
4.3.1. Study location and instrumentation.....	86
4.3.2. Laser Egg air quality monitor.....	88
4.3.3. Data processing and statistical analysis.....	89
4.4. Results and discussion .....	92
4.4.1. Ambient temperature, relative humidity, and solar radiation measurements.....	92
4.4.2. Particulate Matter measurements .....	93
4.4.3. PM <sub>10</sub> measurements.....	93
4.4.4. PM <sub>2.5</sub> measurements .....	97
4.4.5. Impact of aspiration losses on Laser Egg accuracy.....	99

4.4.6.	Impact of aerosol density on Laser Egg accuracy .....	101
4.4.7.	Impact of relative humidity on Laser Egg accuracy.....	103
4.4.8.	Correction for the biases introduced by aspiration losses, particle density, and relative humidity in Laser Egg PM measurements .....	104
4.4.9.	Laser Egg PM measurements post-correction.....	113
4.4.10.	Ambient temperature measurements .....	115
4.4.11.	Relative humidity measurements.....	117
4.5.	Conclusions.....	119
<b>5.</b>	<b>Particulate matter measurements from rural sites in North-West India reveal that local fuel choices and residue management have a disproportionate impact on ambient air quality</b>	<b>122</b>
5.1.	Abstract.....	122
5.2.	Introduction.....	123
5.3.	Materials and Methods .....	124
5.3.1.	Site description and instrumental details for the field calibration.....	124
5.3.2.	Airveda low-cost sensor .....	126
5.3.3.	Metrics to evaluate sensor performance.....	127
5.3.4.	Random Forest algorithm for improving sensor accuracy .....	128
5.3.5.	Retrieval of fire counts and fire radiative power from VIIRS and boundary layer height from ERA5.....	135
5.3.6.	Tools for the source apportionment of ambient particulate matter .....	139
5.4.	Results and discussion .....	141
5.4.1.	Contribution of paddy residue burning to the PM emissions.....	141
5.4.2.	Contribution of domestic biofuel burning to PM mass loadings.....	145
5.4.3.	Source apportionment of PM mass loading .....	146
5.5.	Conclusions.....	152
<b>6.</b>	<b>Conclusions: Major findings and outlook.....</b>	<b>154</b>
<b>7.</b>	<b>List of Figures.....</b>	<b>159</b>
<b>8.</b>	<b>List of Tables .....</b>	<b>165</b>
<b>9.</b>	<b>List of Abbreviations .....</b>	<b>167</b>
<b>10.</b>	<b>Bibliography .....</b>	<b>169</b>

# List of Publications

---

1. **Pawar, H.**, and Sinha, B.: Particulate matter measurements from rural sites in north-west India reveal that local fuel choices and residue management have a disproportionate impact on ambient air quality [Communicated]
2. **Pawar, H.**, and Sinha, B.: Humidity, density, and inlet aspiration efficiency correction improve accuracy of a low-cost sensor during field calibration at a suburban site in the North-Western Indo-Gangetic plain (NW-IGP), *Aerosol Science and Technology*, 54 (6), 685-703, 10.1080/02786826.2020.1719971, 2020.
3. **Pawar, H.**, Garg, S., Kumar, V., Sachan, H., Arya, R., Sarkar, C., Chandra, B. P., and Sinha, B.: Quantifying the contribution of long-range transport to particulate matter (PM) mass loadings at a suburban site in the north-western Indo-Gangetic Plain (NW-IGP), *Atmospheric Chemistry & Physics*, 15 (16), 9501-9520, 10.5194/acp-15-9501-2015, 2015.
4. Hakkim, H., Sinha, V., Chandra, B. P., Kumar, A., Mishra, A. K., Sinha, B., Sharma, G., **Pawar, H.**, Sohpaal, B., Ghude, S. D., Pithani, P., Kulkarni, R., Jenamani, R. K., and Rajeevan, M.: Volatile organic compound measurements point to fog-induced biomass burning feedback to air quality in the megacity of Delhi, *Science of The Total Environment*, 689, 295-304, <https://doi.org/10.1016/j.scitotenv.2019.06.438>, 2019.
5. Chandra, B. P., Sinha, V., Hakkim, H., Kumar, A., **Pawar, H.**, Mishra, A. K., Sharma, G., Garg, S., Ghude, S. D., and Chate, D. M.: Odd-even traffic rule implementation during winter 2016 in Delhi did not reduce traffic emissions of VOCs, carbon dioxide, methane and carbon monoxide, *Current Science*, 114 (6), 1318-1325, 2018.
6. Ghude, S.D., Bhat, G.S., Prabhakaran, T.D., Jenamani, R.K., Chate, D.M., Safai, P.D., Karipot, A., Konwar, M., Pithani, P., Sinha, V., Rao, P.S., Dixit, S.A., Tiwari, S., Todekar, K., Varpe, S., Srivastava, A., Bisht, D.S., Murugavel, P., Ali, K., Mina, U., Dharua, M., Rao, J.S., Padmakumari, B., Hazra, A., Nigam, N., Shende, U., Lal, D.M., Chandra, B.P., Mishra, A.K., Kumar, A., Hakkim, H., **Pawar, H.**, Acharja, P., Kulkarni, R., Subharthi, C., Balaji, B., Varghese, M., Bera, S., & Rajeevan, M.C.: Winter fog experiment over the Indo-Gangetic Plains of India, *Current Science*, 112, 767-784, 2017.

*“Without the dust particles in the atmosphere, there will be no haze, no fog, no clouds, and therefore probably no rain.”*

*- John Aitken, 1880*

# Chapter 1

---

## **Introduction**

### **1.1. Atmospheric composition in the Anthropocene**

The Earth’s climate system is a complex interactive system comprising the atmosphere, the hydrosphere, the land surface, the cryosphere, and the biosphere (Masson-Delmotte, 2018). This system keeps evolving under the influence of intra-component dynamics and external forcing. Since the mid-20<sup>th</sup> century, human activities have had a substantial impact on the components of the Earth system, so much so that a new geological epoch, the “Anthropocene,” was coined to differentiate this period from the preceding epoch, the Holocene (Crutzen, 2006; Lewis and Maslin, 2015).

The Anthropocene is characterized by an accelerated rate of urbanization, industrialization, and land-use changes coupled with an incessant burning of fossil fuels to fulfil the energy demands of a growing population (Zalasiewicz et al., 2015) accompanied by profound changes in the global nitrogen, carbon, and phosphorus cycles. Human actions have adversely affected the climate system and the ecosystems, including but not limited to causing global warming, habitat destruction, climate extremes, microplastic pollution, ocean acidification, wildfires, and extinctions.

Like all components of the Earth system, human actions have also altered the chemistry and composition of the atmosphere. Briefly, the atmosphere is a protective layer of gases that envelops the Earth. Based on the temperature profile, it is divided into five layers, namely, the troposphere, the stratosphere (contains the ozone layer), the mesosphere, the thermosphere, and the exosphere.

This thesis focuses primarily on the troposphere. It is the closest layer to the Earth's surface and is characterized by a decrease in temperature with height. Its vertical extent from the ground varies between 8 km (near poles) to 18 km (near tropics), depending upon the latitude. This layer consists of roughly 80% of the total atmospheric mass and almost all of the water vapor (Showman and Dowling, 2014; Speight, 2017). Temperature gradient and turbulence in this portion of the atmosphere ensure the rapid mixing of air masses due to convection (Boucher, 2015). Most weather phenomena, including cloud formation, are limited to this layer. Most of my work focuses on characterizing sources of aerosol in the boundary layer. The boundary layer is the section of the troposphere most affected by the activities on Earth's surface. It extends roughly 1 to 2 km from the ground level. The vertical extent of the boundary layer evolves in response to the radiative heating from the ground.

The Atmosphere is the recipient of trace gases and tiny particles emitted directly or indirectly from the Earth's surface. Mankind has known about the greenhouse gas effect since the early nineteenth century, thanks to the scientific contributions of Eunice Newton Foote, Joseph Fourier, John Tyndall, and Arrhenius (Ortiz and Jackson; Charlson, 1999). However, it was only in the mid-twentieth century that we acquired a more fundamental and mechanistic understanding of the impact of human activities on atmospheric composition and their repercussions on the terrestrial biosphere, air quality, and climate. For instance, photochemical Los Angeles smog, forming due to compounds in the automobile exhaust (Haagen-Smit, 1952; Whitby et al., 1972), depletion of the ozone layer in the Antarctic due to chlorofluorocarbons in spray cans (Molina and Rowland, 1974; Rowland and Molina, 1975) and the sulfurous London smog (Logan, 1953) were landmark events which triggered extensive research aimed at understanding the changing atmospheric composition. Now, in the early twenty-first century, we know more about the atmosphere than ever because of sophisticated measurement and modelling techniques. While the scientific understanding regarding the sources, fate, and impacts of trace gases is quite robust, our knowledge regarding tiny particles called aerosols and their impact on the hydrological cycle and human health continues to evolve.

This chapter provides detailed insights regarding the significance of studying aerosols and the reasons which make this field challenging.

## **1.2. What are aerosols? What is particulate matter?**

Aerosols are defined as a suspension of solid or liquid particles in a carrier gas with a negligible settling velocity (Boucher, 2015). In Atmospheric Sciences, the term *Particulate Matter* refers to the complex mixture of suspended particles in the atmosphere (or air). A particle

suspended in a medium experiences opposing gravitational and drag force (exerted by the carrier gas). The velocity of the particle when these forces achieve equilibrium is called the terminal settling velocity (Kulkarni et al., 2011). It is governed by the particle size, density, and shape and ultimately decides the atmospheric lifetime of a particle. For instance, the settling velocities of unit density spherical aerosol particles with a diameter of 0.01 and 10  $\mu\text{m}$  are  $6.76 \text{ E }^{-08} \text{ m s}^{-1}$  and  $3.07 \text{ E }^{-03} \text{ m s}^{-1}$ , respectively (Kulkarni et al., 2011). Due to low settling velocities, the lifetime of aerosols in the atmosphere varies between a few hours to a few weeks (Seinfeld and Pandis, 2016). However, for many aerosol particles, the mass and size cannot be measured easily, as the particles are in a constant physical and chemical equilibrium with the carrier gas, the atmosphere. Water vapour and several trace gases (e.g.,  $\text{NH}_3$ ,  $\text{SO}_2$ ,  $\text{NO}_x$ ,  $\text{HNO}_3$ ,  $\text{HCl}$ , and  $\text{H}_2\text{SO}_4$ ) are in a temperature, RH, and pH-dependent equilibrium with the solid /liquid phase of the aerosol.

While atmospheric aerosols are ubiquitous, there is immense variability in their shape, structure, sources, chemical composition, global spatial, and temporal distribution. They are intricately linked to the elements of the Climate system. In comparison to the vastly abundant air molecules, aerosols are trace species; however, they play a central role in regulating the atmospheric chemistry and physics, the biosphere, climate, and human health (Pöschl, 2005). Aerosols act as the site of heterogeneous chemical reactions in the atmosphere. Aerosols affect the precipitation patterns and the radiative budget of the planet. A large aerosol loading impairs visibility. They also cause adverse impacts on human health. Airborne aerosols laden with pathogens can lead to the transmission of several infectious diseases (Jones and Brosseau, 2015; Asadi et al., 2020).

While there is unequivocal evidence regarding the implications of greenhouse gases, the scientific understanding regarding the range of impacts aerosol particles can have on the climate and terrestrial biosphere continues to evolve (Andreae, 2007; Larson and Portmann, 2019; Sorensen et al., 2019). The next few sections shed light on the complexity of aerosol characterization.

### **1.3. Types of atmospheric aerosols**

There is immense diversity in the sources, size range, and formation mechanism of aerosol particles. This section highlights a few of the broad ways to classify atmospheric aerosols and discusses their physical and chemical properties.

Aerosol particles emitted from natural sources like deserts, oceans, volcanoes, natural forest fires, terrestrial vegetation, etc., are called natural aerosols. Aerosol particles emitted from human actions such as fossil fuel combustion, biofuel burning, agricultural, land clearance, and waste fires are termed anthropogenic aerosols.

Primary aerosol particles are directly emitted as particles in the atmosphere, for example, mineral dust, soot from combustion. They can be natural (marine aerosols, mineral dust, biogenic aerosols, forest fire smoke, volcanic dust, and cosmic dust) and anthropogenic (industrial dust, fossil fuel combustion, and biomass burning). The atmospheric loading of primary aerosol particles is regulated by the source strength, transport, and sinks.

Secondary aerosol particles are formed in the atmosphere through physicochemical transformations by a process called gas-to-particle conversion, which involves oxidation of gases into low volatility species, heterogeneous chemical reactions, and condensation onto pre-existing particles. Secondary aerosols can have both natural (dimethyl sulfate (DMS), terpenes) and anthropogenic precursors. The atmospheric loading of secondary aerosols is a function of precursor species, meteorological conditions, and atmospheric oxidants. Figure 1.1 categorizes the major aerosol types into natural and anthropogenic and reports their annual emission budgets.

Anthropogenic aerosols contribute only 10% to the overall aerosol loading; however, they are spatially concentrated over regions of human habitation (Tomasi and Lupi, 2016). The health effects of natural aerosols differ drastically from anthropogenic aerosols. For instance, while the inhalation of sea-spray aerosols is known to relieve asthma (Zajac et al., 2020) and inhibit tumour growth (Asselman et al., 2019), inhalation of ultrafine soot particles exacerbates cardiovascular and pulmonary disorders (Schraufnagel, 2020).

Population growth, rapid urbanization, and industrialization are responsible for further increasing the anthropogenic aerosol loading. Human actions have also altered the emissions of natural aerosols. Land-use changes have accelerated desertification and have facilitated dust mobilization due to plowing, mining, blasting activities, thereby increasing the mineral dust levels (Woodward et al., 2005).

The atmospheric aerosol loading is particularly high over the Indo-Gangetic Plain, home to one-seventh of the global population (Chapter 3). Wide disparities in the socio-economic conditions of the population, absence of stringent emission standards, large-scale development, complex emission sources (solid fuel for cooking, leaf litter burning, agricultural waste burning), and lack of high-density measurements together make pollution mitigation in this region challenging. There is an urgent need for granular mitigation policies to mitigate PM pollution in this region.

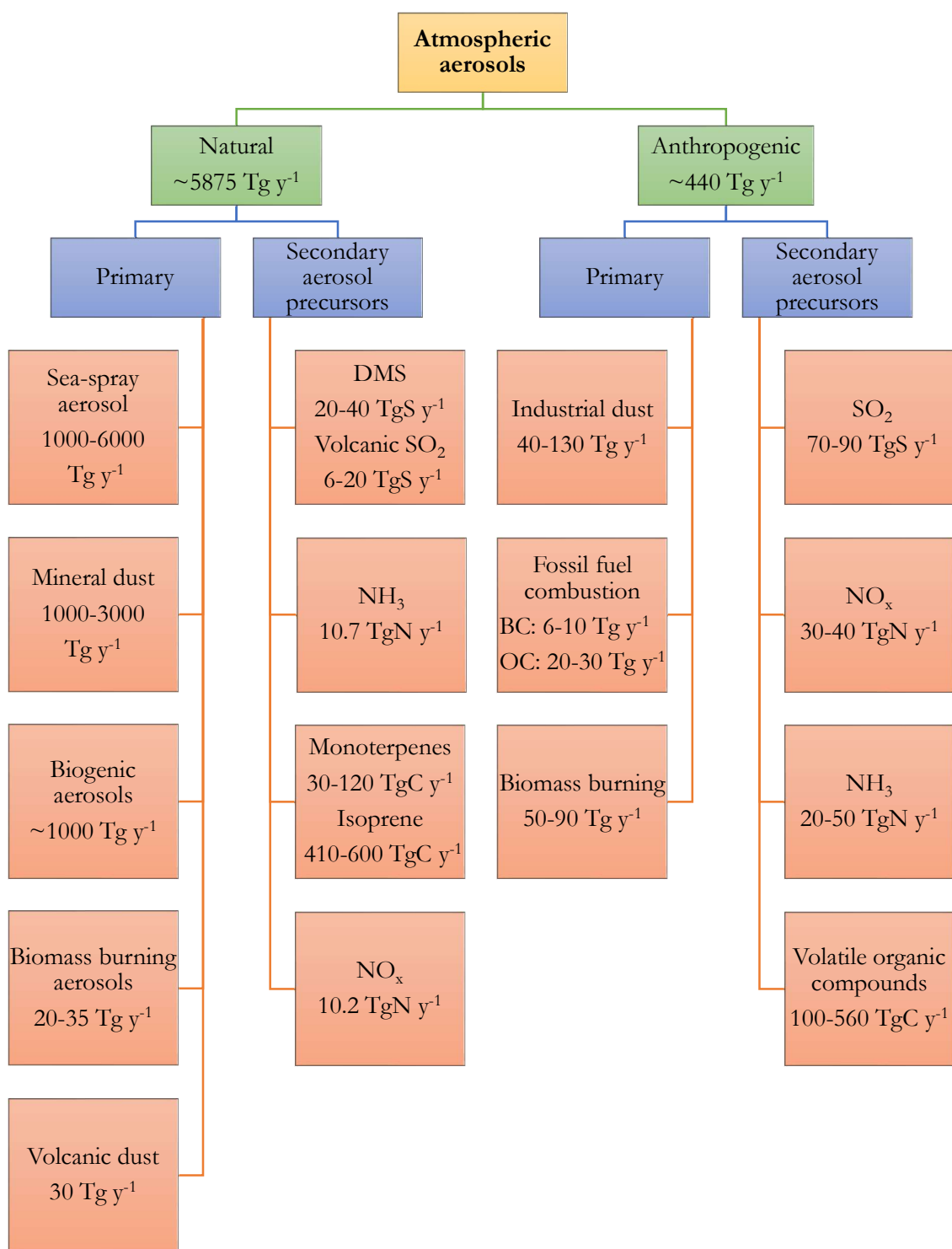


Figure 1.1: Approximate emission fluxes of the primary aerosols and the precursors to secondary aerosols classified based on the source [The estimates in this figure have been compiled from (Boucher, 2015) and (Tomasi and Lupi, 2016)]



## 1.4. Physicochemical properties of atmospheric aerosols

The fate of an aerosol particle in the atmosphere and its effect on human health is dictated by a combination of its physicochemical properties. The important physical properties are size, shape, and density. The concept of size/diameter is tricky for aerosols because of their complex shapes (Figure 1.2) and heterogeneous nature. Therefore, aerosols are assigned a size, termed as the equivalent diameter (Thomas and Charvet, 2017), based on a specific property or behaviour. It is the diameter of a unit density sphere with an identical value of a particular property (identical mass, identical specific area, identical volume, etc.) as that of the particle under consideration (Kulkarni et al., 2011).

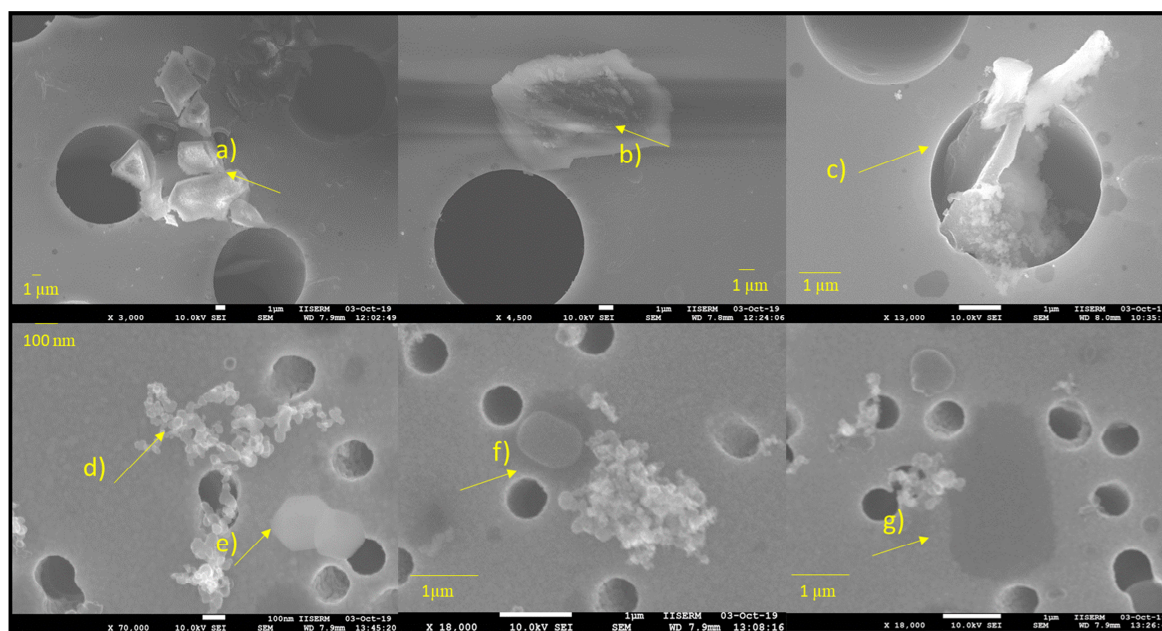


Figure 1.2: Scanning electron micrographs highlighting the diversity in aerosol particles collected at IISER Mohali showing a) Inorganic aerosol particle, b) dust aerosol, c) inorganic aerosol, d) soot particle, e) inorganic aerosol, f) Internally mixed SOA-inorganic particle, g) SOA droplet.

For instance, aerodynamic equivalent diameter is the diameter of a sphere with the same terminal fall velocity (due to gravity) as the particle under consideration. The aerodynamic diameter of a particle guides its behaviour in the respiratory tract and lifetime in the atmosphere. Therefore, global regulations on the exposure limit to particulate matter use this equivalent size. Optical equivalent diameter is the diameter of a polystyrene latex sphere with the same optical scattering properties as the particle under consideration. The volume equivalent diameter of the particle can be larger (e.g., in the case of highly absorbing aerosol) or smaller (in the case of a highly scattering aerosol) than the optical equivalent diameter.

Aerosol density is also an important parameter that affects dry deposition, lung deposition, and cloud scavenging. An aerosol may be externally mixed, such that chemically pure particles are

physically separated from other constituents, or it can be in the form of an internal mixture where each particle is composed of different chemical constituents (Lesins et al., 2002). The effective density of a porous or internally mixed aerosol particle is affected by its shape, morphology, and chemical composition. The density of aerosols varies as  $\sim 1.4 - 1.9 \text{ g cm}^{-3}$  (black carbon), 2 to  $2.7 \text{ g cm}^{-3}$  for mineral dust,  $\sim 1 \text{ g cm}^{-3}$  for primary biogenic aerosols,  $\sim 1.4 \text{ g cm}^{-3}$  for organic aerosols. Absorption of water reduces the effective density of aerosols when the wet diameter is used to measure particle size.

The chemical composition of aerosols affects the aerosol-radiation and aerosol-water interaction and ultimately impacts their environmental and health impacts. Broadly, the chemical composition of aerosols can be split into organic, inorganic, and elemental/black carbon constituents. Organic aerosols are composed of a wide variety of complex carbon-containing species like alkanes, alkenes, alcohols, aromatics, carboxylic acids, amines, etc (Boucher, 2015). They contribute roughly 50% of the total aerosol mass in the submicron range (De Gouw and Jimenez, 2009). Inorganic components in the atmospheric aerosols consist of sulfate, nitrate, chloride, ammonium, sodium, and chloride. Black carbon is the fractal aggregate produced from the inefficient high-temperature combustion of vegetation or fossil fuels.

## 1.5. Atmospheric processing and sinks of aerosols

Atmospheric aerosols are dynamic entities with short lifetimes (varying between few days to few weeks). During their life cycle, several ageing processes constantly modify their physical and chemical properties, which has important implications for the health and climate impacts.

Nucleation refers to a multi-step process of the transformation of individual gas molecules into a molecular cluster called the aerosol nucleus. Only a few nuclei that successfully achieve a critical radius contribute to new particle formation, while others evaporate back to the gaseous phase. While ample evidence of nucleation events was observed in the Finnish boreal forest, new particle formation rarely occurs in the Amazonian rainforest, and potassium-rich salts emitted by the fungi act as the aerosol seed there (Pöhlker et al., 2012).

Semi-volatile compounds may condense onto the surface of the pre-existing aerosol particles in a process called condensation, or two suspended particles could collide and join with each other in a process called coagulation. Condensation and coagulation contribute substantially to submicron aerosol growth. Aerosol particle size also evolves due to changing relative humidity in the atmosphere due to the reversible transfer of water molecules from the gas to particle-phase or vice versa. At deliquescence relative humidity (DRH), aerosol particle changes from a solid to a supersaturated aqueous solution and undergoes a rapid increase in size. Once the particle becomes

a droplet, it acquires additional mass from the aqueous-phase processing of water-soluble gases such as  $\text{NH}_3(\text{g})$ ,  $\text{SO}_2(\text{g})$ , and  $\text{NO}_2(\text{g})$  (Harris et al., 2014). The uptake properties of aerosols are strongly affected by the particle's chemical composition.

Akin to gases, organic and inorganic constituents of atmospheric aerosols are also subject to free radical initiated heterogeneous oxidation in the atmosphere, altering their composition, density, and hygroscopicity and further release toxic gases into the atmosphere (George and Abbatt, 2010).

As a combined result of atmospheric ageing processes, freshly emitted aerosols evolve from a primarily externally mixed state to an internally mixed state (Riemer et al., 2019). This has important implications for the radiative budget. For instance, studies have reported that the radiative forcing of an internally mixed soot particle is larger (factor varying between 1.05 to 3.5) when compared to an externally mixed soot particle (Jacobson, 2001; Shiraiwa et al., 2010; Cappa et al., 2012; Riemer et al., 2019). The hygroscopicity of aerosols also changes because of atmospheric processing. Laboratory studies have reported that either radical initiated oxidation or condensation of sulphuric acid vapors on hydrophobic soot particles makes them hydrophilic, as a result of which they can be scavenged by acting as cloud nuclei or agglomerating with existing liquid cloud droplets (Zuberi et al., 2005; Zhang et al., 2008).

The two major sinks of aerosols in the atmosphere are wet deposition and dry deposition. Wet deposition includes all processes by which aerosol particles are removed from the atmosphere in an aqueous form through hydrometeors (rain, snow, fog, or cloud). It can be split into in-cloud scavenging and below-cloud scavenging. In-cloud scavenging is the process when aerosol particles serve either as cloud condensation nuclei on which water condenses to form cloud droplets or merge in an existing droplet following impaction. Particles are subjected to aqueous-phase chemical reactions in the cloud, which alters their compositional properties. Several laboratory studies have hypothesized that aqueous-phase processing in clouds contributes significantly to secondary organic aerosols in the atmosphere. For instance, volatile organic compounds like isoprene undergo oxidation in the atmosphere to form hydrophilic compounds (aldehydes) that conveniently get condensed on CCN droplets. These compounds further get oxidized to form organic acids with low volatility in the clouds. Clouds that do not precipitate reevaporate back, releasing the chemically modified aerosols back into the atmosphere (Lim et al., 2005; Ervens et al., 2008). Below-cloud scavenging refers to the process when hydrometeors (rain droplets, snow, or hailstorm) physically remove aerosols from the volume of air swept by them while falling from the atmosphere (Duhanyan and Roustan, 2011). There are wide spatial and temporal variations in the distribution of clouds and the amount of rainfall; therefore, the intensity of wet deposition is highly variable. Dry deposition (facilitated by atmospheric turbulence) and sedimentation

(gravitational settling) are the dominant sinks of atmospheric aerosols in the absence of wet scavenging (Wesely and Hicks, 2000; Boucher, 2015). Particle size and composition, nature of the depositing surface, and the scale of turbulence are the major factors governing dry deposition (Seinfeld and Pandis, 2016). Sedimentation is the dominant sink for coarse mode desert dust and sea-spray aerosols. Both dry and wet deposition are quick processes that lead to short atmospheric lifetimes for aerosol particles (a few days to weeks).

Changes in meteorological conditions such as wind speeds, temperature, relative humidity, and boundary layer height also affect aerosol properties. Aerosols emitted from the source can be advected to distant places by winds in the troposphere and long-range transport (Ramachandran, 2018). For instance, anthropogenic black carbon, which can affect the snow albedo, has been detected at pristine mountainous sites in the Himalayas and the Arctic (McConnell et al., 2007; Li et al., 2016). The atmospheric boundary layer, which evolves in response to differential solar heating between day and night, regulates the dispersion of aerosols. For example, to evade regulatory action, agricultural waste burning activities typically occur during the evening to night hours. A shallow boundary layer coupled with temperature inversion restricts the vertical transport of aerosols, thus posing health risks to human health (Ganguly et al., 2006). As relative humidity increases, hygroscopic aerosols swell up due to the condensation of water vapor. This process affects both the physical, chemical, and radiative properties of aerosols. The particle density of water-soluble aerosols decreases from  $\sim 1.80 \text{ g cm}^{-3}$  at 0% RH to  $1.12 \text{ g cm}^{-3}$  at 95% RH. (Ramachandran, 2018) (Nessler et al., 2005).

Thus, atmospheric processing, transport processes, short atmospheric lifetimes, and complex links between the physicochemical properties together make the study of aerosols and their impacts challenging.

## **1.6. Health impacts of atmospheric aerosols**

Exposure to particulate matter causes a substantial disease burden and reduces life expectancy (Pope III and Dockery, 2006; Kim et al., 2015; West et al., 2016). The Meuse Valley fog of 1930 and the London smog of 1952, two major events of the 20<sup>th</sup> century, decisively established the inextricable link between high particulate matter loadings and their adverse impact on human health (Logan, 1953; Nemery et al., 2001; Anderson et al., 2012). Since then, epidemiological and toxicological studies continue to provide evidence of respiratory, cardiovascular, and cerebrovascular morbidities arising from both acute and chronic exposure to particulate matter (Dockery et al., 1993; Pope III et al., 1995; Pope III, 2000b; Brook Robert et al., 2010; Anderson et al., 2012; West et al., 2016). Global modeling findings combined with

epidemiological response functions indicate that exposure to outdoor as well as indoor particulate matter is reported to increase the annual premature death burden by ~1.61 - 4.81 million (Lelieveld et al., 2015) and ~3.54 million (Lim et al., 2012) respectively worldwide.

Ambient particulate matter primarily enters the human respiratory system during inhalation through the nose or the mouth. The fate of an aerosol particle inside the respiratory system is decided by both the lung anatomy and the physicochemical properties of the particle (size, density, shape, chemical composition, electrostatic charge) (Prata, 2018). The human respiratory system can be broadly split into the extrathoracic (nasal-pharyngeal) and intrathoracic regions (consists of tracheobronchial and alveolar regions). Large-sized particles ( $d_{ac} > 10 \mu\text{m}$ ) are mostly filtered by the nose or the mouth and get mechanically removed during sneezing. Beyond that, deposition in the upper bronchial region happens primarily due to inertial impaction as particles with large momentum ( $d_{ac} < 10 \mu\text{m}$ ) continue to follow their trajectory even as the air stream changes direction, causing them to collide with the lung epithelium. A reduction in the air flowrate in the tracheobronchial region increases the residence time of particles allowing them to settle down under the influence of gravity, a process called sedimentation. As the air reaches the alveolar region, diffusion due to random Brownian motion is the major mechanism of particle deposition. Due to negligible air velocity in the alveolar region, small particles ( $d_{ac} < 0.5 \mu\text{m}$ ) behave akin to gas molecules and collide and stick to the surface due to diffusion.

For particles that travel beyond the extrathoracic region, phagocytosis and mucociliary clearance are the two dominant pathways adopted by the respiratory system to eliminate foreign material. The lung epithelia in the thoracic region are covered in tiny hair-like projections called cilia. These cilia nodes form the mucus, which entraps the foreign substance. The orchestrated movement of cilia then propels the mucus upwards, where it is either swallowed or expectorated. The efficiency of mucociliary clearance reduces as the particle size reduces. Phagocytosis is the dominant removal pathway in the alveolar region where the epithelia are not ciliated. Cells called macrophages (a type of white blood cell) encapsulate the foreign material and transport them to the tracheobronchial region, where they are coughed out or transported to the lymphatic system. However, the efficiency of phagocytosis decreases with particle size, and particles with a diameter less than 500 nm may remain deposited in the alveolar region for years as the human body does not have a mechanism to eliminate them.

The ultrafine particles (UFP), which are emitted mostly from anthropogenic sources, owing to their extremely small size ( $< 0.1 \mu\text{m}$ ), cross the air-blood barrier and get translocated to other organs causing damage. A recent study has reported the presence of black carbon particles in the fetal side of the placenta, which was previously considered impenetrable (Bové et al., 2019).

Magnetite nanoparticles that get formed during high-temperature combustion have also been detected in the human brain and cardiovascular system (Maher et al., 2016; Calderón-Garcidueñas et al., 2019). These particles can become activated in the presence of external magnetic fields and can potentially lead to neurodegenerative and cardiovascular disorders. Prolonged exposure to ultrafine particles impairs the ability of alveolar macrophages to eliminate pollutants, thus increasing the risk of future ailments (Lundborg et al., 2001; Renwick et al., 2001).

While most countries currently have air quality standards for PM<sub>10</sub> and PM<sub>2.5</sub>, there is a lack of consensus regarding what mass concentration of particulate matter can be earmarked as safe. India's annual air quality standard for PM<sub>2.5</sub> (annual exposure should be under 60 µg m<sup>-3</sup>) is about five times that of the USA (annual exposure should be under 12 µg m<sup>-3</sup>). A recent study reviewed the hospital admission rates in US cities which were always under compliance with the PM<sub>2.5</sub> air quality standard (annual exposure < 12 µg m<sup>-3</sup>) and reported that an increase in the exposure from 8 µg m<sup>-3</sup> to higher levels led to a 21% increase in respiratory ailment related admission rates (Makar et al., 2017). This shows that mass concentration and size alone are insufficient to gauge the health effects arising from exposure to particulate matter (Al-Kindi et al., 2020), and additional parameters such as chemical composition may be necessary to appreciate the bigger picture.

Using cellular and molecular assays, Park et al. (2018) studied the differential toxicities of PM<sub>2.5</sub> particles generated from combustion and non-combustion sources and assigned toxicity scores based on cellular response. They concluded that for the same mass concentration, diesel exhaust particles were the most toxic (followed by gasoline exhaust and biomass burning), while desert dust, ammonium sulfate, and ammonium nitrate were the least toxic. The link between particulate matter and its health impact is a complex interplay of particle size, surface area, and chemical composition, often referred to as the "PM hydra." While the size of the particles decides the location of deposition and the clearance mechanism, chemical composition affects the biological impact.

Bioaerosols, polycyclic aromatic hydrocarbons (PAHs), quinones, dioxins, etc., present in the particulate matter are known to damage the DNA. At the same time, heavy metals trigger oxidative stress and inflammation. Addressing and mitigating the health impacts arising from exposure to particulate matter requires a multi-disciplinary approach from a global to a molecular scale involving atmospheric chemists, epidemiologists, public health specialists, toxicologists, biochemists, and policymakers. From the point of view of air quality standards, it is not essential to simply reduce the ambient mass concentration of particulate matter but also understand how the disease burden fluctuates with varying chemical composition and size distribution.

## 1.7. Climate impacts of atmospheric aerosols

Aerosol particles impact the climate by altering the incoming solar radiation and outgoing infrared radiation. The optical properties of aerosol particles are a complex function of their size, structure, chemical composition, and mixing state and decide if they have a net cooling or a warming effect. The climate impact due to any environmental perturbation, such as greenhouse gas emissions, or in this case, aerosol particles, is generally quantified in terms of radiative forcing. It is defined as the change in net downward solar radiation in response to a perturbation while keeping the temperature at the surface and in the troposphere fixed. A positive radiative forcing implies a warming impact, while negative denotes a cooling effect. In its 5<sup>th</sup> Assessment Report, the IPCC introduced the concept of effective radiative forcing (ERF), which is defined as the net change in downward solar flux in response to a perturbation while allowing for the temperature and albedo to adjust in response to the perturbation. ERF is a better predictor of resultant surface temperature.

Aerosol interactions can be broadly split into aerosol-radiation and aerosol-cloud interactions. Aerosol-radiation interactions consist of the direct effect (a net reduction in the solar flux reaching the Earth's surface due to scattering and absorption of radiation by aerosols) and the semi-direct effect (the readjustment of the temperature profile in the troposphere in response to the absorption of radiation by aerosol particles). Together, the radiative forcing from the direct and semi-direct effects is called  $ERF_{ari}$  or the Effective Radiative Forcing due to aerosol-radiation interactions. Cloud-active aerosol particles also impact the precipitation patterns by altering the microphysical and radiative properties of clouds (Stevens and Feingold, 2009). The aerosol-cloud interactions can be split into the first indirect effect (decrease in droplet size and increase in cloud reflectivity due to rise in CCN concentration) and the second indirect effect (change in the cloud properties and precipitation patterns to adjust to changes in CCN). Together, the effective radiative forcing from the two indirect effects is represented as  $ERF_{aci}$ , which stands for Effective Radiative Forcing due to aerosol-cloud interactions.

IPCC AR5 estimated  $ERF_{ari}$  and  $ERF_{aci}$  to be approximately  $-0.45 \text{ W m}^{-2}$  ( $-0.95$  to  $+0.05 \text{ W m}^{-2}$ ) and  $-0.45 \text{ W m}^{-2}$  ( $-1.2$  to  $+0.0 \text{ W m}^{-2}$ ), respectively (Myhre et al., 2014). The current understanding points towards a net cooling effect from aerosols, albeit with high uncertainty. Aerosol particles offset about 40% of the warming caused due to greenhouse gases (Allen et al., 2019).

The ERF estimates conceal the high variability in the climate effects of aerosol particles. For instance, the cooling effect of aerosols is not homogeneously observed across the globe when compared to the uniform warming effect caused by the well-mixed greenhouse gases. The  $ERF_{ari}$

is more pronounced in the midlatitudes and densely populated regions of the world than marine and polar regions.  $ERF_{\text{ari}}$  is also highly dependent on the chemical composition of aerosols. Sulfate aerosol particles have an average effective forcing of  $-0.4 \text{ W m}^{-2}$  ( $-0.6$  to  $-0.2 \text{ W m}^{-2}$ ). In comparison, soot particles result in an average effective forcing of  $+0.4 \text{ W m}^{-2}$  ( $+0.05$  to  $+0.8 \text{ W m}^{-2}$ ). These net estimates or radiative forcing undergo continuous revisions while we gain a better understanding of the aerosol processes through more observations and modeling studies. For instance, the  $RF_{\text{ari}}$  reduced from  $-0.50 \text{ W m}^{-2}$  ( $-0.90$  to  $-0.10 \text{ W m}^{-2}$ ) in AR4 [1750-2005] to  $-0.35 \text{ W m}^{-2}$  ( $-0.85$  to  $-0.15 \text{ W m}^{-2}$ ) in AR5 [1750-2011].

However, considering the adverse impact of aerosols on human health and air quality, most countries are taking targeted steps to mitigate ambient PM. Since aerosols have a net cooling impact, it is crucial to understand how cleaner air will impact climate (Andreae et al., 2005). For now, it's almost as if humans have dampened climate change through air pollution (Samset, 2018).

## 1.8. Research motivation and thesis outline

The Indo-Gangetic Plain in India faces severe air pollution. Rapid urbanization and economic growth have further depleted the air quality in this region. Mitigation of severe particulate matter pollution necessitates the formulation of science-based policies. However, the current efforts are vastly limited due to an inadequate understanding of the relative contribution of local, regional, and distant sources to the overall pollutant loading and the limited availability of high-density measurements. In this thesis, I have tried to focus on these issues using a combination of statistical tools and low-cost optical sensors.

The major scientific questions addressed in this thesis are:

- 1) What is the general climatology of the air masses arriving at Mohali?
- 2) What is the contribution of long-range transport to the particulate matter loadings at Mohali, a regionally representative site in north-west India?
- 3) How does long-range transport impact the percentage of days when the receptor site, Mohali, is not in compliance with the national ambient air quality standards of  $PM_{10}$  and  $PM_{2.5}$ ?
- 4) In a complex atmospheric environment with varied meteorology, what factors impact the accuracy of low-cost optical sensors when used for continuous ambient air quality monitoring?
- 5) How to correct the inaccurate particulate matter measurements from the low-cost sensors?
- 6) What is the contribution of paddy residue burning and domestic biofuel burning for heating purposes to the particulate matter loadings in north-west India?



Meteorology and air mass transport impact the particulate matter loading at the measurement site. Therefore, it becomes pertinent to understand the general trends in air-flow patterns. Question 1 has been addressed in Chapter 3 of this thesis using a collection of 72 h back trajectories for Mohali. Further, Questions 2 and 3 have also been discussed in Chapter 3 by evaluating the impact of air-mass patterns on the particulate matter loadings and exceedance events observed in Mohali. This chapter sheds light on the nature of aerosol sources in the North-western Indo-Gangetic Plains by combining the measurements of particulate matter with several gas-phase tracers.

To meet the rising demand for open-source air quality data, low-cost sensors have permeated the consumer market with little to no field validation. However, their cost-effectiveness and compact size make them desirable candidates to integrate with the existing measurement facilities and achieve a denser network of real-time data. Relevant to these issues, Questions 4 to 6 have been addressed in Chapters 4 and 5 of this thesis.

Question 4 seeks to assess the suitability of a low-cost sensor, Laser Egg (customized and calibrated in China), for continuous ambient monitoring by comparing its measurements with an EPA-approved  $\beta$ -attenuation monitor under varied meteorological conditions. I found that the Laser Egg sensor was challenging to use and had poor accuracy. Therefore, I evaluated the Airveda low-cost sensor, which was customized and calibrated specifically for Indian conditions. This sensor was found to be more robust and accurate and thus was used for field deployment.

Question 5, focusing on approaches to correct the inaccurate particulate matter measurements from the low-cost sensors, has been addressed in both Chapters 4 and 5 of this thesis. Specifically, corrections for hygroscopic growth and inlet losses have been mentioned in Chapter 4. Chapter 5 presents a machine learning approach based on the Random Forest algorithm to correct the sensor measurements.

Question 6 has been addressed in Chapter 5, which presents the field application of low-cost sensors. A set of Airveda particulate matter sensors have been used to quantify the contribution of paddy residue burning and domestic biofuel burning for heating purposes to the particulate matter loadings in north-west India.

# Chapter 2

---

## **Quantitative measurement of ambient particulate matter (PM) and approaches to link PM measurements with air-mass history**

Quantification of the physicochemical properties of aerosols (particle mass distribution, number distribution, shape, chemical composition, optical properties, hygroscopicity, mixing state) is essential to fully understand their impact on human health, air quality, and climate. Aerosols properties are dynamic in nature and undergo constant evolution and atmospheric processing. Aerosols also provide surfaces for heterogeneous chemical reactions, which alters the properties of both the reactant and the participating aerosol. Aerosol composition is also affected by land use and land cover (LULC), dominant fuel types, biodiversity, agricultural practices, meteorological conditions, and emission standards of the origin site. This substantial spatio-temporal heterogeneity in aerosol properties makes their accurate quantification extremely challenging. Further, the aerosol loading at any given site is affected by the sources, sinks, and meteorological conditions. Winds can transport the pollutants emitted at a particular site to places downwind. Therefore, assessing the link between air mass history and the locally measured particulate matter loading is crucial to fully understand the factors at play. This chapter discusses the analytical techniques to quantify particulate matter and methods to link air-mass history with the locally measured pollutant concentration.

### **2.1. Analytical techniques to quantify ambient Particulate Matter**

The analytical techniques to characterize and quantify aerosols can be broadly classified into: a) Offline techniques b) Online techniques (Figure 2.1). In offline techniques, the aerosol sample is collected on a filter using a suitable instrumental setup. The sample is then extracted and measured using a suitable analytical technique. In online techniques, the sample is collected and measured simultaneously in near real-time. This category can be further classified into extractive and external sensing techniques. In extractive sampling, the aerosol sample needs to be transported inside the instrument chamber for measurement (examples  $\beta$ -attenuation analyzers, tapered elemental oscillating microbalance), while in external sensing techniques, the aerosol is directly

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

quantified in its natural innate state in the environment (example: Light Detection and Ranging instrument, sunphotometer, satellite-based remote sensing, etc.). Although offline techniques permit several sophisticated aerosol measurements, their temporal resolution is poor. Such instruments provide a time-averaged measure of the aerosol property being quantified. On the other hand, real-time measurements from the online instruments furnish excellent temporal resolution. Online analyzers offer an edge over offline counterparts when linking emission activities or air mass history to the aerosol composition.

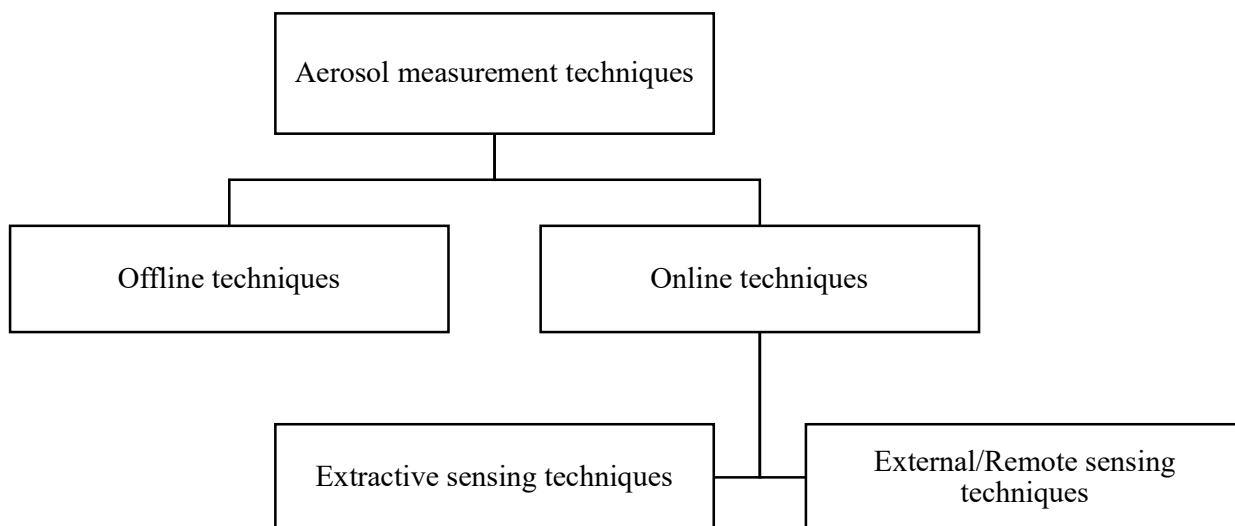


Figure 2.1: Broad classification of analytical techniques for the characterization of ambient particulate matter

The size of an aerosol particle is, by far, one of the most important properties that affect the lifetime, optical properties, fate in the lungs, and, therefore, health impacts. The size of ambient aerosols spans a wide range of approximately five orders of magnitude from  $0.001\mu\text{m}$  (freshly nucleated molecular clusters) to  $100\mu\text{m}$  (raindrops and large dust particles), making it nearly impossible for a single measurement technique to effectively quantify all types of aerosols. For instance, a scale used for measuring the weight of humans (in kilograms) cannot provide adequate accuracy to measure the weight of chemical compounds (in milligrams) in a laboratory. Therefore, the quantification of ambient aerosols is performed using a suite of techniques complementary to the size range of interest.

Measurements of aerosol mass concentration are essential for regulatory and health purposes. The human respiratory tract is an aerodynamic classifying system for airborne particles (Papastefanou, 2008). Therefore, the air quality standard for exposure to particulate matter is based on the mass

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

concentration of particles below the aerodynamic diameter of 10  $\mu\text{m}$  and 2.5  $\mu\text{m}$  for most countries, including India. Mass concentration can be quantified using both manual or offline (filtration followed by gravimetric analysis) and automated methods ( $\beta$ -attenuation, TEOM). In my thesis, I probe the link between the observed aerosol mass loading and air mass history. The PM mass concentration can be quantified using both online and offline techniques. However, the former provides an edge over the latter in terms of better temporal resolution. Generally, measurements with an hourly or finer temporal resolution are needed to effectively resolve the impact of sources on the observed mass loading. Offline instruments only provide daily average measurements, which eclipse the effect of contributing factors. For instance, the contribution of long-range transport dominates during the afternoon hours when the boundary layer height is maximum or during the late-night to early-morning hours when the contribution of the local sources is minimum. Therefore, isolating the contribution of long-range transport would necessitate the availability of PM measurements with at least an hourly temporal resolution, which offline techniques fail to provide.

Cloud condensation nuclei (CCN) counter and condensation particle counters (CPCs) are suitable to perform real-time counting of small-sized particles (10 nm to 3  $\mu\text{m}$ ). In CPCs, particles are first grown via the condensation of supersaturated vapor of solvents like n-butyl alcohol or water. On the other hand, CCN counters specifically report the particles that get transformed into cloud droplets due to water condensation at a specific value or range of supersaturation. A differential mobility analyzer (DMA) resolves 3 nm to 1  $\mu\text{m}$  sized aerosol particles based on their mobility in electric fields. The particle's electrical mobility depends on the particle size, charge on the particle, and viscosity of the gas in which the particle is suspended. The mobility equivalent diameter is equal to the volumetric diameter only for spherically shaped particles. Integrating a CPC with DMA into a single instrument called the Differential Mobility Particle Sizer (DMPS) gives a high-resolution measurement of the size-resolved number-distribution.

Optical particle counters (OPCs), working on the principle of light scattering, are powerful tools for particle sizing and counting in the size range of 0.5  $\mu\text{m}$  to 10  $\mu\text{m}$ . These instruments are factory calibrated using polystyrene latex spheres. The optical equivalent diameter of atmospheric particles has complex dependencies on the particle shape, chemical composition, morphology, and refractive index. OPCs can provide non-invasive high temporal resolution measurements of the size composition of aerosol particles. After laboratory-based calibrations, OPCs can also provide the PM mass concentration subject to assumptions on the aerosol density and scattering characteristics. In my thesis, I demonstrate the applicability of low-cost optical sensors towards achieving high-density PM monitoring. Aerodynamic particle sizers resolve and count the aerosol

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

particles based on their aerodynamic equivalent diameter. They are suited for particles in the range of 0.5 to 20  $\mu\text{m}$ . The aerodynamic equivalent diameter is dependent on the particle size, density, and shape and independent of its optical properties.

The bulk chemical composition of aerosol particles can be analyzed using both offline and online techniques. In offline methods, the aerosol samples are passed through an impactor and collected on a plate or passed through a filter mounted in a filter holder. The substrate is then extracted and analyzed using a suite of chromatographic and mass spectrometric techniques. For instance, the major organic and inorganic ions can be analyzed using Ion Chromatography; the elemental and organic carbon can be quantified using the EC/OC analyzer. Volatilization of semi-volatile species and the absorption of organic species on the filter can result in negative and positive artifacts, respectively. Aethalometer is an online method that reports the mass concentration of black carbon based on the Beer-Lambert's law.

Microphysical properties of aerosols such as hygroscopicity alter their optical and health properties. The setup to analyze the water absorption properties of aerosols consists of twin DMAs. The first DMA allows dry particles of a specific size into a humidifier where aerosols equilibrate with the water. These swollen particles are then passed through a second DMA to obtain a wet size distribution. This complete setup is commonly referred to as the Humidified Tandem Differential Mobility Analyzer (HTDMA).

The aerosol mass spectrometer is an advanced instrument capable of providing real-time measurements of the size and chemical composition of non-refractory (exhibit rapid evaporation at 600  $^{\circ}\text{C}$ ) sub-micron aerosol particles.

## **2.2. Federal Reference and Federal Equivalent Methods to measure the mass concentration of particulate matter**

Measurements of aerosol mass concentration are essential for regulatory purposes. For instance, the air quality standards of most countries are defined in terms of the mass concentration below a specific aerodynamic diameter.  $\text{PM}_{10}$  is the thoracic fraction, while  $\text{PM}_{2.5}$  can even reach the alveolar regions. Therefore, most studies define exposure in terms of mass concentration in this size range. Research facilities or regulatory bodies use either Federal Reference Methods or Federal Equivalent Methods to measure airborne PM. This certification is provided by the United States Environmental Protection Agency based on the Code of Federal Regulations, 40 (EPA, 2015a). The aerosol measurements mentioned in this thesis deal with the mass concentration of particles with an aerodynamic diameter smaller than 2.5  $\mu\text{m}$  and 10  $\mu\text{m}$ , respectively.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

**Federal reference methods:** The primary standard reference methods to measure PM<sub>10</sub> and PM<sub>2.5</sub> are based on gravimetric measurements. They primarily comprise a size-selective inlet, a vacuum pump, and pre-conditioned and pre-weighed 47 mm filters. A user manually inserts an empty filter inside the filter cassette. After that, the inlet selectively allows PM<sub>10</sub> (or PM<sub>2.5</sub> depending upon the inlet) particles to be collected on a pre-weighed and pre-conditioned filter using a vacuum pump (flow rate of 16.67 LPM) for 24 hours. The filter is removed from the analyzer and conditioned and weighed again. The total PM mass collected on the filter divided by the volume of sampled air provides the average PM mass concentration in  $\mu\text{g m}^{-3}$  on the day of sample collection.

**Federal equivalent methods (FEMs):** FEMs are classified into different categories based on their compliance with the FRMs.

**Class I FEM:** Class I FEMs are identical to FRMs in most aspects pertaining to sampling duration, flow rate, inlet specifications, temperature requirements, etc. They are characterized by minor changes relating to sequential sampling to reduce operator intervention. For instance, Thermo Scientific Partisol 2025 Sequential Sampler belonging to Class 1 FEM consists of 16 filter cassettes, allowing for two weeks of unattended daily sampling of particulate matter. Few other examples of commercial Class I FEMs are BGI PQ200-VSCC, PQ200A-VSCC, Patachnick Partisol Model 2000, Thermo Electron RAAS2.5 (100, 200, or 300), etc.

**Class II FEM:** Class II FEMs also perform a 24h sampling on a filter through a size-selective inlet; however, they differ substantially in design when compared to the FRM. Examples include dichotomous samplers, high volume PM<sub>2.5</sub> samplers. The word “high” in a high-volume sampler implies that such analyzers sample  $\sim 1500 \text{ m}^3$  of air in 24 hours compared to reference methods that sample only  $24 \text{ m}^3$  of air per day. The word “dichotomy” means division or a split into two. Dichotomous samplers split total PM<sub>10</sub> into coarse (PM<sub>10-2.5</sub>) and fine (PM<sub>2.5</sub>) fractions using a virtual flow impactor. Few examples of commercial Class II FEMs are Thermo Scientific Partisol-D 2000 and Partisol 2025-D dichotomous samplers, which can simultaneously measure PM<sub>2.5</sub> and PM<sub>10-2.5</sub> (EPA, 2015b).

**Class III FEM:** Class III equivalent method comprises instruments that can provide PM measurements at an hourly or lower temporal resolution. The 24h mean of PM measurements from Class III FEMs is considered equivalent to the daily average measurement reported by an FRM. Most Class III FEMs, which report mass concentration, are either continuous or semi-continuous. They rely on  $\beta$ -attenuation, harmonic oscillation (Tapered elemental oscillating microbalance, TEOM), or optical methods (light scattering) to continuously detect and quantify particulate matter in real-time or near-real-time. The inlets of FEMs follow the same specifications as that of FRMs;

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

however, they do not need filter collection, conditioning, and gravimetric measurement in a laboratory. Few examples of commercial Class III FEMs are Met One BAM 1020 and 1022, Teledyne 602 Beta, Environmental S.A. MP101M, and Grimm EDM 180.

A TEOM instrument is referred to as the “true gravimetry-based” measurement technique, which can report the mass concentration of ambient PM in near real-time. It has two major components, the inlet assembly, and the measurement chamber. The inlet assembly is identical to that of the Federal Reference Method analyzer with a size-selective inlet, a pump operating at a constant flow rate of 16.67 LPM, and a heated inlet to regulate the sample RH. Ambient PM is collected on an exchangeable filter cartridge kept atop a hollow tube (tapered element), continuously oscillating at its natural frequency. This natural frequency ( $f_o$ ) is dependent on the spring constant ( $K$ ), and the mass of the spring ( $m$ ) and is mathematically represented as:

$$f_o = \sqrt{\frac{K}{m}} \quad (1)$$

Deposition of ambient aerosol ( $\Delta m$ ) on the filter cartridge alters the net mass ( $m + \Delta m$ ) which leads to a change in the oscillation frequency  $f_1$  of the element. An electronic control unit senses this change in the frequency and uses it to compute the mass of particulate matter deposited on the filter as:

$$\Delta m = K \left[ \frac{1}{f_o^2} - \frac{1}{f_1^2} \right] \quad (2)$$

### **2.3. PM monitoring at IISER Mohali using 5014i Beta Continuous Ambient Particulate Monitor**

The reference PM<sub>10</sub> and PM<sub>2.5</sub> mass concentrations reported in this thesis were acquired using two separate  $\beta$ -attenuation monitors (5014i Beta Continuous Ambient Particulate Monitor, Figure 2.2) installed at the IISER Mohali Atmospheric Chemistry Facility (30.667 °N, 76.729 °E; 310 m a.m.s.l.). This model has been approved by the US Environmental Protection Agency (US EPA) as a Federal Equivalent Method (Certification number: EQPM-1102-150 for PM<sub>10</sub> and EQPM-0609-183 for PM<sub>2.5</sub>) for 24-hour average PM measurements. The monitors report the real-time mass concentration of ambient PM at a temporal resolution of 1 minute. Both PM<sub>10</sub> and PM<sub>2.5</sub> monitors have similar detection principles but differ in their sampling inlets.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

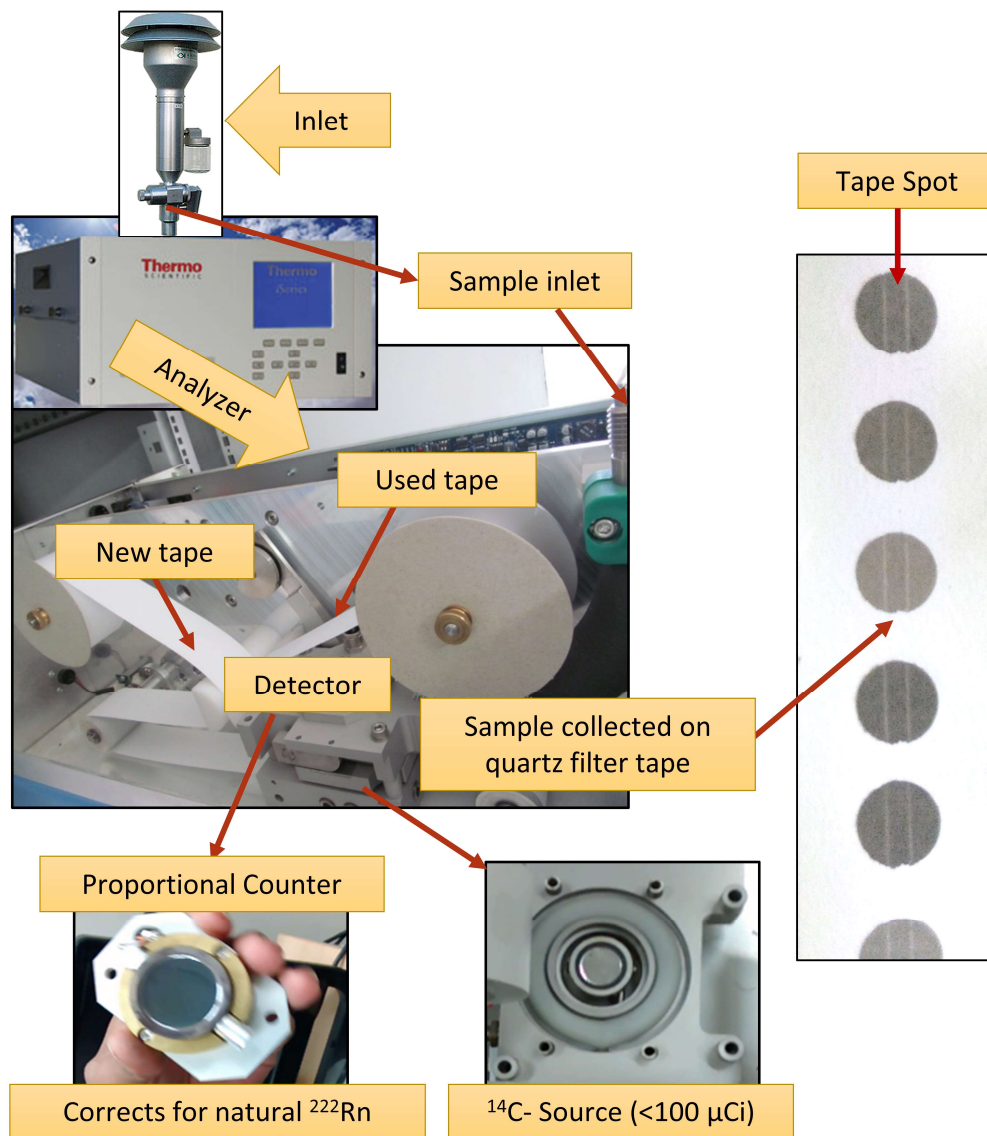


Figure 2.2: Schematic of the 5014i  $\beta$ -attenuation analyzer showing the major internal parts

5014i uses a low-volume US EPA  $\text{PM}_{10}$  inlet comprising of a single-stage size-selective conventional inertial impactor made of anodized aluminium. The top is louvered (curved) at an angle of  $45^\circ$  to the vertical over the inlet mesh (meant to block insects) to reduce the entry of windblown raindrops. Particle-laden air enters the vertically mounted omnidirectional inlet head through the mesh at a flow rate of 16.67 LPM, where a deflector cone directs the airstream to the acceleration nozzle in the upper plenum. Particles with  $d_a > 10 \mu\text{m}$  get collected on the impaction plate by virtue of their higher inertia. In comparison, smaller particles ( $d_a < 10 \mu\text{m}$ ) move upwards along with the airstream and fall into the lower plenum via three-vertical elutriator (particle collection) tubes. Rainwater accumulated on the impaction plate gets collected in a glass bottle attached outside the inlet via a small drain.



## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

PM<sub>2.5</sub> inlet system consists of a dual-stage conventional and cyclone inertial impactor. A Very Sharp Cut Cyclone, TM (BGI), which separates the fine particles (PM<sub>2.5</sub>) from the coarse particles (PM<sub>10-2.5</sub>), is attached at the bottom of the louvered PM<sub>10</sub> US EPA inlet. The air stream from the louvered PM<sub>10</sub> inlet containing particles <10 µm d<sub>a</sub> impinges tangentially on an inner cylindrical surface of the VSCC. It then spirals downwards and then back upwards after reversing direction in the conical section. Only fine particles (<2.5 µm d<sub>a</sub>) can follow the airstream and exit the VSCC via a central tube owing to their lower inertia. In contrast, coarse particles get collected on cylindrical and conical walls from where they fall off into a container referred to as “grit pot.”

After appropriate sizing, the airstream carrying PM<sub>10</sub> or PM<sub>2.5</sub> particles enter the dynamically heated section of the inlet aimed at reducing the particle-bound water. An external assembly shielded from radiation measures the real-time RH and temperature of the ambient air. The heating system consisting of a resistor regulates the RH of sample flow at a pre-set threshold of 40% by modulating its power whenever the ambient RH reported by the external sensor exceeded 40%.

### **Detection Principle**

After passing through a size-selective heated inlet, PM<sub>10</sub> and PM<sub>2.5</sub> particles enter the detection chamber and get deposited on a quartz filter tape (Figure 2.2) in the form of a circular spot placed between a radioactive <sup>14</sup>C source (radioactivity <100 µCi) at the bottom and a proportional detector at the top. The <sup>14</sup>C source, with a half-life of 5700 years, undergoes spontaneous radioactive decay into stable <sup>14</sup>N and, as a result, continuously emits β particles (electrons).



This beam of β particles then travels upwards through the filter tape with an accumulated PM layer. The proportional detector on the top quantifies the incoming flux of electrons.

The number of β particles which get transmitted through this thin layer of particulate matter sample deposited on the filter tape decreases exponentially with an increase in the thickness of the collected sample, which can be represented as:

$$I = I_0 e^{-\mu x} \quad (4)$$

and rearranged to give

$$x = \frac{1}{\mu} \ln \left[ \frac{I_0}{I} \right] \quad (5)$$

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

Here,  $I_o$  is the incident flux of  $\beta$  particles through a particle-free filter tape spot,  $\mu$  is the mass absorption coefficient, and  $x$  is the thickness of the particulate matter sample. The mass coefficient was factory calibrated for PM analysers at  $\sim 7000 \text{ kg m}^{-2}$ .

Mass concentration (c) of ambient PM is then computed by the analyzer as:

$$c \left( \frac{\text{kg}}{\text{m}^3} \right) = \frac{A \text{ (m}^2\text{)}}{Q \left( \frac{\text{m}^3}{\text{s}} \right) \times t \text{ (s)}} \times x \left( \frac{\text{kg}}{\text{m}^2} \right) \quad (6)$$

where  $A$  is the area of filter tape spot,  $Q$  is the volumetric flow rate ( $1000 \text{ m}^3 \text{ s}^{-1}$ ),  $t$  is the time step interval generally kept at 60 seconds, and  $x$  is the mass absorption coefficient. The integrated count rate (I) of the  $\beta$  particles and the sample flow rate (Q) regulated via a proportional valve are continuously measured. The mass concentration is reported at a temporal resolution of 1 minute. After every 8 hours of sampling on a spot, the instrument firmware automatically stops the pump system. A tape motor assembly pulls a fresh filter tape spot forward. Additionally, automatic spot change also happens if  $1500 \mu\text{g}$  of PM mass has deposited on the spot or the flowrate has decreased by more than 5%. At every spot change, the detector measures  $I_o$ , which is the intensity of  $\beta$  particles transmitted through particle-free glass fiber filter tape.

A potential artifact of estimating PM mass using the attenuation of  $^{14}\text{C}$   $\beta$  particles arises from the  $\beta$  particles emitted by the daughter nuclides of Radon 222, often found attached to airborne particulate matter.  $\beta$  particles from this source, in addition to the  $^{14}\text{C}$   $\beta$  particles, would result in an underestimation of the aerosol mass, particularly at low PM mass concentrations. To account for this artifact, the proportional detector used in 5014i also possesses the capability to quantify alpha particles. Naturally occurring noble gas, Radon 222 emits an alpha particle and other nuclides as part of its radioactive decay cycle. The detector measures this natural alpha activity of the deposited aerosol and uses it to estimate the  $\beta$  interference from the daughter nuclide emissions, which are subtracted from the total count rate to provide an accurate measurement.

### **Quality Assurance/Quality Control**

The following steps were undertaken as part of the quality assurance/quality control of the PM measurements acquired from the Thermo Fischer Scientific 5014i  $\beta$ -continuous ambient particulate monitors installed at the IISER Mohali

- i. All instrumental diagnostic parameters such as flow, pressure (flow, vacuum, and barometric), detector counts, RH (ambient and sample RH), and temperature (ambient and flow) were regularly monitored to obtain a timely indication of any issue pertaining to the instrument performance and perform quick troubleshooting to minimize data losses.
- ii. As a part of routine maintenance, inlets of both  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  were disassembled and cleaned once every two months.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

- iii. The mass absorption coefficient of PM analyzers was calibrated bi-annually using a pair of NIST traceable foils. In a typical calibration experiment, a null/blank foil (would cause the same attenuation as that of a particle free filter tape) and a foil which would result in the same attenuation as a known fixed mass (generally varying between 1000 to 2500  $\mu\text{g}$ ) are introduced sequentially in the instrument which then re-calculates the mass absorption coefficient from the observed  $\beta$  counts and the known value of mass. Typically, the old and new mass coefficients varied by less than  $\pm 10\%$ .
- iv. Periodically, the flow coefficient in the instrument was adjusted if the measured value of flow at the inlet of the PM instrument was found to be different from 16.67 LPM. The pump of the PM analyzer was serviced, and spare parts were replaced when the flow measurements were found to be unstable.
- v. Ambient temperature and RH calibrations were performed by adjusting for the offset observed between corresponding measurements from a co-located MetOne meteorological station.
- vi. Occasionally, insects would make their way past the mesh and get trapped inside the optical bench. Such instances could be identified from a sudden increase in the observed mass concentration of up to  $\sim 1500 \mu\text{g m}^{-3}$  with no accompanying increase reported by the other PM analyzer and no indication of a dust storm. Old filter tape spots were inspected for remnants of the insect body, following which the optical assembly was opened and cleaned cautiously.

In my thesis, I use the PM measurements from a spatially fixed FEM analyzer to investigate the link between air-mass history and the observed PM composition. Combining the PM measurements with back-trajectories gave us a rough idea about the potential location of sources on a synoptic scale. However, resolving the exact location of diverse emission sources was not possible with a single instrument, and granular measurements with higher spatio-temporal density were needed.

### **2.4. Optical methods to characterize atmospheric aerosols**

A high-density monitoring network is a pre-requisite for enhancing the spatio-temporal resolution of PM measurements. However, despite high accuracy, deploying FRMs or FEMs for high-density ambient monitoring is unfeasible given the exorbitant financial and infrastructural demands. For instance, Punjab, a state that routinely sees high PM pollution, has only seven government-regulated monitoring stations that provide real-time data. The current measurement

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

facilities are inadequate to resolve the location of hotspots, estimate exposure, and verify the success or failure of a mitigation policy. Therefore, a low-cost measurement alternate is needed.

Recently, the sudden demand for open-source air quality data by concerned citizens spurred technological advancements in the field of low-cost sensors. These sensors rely on optical techniques and detect particles via laser-based light scattering. They also provide prompt feedback on changes in air quality. Although low-cost, rapid response time, and compact size make them fit for high-density monitoring, their accuracy remains a significant concern. Such sensors went directly from the factory or place of development to the end-users with just a laboratory calibration of a brief field evaluation. Further, a standard calibration protocol or a regulatory body that assesses these sensors' performance does not exist.

Using optical techniques to derive the mass concentration of PM remains challenging because the interaction of light with aerosol particles is a complex function of the latter's size, shape, chemical composition, and mixing state of the particle. In this section, I provide a theoretical framework of the light scattering by ambient aerosol particles.

### **2.4.1. Theoretical background**

Light is a form of electromagnetic radiation characterized by electric and magnetic field vectors. Interaction of light waves with heterogeneous obstacles in their path (molecules, aerosol particles, droplets, etc.) brings about a change in the intensity and direction of propagation of light. For a light source of fixed wavelength, this change in properties of the incident light wave is a function of the particle's size, shape, chemical composition, and mixing state. Instruments working on optical techniques try to solve a complicated inverse problem by quantifying this change in light properties and using it to evaluate the concentration and size distribution of particles.

Gas molecules and aerosol particles suspended in the atmosphere scatter, absorb, and transmit solar radiation. Interaction of light waves (a form of energy) with particles (dielectric constant  $\neq 1$ ) causes the particle's dipoles to undergo oscillatory motion. The incident energy from the light wave can be transmitted as such, reradiated along with all different directions without loss of energy (scattering), and/or converted to thermal energy (absorption). Together, absorption and scattering are called extinction.

$$\textit{Extinction} = \textit{Absorption} + \textit{Scattering} \quad (7)$$

Scattering of light can proceed via reflection, refraction, total internal reflection, and/or diffraction. All atmospheric phenomena pertaining to the interaction of light with particles are based on elastic scattering, meaning that the wavelength of the incident light wave is equal to the wavelength of the scattered light wave.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

The scattering and absorption of the light energy by particles is a function of the following factors:

- i. **The size parameter  $\alpha$ :** The diameter ( $d_p$ ) and the wavelength ( $\lambda$ ) of incident light both affect the scattering properties of the particle. Together they are combined into a single quantity called the size parameter denoted as  $\alpha$  (Equation 8). Alpha ( $\alpha$ ) dependency determines the scattering regime of the particles. Particles with  $\alpha \ll 1$ , i.e., particles much smaller than the wavelength of light (for example, gas molecules in the air) undergo Rayleigh scattering; particles with  $\alpha \gg 1$ , i.e., particles much larger than the wavelength of light (cloud droplets) undergo geometric scattering while particles with  $\alpha \sim 1$ , i.e., particles comparable to the wavelength of light undergo Mie scattering. Most aerosols characterized using optical methods fall in the size range of 0.1 to 10  $\mu\text{m}$  shown by the red shaded box in Figure 2.3. Most optical instruments use a laser or a light-emitting diode with the wavelength in the visible to near-infrared range highlighted to simulate scattering as the blue box below. The intersection of the particle size and the wavelength of the light source lies in the Mie scattering regime (Figure 2.3).

$$\text{Size parameter} = \alpha = \frac{\pi d}{\lambda} \quad (8)$$

- ii. **The refractive index of the particle:** The refractive index of a medium is the ratio of the speed of light in a vacuum to the speed of light in the medium. It can be represented as a complex number (Equation 9). The real part represents the non-absorbing component, which affects scattering, and the imaginary part represents the absorbing component. Sodium chloride ( $m_{NaCl}^{\lambda=550\text{nm}} = 1.544 - 0i$ ), ammonium sulphate ( $m_{(NH_4)_2SO_4}^{\lambda=550\text{nm}} = 1.521 - 0i$ ) are non-absorbing aerosols as the imaginary part of their refractive indices is zero while soot is a highly absorbing aerosol particle ( $m_{soot}^{\lambda=550\text{nm}} = 1.95 - 0.79i$ ). The refractive index of a material is always normalized to the refractive index of the medium in which it is suspended; however, it can be neglected for ambient aerosols suspended in the air as its refractive index is almost unity ( $m_{air}^{\lambda=585\text{nm}} = 1.0029$ ).

$$\text{Refractive index} = m(\lambda) = n + ik \quad (9)$$

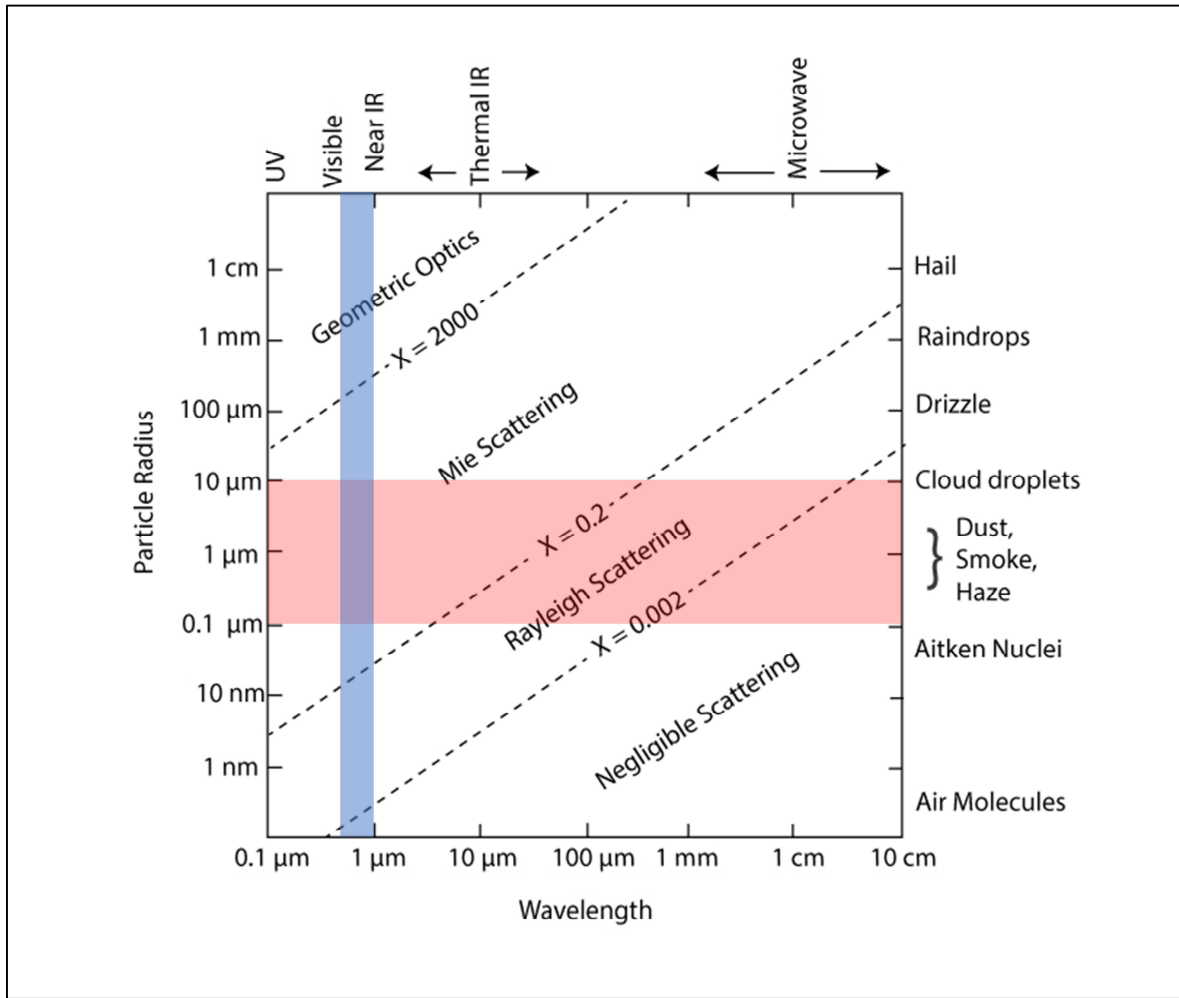


Figure 2.3: Figure showing different regimes of light scattering as a function of particle size and the wavelength of the incident light. [Adopted from (Brasseur and Jacob, 2017)]

- iii. **Particle shape:** Particle shape also significantly affects the scattering patterns (Piedra et al., 2019).

#### 2.4.2. The mathematical formulation of light scattering

The amount of energy per unit area intercepted by a detector/observer in a plane perpendicular to the direction of the wave propagation is called intensity ( $W m^{-2}$ ). The energy scattered ( $F_{sca}$ ) and absorbed ( $F_{abs}$ ) by the particle are proportional to the incident intensity, ( $F_o$ ) of the light wave, which can be represented mathematically as

$$F_{sca} = C_{sca}F_o \quad (10)$$

$$F_{abs} = C_{abs}F_o \quad (11)$$

where  $C_{sca}$  and  $C_{abs}$  are the single-particle scattering cross-section and single-particle absorption cross-section, respectively, with units of  $m^2$ . They represent the “area of the particle,” which interacts with the light wave in a plane perpendicular to the direction of propagation. As per energy

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

conservation (Equation 7), the total extinction can be written as the algebraic sum of the scattered and absorbed energy.

$$F_{ext} = C_{ext} F_o = (C_{sca} + C_{abs}) F_o \quad (12)$$

which implies

$$C_{ext} = C_{sca} + C_{abs} \quad (13)$$

Dividing both sides in Equation 13 with the geometric cross-section of the particle, A (units of  $m^2$ ) results in dimensionless quantities called extinction efficiency ( $Q_{ext}$ ), scattering efficiency ( $Q_{sca}$ ) and absorption efficiency ( $Q_{abs}$ ).

$$Q_{ext} = Q_{sca} + Q_{abs} \quad (14)$$

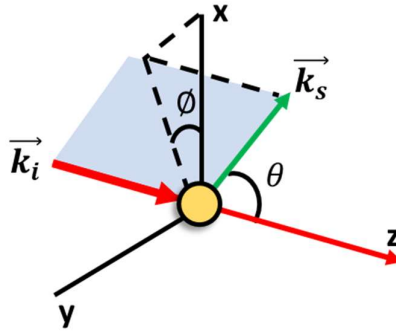


Figure 2.4: Basic scattering geometry depicting the incident wave, the object, and the scattered wave

Figure 2.4 shows a basic scattering geometry. The incident wave  $\vec{k}_i$  is traveling along the z-axis. The angle between the incident wave and the scattered wave is called the scattering angle  $\theta$ . The angular distribution of the scattered wave reveals important information about the particle size. Complete backscattering implies that all the scattering angle is  $\pi$  while total forward scattering implies that the scattering angle is 0. The plane containing the incident and the scattered wave is called the scattering plane. For the sake of simplicity, the total scattered light is split into two polarized components, one whose electric vector is perpendicular to the scattering plane ( $I_{\perp}$ ) and the other whose electric vector is parallel to the scattering plane ( $I_{\parallel}$ ).

Few important parameters to establish the scattering pattern are

- i. **Angular scattering phase function:** Angular scattering phase function describes the angular distribution characteristics of the scattered wave. It is defined as the probability that an incident photon will be scattered along a scattering angle  $\theta$ . Mathematically, it is

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

formulated as the scattered intensity at an angle  $\theta$  normalized with respect to the scattering intensity integrated along with all directions.

$$P(\theta, \alpha, m) = \frac{F(\theta, \alpha, m)}{\int_0^\pi F(\theta, \alpha, m) \sin\theta d\theta} \quad (15)$$

- ii. **Hemispheric backscatter ratio:** In simple words, this is the fraction of intensity that gets backscattered with respect to the total scattered intensity.

$$b(\theta, \alpha, m) = \frac{\int_{\pi/2}^\pi P(\theta, \alpha, m) \sin\theta d\theta}{\int_0^\pi P(\theta, \alpha, m) \sin\theta d\theta} \quad (16)$$

### **Scattering regimes**

Rayleigh scattering is applicable in the cases where the particle size ( $d_p < 0.1 \mu\text{m}$ ) is much smaller in comparison to the wavelength of light, for instance, the scattering of UV and visible radiation by air molecules. For an incident light of intensity  $I_o$ , the intensity of light scattered by a spherical particle can be represented mathematically as a sum of two mutually perpendicular components denoted as,  $I_\perp$  and  $I_\parallel$

$$I_\perp = \frac{I_o \pi^4 d^6}{8R^2 \lambda^4} \left( \frac{m^2 - 1}{m^2 + 2} \right)^2 \quad (17)$$

$$I_\parallel = \frac{I_o \pi^4 d^6}{8R^2 \lambda^4} \left( \frac{m^2 - 1}{m^2 + 2} \right)^2 \cos^2 \theta \quad (18)$$

where  $\theta$  is the angle between the incident light and the detector/observer,  $R$  is the distance between the particle and the detector/observer,  $m$  is the refractive index of the particle,  $d$  is the particle diameter,  $\lambda$  is the wavelength, and  $I_o$  is the intensity of incident light. Rayleigh scattering is isotropic in nature (equal intensities of light is scattered in the forward and backward direction) and proportional to  $\lambda^{-4}$ . It is important to note that as the particle size decreases, the scattering intensity drops sharply (by a factor of  $d^6$  where  $d$  is the particle diameter); therefore, scattering may not be a suitable technique to study small-sized particles. Moreover, the fact that scattering intensity is proportional to  $\lambda^{-4}$  explains why the sky is blue ( $\lambda_{\text{blue}} = 470 \text{ nm}$  and  $\lambda_{\text{red}} = 665 \text{ nm}$ ). Rayleigh scattering efficiency is much smaller than one, implying that particles in this size range are inefficient scatterers compared to their geometric cross-section.

### **Mie scattering**

Mie scattering was formulated in 1908 by Gustav Mie to fill the gap between geometric scattering (applicable to particles much larger than the wavelength of light) and Rayleigh scattering (applicable to particles much smaller than the wavelength of light). Mie scattering completely



## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

solves Maxwell's equations for spherical particles. Although the solution is exact, it requires complex computational resources.

The scattering efficiency of a sphere of radius  $a$ , size parameter  $\alpha$  and refractive index  $m$  can be written as (Bohren and Huffman, 1983)

$$Q_{scat}(m, \alpha) = \frac{2}{\alpha^2} \sum_{k=1}^{\infty} (2k+1) [|a_k|^2 + |b_k|^2] \quad (19)$$

$$Q_{ext}(m, \alpha) = \frac{2}{\alpha^2} \sum_{k=1}^{\infty} (2k+1) \text{Re}[a_k + b_k] \quad (20)$$

$$a_k = \frac{\alpha \psi'_k(y) \psi_k(\alpha) - y \psi'_k(\alpha) \psi_k(y)}{\alpha \psi'_k(y) \zeta_k(\alpha) - y \zeta'_k(\alpha) \psi_k(y)} \quad (21)$$

$$b_k = \frac{y \psi'_k(y) \psi_k(\alpha) - \alpha \psi'_k(\alpha) \psi_k(y)}{y \psi'_k(y) \zeta_k(\alpha) - \alpha \zeta'_k(\alpha) \psi_k(y)} \quad (22)$$

where  $y = \alpha m$ ,  $\psi_k(z)$  and  $\zeta_k(z)$  are Riccati-Bessel functions (expanded in Equation 23 and Equation 24), while  $J_{k+1/2}$  and  $J_{-1-1/2}$  are Bessel functions of the first kind. In Equation 20,  $\text{Re}$  stands for the real part of the complex number. Determination of the coefficients  $a_k$  and  $b_k$  is the most computationally intensive part of solving the above equations.

$$\psi_k(z) = \left(\frac{\pi z}{2}\right)^{1/2} J_{k+1/2}(z) \quad (23)$$

$$\zeta_k(z) = \left(\frac{\pi z}{2}\right)^{1/2} [J_{k+1/2}(z) + i(-1)^k J_{-1-1/2}(z)] \quad (24)$$

Mie scattering is anisotropic in nature. The scattering becomes increasingly forward as the particle size increases.

### **2.4.3. PM monitoring at IISER Mohali using low-cost optical sensors**

Optical instruments to characterize ambient aerosols can be broadly categorized into two categories, optical particle counters and photometers (nephelometers). Optical particle counters study light scattering and extinction from individual particles, while photometers analyze light scattering and extinction from an ensemble of particles. Several commercial research-grade (GRIMM, Sensidyne nephelometer) optical instruments are available today. GRIMM aerosol monitor is approved by the US EPA as a Federal Equivalent Method. However, their application is limited to research and regulatory bodies.

The sudden demand for air quality data by concerned citizens has fuelled the development and propagation of low-cost optical sensors. Low-cost PM sensors are commercial ready-to-use

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

products that detect particles via laser-based light scattering. Sensors belonging to both nephelometric and OPC categories are available. Few examples of nephelometric-based sensors are Nova SDS011, Shinyei and Plantower sensors. Few examples of optical particle counters are AlphaSense OPC-N2 and Alphasense OPC-N3. Currently, a standard calibration protocol or a regulatory body that assesses/certifies the performance of these sensors does not exist.

The factory calibration of OPCs involves establishing a relationship between the scattering intensity and particle size is established, carried out using calibration aerosols with known properties. In contrast, nephelometric sensors are empirically calibrated by relating the total scattering intensity from an ensemble of particles to the mass concentration reported by co-located instruments (Hagan and Kroll, 2020). However, most low-cost sensor manufacturers provide little information on the factory calibration procedures or the choice of test aerosol (Badura et al., 2018; Badura et al., 2019). Nakayama et al. (2018) reported using spherical polystyrene latex spheres (density =  $1.05 \text{ g cm}^{-3}$ ) to calibrate the Panasonic PM<sub>2.5</sub> sensor. Most low-cost sensors use a DC fan to pull in air laden with particles inside the detection chamber. The ambient particles then scatter light, which is sensed by a photodetector. The particle size is estimated using a scattering intensity. The number of particles is calculated based on the pulses detected by the photodetector. Then, the measured volume concentration is converted into the mass concentration using an estimate for aerosol density. Most manufacturers perform either a laboratory calibration or a short-term field calibration by co-locating the sensors next to a Federal Equivalent or Federal Reference Method analyzer. For most low-cost sensors, the exact algorithm that converts the scattering information to mass concentration is proprietary information. The sensors do not have flow control, a size-selective inlet, or humidity control.

While studies are still evaluating the feasibility of low-cost sensors for long-term measurements, they are already being deployed for a wide range of applications ranging from personal to research-based (Kumar et al., 2015). South Coast Air Quality Management District (SCAQMD) has carried out the field evaluation of several low-cost sensors in California, USA, and the preliminary results are freely accessible on the internet (SCAQMD, 2016). The satellite-based PM<sub>2.5</sub> measurements reported by the Multi-Angle Imaging Spectroradiometer onboard the NASA satellite, TESS will be validated by a network of ground-based low-cost PM sensors.

In this thesis, I used the Airveda (nephelometric-based) and Laser Egg (OPC) low-cost optical sensors to measure PM mass concentration. Airveda is a nephelometric-type sensor, while Laser Egg is an optical particle counter (Figure 2.5). Both the sensors report the mass concentration of PM<sub>10</sub> and PM<sub>2.5</sub>. The major difference in these instruments lies in the internal sensor. While Laser Egg uses a Plantower PMS3003 sensor, Airveda used a Nova SDS 011 sensor. Additionally, the

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

former has two fans, one at the inlet and the other at the suction end. In contrast, the latter has only one fan at the suction end. None of these sensors have flow control, size-selective inlet, or humidity control. Laser Egg used Wi-Fi for data transmission, while the Airveda sensor relied on a SIM card with an active internet connection. The Laser Egg sensor was installed from April 2016 till July 2016. The Airveda sensors were installed from November 2018 to March 2020. Often, insect debris or dust would accumulate inside the inlet that adversely affected the data quality. As a way out, a thin mesh was wound to the inlet to safeguard the sensors' detector.

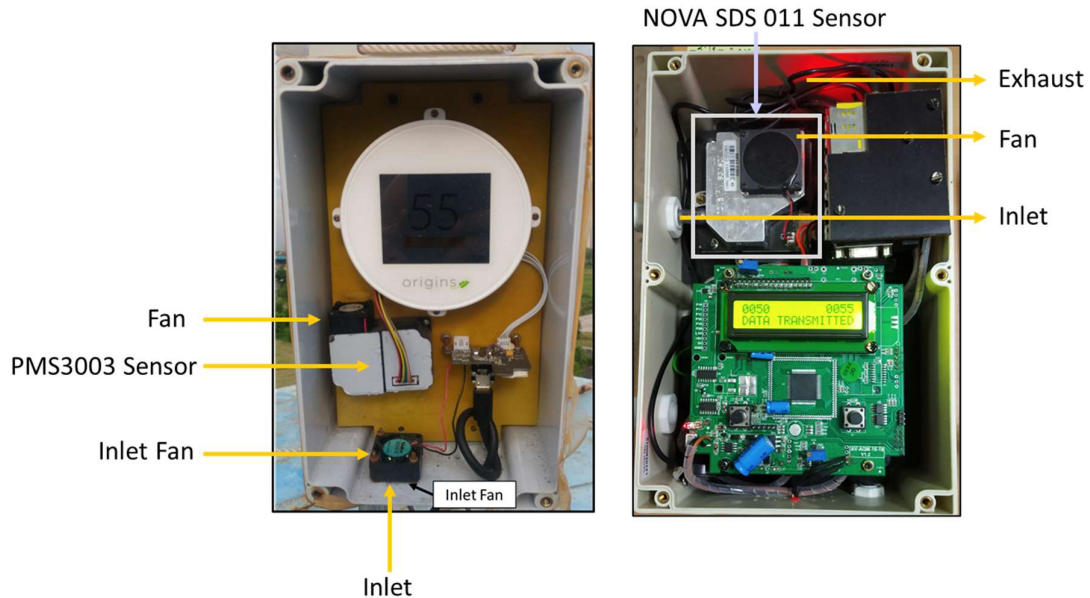


Figure 2.5: Figure showing the major internal components of the modified Laser Egg (left) sensor and the Airveda sensor (right)

Provided the low-cost sensors are accurate, they can be used for formulating effective PM mitigation measures in the Indo-Gangetic plains. Policymakers can dictate steps to mitigate air pollution only on the sources present in their jurisdiction. Therefore, it is essential to know the relative contribution of various sources to the overall pollutant loading. For instance, most of the studies dealing with source-apportionment of PM conclude by blaming the elevated PM levels on generic sources such as industrial, traffic, biomass burning, etc. A blanket policy aimed at reducing emissions from transport or all industries is challenging to implement. On the other hand, supplementing the existing measurement facilities with accurate, low-cost PM sensors and encouraging citizen science participation can reveal the exact spatio-temporal location of pollution micro-hotspots with a relatively lower investment. Thus, a high-density measurement network can help formulate targeted granular policies (for instance, a specific industry in the industrial area) that are convenient to implement and evaluate. Low-cost sensing technologies can also assist in better planning and execution of sophisticated scientific studies. Quantifying the contribution of a

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

distant source (whose exact spatial location is unknown) to the observed speciated pollutant loading using an instrument installed at the receptor site can only provide a limited accuracy. On the other hand, planning a field study at the source (or nearby) only after identifying its exact location using an ensemble of low-cost sensors will furnish more robust results.

After extensive evaluations, I used the Airveda low-cost sensors for quantifying the contribution of paddy residue burning and domestic biofuel burning for heating purposes to the particulate matter loadings in north-west India.

### **2.5. Approaches to link air mass history with atmospheric composition at a receptor site**

The concentration of particulate matter at a site is dependent on the interplay between sources, sinks, and meteorology. The sources may be primary, meaning that the particles are directly released into the atmosphere, or secondary, meaning they are formed in the atmosphere through complex chemical reactions. The sources may be local or distant, in which case pollutants at the source can be transported to sites downwind far away from the source. Policymakers can only target the sources in the area under their jurisdiction. Therefore, they need to know the relative contribution of local versus distant sources to the observed air composition to mitigate air pollution. Thus, identifying the location of sources by assessing the link between air mass history with observed atmospheric composition at a receptor site is pivotal.

Several recent scenarios delve into understanding this link. For instance, every year during the paddy harvest period (October-November), farmers in Punjab, Haryana, and Uttar Pradesh set their fields on fire to eliminate the residue in time for the wheat sowing. This massive activity deteriorates the air quality in the entire north-western Indo Gangetic Plain (Venkataraman et al., 2006; Sahai et al., 2007; Sarkar et al., 2013; Chandra and Sinha, 2016). Massive emissions from paddy residue burning coupled with an already high pollutant loading in metropolitan cities along with decreasing temperatures during this period lead to an air pollution health emergency, particularly in the National Capital Region (NCR) (Kulkarni et al., 2020a). Therefore, it becomes critical to understand the quantitative contribution of emissions from paddy residue burning to Delhi's pollutant load. Several studies have tried to unravel the link between the observed air composition in New Delhi and the history of air masses to present a clearer picture to all the stakeholders involved (Cusworth et al., 2018; Liu et al., 2018; Bikkina et al., 2019; Kulkarni et al., 2020a) but the current estimates are marred with huge uncertainties. Nevertheless, it is critical to understand, establish, and quantify the link between air mass history and atmospheric composition at a receptor site to formulate evidence-based policies and mitigate air pollution.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

The tools suitable to establish the link between air mass history and atmospheric composition are primarily dependent upon the spatial distance between the sources and receptors. For instance, tools based on locally measured wind speed and wind direction are apt to locate nearby sources. In contrast, trajectory-based tools are suitable for identifying distant sources that involve synoptic-scale air transport. The section below provides a brief description of a few tools used to assess the link between air mass history and atmospheric composition.

### **2.5.1. Wind speed and wind direction measurements**

Locally measured wind direction and wind speed measurements are routinely used to plot wind rose plots, which summarize the meteorological data and assess the general distribution of flow patterns. Such plots are often useful to discern local influences that affect the air parcel roughly a few hours before their arrival at the receptor site.

To visualize the location of nearby sources, several informative plots such as conditional probability, polar plot, polar annulus plot, etc., can be constructed by combining the measurements of the pollutant with local meteorological parameters, wind speed, and wind direction. These plots provide a clearer picture of the spatial location of the sources with respect to the receptor. However, they do not give any information about the synoptic scale of air transport.

A conditional probability plot reveals the dominant direction associated with high pollutant loading as a probability function (Ashbaugh et al., 1985; Kim and Hopke, 2004). To plot a conditional probability plot, we need simultaneous measurements of wind direction and mixing ratio of the pollutant of interest,  $P$ . The entire dataset is then split into wind sectors of suitable width, referred to here as  $\theta$ . For each bin,  $\theta$ , conditional probability function (CPF) is then calculated as

$$CPF_{\theta} = \frac{m_{\theta}}{n_{\theta}} \quad (25)$$

where  $n_{\theta}$  is the total number of available measurements for that wind sector and  $m_{\theta}$  is the number of measurements when the atmospheric loading of pollutant  $P$  exceeded a threshold, usually defined as a percentile. The higher the value of CPF for a wind sector  $\theta$ , the more is the probability of a source being present there. Although wind speed measurements are not explicitly needed for CPF plots, Kim and Hopke (2004) recommend to filter out the measurements when wind speed was less than  $1 \text{ m s}^{-1}$  to reduce artifacts from the isotropic behavior of wind vanes under calm conditions. CPF is frequently used in tandem with source apportionment studies of gas-phase VOCs (Sarkar et al., 2017; Pallavi et al., 2019) as well as particulate matter (Gildemeister et al., 2007; Kim et al., 2007; Begum et al., 2010; Oh et al., 2011; Ofosu et al., 2012) based on Positive Matrix Factorisation to observe the directional dependence of each source.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

The polar plot is a bivariate plot that displays the concentration of a pollutant modeled as a function of wind speed and wind direction on a smooth continuous surface (Carslaw et al., 2006; Carslaw and Ropkins, 2012; Carslaw and Beevers, 2013). While wind direction reveals the spatial location of a potential source, wind speed too can impact the particulate matter loadings observed at the receptor site in several ways. For instance, re-suspension of particulate matter, generation of sea-spray aerosol, and long-range transport of particulates increase with increasing wind speed. To make such a plot, the wind speed, wind direction is partitioned into bins of appropriate width, and the average pollutant loading,  $\bar{p}$  for each bin is computed. Wind direction is generally split into bins of 10 degrees width, while the wind speed data is divided into 30 bins. The average wind speed  $\bar{u}$  and the wind direction,  $\theta$  are converted from cartesian coordinates into polar coordinates with wind components  $u$  and  $v$  as

$$u = \bar{u} \cdot \sin\left(\frac{2\pi}{\theta}\right) \quad (26)$$

$$v = \bar{u} \cdot \cos\left(\frac{2\pi}{\theta}\right) \quad (27)$$

Now, to the observed values of  $u$ ,  $v$ , and  $\bar{p}$  a General Additive Model (GAM) is applied, which models the surface to describe the concentration as a function of wind vectors (Carslaw and Beevers, 2013). Several studies (Zhou et al., 2018; Jandacka and Durcanska, 2019; Hama et al., 2020; Kumar et al., 2020) have used bivariate polar plots to locate the spatial location of sources of ambient particulate matter.

Polar annulus plot is another type of bivariate plot. In addition to the two variables (wind direction and pollutant concentration) used in CPF plots, polar annulus plots allow us to visualize the temporal variations (daily, day of the week, the month of the year) of the source in a concise manner (Carslaw and Ropkins, 2012). This plot is shaped like an annulus or a donut with a hollow circular cavity in the middle. The earliest time index is at the inner edge of the annulus, while the latest index lies at the outer edge of the annulus. In order to make a polar annulus plot, the available data is split into the different wind and temporal (hourly, day of the week, or monthly depending upon the objective) sectors, and the average pollutant loading is evaluated for each sector. The missing data is then interpolated and smoothed using a Kriging technique to give a Polar annulus plot (Carslaw et al., 2006). Harrison et al. (2012) used a polar annulus plot (plotted as the hour of the day) to conclude that while the morning rush hour traffic from a major road in the north-northeast sector contributes significantly to  $\text{NO}_x$ , its contribution to ambient  $\text{PM}_{2.5}$  measured at the receptor site in the UK is minimal. Suppose there are several sources of a pollutant present in the vicinity of the receptor. In that case, the temporal disparities among them can throw some light on the possible identity of the sources. For instance Sarkar et al. (2016) distinguished the biogenic

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

source of isoprene from the combustion related sources using a polar annulus plot during a field campaign carried out in Kathmandu, Nepal.

To summarize, tools based on locally measured meteorological parameters coupled with the pollutant's concentration are convenient to use and suitable for locating nearby sources. However, due to the turbulent nature of wind, the assumption that the in-situ wind direction represents the synoptic-scale flow may not be accurate.

### **2.5.2. Trajectory based models**

Atmospheric motion can be visualized using two distinct methodologies, namely the *Eularian* formulation and the *Lagrangian* formulation. Eulerian models are grid-based; that is, they use a frame of reference fixed in space, and calculation of dependant variables is done with respect to this fixed grid coupled to rotating earth. On the other hand, a Lagrangian model uses a moving frame of reference, which traces the movement of a displaced air particle or an ensemble of particles from the source to the receptor. A trajectory is the simplest form of a Lagrangian model.

A trajectory traces the motion of an infinitesimally small air parcel or an ensemble of particles either forward or backward in time. Back-trajectories, which estimate the path of an air parcel from the receptor to the source backward in time, have been a popular tool to identify the spatial location of sources. A simple trajectory-based model was first developed graphically by Petterssen (1956).

Using the Lagrangian formulation, the path of an infinitesimally small air parcel can be expressed mathematically as a differential *trajectory equation* given by

$$\frac{dX}{dt} = \dot{X}[X(t)] \quad (28)$$

where  $X$  is the position of the air parcel,  $t$  is the time, and  $\dot{X}$  represents the wind velocity vector with 'u' and 'v' as the horizontal components and 'w' as the vertical component.

Provided the initial position of the air parcel,  $X_o$  at time  $t = t_o$  is known, then its position  $X$  at a later time  $t$  can be evaluated as

$$X(t) = X(X_o, t_o) \quad (29)$$

Similarly, if the position of particle  $X$  is known at a time  $t$ , then its initial position  $X_o$  can be evaluated at an earlier time  $t_o$  as per the following inverse expression

$$X_o(t_o) = X_o(X, t) \quad (30)$$

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

Equation 29 allows for the calculation of forward trajectories, which trace the path of the air parcel forward in time, while Equation 30 allows for the calculation of backward trajectories, which track the position of air parcels backward in time.

The motion of air mass trajectories is modelled on the assumption that two particles that were adjacent at the beginning ( $t_o$ ), continue to remain next to each other during their journey from the source to the receptor site. This assumption can be represented mathematically as

$$\lim_{\Delta X_o \rightarrow 0} |X(X_o + \Delta X_o, t) - X(X_o, t_o)| = 0 \quad (31)$$

However, the above assumption is rarely valid in real life because air parcel becomes distorted under the action of turbulent forces in the atmosphere.

The *trajectory equation* can be solved in the forward direction by applying a Taylor expansion on  $X(t)$  about  $t = t_o$  and evaluating at  $t_1 = t_o + \Delta t$  which gives

$$X(t_1) = X(t_o) + (\Delta t) \left. \frac{dX}{dt} \right|_{t_o} + \frac{1}{2} (\Delta t)^2 \left. \frac{d^2X}{dt^2} \right|_{t_o} + \dots \quad (32)$$

Considering only the first two terms in Equation 32 gives us an approximate, *zero-acceleration solution* to the trajectory equation, as shown in Equation 33. This is called a *zero-acceleration solution* because we assume that the second derivative of position vector  $X$  with respect to time is zero. This solution requires significantly less computational resources and is reasonably accurate provided the integration step length denoted by  $\Delta t$  is small and the initial position  $X(t_o)$  along with wind field is known

$$X(t_1) \approx X(t_o) + (\Delta t) \dot{X}(t_o) \quad (33)$$

Likewise, the *trajectory equation* can be expanded about  $t = t_1$  and evaluated at  $t_o$

$$X(t_o) = X(t_1) + (\Delta t) \left. \frac{dX}{dt} \right|_{t_1} + \frac{1}{2} (\Delta t)^2 \left. \frac{d^2X}{dt^2} \right|_{t_1} + \dots \quad (34)$$

Combining Equation 32 and 34 gives us

$$X(t_1) = X(t_o) + \frac{1}{2} (\Delta t) [\dot{X}(t_o) + \dot{X}(t_1)] + \frac{1}{4} (\Delta t)^2 \left[ \left. \frac{d\dot{X}}{dt} \right|_{t_o} - \left. \frac{d\dot{X}}{dt} \right|_{t_1} \right] + \dots \quad (35)$$

Considering only the first two terms of Equation 35 and ignoring the higher-order terms gives us the *constant-acceleration solution* to the *trajectory equation*, as shown in Equation 36.

$$X(t_1) \approx X(t_o) + \frac{1}{2} (\Delta t) [\dot{X}(t_o) + \dot{X}(t_1)] \quad (36)$$



## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

This is called the constant-acceleration solution because the derivative of the velocity vector  $\dot{X}$  is the same when evaluated at the times  $t_o$  and  $t_1$  implying that the acceleration does not change with time. In principle, the second-order in Equation 32 can also be retained to give a variable acceleration solution; however, not only is that computationally expensive but using wind velocity data which is available at a low temporal resolution to compute acceleration at two separate times  $t_o$  and  $t_1$  can result in inaccuracy (Stohl, 1998).

### **Pre-requisite meteorological data for computation of trajectories**

As per the zero-acceleration and constant-acceleration solution to the trajectory equation mentioned in Equation 33 and Equation 36, we need the value of wind velocity,  $\dot{X}$  at  $t_o + \Delta t$  where  $\Delta t$  is the time integration step to compute the position of the air parcel, both in the past (back-trajectory) and in the future (forward trajectory). It is evident that the smaller the length of the integration step, the higher will be the accuracy of the computed trajectory. Therefore, we need high spatial and temporal resolution data of wind vectors. Ground-based measurements of wind speed are often limited both spatially and temporally and hence, insufficient for synoptic-scale level calculations. Moreover, the measurements of the vertical component of the velocity, 'w' are even more scarce. Therefore, most air trajectory models retrieve the required wind vector information by interpolating data from an archived gridded dataset available from Numerical Weather Prediction (NWP) models.

The primary sources of errors in trajectory computation arise from the truncation of higher-order terms in the differential solution of trajectory equation (called truncation errors), inherent errors in the archived meteorological data fields, interpolation of the gridded meteorological data to approximate the wind vector at trajectory location, the paucity of measurements and validation of the vertical component of wind vector and dissimilarities in model and actual topography (Stohl, 1998). The truncation error is proportional to  $\Delta t$  in the case of the zero-acceleration solution and to  $\Delta t^2$  if constant-acceleration solution is being used, where  $\Delta t$  represents the time integration step. The magnitude of this truncation error can therefore be reduced by keeping this integration step as small as possible. Inaccurate wind data and interpolation are the largest sources of error in predicting the path of a trajectory. A study by Baumann and Stohl (1997) reported that after 96 hours of run time, the trajectory position could be erroneous by thousands of kilometers. It is equally important to assess which among the freely available gridded datasets performs better for the chosen receptor site. Kahl et al. (1989) reported that trajectories computed using two different gridded meteorological datasets (the European Centre for Medium-Range Weather Forecasts (ECMWF) and by the National Meteorological Centre, currently known as the National Centre for

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

Environmental Prediction) were ~1000 km apart from one another after a travel time of 5 days. Another study performed by Cabello et al. (2008) for a site in Spain presented a similar result. The median distance between the computed trajectories grew linearly up to a travel time of 72 hours (3 days), following which the differences started to increase at a much faster rate when the trajectory model was initialized using two different meteorological gridded datasets (NCEP Reanalysis and FNL). Gebhart et al. (2005) also reported that the trajectories computed using two meteorological datasets in the HYSPLIT model could be 180° apart for some episodes. The net aggregate error in the final trajectory output has contributions from all possible sources of error mentioned above, due to which a computed trajectory may only be representative of the path of an air parcel for few days. Rather than using a single trajectory as an attempt to mimic the path of the air parcel, Stohl et al. (1995) recommended initializing multiple trajectories (called an ensemble) from the starting location by inculcating random errors in the wind field at every time-step to evaluate the extent of inaccuracy introduced by interpolation or sparse meteorological data and the turbulent nature of the atmosphere.

To validate a computed trajectory, we need to have a “reference” or an “observed” trajectory to compare the former against the latter. This is usually carried out by performing tracer experiments. Ideally, a tracer should be chosen such that it is conserved along the path of the trajectory. Most studies in the past have utilized either balloons, material tracers, or dynamical tracers to establish a “true” trajectory (Stohl, 1998; Riddle et al., 2006). Constant level balloons follow the horizontal motion of the air perfectly; however, they are unable to reveal the potential inaccuracies arising from the error-prone vertical component of the velocity in meteorological datasets (Riddle et al., 2006). A chemically stable chemical compound that does not undergo atmospheric loss processes is readily available, can be conveniently detected, does not harm the environment, and its level in the atmosphere is negligible has the potential to be used as a material tracer. These tracers can be categorized into two categories, the first being tracer experiments where inert compounds are released as part of planned experiments. The second type is tracers of opportunity when an accidentally emitted species fulfils the necessary pre-requisites can serve as a tracer. Field campaign studies trying to validate trajectory motion on synoptic scales have used perfluorocarbon compounds as the material tracer in the past (Gebhart et al., 2005). Volcanic ash or the release of radioactive waste from the Chernobyl explosion have been used as tracers of opportunity (Stohl, 1998). Potential temperature ( $\theta$ ) and isentropic potential vorticity (PV) are the two meteorological variables that fall under dynamical tracers. Potential temperature ( $\theta$ ) is defined as the temperature achieved by the air parcel would if it were brought down from an initial pressure P to the reference

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

pressure  $P_o$  (1000 hPa) adiabatically, i.e., without exchanging temperature/energy with the surroundings. It is defined as

$$\theta = T \left[ \frac{P_o}{P} \right]^{\frac{R}{C_p}} \quad (37)$$

In the above equation, T is the initial temperature of the air parcel in Kelvin, R (287 J K<sup>-1</sup> kg<sup>-1</sup>) is the universal gas constant, and C<sub>p</sub> (1005 J K<sup>-1</sup> kg<sup>-1</sup>) is the specific heat of dry air at constant pressure. Unlike temperature, the potential temperature of an air parcel remains conserved during an adiabatic displacement and hence can be used as a tracer.

### **2.5.3. Lagrangian particle dispersion models**

Lagrangian particle dispersion models (LPDMs) are stochastic models that simulate the motion (transport and dispersion) of an ensemble of infinitesimally small air parcels by computing random trajectories in a turbulent flow (Brasseur and Jacob, 2017). Unlike Eulerian models, LPDMs do not use a fixed grid, so their theoretical resolution can be modified depending upon the need and computational resources (Stohl et al., 2005). Also, LPDMs do not have numerical diffusion (the tendency of particles to diffuse more than they should), unlike Eulerian models, which maintains the gradients. FLEXPART is one of the most popularly used LPDMs for particulate matter dispersion modeling.

### **2.5.4. Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) Model**

HYSPLIT is one of the most widely used tools for studying and modeling atmospheric transport, deposition, dispersion, and transformation (Draxler and Hess, 1998; Draxler and Rolph, 2013). I use this model in the computation of air mass back trajectories for Mohali, India. This model is provided by the National Oceanographic and Atmospheric Administration (NOAA) Air Resource Laboratory (ARL). The acronym HYSPLIT stands for Hybrid Single-Particle Lagrangian Integrated Trajectory. The word “Hybrid” sheds light on the hybrid nature of the model as it uses a hybrid approach involving both Eulerian and Lagrangian formulations. Computation of air mass trajectories from the source to the receptor site is performed using a Lagrangian approach, which employs a moving frame of a reference. In contrast, the computation of the concentration profile of pollutants is performed using an Eulerian approach, with respect to a fixed grid.

Computation of air mass trajectories using gridded archived meteorological datasets is the most popular application of HYSPLIT. This is performed via an online platform provided by NOAA’s ARL ([https://www.ready.noaa.gov/HYSPLIT\\_traj.php](https://www.ready.noaa.gov/HYSPLIT_traj.php)). The first step is choosing the number of starting locations where trajectories are to be initialized. In the case of backward trajectories, the

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

starting location is the receptor, while in the case of forward trajectories, the starting location is the source. Using the online platform, up to 3 locations can be chosen as the starting point of the trajectories. The next step is choosing among the trajectory types: single, ensemble, matrix, or frequency. Selecting the single option runs one trajectory from each of the selected starting points. Matrix option runs uniformly spread-out trajectories within a grid box, the limiting coordinates of which provided by the user. The ensemble option starts multiple trajectories from the specified starting point by offsetting the input meteorological data  $\pm 1$  grid point in the horizontal direction and  $\pm 0.01$  sigma units in the vertical direction, which results in a total of 27 trajectories (Draxler, 2003). In my thesis, I use the HYSPLIT model in ensemble mode to compute 3-day back trajectories for Mohali. However, not all 27 trajectories were consistent with the local topography. I only selected those trajectories which correctly attributed the altitudes of Mohali being located in plains ( $< 400$  m a.m.s.l.) and Shimla, a mountain site 60 km north-east of this site, being located in the mountains ( $> 400$  m a.m.s.l.).

The frequency option computes a single trajectory from the specified starting location periodically every 6 hours and then plots the probability of a trajectory/trajectory endpoint passing through a grid cell with respect to the total number of trajectories/trajectory endpoints.

The next step in trajectory calculation is selecting an archived gridded meteorological dataset with an appropriate resolution from the available options. HYSPLIT uses a processed version of the output meteorological fields from numerical weather prediction models. At the outset, HYSPLIT interpolates the dataset's vertical coordinate system into terrain-following ( $\sigma$ ) so that all heights are reported with respect to the local terrain height.

$$\sigma = 1 - \frac{z}{Z_{top}} \quad (38)$$

In the above equation,  $z$  is the height above ground level and  $Z_{top}$  is the highest point in HYSPLIT's model domain. The value of  $\sigma$ , therefore, varies from 0 (at  $z = Z_{top}$ ) to 1 (at the surface). This solves the issue of the intersection with the ground in undulating terrain. Therefore, for the trajectory output to be accurate, the terrain must be accurately represented in the archived meteorological dataset. The horizontal wind vectors,  $u$  and  $v$ , temperature, height, and pressure are variables necessary to initialize the HYSPLIT trajectory model. The vertical velocity component is available in terms of pressure in most datasets. Therefore, the default trajectory calculation is carried out using a kinematic approach, i.e., using three velocity vectors. In case the vertical velocity data is missing, then HYSPLIT provides the option of assuming that the trajectory is traveling on a fixed pressure or on a fixed potential temperature surface, allowing users to compute isobaric or isentropic trajectories, respectively.

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

HYSPLIT provides several archived gridded datasets to choose from, out of which GDAS (Global Data Acquisition System) and Reanalysis datasets have a global domain. Data assimilation is a methodology by which observational data is integrated into a model to improve the accuracy of the forecast. GDAS is used by National Centre for Environmental Prediction (NCEP) to incorporate observational data from balloons, radars, satellites, and buoys into a gridded format for the initialization of the Global Forecast Model (GFS) (Kanamitsu, 1989). This dataset is available at a 1° spatial resolution from December 2004 till present and at a 0.5° spatial resolution from September 2017 till June 2019. Reanalysis simulations perform data assimilation of historical observations in a recent numerical weather prediction model using a single consistent scheme to provide a gridded dataset of the past climate. HYSPLIT can also run a trajectory model using the Reanalysis dataset with a reported spatial and temporal resolution of 2.5° and 6 hours, respectively, available from 1948 till the present (Kalnay et al., 1996). This dataset is jointly provided by NCEP and the National Centre for Atmospheric Research (NCAR).

While selecting from the available datasets, it should be ensured that the terrain of the receptor location is adequately represented in the gridded data, as this can have implications on the trajectory accuracy. For instance, my study site, Mohali, lies close to the forested foothills of the Himalayas at a terrain height of 310 m a.m.s.l. However, the modelled terrain height computed through HYSPLIT using the GDAS and reanalysis meteorological data fields was 667.6 m a.m.s.l. and 1249.9 m a.m.s.l., respectively. Likewise, when HYSPLIT was run in the ensemble mode, Mohali's modelled terrain height varied between 200 and 3500 m a.m.s.l. for the GDAS meteorological data field and between 240 and 5100 m a.m.s.l. for the reanalysis meteorological data field. Therefore, the GDAS dataset better resolved the topography than the reanalysis dataset and hence was used for further analysis.

The next step in trajectory computation is entering the details about the trajectory direction (forward/backward), start time, run time in hours, vertical motion (model velocity/isobaric/isentropic), and the starting height of the trajectory. HYSPLIT also gives the option to retrieve additional variables like terrain height, potential temperature, ambient temperature, rainfall, mixed layer depth, relative humidity, and solar radiation flux along with all the trajectory endpoints in the output file.

As part of the first Aerosol Characterization Experiment carried out in 1995, three constant-level balloons were released from Tasmania (Bates et al., 1998), and their movement was tracked for 24 hours. The path followed by these balloons were used to validate the trajectories computed using HYSPLIT. The model was run in forward mode, at the height of 200m above ground level, using two different input meteorological datasets to simulate the path of this balloon. After a run

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

time of 24 hours, it was observed that the simulated trajectory was able to successfully track the path of the balloon within an error range of 10 to 20% (Draxler and Hess, 1998)

### **2.5.5. Chemistry transport models**

Neither trajectory-based nor particle dispersion models can distinguish the primary PM in the atmosphere from the secondary PM. While they may tell us about the geographical location of a potential source in terms of probability, they cannot ascertain if that source is directly emitting PM (primary emissions) in the atmosphere or if it is emitting precursors, which later form secondary PM. Moreover, simple models cannot account for the chemical evolution of ambient PM during transport from the source to the receptor site. Chemical transport models (CTM) address these limitations by having a quantitative mathematical framework for all-important atmospheric processes. CTMs integrate meteorology and emissions (in the form of emission inventories) with chemical reactions and kinetics. After assimilating all pre-requisite information, a CTM solves complex numerical equations to yield pollutant loading as a function of space and time.

## **2.6. Approaches to link air mass trajectory with atmospheric composition**

A single trajectory is insufficient to trace the path of an air parcel due to the underlying simplified assumptions involved in computing trajectories (Kahl, 1993; Stohl, 1998; Stohl et al., 2002). However, using an ensemble of trajectories causes the errors associated with individual trajectories to average out, paving the way for more meaningful analysis (Fleming et al., 2012). A collection of back trajectories remains a popular tool to establish synoptic-scale climatology, identify distant sources of pollutants, investigate trans-boundary pollution (Pérez et al., 2015). This section describes a few of the widely used methods to discern air mass history using a collection of trajectories. They can be broadly classified into geographical sector classification and clustering methods.

### **2.6.1. Geographical sector classification**

In geographical sector classification, the group of trajectories or the collection of trajectory endpoints is resolved into spatially fixed topographical sectors. A grid sector is termed as a source when a trajectory passing through that sector results in high pollutant loading at the receptor site. A percentile-based benchmark is used to split pollutant loadings into high and low.

#### **2.6.1.1. Potential Source contribution function (PSCF)**

PSCF is a type of conditional probability that helps us identify the geographical location of sources by combining air mass back trajectories with the measured pollutant loading at the

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

receptor site. It is identical to the Conditional Probability Function, which uses locally measured wind direction instead of trajectories. While CPF is used to discern the directional location of local sources, PSCF is more relevant for distant or far-off sources of air pollution.

To evaluate PSCF, a latitude-longitude grid of suitable resolution is superimposed on the area of interest around the receptor site. In simple terms, the PSCF value denotes the probability of a potential source being present at a grid cell with the coordinates  $i$  and  $j$ . It is calculated as:

$$PSCF_{ij} = \frac{m_{ij}}{n_{ij}} \quad (39)$$

where  $n_{ij}$  is the total number of trajectories which passed over the  $ij$  grid cell and  $m_{ij}$  represents the number of times when trajectory crossing the  $ij$  grid cell coincided with the pollutant loading at the receptor site exceeding a pre-defined threshold. Usually, the threshold is considered in terms of percentile, and most studies use the 90<sup>th</sup> percentile as the exceedance criteria. Another way to compute PSCF values is by using the number of trajectory endpoints falling in a grid cell, generally referred to as the residence time rather than the number of trajectories. Since the trajectory speed is inversely proportional to the number of trajectory endpoints falling in a grid cell, it suggests that a slow-moving trajectory (higher number of trajectory endpoints in a grid cell) will accumulate more pollutants than a faster one (lower number of trajectory endpoints in a grid cell). PSCF values for all grid cells are evaluated and displayed using a suitable colour scale. Sources are the grid cells that have high PSCF values. To reduce the error associated with grid cells with a low number of trajectory endpoints, it is advised to use a weighting function.

PSCF cannot resolve local resources in the vicinity of the receptor as all trajectories converge there. Moreover, all grid cells along the path of a trajectory are assigned equal probabilities, resulting in “type 1” errors (false positives) with high PSCF values upwind and downwind from the actual source location. This is also called the “tailing effect.”

### **2.6.1.2. Concentration Field (CF)**

In PSCF, the value of probability is independent of how high or how close the value of measured pollutant loading is from the chosen threshold, meaning a strong and weak source are both attributed to the same value of PSCF. The concentration field or CF helps to overcome this limitation of PSCF and thus reveals the strength of sources. The CF value of a cell with coordinates  $i$  and  $j$  is:

$$\ln(\overline{C}_{ij}) = \frac{1}{\sum_{k=1}^N \tau_{ijk}} \sum_{k=1}^N \ln(C_k) \tau_{ijk} \quad (40)$$

## 2. Quantitative measurement and source-apportionment of ambient particulate matter (PM)

where  $i$  and  $j$  are the coordinates of the grid cell,  $k$  is the index number of trajectory,  $N$  is the total number of trajectories being analyzed,  $\tau_{ijk}$  is the residence time (number of trajectory segments or endpoints) of the  $k^{\text{th}}$  trajectory in the  $ij^{\text{th}}$  grid cell, and  $C_k$  is the pollutant loading observed at the receptor site at the arrival of  $k^{\text{th}}$  trajectory. Instead of using a single threshold for exceedance, like in the case of PSCF, in CF, we are weighting the residence time by the measured concentration. A higher value of  $C_{ij}$  indicates that a trajectory passing over that grid cell results in high pollutant loading at the measurement (receptor site)

### **2.6.2. Clustering methods**

Clustering methods split the collection of trajectories into distinct groups called clusters. Then the average pollutant loading associated with each is calculated. Hypothesis testing can be applied to examine statistical significance in the differences observed in the chemical composition of clusters.

Classification of air mass trajectories into representative groups is frequently performed using cluster analysis. It is a multivariate technique that seeks to find a structure in an ensemble of objects based on a measure of similarity (or dissimilarity). It works by splitting the data into meaningful groups such that the members of a group are more related to each other than those allotted to other groups. It is an unsupervised learning method as no pre-conceived information about the properties of the groups is passed on to the algorithm, or in simpler words, the cluster labels are not known a priori. This technique can reveal complex relationships in large datasets that may be hitherto known.

Cluster analysis algorithms are broadly classified into hierarchical and non-hierarchical methods, which have been described in the sections below.

#### **2.6.2.1. Hierarchical clustering methods**

In hierarchical clustering, the data is fragmented into clusters in a hierarchical or tree-like manner, which allows us to discover relationships among elements. It can be of two types: top-down and bottom-up, also referred to as divisive and agglomerative. In the initial step of a divisive approach, all elements belong to the same cluster, while in the agglomerative approach, every single element belongs to a cluster of its own. In successive steps, the hierarchical method forms cluster within the cluster. In a multi-variate analysis performed by Kassomenos et al. (2010) on a 4-year dataset of 5-day back-trajectories arriving at Athens (Greece), the output of the hierarchical clustering algorithm was found to be most sensitive to the trajectory arrival height. Hierarchical clustering techniques are both time and computational resource intensive.



### **2.6.2.2. Non-hierarchical clustering methods**

Non-hierarchical clustering methods seek to directly divide the dataset of trajectories into a distinct number of clusters while maximizing the similarity between the members of each cluster and minimizing the similarity between different clusters. Such algorithms establish all clusters iteratively in a single step and do not provide any information about the hierarchy involved among the data elements.

k-means clustering is one of the most popular algorithms among non-hierarchical methods. It splits a dataset of  $N$  objects into  $k$  clusters. The number of cluster  $k$  must be specified by the user. First,  $k$  centroids are randomly selected from the dataset. Then all others are allocated to any one of the  $k$  clusters depending upon the Euclidean distance. The centroid of each cluster is re-calculated, and re-assignment happens again until convergence. This algorithm is sensitive to the input order of data; therefore, it is recommended to initialize the clustering in several different orders.

Kassomenos et al. (2010) found that when a non-hierarchical clustering algorithm (k-means clustering) was applied to a 4-year dataset of 5-day back trajectories, the output was least sensitive to the arrival height of trajectories in comparison to when a hierarchical clustering algorithm was applied on the same dataset. Further, non-hierarchical methods are also faster than hierarchical and require less computational resources. Therefore, I used the k-means clustering algorithm to split the ensemble of 3-day back trajectories arriving at Mohali into distinct clusters. Since this algorithm is sensitive to the input order of the data, I initialized the clustering software using 11 differently sorted input files.

# Chapter 3

---

## **Quantifying the contribution of long-range transport to particulate matter (PM) mass loadings at a suburban site in the north-western Indo-Gangetic Plain (NW-IGP)**

### **3.1. Abstract**

The levels of an atmospheric pollutant at any given site are governed by sources (primary emissions or chemical formation), sinks, and transport via air movement. Rapid urbanisation, industrialization, and population growth witnessed by the Indo-Gangetic Plain in the last few decades have led to increased PM emissions. As a result, many sites here frequently exceed the national ambient air quality standard (NAAQS) of  $100 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{10}$  and  $60 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{2.5}$  mass loadings, exposing residents to hazardous levels of particulate matter (PM) throughout the year. In this chapter, I seek to quantify the contribution of long-range transport to elevated PM levels and the number of exceedance events compared to the regional sources of PM present in the North-West Indo Gangetic Plain. I perform a back-trajectory climatology analysis on the air masses arriving at the IISER Mohali Atmospheric Chemistry facility ( $30.667^\circ\text{N}$ ,  $76.729^\circ\text{E}$ ; 310 m a.m.s.l.) for the period August 2011–June 2013. Air masses arriving at the receptor site were classified into six clusters representing the synoptic-scale air-mass transport patterns. Long-range transport from the west leads to significant enhancements in the average fine- and coarse-mode PM mass loadings during all seasons. The contribution of long-range transport from the west and south-west (source regions: Arabia, Thar Desert, Middle East, and Afghanistan) to coarse-mode PM varied between 9 and 57% of the total  $\text{PM}_{10-2.5}$  mass. Local pollution episodes (wind speed  $< 1 \text{ m s}^{-1}$ ) contributed to enhanced  $\text{PM}_{2.5}$  mass loadings during both the winter and summer seasons and to enhanced coarse-mode PM only during the winter season. South-easterly air masses (source region: eastern IGP) were associated with significantly lower fine- and coarse-mode PM mass loadings during all seasons.

This chapter has been published in "Atmospheric Chemistry and Physics" with S. Garg, V. Kumar, H. Sachan, R. Arya, C. Sarkar, B. P. Chandra, and B. Sinha as co-authors (Pawar et al. 2015). I wrote the draft of the manuscript. All statistical analysis figures and tables presented in the manuscript and used in this thesis are mine. S. Garg, V. Kumar, H. Sachan & R. Arya, helped downloading back trajectories using NOAA HYSPLIT and performed preliminary back trajectory analysis for one season under my guidance as part of a summer project. C. Sarkar, V. Kumar & B. P. Chandra acquired the primary data used in this analysis and performed basic QA/QC of the primary dataset. Dr. B. Sinha provided overall guidance and commented on the draft.

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

The fraction of days in each season during which the PM mass loadings exceeded the NAAQS was controlled by long-range transport to a much lesser degree. For the local cluster, which represents regional air masses (source region: NW-IGP), the fraction of days during which the NAAQS of  $60 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{2.5}$  was exceeded varied between 36% of the days associated with this synoptic-scale transport during the monsoon, and 95% during post-monsoon and winter seasons; the fraction of days during which the NAAQS of  $100 \mu\text{g m}^{-3}$  for the 24h average  $\text{PM}_{10}$  was exceeded, varied between 48% during the monsoon and 98% during the post-monsoon season.

Long-range transport was responsible for both, bringing air masses with a significantly lower fraction of exceedance days from the eastern IGP and air masses with a moderate increase in the fraction of exceedance days from the west (source regions: Arabia, Thar Desert, Middle East, and Afghanistan). In order to bring PM mass loadings into compliance with the NAAQS and to reduce the number of exceedance days, mitigation of regional combustion sources in the NW-IGP needs to be given the highest priority.

## **3.2. Introduction**

India is a rapidly developing nation. Over the past few decades, it has witnessed large-scale urbanisation, industrialization accompanied by an increasing population. These developments have led to increased emissions resulting in PM mass loadings that frequently exceed the NAAQS of  $100 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{10}$  and  $60 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{2.5}$  mass loadings. This exposes the residents to hazardous levels of PM throughout the year.

Daily PM mass loadings show a clear correlation with daily mortality and morbidity from respiratory and cardiovascular diseases (Englert, 2004; Kappos et al., 2004; Pope III and Dockery, 2006). The correlation between extreme PM mass loadings and mortality was recognised early in the history of air pollution research (Firket, 1931; Schrenk et al., 1949; Nemery et al., 2001) and the predicted disaster of the “London Fog” (Logan, 1953; Bell and Davis, 2001) resulted in first efforts to combat PM air pollution through legislation and regulatory intervention. However, the effect of moderate to low PM mass loadings on human health was recognised much later (Shy, 1979; Ware et al., 1981; Dockery et al., 1989; Dockery et al., 1993; Schwartz, 1994; Pope III et al., 1995; Pope III, 2000a) and the accumulated evidence has resulted in a revision of the air quality standards in many countries including India (NAAQS, 2009).

Many sites in the densely populated IGP violate the NAAQS throughout the year except during the monsoon season when removal through wet scavenging brings PM levels into compliance.

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

Enhanced PM is associated with increased hospital visits/admission of patients with respiratory symptoms and increased mortality (Mohanraj and Azeez, 2004; Nag et al., 2005; Pandey et al., 2005; Kaushik et al., 2006). The complex interplay of natural windblown dust, transboundary air pollution, and local sources impose severe challenges on the Central Pollution Control Board, the local regulatory body responsible for enforcing the NAAQS. While it is clearly inappropriate to blame individual industrial units for natural windblown dust or trans-boundary air pollution, the contribution of these two factors to extreme events should not be used as an excuse to avoid all regulatory action. In this study, I seek to quantify the effect of long-range transport of both natural windblown dust and anthropogenic PM to the regional background PM mass loadings and establish a baseline against which the enhancement due to local sources can be measured.

Back-trajectory models use archived meteorological data and allow for the identification of source regions of pollutants measured at a receptor site. Air-mass trajectories are defined as the path of an infinitesimally small air parcel. Back trajectories trace the air mass back in time and describe where the air mass reaching a receptor site originated.

Statistical analysis of large data sets has been a popular tool for identifying source regions of particulate matter. Such analysis attributes all changes in particulate matter mass loading at a receptor site to spatially fixed sources and seeks to identify those sources by investigating the statistical correlation between air-mass origin and the particulate mass loadings observed at the receptor site. While wind rose or pollution rose plots are most appropriate for identifying local sources (Fleming et al., 2012), analysis of a large set of back trajectories (Stohl, 1996, 1998) has been a popular tool for identifying distant source regions of particulate matter (Buchanan et al., 2002; Abdalmogith and Harrison, 2005; Borge et al., 2007; Nyanganyura et al., 2008) and investigating trans-boundary particulate matter pollution (Borge et al., 2007; Grivas et al., 2008; Miller et al., 2010). Cluster analysis is a multivariate statistical technique that splits the data into several groups while maximising the homogeneity within each group and maximising the distance between groups.

The aim of this chapter is to better understand the conditions under which PM mass loadings exceeding the national ambient air quality standard (NAAQS) of  $100 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{10}$  and  $60 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{2.5}$  (NAAQS, 2009) occur in the NW-IGP and to quantify the contribution of long-range transport to those exceedance events. I quantify the contribution of long-range transport to fine ( $\text{PM}_{2.5}$ ) and coarse ( $\text{PM}_{10-2.5}$ ) PM using back-trajectory cluster analysis, pinpoint potential source regions of enhanced background PM mass loadings, and further attempt to constrain the origin of the particulate matter by correlating the observations with those

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

of gas-phase combustion tracers (CO, NO<sub>2</sub>, benzene, and acetonitrile). I analyse a 2-year data set (August 2011 till June 2013) measured at the Atmospheric Chemistry facility of the Indian Institute for Science Education and Research (IISER) Mohali.

## 3.3. Materials and Methods

### 3.3.1. Study location, air quality data, and general meteorology

In this chapter, I use a 2-year (August 2011 till June 2013) dataset of hourly averaged fine (PM<sub>2.5</sub>) and coarse (PM<sub>10-2.5</sub>) PM and gas-phase combustion tracers (CO, NO<sub>2</sub>, benzene, and acetonitrile) measured at the Atmospheric Chemistry facility of the Indian Institute of Science Education and Research Mohali. Figure 3.1 shows the location of the city of Mohali in the north-western Indo-Gangetic Plain in the Indian state of Punjab, close to the forested slopes of the foothills of the Himalayan mountain range on the left side. The measurement facility is located at a suburban site, south-west of the city centre of the “tri-city” – an urban agglomeration of the three cities of Chandigarh, Mohali, and Panchkula – inside the residential campus of IISER Mohali (30.667 °N, 76.729 °E; 310 m a.m.s.l.). On the right side, Figure 3.1 shows a close-up illustrating the exact location of the measurement facility and its spatial relationship with respect to the nearby cities and potential local point sources of particulate matter and the mountain range. The inlets of all instruments and the meteorological sensors are co-located and placed at a measurement height of 20 m a.g.l. (Sinha et al., 2014a).

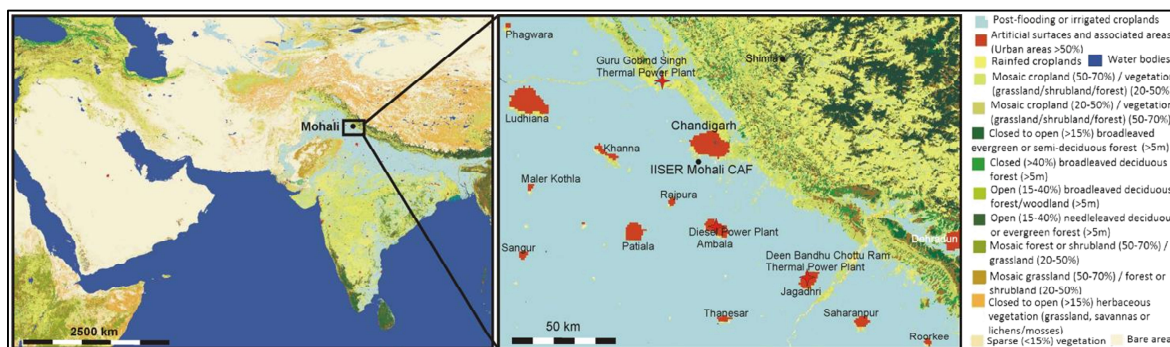


Figure 3.1: (Left) Location of Mohali on a land classification map (courtesy ESA GlobCover 2009 project). The site is located in the north-western Indo-Gangetic Plain, close to the forested slopes of the foothills of the Himalayan mountain range. (Right) The exact location of the measurement facility and its spatial relationship with respect to the nearby cities, the mountain range, and potential local point sources of particulate matter.

Particulate matter (PM<sub>10</sub> and PM<sub>2.5</sub>) mass concentrations were measured using separate Thermo Fischer Scientific 5014i beta-continuous ambient particulate monitors working on the principle of  $\beta$ -attenuation. A vacuum pump operating at a constant flow rate of 16.67 LPM was used to draw in ambient air carrying particles through size-selective inlets. The attenuation of  $\beta$  rays due to

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

ambient particles collected on a quartz fibre filter tape was used to estimate the mass concentration of airborne PM in real-time. A dynamic inlet heater regulated the sample RH at 40% whenever the ambient RH was more than 40%. Under dry conditions (ambient RH < 40%), the heater did not operate, and the measurement was performed under ambient conditions. To ensure data quality of PM measurements, mass coefficient calibration of analyzers was performed periodically (twice in a year) using NIST- traceable foil sets.

NO<sub>2</sub> measurements were performed using a Thermo Fischer Scientific 42i trace-level analyser, which works on the principle of chemiluminescence. A photomultiplier tube (PMT) in the instrument detects and quantifies the radiation emitted by electronically excited NO<sub>2</sub> molecules (far visible to short infrared; wavelength range 600-3000 nm) formed from the reaction between NO and ozone. An “Ozonator” in the instrument produces ozone in excess to react with NO. The instrument operates in two modes, namely NO and NO<sub>x</sub> mode. In the NO mode, the NO in the sample directly reacts with ozone to give the NO mixing ratio. In the NO<sub>x</sub> mode, NO<sub>2</sub> in the sample line is first converted to NO via reaction with a molybdenum converter maintained at 325 °C. This total NO (original NO in the sample along with NO converted from NO<sub>2</sub>) then reacts with ozone to give the mixing ratio of NO<sub>x</sub>. The difference between NO<sub>x</sub> and NO mixing ratios gives the ambient NO<sub>2</sub> mixing ratio. Apart from NO<sub>2</sub>, some other ambient trace species containing active nitrogen (nitric acid, nitrous acid, nitrate radical, alkyl nitrates, etc.) also get converted to NO after reacting with a molybdenum converter. Therefore, NO<sub>2</sub> reported in this chapter much be considered as an upper limit to ambient NO<sub>2</sub>.

Carbon monoxide (CO) was measured using a Thermo Fischer Scientific 48i trace-level enhanced analyser that works on the principle of gas filter correlation (GFC) nondispersive infrared (NDIR) absorbance. The detection principle is based on the absorbance of infrared radiation (4.6 μm) by CO molecule. The radiation from the IR source is chopped and passed through a gas filter alternating between CO and N<sub>2</sub>. The beam passing through the CO filter is completely attenuated and thus cannot be further absorbed by the CO in the sample. This beam serves as the reference beam. On the other hand, N<sub>2</sub> being transparent to infrared radiation, the beam passing through the N<sub>2</sub> filter is absorbed by the CO in the sample. This beam serves as the measurement beam. Thus, the periodic modulation observed in the detector signal due to the gas filter correlation wheel can solely be attributed to CO in the sample as other gases absorb radiation equal radiation from both the measurement and the reference beam.

Mixing ratios of benzene and acetonitrile were determined using a high-sensitivity proton transfer reaction quadrupole mass spectrometer (HS model 11-07HS-088; Ionicon Analytik Gesellschaft,

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

Austria). Briefly, an ion source in the PTR-MS produces high purity hydronium ions. In the reaction chamber, VOCs in the ambient air having a proton affinity higher than that of water get protonated after reacting with hydronium ions. A quadrupole then separates the protonated ions based on their mass to charges ( $m/z$ ) ratio, where a secondary electron multiplier detector quantifies the ion signal.



Benzene and acetonitrile were detected at their respective protonated organic ions signals at  $m/z$  79 and  $m/z$  42.

Figure 3.2 shows wind rose plots for the winter (December–February), summer (March–June), monsoon (July–September), and post-monsoon (October and November) seasons. Most air masses impacting the site travel parallel to the mountain range and reach the facility from a north-westerly or south-easterly direction.

Periods of calm (wind speed  $< 1 \text{ m s}^{-1}$ ) account for only 4.5, 2.5, 5.2, and 8.7% of the total time during the winter, summer, monsoon, and post-monsoon seasons respectively, and slow transport (wind speed  $1\text{--}5 \text{ m s}^{-1}$ ) was observed 64.1, 48.7, 56.1 and 71.4% of the total time, respectively. The high frequency with which rapid transport of air masses towards the facility (wind speed  $> 5 \text{ m s}^{-1}$ ) was observed (31.4, 48.8, 38.7, and 19.9% of the total time during the winter, summer, monsoon, and post-monsoon seasons respectively) indicates that long-range transport potentially plays a significant role in determining pollutant loadings at the site.

The general meteorology of the site is as follows: During the winter season, weak northerlies or north-westerlies and a weak, low-level anti-cyclonic circulation prevail in the NW-IGP. The surface pressure map, the surface winds, and 700 hPa winds of NCEP Reanalysis derived data provided by the NOAA Physical Science Division, Boulder, Colorado, USA, are shown in Figure 3.3a–c. Wintertime fog occurs frequently, and fog formation is favored by the subsidence of air masses over the IGP, low temperatures, high relative humidity, and low wind speeds ( $< 5 \text{ m s}^{-1}$ ). However, ground-level wind speeds at this site are generally not as low as the surface wind speeds of the NCAR Reanalysis data set would suggest. The seasonally averaged experimentally observed wind speed is  $4.4 \text{ m s}^{-1}$  and not  $< 2 \text{ m s}^{-1}$ .

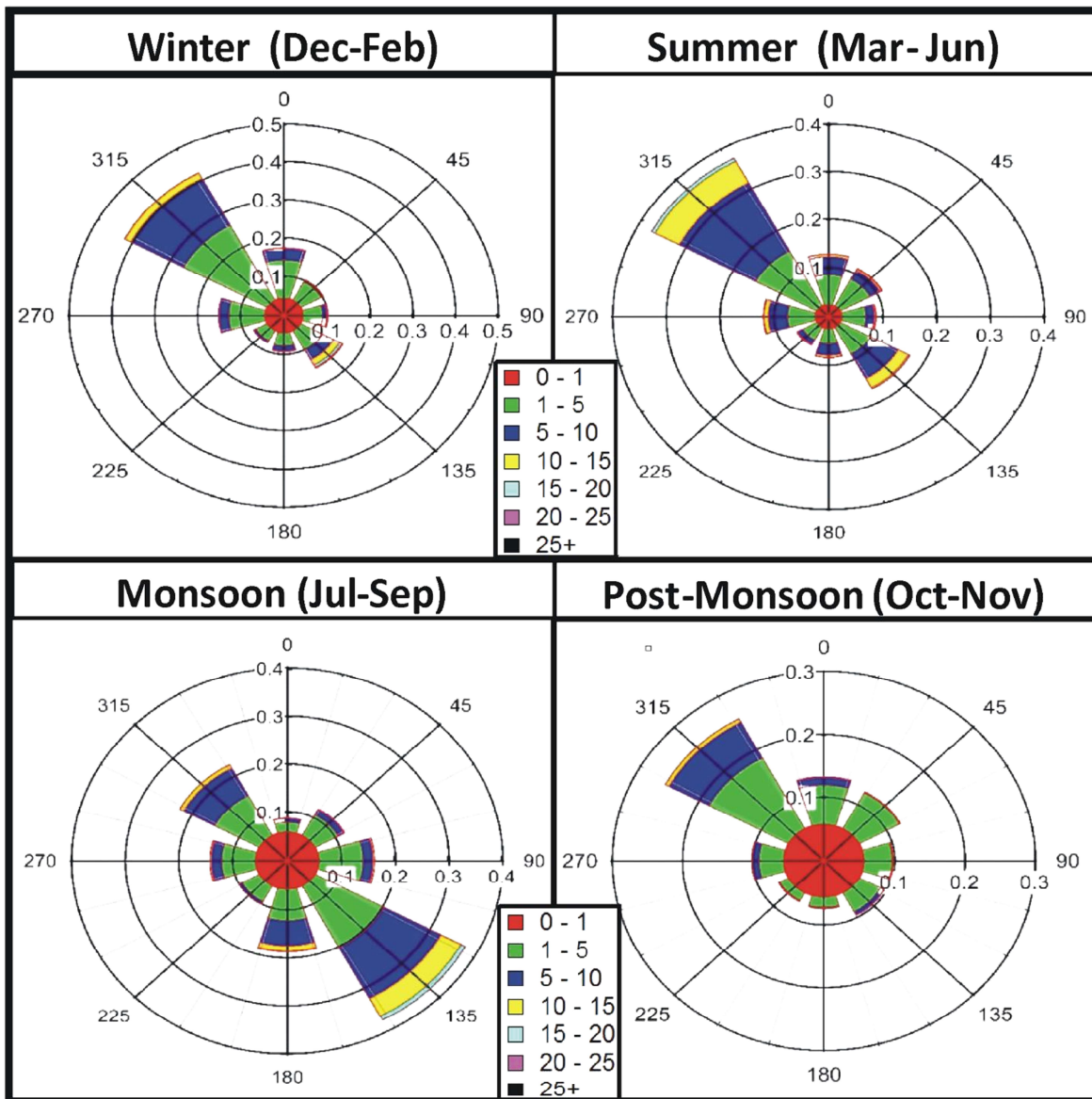


Figure 3.2: Wind rose plot for the measurement site for the winter (December–February), summer (March–June), monsoon (July–September), and post-monsoon (October and November) seasons. Wind speed and wind direction were measured at the height of 20 m a.g.l.

This underestimation of the surface wind speeds in the NW-IGP is not unique to this particular model, and meteorological data set but also applies to, e.g., Figure 6 in Lawrence and Lelieveld (2010). Sporadic winter rains are generally associated with western disturbance (Pisharoty and Desai, 1956; Mooley, 1957; Agnihotri and Singh, 1982; Dimri, 2004). The western disturbance is a terrain-locked low-pressure system that forms when an upper-level extratropical storm originating over the Mediterranean passes over the notch formed by the Himalayas and the Hindu Kush mountains. The resulting notch depression is small, 5° latitude/longitude in size. It develops within an existing trough in the belt of subtropical westerly wind. South-westerly wind ahead of



### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

the trough brings moisture from the Arabian Sea, which encounters the Western Himalayas that lie almost normal to this moist wind. Part of the wind is channelled into the IGP, which subsequently reaches the receptor site from the south-east. Fast westerly winds in winter are typically associated with a strong subtropical jet stream poised over westerly troughs.

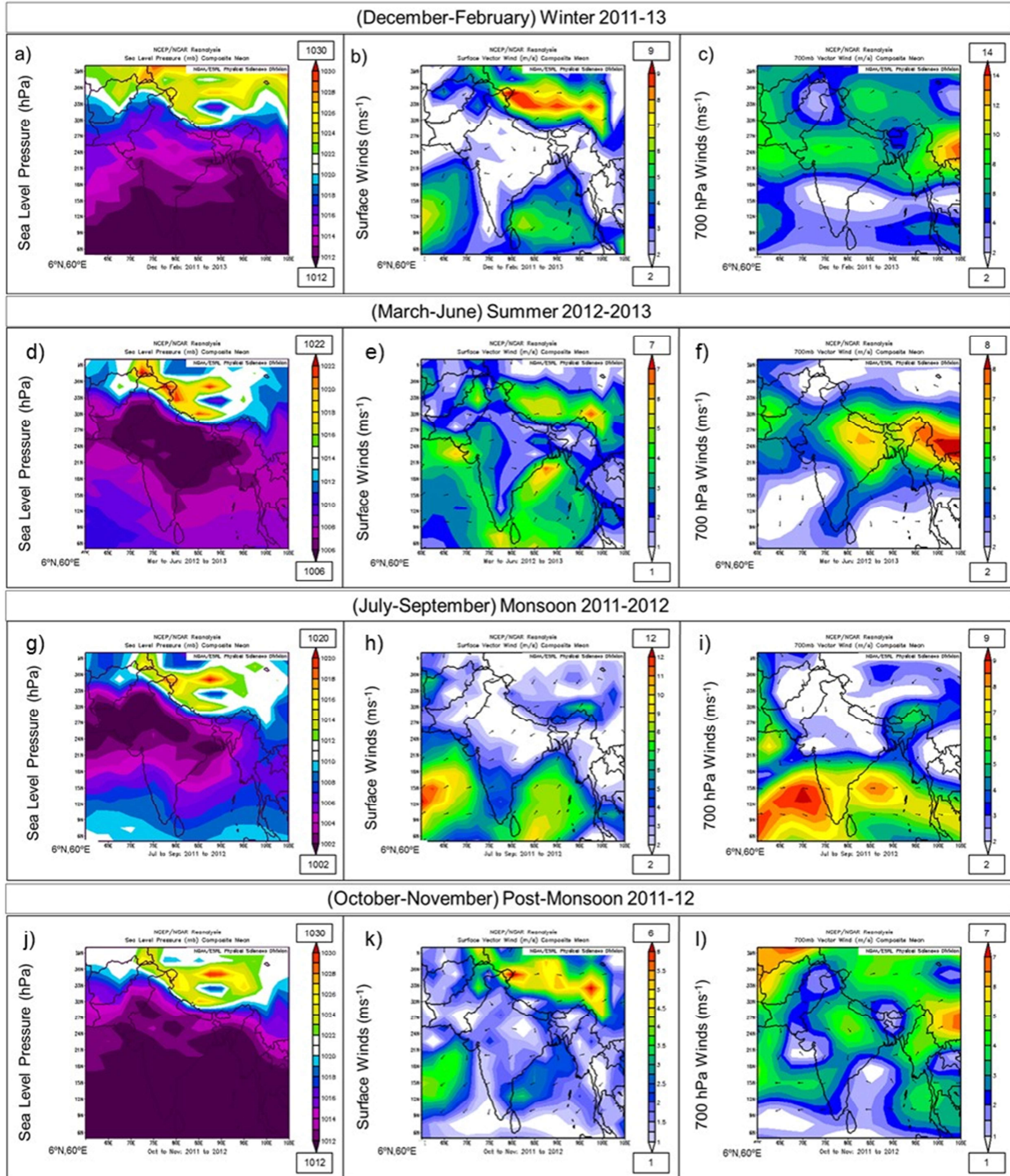


Figure 3.3: Seasonal composites of sea level pressure, surface winds, and 700hPa winds. The values on the top and bottom of the legend in each figure show the maximum and minimum, respectively. (Images provided by the NOAA-SRL Physical Sciences Division, Boulder Colorado from their Web site at <http://www.esrl.noaa.gov/psd/>)

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

During the summer season, the prevailing wind direction is north-westerly. Tropospheric subsidence over north-western India due to the “heat low” associated with westerly flow across Afghanistan and Pakistan channels air masses originating in the Middle East into the IGP (Figure 3.3d–f). The boundary layer experiences a strong temperature inversion (Das, 1962) due to dust-induced cooling in the upper layers. The resulting steep horizontal pressure gradient is responsible for strong surface winds that carry dust and sand storms (Bryson and Swain, 1981). These loo winds are extremely hot and dry and are not adequately resolved by the NCEP Reanalysis meteorological data set (Figure 3.3d–f), which shows a very moderate average surface wind speed of 1.5–2.5 m s<sup>-1</sup> over the NW-IGP that stands in stark contrast to the observed seasonally average wind speed of 5.6 m s<sup>-1</sup>. During April, the centre of the subtropical jet stream is located over northern India. It gives rise to cold subtropical westerlies in the upper troposphere. Simultaneously, the lower troposphere reaches very high temperatures due to the quick response of the land to the overhead Sun. This favours severe thunderstorms with strong squalls at the leading edge of the downdraft during the summer season. Such thunderstorms can cause convective dust storms, locally known as Aandhi (Ramaswamy, 1956; Joseph, 1982). In March and April, south-easterly winds are generally associated with the western disturbance. The climatology of the western disturbance shows a significant inter-annual variability. It is correlated with the polar/Eurasia teleconnection pattern. A weak circumpolar vortex can enhance the number of notch depressions and late “winter storms” in March and April (Lang and Barros, 2004). South-easterly winds in June are generally associated with an earlier than usual onset of the monsoon. In normal years, the monsoon reaches Punjab in the first week of July.

During the monsoon season, the surface heat low located over the Pakistan region and the monsoon trough stretching from the NW-IGP to the Bay of Bengal dominate the general circulation over the IGP (Figure 3.3g–i). The strength and position of the monsoon trough drive the “active-break” cycles of the rains on the intra-seasonal scale (Sikka and Gadgil, 1980; Goswami, 1998; Goswami et al., 2006a; Goswami et al., 2006b). During “active” spells, the trough is located over southern or central India, and cyclonic swirls form all across the trough. The prevailing wind direction during active spells is south-easterly. During “break” spells, the trough is located over the foothills of the Himalayas, the low-level jet originating off the coast of Somalia enters the IGP through the Indus Valley (Joseph and Raman, 1966). Cyclonic swirls are mostly absent, and rains are suppressed everywhere except over the foothills (Joseph and Sijikumar, 2004). During break spells of the monsoon circulation, the prevailing wind direction is north-westerly. Break spells are associated with lower tropospheric inversions, dusty winds, and lower troposphere anti-cyclonic

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

vorticity over the IGP (Sikka, 2003; Rao and Sikka, 2005; Bhat, 2006). Both the observed wind direction and wind speeds of the surface winds during the monsoon season are poorly resolved by the meteorological data set. While the model suggests low wind speed ( $< 2 \text{ m s}^{-1}$ ), southwesterly winds dominate (Figure 3.3g–i), actual observations show that north-westerly winds dominate during break spells, and south-easterly winds during active spells. The average wind speed is  $5.1 \text{ m s}^{-1}$ . Most rainfall events occur during the “active spells” when the wind direction is south-easterly. However, occasional nighttime rainfall events are observed even during break spells in Punjab.

During the post-monsoon season, the prevailing wind direction is north-westerly. The retreating monsoon brings subsidence of dry central Asian air masses over the NW-IGP. In particular, during nighttime, katabatic winds reach the receptor site from the northern to eastern wind sectors. Winds are generally weak; the wind speed is less than  $5 \text{ m s}^{-1}$  for more than 80% of the time. The post-monsoon season shows the lowest discrepancy between modelled surface winds ( $1.5 - 2 \text{ m s}^{-1}$ ) and the observed wind speed ( $3.4 \text{ m s}^{-1}$ ) (Figure 3.3j–l). Surface wind vectors represent the observed nighttime flow at this site better, while the 700hPa wind vectors represent daytime observations better.

Since the purpose of this study is to investigate the contribution of long-range transport to PM pollution, I restricted my analysis to measurements obtained between 12:00 and 16:00LT (UTC+05:30) during the day and between 03:00 and 06:00LT (UTC+05:30) at night as depicted in Figure 3.4 and consider calm conditions with wind speeds of less than  $1 \text{ m s}^{-1}$  separately. The daytime period was selected because the local daytime boundary layer reaches its maximum height around 14:00, and the contribution of long-range transport to PM is highest in the period centered on this time. The nighttime period was selected because the contribution of local sources, in particular the contribution of local traffic, construction activity, and biomass combustion to air pollution, is least during late night and early morning hours.

Figure 3.4 shows that both fine-mode PM ( $\text{PM}_{2.5}$ , bottom panels) and coarse-mode PM ( $\text{PM}_{10-2.5}$ , top panels) show a bimodal distribution with a peak in the morning and the evening due to local traffic and biomass combustion emissions and a low in the mid-day hours and the early morning hours during all seasons (Figure 3.4). The bimodal behavior is most pronounced during the post-monsoon season when low wind speeds prevail, and local and regional sources dominate the aerosol mass loading. The bimodal behavior is weakest during the monsoon season when active convection and high wind speeds reduce the influence of local sources. Fine-mode PM shows the lowest mass loadings during mid-day, when the boundary layer is highest, while coarse-mode PM

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

shows the lowest mass loadings in the early morning hours. During the summer and monsoon seasons, there is a pronounced difference between mean and median coarse-mode PM loadings. This discrepancy is caused by the contribution of episodic events (dust storms) to coarse-mode PM in both seasons.

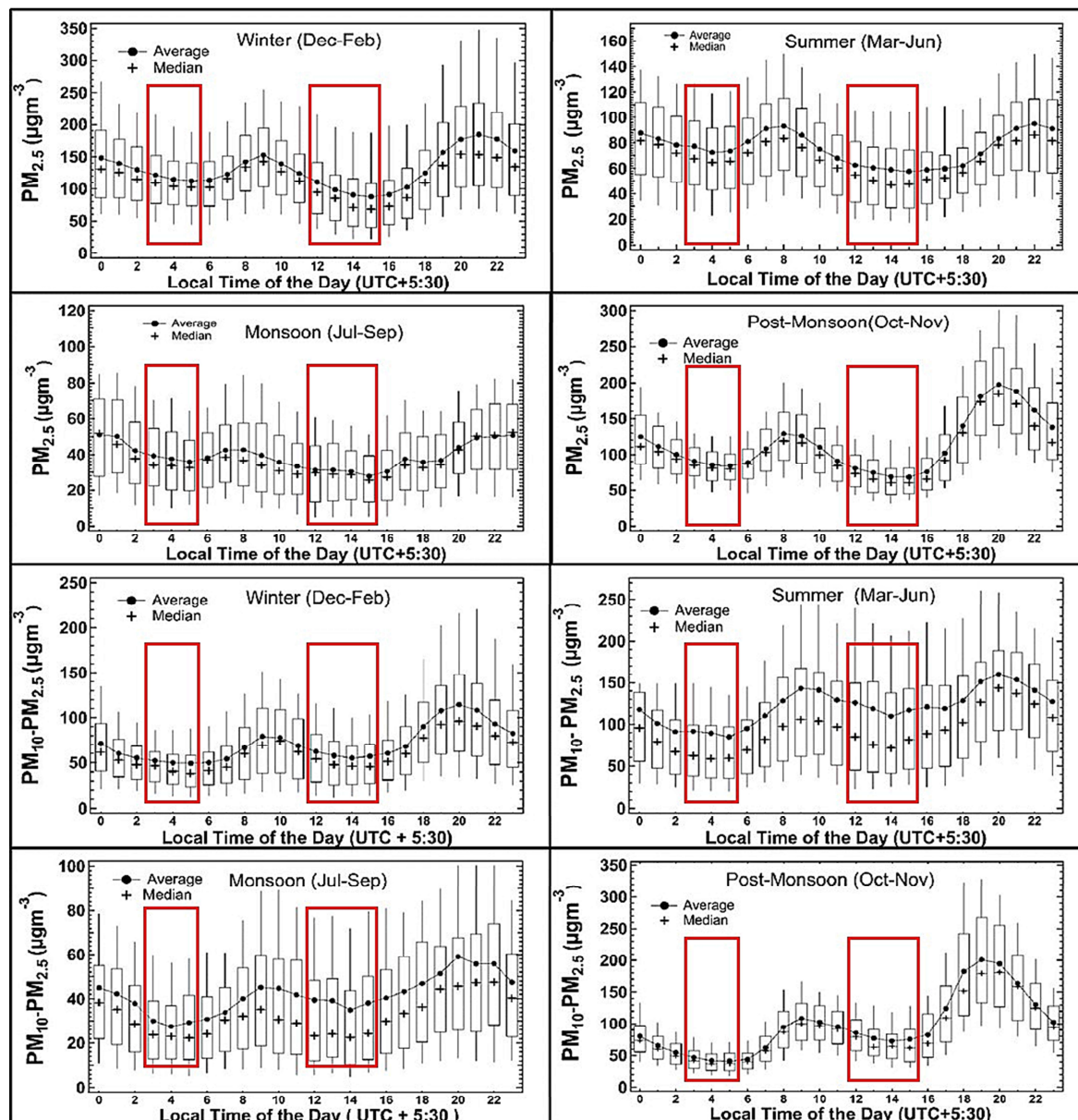


Figure 3.4: Diel box and whisker plots for fine-mode (top four panels) and coarse-mode (bottom four panels) particulate matter for the winter, summer, monsoon, and post-monsoon seasons for the period November 2011 to August 2013, respectively. The box indicates the upper and lower quarter values; the cross indicates the median, and the dots connected by lines provide the mean. The whiskers indicate the 5<sup>th</sup> and 95<sup>th</sup> percentiles, respectively. Periods of calm ( $< 1 \text{ m s}^{-1}$ ) have been excluded while preparing the graph. The interval highlighted in red shows the daytime low (12:00 to 16:00LT) and nighttime low (03:00 to 06:00LT).

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

For the back-trajectory analysis, I calculated the median PM mass loading and gas-phase mixing ratio of CO, benzene, and acetonitrile for the period in question. For rainfall, I provided the sum (total precipitation in the time window in question) rather than the median, as I consider the total rainfall a better indicator of wet scavenging.

#### **3.3.2. Back trajectory modelling**

I computed 3-day (72 h) backward trajectories using the HYSPLIT\_4 (HYbrid Single-Particle Lagrangian Integrated Trajectory) model in ensemble mode using the National Oceanic and Atmospheric Administration's Global Data Acquisition System (GDAS) meteorology (Draxler and Hess, 1998; Draxler and Rolph, 2013) as input. I calculated ensemble runs for air masses arriving at the site (30.667° N, 76.729° E) at 20m above the ground (the approximate sampling height for all instruments). The model was run twice daily with an arrival time of 09:00 UTC (14:30 local daytime) and 23:00 UTC (04:30 local nighttime). Due to the proximity of the site to the Himalayan mountain range, the trajectory output was susceptible to the model's input data. The IISER Mohali air quality station (30.667° N, 76.729° E) is located in the Indo-Gangetic Plain (IGP) at an altitude of 310 m a.m.s.l. approximately 20 km south-west of the Shivalik hills, but the model's terrain height at the receptor site is 667.6 m a.m.s.l. for the GDAS meteorological data field and 1249.9 m a.m.s.l. for the reanalysis meteorological data field for a single trajectory arriving at the site. I calculated trajectory ensembles for this site using both data sets. The trajectory ensemble option starts multiple trajectories from the selected starting location. Each member of the trajectory ensemble is calculated by offsetting the meteorological data by a fixed grid factor (1 meteorological grid point in the horizontal and 0.01 sigma units in the vertical), which results in 27 trajectories. The model's terrain height varied between 200 and 3500 m a.m.s.l. for individual runs of the trajectory ensemble for the GDAS meteorological data field and between 240 and 5100 m a.m.s.l. for the reanalysis meteorological data field. I conclude that the GDAS meteorological data field performs better in modelling the terrain at this site. All further analysis in this work uses this meteorological data field.

Out of the 27 trajectories in the ensemble run, only 3 trajectories are consistent with the measurement site, Mohali, being located in the plain (< 400 m a.m.s.l.) and Shimla, a mountain site 60 km north-east of this site (Figure 3.1) at 31.103° N, 77.172° E, being located in the mountains (> 400 m a.m.s.l.). To ensure that this sensitivity of the model does not affect the results of this study, I selected these three trajectories for further analysis.

Borge et al. (2007) have recently emphasized the importance of specifying adequate arrival heights for the outcome due to possible considerable variation in the wind speed and direction with

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

height above the ground. Kassomenos et al. (2010), too, found a dependency of clustering outcome on the arrival heights of the trajectories. At this site, I find that the Himalayan mountain range acts as a great barrier, and air masses are funneled into the IGP. Consequently, there is a little dependency of the trajectory run outcome on arrival heights  $\leq 500$  m a.g.l., as long as the meteorological input data set represents the Himalayan mountain range topography adequately. The local wind direction observed at the site generally agrees with the wind vector at 10 m a.g.l. used in the GDAS meteorological data field.

#### 3.3.3. Back trajectory cluster analysis

Cluster analysis is a statistical method used to group data in large data sets into a small number of groups of similar data known as clusters. In this work, I have used air masses' trajectory coordinates (time steps) as the clustering variables. A non-hierarchical method known as the  $k$ -means procedure has been used in this study. The air mass back trajectories were subjected to  $k$ -means clustering using a freeware called PAST (PAleontological STatistics). The number of clusters " $k$ " is to be specified by the user prior to clustering. The assignment of back trajectories to clusters is initially random. In an iterative procedure, trajectories are then moved to the cluster which has the closest cluster mean in Euclidean distance, and the cluster means are updated accordingly (Bow, 1984). This process continues until the process of hopping from one cluster to another ceases. As a typical artifact of the  $k$ -means clustering algorithm, the result depends on the seed used for clustering (Kassomenos et al., 2010). Therefore, to get a robust result, clustering was initialized with 11 different trajectory orders. I selected the clustering result with the lowest root mean square difference for each predefined number of clusters.

The root mean square difference for an individual latitude and longitude value is given by:

$$RMSD = \sqrt{(x_i - \bar{x})^2} + \sqrt{(y_i - \bar{y})^2} \quad (42)$$

where  $x_i$  and  $y_i$  stand for the latitude and longitude of the individual trajectory for a given hour and  $\bar{x}$  and  $\bar{y}$  for the cluster mean latitude and longitude of the same hour for the cluster to which that trajectory belongs. The total root mean square difference t-RMSD is the sum over all the RMSD values for a given set of back trajectories.

I selected the optimal number of clusters that best describe the different air-flow patterns to Mohali site by computing the change in the minimum t-RMSD while increasing the number of clusters from  $n$  to  $n+1$ . This change in the minimum t-RMSD decreases abruptly as clusters of trajectories that are significantly different in terms of wind directions, and speeds are separated from each other (Dorling et al., 1992). A threshold of 5% change has been adopted (Dorling et al., 1992; Brankov et al., 1998) to indicate the number of clusters to be retained.

## 3.4. Results and discussion

### 3.4.1. Optimisation of the number of clusters

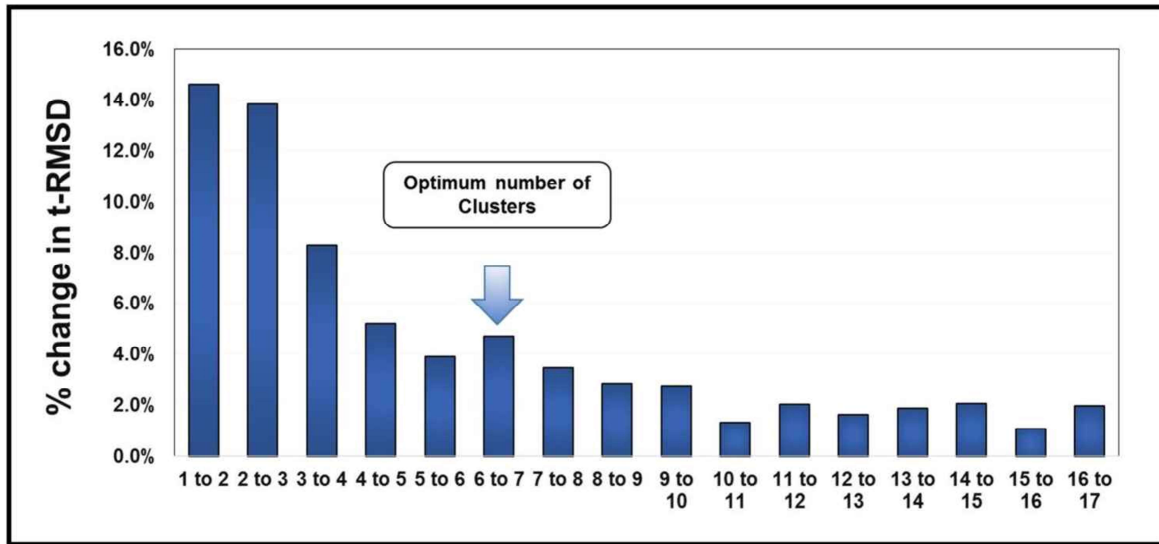


Figure 3.5: Percentage change in t-RMSD with increase in cluster number.

The number of clusters was optimised by minimising the t-RMSD change for an increase in the number of clusters, while retaining as few clusters as possible. Figure 3.5 shows the percent change in t-RMSD for a subsequent increase in the number of clusters. The largest % decrease in t-RMSD is observed when the number of clusters is increased to 2 and subsequently 3. This corresponds to two major airflow corridors (south-easterly flow and westerly flow) and a local cluster. Increasing the number of clusters beyond this number allows splitting of the air masses in the western and south-eastern corridors into several groups according to their transport speed and introduces a south-westerly cluster.

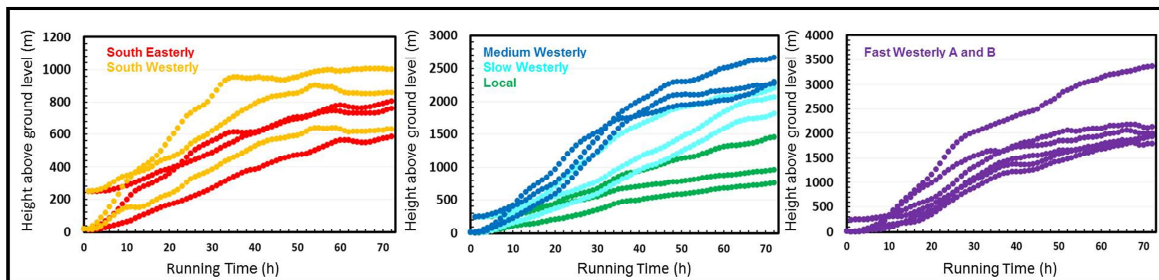


Figure 3.6: Mean height (a.g.l.) of all trajectories in an individual cluster as a function of trajectory running time (72h) for trajectories arriving at 09:00 and 23:00 UTC.

Initially, I identified the optimum number of clusters as seven; however, two of these seven clusters (fast north-westerly flow A and B) were classified into two different groups only, because, for the fast north-westerly flow A, all three trajectories of the ensemble run showed equal transport speed.

In contrast, for fast north-westerly flow B, two of the three trajectory solutions in the ensemble indicated slow air mass transport (trajectories arriving at 20 m a.g.l., Figure 3.6), while one solution supported rapid transport (trajectory arriving at 250 m a.g.l., Figure 3.6). Locally measured meteorological parameters indicated that both clusters are associated with hot, dry loo winds and dust storms during summer and above-average wind speeds during other seasons. Therefore, both clusters were combined into one fast westerly cluster, and the final number of optimum clusters was six.

#### **3.4.2. Spatial and dynamic patterns of the airflow associated with the clusters**

Figure 3.7a–f shows the trajectories of the cluster averages and individual trajectories associated with each cluster identified by  $k$ -means clustering superimposed on a land classification map (courtesy of the ESA GlobCover 2009 project). The length of each mean trajectory is 3 days, and the distance between two successive data points represents the 1-hour interval. I find six distinct flow patterns, south-easterly flow ( $N = 263$ ), south-westerly flow ( $N = 79$ ), fast ( $N = 78$ ), medium ( $N = 83$ ) and slow ( $N = 305$ ) westerly flow, and a local cluster ( $N = 556$ ). Although considerable variability is observed between the individual trajectories contributing to each of the clusters, particularly for the “local” cluster (Figure 3.7d), each cluster of trajectories represents a distinct air-mass fetch region. Figure 3.6 depicts the mean height (m a.g.l.) of the mean trajectories during the 3 days before they arrived at the receptor site. Trajectories from the slow, medium, and fast westerly clusters descend from the free troposphere to reach the receptor site. Fast and medium westerly air masses show a rapid descent in the last 30h before reaching the receptor site. They are generally associated with high wind speeds, while slow westerly air masses display gradual subsidence. The south-easterly, south-westerly and local clusters remain within the convective boundary layer throughout. Figure 3.8 shows the temporal distribution of the air flows for the six clusters determined with respect to the different seasons, and Table 3.1 presents the locally measured meteorological parameters for daytime/nighttime for each cluster and season.



### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

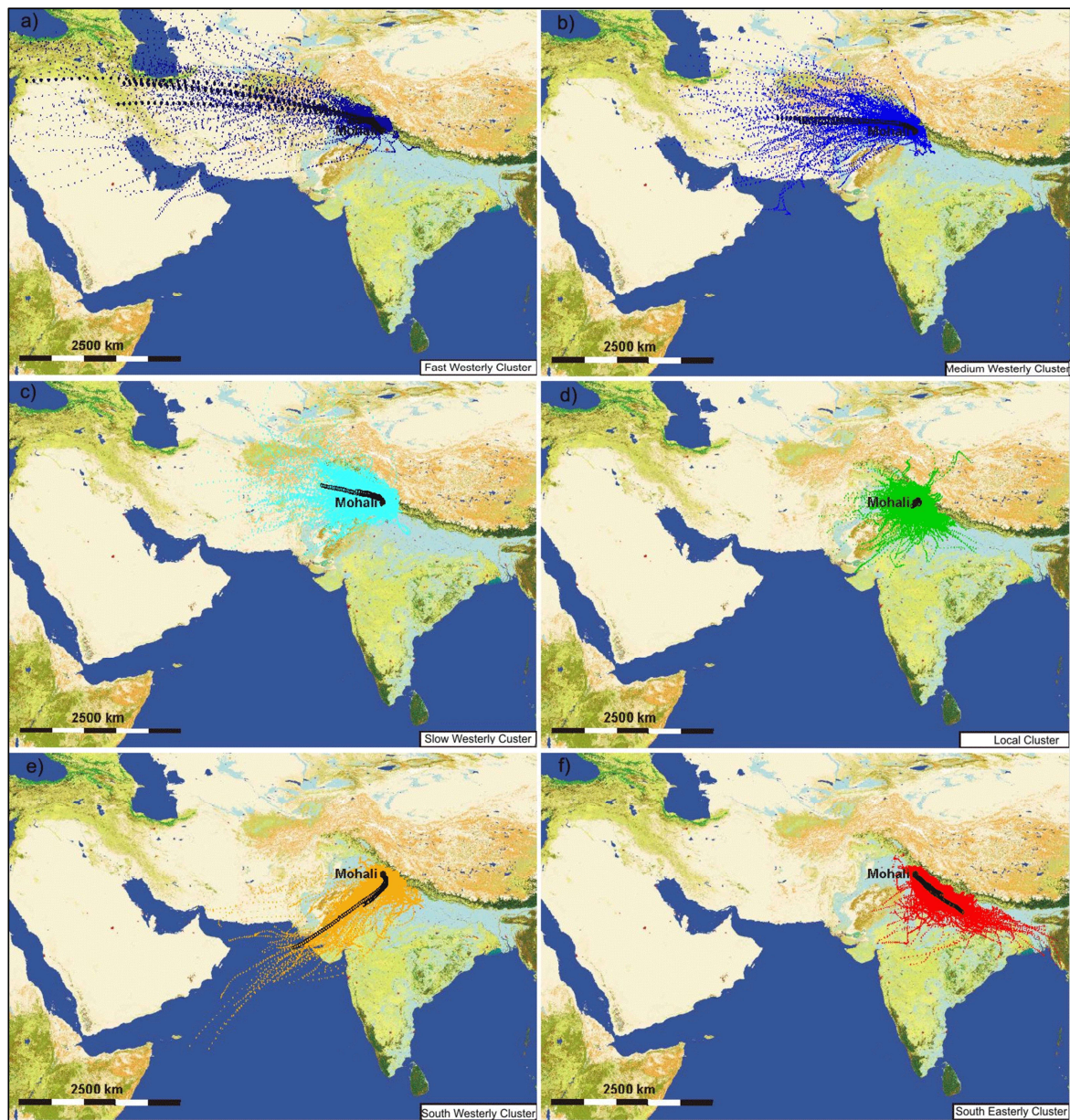


Figure 3.7: (a–f) All individual trajectories that contributed to each of the clusters and the cluster mean, superimposed on a land classification map (courtesy of the ESA GlobCover 2009 project). The length of each mean trajectory is 3 days, and the distance between two successive data points represents a 1h interval. The average trajectory of each cluster has been superimposed using circles with a black outline.

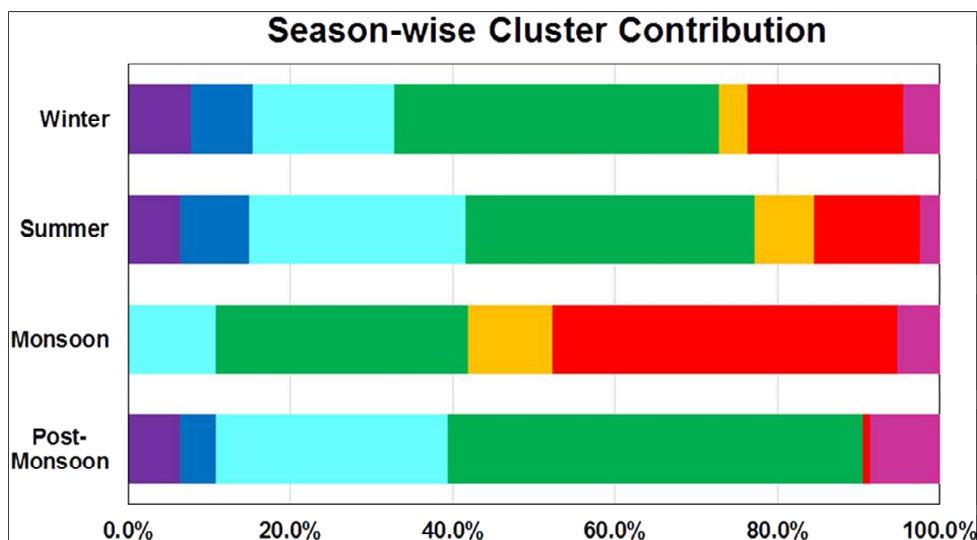


Figure 3.8: Contribution of individual clusters to air-mass flow for all four seasons. Magenta: calm; red: south-easterly cluster; orange: south-westerly cluster; green: local cluster; light blue: slow westerly cluster; dark blue: medium westerly cluster; purple: fast westerly cluster.

The *local cluster* accounts for 40.1, 35.7, 31.1, and 51.3% of the air mass trajectories during winter (December to February), summer (March to June), monsoon (July to September), and post-monsoon (October and November) season, respectively. This air mass transport corresponds to times when weak northerlies or north-westerlies and a weak, low-level anti-cyclonic circulation prevail in the NW-IGP during winter, summer, and post-monsoon season. The predominant locally measured wind direction for this cluster during winter, summer, and the post-monsoon season is west to north-west (47% of the time) with katabatic winds (320 – 120° wind sector) accounting for most of the remainder (31%). South easterly winds (8%) and south westerly winds (14%) account only for a minor fraction of the locally observed wind direction. During summer season and monsoon season convective activity above the site is also attributed to this cluster. Strong squalls leading to “Aandhi” type convective dust storms (Joseph, 1982) are observed occasionally, both in late summer and early in the monsoon season. During monsoon season, locally observed wind direction associated with this cluster is variable (35% north west, 29% south east, 18% south west, 18% katabatic flow). The local cluster is associated with average temperatures and wind speeds at the lower end of those observed for the different clusters during day and night in all seasons. The absolute humidity is higher than the absolute humidity observed for medium and fast westerly flows and lower than the absolute humidity observed for south easterly and south westerly flow during all seasons. Occasionally rain events occur in all seasons. However, the total number of rain events associated with the local cluster is low except during the monsoon season.

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

	Fast westerly	Medium westerly	Slow westerly	Local	South westerly	South-easterly	Calm
WINTER (Dec-Feb)							
T (°C)	18.9/8.5	17.7/9.4	17.5/9.3	17.7/9.8	15.6/12.7	18.3/12.0	16.6/10.4
RH (%)	43.6/82.5	49.5/79.6	50.0/79.5	53.7/79.9	57.8/74.5	57.2/77.8	62.2/79.7
Wind speed (m s <sup>-1</sup> )	7.0/5.0	8.2/3.2	5.7/3.0	5.4/3.5	5.4/6.0	6.9/4.8	0.8/0.7
Wind direction	302/319	270/303	306/294	292/304	298/240	241/188	235/139
Absolute humidity (g m <sup>-3</sup> )	7.4/7.2	7.9/7.3	7.8/7.3	8.5/7.6	8.1/8.6	9.4/8.5	9.2/7.9
Solar radiation (W m <sup>-2</sup> )	414	331	381	376	362	338	307
Rain (mm)	-	-	-	7.7 (2)	-	50.7(11)	27.3(8)
SUMMER (Mar-Jun)							
T (°C)	31.7/19.3	32.6/20.3	35.9/26.5	32.1/22.7	32.5/25.2	29.5/22.9	34.2/23.3
RH (%)	23.4/58.8	22.1/50.7	24.3/44.7	28.7/51.8	36.2/52.5	44.1/64.1	29.0/53.3
Wind speed (m s <sup>-1</sup> )	9.1/5.6	7.6/3.6	7.7/4.5	7.2/4.3	7.9/4.5	7.0/5.6	0.9/0.7
Wind direction	312/308	306/124	288/280	298/264	182/135	171/152	209/186
Absolute humidity (g m <sup>-3</sup> )	8.1/10.3	8.0/9.4	10.3/11.7	10.0/11.0	12.9/12.8	13.5/13.8	11.3/11.7
Solar radiation (W m <sup>-2</sup> )	633	607	593	586	519	569	548
Rain (mm)	-	-	8.5 (3)	28.5 (5)	18.6 (2)	35.8 (10)	27 (27)
MONSOON (Jul-Sep)							
T (°C)	-	-	32.9/25.6	32.4/26.3	30.6/26.0	31.3 /27.1	32.3/26.5
RH (%)	-	-	50.5/80.1	57.5/82.0	68.0/85.2	64.9/81.8	61.7/83.6
Wind speed (m s <sup>-1</sup> )	-	-	8.5/4.5	6.4/3.2	5.4/4.1	6.3/4.2	0.8/0.7
Wind direction	-	-	309/191	280/148	182/124	166/128	206/132
Absolute humidity (g m <sup>-3</sup> )	-	-	18.4/19.9	20.4/21.2	22.0/21.7	21.7/22.1	21.8/21/8
Solar radiation (W m <sup>-2</sup> )	-	-	565	515	430	472	499
Rain (mm)	-	-	0.4 (1)	33.8 (4)	74.1 (3)	142.2 (9)	36 (29)
POST-MONSOON (Oct-Nov)							
T (°C)	23.3/12.8	23.6/12.9	28.0 /18.2	27.3/17.3	-	-	28.1/17.4
RH (%)	35.7/72.9	22.3/74.2	34.1/64.2	34.9/67.7	-	-	34.6/68.0
Wind speed (m s <sup>-1</sup> )	6.3/2.9	9.2/3.9	5.5/2.7	5.3/2.6	-	-	0.7/0.7
Wind direction	310/175	313/315	310/71	304/174	-	-	214/115
Absolute humidity (g m <sup>-3</sup> )	7.7/8.5	8.0/8.7	9.7/10.5	9.5/10.4	-	-	9.8/10.6
Solar radiation (W m <sup>-2</sup> )	409	417	447	422	-	-	461
Rain (mm)	-	-	-	3.1 (2)	-	-	0.2 (2)

Table 3.1: Average of the locally measured meteorological parameters for daytime/nighttime for the different clusters and seasons. For solar radiation, I provided the daytime average only. For rain, I calculated the sum of the rainfall instead of the average, and the numbers in brackets represent the number of rain events.

The *slow westerly* cluster is associated with the same general meteorology (weak northerlies or north-westerlies and a weak, low-level anti-cyclonic circulation) as the local cluster, and most of the locally observed parameters are also similar to those of the local cluster. However, the fetch region of the air masses is larger. The predominant local wind direction for this cluster is west to north-west during all seasons, including monsoon (50%) and katabatic winds from the north-northwest to east-southeast sector (28%) account for most of the remainder. South easterly winds (10%) and south westerly winds (12%) account only for a minor fraction of the locally observed wind direction each. The main differences between air masses associated with the local cluster and air masses associated with the slow westerly cluster are as follows: slightly higher temperatures and wind speeds are observed for the slow westerly cluster during most seasons, and both relative and absolute humidity are lower for air masses associated with the slow westerly cluster due to the recent descend of the air masses from the free troposphere. Air masses related to the slow westerly cluster have a shorter residence time in the convective boundary layer over the irrigated fields in

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

the IGP and consequently contain less moisture. The slow westerly cluster accounts for 17.4, 26.6, 10.7, and 28.4% of the air mass transport to the site during winter, summer, monsoon, and post-monsoon season, respectively. Rain events are associated with this cluster only rarely.

The *medium westerly* cluster is observed only during winter, summer, and post-monsoon seasons, and accounts for 7.7, 8.5, and 4.5% of the air masses, respectively. The clusters are associated with a strong subtropical jet stream poised over westerly troughs and show higher than average wind speeds. The predominant local wind direction for this cluster is west to northwest during all seasons (44%), and katabatic winds from the north-northwest to east-southeast sector (38%) account for most of the remainder of the flow. South westerly (13%) and south easterly winds (6%) account for only a minor fraction of the locally observed wind direction each. Air masses associated with this cluster descended from the free troposphere less than 30h before they arrived at the receptor site and had significant residence time over arid regions west of India. Consequently, they are associated with low relative and absolute humidity and do not bring rain. The medium westerly cluster is typically observed shortly before the arrival of a western disturbance.

The *fast westerly* cluster is observed only during winter, summer, and post-monsoon seasons. It accounts for 7.7, 6.4, and 6.4% of the air masses, respectively. The cluster is associated with a strong subtropical jet stream poised over westerly troughs and higher than average wind speeds. The predominant local wind direction for this cluster is west to northwest during all seasons (60%), and katabatic winds from the north-northwest to east-southeast sector (30%) account for most of the remainder. South westerly (6%) and south easterly winds (4%) account only for a minor fraction of the locally observed wind direction each. Since air masses associated with this cluster descended from the free troposphere less than 30h before they arrived at the receptor site and had significant residence time over arid regions west of India, they are associated with low relative and absolute humidity and do not bring rain. The fast westerly cluster is most frequently observed during winter and early summer season 2–3days before the arrival of a western disturbance.

The *south easterly* cluster is associated with the passage of a western disturbance in winter and summer and with active spells of the monsoon during monsoon season. It accounts for 19.3, 13.1, and 42.6% of the flow respectively. It is generally not observed during the post-monsoon season. The western disturbance is responsible for most of the wintertime and summertime rain events (Table 3.1). During the winter and summer seasons, the predominant local wind direction for this cluster is south-easterly (38%). Katabatic winds from the north-northwest to east-southeast sector account for 27%, south westerly winds for 17%, and north westerly winds for 18% of the locally

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

observed wind direction. Temperatures and wind speeds associated with this cluster are above average in winter and below average in summer. The relative and absolute humidity of air masses associated with this cluster are always high during both seasons. During monsoon season, the “Bay of Bengal branch” of the monsoon circulation brings warm and moist air masses to the receptor site. The absolute humidity is high, and the highest total amount of rainfall is observed for this cluster. The predominant local wind direction for this cluster is south east (53%). Katabatic winds from the north-northwest to east-southeast sector account for 24%, south westerly winds for 12%, and north westerly winds for 11% of the locally observed wind direction.

The *south westerly* cluster is associated with the passage of a western disturbance in winter and summer and with “break” spells of the monsoon during monsoon season. It accounts for 3.4, 7.3, and 10.4% of the flow, respectively. It is not observed during the post-monsoon season. During winter and summer, the south-westerly cluster is usually observed in association with a weakening western disturbance or when the centre of the low-pressure system is above or close to the receptor site. The predominant local wind directions are west to the northwest (42%) and southeast (35%). South westerly winds (13%) and katabatic flow (10%) account only for a minor fraction of the locally observed wind direction. During monsoon, this cluster is associated with “break” spells. “Break” spells occur when the monsoon trough is located over the foothills of the Himalayas, and the low-level jet originating off the coast of Somalia enters the IGP through the Indus valley. The local wind direction is variable: 46% southeast, 21% northwest, 21% katabatic flow, and 11% south westerly winds. The absolute humidity of air masses associated with this cluster is high, although rainfall events occur only rarely. However, extreme rainfall events are associated more frequently with this cluster.

Calm conditions ( $WS < 1\text{ m s}^{-1}$ ) account for only 4.5, 2.5, 5.2, and 8.7% of the total time during winter, summer, monsoon, post-monsoon season, respectively. They occur more frequently at night (60%) and less frequently during the day (40%). The local wind direction during periods with low wind speed is variable: 36% south west, 33% katabatic flow, 19% south east, and 12% north west.

#### **3.4.3. Impact of air-mass transport on particulate matter (PM) mass loadings**

To quantify the contribution of long-range transport to particulate matter mass loadings at the receptor site, I calculated the cluster average mass loadings of coarse- and fine-mode particulate matter at the receptor site (Figure 3.9) and the enhancement of PM mass loadings above the levels observed for the “local” cluster which represents the regional background pollution in the NW-IGP best (Table 3.2).

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

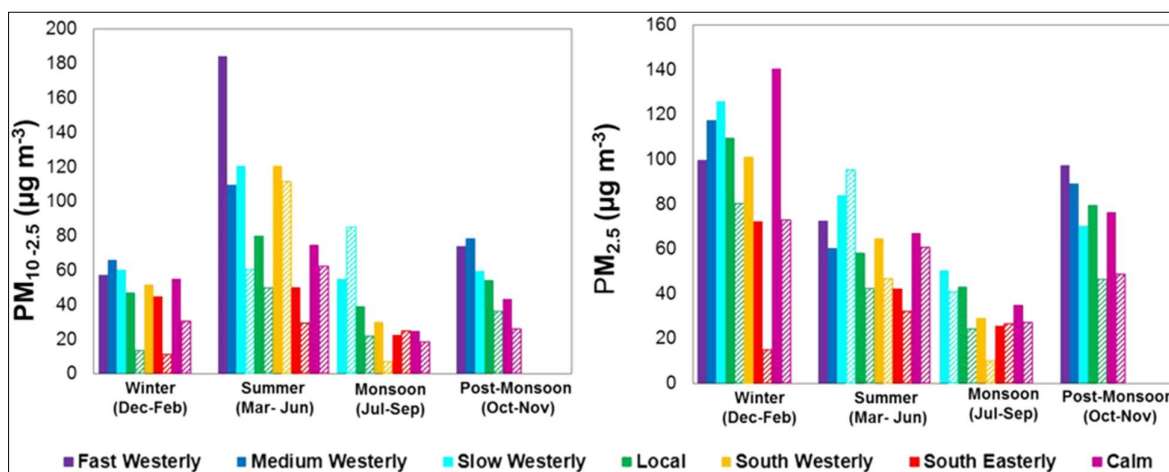


Figure 3.9: Mean coarse-mode ( $PM_{10-2.5}$ ) and fine-mode ( $PM_{2.5}$ ) mass loading for each air-mass cluster and season at the IISER Mohali air quality station. Hatched bars indicate coarse-mode and fine-mode PM mass loadings observed during rain events.

The enhancement is expressed in % of the total PM mass loading observed for the respective cluster. I determined whether the differences in PM mass loadings between the different clusters are significant using Levene’s test for homogeneity of variance based on means and used the pairwise comparison based on Tukey’s studentised HSD test (honestly significant differences) for assessing the statistical significance of the difference of the mean for each pair of clusters and each season (Table 3.2).

	Fast westerly	Medium westerly	Slow westerly	South-westerly	South-easterly	Calm
$PM_{2.5}$						
Winter	Negative	$7 \pm 4$	$13 \pm 9$	Negative	Negative	$22 \pm 15$
Summer	$20 \pm 15$	$4 \pm 3$	$31 \pm 21$	$10 \pm 8$	Negative	$13 \pm 8$
Monsoon	Negative	Negative	$15 \pm 11$	Negative	Negative	Negative
Post-Monsoon	$18 \pm 10$	$11 \pm 7$	Negative	Negative	Negative	Negative
$PM_{10-2.5}$						
Winter	$18 \pm 8$	$28 \pm 16$	$22 \pm 15$	$9 \pm 5$	Negative	$14 \pm 10$
Summer	$57 \pm 49$	$27 \pm 21$	$34 \pm 28$	$34 \pm 26$	Negative	Negative
Monsoon	Negative	Negative	$29 \pm 11$	Negative	Negative	Negative
Post-Monsoon	$27 \pm 18$	$31 \pm 21$	$9 \pm 6$	Negative	Negative	Negative

Table 3.2: Lower limit for the contribution of long-range transport and local pollution events to PM mass loadings in %  $\pm 1\sigma$  of the total PM. “Negative” indicates that the PM mass loadings are not enhanced compared to the local cluster, which represents the regional background levels.

#### 3.4.3.1. Winter season

During the winter season, both long-range transport from the west and south-west and local pollution episodes lead to enhanced coarse-mode PM mass loadings (Table 3.2). The contribution of long-range transport to coarse-mode PM loadings varies from 9% for the south-westerly cluster to 28% for the medium westerly cluster. Local pollution episodes contribute 14% on average to the coarse-mode PM observed under calm conditions. Even though the average coarse-mode PM

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

varies from  $45 \mu\text{g m}^{-3}$  for the south-easterly cluster to  $66 \mu\text{g m}^{-3}$  for the medium westerly cluster (Figure 3.9), the difference of the average is not statistically significant for any of the cluster pairs due to the high intra-cluster variance of coarse-mode PM during the winter season (Table 3.2). Figure 3.10 shows the correlation of CO with coarse-mode PM ( $\text{PM}_{10-2.5}$ ) as a function of meteorological conditions. This figure suggests that aqueous-phase processing of gas-phase precursors and dust contribute prominently to coarse-mode PM mass loading during winter. Aqueous-phase oxidation of gas-phase precursors refers to a process wherein gases (e.g.,  $\text{NH}_3(\text{g})$ ,  $\text{SO}_2(\text{g})$ ,  $\text{NO}_2(\text{g})$ ) are taken up by aqueous-phase aerosol and subsequently undergo reactions that change the oxidation state, resulting in the formation of ammonium, sulfate, and nitrate ions. Upon drying, these precipitate out as salts, which can be of coarse mode in size. Aqueous-phase processing of gas-phase precursors emitted during combustion leads to a high degree of correlation between coarse-mode particulate matter and CO at high relative humidity ( $> 70\%$ ,  $r = 0.55$ ), while dust, both dust from long-range transport and locally suspended dust contribute significantly to coarse-mode PM at  $\text{RH} < 50\%$  and high wind speeds (Figure 3.10). Dey and Tripathi (2007) reported that in wintertime in Kanpur, more than 75% of coarse-mode particulate matter consisted of water-soluble salts, and only less than 25% of coarse-mode PM consisted of mineral dust. Their findings are in line with this study's observations that aqueous-phase processing of gas-phase precursors is responsible for a significant fraction of coarse-mode PM during the winter season (Figure 3.10). However, the complex interplay of meteorology dependent emissions and aqueous-phase processing leads to a high intra-cluster variance of PM mass loadings and obscures the contribution of long-range transport to PM levels.

The influence of wet scavenging on PM mass loadings, however, is statistically significant. During rain events, coarse-mode PM mass loadings drop to  $30 \mu\text{g m}^{-3}$  under calm conditions ( $-47\%$ ),  $13 \mu\text{g m}^{-3}$  for the local cluster ( $-72\%$ ) and  $11 \mu\text{g m}^{-3}$  for the south-easterly cluster ( $-78\%$ ), and the magnitude of the drop depends only weakly on the total amount of rainfall. Fine PM mass loadings drop to  $73 \mu\text{g m}^{-3}$  under calm conditions ( $-48\%$ ),  $80 \mu\text{g m}^{-3}$  for the local cluster ( $-27\%$ ), and  $15 \mu\text{g m}^{-3}$  for the south-easterly cluster ( $-80\%$ ). This clearly demonstrates the profound influence of wet scavenging on fine-mode PM mass loadings during winter. The percent decrease in fine-mode PM mass loadings during rain events scales perfectly linearly with the total rainfall for each cluster (1.3% decrease in  $\text{PM}_{2.5}$  per mm of rainfall,  $r^2 = 0.99$ ). The fact that the drop in coarse-mode PM is independent of the total amount of rain while the drop in fine-mode PM strongly depends on the total amount of rainfall could be an indicator that, during wintertime, soluble coarse-mode PM

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

(large salts) plays a crucial role in initiating rainfall as giant cloud condensation nuclei. In contrast, fine-mode PM is mostly scavenged by below cloud scavenging.

	Calm	South-easterly	South-westerly	Local	Slow-westerly	Medium westerly	Fast westerly
WINTER (Dec-Feb)							
Calm	–						
South-easterly	3 $\sigma$ (3 $\sigma$ )	–					
South-westerly	1 $\sigma$ (1 $\sigma$ )		–				
Local	– (1 $\sigma$ )	1 $\sigma$		–			
Slow westerly		2 $\sigma$ (1 $\sigma$ )			–		
Medium westerly		1 $\sigma$				–	
Fast westerly	1 $\sigma$ (1 $\sigma$ )						–
SUMMER (Mar-Jun)							
Calm	–						4 $\sigma$ (3 $\sigma$ )
South-easterly	2 $\sigma$ (2 $\sigma$ )	–	2 $\sigma$ (1 $\sigma$ )		2 $\sigma$ (1 $\sigma$ )	1 $\sigma$ (1 $\sigma$ )	4 $\sigma$ (4 $\sigma$ )
South-westerly		1 $\sigma$ (1 $\sigma$ )	–				1 $\sigma$ (1 $\sigma$ )
Local				–			4 $\sigma$ (3 $\sigma$ )
Slow westerly		4 $\sigma$ (4 $\sigma$ )	1 $\sigma$ (1 $\sigma$ )	2 $\sigma$ (2 $\sigma$ )	–		1 $\sigma$ (1 $\sigma$ )
Medium westerly		1 $\sigma$			1 $\sigma$ (1 $\sigma$ )	–	2 $\sigma$ (2 $\sigma$ )
Fast westerly		2 $\sigma$ (2 $\sigma$ )					–
MONSOON (Jul-Sep)							
Calm	–			1 $\sigma$	3 $\sigma$ (2 $\sigma$ )		
South-easterly	– (1 $\sigma$ )	–		1 $\sigma$ (1 $\sigma$ )	4 $\sigma$ (3 $\sigma$ )		
South-westerly			–		2 $\sigma$ (1 $\sigma$ )		
Local		2 $\sigma$ (2 $\sigma$ )	2 $\sigma$ (1 $\sigma$ )	–	1 $\sigma$		
Slow westerly	2 $\sigma$ (1 $\sigma$ )	4 $\sigma$ (4 $\sigma$ )	3 $\sigma$ (2 $\sigma$ )		–		
Medium westerly						–	
Fast westerly							–
POST-MONSOON (Oct-Nov)							
Calm	–					3 $\sigma$ (3 $\sigma$ )	2 $\sigma$ (2 $\sigma$ )
South-easterly		–					
South-westerly			–				
Local				–		2 $\sigma$ (1 $\sigma$ )	1 $\sigma$ (1 $\sigma$ )
Slow westerly					–	1 $\sigma$ (1 $\sigma$ )	
Medium westerly						–	
Fast westerly	1 $\sigma$ (1 $\sigma$ )			1 $\sigma$ (1 $\sigma$ )	2 $\sigma$ (2 $\sigma$ )		–

Table 3.3: Statistical significance of the difference of the mean for each pair of clusters. Values to the right of the principal diagonal denote significance among PM<sub>10-2.5</sub> pairs. In contrast, values to the left of the principal diagonal denote significance among PM<sub>2.5</sub> pairs. Pairwise comparison based on Tukey’s studentised HSD (honestly significant differences) test was used to assess the statistical significance of the difference of the mean for each pair of clusters and each season. Values in brackets indicate the statistical significance after all rain events were removed from the data set.

For fine-mode PM, local pollution episodes lead to the highest enhancements in fine-mode PM mass loadings (22%). Long-range transport from the west contributes only moderately to fine-mode PM (7 and 13% for the medium westerly and slow westerly clusters, respectively). The highest fine-mode PM mass loadings are observed under calm conditions during local pollution episodes (141  $\mu\text{g m}^{-3}$ ). The enhancement is significant when compared to the south-easterly (72  $\mu\text{g m}^{-3}$ ), south-westerly (101  $\mu\text{g m}^{-3}$ ), and fast westerly (100  $\mu\text{g m}^{-3}$ ) clusters. When only dry days are considered, the difference between local pollution episodes (146  $\mu\text{g m}^{-3}$  on dry days) and the local cluster (110  $\mu\text{g m}^{-3}$  on dry days), which represents the regional air pollution, also becomes significant (Table 3.3).



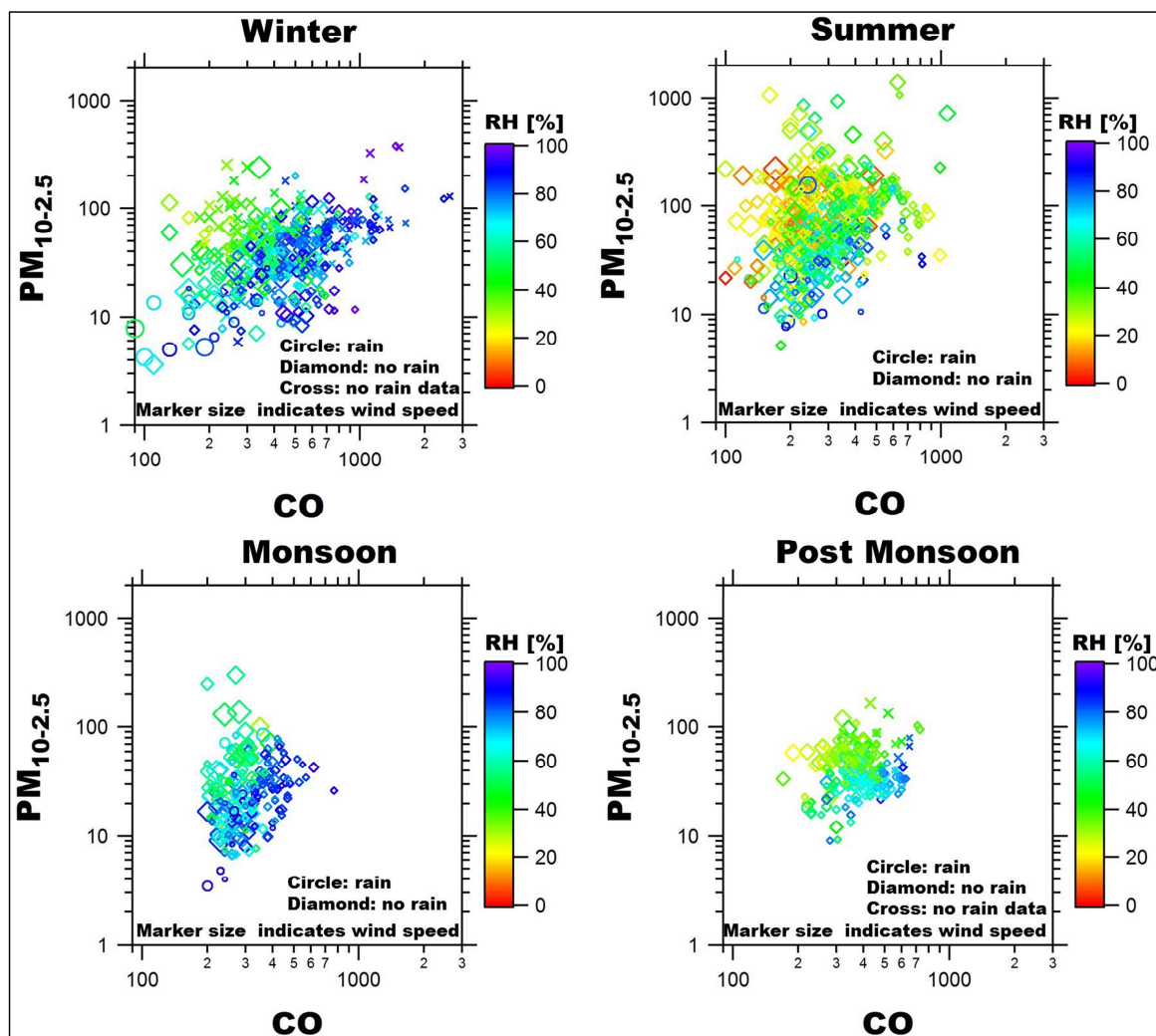


Figure 3.10: Dependence of coarse-mode PM mass loadings on the emission of gas-phase precursors and meteorological parameters for the different seasons. The marker shape distinguishes PM mass loadings measured during rain events (circles) and under dry conditions (diamonds); data points obtained while the rain gauge was not working are marked with crosses. Marker size is proportional to wind speed. The smallest markers indicate  $WS \leq 1 \text{ m s}^{-1}$ , the largest markers  $WS \geq 15 \text{ m s}^{-1}$ . Markers are colour coded with relative humidity.

The lowest fine-mode PM mass loadings are observed for the south-easterly cluster ( $72 \mu\text{g m}^{-3}$ ) associated with the western disturbance and has significantly lower mass loadings than the local, slow westerly and medium westerly clusters. The fast westerly cluster, which is usually observed shortly before a western disturbance establishes itself over India, has the second-lowest fine-mode mass loadings ( $100 \mu\text{g m}^{-3}$ ). However, the difference is not statistically significant with respect to the other clusters due to large intra-cluster variability in the fine-mode PM mass loadings.

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

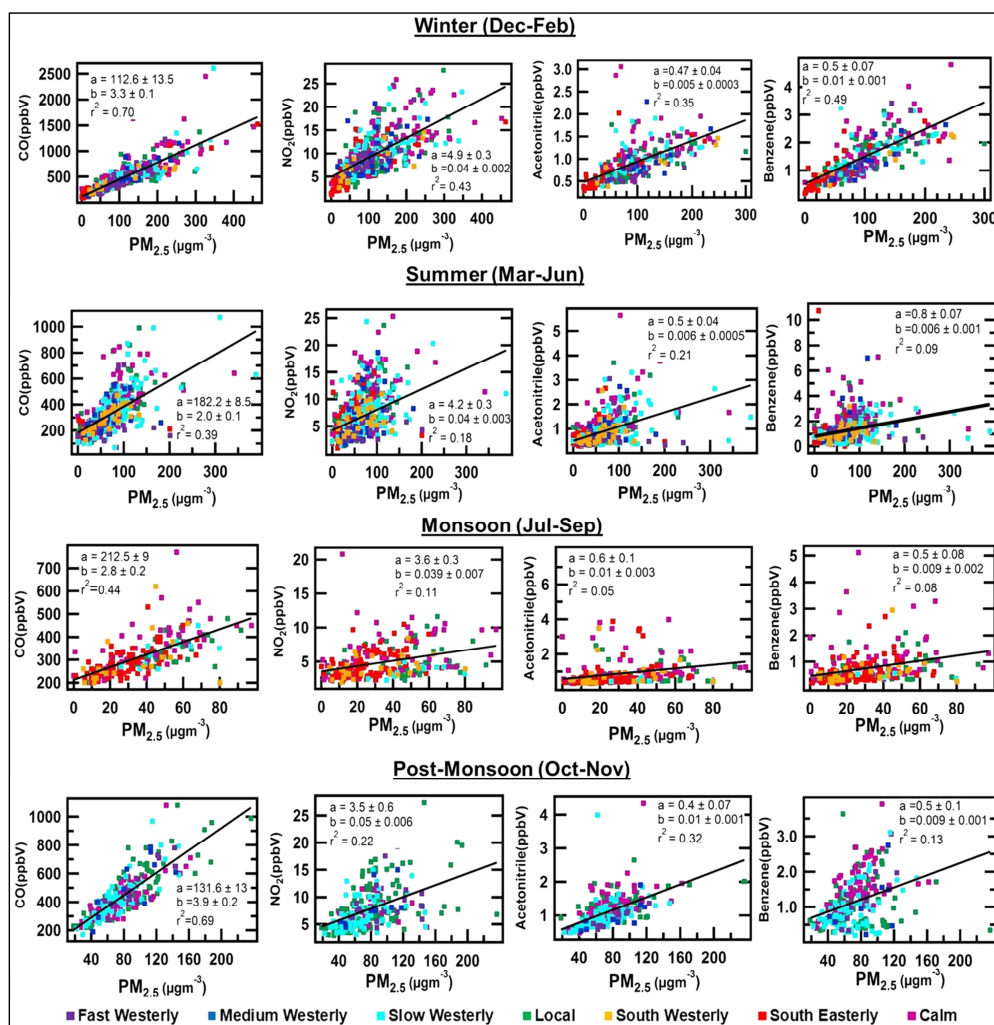


Figure 3.11: Scatter plots of fine-mode PM with CO, acetonitrile, benzene and  $NO_2$  for the winter, summer, monsoon, and post-monsoon seasons. ‘a’ stands for intercept and ‘b’ stands for slope in the linear regression equation.

During wintertime, the emission of fine-mode particulate matter is driven by combustion. Correlation plots of fine PM with CO ( $r^2 = 0.70$ ), acetonitrile, a biomass combustion tracer ( $r^2 = 0.35$ ), benzene ( $r^2 = 0.49$ ) and  $NO_2$ , a tracer for high temperature combustion ( $r^2 = 0.43$ ), Figure 3.11, clearly indicate that, at the receptor site, combustion is the predominant source of wintertime fine-mode PM across all clusters. Due to low ambient temperatures in the winter months, particularly in the surrounding mountain regions, those who cannot afford electric heaters burn dry leaves, wood, coal, agricultural residues, and cow dung often mixed with garbage to keep themselves warm. This practice prevails in the whole of South Asia. It explains the simultaneous increase in fine-mode PM and acetonitrile during the winter season. High emissions of benzene have previously been observed during biomass combustion episodes in the region (Sarkar et al., 2013), and inefficient combustion in open fires or simple stoves is known to cause high PM mass

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

loadings (Habib et al., 2004; Venkataraman et al., 2005; Massey et al., 2009; Akagi et al., 2011). While there is a clear correlation between both benzene and acetonitrile and PM<sub>2.5</sub>, the lower  $r^2$  for acetonitrile ( $r^2 = 0.35$ ) compared to the higher  $r^2$  for benzene ( $r^2 = 0.49$ ) and CO ( $r^2 = 0.70$ ) indicates that mixtures of fuels with variable biomass content are used for domestic heating purposes. The scatter plot between PM<sub>2.5</sub> and benzene (Figure 3.11) also suggests that there may be regional preferences with respect to the fuel mixture as the emission ratios of the south-easterly cluster usually fall below the fit line, while those for the south-westerly cluster fall above it. The largest scatter and hence the variation in fuel type is observed under calm conditions and for the local and slow westerly cluster. The high mass loading of fine-mode particulate matter coupled with the high relative humidity, which frequently reaches values above 75%, in particular during the night, leads to the formation of persistent fog and haze during wintertime. The uptake of water-soluble organic and inorganic gas-phase species into the aqueous-phase and the subsequent chemical reactions result in a fine-mode aerosol that contains a large mass fraction of water-soluble inorganic species (Kumar et al., 2007) and acts as a very efficient cloud condensation nuclei (CCN). Repeated fog processing also leads to the formation of coarse-mode inorganic salt particles (Kulshrestha et al., 1998; Kumar et al., 2007). Kaskaoutis et al. (2013) reported a bi-modal volume size distribution for wintertime aerosol in Kanpur with a first, higher peak between 200 and 300 nm and a second peak between 3 and 4  $\mu\text{m}$  optical equivalent diameter. The ratio of coarse-mode to fine-mode PM observed at Mohali site agrees well with the ratio of coarse-mode to fine-mode PM observed in their study. Kulshrestha et al. (1998) reported a bimodal size distribution peaking at 1 and 5  $\mu\text{m}$  aerodynamic equivalent diameter for wintertime aerosol in Agra and found that ammonium sulfates, ammonium nitrate and potassium sulfate dominated water-soluble salts in the fine mode, while sulfates, nitrates and chlorides of sodium, calcium and magnesium dominated coarse-mode aerosol.

#### 3.4.3.2. Summer season

During the summer season, long-range transport from the west and south-west contributes significantly to enhanced coarse-mode PM mass loadings (Table 3.2). Long-range transport contributes approximately 30% to coarse-mode PM in air masses associated with the south-westerly, slow westerly, and medium westerly clusters each and 57% to coarse-mode PM in air masses associated with the fast westerly cluster. Air masses related to the south-easterly cluster ( $50 \mu\text{g m}^{-3}$ ; Figure 3.9) show significantly lower coarse-mode PM mass loadings than south-westerly, slow westerly, medium, and fast westerly clusters and also compared to the local air masses observed under calm conditions. Only the difference with respect to the local cluster ( $80 \mu\text{g m}^{-3}$ ), which represents regional air masses, is not significant, mainly due to the high variance

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

of coarse-mode PM mass loadings of air masses attributed to the local cluster. The variance is caused by convective dust storms (Joseph, 1982). It is very interesting to note that air masses that have crossed the entire, densely populated IGP show the lowest PM mass loadings even when compared with the local cluster, which represents regional air masses or when compared to air masses representing a local fetch region observed under calm conditions ( $75 \mu\text{g m}^{-3}$ ). This is true during both rain events and on dry days.

The highest cluster average is observed for the fast westerly cluster ( $184 \mu\text{g m}^{-3}$ ). The coarse-mode PM ( $\text{PM}_{10-2.5}$ ) mass loadings for this cluster are significantly enhanced above the coarse-mode PM mass loadings observed in all other clusters and under calm conditions (Table 3.3), and 57% of the average PM mass is due to long-range transport for this cluster. The coarse PM enhancement for the fast westerly cluster is associated with dust storms originating in the Middle East that reach Mohali site from the west (Pandithurai et al., 2008). The slow and medium westerly clusters and the south westerly cluster show enhanced coarse-mode PM ( $\text{PM}_{10-2.5}$ ) mass loadings as well, though the difference is statistically significant only with respect to the south-easterly cluster (Table 3.2). PM enhancements for the south-westerly cluster are associated with dust storms originating from the Thar Desert (Sharma et al., 2012) or the Arabian Peninsula that reach Mohali site through the Indus Valley. During the summer season, maximum rainfall is observed for the south-easterly, local cluster, south-westerly, and slow westerly cluster in descending order of the absolute rainfall amount. Even when rain events, characterized by average coarse-mode PM mass loadings of  $50 \mu\text{g m}^{-3}$  ( $-38\%$  in average  $\text{PM}_{10-2.5}$  mass loading),  $29 \mu\text{g m}^{-3}$  ( $-41\%$  in average  $\text{PM}_{10-2.5}$  mass loading) and  $60 \mu\text{g m}^{-3}$  ( $-50\%$  in average  $\text{PM}_{10-2.5}$  mass loading) for the local, south-easterly and slow westerly cluster respectively and  $62 \mu\text{g m}^{-3}$  ( $-15\%$  in average  $\text{PM}_{10-2.5}$  mass loading) for periods of calm are removed, the differences outlined above remain significant. The south-westerly brings moisture from the Arabian sea and dust from the Arabian Peninsula (Pease et al., 1998); consequently, the average coarse-mode PM during rain is comparable to the average coarse-mode PM on dry days. It is interesting to note that the slow westerly cluster shows an increment in fine PM values on rainy days ( $95 \mu\text{g m}^{-3}$ ) compared to dry days ( $84 \mu\text{g m}^{-3}$ ), indicating that rainfall for this cluster is associated with convective dust storms. For fine-mode PM, the slow westerly ( $84 \mu\text{g m}^{-3}$ ) cluster shows significantly (Table 3.3) enhanced fine PM mass loadings, and approximately 31% of the fine PM for this cluster is contributed by transport from the west (Table 3.2). For the slow westerly cluster, the differences are significant with respect to the south-easterly ( $42 \mu\text{g m}^{-3}$ ), south westerly ( $65 \mu\text{g m}^{-3}$ ), local ( $58 \mu\text{g m}^{-3}$ ) and medium westerly ( $60 \mu\text{g m}^{-3}$ ; Figure 3.9) clusters. Local pollution episodes lead to a 13% increase above the regional

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

PM<sub>2.5</sub> background. For the fast westerly cluster ( $73 \mu\text{g m}^{-3}$ ), 20% of the fine-mode PM is contributed by long-range transport, but the difference is only significant with respect to the south-easterly cluster. Just like for coarse-mode PM, the lowest fine-mode PM mass loadings are observed for the south-easterly cluster. The difference is significant with respect to the south-westerly, slow westerly, and fast westerly cluster and with respect to the local pollution episodes observed under calm conditions (Table 3.3). Overall fine-mode PM mass loadings in summer are lower than during wintertime.

During the summer season, the comparison of the emission ratios of acetonitrile, benzene, CO, and NO<sub>2</sub> with fine-mode particulate matter patterns indicate that several sources drive fine-mode PM (Figure 3.11). While there is still a reasonable correlation between CO and PM<sub>2.5</sub> ( $r^2 = 0.39$ ), most other combustion tracers have a poor coefficient of correlation with PM<sub>2.5</sub>. The scatter plots indicate a spread between at least two types of combustion. One type is characterised by high acetonitrile, benzene, and NO<sub>2</sub> emissions but relatively low PM<sub>2.5</sub> mass loadings and is probably associated with wheat residue burning in Punjab, while the other type is characterized by lower benzene, NO<sub>2</sub>, and acetonitrile mixing ratios but higher PM<sub>2.5</sub> mass. This second source is probably traffic, which in arid regions during the summer season is responsible for significant (re-)suspension of dust, in particular when wind speeds are high ( $>5 \text{ m s}^{-1}$ ). Several authors reported that during the summer season, coarse-mode mineral dust with a single peak at 3 – 4  $\mu\text{m}$  optical equivalent diameter dominates PM mass loadings in the IGP (Gautam et al., 2011; Kaskaoutis et al., 2013); however, I find that fine-mode particulate matter (PM<sub>2.5</sub>) contributes almost equally to PM mass loadings. A significant fraction of PM<sub>2.5</sub> mass is still combustion derived. Only Jethva et al. (2005) reported a bimodal volume distribution for dust storms, with one peak at 3–4  $\mu\text{m}$  and a second peak at 1.5  $\mu\text{m}$ , which agrees well with this study's findings. The peak at 1.5  $\mu\text{m}$  corresponds to the clay fraction of mineral dust and is frequently found to be strongly enriched in mineral dust plumes after extended long-range transport (Pöschl et al., 2010). At this site, I find that the coarse-mode PM fraction in individual dust storm events varies between 45 and 92%, with the highest coarse-mode fraction typically recorded for dust storms originating in the Thar Desert. While during dust storms, windblown dust contributes significantly to fine-mode PM, PM<sub>2.5</sub> is usually dominated by combustion-derived aerosols at this site (Figure 3.11).

### 3.4.3.3. Monsoon season

During the monsoon season, the effect of wet scavenging of coarse-mode PM mass loadings can be clearly seen in the low average coarse PM mass loadings. Qualitatively, the average mass loading is anti-correlated with rainfall. The lowest coarse-mode PM mass loadings are observed for the south easterly ( $22 \mu\text{g m}^{-3}$ ) and south-westerly ( $30 \mu\text{g m}^{-3}$ ) clusters.

The slow westerly cluster ( $55 \mu\text{g m}^{-3}$ ) shows significant enhancement over all other clusters and the calm periods. Long-range transport from the west contributes approximately 30% to enhanced coarse-mode PM mass loadings in the slow westerly cluster (Table 3.2). However, when rain events are removed, the enhancement over the local cluster is no longer significant.

The local cluster ( $39 \mu\text{g m}^{-3}$ ) shows enhancements over the south-easterly cluster ( $22 \mu\text{g m}^{-3}$ ) and periods with calm conditions ( $25 \mu\text{g m}^{-3}$ ); however, the enhancement over calm conditions is no longer significant when rain events with average coarse-mode PM mass loadings of 22, 25 and  $18 \mu\text{g m}^{-3}$  respectively are removed from the three clusters (Figure 3.9, Table 3.3).

During monsoon seasons, most coarse-mode PM is derived from aqueous-phase processing of gas-phase precursors (Figure 3.10), a process that is extremely efficient at  $\text{RH} > 75\%$ , and the removal is controlled by wet scavenging. Dust storms contribute only occasionally to coarse-mode PM.

For fine-mode PM ( $\text{PM}_{2.5}$ ), the south-easterly cluster ( $26 \mu\text{g m}^{-3}$ ) shows the lowest mass loadings. The difference is significant with respect to the local ( $43 \mu\text{g m}^{-3}$ ) and slow westerly ( $51 \mu\text{g m}^{-3}$ ) clusters (Figure 3.9). The difference between south-easterly cluster and calm pollution episodes becomes significant when only dry days with an average fine PM loading of 26 and  $38 \mu\text{g m}^{-3}$ , respectively, are considered (Table 3.3). The south-westerly cluster ( $29 \mu\text{g m}^{-3}$ ), too, shows significantly lower fine-mode PM when compared to the local ( $43 \mu\text{g m}^{-3}$ ) and slow westerly ( $51 \mu\text{g m}^{-3}$ ) clusters (Figure 3.9). The slow westerly cluster shows significant enhancements of fine-mode PM over all other clusters except the local cluster and significant enhancement over calm periods. Transport contributes approximately 15% to the fine-mode PM for this cluster.

During the monsoon season, the correlation of acetonitrile, benzene, CO, and  $\text{NO}_2$  with fine-mode particulate matter indicates that multiple combustion sources drive fine mode PM (Figure 3.11). While there is still a correlation with CO ( $r^2 = 0.44$ ), the coefficient of correlation of acetonitrile, benzene, and  $\text{NO}_2$  with  $\text{PM}_{2.5}$  is low. The largest scatter due to biomass combustion derived  $\text{PM}_{2.5}$  (associated with high acetonitrile, benzene, and  $\text{NO}_2$ ) is observed under calm conditions.  $\text{PM}_{2.5}$

enhancements for the slow westerly cluster, on the other hand, are accompanied by low acetonitrile, benzene, and NO<sub>2</sub> mixing ratios and are possibly caused by traffic.

#### **3.4.3.4. Post-monsoon season**

During the post-monsoon season, air masses reaching the site from the west (slow, medium, and fast westerly clusters) show higher coarse PM mass loadings compared to the local cluster and air masses observed under calm conditions. Transport from the west contributes approximately 30% each to the coarse-mode PM mass loadings of the medium westerly and fast westerly clusters and approximately 10% to the coarse-mode PM mass loadings of the slow westerly cluster. The highest coarse-mode PM is observed for the fast and medium westerly clusters. The enhancement in coarse-mode PM observed for the medium westerly cluster is statistically significant with respect to all other clusters, including the slow westerly cluster.

The enhancement observed for the fast westerly cluster is statistically significant only with respect to the local cluster and calm conditions. Results remain significant even when rain events are removed from both. Calm episodes have significantly lower coarse-mode PM mass loadings ( $43 \mu\text{g m}^{-3}$ ) than the medium and fast westerly clusters, indicating that local pollution episodes are not a significant source of coarse-mode PM during the post-monsoon season, while the fetch regions of the westerly clusters are. The highest fine-mode PM during the post-monsoon season is observed in the fast westerly cluster ( $97 \mu\text{g m}^{-3}$ ) and transport contributes 18% to the PM mass loading associated with this cluster (Table 3.2). Fine PM mass loadings for this cluster are significantly enhanced compared to calm conditions ( $76 \mu\text{g m}^{-3}$ ) and the local ( $80 \mu\text{g m}^{-3}$ ) and slow westerly ( $70 \mu\text{g m}^{-3}$ ) clusters. All differences discussed above remain significant when rain events are removed from the data set, and only dry days are considered. The second highest fine-mode PM mass loadings are observed for the medium westerly cluster. Transport contributes 11% to the PM mass loadings observed for this cluster, and the source characteristics are similar to those observed for the fast westerly cluster (Figure 3.11).

For the slow westerly and local clusters, smoke produced by crop residue burning is a significant source of PM during this season as crop residue burning is practiced in most of Punjab. Consequently, all air masses reaching the receptor site from the west are impacted by this source. Air masses attributed to the slow westerly and local clusters and air masses observed under calm conditions show a clear enhancement in the PM to acetonitrile ratio above the line fit representing the regional background on days when fresh crop residue burning plumes impact the site. Paddy residue burning leads to massive enhancements in acetonitrile and benzenoids (Sarkar et al., 2013) and equally large emissions of PM. Singh et al. (2010) reported monthly average SPM of 400–500

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

$\mu\text{g m}^{-3}$  for a village site near Patiala during October and November and monthly average SPM of  $300 - 400 \mu\text{g m}^{-3}$  at a suburban site (residential campus of Punjabi University) during the same 2 months. At this suburban receptor site, further downwind of the burning fields, I find that  $\text{PM}_{10}$  generally ranges between 100 and  $200 \mu\text{g m}^{-3}$  and exceeds  $200 \mu\text{g m}^{-3}$  only during a few episodes. Badarinath et al. (2009) showed that the crop residue burning smoke is mostly channeled into the IGP, and Mishra and Shibata (2012) showed that the crop residue burning plumes impact sites as far downwind as Kanpur. During its journey, the smoke ages and aerosol size distributions are modified. Kaskaoutis et al. (2013) reported a bimodal volume size distribution for post-monsoon aerosol in Kanpur with a first peak between 200 and 300 nm and a second peak at 3–4 $\mu\text{m}$ . In Kanpur, coarse-mode aerosol exceeded fine-mode PM (by a factor of 1.3), while at Mohali receptor site closer to the burning fields, fine-mode PM exceeds coarse-mode PM by a factor of 1.5. This indicates that during the 720km journey from Punjab to Kanpur approximately 30% of the fine particulate matter mass is transformed into coarse-mode PM through repeated fog processing.

#### **3.4.4. Impact of air-mass transport on particulate matter (PM) exceedance events**

The mean PM mass loadings of an air-mass cluster represent a poor proxy for the number of exceedance events, that is, the number of days on which the 24h average of  $\text{PM}_{2.5}$  or  $\text{PM}_{10}$  exceeded the NAAQS of  $60 \mu\text{g m}^{-3}$  for 24h average  $\text{PM}_{2.5}$  or  $100 \mu\text{g m}^{-3}$  for the 24h average  $\text{PM}_{10}$ , respectively. During the winter season on ~160 days (out of 180 days), the NAAQS of both  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  was exceeded. For the summer season, exceedance days associated with  $\text{PM}_{10}$  were more frequent (203 out of 243 days) than those associated with  $\text{PM}_{2.5}$  (157 out of 243). During the monsoon season, the receptor site received cleaner air masses, with only 13 and 21 exceedance days (out of 92) of  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$ , respectively.

The post-monsoon season had frequent exceedance days, with NAAQS of  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  being exceeded on ~110 out of 114 days. While individual pollution episodes with extremely high PM mass loadings such as dust storms can profoundly influence the cluster mean, they barely affect the number of exceedance days, as such events are rare.

##### **3.4.4.1. Winter season**

During the winter season, emissions of gas-phase precursors and particulate matter from local and regional sources are so high, and the conversion of gas-phase precursors to both  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  is so efficient, that the NAAQS for both  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  is exceeded 95% (Figure 3.12) of the days associated with the local cluster, i.e., air masses that had been confined over the NW-IGP 3 days prior to their arrival at the receptor site. Despite the fact that transport from the West



### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

enhances PM<sub>10</sub> mass loadings for the slow, medium, and fast westerly clusters, it increased the fraction of days during which PM mass loadings exceed the NAAQS by a maximum of 5%. The largest increase in the fraction of exceedance days is observed for the south-westerly and medium westerly clusters (from 95 to 100% of the days associated with this synoptic-scale transport for both PM<sub>2.5</sub> and PM<sub>10</sub>) and the slow westerly cluster (from 95 to 100% for PM<sub>10</sub>). Significantly cleaner air masses with a lower fraction of exceedance events are usually associated with wet scavenging and/or air masses brought by a western disturbance (south-easterly cluster: 56% of the days related to this synoptic-scale transport for PM<sub>2.5</sub> and 58% for PM<sub>10</sub>). Calm conditions, on the other hand, do not increase the percentage of exceedance days for either PM<sub>10</sub> or PM<sub>2.5</sub> when compared to the local cluster.

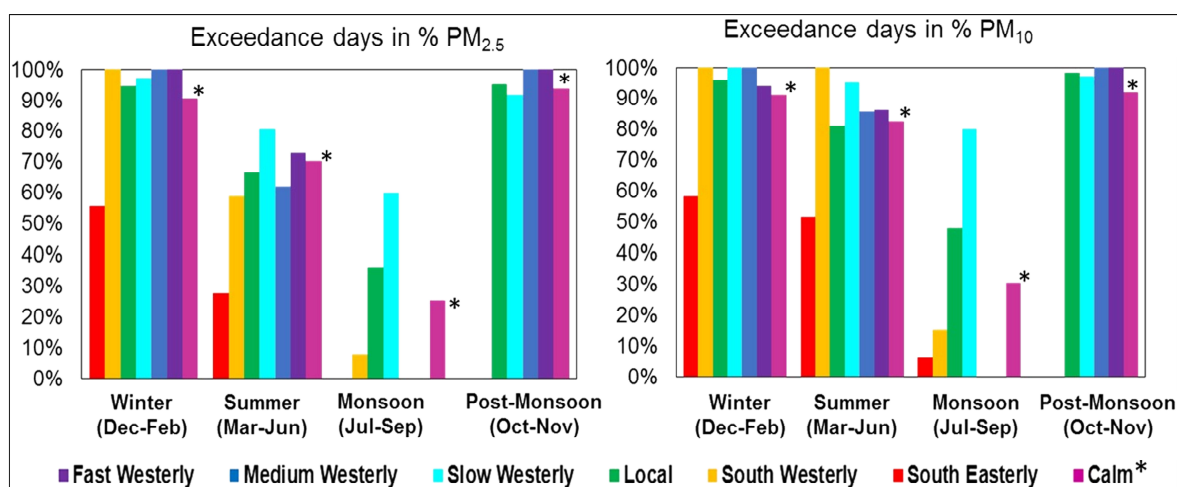


Figure 3.12: Percentage of days where the 24h average PM<sub>10</sub> and PM<sub>2.5</sub> mass loading (in  $\mu\text{g m}^{-3}$ ) exceed the national ambient air quality standard for each air-mass cluster and season. \*For calm conditions, PM<sub>2.5</sub> or PM<sub>10</sub> were averaged for all times during a 24h interval that had WS < 1 m s<sup>-1</sup> only. The fraction of exceedance days is calculated based on this average and is not based on a genuine 24h average, as wind speeds do not remain low continuously.

#### 3.4.4.2. Summer season

As discussed in Sect. 3.3.2, during the summer season, the ambient RH is usually < 60% at night and 20–30% during the day (Table 3.1), and aqueous-phase processing contributes less to PM<sub>2.5</sub> and PM<sub>10</sub> mass loadings. Instead, PM mass is dominated by direct emissions, dust, and photochemistry. Despite frequent dust storms, exceedance events are less frequent during the summer season than during the winter season. The NAAQS for PM<sub>10</sub> is exceeded on 80% of the days associated with the local cluster. The NAAQS for PM<sub>2.5</sub> is surpassed on 70% of the days related to this synoptic-scale transport. While dust storms – episodic events during which PM<sub>10</sub> mass loadings can reach up to 3000  $\mu\text{g m}^{-3}$  have a substantial impact on the cluster mean, particularly for the fast westerly cluster, they barely affect the number of exceedance events. This

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

is particularly true for the fast and medium westerly clusters. Only for the south-westerly cluster do dust storms increase the number of exceedance events compared to the local cluster, from 80 to 100% of the days associated with this synoptic-scale transport for  $PM_{10}$ . The highest increase in the number of exceedance events for  $PM_{2.5}$  is observed for the slow westerly cluster, which is most strongly affected by wheat residue burning in Punjab. Wheat residue burning increases the number of exceedance events observed for the slow westerly cluster compared to the local cluster from 70 to 80% of the days associated with this synoptic-scale transport for  $PM_{2.5}$  and from 80 to 95% for  $PM_{10}$ . The fraction of exceedance events for the south-easterly cluster, both for  $PM_{10}$  and  $PM_{2.5}$  (50 and 30% of the days with this synoptic-scale transport, respectively), is associated with cleaner air masses reaching the receptor site from the eastern IGP. Calm conditions barely increase the number of exceedance days.

#### 3.4.4.3. Monsoon season

Figure 3.12 shows that the frequency of  $PM_{2.5}$  exceedance days for each cluster is anti-correlated with the total rainfall observed for the respective cluster. For the local cluster, the NAAQS for  $PM_{10}$  is exceeded on 48% of the days, and the NAAQS for  $PM_{2.5}$  is exceeded on 36% of the days associated with this synoptic-scale transport. The slow westerly cluster, which is associated with occasional dust storms, increases the percentage of exceedance days from 36 to 60% for  $PM_{2.5}$  and from 48 to 80% for  $PM_{10}$ . The fraction of exceedance events for the south-easterly cluster, both for  $PM_{10}$  (6% of the days associated with this synoptic-scale transport) and  $PM_{2.5}$  (0% of the days associated with this synoptic-scale transport), is associated with extremely clean air masses reaching the receptor site from the eastern IGP.

#### 3.4.4.4. Post-monsoon season

As discussed in Sect. 3.3.4, crop residue burning coupled with aqueous-phase processing of gas-phase precursors is responsible for high PM mass loading during the post-monsoon season. It can be seen from Figure 3.12 that it also contributes to a high frequency of exceedance events. The NAAQS for both  $PM_{2.5}$  and  $PM_{10}$  is exceeded on 90–100% of the days associated with the synoptic-scale transport for the local, slow westerly, medium westerly, and fast westerly clusters and under calm conditions. Transport barely leads to an increase in the fraction of exceedance days.

## 3.5. Conclusions

In this chapter, I investigated the contribution of long-range transport and local pollution episodes to the average coarse- and fine-mode PM mass loadings at Mohali receptor site using 2 years of high temporal resolution data. The study yielded several results as follows.

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

---

1. Long-range transport from the west (source regions: Arabia, Thar Desert, Middle East, and Afghanistan) leads to significant enhancements in the average coarse-mode PM mass loadings during all seasons. The contribution of long-range transport from this source region to coarse-mode PM varied between 9 and 57% of the total PM<sub>10-2.5</sub> mass.
2. For fine-mode PM, the situation is more complicated. The fast westerly cluster is associated with a 20% increase in fine-mode PM during the summer and post-monsoon seasons but cleaner air masses during the winter season. The medium westerly cluster shows moderately enhanced PM mass loadings during all seasons, while slow westerly transport leads to enhanced PM<sub>2.5</sub> mass loadings during the winter, summer, and monsoon seasons, but not during the post-monsoon season.
3. Local pollution episodes (wind speed < 1 m s<sup>-1</sup>) contributed to enhanced PM<sub>2.5</sub> mass loadings during both the winter and summer seasons and to enhanced coarse-mode PM only during the winter season.
4. The south-easterly cluster (source region: eastern IGP) is associated with significantly lower fine- and coarse-mode PM mass loadings during all seasons.
5. The number of days during which PM mass loadings exceed the national ambient air quality standard (NAAQS) of 100 µg m<sup>-3</sup> for 24h average PM<sub>10</sub> and 60 µg m<sup>-3</sup> for 24h average PM<sub>2.5</sub> (NAAQS, 2009), however, is controlled by long-range transport to a much lesser degree. For the local cluster, which represents regional air masses (source region: NW-IGP), the fraction of days during which the national ambient air quality standard (NAAQS) of 60 µg m<sup>-3</sup> for 24h average PM<sub>2.5</sub> was exceeded varied between 36% of the days associated with this synoptic-scale transport during the monsoon season and 95% during the winter and post-monsoon seasons respectively; the fraction of days during which the national ambient air quality standard (NAAQS) of 100 µg m<sup>-3</sup> for the 24h average PM<sub>10</sub> was exceeded varied between 48% during the monsoon season and 98% during the post-monsoon season.
6. Long-range transport was responsible for bringing both air masses with a significantly lower fraction of exceedance days from the eastern IGP and air masses with a moderate increase in the fraction of exceedance days from the west (source regions: Arabia, Thar Desert, Middle East, and Afghanistan). The south-easterly cluster (source region: eastern IGP) is always associated with a significantly lower fraction of exceedance days, and the south-westerly cluster also leads to a lower fraction of exceedance days during the monsoon season. Whenever long-range transport increases the fraction of exceedance days, the increase varies between a few percent and at most 20%.

---

### 3. Quantifying the contribution of long-range transport to PM loadings at Mohali

7. Fine-mode PM ( $PM_{2.5}$ ) contributes most to PM exceedance events at a regional level, and  $PM_{2.5}$  mass loadings are primarily controlled by combustion sources during all seasons. Primary emission and gas to particle conversion of gas-phase precursors emitted during the combustion both contribute to the final mass loadings in varying proportions.

In order to bring PM mass loadings into compliance with the national ambient air quality standard (NAAQS) and to reduce the number of exceedance days, mitigation of regional combustion sources needs to be given the highest priority as the number of exceedance days for air masses associated with the source region NW-IGP is already extremely high. To devise efficient mitigation strategies targeted at bringing down the number of PM exceedance events, a more extensive set of tracers needs to be incorporated and alternate source receptor modelling approaches, e.g., PMF modelling targeted specifically towards identifying local and regional combustion sources contributing towards the emissions of PM and towards the emission gas-phase aerosol precursors, need to be adopted.

# Chapter 4

---

## **Humidity, density, and inlet aspiration efficiency correction improve accuracy of a low-cost particulate matter sensor during field calibration at a suburban site in the north-western Indo-Gangetic plain (NW-IGP)**

### **4.1. Abstract**

Low-cost particulate matter sensors are now widely used by concerned citizens to monitor PM exposure despite poor validation under field conditions. They offer several advantages over conventional research-grade instruments to the end-user, such as lower cost, more portable, and ease of use. It is important to characterize the performance of such sensors under diverse field conditions before their widespread penetration among consumers. Provided the measurements from such sensors are accurate and precise, they could potentially be used to increase the currently limited spatio-temporal network of PM monitoring. In this chapter, I report the field calibration of a modified version of the Laser Egg (LE), a commercial low-cost PM sensor against Class III US EPA Federal Equivalent Method PM<sub>10</sub> and PM<sub>2.5</sub>  $\beta$ -attenuation analyzers. The calibration was performed at IISER Mohali, a suburban site in the north-western Indo-Gangetic Plain, from April 27, 2016, to July 25, 2016. At ambient PM mass loadings ranging from  $< 1\text{-}838 \mu\text{g m}^{-3}$  and  $< 1\text{-}228 \mu\text{g m}^{-3}$  for PM<sub>10</sub> and PM<sub>2.5</sub>, respectively, measurements of PM<sub>10</sub>, PM<sub>2.5</sub> from the LE were precise, with a Pearson correlation coefficient ( $r$ )  $> 0.9$  and a percentage coefficient of variance (CV)  $< 12\%$ . The original Mean Bias Error (MBE) of  $\sim -90 \mu\text{g m}^{-3}$  decreased to  $-30.9 \mu\text{g m}^{-3}$  (Sensor 1) and  $-23.2 \mu\text{g m}^{-3}$  (Sensor 2) during the summer period (April 27 - June 15, 2016) after correcting for particle density and aspiration losses. During the monsoon period (June 16 - July 25, 2016), the MBE of the PM<sub>2.5</sub> measurements decreased from  $19.1 \mu\text{g m}^{-3}$  to  $8.7 \mu\text{g m}^{-3}$  and from  $28.3 \mu\text{g m}^{-3}$  to  $16.5 \mu\text{g m}^{-3}$  for Sensor 1 and Sensor 2, respectively after correcting for particle density and hygroscopic growth. The corrections reduced the overall MBE to  $< -20 \mu\text{g m}^{-3}$  for

This chapter was published in the journal "Aerosol Science and Technology" with B. Sinha as co-author (Pawar & Sinha 2020). I acquired the dataset, performed all the QA/QC, and wrote the draft of the manuscript. Dr. B. Sinha provided guidance and commented on the draft. All statistical analysis, figures, and tables presented in the manuscript and used in this thesis are mine

PM<sub>10</sub> and  $< 3 \mu\text{g m}^{-3}$  for PM<sub>2.5</sub>, indicating that the modified version of the LE could be used for ambient PM monitoring with appropriate correction and meteorological observations. However, users of the original product may underestimate their PM<sub>10</sub> exposure, particularly under conditions characterized by high dust loading in the air.

## 4.2. Introduction

Prolonged exposure to particulate matter is associated with adverse impacts on human health and can lead to pulmonary and cardiovascular diseases (Pope III, 2000b; Lin et al., 2002; Englert, 2004; Brunekreef and Forsberg, 2005; Pope III and Dockery, 2006; Adar et al., 2014). Rapid economic growth accompanied by a spur in industrialization, urbanization and energy consumption has led to increased PM emissions, which has put the health of citizens at stake. In 2012 alone, 0.6 million premature deaths and a loss of 25 million disability-adjusted life years have been attributed to ambient air pollution in India (Lim et al., 2012).

Traditionally, ambient PM monitoring is carried out at sparsely located research facilities or government environmental monitoring agencies. For instance, in India, under the purview of the National Air Quality Monitoring Program (NAMP), as of October 2019, the Central and State Pollution Control Boards monitor the levels of PM<sub>10</sub> and PM<sub>2.5</sub> at ~793 stations twice a week (24-hour sampling at an 8-hour sampling interval) resulting in a mere 104 annual observations per station (CPCB, 2003, 2019b). As of September 2019, only 200 government-owned continuous air quality monitoring stations in India broadcast real-time PM measurements accessible over the internet (CPCB, 2019a), a degree of coverage inadequate for the purposes of locating pollution point sources, gauging spatio-temporal variations in PM, accurately estimating exposure for a population of 1.25 billion, and devising efficient strategies to reduce ambient PM levels. High investment costs incurred during installation and maintenance of PM analyzers have hindered extensive coverage and widespread availability of measurements.

Recently, due to frequent media attention and growing public awareness, there is a surge in demand for real-time air quality data by concerned people who wish to monitor and regulate their exposure to ambient pollutants and procure low-cost sensors for their personal use. In the last few years, immense progress has been made in the development of portable, low-cost sensors by small and medium-sized enterprises for providing real-time information on PM levels (Snyder et al., 2013; Kumar et al., 2015).

Most of these low-cost PM sensors mentioned above detect particles via a light scattering method (Jiao et al., 2016). These sensors have garnered widespread attention because of their affordability

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

and their ability to quantify PM concentrations at a high spatial and temporal resolution. The data from such sensors are readily available to the user. Without a doubt, low-cost PM sensors have many promising applications. However, more validation data that allows users to make an informed decision about the quality of data they can expect in a given atmospheric environment is required. In the last two years, several studies have evaluated the performance of a few commercially available particle-sensors in the laboratory. Wang et al. (2015) assessed the performance of three low-cost PM sensors, Shinyei PPD42NS, Samyoung DSM501A, and Sharp against US EPA certified methods under laboratory conditions and reported a high dependence of the performance on particle composition, particle size, and relative humidity (RH). Manikonda et al. (2016) also evaluated the performance of four low-cost PM sensors (Speck, Dylos, TSI AirAssure, and UB AirSense) using cigarette smoke and Arizona test dust under standard RH and temperature conditions and found adequate precision for monitoring PM exposure in indoor environments. Austin et al. (2015) evaluated the performance of the Shinyei PPD42NS under laboratory conditions using monodisperse polystyrene spheres and found the sensor appropriate for low to medium concentrations of respirable particles ( $< 100 \mu\text{g m}^{-3}$ ). Lab evaluation of PM<sub>2.5</sub> measurements from 241 low-cost PMS3003 sensors (Sayahi et al., 2019b) manufactured in two separate batches revealed that the sensors belonging to the first batch (n=154) overestimated while those belonging to the second batch (n=88) underestimated ammonium nitrate mass concentration in the chamber when compared against the DustTrak (a light-scattering laser photometer), highlighting that manufacturing level differences also impact the accuracy of low-cost sensors.

Several studies have also evaluated the performance of a few lost-cost PM sensors under ambient conditions. Mukherjee et al. (2017) compared the performance of the low-cost AlphaSense Optical Particle Counter (OPC) with two reference analyzers, namely the GRIMM 11-R OPC and a  $\beta$ -attenuation Monitor (BAM-1020, Met One Instruments), over a 12-week period at a site affected by wind-blown dust in California. Although the sensors demonstrated a good correlation ( $r^2 = 0.6$  to  $0.76$ , hourly average PM<sub>10</sub> between  $20$  to  $700 \mu\text{g m}^{-3}$ ), they reported only a small fraction ( $\sim 20\%$ ) of reference (BAM-1020) measured PM<sub>10</sub>. Holstius et al. (2014) evaluated the performance of a custom-built platform that employed a Shinyei PPD42NS using the BAM-1020 as a reference at a regulatory monitoring site in California with low ambient PM<sub>2.5</sub> mass concentration between  $2$  to  $21 \mu\text{g m}^{-3}$ . They were able to explain 72% of the variance observed in 24-hour PM<sub>2.5</sub> data based on linear corrections. Jiao and co-workers (Jiao et al., 2016) tested five different types of low-cost PM sensors, using the BAM-1020 as a reference, at a site in the south-eastern US with low ambient PM<sub>2.5</sub> levels of  $\sim 10 \mu\text{g m}^{-3}$ ; ordinary least squares (OLS) regression between the data sets revealed

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

that only three sensors, namely the Dylos, the Shinyei PPD60PV, and the Shinyei PPDD42NS had an  $r$  (Pearson correlation coefficient) value more than 0.5. Nakayama et al. (2018) developed and tested the efficacy of Panasonic  $PM_{2.5}$  optical sensors by comparing year-round observations at four urban and suburban sites in Japan with Federal Equivalent Methods (FEM) deployed at observatories  $\sim 1.7$  to 4 km away and found a good correlation with slopes of 0.97 to 1.23 and an  $r$ -value of 0.89-0.95 when the daily averaged  $PM_{2.5}$  concentration varied between 5 to 55  $\mu\text{g m}^{-3}$ . They also reported that the low-cost sensors tended to overestimate  $PM_{2.5}$  at  $RH > 70\%$  compared to the reference, probably because of the hygroscopic growth of particles. Crilley et al. (2018) also demonstrated a significant positive artifact in  $PM_{2.5}$  mass concentrations measured by the low-cost AlphaSense OPC PM sensor when compared with the reference (GRIMM) at  $RH > 85\%$ .

There is an urgent need to test the efficacy of such low-cost PM sensors for regular ambient air quality monitoring in extremely polluted environments with high levels of PM and strong seasonality in meteorological conditions to evaluate whether the sensors with their factory calibration can be used in such environments. Evaluating the factory calibration is essential as the targeted end-user of the product does not have the scientific training and experience to perform a site-specific calibration or carry out post-processing of data.

In this chapter, I focus on the field calibration of a modified version of the Laser Egg (LE), a commercially available low-cost PM sensor that works on the principle of laser-based light scattering. The Laser Egg is designed, manufactured, and distributed by Kaiterra (previously Origins Technology, Beijing, China). It is currently priced at around 126 USD. My analysis focuses mainly on the implication of the assumption that aerodynamic and optical diameters are identical, which I evaluate under both high and low dust conditions, and on the performance of the sensor under high wind speed conditions. For this purpose, two identical LE monitors with modified casings to shield them from rain and solar radiation and an accessory inlet fan were co-located next to the inlets of Class III US EPA FEM compliant  $PM_{10}$  and  $PM_{2.5}$   $\beta$ -attenuation analyzers at the IISER Mohali Atmospheric Chemistry Facility, a suburban site in the north-western Indo-Gangetic Plain (NW-IGP) during a 3-month period from April 27, 2016, to July 25, 2016. The performance of the LE sensors was evaluated for regular ambient usage under varied meteorological conditions. The LE monitors also report the levels of ambient temperature and RH, which were contrasted with measurements from the Met One 064 Air Temperature sensor and the Met One 083E RH sensor, respectively.



## 4.3. Materials and methods

### 4.3.1. Study location and instrumentation

The field calibration of the LE air quality monitor was carried out at the IISER Mohali Atmospheric Chemistry Facility (30.667°N, 76.729°E, 310 m above sea level), a suburban site in the NW-IGP (Figure 4.1). Detailed site description and the measurement techniques, including data quality assurance protocols, can be found in Sinha et al. (2014b), whereas prevalent meteorology for summer and monsoon season has been described in Chapter 3.

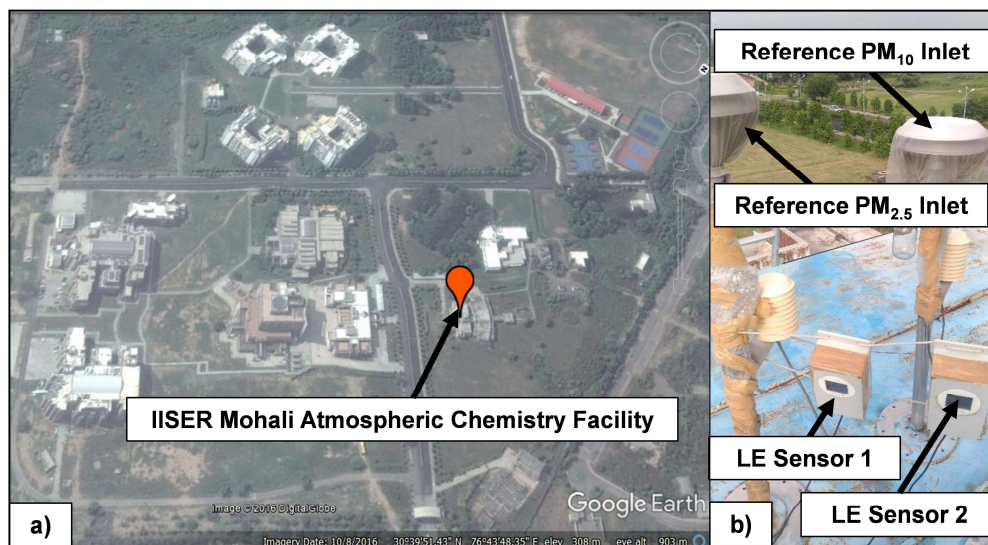


Figure 4.1: a) Location of IISER Mohali Atmospheric Chemistry Facility (red marker) inside the campus of Indian Institute of Science Education and Research (IISER) (30.667 N, 76.729 E, 310m a.s.l.) shown in a Google Earth imagery acquired on August 10, 2016 (Map image: ©2016 DigitalGlobe, Google Earth). b) Location of the two modified Laser Egg PM sensors denoted as LE Sensor 1 and LE Sensor 2 next to the inlets of reference PM<sub>10</sub> and PM<sub>2.5</sub>  $\beta$ -attenuation analyzers at the IISER Mohali Atmospheric Chemistry Facility.

Reference particulate matter (PM<sub>10</sub> and PM<sub>2.5</sub>) mass concentrations were measured using separate Thermo Fisher Scientific 5014i  $\beta$ -continuous ambient particulate monitors working on the principle of  $\beta$ -attenuation. This instrument has been certified by the US EPA as a Class III Automated Equivalent method: EQPM-1102-50 (for PM<sub>10</sub>) and EQPM-0609-183 (for PM<sub>2.5</sub>) for 24-hour average measurements (EPA, 2019). Ambient air is drawn through a size-selective inlet at a constant flow rate of 16.67 liters per minute. The constituent particles are deposited onto a quartz fiber filter tape. The attenuation of  $\beta$  rays emanating from a <sup>14</sup>C source by the deposited particles is used as a metric to estimate PM mass concentration. When the ambient RH is low (< 40%), the measurements are made under ambient conditions. However, under high ambient RH conditions (> 40%), an inlet heating system in the reference analyzer regulates the sample RH at ~40% to remove particle-bound water and measures the dry mass of the aerosol. Apart from

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

moisture removal, inlet heating can also result in the loss of semi-volatile aerosol species such as ammonium nitrate and ammonium chloride, which are in a temperature and RH-dependent equilibrium with the gas phase (Eatough et al., 2003; Triantafyllou et al., 2016). These analyzers are accurate up to  $\pm 5\%$  and calibrated using a National Institute of Standards and Technology (NIST) traceable mass foil set. They offer a resolution of  $\sim 0.1 \mu\text{g m}^{-3}$  and precision of  $\pm 2 \mu\text{g m}^{-3}$  (ambient PM  $< 80 \mu\text{g m}^{-3}$ ) and  $4\text{-}5 \mu\text{g m}^{-3}$  (ambient PM  $> 80 \mu\text{g m}^{-3}$ ) (Table 4.1). As a regular maintenance protocol, the inlets of both PM<sub>10</sub> and PM<sub>2.5</sub> analyzers were cleaned on April 28, 2016, near the beginning of the 3-month calibration period on April 27, 2016. Span calibration for the mass absorption coefficient of the PM analyzers was carried out on May 5, 2016, and the difference obtained was less than 2%.

Parameter	low-cost Laser Egg (LE) air quality monitor	reference Thermo Fisher 5014 $\beta$ -attenuation PM analyzer
Measurement principle	Laser-based light scattering	$\beta$ -attenuation
Measurable particle size	0.3 $\mu\text{m}$ -10 $\mu\text{m}$	PM <sub>2.5</sub> (aerodynamic diameter $< 2.5 \mu\text{m}$ ) PM <sub>10</sub> (aerodynamic diameter $< 10 \mu\text{m}$ )
Accuracy	$\pm 10\%$	$\pm 5\%$
Resolution	1 $\mu\text{g m}^{-3}$	0.1 $\mu\text{g m}^{-3}$
Precision	-	$\pm 2 \mu\text{g m}^{-3}$ ( $< 80 \mu\text{g m}^{-3}$ ) $\pm 4\text{-}5 \mu\text{g m}^{-3}$ ( $> 80 \mu\text{g m}^{-3}$ )
Response time	0.1 second	1 second
Measurement range	1-999 $\mu\text{g m}^{-3}$	1-10000 $\mu\text{g m}^{-3}$

Table 4.1: Characteristics of the low-cost Laser Egg (LE) air quality monitor and the reference Thermo Fisher 5014i  $\beta$ -attenuation PM analyzer.

A dedicated meteorological station (Met One Instruments Inc., Rowlett, USA) provided measurements of wind direction, wind speed, ambient temperature, RH, solar radiation (SR), and rainfall at a temporal resolution of 1 minute. Wind speed and wind direction were measured using the Met One 034B Wind Sensor with an accuracy of  $\pm 1.1\%$  and  $\pm 4^\circ$ , respectively. Ambient temperature was measured using the Met One 064 Air Temperature sensor with an accuracy of

$\pm 0.1^\circ\text{C}$ . The sensor contains a multi-element thermistor that produces a change in resistance in response to changes in ambient temperature. Ambient reference RH was measured using the Met One 083E sensor (stated accuracy:  $\pm 2\%$ ), which measures variance in the capacitance of a  $1\ \mu\text{m}$  thick dielectric polymer layer in response to changes in RH. SR was measured using a Model 094 Pyranometer comprising a multi-junction differential thermopile, which produces a voltage difference upon absorption of incident photons. Rainfall was measured using the Met One 364-1 precipitation gauge (detection limit: 0.1 mm) with a dual-chambered tipping bucket.

#### 4.3.2. Laser Egg air quality monitor

The LE air quality monitor, designed initially for indoor PM monitoring (Kaiterra, previously Origins Technology Beijing, China), is a small portable device that provides real-time measurements of  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  mass concentrations and the corresponding Air Quality Index (AQI). Briefly, the monitor works on the principle of laser-based light scattering (Mie). A small fan draws in ambient air laden with particles that scatter the laser beam (650 nm wavelength). A photodiode positioned at  $90^\circ$  to the beam detects the scattered light and converts it to particle size and number concentration based on a series of proprietary algorithms, and the monitor eventually reports PM mass concentration and AQI (Zuo et al., 2018). As per the manufacturer's specifications, the monitor can detect particles in a size range of  $0.3\ \mu\text{m}$  -  $10\ \mu\text{m}$ . Table 4.1 provides further details about the LE monitor. The South Coast Air Quality Management District (SCAQMD), a regulatory body that carried out field and laboratory evaluation of several low-cost PM sensors, claims that the LE monitors use Plantower PMS3003 sensors to measure PM. However, this information has not been corroborated by the manufacturer (SCAQMD, 2016). The monitors require an active Wi-Fi connection to transmit and store the data on a server. They do not possess a secondary data storage option. The real-time data from the sensors can also be remotely accessed via a mobile application called "Breathing Space."

For this field study, the original LE monitor was modified to make it more rugged for continuous usage under ambient conditions. The sensing assembly and the display unit were placed in an Acrylonitrile Butadiene Styrene (ABS) box to shield them from sunlight and rain. An accessory fan (NIDEC DF251R, flowrate  $0.04\text{m}^3\ \text{min}^{-1}$ ) was placed at the inlet facing downwards (Figure 4.2) to increase the airflow into the box. The sensing assembly was placed above the outflow of the inlet fan. Hereafter, "LE" in this chapter refers to the modified version of the air quality monitor. Two identical LE monitors were co-located next to the inlets of reference  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  analyzers (Figure 4.1b) to minimize artifacts arising from the heterogeneous nature of the ambient air.



Figure 4.2: Figure showing the external casing, accessory inlet fan, and internal circuitry of the modified Laser Egg (LE) PM sensor.

#### 4.3.3. Data processing and statistical analysis

The LE sensors were deployed from April 27, 2016, to July 25, 2016. The data from the sensors were initially provided in JSON (JavaScript Object Notation) format, which was converted to CSV (comma-separated value) format using a freeware called Opal-Convert. The LE monitor gives the measurement of PM<sub>10</sub>, PM<sub>2.5</sub>, RH, and ambient temperature at a temporal resolution of 5 minutes. Each parameter was averaged hourly, provided a minimum of 75% data was available for that hour (a minimum of nine 5-minute measurements per hour); else, that value was considered missing (Mukherjee et al., 2017; Badura et al., 2018; Wang et al., 2019). The availability of hourly averaged data points in % during the calibration period was only 48% and 28% for Sensor 1 and Sensor 2, respectively, because the Wi-Fi connection necessary for transferring data to the cloud server would self-terminate and had to be reset frequently.

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

Reference PM<sub>10</sub> and PM<sub>2.5</sub> analyzers, ambient temperature, and RH sensors yielded data at a temporal resolution of 1 min, which was averaged hourly. The availability of hourly averaged data points in % was 96% and 93% for the PM<sub>10</sub>, and PM<sub>2.5</sub> analyzers, respectively, and 100% for the reference ambient temperature and RH sensors.

The PM<sub>10</sub>, PM<sub>2.5</sub>, RH, and temperature measurements obtained from the LE air quality monitors were then assessed for precision, the linearity of response, and accuracy by computing a suite of statistical parameters as described below.

Percentage coefficient of variance (CV) (Jiao et al., 2016; Sousan et al., 2016; Sousan et al., 2017; Badura et al., 2018; Crilley et al., 2018; Levy Zamora et al., 2019) was used as an index of the LE monitors' precision and reproducibility, which were evaluated as per the expression:

$$CV (\%) = \frac{1}{n} \sum_{i=1}^n CV_i \quad (43)$$

where  $CV_i$  is given by

$$CV_i (\%) = \left( \frac{\sigma (LE1_i, LE2_i)}{\mu (LE1_i, LE2_i)} \right) \times 100 \quad (44)$$

Here,  $\sigma$  and  $\mu$  represent the standard deviation and mean of the  $i^{\text{th}}$  measurement obtained from Sensor 1 (LE1) and Sensor 2 (LE2).

The LE monitor's linearity of response for PM measurements was assessed from the slope and the Pearson correlation coefficient,  $r$ , obtained after performing a Reduced Major Axis (RMA) regression between the measurements from the low-cost sensor and the corresponding reference instrument using a freeware called Paleontological Statistics (PAST) (Hammer et al., 2001a) as per the algorithm described in Warton et al. (2006). Errors associated with the reference PM analyzer would be particularly high when ambient RH exceeds 40% (~71% of the time during the field study), resulting in loss of semi-volatile species due to dynamic heating of sampled air, a feature absent in the LE monitor. Hence, RMA regression was preferred over the OLS regression for an unbiased assessment (Ayers, 2001) as it accounts for errors in both the independent (reference PM) and dependent (LE PM) variables. For inter-comparing, the LE monitor's RH and ambient temperature measurements with the reference, OLS regression was used.

Karagulian et al. (2019) and Williams et al. (2019) highlighted that slope and correlation coefficient alone is insufficient to comprehensively assess the low-cost sensor performance. Therefore, I also computed the Root Mean Square Error (RMSE) and normalized RMSE (nRMSE)

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

(Zíková et al., 2017; Sayahi et al., 2019a) to determine the accuracy of the PM<sub>10</sub> and PM<sub>2.5</sub> measurements obtained from the two LE sensors as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (LE_i - Ref_i)^2} \quad (45)$$

and

$$nRMSE = \left( \frac{RMSE}{\frac{1}{n} \sum_{i=1}^n Ref_i} \right) \times 100 \quad (46)$$

where  $LE_i$  and  $Ref_i$  refer to the  $i^{\text{th}}$  measurement obtained from the LE and reference analyzer.

Mean Bias Error (MBE), evaluated as below, was used as an estimate of average bias in the LE sensor's measurements (Cross et al., 2017; Zimmerman et al., 2018)

$$MBE = \frac{1}{n} \sum_{i=1}^n (LE_i - Ref_i) \quad (47)$$

where  $LE_i$  and  $Ref_i$  refer to the  $i^{\text{th}}$  measurement obtained from the LE and reference analyzer.

Additionally, the coefficient of divergence (COD) was calculated to estimate the heterogeneity in PM mass concentration from the reference and the LE according to the expression below

$$COD = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{PM_i^{Ref} - PM_i^{LE}}{PM_i^{Ref} + PM_i^{LE}} \right)^2} \quad (48)$$

where  $PM_i^{Ref}$  refers to  $i^{\text{th}}$  PM (PM<sub>10</sub> or PM<sub>2.5</sub>) measurement from the reference analyzer and  $PM_i^{LE}$  refers to the corresponding  $i^{\text{th}}$  PM measurement from the LE sensor. The value of the COD varies from 0 to 1 with a low value (< 0.2), indicating high homogeneity and vice versa. The COD has been used as a metric in several studies to describe intra-urban heterogeneity in PM concentrations (Wilson et al., 2005; Wilson et al., 2006; Massoud et al., 2011). Here, I use the value of COD to assess the homogeneity in PM measurements from the reference analyzers and LE air quality monitors. I compute the COD for each hourly averaged measurement point and then study the variation in the daily average value of the COD throughout the calibration period.

## 4.4. Results and discussion

### 4.4.1. Ambient temperature, relative humidity, and solar radiation measurements

Figure 4.3 shows the hourly averaged values of locally measured meteorological parameters: wind speed, ambient temperature, SR,  $RH^{Ref}$ , and absolute humidity ( $AH^{Ref}$ ), the hourly median value of wind direction, and the hourly sum of rainfall recorded during the calibration period from April 27, 2016, to July 25, 2016. The ambient temperature during the calibration period varied from a minimum of 21°C to a maximum of 42°C. The average wind speed observed between April 27, 2016, and July 25, 2016, was in the range of 4.1 to 6.8  $m\ s^{-1}$ . Episodic dust storms were also observed when the wind speeds increased from  $\sim 12\ m\ s^{-1}$  to  $15\ m\ s^{-1}$ . The maximum daytime SR was  $\sim 400\ W\ m^{-2}$  on cloudy days and  $\sim 700\ W\ m^{-2}$  on clear days.

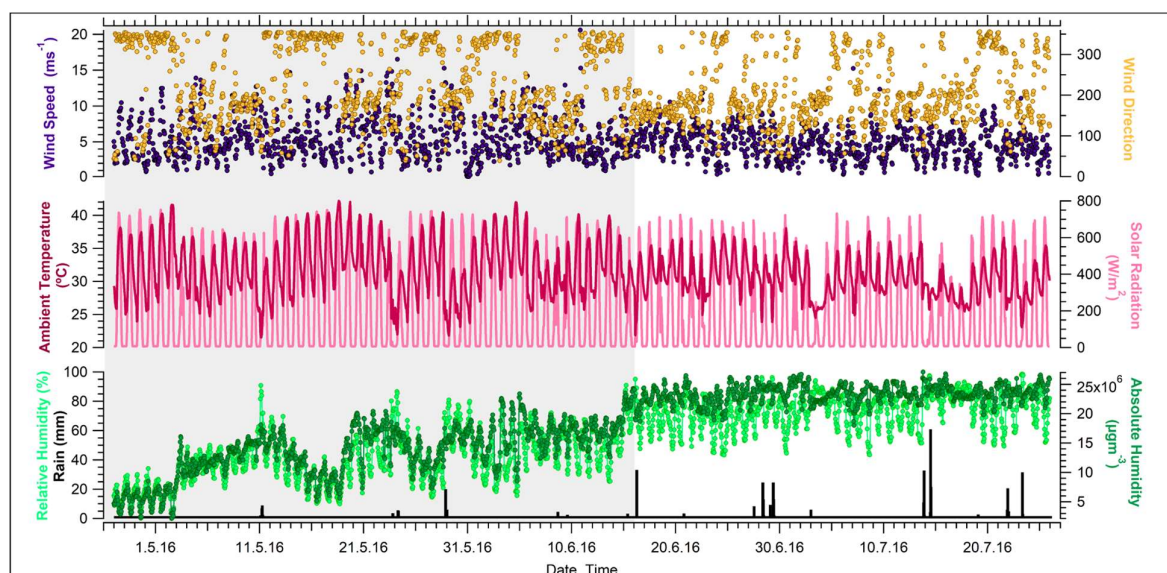


Figure 4.3: Hourly averaged values of locally measured meteorological parameters: wind speed, ambient temperature, solar radiation, relative humidity and absolute humidity, the hourly median value of wind direction and the hourly sum of rainfall recorded at IISER Mohali Atmospheric Chemistry Facility during the three-month calibration period from April 27, 2016, till July 25, 2016. The shaded portion represents the dry summer season (April 27, 2016 – June 15, 2016), while the unshaded portion represents the monsoon season (June 16, 2016 – July 25, 2016).

The period of April to June had different meteorology characterized by higher wind speeds, higher ambient temperature, and lower  $RH^{Ref}$  in comparison to July. From April 27, 2016, to May 2, 2016, dry conditions were observed with average ambient  $RH^{Ref}$  varying between 9% and 23% and an  $AH^{Ref}$  of 4 to 7  $g\ m^{-3}$ . From May 3, 2016, to June 15, 2016, average ambient  $RH^{Ref}$  varied between 28% and 58%, and  $AH^{Ref}$  fluctuated between 11 and 18  $g\ m^{-3}$ . From June 16, 2016, to July 25, 2016, humid conditions were observed during which  $AH^{Ref}$  fluctuated between 21 and 24  $g\ m^{-3}$ ,

and the average  $RH^{Ref}$  varied from 61% to 85%. Thus, the 3-month calibration period allowed for the testing of the efficacy of the LE monitors under varied meteorological conditions.

#### 4.4.2. Particulate Matter measurements

##### 4.4.3. $PM_{10}$ measurements

Figure 4.4a (top panel) shows the hourly averaged  $PM_{10}$  mass concentrations obtained from the reference  $\beta$ -attenuation analyzer (circular markers in orange) and the two LE air quality monitors (green triangular and blue square markers) for the period from April 27, 2016, to July 25, 2016.

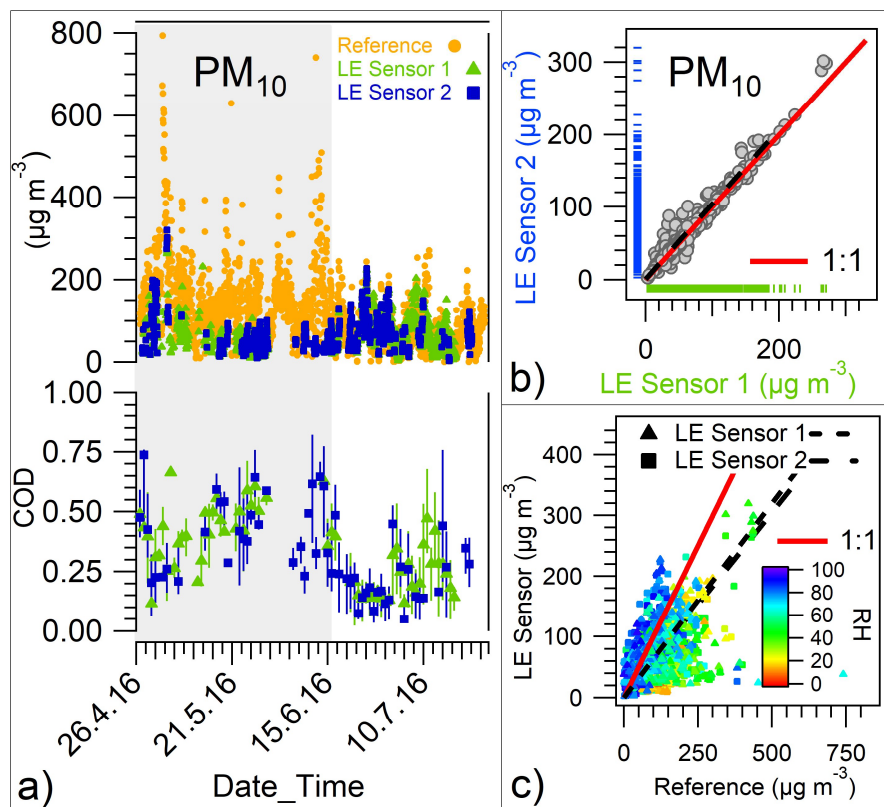


Figure 4.4: Time series plot of hourly averaged  $PM_{10}$  mass concentration from reference analyzer (circular markers in orange) and two Laser Egg (LE) sensors (green triangular and blue square markers respectively) for the period from April 27, 2016, to July 25, 2016. The shaded portion represents the dry summer period (April 27, 2016 – June 15, 2016) with average RH varying from 9% to 58%, while the unshaded portion represents the wet monsoon period (June 16, 2016 – July 25, 2016) with average RH varying from 61% to 85%. (a, bottom) The daily average value of the coefficient of divergence (COD) for raw  $PM_{10}$  measurements from the LE sensors. Vertical bars represent the daily variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles of the COD. b) Reduced Major Axis (RMA) regression of  $PM_{10}$  from LE Sensor 2 versus Sensor 1. Marginal rugs have been added to show the distribution of data c) RMA regression of  $PM_{10}$  from the LE sensors versus the reference analyzer. Markers are color-coded according to ambient RH.

The  $PM_{10}$  mass loadings reported by the reference analyzer and LE Sensors 1 and 2 varied from  $< 1$  to  $838 \mu g m^{-3}$ ,  $3.0$  to  $270.3 \mu g m^{-3}$ , and  $2.8$  to  $319.8 \mu g m^{-3}$ , respectively. The average (1 $\sigma$  ambient variability) of  $PM_{10}$  during the calibration period was  $121.0$  ( $87.1$ )  $\mu g m^{-3}$ ,  $62.8$  ( $41.3$ )  $\mu g$



#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

$\text{m}^{-3}$ , and 69.8 (45.2)  $\mu\text{g m}^{-3}$  as measured by the reference analyzer, LE Sensor 1 and Sensor 2 respectively (Table 4.2).

When hourly averaged  $\text{PM}_{10}$  measurements from three identical original LE sensors were compared against a BAM for two months as part of a field evaluation carried out by SCAQMD at an ambient monitoring site in southern California, a similar underestimation, but with a lower  $r$ -value (0.2) was reported (SCAQMD, 2016). Though the results of the SCAQMD study are preliminary, they reinforce the point that LE monitors do not provide a very accurate measure of ambient  $\text{PM}_{10}$  mass concentration. Another low-cost sensor, the AlphaSense OPC-N2, underestimated  $\text{PM}_{10}$  (slope of  $\sim 0.2$  from OLS,  $r = 0.7-0.9$ ) compared to reference  $\beta$ -attenuation monitors during a field-calibration at a site characterized by heavy dust loading in California (Mukherjee et al., 2017).

	$\text{PM}_{10}$ ( $\mu\text{g m}^{-3}$ )	$\text{PM}_{2.5}$ ( $\mu\text{g m}^{-3}$ )	Ambient temperature ( $^{\circ}\text{C}$ )	Relative humidity (%)	Absolute humidity ( $\text{g m}^{-3}$ )
<b>Reference</b>	121.0 (87.1)	50.0 (33.0)	31.5 (3.9)	55.0 (22.4)	17.9 (6.0)
<b>Laser Egg Sensor 1</b>	62.8 (41.3)	51.6 (32.6)	32.8 (7.0)	23.4 (3.0)	7.4 (0.9)
<b>Laser Egg Sensor 2</b>	69.8 (45.2)	62.1 (39.5)	35.1 (7.5)	27.0 (3.4)	8.9 (1.0)

Table 4.2: Average ( $1\sigma$  ambient variability) of  $\text{PM}_{10}$ ,  $\text{PM}_{2.5}$ , ambient temperature, relative humidity, and absolute humidity from the reference analyzers and the two Laser Egg (LE) sensors for the calibration period from April 27, 2016, until July 25, 2016. The number in the bracket denotes one sigma ambient variability.

During the two severest dust storms observed during the calibration period on May 2, 2016, 21:00 Local Time (LT) and June 11, 2016, 21:00 LT, the hourly averaged  $\text{PM}_{10}$  mass concentration reported by the reference analyzer was 794  $\mu\text{g m}^{-3}$  and 741  $\mu\text{g m}^{-3}$  respectively while the LE sensors reported 120  $\mu\text{g m}^{-3}$  and 38.4  $\mu\text{g m}^{-3}$  (Figure 4.4c) respectively. Likewise, a field study carried out in Salt Lake City, Utah reported that Plantower PMS sensors that also work on the principle of laser-based light scattering are not very accurate at measuring  $\text{PM}_{10}$  when the ambient PM contains a higher coarse fraction (Sayahi et al., 2019a), particularly during dust storm events,

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

when the reference Tapered Element Oscillating Microbalance (TEOM) reported  $\sim 472 \mu\text{g m}^{-3}$  while the low-cost sensor reported a mass concentration  $< 50 \mu\text{g m}^{-3}$ .

A scatter plot of  $\text{PM}_{10}$  measurements from the two LE sensors reveals a slope of 1.04 (Table 4.4) and an  $r$ -value of 0.98, indicating high precision, while a low average CV of 7.9% showed that the measurements were highly reproducible and in compliance with the US EPA standard of CV less than 10% between identical units (EPA, 2015a). The LE  $\text{PM}_{10}$  measurements were more precise than those of another low-cost sensor, the Alphasense OPC-N2, which reported an average CV of  $\sim 22\%$  during a field campaign in the UK (Crilley et al., 2018). The bottom panel of Figure 4.4a shows the daily average value of the COD for the two LE sensors. Vertical bars represent the daily variability of the COD as the 75<sup>th</sup> and 25<sup>th</sup> percentiles. The daily average value of the COD for  $\text{PM}_{10}$  from the two low-cost sensors varied from 0.3 to 0.6 during the days (April 27, 2016, to June 15, 2016) with low to moderate levels of ambient  $\text{RH}^{\text{Ref}}$  and high wind speeds (Figure 4.3). However, from June 16, 2016, with an increase in  $\text{RH}^{\text{Ref}}$  and  $\text{AH}^{\text{Ref}}$  and a decrease in wind speeds,  $\text{PM}_{10}$  reported by the low-cost sensors became more homogeneous to the reference analyzer with the daily-average COD varying from 0.1 to 0.3, which indicates that the accuracy of  $\text{PM}_{10}$  measurements from the LE was affected by particle size distribution, possibly wind speed and ambient RH.

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

<b>PM<sub>10</sub></b>						
	<b>Full Calibration Period</b> (April 27 – July 25, 2016)		<b>Summer season</b> (April 27 – June 15, 2016)		<b>Monsoon season</b> (June 16 – July 25, 2016)	
	Raw (LES1/LES2)	Post-correction (LES1/LES2)	Raw (LES1/LES2)	Post-correction (LES1/LES2)	Raw (LES1/LES2)	Post-correction (LES1/LES2)
<b>RMSE</b> ( $\mu\text{g m}^{-3}$ )	75.1/84.3	58.0/69.8	110.7/115.9	77.0/93.1	43.8/40.8	43.8/40.4
<b>nRMSE</b> (%)	73.6/72.3	56.8/59.9	75.3/77.7	52.4/61.8	57.0/46.7	57.0/46.1
<b>MBE</b> ( $\mu\text{g m}^{-3}$ )	-38.2/-44.9	-20.0/-18.1	-91.7/-87.3	-30.9/-23.2	-8.1/-8.6	-13.9/-13.6

<b>PM<sub>2.5</sub></b>						
	<b>Full Calibration Period</b> (April 27 – July 25, 2016)		<b>Summer season</b> (April 27 – June 15, 2016)		<b>Monsoon season</b> (June 16 – July 25, 2016)	
	Raw (LES1/LES2)	Post-correction (LES1/LES2)	Raw (LES1/LES2)	Post-correction (LES1/LES2)	Raw (LES1/LES2)	Post-correction (LES1/LES2)
<b>RMSE</b> ( $\mu\text{g m}^{-3}$ )	31.0/37.9	27.5/32.2	26.1/35.3	29.4/34.4	34.1/40.7	26.1/29.6
<b>nRMSE</b> (%)	69.7/75.5	61.8/64.1	47.3/60.9	53.4/59.4	91.8/97.4	70.1/70.9
<b>MBE</b> ( $\mu\text{g m}^{-3}$ )	6.6/11.2	-1.9/1.3	-11.3/-4.2	-17.2/-12.4	19.1/28.3	8.7/16.5

Table 4.3: Table showing the Root Mean Square Error (RMSE), normalized Root Mean Square Error (nRMSE), and Mean Bias Error (MBE) computed for raw and post-correction PM<sub>10</sub> and PM<sub>2.5</sub> measurements obtained from Laser Egg sensors denoted as LES1 and LES2.

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

	$PM_{10}$ ( $\mu\text{g m}^{-3}$ )	$PM_{2.5}$ ( $\mu\text{g m}^{-3}$ )	Ambient temperature ( $^{\circ}\text{C}$ )	Relative humidity (%)	Absolute humidity ( $\text{g m}^{-3}$ )
<b>Comparison of Laser Egg (LE) Sensor with Reference (Ref) sensors</b>					
	RMA	RMA	OLS	OLS	OLS
<b>LE Sensor 1</b>	$S = 0.62 \pm 0.01$	$S = 1.13 \pm 0.02$	$S = 1.55 \pm 0.03$	$S = 0.08 \pm 0.00$	$S = -0.05 \pm 0.00$
<b>= <math>S \times \text{Ref} + I</math></b>	$I = 0$	$I = 0$	$I = -15.29 \pm 0.95$	$I = 18.63 \pm 0.23$	$I = (8.32 \pm 0.9) \times 10^6$
	$r = 0.43$	$r = 0.54$	$r = 0.85$	$r = 0.56$	$r = -0.31$
	RMA	RMA	OLS	OLS	OLS
<b>LE Sensor 2</b>	$S = 0.61 \pm 0.01$	$S = 1.23 \pm 0.03$	$S = 1.64 \pm 0.04$	$S = 0.10 \pm 0.00$	$S = -0.04 \pm 0.01$
<b>= <math>S \times \text{Ref} + I</math></b>	$I = 0$	$I = 0$	$I = -16.76 \pm 1.16$	$I = 21.57 \pm 0.28$	$I = (9.62 \pm 0.13) \times 10^6$
	$r = 0.41$	$r = 0.51$	$r = 0.88$	$r = 0.66$	$r = -0.23$
<b>Inter Laser Egg (LE) Sensor Comparison</b>					
	RMA	RMA	OLS	OLS	OLS
<b>LE Sensor 2</b>	$S = 1.04 \pm 0.01$	$S = 1.13 \pm 0.01$	$S = 0.99 \pm 0.01$	$S = 1.07 \pm 0.01$	$S = 1.17 \pm 0.01$
<b>= <math>S \times \text{LE}</math></b>	$I = 0$	$I = 0$	$I = 1.76 \pm 0.21$	$I = 2.52 \pm 0.26$	$I = (0.1 \pm 0.1) \times 10^6$
<b>Sensor 1 + I</b>	$r = 0.98$	$r = 0.98$	$r = 0.99$	$r = 0.98$	$r = 0.97$
	$\text{CV} = 7.9\%$	$\text{CV} = 11.2\%$	$\text{CV} = 3.2\%$	$\text{CV} = 11.8\%$	$\text{CV} = 11.8\%$

Table 4.4: Value of the slope (S), intercept (I), and Pearson correlation coefficient (r) obtained from the regression of Laser Egg sensors (LE Sensor 1 and 2) with Reference (Ref) analyzers and inter-comparison of two LE sensors. % CV (coefficient of variation) depicts the relative precision between two LE sensors. RMA denotes Reduced Major axis regression, and OLS represents Ordinary Least Squares regression.

#### 4.4.4. $PM_{2.5}$ measurements

Figure 4.5a (top panel) shows the hourly averaged  $PM_{2.5}$  mass concentration obtained from the reference  $\beta$ -attenuation analyzer (circular markers in orange) and the two LE sensors (green triangular and blue square markers) for the period of April 27, 2016, to July 25, 2016. The ambient  $PM_{2.5}$  mass loadings reported by the reference analyzer and LE Sensors 1 and 2 varied from  $< 1$  to  $228 \mu\text{g m}^{-3}$ ,  $2.8$  to  $211.2 \mu\text{g m}^{-3}$ , and  $3.2$  to  $267.8 \mu\text{g m}^{-3}$ , respectively. The average ( $1\sigma$  ambient variability) for  $PM_{2.5}$  during the calibration period was  $50.0$  ( $33.0$ )  $\mu\text{g m}^{-3}$ ,  $51.6$  ( $32.6$ )  $\mu\text{g m}^{-3}$ , and  $62.1$  ( $39.5$ )  $\mu\text{g m}^{-3}$ , as measured by the reference analyzer and LE Sensor 1 and Sensor 2,

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

respectively (Table 4.2). The RMA of all  $PM_{2.5}$  measurements acquired from LE Sensors 1 and 2 versus  $PM_{2.5}$  from the reference analyzer during the calibration period reveal a moderate correlation with an r-value of 0.54 and 0.51 (Table 4.4, Figure 4.5c), respectively, which was comparable to few low-cost PM sensors namely Dylos, Shinyei PPD60PV, Shinyei PPDD42NS (Jiao et al., 2016) and AlphaSense (Crilley et al., 2018) but lower than many optical sensors such as Panasonic  $PM_{2.5}$  sensor (Nakayama et al., 2018), Nova SDS011 (Liu et al., 2019a), Shinyei PPD42NS (Holstius et al., 2014), Plantower PMS7003 (Wang et al., 2019), Plantower PMS3003 (Zheng et al., 2018). When hourly averaged  $PM_{2.5}$  measurements from three identical original LE sensors were compared against BAM for two months as a part of field evaluation carried out by SCAQMD at an ambient monitoring site in southern California, a higher r-value of 0.8 was reported (SCAQMD, 2016).

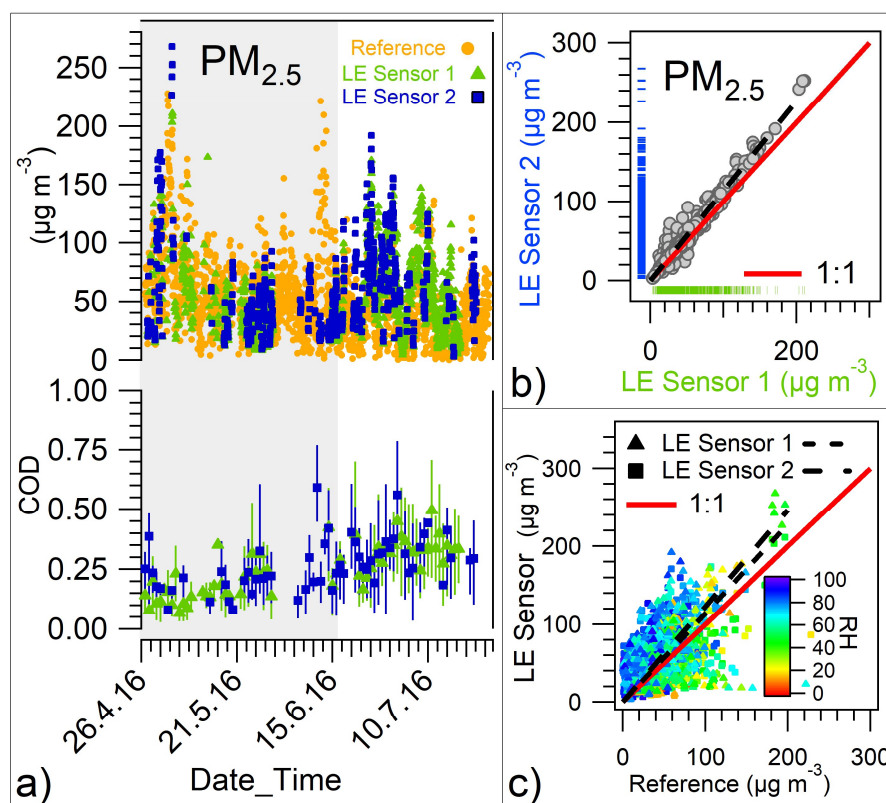


Figure 4.5 (a, top) Time series plot of hourly averaged  $PM_{2.5}$  mass concentration from the reference analyzer (circular markers in orange) and two Laser Egg (LE) sensors (green triangular and blue square markers respectively) for the period from April 27, 2016, to July 25, 2016. The shaded portion represents the dry summer period (April 27, 2016 – June 15, 2016) with average RH varying from 9% to 58%, while the unshaded portion represents the wet monsoon period (June 16, 2016 – July 25, 2016) with average RH varying from 61% to 85%. (a, bottom) The daily average value of the coefficient of divergence (COD) for raw  $PM_{2.5}$  measurements from the LE sensors. Vertical bars represent the daily variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles of the COD. b) Reduced Major Axis (RMA) regression of  $PM_{2.5}$  from LE Sensor 2 versus

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

Sensor 1. Marginal rugs have been added to show the distribution of data. c) RMA regression of  $PM_{2.5}$  from the LE sensors versus the reference analyzer. Markers are color-coded according to ambient RH.

A scatter plot of  $PM_{2.5}$  measurements from the two LE sensors reveals a slope of 1.13 and an  $r$ -value of 0.98, indicating that Sensor 2 moderately overestimated  $PM_{2.5}$  compared to Sensor 1 (Table 4.4, Figure 4.5b). In this context, it is worth noticing that the  $AH^{LE}$  within the outer casing of the LE is always significantly lower than the ambient  $AH^{Ref}$  (Figure 4.15), indicating that the outer casing may be acting as a dryer and that the outer casing of Sensor 2 had a 5-10 % higher RH inside the unit than Sensor 1 (Figure 4.15a). Hence the difference might primarily be driven by differences in the aerosol wet diameter within the two units. A scatter plot of  $PM_{2.5}$  measurements from the two LE sensors versus the reference analyzer reveals a slope of 1.13 and 1.23 (Table 4.4), indicating overestimation. Field evaluations carried out in India (Zheng et al., 2018), China (Barkjohn et al., 2018), and the USA (Kelly et al., 2017; SCAQMD, 2019) also report the tendency of PMS3003 sensors to overestimate ambient  $PM_{2.5}$  mass concentrations when compared to FEM measurement techniques.

A CV of 11.2% between the  $PM_{2.5}$  measurements from the two LE sensors was comparable to that obtained from the Plantower PMSA003 during a field study in Baltimore, Maryland (Levy Zamora et al., 2019) and lower (more precise) than that of another low-cost sensor, the Alphasense OPC-N2, which reported an average CV of ~25% during a field campaign in the UK (Crilley et al., 2018). However,  $PM_{2.5}$  measurements from the LE monitors did not meet the US EPA standard of CV less than 10% between identical units (EPA, 2015a). It seems that the sensor performance could be improved by placing a dryer in the outer casing of both units.

The bottom panel of Figure 4.5a shows the daily average value of COD for the two LE sensors. Vertical bars represent the daily variability of the COD as the 75<sup>th</sup> and 25<sup>th</sup> percentiles. The daily average value of COD for the two LE sensors varied from 0.1 to 0.3 during the days (April 27, 2016, to June 15, 2016) with low levels (< 40%) of  $RH^{Ref}$  (Figure 4.3), indicating high homogeneity in  $PM_{2.5}$  measurements. However, from June 16, 2016, with an increase in  $RH^{Ref}$  and  $AH^{Ref}$ , the  $PM_{2.5}$  measurements from the LE became more heterogeneous in comparison to the reference, with the daily average COD varying from 0.2 to 0.4, which indicates that ambient RH influences the accuracy of  $PM_{2.5}$  measurements from LE monitors.

##### 4.4.5. Impact of aspiration losses on Laser Egg accuracy

Particles entering the outer casing of the LE must make a 90° turn with respect to the ambient horizontal wind direction, as is typical for all down-facing inlets. Particle losses are proportional

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

to the particle Stokes number ( $Stk$ ) and the ratio between the external ( $U_o$ ) and internal ( $U$ ) velocity of the carrier gas and are hence expected to be larger for larger particles and higher wind speeds.

Hangal and Willeke (1990) gave the following expression for the aspiration efficiency of a 90° sampling inlet:

$$\text{Aspiration efficiency} = 1 - 3 \times Stk \sqrt{\frac{U}{U_o}} \quad (49)$$

where  $Stk$  is the particle Stokes number given as

$$Stk = \frac{\rho_p d_p^2 U_o}{18\mu D} \quad (50)$$

and  $U$  is the inlet velocity,  $U_o$  is the wind speed,  $d_p$  is the particle diameter,  $\rho$  is the particle density,  $\mu$  is the air viscosity, and  $D$  is the characteristic inlet dimension.

Figure 4.6 shows the aspiration efficiency of the LE inlet as a function of the wind speed for particles with varying aerodynamic diameters. It is evident that smaller particles have larger aspiration efficiency at any given wind speed. Moreover, for the inlet sampling configuration of the LE monitor, a particle with an aerodynamic diameter of 10  $\mu\text{m}$ , which corresponds to an optical equivalent diameter of 6 to 7  $\mu\text{m}$  for mineral dust, will not be aspirated into the detection chamber beyond external wind speeds of 6  $\text{ms}^{-1}$ . Therefore, the LE monitor is expected to underestimate  $\text{PM}_{10}$  compared to the reference analyzer (which has close to 100% inlet efficiency for the full ambient wind speed range), particularly when the  $\text{PM}_{10}$  aerosol is dominated by coarse mode dust particles and wind speeds are high. However, at the average wind speeds (4.1 to 6.8  $\text{ms}^{-1}$ ) observed in this study period, the LE measures  $\text{PM}_{10}$  thanks to its modified inlet. So far, only two other studies (Mukherjee et al., 2017; Sayahi et al., 2019a) have looked at the impact of wind speed on PM measurements from low-cost sensors. Mukherjee et al. (2017) reported that the Airbeam (a low-cost OPC) underestimated  $\text{PM}_{2.5}$  compared to the GRIMM (reference OPC) when wind speeds were between 1 to 3  $\text{ms}^{-1}$  but overestimated at higher wind speeds, which is opposite to what I observe in this study, and possibly related to the blunt shape of the Airbeam sensor. Vanderpool et al. (2018) reported sampling effectiveness above 100% for super isokinetic sampling and below 100% for sub isokinetic sampling with a blunt PQTSP sampler in a wind tunnel study. Another study (Sayahi et al., 2019a) did not explicitly analyze the impact of wind speed on the accuracy of PM measurements but instead observed the direct correlation between hourly averaged  $\text{PM}_{2.5}$  and  $\text{PM}_{10}$  measured from Plantower PMS (PMS1003 and PMS5003) sensors and wind speed (varied between 0.08 to 5.9  $\text{ms}^{-1}$ ) and found no discernible trend ( $r^2 < 0.138$ ) during a long-term field evaluation carried out at Salt Lake City, Utah from 2016 to 2017.

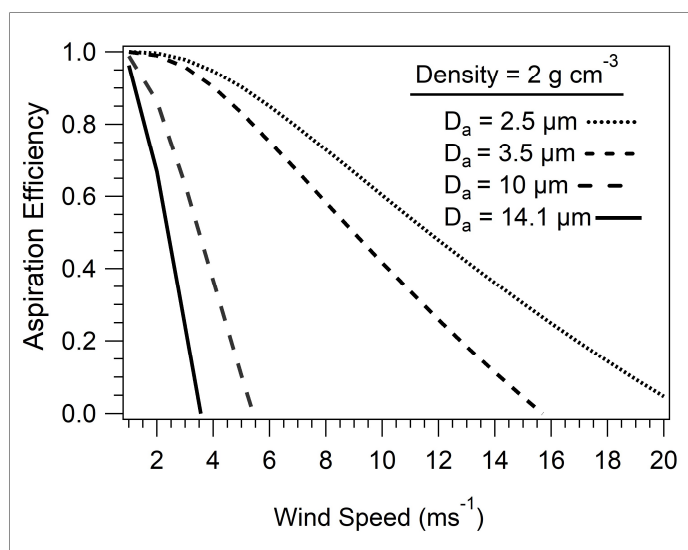


Figure 4.6: Aspiration efficiency of the Laser Egg (LE) inlet as a function of external wind speed ( $U_o$ ) for particles with varying aerodynamic diameters ( $D_a$ ) and a density of  $2 \text{ g cm}^{-3}$ .

In a field evaluation study of several low-cost  $\text{PM}_{2.5}$  sensors carried out in Atlanta (USA) and Hyderabad (India), Johnson et al. (2018) used a single 25mm DC fan (reported flow of 67 LPM) at the inlet of a box housing a single sensor, and three 25mm fans for another box housing five low-cost sensors. However, they did not evaluate the impact of fan flow rate and ambient wind speed on particle losses. As per the expression by Hangal and Willeke (1990), assuming similar fan dimensions, a particle of  $2.5\mu\text{m}$  aerodynamic diameter would still have an aspiration efficiency of 0.44 in the Johnson et al. (2018) inlet at high wind speeds of  $\sim 13 \text{ ms}^{-1}$  with a fan flow rate of 67 LPM. Still, in this study, which uses an accessory fan with a lower flow rate of 40 LPM, a particle of  $2.5\mu\text{m}$  diameter has an aspiration efficiency of only 0.18 at that wind speed.

#### 4.4.6. Impact of aerosol density on Laser Egg accuracy

Figure 4.7 shows the hourly averaged  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  measurements from the reference and the two LE sensors during two dust storms observed from May 3, 2016, 17:00:00 to May 4, 2016, 08:00:00 and May 4, 2016, 17:00:00 to May 5, 2016, 02:00:00, respectively, with a peak  $\text{PM}_{10}$  mass loading of  $\sim 450 \mu\text{g m}^{-3}$ . During the first storm, high dust loading due to long-range transport of desert dust coincided with low local wind speeds, while the second storm was characterized by higher wind speeds. The average wind speeds during the first dust storm ( $4\text{--}6 \text{ ms}^{-1}$ ) ensured that the optical system received only  $\text{PM}_{10}$  aerosol, as the aerodynamic cut-off of the modified LE inlet system is  $10\mu\text{m}$  (Figure 4.6). However, an overestimation of  $\text{PM}_{2.5}$  mass concentration and simultaneous underestimation of  $\text{PM}_{10}$  by the LE monitor is evident during the first dust storm under conditions with no confounding errors from wind speed or RH.



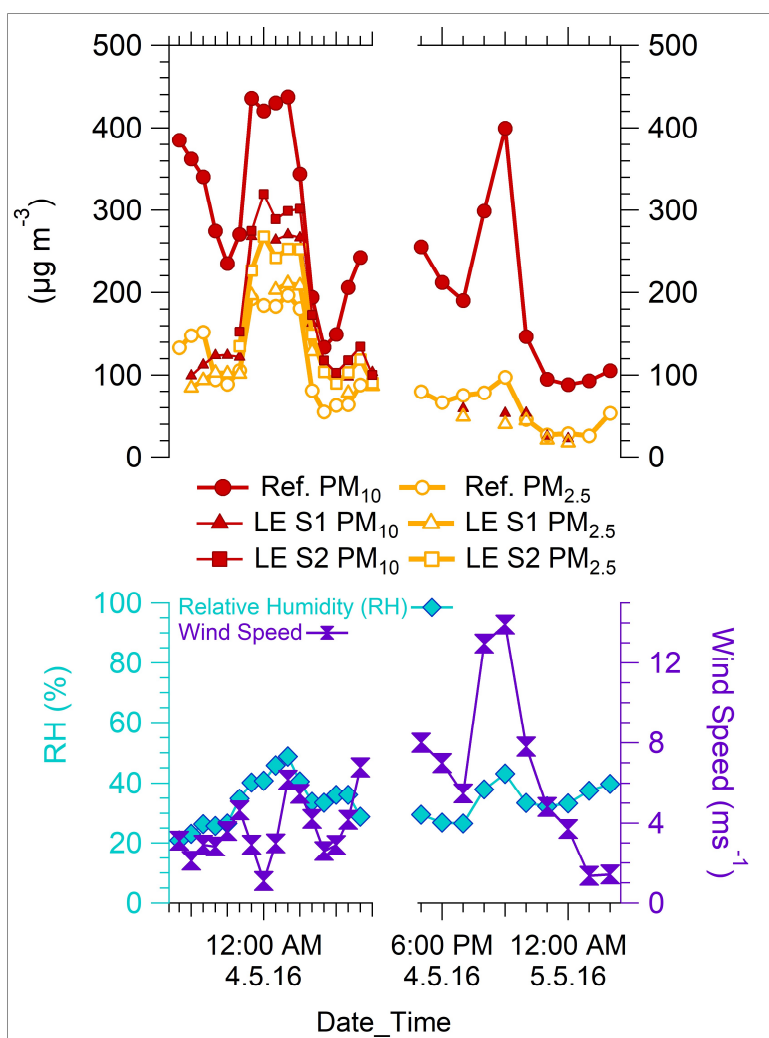


Figure 4.7: (Top)  $\text{PM}_{10}$  (solid markers) and  $\text{PM}_{2.5}$  (hollow markers) measurements from the reference and the two Laser Egg (LE) sensors during two dust storms observed from May 3, 2016, 17:00:00 to May 4, 2016, 08:00:00 and May 4, 2016, 17:00:00 to May 5, 2016, 02:00:00 respectively (Bottom) Relative Humidity (RH) and wind speed during the two dust storm periods.

It appears that the sensor algorithm used in the LE sensor assumes the optical equivalent diameter ( $d_o$ ) to be the same as the aerodynamic equivalent diameter ( $d_a$ ), and therefore ends up attributing a higher fraction of aerosol mass to the  $\text{PM}_{2.5(d_a)}$  size fraction than the reference analyzer, which cuts based on aerodynamic diameter. This error probably stems from the fact the low-cost sensors are generally factory calibrated using polystyrene spherical latex particles, which have a density of  $1.05 \text{ g cm}^{-3}$  (Mukherjee et al., 2017; Nakayama et al., 2018; Zhang et al., 2018). During pre-monsoon season, particles in size range  $0.7\text{-}2.5\mu\text{m}$  are dominated by clay minerals (Singh et al., 2004) with an effective density of  $\sim 2 \text{ g cm}^{-3}$ .

Assuming the optical diameter ( $d_o$ ) to be the same as the volume equivalent diameter ( $d_{veq}$ ), the aerodynamic diameter ( $d_a$ ) of a  $2.5\mu\text{m}$  clay particle is  $\sim 3.5\mu\text{m}$ , as per the following expression (Murphy et al., 2004; Chien et al., 2016)

$$d_a = d_o\sqrt{\rho} \quad (51)$$

where  $\rho$  is the particle density.

The LE measures  $\text{PM}_{3.5}$  but converts volume to mass using a density of  $1.05\text{ g cm}^{-3}$  instead of  $2\text{ g cm}^{-3}$ . The two errors partially cancel each other out, resulting in what appears to be a better accuracy for LE  $\text{PM}_{2.5}$  measurements under dry high dust conditions. Nevertheless, it can be seen from Figure 4.7 that at average wind speeds, the clay mode with an aerodynamic diameter of  $2.5\text{--}3.5\mu\text{m}$  and optical diameter of less than  $2.5\mu\text{m}$  results in a slight overestimation of the  $\text{PM}_{2.5}$  mass by the LE. At high wind speeds ( $>10\text{ ms}^{-1}$ ), the aerodynamic inlet cut-off of the modified LE drops below  $3.5\mu\text{m}$  and approaches  $2.5\mu\text{m}$  for the peak wind speeds of  $12\text{ ms}^{-1}$  observed during the second dust storm. Under these conditions, the LE underestimates the  $\text{PM}_{2.5}$  mass thanks to the lower density of  $1.05\text{ g cm}^{-3}$ , rather than  $2\text{ g cm}^{-3}$ , used to convert volume to mass, even though both the LE and the reference sensors measure  $\text{PM}_{2.5}$  at high wind speeds.

Serious underestimation of the  $\text{PM}_{10}$  mass under conditions with average wind speeds can also be explained by assuming that the LE monitor uses a sub-optimal value of aerosol density. The inlet cut-off at moderate wind speeds is close to  $d_a=10\mu\text{m}$ , and hence both the reference and the low-cost sensors measure  $\text{PM}_{10}$ . However, the LE converts volume to mass using a density of  $1.05\text{ g cm}^{-3}$  rather than  $2\text{ g cm}^{-3}$  and underestimates the  $\text{PM}_{10}$  mass.

#### **4.4.7. Impact of relative humidity on Laser Egg accuracy**

The reference PM analyzers use an inlet heating system to regulate the sample RH to  $\sim 40\%$  and, therefore, measure the dry mass of the aerosol. However, semi-volatile aerosol species such as ammonium nitrate and ammonium chloride, which are in a temperature- and RH-dependent equilibrium with the gas phase, can partition from the aerosol phase into the gas phase due to this heating. On the other hand, in the absence of humidity control in the LE monitor, the technique determines aerosol mass using the wet diameter of the aerosol, which can be up to three times larger than the dry diameter and varies as a function of RH. Therefore, a variation in the accuracy of the LE monitor's PM measurements with varying RH is expected because of the fundamental difference in both techniques.

Several field studies on low-cost optical sensors have reported an inherent positive bias in PM measurements with respect to FEM analyzers at high RH in the past (Badura et al., 2018; Crilley

et al., 2018; Jayaratne et al., 2018; Nakayama et al., 2018; Zheng et al., 2018; Levy Zamora et al., 2019; Liu et al., 2019a). Figures 4.4c and 4.5c show that as the  $RH^{Ref}$  increases to more than 80%, the LE monitor overestimates both  $PM_{10}$  and  $PM_{2.5}$  compared to the reference analyzers, indicating the impact of RH on the LE monitor's accuracy. Likewise, the markers representing low ambient  $RH^{Ref}$  conditions lie mostly below the solid (red) 1:1 line in the scatter plot.

#### 4.4.8. Correction for the biases introduced by aspiration losses, particle density, and relative humidity in Laser Egg PM measurements

The following corrections were applied successively to correct the LE monitors' PM measurements:

**Particle density correction:** The calibration period was split into summer (April 27 to June 15, 2016) and monsoon (June 16 to July 25, 2016) based on the levels of ambient RH. Assuming an average particle density of  $2 \text{ g cm}^{-3}$  and  $1.2 \text{ g cm}^{-3}$  during the summer and the monsoon, respectively, the raw  $PM_{10}$  and  $PM_{2.5}$  measurements were corrected as:

$$\text{During summer: Laser Egg PM (Density corrected)} = \frac{\text{Raw PM} \times 2 \text{ g cm}^{-3}}{1.05 \text{ g cm}^{-3}} \quad (52)$$

$$\text{During monsoon: Laser Egg PM (Density corrected)} = \frac{\text{Raw PM} \times 1.2 \text{ g cm}^{-3}}{1.05 \text{ g cm}^{-3}} \quad (53)$$

Where  $1.05 \text{ g cm}^{-3}$  represents the density of polystyrene latex spheres commonly used to calibrate optical sensors. Nakayama et al. (2018) used a factor of 1.3 to correct for the density of  $PM_{2.5}$  measurements from an optical sensor by assuming particle density between  $1.2$  and  $1.8 \text{ g cm}^{-3}$ . The singular density used for the correction during monsoon season ( $1.2 \text{ g cm}^{-3}$ ) represents a compromise between the best average density for the secondary material ( $\sim 1.5 \text{ g cm}^{-3}$ ) and the primary biological aerosol ( $1 \text{ g cm}^{-3}$ ) that can have a big impact on the volume to mass conversions during certain days in monsoon season. During summer, the coarse mode is usually present and dominated by dust ( $2\text{-}2.8 \text{ g cm}^{-3}$ ), so the average density has been chosen to represent the clay mode of dust ( $2 \text{ g cm}^{-3}$ ).

**Hygroscopic growth correction:** To correct for the positive bias in PM measurements of the LE under high RH, it was assumed that the ambient aerosol follows a growth curve identical to that of secondary organic aerosol formed from the oxidation of cyclopentene, which has a growth factor of 1.20 at 90% RH (Varutbangkul et al., 2006).

$$PM (RH \text{ corrected}) = \frac{PM (Density \text{ corrected})}{[1 + (1 - a_w)^{-0.4428} \times 0.0841 \times a_w^{1.2347}]^3} \quad (54)$$

$$\text{where } a_w = \frac{RH^{Ref}}{100} \quad (55)$$

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

A study carried over the Bay of Bengal (Boreddy et al. 2016) reported an average growth factor of 1.43 for aerosols collected over the northern Bay of Bengal during the winter season (27th December 2008 to 30th January 2009). Another study carried over Ahmedabad, India (Sarangi et al. 2019) during February 2018 reported an average growth factor of 1.39 for Aitken mode size ( $\leq 100$  nm) particles and 1.6 for accumulation mode size ( $>100$  nm (between 200 to 300nm)) particles at 90% RH. Given that these two sites are much more affected by proximity to the ocean (i.e., sea salt aerosol) than the inland site, it is safe to assume that the aerosol over the North West Indo-Gangetic Plain has lower growth factors than the ones reported in these two studies. However, the correction was determined experimentally by using a variety of growth factors and identifying the one that results in the largest improvements of the sensor performance.

**Correction for aspiration losses and incorrect cut-off diameter:** The interaction between external wind speeds and particle size distribution decreases the accuracy of the modified LE PM sensor. Figures 4.8a and 4.8b show the observed ratio of the LE PM (after density and RH correction) to the reference PM plotted as a function of external wind speeds for the summer season (April 27-June 15, 2016) during the calibration period. To correct the aspiration losses and incorrect cut-off diameter, I assume a bimodal aerosol volume size distribution (Figure 4.8c), which peaks at  $0.25 \mu\text{m}$  and  $2.5 \mu\text{m}$  optical diameter and has a uniform density of  $2\text{g cm}^{-3}$  during the summer season. This assumption is reasonable and represents a substantial clay burden present in the coarse mode aerosol during the pre-monsoon season. The assumed particle size distribution fits the observed  $\text{PM}_{2.5}/\text{PM}_{10}$  ratio of the reference analyzer during dust episodes. The equivalent optical cut-off diameters for the reference  $\text{PM}_{10 \text{ (da)}}$  and  $\text{PM}_{2.5 \text{ (da)}}$  analyzers are  $7.07 \mu\text{m}$  and  $1.77 \mu\text{m}$  for a particle density of  $2 \text{g cm}^{-3}$ . Thus, while the reference analyzers measure the area under the black curve until  $1.77 \mu\text{m}$  ( $\text{PM}_{2.5}$ ) and  $7.07 \mu\text{m}$  ( $\text{PM}_{10}$ ), the LE sensors misattribute extra scattering intensity from  $1.77 \mu\text{m}$  to  $2.5 \mu\text{m}$  as  $\text{PM}_{2.5}$  and from  $7.07 \mu\text{m}$  to  $10 \mu\text{m}$  as  $\text{PM}_{10}$ . Moreover, while the reference analyzer is immune to aspiration losses, the various colored curves represent the fraction of aerosol aspirated into the LE monitors at different wind speeds.

Thus, the ratio of the area under the curves of the assumed aerosol size distribution (Figure 4.8c) as measured by the LE monitor and reference analyzer in accordance with their respective cut-offs and the aspiration losses (wind speed) gives us a theoretical ratio shown by the maroon markers (Figure 4.8a and b). The LE PM measurements are then corrected for both aspiration-related losses and incorrect cut-off diameter as:

$$\text{PM corrected for aspiration loss and cut - off diameter} = \frac{\text{Density and RH corrected PM}}{f(\text{Wind speed})} \quad (56)$$

---

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

where  $f(\text{Wind speed})$  has been calculated after fitting a sigmoidal curve (maroon curve in Figure 4.8a and b) through the theoretical aspiration losses derived from the assumed aerosol size distribution (Figure 4.8c). The red and blue curves in Figures S4a and S4b show similar theoretical curves for aspiration losses derived from particle size distributions observed at Kanpur (Figure 4.8d) (Kaskaoutis et al., 2012) and Gual Pahari (Figure 4.8e) (Hyvärinen et al., 2011b) in India, while the yellow curve represents another theoretical correction curve derived from a different assumed low-dust size distribution shown in Figure 4.8f.

I performed a sensitivity analysis of the correction by repeating the above analysis using a lower ( $1.2 \text{ g cm}^{-3}$ ) and a higher value ( $2.7 \text{ g cm}^{-3}$ ) of aerosol density and by varying the position of the coarse mode in the size distributions. It can be seen from Figures 4.8a-f and the accompanying Table 4.5 that the exact position of the coarse mode barely impacts the correction for  $\text{PM}_{2.5}$  shown by the overlapping blue, red, and maroon curves in Figure 4.8b. The correction for  $\text{PM}_{2.5}$  in the presence of a significant coarse mode is impacted only by the assumed density, which widens or closes the gap between the  $d_o=2.5 \mu\text{m}$  and  $d_a=2.5 \mu\text{m}$  as can be seen in Figures 4.9b and 4.10b, which assume a density of  $1.2 \text{ g cm}^{-3}$  and  $2.7 \text{ g cm}^{-3}$  for the ambient aerosol. The correction for  $\text{PM}_{10}$  is overly sensitive to the exact position of the coarse mode peak, as can be seen by the correction curves derived from different size distributions in Figure 4.8a, limiting the usefulness of the LE monitor for measuring  $\text{PM}_{10}$ .

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

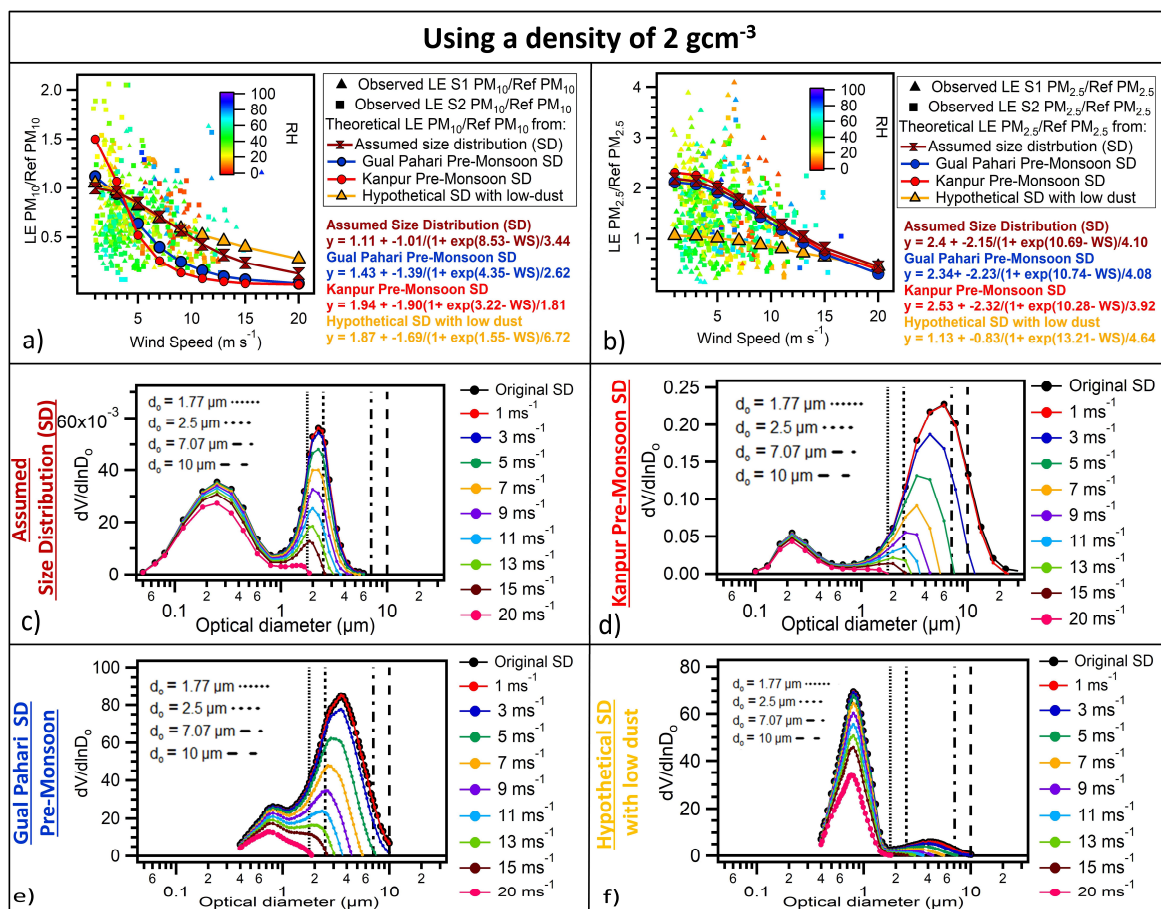


Figure 4.8: (a and b) Markers (color-coded according to ambient relative humidity) showing the observed ratio of Laser Egg (LE) PM (after density and RH correction) to reference PM versus wind speed. The maroon, red, blue, and yellow lines (in a and b) represent the theoretical ratio of LE PM to reference PM derived from the assumed aerosol volume size distribution for summer (April 27, 2016, to June 15, 2016) shown in (c), two observed size distributions from Kanpur (Kaskaoutis et al., 2012) (d) and Gual Pahari (Hyvärinen et al., 2011a) (e) and another assumed “low dust” distribution (f). A density of 2 g cm<sup>-3</sup> has been assumed for the ambient aerosol.

During the monsoon season, aerosol loading comprising primarily combustion-derived aerosol, primary and secondary organic aerosol, secondary inorganic aerosol, and a small coarse mode dominated by primary biological particles with an average density of 1.2 g cm<sup>-3</sup> can be represented using a hypothetical size distribution with low dust loading (Figure 4.9f). Aspiration losses occurring during the monsoon season were corrected using a similar algorithm as that for the summer season using the size distribution depicted in Figure 4.9f.

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

##### Using a density of $2 \text{ g cm}^{-3}$

PM <sub>10</sub>	Slope		r		nRMSE (%)		MBE ( $\mu\text{g m}^{-3}$ )	
	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2
Raw	0.62	0.61	0.43	0.41	73.6	72.3	-38.2	-44.9
Assumed Size Distribution (SD)	0.83	0.89	0.67	0.61	56.8	59.9	-20.0	-18.1
Gual Pahari Pre-Monsoon SD	1.37	1.29	0.48	0.52	117	98	0.5	4.5
Kanpur Pre-Monsoon SD	2.27	2.14	0.38	0.38	227	203	34.4	45.5
Hypothetical SD with low dust	0.78	0.85	0.68	0.61	58	62	-26.8	-25.6

PM <sub>2.5</sub>	Slope		r		nRMSE (%)		MBE ( $\mu\text{g m}^{-3}$ )	
	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2
Raw	1.13	1.23	0.53	0.51	69.7	75.5	6.6	11.2
Assumed Size Distribution (SD)	0.94	1.03	0.55	0.51	61.8	64.1	-1.9	1.3
Gual Pahari Pre-Monsoon SD	0.95	1.04	0.57	0.53	60.5	63.2	-1.4	2.2
Kanpur Pre-Monsoon SD	0.93	1.005	0.54	0.51	61.9	63.2	-2.6	0.5
Hypothetical SD with low dust	1.28	1.54	0.68	0.61	71.6	100.4	11.1	21.9

Table 4.5: Impact of using various size distributions and a density of  $2 \text{ g cm}^{-3}$  shown in Figure 4.8c-f for correction on the slope, correlation coefficient (r), normalized Root Mean Square Error (nRMSE), and Mean Bias Error (MBE) of the PM<sub>10</sub> and PM<sub>2.5</sub> measurements from the two Laser Egg (LE) sensors.

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

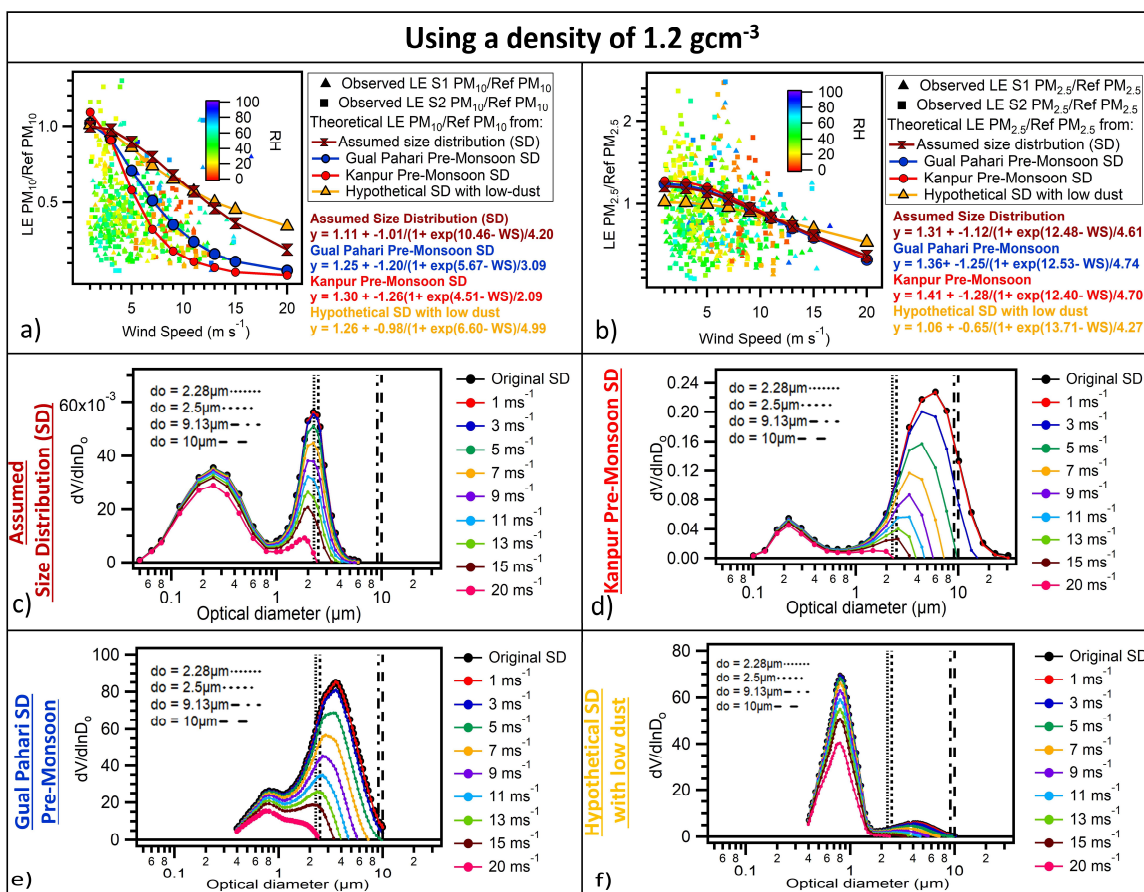


Figure 4.9: (a and b) Markers (color-coded according to ambient relative humidity) showing the observed ratio of Laser Egg (LE) PM (after density and RH correction) to reference PM versus wind speed. The maroon, red, blue, and yellow lines (in a and b) represent the theoretical ratio of LE PM to reference PM derived from the assumed aerosol volume size distribution for summer (April 27, 2016, to June 15, 2016) shown in (c), two observed size distributions from Kanpur (Kaskaoutis et al., 2012) (d) and Gual Pahari (Hyvärinen et al., 2011a) (e) and another assumed “low dust” distribution (f). A density of 1.2 g cm<sup>-3</sup> has been assumed for the ambient aerosol.



#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

##### Using a density of $1.2 \text{ g cm}^{-3}$

<b>PM<sub>10</sub></b>	<b>Slope</b>		<b>r</b>		<b>nRMSE (%)</b>		<b>MBE (<math>\mu\text{g m}^{-3}</math>)</b>	
	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2
Raw	0.62	0.61	0.43	0.41	73.6	72.3	-38.2	-44.9
Assumed Size Distribution (SD)	0.57	0.58	0.60	0.55	70	69	-43.3	-48.6
Gual Pahari Pre-Monsoon SD	0.73	0.71	0.57	0.57	68	64	-33.0	-37.0
Kanpur Pre-Monsoon SD	1.20	1.07	0.42	0.45	109	88	-13.7	-15.3
Hypothetical SD with low dust	0.57	0.59	0.61	0.56	69	68	-42.7	-47.7

<b>PM<sub>2.5</sub></b>	<b>Slope</b>		<b>r</b>		<b>nRMSE (%)</b>		<b>MBE (<math>\mu\text{g m}^{-3}</math>)</b>	
	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2
Raw	1.13	1.23	0.53	0.51	69.7	75.5	6.6	11.2
Assumed Size Distribution (SD)	0.96	1.06	0.58	0.55	59.5	63.2	-1.2	2.8
Gual Pahari Pre-Monsoon SD	0.95	1.05	0.58	0.54	59.8	63.0	-1.4	2.5
Kanpur Pre-Monsoon SD	0.84	0.95	0.61	0.57	56.7	58.5	-6.3	-2.7
Hypothetical SD with low dust	1.00	1.14	0.62	0.58	58.1	66.0	0.8	6.1

Table 4.6: Impact of using various size distributions and a density of  $1.2 \text{ g cm}^{-3}$  shown in Figure 4.9c-f for corrections on the slope, correlation coefficient (r), normalized Root Mean Square Error (nRMSE), and Mean Bias Error (MBE) of the PM<sub>10</sub> and PM<sub>2.5</sub> measurements from the two LE sensors.

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

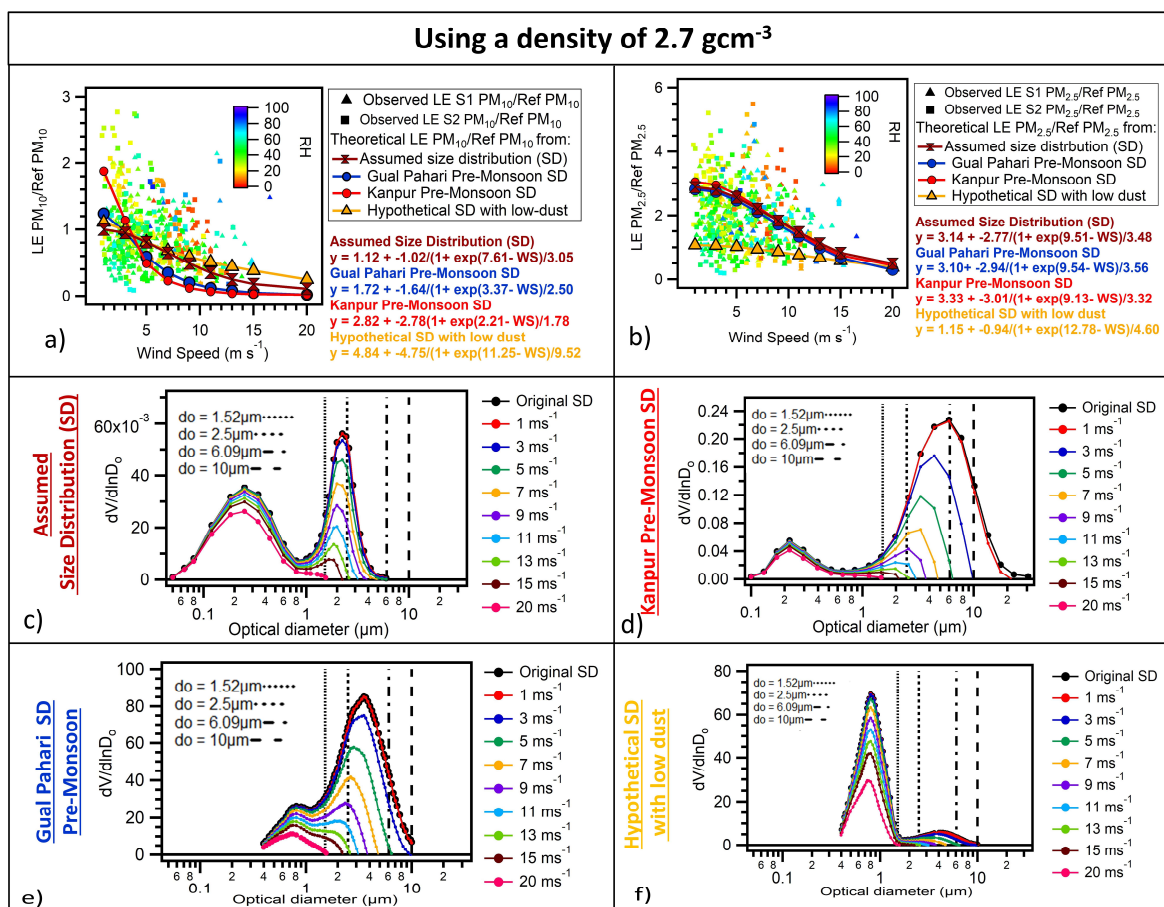


Figure 4.10: (a and b) Markers (color-coded according to ambient relative humidity) showing the observed ratio of Laser Egg (LE) PM (after density and RH correction) to reference PM versus wind speed. The maroon, red, blue, and yellow lines (in a and b) represent the theoretical ratio of LE PM to reference PM derived from the assumed aerosol volume size distribution for summer (April 27, 2016, to June 15, 2016) shown in (c), two observed size distributions from Kanpur (Kaskaoutis et al., 2012) (d) and Gual Pahari (Hyvärinen et al., 2011a) (e) and another assumed “low dust” distribution (f).

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

**Using a density of 2.7 g cm<sup>-3</sup>**

PM <sub>10</sub>	Slope		r		nRMSE (%)		MBE (µg m <sup>-3</sup> )	
	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2
Raw	0.62	0.61	0.43	0.41	73.6	72.3	-38.2	-44.9
Assumed Size Distribution (SD)	1.07	1.15	0.65	0.60	70	77	-8.2	-1.4
Gual Pahari Pre-Monsoon SD	1.62	1.64	0.52	0.51	136	132	18.1	29.8
Kanpur Pre-Monsoon SD	3.38	3.24	0.36	0.36	359	332	82.0	105.0
Hypothetical SD with low dust	0.96	1.07	0.69	0.61	62	72	-14.2	-8.2

PM <sub>2.5</sub>	Slope		r		nRMSE (%)		MBE (µg m <sup>-3</sup> )	
	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2	LE Sensor 1	LE Sensor 2
Raw	1.13	1.23	0.53	0.51	69.7	75.5	6.6	11.2
Assumed Size Distribution (SD)	0.95	1.04	0.56	0.52	60.8	63.4	-1.2	2.4
Gual Pahari Pre-Monsoon SD	0.97	1.07	0.57	0.53	61.0	64.0	-0.4	3.5
Kanpur Pre-Monsoon SD	0.95	1.03	0.55	0.51	61.9	63.6	-1.6	1.7
Hypothetical SD with low dust	1.57	1.96	0.66	0.60	102.5	146.1	20.9	36.8

Table 4.7: Impact of using various size distributions and a density of 2.7 g cm<sup>-3</sup> shown in Figure 4.10c-f for corrections on the slope, correlation coefficient (r), normalized Root Mean Square Error (nRMSE), and Mean Bias Error of the PM<sub>10</sub> and PM<sub>2.5</sub> measurements from the two LE sensors.

#### 4.4.9. Laser Egg PM measurements post-correction

Figure 4.11 shows the correlation plots of the LE PM versus the reference PM obtained after performing RMA regression for raw LE measurements (a and d), after RH and density correction (b and e), and after correcting for aspiration losses and cut-off diameter (c and f). The solid (red) line indicates the 1:1 line.

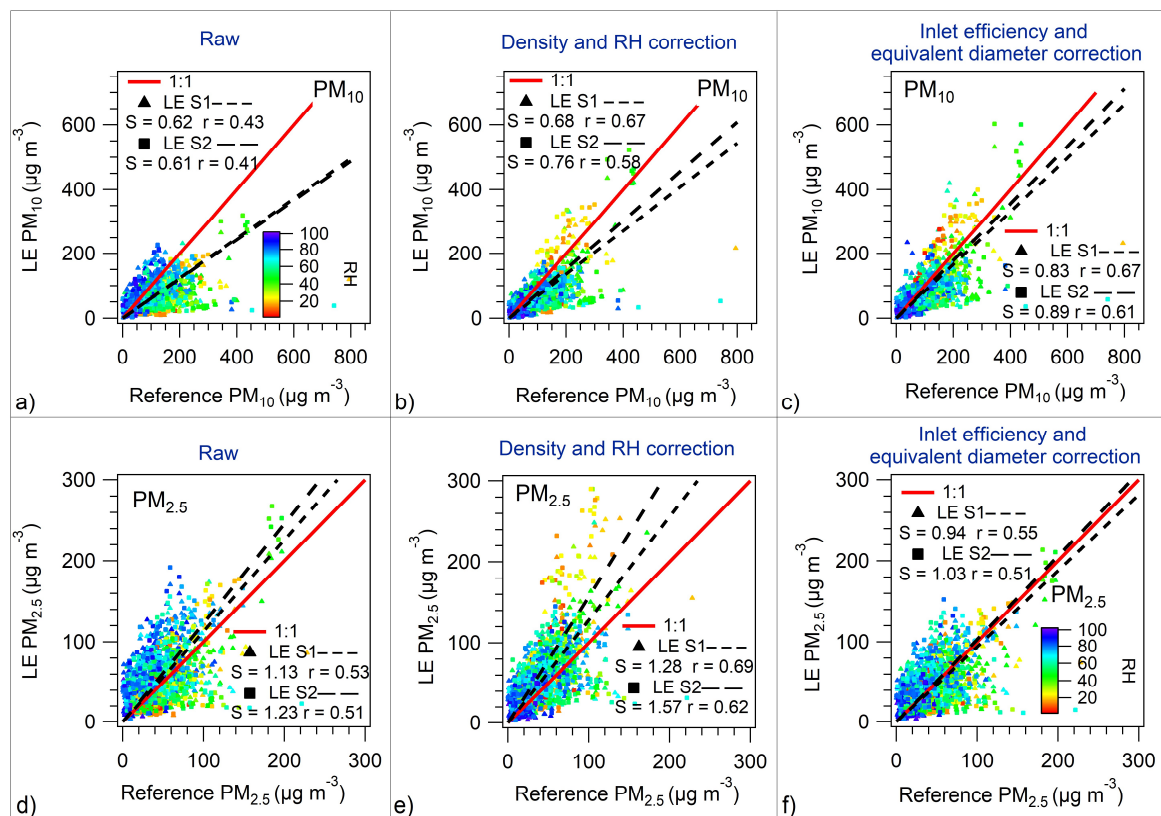


Figure 4.11: Correlation plots of the Laser Egg (LE) PM versus the reference PM obtained after performing reduced major axis (RMA) regression for raw LE measurements (a and d), after RH and density correction (b and e), and after correcting for aspiration losses and cut-off diameter (c and f). The solid (red) line indicates the 1:1 line. In each of the above plots, “S” refers to the slope of the best-fit line, and “r” refers to the Pearson correlation coefficient. The markers are color-coded according to ambient RH.

Figure 4.12 shows the time series of hourly averaged PM<sub>10</sub> and PM<sub>2.5</sub> measured by the reference and LE sensors for the period between April 27, 2016, and July 25, 2016, before and after correction. Corrections reduced the RMSE by  $\sim 22.8\%$  (from  $75.1 \mu\text{g m}^{-3}$  to  $58.0 \mu\text{g m}^{-3}$ ) and  $17.2\%$  (from  $84.3 \mu\text{g m}^{-3}$  to  $69.8 \mu\text{g m}^{-3}$ ) for PM<sub>10</sub> measurements from LE Sensor 1 and Sensor 2 respectively for the entire calibration period. The MBE in the LE monitor’s PM<sub>10</sub> measurements decreased from  $-38.2 \mu\text{g m}^{-3}$  and  $-44.9 \mu\text{g m}^{-3}$  to  $\sim -20.0 \mu\text{g m}^{-3}$  and  $\sim -18.1 \mu\text{g m}^{-3}$  for Sensor 1 and Sensor 2, respectively (Table 4.3).

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

The slope obtained from the RMA regression of the LE monitor's  $PM_{10}$  versus the reference  $PM_{10}$  increased from 0.6 (for raw measurements) to 0.83 and 0.89 for Sensor 1 and Sensor 2, respectively, after both corrections were applied. Simultaneously, the r-value increased from 0.43 and 0.41 to 0.67 and 0.61 (Figure 4.11), indicating an improvement in both linearity and accuracy. Specifically, during the summer period (April 27 to June 15, 2016) characterized by low RH and high dust loading (average  $PM_{10} = 149 \mu\text{g m}^{-3}$ ), the MBE in the LE monitor's  $PM_{10}$  measurements decreased from  $\sim -90 \mu\text{g m}^{-3}$  to  $-30.9 \mu\text{g m}^{-3}$  (Sensor 1) and  $-23.2 \mu\text{g m}^{-3}$  (Sensor 2) after correcting for density and aspiration losses (Table 4.3).

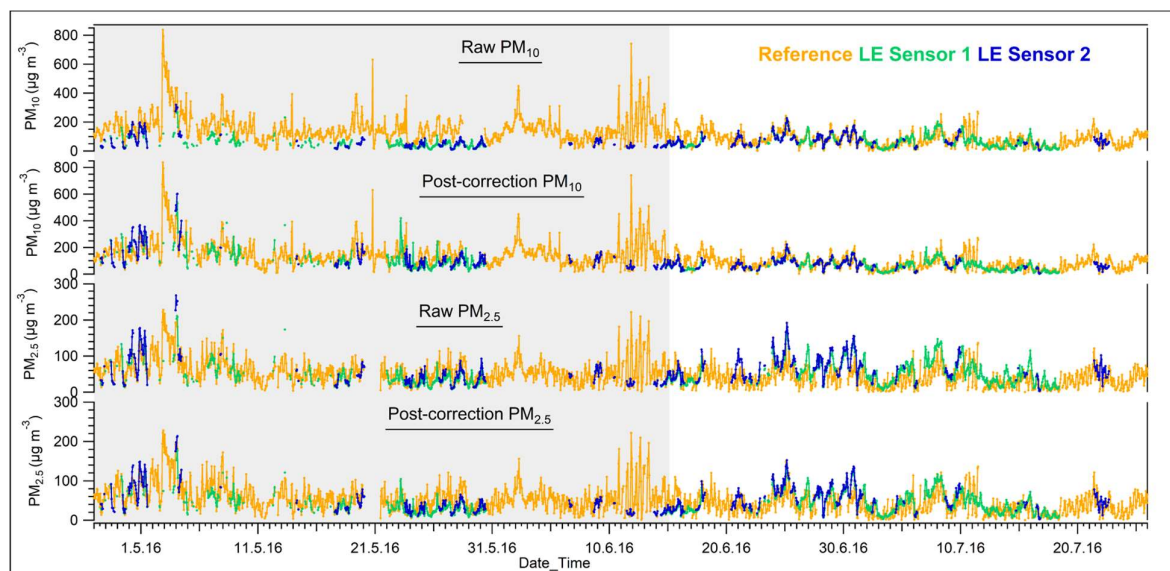


Figure 4.12: Time series of hourly averaged  $PM_{10}$  and  $PM_{2.5}$  mass concentration measured by the reference and Laser Egg (LE) sensors, pre- and post-correction for the calibration period from April 27, 2016, to July 25, 2016. The shaded portion represents the dry summer season (April 27, 2016 – June 15, 2016), while the unshaded portion represents the monsoon season (June 16, 2016 – July 25, 2016).

Overall, after the application of both corrections, the RMSE in the LE monitor's  $PM_{2.5}$  measurements decreased by  $\sim 11.3\%$  (from  $31.0 \mu\text{g m}^{-3}$  to  $27.5 \mu\text{g m}^{-3}$ ) and  $\sim 15\%$  (from  $37.9 \mu\text{g m}^{-3}$  to  $32.2 \mu\text{g m}^{-3}$ ) for Sensor 1 and Sensor 2, respectively. The slope obtained from the RMA regression of  $PM_{2.5}$  measurement from the LE monitor versus the reference decreased from 1.13 and 1.23 for raw measurements to 0.94 and 1.03 post-correction for LE Sensor 1 and Sensor 2, respectively (Figure 4.11). Specifically, during the monsoon period (June 16 to July 25, 2016) characterized by high RH, where the LE monitor tends to overestimate  $PM_{2.5}$  due to lack of inlet heating, the MBE in the LE monitor's  $PM_{2.5}$  measurements decreased from  $19.1 \mu\text{g m}^{-3}$  to  $8.7 \mu\text{g m}^{-3}$  and  $28.3 \mu\text{g m}^{-3}$  to  $16.5 \mu\text{g m}^{-3}$  for Sensor 1 and Sensor 2, respectively, after corrections (Table 4.3).

Overall, post-corrections, the accuracy of the LE monitor's PM measurements increased, indicated by a slope varying between 0.83 and 1.03 (Figure 4.11) obtained after performing RMA regression with the reference PM and a low MBE of  $< 20 \mu\text{g m}^{-3}$  ( $\text{PM}_{10}$ ) and  $< 3 \mu\text{g m}^{-3}$  ( $\text{PM}_{2.5}$ ) (Table 4.3). The modified version of the LE monitor could thus be used for ambient monitoring provided accurate RH and wind speed measurements and knowledge about the site's aerosol size distribution during high dust periods are available.

#### **4.4.10. Ambient temperature measurements**

The average ( $1\sigma$  ambient variability) of ambient temperature during the calibration period (April 27, 2016, to July 25, 2016) was  $31.5 (3.9) ^\circ\text{C}$ ,  $32.8 (7.0) ^\circ\text{C}$ , and  $35.1 (7.5) ^\circ\text{C}$  as measured by the reference Met One 064 Air Temperature sensor, and the LE Sensor 1 and Sensor 2 respectively (Table 4.2). Inter-comparison of the temperature measurements reveals high precision in the LE sensors ( $r = 0.99$ ,  $\text{CV} = 3.2\%$ ); however, they had an offset of  $\sim 1.8 ^\circ\text{C}$  among themselves (Table 4.4).

Figure 4.13a shows the diel plot of ambient temperature measurements obtained from the reference sensor (circular markers in orange) and the two LE monitors (green triangular and blue square markers) during the calibration period. The markers connected by lines indicate the hourly average. The shaded portion represents the ambient variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles. Qualitatively, the diurnal trend of temperature from the reference sensor and the LE monitors looks similar, characterized by a daytime maximum between 14:00 to 15:59 LT and minima between 05:00 to 05:59 LT. However, it is evident that the LE overestimates the ambient temperature by  $\sim 5$  to  $10 ^\circ\text{C}$  between 06:00 to 18:59 LT and underestimates the temperature by  $\sim 0.5$  to  $2 ^\circ\text{C}$  between 19:00 to 05:59 LT. In the LE monitor, the temperature sensor lies inside the body/chamber, making it susceptible to heat given off by the circuitry, which the manufacturer blames for the discrepancies between the reference and the LE temperature measurements. However, it seems that the overestimation in temperature is due to the absorption of SR by the LE monitor's casing. This hypothesis is validated by looking at a time series plot of hourly averaged temperature measurements from the reference, and LE sensors plotted along with SR on the secondary axis (Figure 4.13b) from July 13, 2016, to July 15, 2016. July 13, 2016, being a clear day with maximum daytime SR  $\sim 700 \text{ W m}^{-2}$ , the LE monitor overestimated the daytime temperature by  $\sim 9^\circ\text{C}$ , while on July 14, 2016, an overcast day with average daytime SR of only  $\sim 300 \text{ W m}^{-2}$ , the difference between the LE monitor and the reference temperature sensor decreased to less than  $2 ^\circ\text{C}$ . July 15, 2016, was a partially overcast day, and as the average SR increased until 11:00 LT, the sensors overestimated temperature by  $\sim 6$  to  $8 ^\circ\text{C}$ . However, with a sharp dip in SR from 550

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

$\text{W m}^{-2}$  to  $180 \text{ W m}^{-2}$  because of cloud cover between 11:00 to 12:00 LT, the LE sensors simultaneously cooled down, and the deviation in temperature decreased to  $\sim 2 \text{ }^\circ\text{C}$ . After 12:00 LT, as the SR increased, the LE temperature simultaneously heated up again.

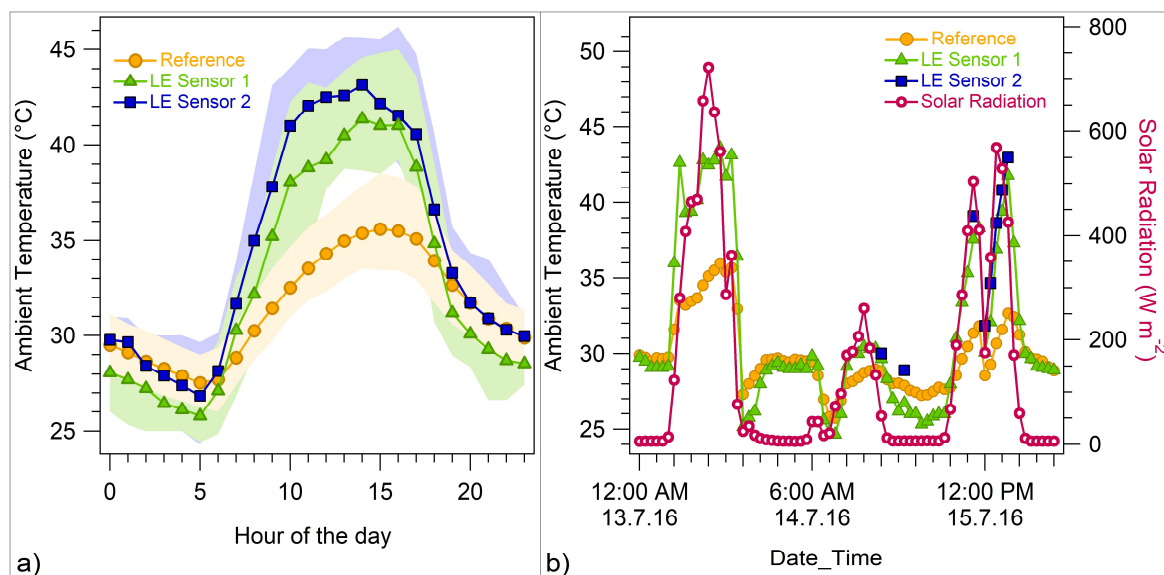


Figure 4.13: a) Diel plot of ambient temperature from the reference analyzer (circular markers in orange) and two Laser Egg (LE) sensors (green triangular and blue square markers, respectively) for the calibration period. The markers connected by lines indicate the hourly averages, and the shaded portion represents the ambient variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles. b) Time series plot of hourly averaged ambient temperature from the reference and the LE sensor plotted along with solar radiation measurements (hollow circular markers) from July 13, 2016, to July 15, 2017.

A field study carried out by SCAQMD (SCAQMD, 2019) in California, United States, reported better accuracy and an r-value of 0.98 between temperature measurements obtained from the Laser Egg 2+ (a new model of LE) and the SCAQMD meteorological station when the ambient temperature varied between  $\sim 2^\circ\text{C}$  and  $35^\circ\text{C}$ . This behaviour confirms that the absorption of SR by the modified casing of the LE monitor is the primary cause for higher-than-average daytime temperature measurements.

Figure 4.14 shows that the temperature deviation of the LE sensor can be explained as a function of SR during the day and ambient RH (a proxy for evaporative cooling) at night. From sunrise until noon (06:00 to 12:00 LT), as the solar radiation increases from  $100 \text{ W m}^{-2}$  to  $600 \text{ W m}^{-2}$ , the deviation in average LE temperature from the reference temperature increases from  $\sim 0 \text{ }^\circ\text{C}$  (at 06:00 LT) to  $8.5 \text{ }^\circ\text{C}$  (12:00 LT) for both sensors. From 13:00 LT until 18:59 LT, as the average solar radiation declines from  $\sim 500 \text{ W m}^{-2}$  to  $\sim 80 \text{ W m}^{-2}$  (Figure 4.14a), the sensors rapidly begin to cool off, and their deviation decreases from  $\sim 7 \text{ }^\circ\text{C}$  (at 13:00 LT) to  $\sim 2 \text{ }^\circ\text{C}$  (at 18:00 LT).

From 19:00 LT until 05:59 LT, the net available solar radiation is less than  $20 \text{ W m}^{-2}$ , and the LE sensors' temperature drops below the reference ambient temperature (Figure 4.14b). With a dip in ambient temperature and rise in relative humidity from 50 % (at 19:00 LT) to ~66 % (at 05:00 LT), the difference in the temperature reported by LE sensors and reference sensor becomes negative ( $-0.5 \text{ }^{\circ}\text{C}$  to  $-2 \text{ }^{\circ}\text{C}$ ) indicating that evaporative cooling due to evaporation of the aerosol liquid water may be taking place within the LE.

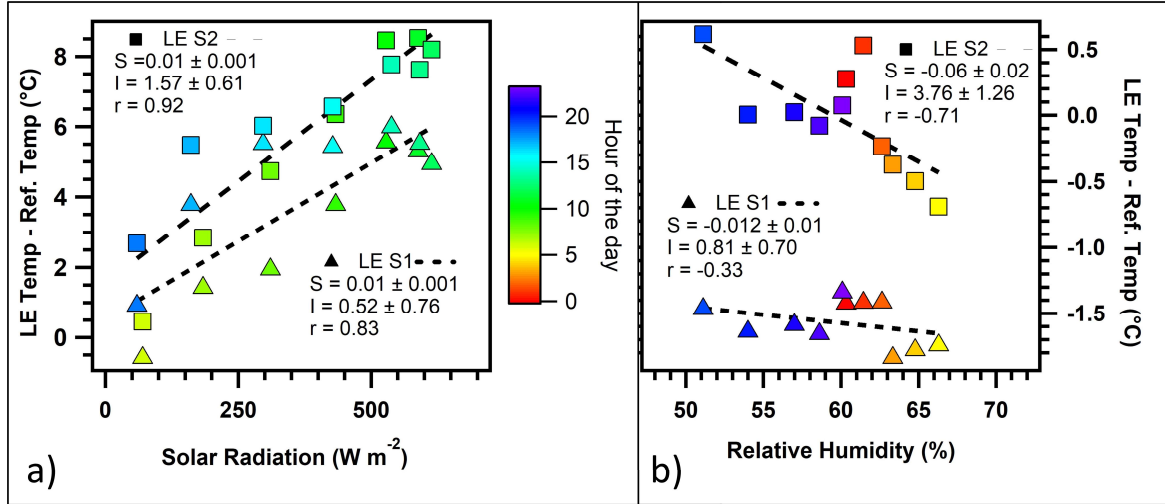


Figure 4.14: The deviation in Laser Egg (LE) temperature from the reference (Ref.) ambient temperature sensor explained as a function of (a) Solar Radiation (SR) during the day [06:00 to 18:59 LT] and as a function of (b) Relative Humidity (RH) during the night [19:00 to 05:59 LT]. Markers, color-coded according to the hour of the day, represent the averaged values of hourly bins derived using the measurements from the three-month calibration period (April 27, 2016, till July 25, 2016). In each of the above plots, “I” refers to the intercept, “S” refers to the slope of the best – fit line, and “r” refers to the Pearson correlation coefficient obtained from the ordinary linear regression of the deviation in temperature versus SR (Figure 4.14a) and RH (Figure 4.14b).

Provided accurate measurements of SR and  $\text{RH}^{\text{Ref}}$  from co-located sensors are available, the biased temperature measurements  $T^{\text{LES1}}$ ,  $T^{\text{LES2}}$  from the LE Sensor 1 and Sensor 2, respectively, can be corrected using the following multiple linear regression equations:

$$T^{\text{Ref}} = 0.43 \times T^{\text{LES1}} - 0.003 \times \text{SR} - 0.064 \times \text{RH}^{\text{Ref}} + 21.50 \quad (r_{\text{adj}}^2 = 0.83) \quad (57)$$

$$T^{\text{Ref}} = 0.47 \times T^{\text{LES2}} - 0.003 \times \text{SR} - 0.049 \times \text{RH}^{\text{Ref}} + 18.77 \quad (r_{\text{adj}}^2 = 0.84) \quad (58)$$

#### 4.4.11. Relative humidity measurements

Figure 4.15a and b show the hourly averaged time series and diel profile of the RH measurements from the reference Met One 083E RH sensor and the two LE sensors. During the calibration period, the average RH ( $1\sigma$  ambient variability) was 55.0 (22.4) % from the reference and 23.4 (3.0) % and 27.0 (3.4) % from the two LE sensors, respectively (Table 4.2). The LE underestimated RH by more than 50%. OLS of RH from the LE sensors versus the reference reveals only a moderate



#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

correlation, with an r-value of 0.56 and 0.66 (Table 4.4). Inter-comparison of the RH measurements from the two LE sensors shows moderate precision, with CV = 11.8% (Table 4.4). The SCAQMD field study (SCAQMD, 2019) reported better accuracy and an r-value of 0.99 between RH measurements obtained from the Laser Egg 2+ and the SCAQMD meteorological station when the ambient RH varied between ~15% to 100 %.

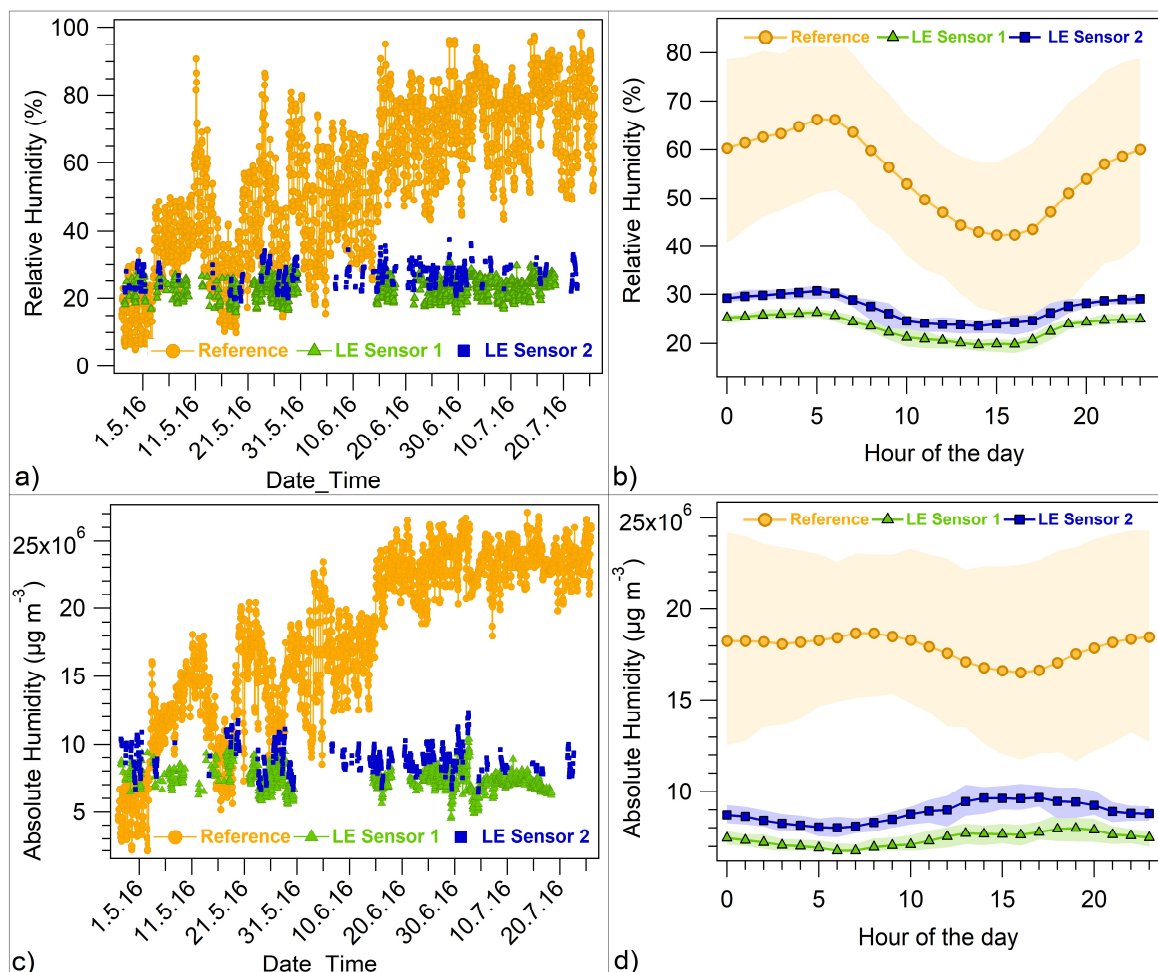


Figure 4.15: Time series plot of hourly averaged (a) relative humidity (RH) and (c) absolute humidity reported by reference RH sensor (orange circular markers) and two Laser Egg (LE) sensors (green triangular and blue square markers, respectively) for the calibration period from April 27, 2016, until July 25, 2016. A diel plot of (b) relative humidity and (d) absolute humidity for the calibration period. The markers connected by lines indicate the hourly averages, while the shaded portion represents ambient variability as 75<sup>th</sup> and 25<sup>th</sup> percentiles.

To ascertain if the LE sensors' inaccurate RH levels were caused by their inaccurate temperature measurements, I converted the  $RH^{Ref}$  and  $RH^{LE}$  to corresponding  $AH^{Ref}$  and  $AH^{LE}$  using the accurate reference temperature ( $T^{Ref}$ ) in both cases and compared the two.

$$AH^{Ref} (g m^{-3}) = \frac{e_s \times RH^{Ref} \times 18 g mol^{-1}}{100 \times (8.314 J K^{-1} mol^{-1} \times (T^{Ref} + 273.15))} \quad (59)$$

$$AH^{LE} (g m^{-3}) = \frac{e_s \times RH^{LE} \times 18 g mol^{-1}}{100 \times (8.314 J K^{-1} mol^{-1} \times (T^{Ref} + 273.15))} \quad (60)$$

where  $e_s$  is the saturation vapor pressure estimated as per the expression by Alduchov and Eskridge (1996)

$$e_s = 610.94 \times \exp\left(\frac{17.625 \times T^{Ref}}{243.04 + T^{Ref}}\right) \text{ Pa} \quad (61)$$

Figures 4.15c and 4.15d show the hourly averaged time series and diel profile of the AH estimated from the reference RH sensor and LE monitors, respectively. While the reference analyzer reported a wide variation in AH from 5 to 25 g m<sup>-3</sup>, the AH from the LE monitors varied only between 2 and 12 g m<sup>-3</sup> during the calibration period. It is again evident that the LE underestimated AH by almost 50% (Table 4.2) and did not capture its ambient variability. This suggests that the bias in the LE monitor's RH sensor was independent of its biased temperature sensor. The RH sensor's placement inside the outer casing could be responsible for the lower accuracy of RH measurements. It is possible that the outer casing not only absorbs solar radiation but is also hygroscopic and acts as a dryer, as the LE Sensor 1 with the lower measured AH (Figure 4.15c) also showed a lower overestimation of the measured PM<sub>2.5</sub> due to aerosol hygroscopic growth (Figure 4.11d). OLS between the reference AH and the LE monitor AH revealed a poor anti-correlation, with an r-value of -0.31 and -0.23 (Table 4.4).

As such, RH measurements from the LE do not quantitatively correspond to the ambient RH levels.

## 4.5. Conclusions

In this chapter, I focus on the field calibration of a modified version of LE, a commercially available low-cost PM sensor that works on the principle of laser-based light scattering. Two identical LE monitors with modified casings to shield them from sunlight and rain, and accessory inlet fans were co-located next to the inlets of PM<sub>10</sub> and PM<sub>2.5</sub>  $\beta$ -attenuation reference analyzers at the IISER Mohali Atmospheric Chemistry Facility, a suburban site in the NW-IGP during a 3-month period from April 27, 2016, to July 25, 2016, and their performance was evaluated for regular ambient usage.

---

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

1. Measurements of PM<sub>10</sub>, PM<sub>2.5</sub>, ambient temperature, and RH from the LE monitor were always precise, with  $r > 0.9$  and  $CV < 12\%$  for each inter-species correlation among the two sensors.
2. The average ( $1\sigma$  ambient variability) of PM<sub>10</sub> during the calibration period was 121.0 (87.1)  $\mu\text{g m}^{-3}$ , 62.8 (41.3)  $\mu\text{g m}^{-3}$ , and 69.8 (45.2)  $\mu\text{g m}^{-3}$  as measured by the reference analyzer, and the LE Sensor 1 and Sensor 2 respectively. High nRMSE ( $\sim 73\%$  for both sensors) and an MBE of  $-38.2 \mu\text{g m}^{-3}$  (Sensor 1) and  $-44.9 \mu\text{g m}^{-3}$  (Sensor 2) indicated low accuracy and underestimation of PM<sub>10</sub> measurements by the LE monitor.
3. This study highlights the effects of hygroscopic growth, aerosol density, aspiration losses of particles at high wind speeds, and the implications of assuming optical equivalent diameter to be the same as aerodynamic diameter on the accuracy of PM measurements acquired from low-cost sensors. While the study also presents methods to correct the biases mentioned above, the end goal should be to ensure that the low-cost sensor reports accurate PM measurements to the targeted end-user rather than requiring complicated post-processing. Assessing personal exposure to PM using a sensor that underestimates ambient PM levels under certain conditions can also have a detrimental effect on people's health. For instance, in the dust storms (average wind speeds of  $9 \text{ ms}^{-1}$ , solar radiation of  $400 \text{ Wm}^{-2}$  and reference PM<sub>10</sub>  $> 200 \mu\text{g m}^{-3}$ ) observed during daytime in the calibration period, the average LE PM<sub>10</sub> levels ( $59 \mu\text{g m}^{-3}$ ) were about five times lower than those of the reference PM<sub>10</sub> analyzer ( $278 \mu\text{g m}^{-3}$ ) thus creating a false sense of air being clean to a person relying on LE for personal air quality monitoring. This false perception of clean and fresh air may be further amplified by dips in temperature and increases in wind speed commonly observed during dust storms, possibly prompting the user to schedule outdoor activity during such hours. Hence future development of low-cost PM sensors must focus on an inlet design that minimizes aspiration losses, better placement of the RH and temperature sensor outside the casing to avoid artifacts and using the sensor location and additional information (e.g., from nearby pollution monitoring stations with data in the public domain) to choose the right data processing algorithm in terms of the appropriate density for volume-to-mass conversion.
4. Overall, post-corrections, the accuracy of the LE monitors' PM measurements increased, indicated by a slope varying between 0.83 and 1.03 obtained after performing RMA regression with the reference PM and a low MBE of  $< 20 \mu\text{g m}^{-3}$  for PM<sub>10</sub> and  $< 3 \mu\text{g m}^{-3}$  for PM<sub>2.5</sub>. The modified version of the LE could thus be used for ambient monitoring provided accurate RH and wind speed measurements and knowledge about the site's

---

#### 4. Correcting PM measurements of the Laser Egg low-cost PM sensor

- aerosol size distribution during high dust periods are available. However, the correction for aspiration losses of PM<sub>10</sub> is highly problematic and should be avoided by better inlet design.
5. Installing an exhaust fan with a higher flow rate, as has been done by Badura et al. (2018) and Holstius et al. (2014), could reduce aspiration efficiency-related artifacts.
  6. To further improve the accuracy of the low-cost sensors, a size-selective inlet specifically for PM<sub>10</sub> or PM<sub>2.5</sub> could be incorporated.
  7. Temperature and RH measurements from the LE sensors had several issues. Absorption of solar radiation by the casing of the LE caused an overestimation of temperature (by ~5 °C to 10 °C) during the daytime, while evaporative cooling at night resulted in temperature readings that were lower than the reference ambient temperature by 0.5 °C to 2 °C.
  8. Both the sensors grossly underestimated RH and AH by almost 50% throughout the calibration period. This and the fact that evaporative cooling was observed indicates that the air inside the LE is drier than the ambient air. Dry air inside the LE is not undesirable. In fact, further drying could reduce moisture-based measurement artifacts in the LE. However, the drying seems to be an unintended side effect of the nature of the polymer chosen as casing material and is not equal for both sensors used in this study. Inlet heating may provide a better tool for accurate RH control and may be relatively easy to implement, even in low-cost sensors.
  9. Even though the monitors were continuously charged and there was no power shortage, the data availability was only 48% and 28% for Sensors 1 and 2, respectively. The major issue was that the internet connection from Wi-Fi would self-terminate and had to be reset several times a day. On other occasions, the LE sensors failed to recognize Wi-Fi signals. Both these issues need to be taken care of to ensure continuous monitoring. The LE should incorporate a secondary data storage option to minimize data losses.
  10. Overall, the results of this study could help design better, more accurate PM sensors. In the future, to effectively utilize low-cost sensors for increasing the spatio-temporal resolution of PM measurements, concerted efforts that go beyond analyzing  $r^2$  values and instead delve into reasons behind observed inaccuracies are urgently needed.

# Chapter 5

---

## **Particulate matter measurements from rural sites in North-West India reveal that local fuel choices and residue management have a disproportionate impact on ambient air quality**

### **5.1. Abstract**

November onwards, the poor air quality over northwest India is blamed on the large-scale paddy residue burning in Punjab and Haryana. However, the emission strength of this source remains poorly constrained due to the lack of ground-based measurements over rural areas and issues in the satellite detection of paddy residue fires. In this chapter, I report the first-ever particulate matter (PM) measurements at Nadampur, a village in the Sangrur district with the highest reported paddy residue fires, from 1 October to 19 December 2019, using the Airveda low-cost PM sensors. The daily average PM<sub>10</sub> and PM<sub>2.5</sub> mass concentration at Nadampur correlated well with the daily sum of Visible Infrared Imaging Radiometer Suite (VIIRS) fire counts ( $r > 0.7$ ) in a 50 km × 50 km area surrounding the village. Agreement of the Coefficient of Emissions (Ce) estimated in this chapter ( $0.038 \text{ kg MJ}^{-1}$ ) with the reported value ( $0.04 \text{ kg MJ}^{-1}$ ), and a disagreement of the top-down estimate of PM emission factors with the laboratory reported values indicates an underdetection of paddy residue fires. Residential burning of solid fuels such as cow-dung cakes and fuelwood for space heating triggered by a dip in the temperature led to poor air quality from 20 November onwards. Source apportionment performed using Multiple Linear Regression (MLR) and Positive Matrix Factorization (PMF) revealed that agricultural activities (the sum of paddy harvest and paddy residue burning from 20 October-19 November 2019) and residential heating (from 20 November to 19 December 2019) increased the PM<sub>10</sub> (PM<sub>2.5</sub>) at Nadampur by  $141.7 \pm 38.3 \mu\text{g m}^{-3}$  ( $73.5 \pm 11.6 \mu\text{g m}^{-3}$ ) and  $151.2 \pm 47.2 \mu\text{g m}^{-3}$  ( $120.1 \pm 8.8 \mu\text{g m}^{-3}$ ), respectively. Unlike agricultural activities, which typically affect the air quality for roughly one month, heating-related emissions profoundly impact the air quality for multiple months.

## **5.2. Introduction**

November onwards, New Delhi and the entire northwest Indo-Gangetic Plain (NW-IGP) experience severe air pollution episodes that persist throughout the winter season (December to February) (Aneja et al., 2001; Kandlikar, 2007; Guttikunda and Gurjar, 2012; Mandal et al., 2014; Tiwari et al., 2014; Pant et al., 2015; Bikkina et al., 2019). The onset of air quality deterioration coincides with the beginning of large-scale paddy residue burning in the NW-IGP (Chandra and Sinha, 2016). As a result, paddy residue burning is blamed for the poor air quality in the National Capital Region (NCR) during post-monsoon (October to November) and winter seasons (Nagar et al., 2017), even though this activity typically subsides by the end of November.

Limited studies have quantified the contribution of paddy residue burning to the Particulate Matter (PM) in the NCR. Kulkarni et al. (2020b) attributed 20% of the post-monsoon PM<sub>2.5</sub> in Delhi to paddy residue burning using the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) and FINNv1.5 fire emission inventory, while Cusworth et al. (2018) estimated the contribution to be 25% using the Stochastic Time-Inverted Lagrangian Transport (STILT) model. Both studies raised the issue of a potentially severe underestimation of the impact of paddy residue burning on PM levels in the NCR region due to the failure of satellites to detect fire counts of small fires with short durations (Liu et al., 2019b). However, when STILT is used as a tool to estimate the impact of a faraway source on a downwind site, an overestimation when high fire counts, Diwali festival, and the onset of shallow nighttime inversion layers coincide is equally plausible. Both studies hypothesized that their estimate provides a lower boundary for the true impact of paddy residue burning on the air quality in the NW-IGP. However, this hypothesis has not been tested due to a lack of ground-level PM measurements in rural sites, which are hardly impacted by other sources.

Due to low ambient temperatures in the winter season, people with limited resources tend to burn solid fuels such as cow dung cakes, firewood, dry waste, leaf litter, and agricultural residue to keep warm (Chapter 3). Increased biomass burning for heating purposes has been identified as a significant contributor to the wintertime VOC burden, even in urban areas (Hakkim et al., 2019). Bikkina et al. (2019) reported that ~42% of post-monsoon and wintertime black carbon in New Delhi could be attributed to a combination of paddy-residue burning and wood burning in rural areas. However, no study has quantified the impact of increased wintertime residential heating emissions on the rural air quality.

This chapter investigates the trends in the PM levels across four sites with variable socioeconomic levels and land-use categories in Punjab from 1 October to 19 December 2019,

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM covering the post-monsoon paddy residue burning and part of the winter season. This chapter reports the first-ever dataset of PM measurements over Nadampur, a rural village in Punjab, using the Airveda low-cost PM sensor. The satellite detection of short-term paddy residue fires over this region was constrained by combining ground-based PM measurements over Nadampur with the Fire Radiative Power (FRP) detected by the Visible Infrared Imaging Radiometer Suite (VIIRS) sensor. Simultaneously, the unique observations of PM enhancements in a village within the source region provide an upper boundary to the PM enhancements at a downwind site, which can be blamed on the paddy residue burning in Punjab. I also quantify the effects of increased residential burning of solid fuels caused by a dip in the temperature on the ambient PM levels. Finally, the contribution of agricultural sources (paddy harvest, local and regional paddy residue burning) and residential heating emissions to the overall PM mass loadings was quantified using two independent tools. This chapter is the first to quantify the magnitude of PM enhancement due to increased heating demand in rural and urban environments.

### 5.3. Materials and Methods

#### 5.3.1. Site description and instrumental details for the field calibration

The variations in the PM<sub>10</sub> and PM<sub>2.5</sub> levels were investigated across four sites in Punjab (Nadampur, Kalal Majra, Mohali, and Patiala) with varying socioeconomic levels and land-use categories from 1 October 2019 to 19 December 2019 (Figure 5.1).

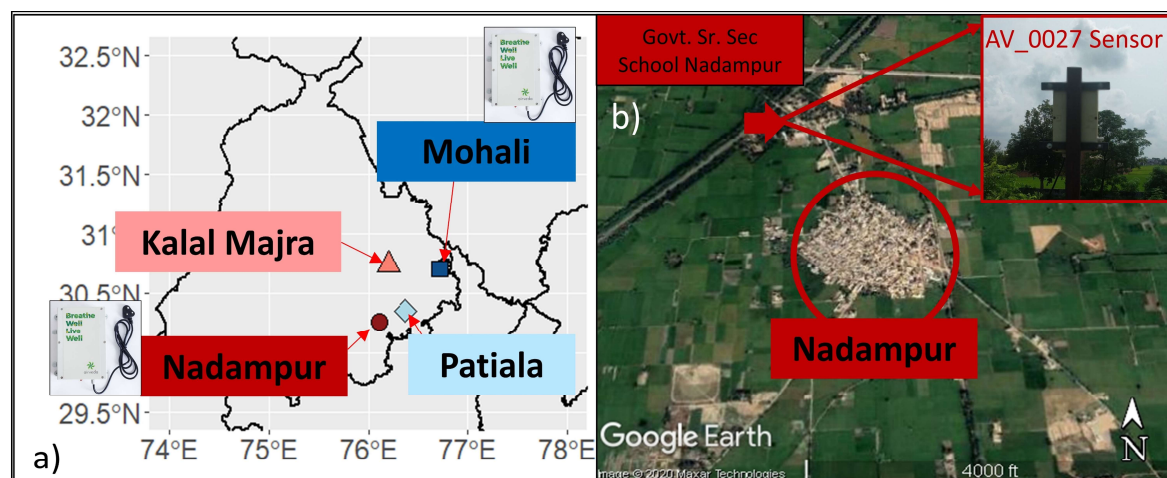


Figure 5.1: a) Map showing the locations of the Nadampur, Mohali, Patiala, and Kalal Majra in the state of Punjab, India. (b) Google Earth imagery (captured on 3 October 2020, © Google Earth and Maxar Technologies) of Nadampur village showing agricultural fields in all directions and the Airveda AV\_0027 sensor installed at the Government Senior Secondary School, Nadampur.

Nadampur is a village in the Sangrur district of Punjab, which had the highest district-wise crop production and agricultural residue fires in Punjab in 2017-18 (Singh et al., 2020). The village has

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM an area of 707 hectares and a population of 2772 as per the 2011 Census of India [*Census of India*, 2011]. It is surrounded by agricultural fields in all directions. The closest cities, Patiala and Sangrur, lie at 30 km and 25 km in the northeast and west directions.

Mohali is an urban city in the SAS Nagar district of Punjab, which had the lowest district-wise crop production and lowest number of agricultural residue fires in 2017-18 (Singh et al., 2020).

PM levels at Nadampur and Mohali were monitored using the outdoor Airveda low-cost PM sensors, which convert the light scattered by airborne particles into the mass concentration of PM<sub>10</sub> and PM<sub>2.5</sub> in real-time using proprietary algorithms. A detailed description of the Airveda sensor can be found in Section 5.3.2. Airveda sensor units, AV\_0027 and AV\_0013, were installed at the Government Senior Secondary School in Nadampur (30.254 °N, 76.117 °E, Figure 5.1b) and the Indian Institute of Science Education and Research (IISER) in Mohali (30.667 °N, 76.729 °E), respectively, from 1 October to 19 December 2019.

Kalal Majra is a village in the Ludhiana district of Punjab, which had the fourth-highest district-wise crop production out of 22 districts in Punjab (Singh et al., 2020). The village is surrounded by several industries (cattle feed, fabrication, flour mill, frozen food) and an educational institute. Kalal Majra is 40 km from the industrial city and district headquarters Ludhiana. Kalal Majra has an area of 306 hectares and a population of 1874 as per the 2011 Census of India [*Census of India*, 2011].

Patiala, an urban city in south-eastern Punjab, is the administrative capital of the Patiala district. The city is one of the leading trading and industrial hubs of Punjab. Agriculture is widely practiced in the surrounding regions of Patiala city.

Ambient PM levels at Kalal Majra and Patiala were monitored by the Punjab Pollution Control Board (PPCB) using reference-grade instruments. PM<sub>10</sub> and PM<sub>2.5</sub> measurements from 1 October to 19 December 2019 were obtained from the Continuous Ambient Air Quality Monitoring Station (<https://app.cpcbcr.com/ccr/#/caaqm-dashboard-all/caaqm-landing>), a publicly accessible data repository maintained by the Central Pollution Control Board.

For a more comprehensive analysis of the sources at the four sites, the study period (1 October to 19 December 2019) was split into two intervals based on the daily average temperature: a non-heating and a heating period. The interval from 1 October to 19 November 2019 was termed the non-heating period due to high ambient temperatures (daily average  $22.2 \pm 2.4$  °C). This period is generally characterized by large-scale paddy residue burning (Chandra and Sinha, 2016). In the



5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM heating period from 20 November to 19 December 2019, the daily average temperature dropped from 17.0 °C to 9.9 °C.

### 5.3.2. Airveda low-cost sensor

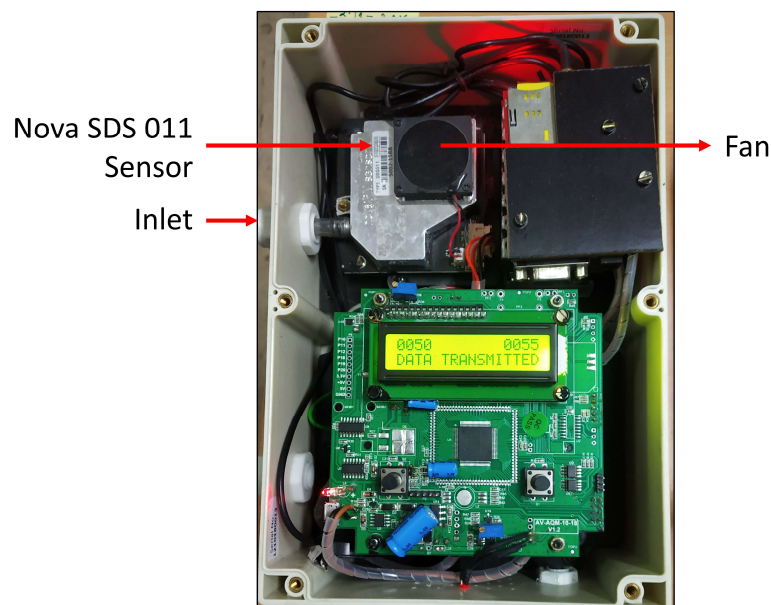


Figure 5.2: The inside and the inlet of the low-cost Airveda PM sensor

This chapter uses the outdoor Airveda low-cost PM sensors (Figure 5.2), which convert the light scattered by airborne particles into the mass concentration of PM<sub>10</sub> and PM<sub>2.5</sub> in real-time using proprietary algorithms. These monitors use a NOVA PM sensor, SDS011 sensor, to detect and quantify the mass concentration of ambient particulate matter. The sensors used in this chapter transmitted data to a local server using an internet connection from a SIM card placed in the instrument. The company also mentions that the sensors regulate the humidity of the air, but no specific details are disclosed to the user. Table 5.1 lists the major difference between the Airveda low-cost sensors and the reference analyzers. These sensors can measure PM<sub>10</sub> and PM<sub>2.5</sub> mass concentrations up to 1999 and 999  $\mu\text{g m}^{-3}$ , respectively.

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

Parameter	Airveda air quality monitor	Thermo Fisher 5014 $\beta$ -attenuation PM analyzer
Measurement principle	Laser-based light scattering	$\beta$ -attenuation
Measurable particle size	Particles larger than 0.3 $\mu\text{m}$	PM <sub>2.5</sub> (aerodynamic diameter < 2.5 $\mu\text{m}$ ) PM <sub>10</sub> (aerodynamic diameter < 10 $\mu\text{m}$ )
Accuracy	$\pm 10\%$	$\pm 5\%$
Measurement range	PM <sub>10</sub> : 1-1999 $\mu\text{g m}^{-3}$ PM <sub>2.5</sub> : 1-999 $\mu\text{g m}^{-3}$	1-10000 $\mu\text{g m}^{-3}$

Table 5.1: Major differences in the Airveda PM sensor and the reference Thermo Fisher 5014  $\beta$ -attenuation PM analyzer

### 5.3.3. Metrics to evaluate sensor performance

The following metrics were used to evaluate the performance of Airveda low-cost sensors against the reference analyzers:

1. Root Mean Square Error (RMSE): An indicator of the error between the reference mass concentration and the low-cost sensor concentration. It is computed as:

$$RMSE = \sqrt{\left(\frac{\sum_{i=1}^n (AV_i - Ref_i)^2}{n}\right)} \quad (62)$$

where  $AV_i$  and  $Ref_i$  refer to the  $i^{\text{th}}$  PM mass concentration from the Airveda sensor and the reference analyzer, respectively.

2. Pearson correlation coefficient ( $r$ ): This quantifies the strength of the linear relationship between variables. I performed Reduced Major regression instead of Ordinary Least Squares regression to account for the errors in both the dependent (Airveda PM) and the independent variable (Reference PM). However, the choice of regression does not affect the value of the correlation coefficient.

3. Mean Bias Error (MBE): An indicator of the average bias in the sensor PM measurements. It is computed as:

$$MBE = \frac{1}{n} \sum_{i=1}^N (AV_i - Ref_i) \quad (63)$$

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM where  $AV_i$  and  $Ref_i$  refer to the  $i^{\text{th}}$  PM mass concentration from the Airveda sensor and the reference analyzer, respectively. MBE has a negative value in the case of underestimation and a positive value in the case of overestimation.

4. Coefficient of Efficiency (COE): COE is used as a metric to evaluate model performance (Legates and McCabe Jr, 1999; Legates and McCabe, 2013). A perfect model has a COE of 1. It is computed as:

$$COE = 1.0 - \left( \frac{\sum_{i=1}^n |AV_i - Ref_i|}{\sum_{i=1}^n |Ref_i - \overline{Ref}|} \right) \quad (64)$$

where  $AV_i$  and  $Ref_i$  refer to the  $i^{\text{th}}$  PM mass concentration from the Airveda sensor and the reference analyzer, respectively. The value of COE varies between negative infinity to 1, with a higher number indicating higher agreement.

5. Accuracy: The accuracy in % was computed as (Levy Zamora et al., 2019)

$$Accuracy (\%) = 100 - \left( \left( \frac{|PM_{Ref} - PM_{Airveda}|}{PM_{Ref}} \right) \times 100 \right) \quad (65)$$

All the above parameters were computed using the openAir package (Carslaw and Ropkins, 2012) in the R programming environment.

### **5.3.4. Random Forest algorithm for improving sensor accuracy**

The accuracy of the low-cost Airveda sensors was compared by co-locating them with separate  $PM_{10}$  and  $PM_{2.5}$   $\beta$ -attenuation analyzers (Thermo Fischer Scientific 5014i) at the IISER Mohali Atmospheric Chemistry Facility (30.667°N, 76.729°E, 310 m above sea level) before (20 September to 27 September), during (1 October to 19 December), and after (31 December 2019 to 6 January 2020 and 8 February 2020 to 18 February 2020) the field deployment. The calibration site, IISER Mohali, lies in the SAS Nagar district to the southwest of Chandigarh, Panchkula, and Mohali (Chapter 3). Further details on the working principle and QA/QC of the reference analyzers have been mentioned in Chapter 4.

The raw Airveda PM measurements (temporal resolution 30 min) exhibited a moderate to a high degree of correlation with the reference PM analyzer, as indicated by Pearson's  $r$  values varying between 0.6 and 0.8. However, the sensors demonstrated a positive bias under high relative humidity (RH) conditions, as is typical of low-cost sensors without humidity control, and a non-linear behavior under high PM loading. In order to enhance the Airveda sensor's accuracy, I used

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM  
the Random Forest (RF) machine-learning algorithm to predict the reference PM measurements using the Airveda PM, ambient RH, and ambient temperature predictors (Liu et al., 2019a). To enhance the accuracy of the Airveda sensor, I used the Random Forest (RF) machine-learning algorithm to predict the reference PM measurements.

Random Forest is a non-parametric machine learning algorithm that can perform classification and regression predictions on categorical and continuous variables. Briefly, the RF model consists of an ensemble of classification and regression trees (CARTs) whose aggregated results decide the final prediction. The original dataset is randomly split into training and testing subsets where the former is used for training the model while the latter helps in model validation. An RF model generates subsets of the training data by a bootstrapping technique (random sampling with replacement). Roughly 63% of the observations in the training dataset occur at least once in the bootstrapped samples. The remaining 37% observations are termed as the “out-of-bag” observations and are used for model validation. A regression tree, grown by binary partitioning technique, is fit to each of the bootstrapped training data samples. By default, RF uses one-third of randomly selected variables from the total predictor variables at each node. The outcome is the average of the predictions made by the individual regression trees in the forest. The fully built RF model is validated by making predictions on the out-of-bag observations in the testing dataset. The accuracy of the prediction is quantified by the mean square error and the percentage of variance explained. Since the out-of-bag records were not a part of the training data, the predictions made on them essentially serve as cross-validated accuracy estimates. A limitation of the RF model is its inability to extrapolate. The range of predictions made by an RF model varies between the maximum and minimum of the dependent variable in the training dataset.

Liu et al. (2019a) reported that the accuracy of the NOVA SDS011 sensor, which is also used in the Airveda monitor, is affected by both ambient relative humidity and temperature. Therefore, I trained the RF model to predict the reference PM mass concentration using the Airveda PM, ambient relative humidity, and ambient temperature as predictors.

50% of the multi-season calibration measurements were randomly selected and used as the training dataset, while the remaining 50% of the dataset was withheld from the model for model validation. This cut-off was decided by sequentially increasing the percentage of the training dataset from 10% to 90% in increments of 10% and tracking the resulting decrease in the mean of squared residuals. It was also verified that the testing and training datasets had identical distributions. I used default values of the parameters, *n*tree (number of trees in the RF, 500), default terminal node size (5), and *m*try (number of variables tried at each split) in this chapter. All computations were

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM carried out in the R programming language using the randomForest package. The fully built RF model needs RH and temperature apart from the sensor PM concentration to predict the corrected PM concentration.

Experiment	Duration	Sensor ID and Location	Remarks
Pre-deployment calibration	field 20 Sep 2019 to 27 Sep 2019	<u>AV_0046 at CAF</u> <u>AV_0027 at CAF</u> <u>AV_0013 at CAF</u>	
Quantifying the contribution of regional crop-residue burning fires to the overall PM mass loadings	1 Oct 2019 to 19 Dec 2019	AV_0027 at Nadampur <u>AV_0013 at CAF</u> <u>AV_0046 at CAF</u>	
Post-deployment calibration	field 31 Dec 2019 to 6 Jan 2020	<u>AV_0013 at CAF</u> <u>AV_0046 at CAF</u> <u>AV_0105 at CAF</u> AV_0027 at CAF	The inlet of the AV_0027 sensor was infested with spiders and cobwebs during this period
Field calibration	8 Feb 2020 to 18 Feb 2020	<u>AV_0013 at CAF</u> <u>AV_0046 at CAF</u> <u>AV_0105 at CAF</u> <u>AV_0027 at CAF</u>	Inlets of all analyzers were cleaned before the co-location. AV_0027 was carefully cleared of the cobwebs and spiders.

Table 5.2: Location of the Airveda low-cost sensors and the duration of the experiment. The sensors with names underlined were used as part of the training dataset for the Random Forest machine learning algorithm.

The RF model (denoted as the “RF Model I”) was trained using a comprehensive multi-sensor dataset spanning 20 September 2019 to 18 February 2020, comprising the reference and the Airveda PM measurements when they were co-located, and the RH and the temperature recorded at the IISER Mohali Atmospheric Chemistry Facility at a temporal resolution of 30 min. The training dataset captured the variability in PM mass loadings and ambient meteorological conditions before, during, and after the field deployment. Table 5.2 provides further details on the Airveda sensors used in the training dataset.

Post RF correction, the accuracy of the AV\_0027 sensor improved from a root mean square error (RMSE) of 50 and 62  $\mu\text{g m}^{-3}$  to 19 and 32  $\mu\text{g m}^{-3}$ , while the mean bias error (MBE) changed from 5 and 29  $\mu\text{g m}^{-3}$  to 10 and 6  $\mu\text{g m}^{-3}$  for the PM<sub>2.5</sub> and PM<sub>10</sub> measurements, respectively (Table 5.3). For the AV\_0013 sensor, the RMSE decreased from 55 and 77  $\mu\text{g m}^{-3}$  to 32 and 46  $\mu\text{g m}^{-3}$ , while

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM the mean bias error (MBE) changed from -16 and 24  $\mu\text{g m}^{-3}$  to -4 and -2  $\mu\text{g m}^{-3}$  for the PM<sub>2.5</sub> and PM<sub>10</sub> measurements, respectively. At the same time, the Pearson correlation coefficient (r) improved from < 0.8 to > 0.9, while the coefficient of efficiency (COE) increased from < 0.4 to > 0.6 for both sensors (Table 5.3), indicating the enhancement in sensor accuracy.

Sensor ID	Slope/Intercept		Correlation coefficient (r)		RMSE ( $\mu\text{g m}^{-3}$ )		MBE ( $\mu\text{g m}^{-3}$ )		COE		n	
	Raw	RF	Raw	RF	Raw	RF	Raw	RF	Raw	RF		
PM <sub>10</sub>	AV_0013	0.9/35.2	0.8/24.7	0.8	0.9	77	46	24	-2	0.4	0.7	2957
	AV_0027	1.1/-13.2	1.0/4.3	0.7	0.9	62	32	29	6	0.2	0.6	894
	AV_0046	0.9/48.0	0.9/28.1	0.8	0.9	89	53	26	2	0.4	0.7	1686
	AV_0105	0.7/68.9	0.8/39.5	0.8	0.9	58	38	19	1	0.3	0.6	865
PM <sub>2.5</sub>	AV_0013	0.7/17.3	0.8/17.8	0.7	0.9	55	32	-16	-4	0.3	0.7	4356
	AV_0027	0.7/33.1	1.0/8.9	0.6	1	50	19	5	10	0.2	0.7	859
	AV_0046	0.8/24.3	0.8/28.5	0.6	0.9	38	21	-1	8	0.2	0.6	1591
	AV_0105	0.5/34.6	0.9/30.8	0.8	0.9	52	33	-28	16	0.4	0.5	835
PM <sub>10</sub>	AV_0027*	0.2/3.5	0.8/30.3	0.8	0.9	191	45	-163	-2	-0.9	0.7	287
PM <sub>2.5</sub>	AV_0027*	0.1/0.6	0.9/17.9	0.8	1	160	30	-138	-2	-0.9	0.7	286

Table 5.3: Slope/Intercept and correlation coefficient (r) obtained from the RMA regression between the Airveda sensor and reference PM between 20 September 2019 to 18 February 2020 when they were co-located. AV\_0027\* indicates the measurements obtained during the post-Nadampur deployment field calibration of the sensor when its inlet was infested with spiders (performed between 31 December 2019 and 6 January 2020). A separate RF model was used to correct the measurements of this period. RMSE, MBE, and COE stand for root mean square error, mean bias error, and coefficient of efficiency, respectively. 'n' refers to the number of available measurement pairs used for regression evaluation.

Figures 5.3a and 5.3c show the scatter plot of the daily averaged raw and RF corrected AV\_0013 PM (the sensor co-located with the reference analyzer for the most extended duration, Table 5.2) versus the reference PM measurements acquired between 20 September 2019 and 18 February 2020. Overall, the diel profiles in Figures 5.3b and 5.3d show that the RF-corrected PM measurements from the low-cost sensors agreed with the reference within 82% to 93%, indicating their suitability for field-deployment in remote places in north-west India in the post-monsoon and winter seasons.

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

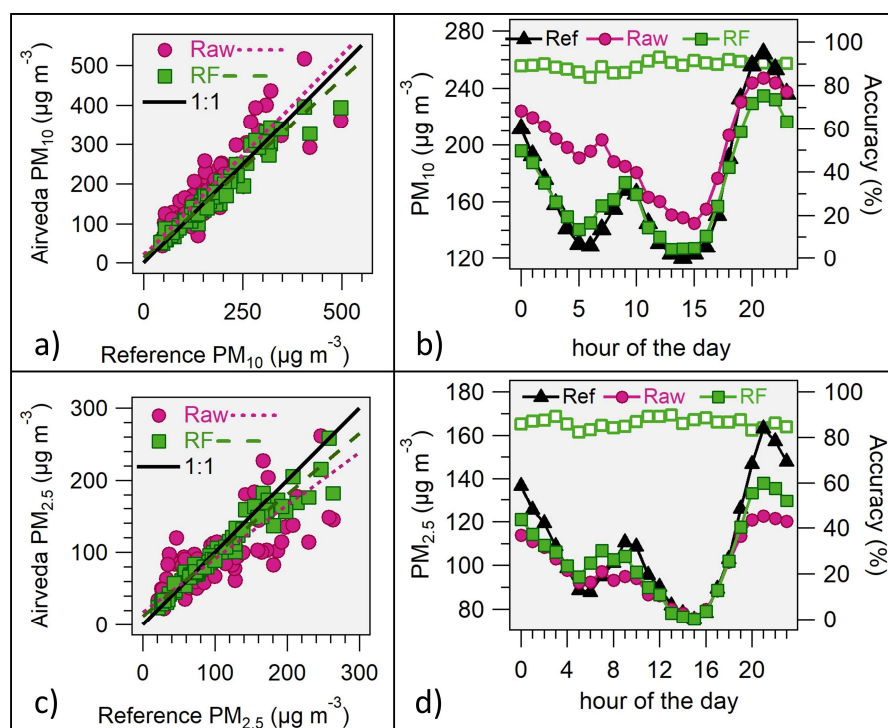


Figure 5.3: (a and c) Scatter plot of the daily averaged raw and corrected Airveda AV\_0013 PM (using “RF Model I”) versus the reference PM obtained in the field calibrations performed between 20 September 2019 and 18 February 2019. (b and d) Diel profile of the Airveda (raw and corrected) and reference PM. The solid markers connected by dots indicate hourly averaged mass concentration, while the hollow markers indicate the hourly median accuracy.

The PM measurements of the AV\_0013 sensor, deployed at Mohali from 1 October to 19 December 2019, were corrected using the “RF model I.” The RH and temperature recorded at the IISER Mohali Atmospheric Chemistry Facility were used to predict the corrected PM measurement. The Airveda sensor AV\_0027 was deployed at Nadampur from 1 October to 19 December 2019. PM measurements from 1 October to 6 November 2019 were corrected using the “RF model I” described above. The RH and temperature measurements recorded at the Patiala PPCB station, a site 30 km away from Nadampur, were used for predicting corrected PM measurements.

From 7 November onwards, the inlet of AV\_0027 (the sensor used during field deployment at Nadampur) got infested with spiders (Figure 5.4), resulting in data losses until 13 November and affecting the sensor’s measurements thereafter until 19 December 2019. After the termination of the field deployment at Nadampur, AV\_0027 was brought back to the IISER Mohali Atmospheric Chemistry Facility and co-located with the PM<sub>10</sub> and PM<sub>2.5</sub> reference analyzers from 31 December 2019 to 6 January 2020, without disturbing the inlet and the internal circuitry. A separate RF model

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM (denoted as “RF Model II”) was built based on the post-deployment calibration measurements, comprising the sensor PM, the reference PM, RH, and the temperature (Figure 5.5).



Figure 5.4: Spider webs outside a) and in the inlet b) of the Airveda low-cost sensor (AV\_0027) installed at Nadampur. The images were captured on 20 December 2019 at the end of the field deployment at Nadampur.

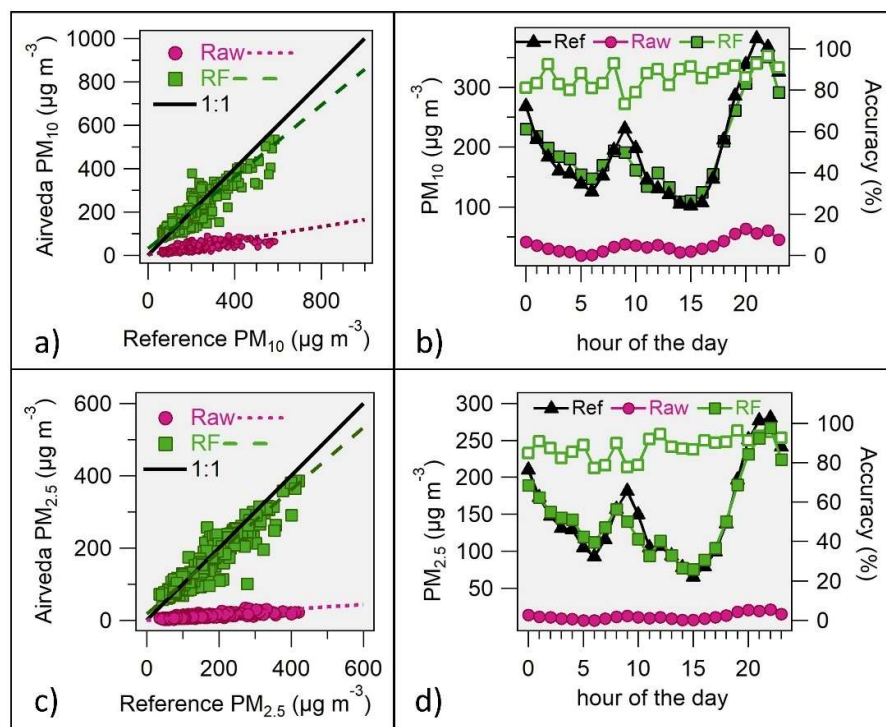


Figure 5.5: (a and c) Scatter plot of daily averaged raw and corrected 30-min averaged Airveda AV\_0027 PM versus the reference PM obtained in the field calibration performed from 31 December 2019 to 6 January 2020. (b and d) Diel profile (b and d) of the Airveda (raw and corrected) and reference PM. The solid markers connected by dots indicate hourly averaged mass concentration, while the hollow markers indicate the hourly median accuracy.

Figure 5.5 shows the scatter plot (a and c) and the diel profile (b and d) of the raw and RF corrected AV\_0027 PM measurements acquired during the post-deployment calibration while the sensor was still infested with spiders. The RF correction using “RF Model II” improved the accuracy of the AV\_0027 sensor (Table 5.3). For PM<sub>10</sub>, the RF correction increased the slope (from 0.2 to



5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM<sub>0.8</sub>, r-value (from 0.8 to 0.9, Table 5.3) and reduced the RMSE (from 191  $\mu\text{g m}^{-3}$  to 44  $\mu\text{g m}^{-3}$ ). Likewise, for PM<sub>2.5</sub>, the RF correction increased the slope (from 0.1 to 0.9), r-value (from 0.8 to 1) and reduced the RMSE from 160  $\mu\text{g m}^{-3}$  to 30  $\mu\text{g m}^{-3}$ . It was ascertained that the spider webs were the sole cause of the low PM mass concentration by cleaning the inlet and the instrument chamber and performing a second round of calibration from 8 February to 18 February 2020. After cleaning, the performance of AV\_0027 was found to be comparable to other Airveda sensors and the reference analyzer.

Corrected AV\_0027 measurements from 14 November to 19 December were predicted using the spider-infested sensor measurements and the RH and temperature recorded at the PPCB station at Patiala based on the “RF model II.”

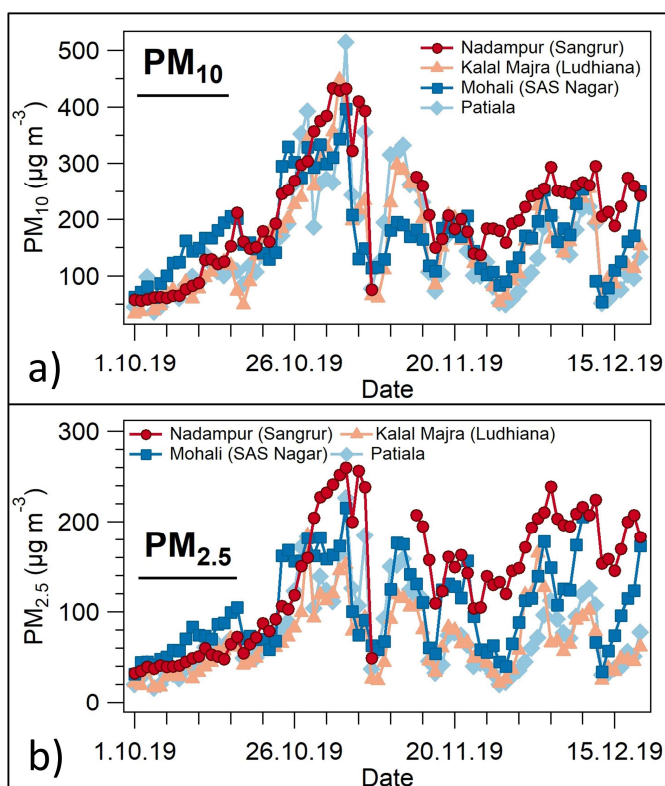


Figure 5.6: Daily average measurements of PM<sub>10</sub> (a) and PM<sub>2.5</sub> (b) at four sites in Punjab: Nadampur, Mohali, Patiala, and Kalal Majra from 1 October 2019 to 19 December 2019. For Nadampur and Mohali, the line connected by markers indicates the Random Forest-corrected PM measured using the Airveda low-cost PM sensors AV\_0027 and AV\_0013, respectively. The Punjab Pollution Control Board (PPCB) measured PM at Kalal Majra and Patiala, and the measurements (in the public domain) were obtained from <https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing>.

Figure 5.6a and 5.6b show that the RF corrected daily average PM at Nadampur and Mohali, measured using the low-cost Airveda sensors, qualitatively matched the trends in PM at Patiala

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM and Kalal Majra, an urban and a rural site at 30 km and 55 km from Nadampur in the north to the northeast directions, respectively.

### 5.3.5. Retrieval of fire counts and fire radiative power from VIIRS and boundary layer height from ERA5

I retrieved the active fires, and the FRP detected by the VIIRS sensor with a spatial resolution of 375 m, aboard the joint NASA/NOAA Suomi-National Polar-orbiting (Suomi NPP) satellite (Schroeder et al., 2014; Oliva and Schroeder, 2015; Vadrevu and Lasko, 2018; Vadrevu et al., 2019). A total of 69177 fires were detected in the NW-IGP (region spanning from 25 °N to 33 °N and 72 °E to 82 °E) between 1 October and 19 December 2019. Due to the annual periodicity of increased fire activity in this time of the year over this region, it can be safely assumed that all fires detected were paddy residue fires (Chandra and Sinha, 2016; Kumar et al., 2016; Jethva et al., 2019; Vadrevu et al., 2019).

Evaluating the ideal fetch region around a receptor site to investigate the impact of satellite-detected local fires on the receptor's PM concentration involves a trade-off. Satellite detection of small-scale paddy residue fires suffers from massive under detection due to the short duration of fires, partial-field burning practices, frequent haze and cloudy conditions in the NW-IGP, and a limited satellite overpass time (Liu et al., 2020). Therefore, the smaller the fetch region around the receptor, the lower the number of fires detected in that region. However, the impact of distant fires on the receptor's PM concentration reduces as the fetch region grows larger, particularly in October and November, characterized by a higher prevalence of periods with low wind speeds (Chapter 3).

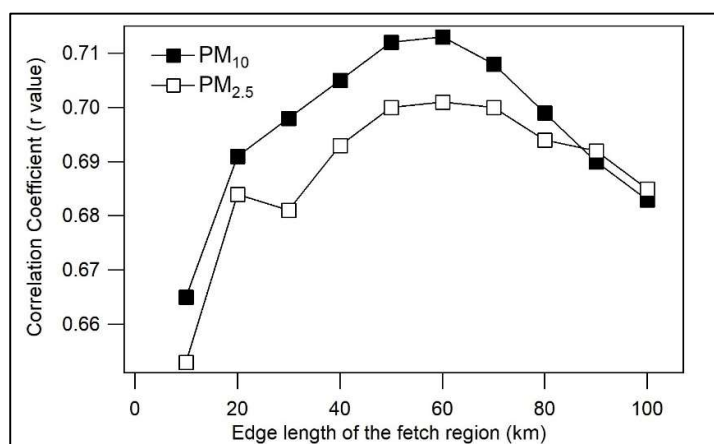


Figure 5.7: The correlation coefficient between the daily average PM at Nadampur with the daily sum of VIIRS detected Fire Radiative Power (FRP) in fetch regions with edge length varying from 10 km to 100 km.

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

In order to evaluate the ideal local fetch region, I tracked the correlation between the RF-corrected daily averaged PM mass concentration at Nadampur and the daily sum of VIIRS-detected FRP in fetch-regions with edge lengths varying from 10 km to 100 km in steps of 10 km. A higher correlation coefficient was achieved with edge lengths of 50 km and 60 km (Figure 5.7), out of which the former was chosen as the local fetch region in this chapter. A total of 2899, 2140, 1146, and 602 fires were detected in the 50 km × 50 km fetch region surrounding Nadampur, Patiala, Kalal Majra, and Mohali, respectively (Figure 5.8). Clouds and/or haze severely obscured the fire detection over the study area on 2<sup>nd</sup> November, 7<sup>th</sup> November, and 13<sup>th</sup> to 15<sup>th</sup> November 2019. The same was confirmed by the imagery obtained from the Worldview tool of NASA's Earth Observing System Data and Information System (EOSDIS) (Figure 5.9).

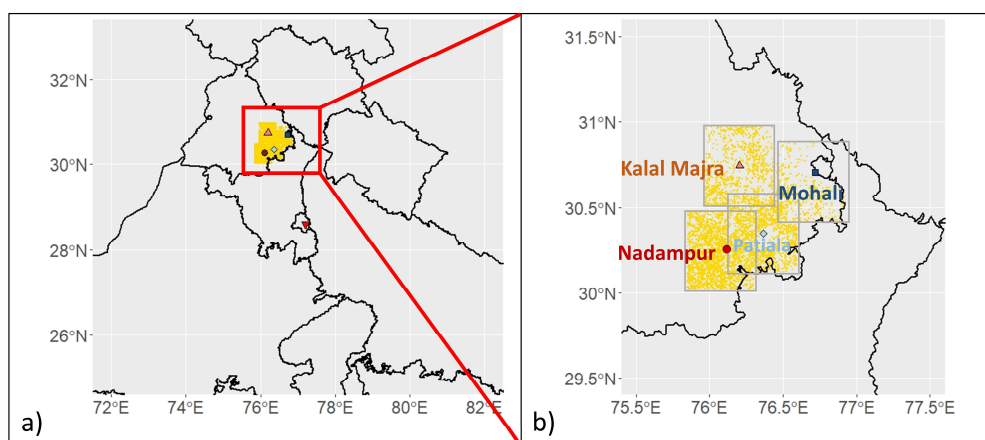


Figure 5.8: Active fire detections (yellow dots) from the Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) captured between 1 October 2019 and 19 December 2019 in a 50 km × 50 km grid box surrounding Nadampur (n = 2899), Mohali (n = 602), Kalal Majra (n = 1146), and Patiala (n = 2140).

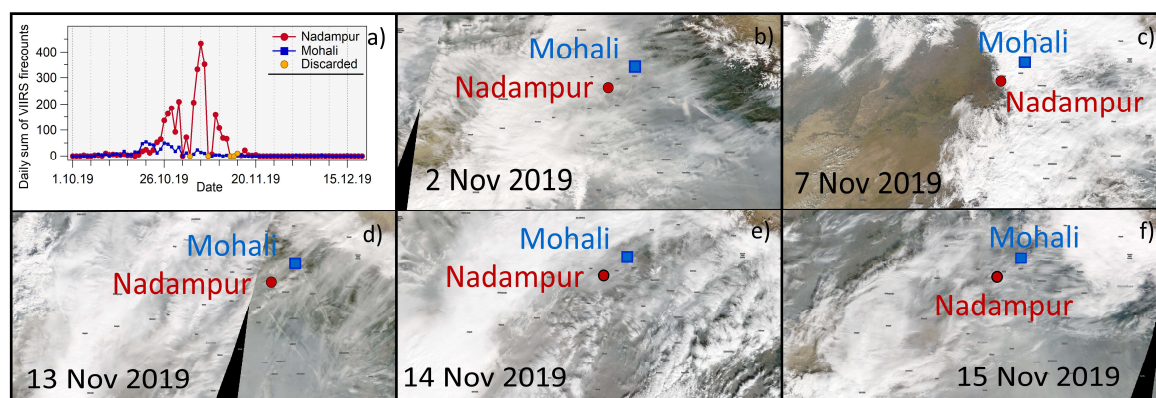


Figure 5.9: a) Daily sum of active fires detected by the VIIRS in a 50km × 50 km grid box surrounding Mohali and Nadampur from 1 October 2019 to 19 December 2019. The orange markers show the days with dense cloud cover over the study area. [b) to f)] Satellite images (The Visible Infrared Imaging Radiometer Suite (VIIRS) Corrected Reflectance imagery) showing the cloud cover over the study area on 2, 7, and 13-15 November 2019.

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM  
 Fire Radiative Energy (FRE) is FRP integrated over space and time (Vadrevu and Lasko, 2018). The daily sum of FRP was used to estimate the FRE assuming that all fires in this period were paddy residue fires that lasted for one hour (Thumaty et al., 2015).

$$\text{Fire Radiative Energy (MJ)} = \text{Fire Radiative Power (MJ s}^{-1}\text{)} \times 3600 \text{ s} \quad (66)$$

The mass of aerosol emission ( $M_{\text{esa}}$ ) can be directly inferred from the FRE as:

$$M_{\text{esa}} \text{ (kg)} = C_e \text{ (kg MJ}^{-1}\text{)} \times \text{FRE (MJ)} \quad (67)$$

where  $C_e$  is the Coefficient of Emissions, usually derived from coincident satellite measurements of Aerosol Optical Thickness (AOT) and FRP over a region (Ichoku and Ellison, 2014; Vadrevu and Lasko, 2018). This approach provides a direct estimate of the PM emissions without invoking the concept of emission factors.

Finally, the mass of paddy residue burnt was estimated from the FRP, assuming all fires detected in the study period were paddy residue fires, lasting for one hour (Thumaty et al., 2015).

$$\text{Daily fuel burnt (kg)} = \frac{\text{Daily sum of Fire Radiative Power (MJ s}^{-1}\text{)} \times 3600 \text{ s}}{14.57 \text{ (MJ kg}^{-1}\text{)}} \quad (68)$$

In the above equation,  $14.57 \text{ MJ kg}^{-1}$  is the calorific value of paddy straw (Van Hung et al., 2020). This approach provides an estimate of the amount of residue burned during the satellite overpass time within the  $50 \times 50 \text{ km}$  grid but does not provide any estimate of the total area burned or the fraction of the residue burned on any particular plot.

The ventilation coefficient (a product of wind speed and boundary layer height) that affects the dispersion of pollutants has a strong impact on the ambient mass concentration of PM. The daily average VC was computed as (Sujatha et al., 2016; Saha et al., 2019):

$$VC \text{ (m}^2\text{s}^{-1}\text{)} = \text{Boundary Layer Height (m)} \times \text{Wind Speed (ms}^{-1}\text{)} \quad (69)$$

The daily averaged BLH was computed from the hourly averaged BLH for the four sites of Nadampur Mohali, Kalal Majra, and Patiala. The hourly averaged BLHs were retrieved from the fifth-generation reanalysis dataset (ERA5) of the global climate from the ECMWF (European centre for medium range weather forecast) (C3S, 2017).

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

To estimate the emission strength per fire, I assume the boundary layer to be well mixed and compute the daily mass of PM in the 50km ×50km grid box surrounding the receptor for which FRE and amount of fuel burned were calculated as:

$$\text{Mass of PM (g)} = \text{PM } (\mu\text{g m}^{-3}) \times \text{Daily average BLH(m)} \times 2.5 \times 10^{12} (\text{m}^2) \times 10^{-6} \quad (70)$$

PM here refers to the daily averaged RF-corrected PM measured using the Airveda sensors for Nadampur ad Mohali. For Patiala and Kalal Majra, the daily-averaged PM measurements reported by the PPCB were used.

The ambient temperature range between 25 °C to 30 °C is referred to as the “thermoneutral zone” in which humans can regulate their core body temperature by adjusting the skin blood flow (Kingma et al., 2014; Kuht and Farmery, 2014). A dip in temperature below 25 °C activates thermoregulatory metabolic response mechanisms to limit heat loss. A further drop in ambient temperature generates the need for supplemental heating to help maintain the body temperature. However, the body temperature is also affected by additional factors, including RH (evaporative heat loss) and solar radiation (radiative heat gain) (Kuht and Farmery, 2014).

The daily averaged additional heating demand that factors in the effect of ambient temperature, solar radiation, and RH was estimated as:

$$\text{Heating Demand } (^\circ\text{C}) = 25^\circ\text{C} - \text{Globe Temperature} \quad (71)$$

where 25 °C is the lower end of the “thermoneutral” zone.

The globe temperature ( $T_g$ ) is a physical parameter defined as the heat transfer between the body and its environment by radiation and convection (Oliveira et al., 2019). It was calculated based on a parametrization reported in Carter et al. (2020) as:

$$\text{Globe Temperature } (T_g) = 0.009624 \times \text{SR} + 1.10 \times \text{AT} - 0.00404 \times \text{RH} \quad (72)$$

where SR refers to the solar radiation, AT refers to the ambient temperature, and RH refers to the relative humidity. The globe temperature for the four sites in Punjab was calculated using the SR, AT, and RH measurements at a temporal resolution of 15 minutes, which was then converted to daily averaged values. The AT, SR, and RH measurements for Mohali were obtained from the IISER Mohali Atmospheric Chemistry Facility (Sinha et al., 2014b), while those for Kalal Majra and Patiala were obtained from the PPCB stations at these sites. Since co-located measurements of AT, SR, and RH were not available for Nadampur, the measurements acquired at the closest PPCB station at Patiala (~30 km away from Nadampur) were used.

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

### **5.3.6. Tools for the source apportionment of ambient particulate matter**

Source apportionment tools provide information about the pollution sources and quantify their contribution to the overall PM mass loadings (Belis et al., 2014). In this chapter, the source apportionment of the daily averaged PM mass loadings at four receptor sites (Nadampur, Mohali, Kalal Majra, and Patiala) was carried out using two independent tools: multiple linear regression (MLR) and positive matrix factorization (PMF).

MLR is a multivariate technique that can be used to relate the aerosol mass to known tracers or suitable proxies (Shah et al., 1985; Chan et al., 1999). An MLR model for expressing PM mass concentration as a linear combination of its sources can be formulated as:

$$PM = \sum_{i=1}^P a_i C_i + a_o \quad (73)$$

where PM is the PM<sub>10</sub> or PM<sub>2.5</sub> mass concentration, P is the number of sources, C<sub>i</sub> is the concentration/loading of the i<sup>th</sup> tracer, and a<sub>i</sub> is the i<sup>th</sup> coefficient to be determined using MLR. The contributions of only those sources can be quantified using the MLR model that have a suitable proxy.

The contribution of the following sources to the PM loadings at the receptor sites was quantified in the MLR model as:

1. Local paddy residue fires: The daily sum of fires observed in a 50 km × 50 km area surrounding the receptor was used as a proxy for PM emissions from local paddy residue fires.
2. Harvesting activities: Most studies account for agricultural emissions only from residue burning. However, harvesting operations could potentially lead to PM emissions (Péterfalvi et al., 2018; Maffia et al., 2020). The mechanical turbulence in the soil due to the movement of combines along with the crop processing can lead to the emission of grain dust, plant debris, bioaerosols (Pattey and Qiu, 2012). As most fields are cleared by residue burning (Lohan et al., 2018), large fire counts on a particular day imply that a large quantity of paddy was harvested in the past. Therefore, I used the daily sum of fire counts in a 50 km × 50 km area surrounding the receptor with a negative lag as a proxy for harvest activity and determined an experimental lag using cross-correlation analysis. Most farmers in the study region used long-duration rice cultivars that are harvested when the optimum date for wheat sowing (1 November) is about to arrive or has already passed. Consequently, there appears to be a relatively short 5-day lag between harvest and burning.

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

3. Regional fires: About 40 to 50% of the air masses spend 3-days over the NW-IGP in the post-monsoon (October-November) and winter (December to February) seasons (Pawar et al., 2015). Therefore, the 3-day rolling sum of regional fires in the NW-IGP was used as a proxy to quantify their contribution to the PM loadings at the receptor site.
4. Heating demand: The daily average heating demand was used as a proxy for PM emissions from increased heating demand at the receptor site.
5. Ventilation: The daily averaged VC was used as a proxy for the effect of dispersion on the PM mass concentration at the receptor site.

MLR for the four receptor sites was performed using a freeware called Paleontological Statistics (PAST) (Hammer et al., 2001b). The input file comprised the daily averaged PM mass concentration from 1 October to 19 December 2019, along with the relevant proxies for the different sources described above. The days where cloud cover obscured fire detection were not included in the input dataset of MLR. The quality of the fit was assessed using the adjusted r-square values, which represent the proportion of the variance of the dependent variable (PM mass concentration) explained by the independent variables (sources)

MF is a multivariate receptor modeling tool that decomposes the PM measurements into the sum of contributions from time-invariant source profiles (Paatero and Tapper, 1994; Lee et al., 1999; Reff et al., 2007; Viana et al., 2008). PMF is one of the most extensively used receptor models for ambient PM source apportionment (Hopke, 2016). In this chapter, I used the US EPA PMF 5.0 in the constrained mode for the PM source apportionment at four receptor sites (Norris et al., 2014). A detailed description of the technique is available in Brown et al. (2015). Briefly, the mathematical formulation of the PMF model can be represented as:

$$x_{ij} = \sum_{k=1}^p g_{ik} f_{kj} + e_{ij} \quad (74)$$

where  $x_{ij}$  is the input data matrix containing  $i$  observations and  $j$  species, and  $P$  is the number of factors contributing to PM mass loadings. PMF decomposes the input matrix into a linear combination of a factor contribution matrix ( $g_{ik}$ ), a factor profile matrix ( $f_{kj}$ ), and a residual matrix ( $e_{ij}$ ).

The input PMF files contained 80 observations for Patiala and Kalal Majra and 73 and 79 observations for Nadampur and Mohali. Since the unexplained PM mass gets included in the residuals, the cloudy days with sub-optimal fire detection were also included in the input file, unlike the MLR analysis where such days must be removed from the input. The input file for each

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM site contained nine species, including the daily averaged mass concentrations of PM<sub>10</sub> and PM<sub>2.5</sub>, the daily sum of local paddy residue fires and harvest activity, the 3-day rolling sum of regional fires, daily averaged heating demand, solar radiation, and RH, and the stagnation coefficient (defined as the inverse of the VC). Increased ventilation leads to a reduction in the PM levels, an inverse relationship that cannot be explained by PMF due to its positive constraints. Therefore, the VC used in MLR was replaced with its inverse, which I call the stagnation coefficient, in PMF. This approach allows the PMF to pick up the accumulation of PM due to stagnant conditions.

$$\text{Stagnation Coefficient } (m^{-2}s^1) = \frac{1}{\text{Ventilation Coefficient}} = \frac{1}{\text{Wind Speed} \times \text{BLH}} \quad (75)$$

## 5.4. Results and discussion

### 5.4.1. Contribution of paddy residue burning to the PM emissions

The highest correlation between the PM (PM<sub>10</sub>, PM<sub>2.5</sub>) measurements and the daily local paddy residue fires detected by VIIRS was observed for Nadampur (0.73, 0.74), followed by Patiala (0.67, 0.70), Kalal Majra (0.54, 0.55), and Mohali (0.33, 0.21) in the non-heating period. Sangrur district, which contains the village Nadampur, has the highest reported crop-production and active fire detections in Punjab (Singh et al., 2020). The first-ever ground-based PM measurements acquired over Nadampur in this chapter establish a strong link between PM emissions in rural Punjab and satellite-detected paddy residue fires. Lower R values of the correlation coefficient between PM levels and paddy residue fires at other sites possibly suggest that sources other than paddy residue fires contributed to the PM or the paddy residue was managed by methods other than burning.

In order to constrain the PM emissions derived using satellite-based fire-detection methods over this region, I estimated the value of the Ce using the PM measurements obtained in this chapter. The satellite-derived Ce over Punjab is reported to be 0.04 kg MJ<sup>-1</sup> for PM<sub>10</sub> (Ichoku and Ellison, 2014; Vadrevu and Lasko, 2018). The measured PM mass concentrations at Nadampur, a site characterized primarily by agricultural activities, were converted to the absolute PM mass (kg) and plotted against the daily sum of FRE (MJ) (Figure 5.10b and 5.10d). The value of Ce, estimated from the slope of the best-fit line, was 0.038 ± 0.004 kg MJ<sup>-1</sup> and 0.024 ± 0.001 kg MJ<sup>-1</sup> for PM<sub>10</sub> and PM<sub>2.5</sub>, respectively, which agreed with the Ce value reported for PM<sub>10</sub> emissions over this region.

Bottom-up PM emission inventories rely on the laboratory-measured emission factors to convert the amount of fuel (agricultural residue) into corresponding pollutant emissions. For paddy



5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM residue, the measured PM<sub>2.5</sub> EF varies from 5.86 to 20.67 g kg<sup>-1</sup> depending upon the moisture content of the residue (Sanchis et al., 2014). Combining the daily average PM mass over Nadampur (derived from ground-based PM measurements) with the top-down estimate of paddy residue burnt (derived from the FRP detected by VIIRS in a 50 km × 50 km area around Nadampur), a top-down EF was estimated for PM<sub>10</sub> and PM<sub>2.5</sub> emissions from paddy residue burning (Figure 5.10a and 5,10c). Top-down EFs of 561.4 g kg<sup>-1</sup> for PM<sub>10</sub> and 351.6 g kg<sup>-1</sup> for PM<sub>2.5</sub> were obtained for PM emissions from paddy residue burning. The PM<sub>2.5</sub> EF of paddy residue burning estimated in this chapter is more than 17 times larger than the EFs measured in laboratory studies (Sanchis et al., 2014).

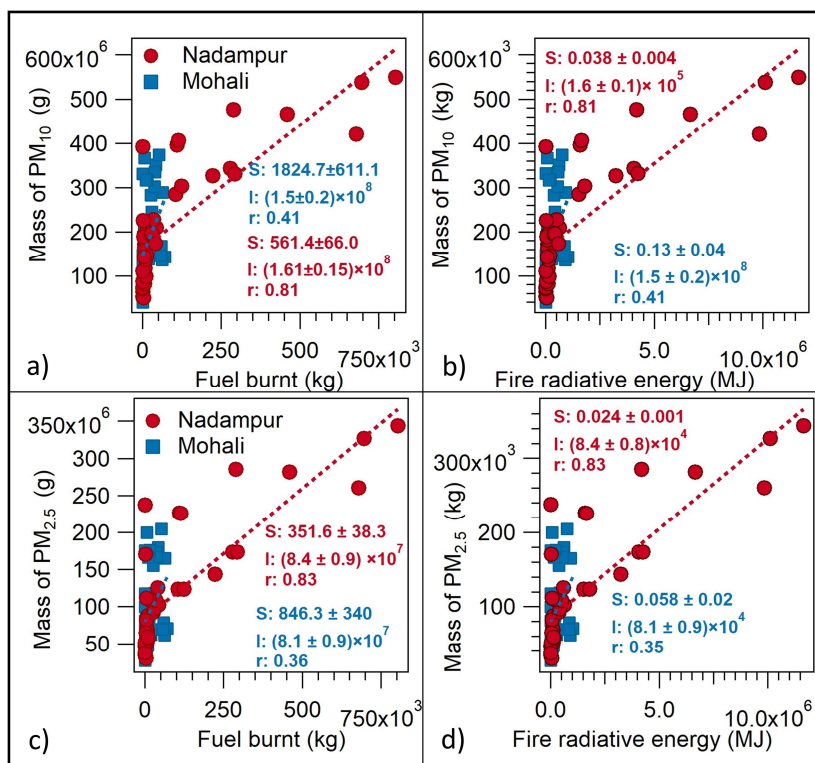


Figure 5.10: Scatter plot of the daily sum of paddy residue burnt versus the mass of a) PM<sub>10</sub> and c) PM<sub>2.5</sub>. Scatter plot of the daily sum of fire radiative energy versus the mass of b) PM<sub>10</sub> and d) PM<sub>2.5</sub>.

The fact that the Ce value estimated in this chapter agrees with the reported top-down estimates, while the EF derived from the top-down estimate of paddy residue burnt does not agree with laboratory-measured EFs, points towards the underdetection of fire counts and FRP by VIIRS. Liu et al. (2020) also reported the massive underdetection of paddy residue fires by satellites, primarily due to the limited overpass time (1:30 LT for VIIRS) that is out of sync with the fire activity, partial burning practices which aggravate underdetection, frequent haze, and cloudy conditions, and the moderate spatial resolution of the satellites. The top-down FRP-derived estimate of the

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM paddy-residue burnt in 2019 in Punjab (the region between 29 °N to 33 °N and 74 °E to 77 °E), assuming all fires to be agricultural, was less than 2% (0.09 million tonnes) of the bottom-up estimates (~7-8 million tonnes) (Kumar et al., 2019; Kulkarni et al., 2020b), further confirming the underdetection of fire activity over this region. This confirms that the PM EFs estimated in this chapter are an order of magnitude higher due to the suboptimal FRP and active fire detection. If the same is reconciled, the estimates of EF values will fall well within the laboratory reported range. Several top-down emission inventories compensate for this underdetection by assuming a 24-hour fire duration for every detected fire (Wiedinmyer et al., 2006; Wiedinmyer et al., 2011), even though most crop residue burning fires are short-lived and burn hardly for more than one hour (Thumaty et al., 2015; Liu et al., 2019b).

I used the estimated PM<sub>10</sub> and PM<sub>2.5</sub> EFs to calculate the corresponding emissions from paddy residue burning over the NW-IGP (25° E to 33° E and 72 °N to 82°N). A total of 61.1 Gg of PM<sub>10</sub> and 38.3 Gg of PM<sub>2.5</sub> was emitted from 1 October 2019 to 19 December 2019 (Table 5.4) over the NW-IGP. Out of the total PM emissions from paddy residue burning over the NW-IGP, roughly 80% was emitted from Punjab and northern Haryana (29 °N to 33 °N and 74 °E to 77 °E) (Figure 5.11, Table 5.4). Using a top-down analysis, Vadrevu and Lasko (2018) reported that the total TPM emissions over Punjab varied between 42.9 Gg to 66.1 Gg in the November months of 2012 to 2016. I estimated 29.6 Gg of PM<sub>10</sub> emissions from paddy residue burning over Punjab during November 2019. The total number of fires detected by VIIRS in this region in November 2019 (28960) was half of the fires detected in November 2016 (59979), explaining the lower PM<sub>10</sub> emissions.

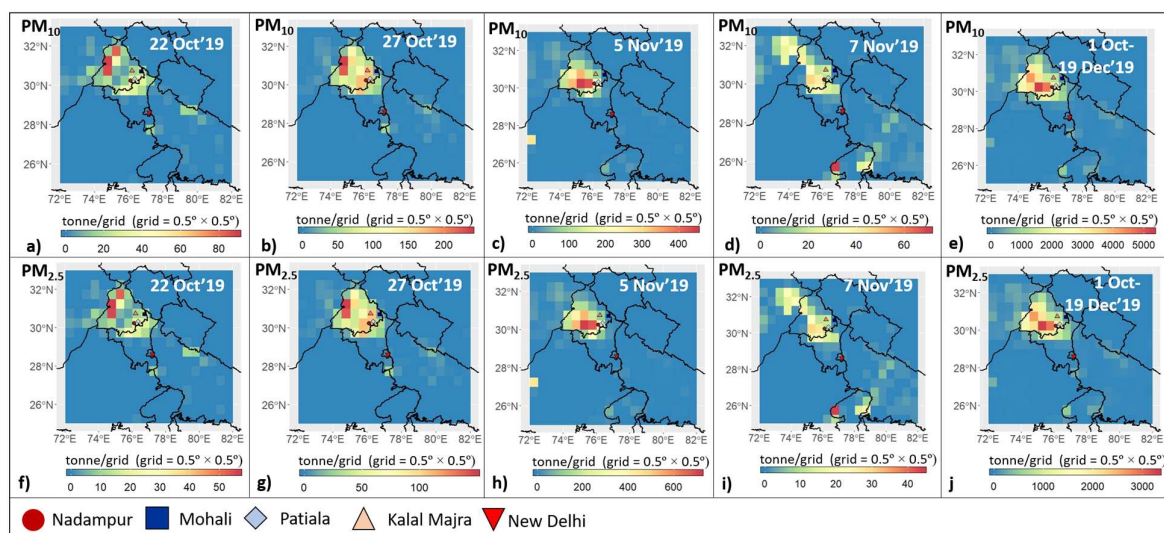


Figure 5.11: PM<sub>10</sub> and PM<sub>2.5</sub> emissions in the NW-IGP at a spatial resolution of 0.5° by 0.5° from paddy residue fires detected by the VIIRS.

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

A recent hybrid emission inventory reported that the agricultural residue burning led to the emissions of 142 Gg (Beig et al., 2020) of PM<sub>2.5</sub> in Punjab and Haryana alone in 2018, which is 3.7 times larger than my estimate of PM<sub>2.5</sub> emissions over the entire NW-IGP (Table 5.4). Another bottom-up study using crop production, coupled with assumptions regarding the harvest index and the fraction of crop residue burned, reported that agricultural residue burning emitted 194 Gg of PM<sub>2.5</sub> annually in Punjab and Haryana in 2017-18. Both emissions estimates are larger than the 2008-09 combined wheat and paddy residue burning emissions of 118.25 Gg of PM<sub>2.5</sub> in Punjab and Haryana estimated using the IPCC method (Jain et al., 2014). Although there are variations in the bottom-up estimates of PM emissions, this chapter's results suggest that the bottom-up approaches may be overestimating the total emissions, possibly due to recent changes in the residue management strategies.

Spatial Region	Type	Duration	PM <sub>10</sub> (Gg)	PM <sub>2.5</sub> (Gg)	Reference
NW-IGP		Oct'19	23.4	14.7	
(25° E to 33° E, and 72°N to 82°N)	Top-down	Nov'19	36.0	22.6	This chapter
		1-19 Dec'19	1.7	1.1	
Punjab and Northern Haryana		Oct'19	19.8	12.4	
(29 °N to 33 °N and 74 °E to 77 °E)	Top-down	Nov'19	29.6	18.5	This chapter
		1-19 Dec'19	0.7	0.4	
Punjab	Top-down	November	42.9 to 66.1		(Vadrevu and Lasko, 2018)
Punjab and Haryana	Bottom-up	Oct'18 - Nov'18		142.00	(Beig et al., 2020)
Punjab and Haryana	Bottom-up	Jul'17 to Jun'18	235.80	194.10	(Singh et al., 2020)
Punjab and Haryana	Bottom-up	2008-2009		118.25	(Jain et al., 2014)

Table 5.4: Estimates of emissions of PM<sub>10</sub> and PM<sub>2.5</sub> from agricultural residue burning compiled from different studies.

In October, the onset of paddy residue burning was observed in the north-western districts of Punjab (India), namely Amritsar, Gurdaspur, and Tarn Taran, and Pakistan (Figure 5.11a and 5.11f). The paddy residue fires peaked in the first week of November (Figure 5.11c and 5.11h), with southern Punjab consisting of Sangrur, Patiala, and Bhatinda districts reporting maximum emissions of PM<sub>10</sub> and PM<sub>2.5</sub>. Delayed burning was observed in Uttar Pradesh and Madhya Pradesh (Figures 5.11d and 5.11i).

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

### 5.4.2. Contribution of domestic biofuel burning to PM mass loadings

In 2011-12, roughly 60.8% of rural households in Punjab relied on solid fuels, including firewood and cow-dung cakes, for cooking (NSSO, 2011). In the heating period from 20 November to 19 December 2019, the daily average temperature dropped from  $\sim 17.0$  °C to  $\sim 9.9$  °C. A drop in temperature results in the increased use of solid fuels, such as firewood, cow-dung cakes, and dry waste, for space heating and water heating, even in urban India (Hakkim et al., 2019).

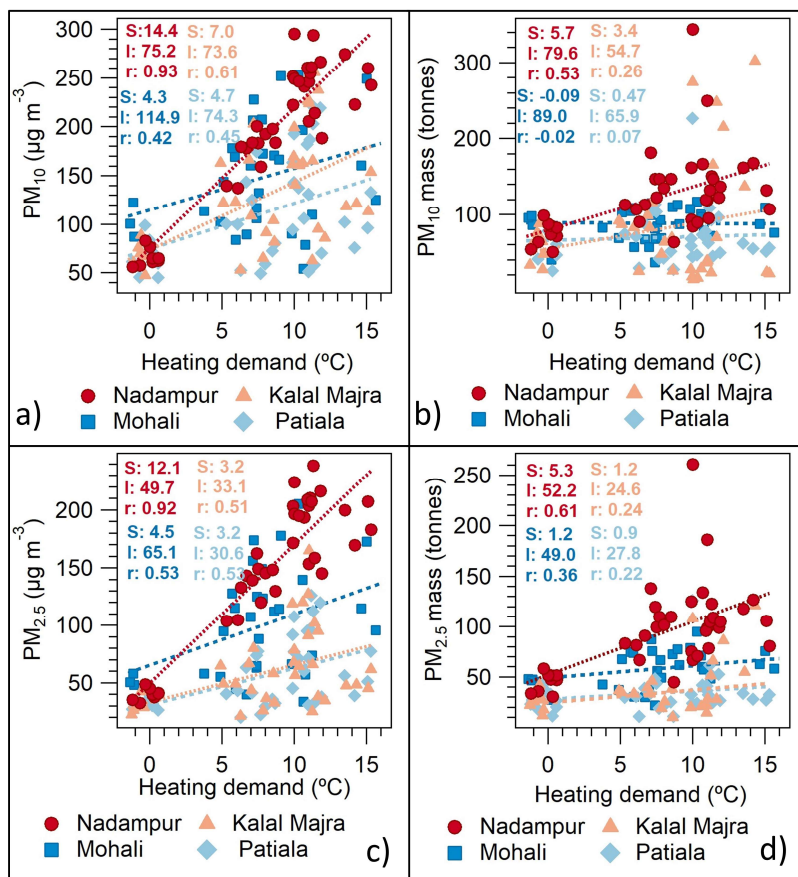


Figure 5.12: Scatter plot between the daily average heating demand versus the PM mass concentration (a and c) and absolute mass of PM (b and d) observed at Nadampur, Mohali, Kalal Majra, and Patiala.

Figure 5.12c shows a plot of the daily average  $PM_{2.5}$  mass concentration versus the heating demand for urban (Mohali and Patiala) and rural sites (Nadampur and Kalal Majra) during the heating period. As the heating demand increased by 1°C, the  $PM_{2.5}$  mass concentration at Nadampur increased by  $12.1 \mu g m^{-3}$  ( $r = 0.92$ ). The correlation between the  $PM_{2.5}$  mass concentration and the heating demand was less significant at other sites, as indicated by the lower Pearson's r values. An increase in the heating demand by 1°C, increased the  $PM_{2.5}$  mass concentration by  $4.5 \mu g m^{-3}$  ( $r = 0.53$ ),  $3.2 \mu g m^{-3}$  ( $r = 0.51$ ), and  $3.2 \mu g m^{-3}$  ( $r = 0.53$ ) at Mohali, Kalal Majra, and Patiala sites,

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM respectively. A similar enhancement with heating demand was also observed in the PM<sub>10</sub> mass concentration (Figure 5.12a).

The temperature drop was accompanied by a drop in the daily average planetary BLH and VC from 335 to 251 m and from 195 to 149 m<sup>2</sup> s<sup>-1</sup>, respectively. A reduction in the BLH under low temperatures amplifies the effect of increased emission rates. I calculated the PM<sub>2.5</sub> and PM<sub>10</sub> emissions using the daily average BLH and examined their relationship with the heating demand to address this issue. Heating demand neither dominated wintertime PM<sub>10</sub> (-0.1 t °C<sup>-1</sup>, r = -0.02) and PM<sub>2.5</sub> (1.2 t °C<sup>-1</sup>, r = 0.36) emissions in Mohali, nor wintertime PM<sub>10</sub> (0.5 t °C<sup>-1</sup>, r = 0.1) and PM<sub>2.5</sub> (0.9 t °C<sup>-1</sup>, r = 0.2) emissions in Patiala. Both cities are tier-1 cities (Figure 5.12b and 5.12d).

However, every 1 °C increase in the heating demand led to the emission of 5.7 tonnes of PM<sub>10</sub> (r = 0.53) and 5.3 tonnes (r = 0.61) of PM<sub>2.5</sub> in Nadampur (Figures 5.12b and 5.12d). In comparison, heating demand did not dominate wintertime PM<sub>10</sub> (3.4 t °C<sup>-1</sup>, r = 0.26) and PM<sub>2.5</sub> (1.2 t °C<sup>-1</sup>, r = 0.24) emissions in Kalal Majra, a village that is also home to several industries. The fuel used by the residents for space heating and cooking is typically dependent on their income. Solid fuels such as fuelwood and cow dung cakes are available at a nominal cost, while the use of electric geysers, room heaters, and gas stoves costs more. Out of the 22 districts in Punjab, Ludhiana (containing Kalal Majra) ranked second while Sangrur (containing Nadampur) ranked tenth in terms of the per-capita income in 2015-16 (ESOPB, 2015), possibly explaining higher heating-related PM emissions in Nadampur.

### 5.4.3. Source apportionment of PM mass loading

MLR results indicate that paddy harvest, local and regional paddy residue burning, heating demand, and ventilation could explain 79% and 84% of the variability in PM<sub>10</sub> and PM<sub>2.5</sub> measurements from 1 October to 19 December 2019 at Nadampur. The observed PM at Nadampur was modeled using MLR as:

$$PM_{10} = 0.50 \times HV + 0.54 \times LF + 0.004 \times RF + 11.40 \times HD - 0.14 \times VC + 122.85 \quad (r_{adj}^2 = 0.79, n = 70) \quad (76)$$

$$PM_{2.5} = 0.26 \times HV + 0.34 \times LF + 0.002 \times RF + 12.50 \times HD - 0.10 \times VC + 60.53 \quad (r_{adj}^2 = 0.84, n = 70) \quad (77)$$

where HV stands for harvest, LF stands for the daily sum of local paddy residue fires, RF stands for the 3-day rolling sum of regional fires, and VC stands for the ventilation coefficient. However, these sources explained a lower proportion of the PM variability at other sites. Sources mentioned

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM above explained 55% and 57%, 53% and 42%, and 28% and 34%, of the variability in daily averaged PM<sub>10</sub> and PM<sub>2.5</sub> at Patiala, Kalal Majra, and Mohali, respectively, eliciting that different sources dominate the PM mass loading in different rural and urban environments.

A 5-factor PMF solution in constrained mode (%dQ < 5) was found to optimally resolve the contribution of paddy harvest, local paddy residue fires, regional paddy residue fires, heating demand, and other sources (including stagnation) across all sites. Figure 5.13 shows the factor fingerprints obtained from the PMF runs for the four receptor sites.

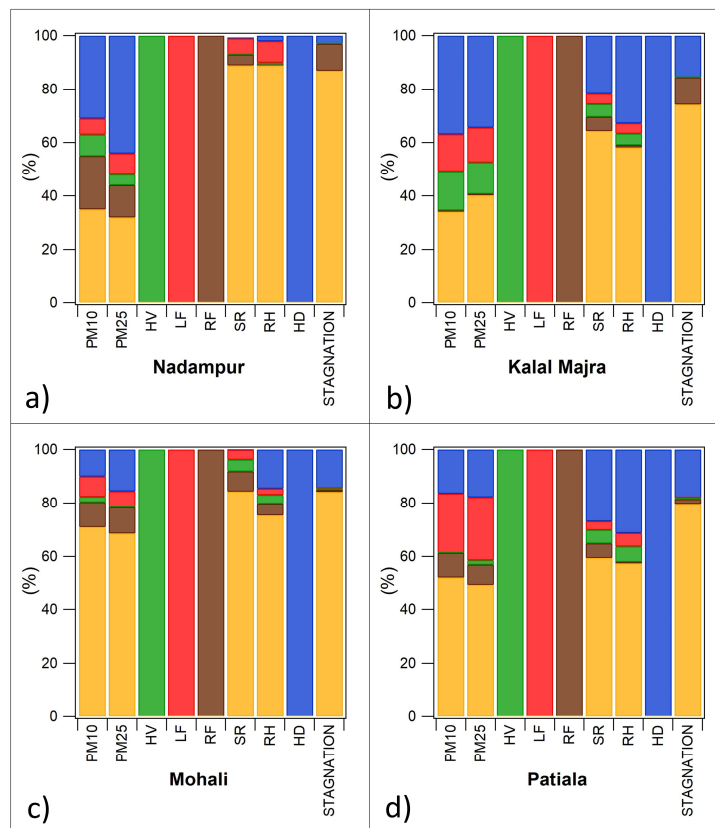


Figure 5.13: Factor fingerprints obtained from the constrained 5-factor PMF run for the four sites. HV refers to paddy harvest, LF refers to local paddy residue fires, RF refers to regional paddy residue fires, HD refers to heating demand, SR refers to solar radiation, and RH refers to relative humidity. The five factors are local paddy residue fires (red), paddy harvest (green), regional paddy residue fires (brown), heating demand (blue), and others (orange).

The source apportioned PM was computed using the modeled coefficients (in the case of MLR), and the factor contribution and factor profiles (in the case of PMF) for the four sites, and the overall sum was compared with the measured PM (Figures 5.14 and 5.15).

From 1 October to 19 October 2019, low PM levels were observed at four sites than the corresponding PM levels in November and December. In the descending order, daily average  $\pm$

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM ambient variability PM ( $PM_{10}$  and  $PM_{2.5}$ ) levels of  $136.8 \pm 47.3 \mu\text{g m}^{-3}$  and  $66.9 \pm 20.4 \mu\text{g m}^{-3}$ ,  $100.9 \pm 45.5 \mu\text{g m}^{-3}$  and  $48.4 \pm 11.1 \mu\text{g m}^{-3}$ ,  $87.0 \pm 28.9 \mu\text{g m}^{-3}$  and  $43.3 \pm 17.7 \mu\text{g m}^{-3}$ , and  $75.7 \pm 28.7 \mu\text{g m}^{-3}$  and  $37.7 \pm 15.7 \mu\text{g m}^{-3}$  were observed at Mohali, Nadampur, Patiala, and Kalal Majra, respectively (Figures 5.14 and 5.15). In this period, “Other sources,” including cooking, waste disposal, transport sector, and industries which are not explained by both MLR and PMF models, contributed more than 60% to the PM mass concentration at all four sites.

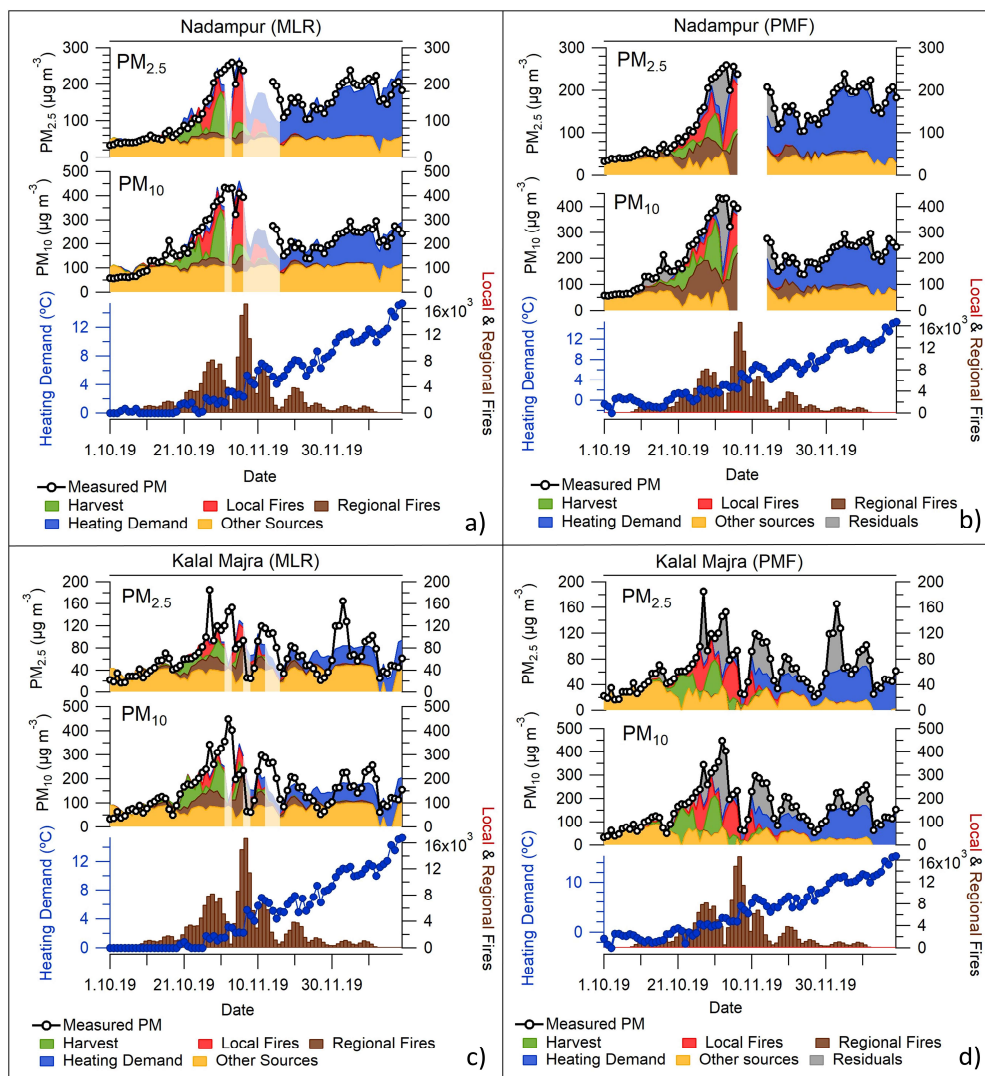


Figure 5.14: Measured daily averaged PM mass concentration and the daily contribution of different sources to the overall PM mass loading at Nadampur (a and b) and Kalal Majra (c and d) from 1 October to 19 December 2019, evaluated using Multiple Linear Regression (MLR) and Positive Matrix Factorization (PMF). The predicted source-contribution on days with cloud cover/haze and data gaps, which were not included while evaluating the MLR regression, are shown in lighter shades in a) and c)

The period from 20 October to 19 November 2019 was dominated by agricultural activities such as paddy harvest and local and regional paddy residue burning. As per the MLR results, local

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM  
paddy residue fires contributed on average of  $49.3 \mu\text{g m}^{-3}$  and  $30.8 \mu\text{g m}^{-3}$  to the  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  loadings at Nadampur. In comparison, this source contributed an average of  $32.9 \mu\text{g m}^{-3}$  and  $29.5 \mu\text{g m}^{-3}$  to the  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$ , respectively, based on the PMF analysis. Combining both, I estimate that on average,  $41.1 \pm 11.7 \mu\text{g m}^{-3}$  and  $30.2 \pm 0.9 \mu\text{g m}^{-3}$   $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  at the receptor site, respectively, can be attributed to local paddy residue fires.

Paddy harvesting, an essential agricultural activity primarily carried out using mechanized combine harvesters, increased the  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  mass concentration at Nadampur by  $44.8 \pm 1.7 \mu\text{g m}^{-3}$  and  $20.1 \pm 5.2 \mu\text{g m}^{-3}$ , respectively. It is noteworthy that the contribution of paddy harvesting to  $\text{PM}_{10}$  was comparable to that of local paddy residue fires at Nadampur. No studies over India have evaluated the impact of harvesting on PM emissions. However, a study investigating the contribution of different agricultural activities to the PM emissions from the agriculture sector in Canada reported that harvesting contributed 18.4 % and 13.9 % to  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$ , respectively (Pattey and Qiu, 2012).

Local paddy residue fires contributed (MLR/PMF) a maximum of  $234 \mu\text{g m}^{-3}/141.0 \mu\text{g m}^{-3}$  and  $146.0 \mu\text{g m}^{-3}/126.8 \mu\text{g m}^{-3}$  to the  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  loadings at Nadampur on 5 November 2019 when 433 active fires were detected by VIIRS (Figure 5.14). Considering that the air parcel undergoes dilution while traveling from the source to a distant receptor site, it is likely that the STILT derived maximum  $\text{PM}_{2.5}$  enhancement of  $255 \mu\text{g m}^{-3}$  at Delhi (Cusworth et al., 2018) overestimated the impact of distant paddy burning fires on the Delhi air quality, given the maximum contribution of paddy fires on the day with the highest impact was  $\sim 146 \mu\text{g m}^{-3}$  in the source region itself.

A large proportion of residuals in the PMF solution for Nadampur (between  $103.2$  and  $286 \mu\text{g m}^{-3}$  for  $\text{PM}_{10}$  and between  $79.6$  to  $154.0 \mu\text{g m}^{-3}$  for  $\text{PM}_{2.5}$ ) from 1 November to 3 November 2019 was possibly caused due to suboptimal fire detection as a result of widespread haze and cloudy conditions (Figure 5.14). Liu et al. (2020) corrected the haze/cloud-related dips in the daily paddy residue fire activity by assuming that the spatially and temporally aggregated fires together can be assumed as a long-term fire with a gaussian-like distribution.



## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

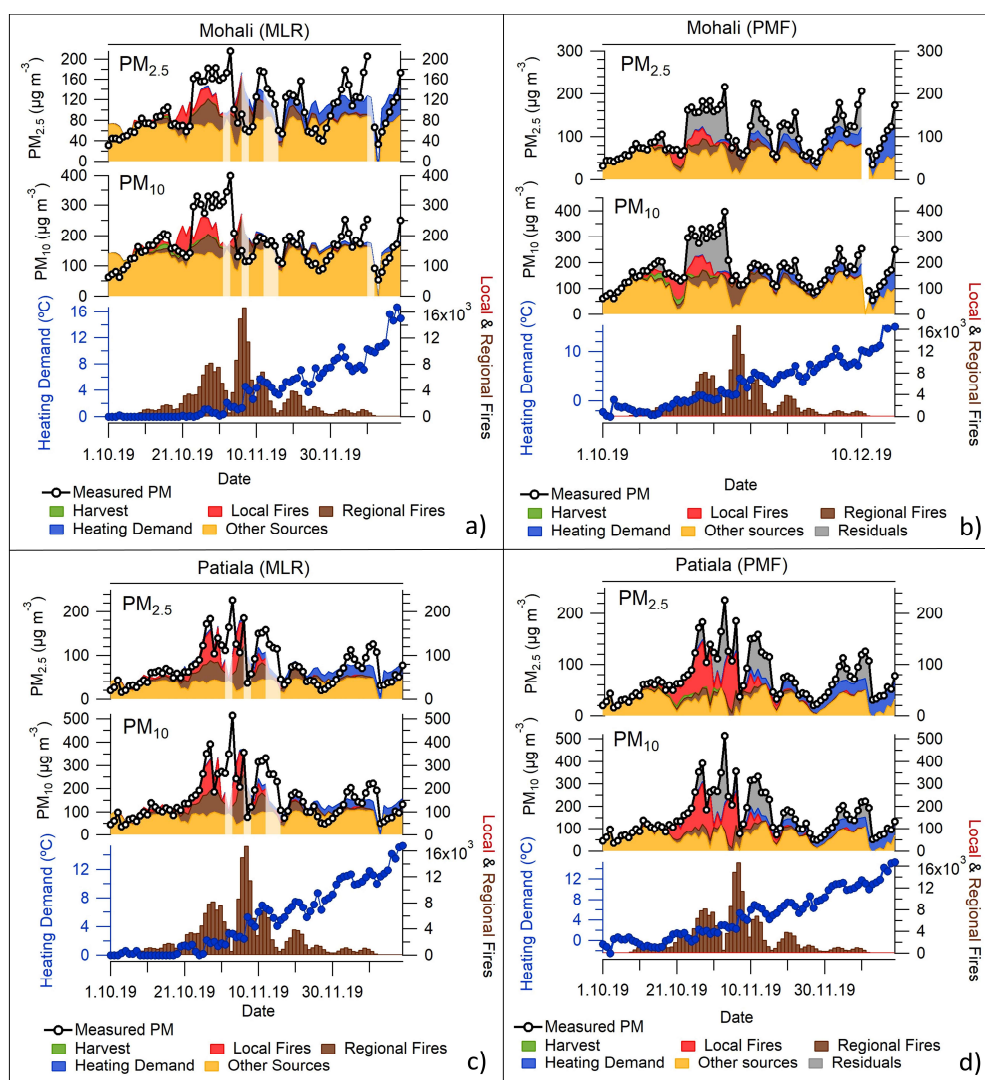


Figure 5.15: Measured daily averaged PM mass concentration and the daily contribution of different sources to the overall PM mass loading at Mohali (a and b) and Patiala (c and d) from 1 October to 19 December 2019, evaluated using Multiple Linear Regression (MLR) and Positive Matrix Factorization (PMF). The predicted source-contribution on days with cloud cover/haze and data gaps, which were not included while evaluating the MLR regression, are shown in lighter shades in a) and c).

Even though there were differences in the PM mass concentration attributed to paddy harvest and local and regional paddy residue burning by PMF and MLR techniques, the total PM emitted by agricultural activities was comparable between the two approaches. In the descending order, paddy harvest enhanced the PM<sub>10</sub> (PM<sub>2.5</sub>) mass concentration in Nadampur and Kalal Majra by  $44.8 \pm 1.7 \mu\text{g m}^{-3}$  ( $20.1 \pm 5.2 \mu\text{g m}^{-3}$ ),  $41.1 \pm 7.0 \mu\text{g m}^{-3}$  ( $12.0 \pm 4.4 \mu\text{g m}^{-3}$ ), respectively (Figures 5.14 and 5.15) from 20 October to 19 November 2019. In contrast, paddy harvest contributed  $< 5 \mu\text{g m}^{-3}$  to the PM levels at Mohali and Patiala. Local and regional paddy residue fires enhanced the PM<sub>10</sub> (PM<sub>2.5</sub>) mass concentration in Patiala, Nadampur, Mohali, and Kalal Majra by  $103.1 \pm 13.7 \mu\text{g m}^{-3}$  ( $50.1 \pm 7.6 \mu\text{g m}^{-3}$ ),  $97.0 \pm 36.6 \mu\text{g m}^{-3}$  ( $53.4 \pm 16.8 \mu\text{g m}^{-3}$ ), and  $58.1 \pm 6.6 \mu\text{g m}^{-3}$  ( $34.1 \pm$

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM  $8.7 \mu\text{g m}^{-3}$ ),  $55.4 \pm 12.8 \mu\text{g m}^{-3}$  ( $22.1 \pm 5.6 \mu\text{g m}^{-3}$ ) respectively, from 20 October to 19 November 2019 (Figures 5.14 and 5.15). Interestingly, the contribution of paddy harvest to the PM mass concentration in the Kalal Majra (in Ludhiana district) was comparable to that of Nadampur, but the combined contribution of local and regional fires was lower. Punjab Agricultural University in Ludhiana has successfully promoted the use of the Happy Seeder, a machine that sows wheat through the paddy residue (Sidhu et al., 2015; Mahal et al., 2016). The impact of increased Happy Seeder penetration in Kalal Majra is apparent from the 60% lower fire counts and 47% lower ambient PM<sub>2.5</sub> ( $86 \mu\text{g m}^{-3}$ ) than Nadampur village during the paddy residue burning period. The highest contribution of agricultural activities to the PM loadings in south-western Punjab (Nadampur and Patiala) accentuates the need for expanding the spatio-temporal scale of ground-based measurements, particularly over regions with large-scale agricultural activities.

Even though the paddy residue fires had subsided after 20 November 2019, the air quality continued to be poor. The dip in the daily average temperature from  $17.0 \text{ }^\circ\text{C}$  to  $9.9 \text{ }^\circ\text{C}$  increased the heating demand, enhancing the residential burning of solid fuels for space and water heating. In the descending order, the increased heating demand enhanced the PM<sub>10</sub> (PM<sub>2.5</sub>) mass concentrations in Nadampur, Kalal Majra, Patiala, and Mohali by  $151.2 \pm 47.2 \mu\text{g m}^{-3}$  ( $120.1 \pm 8.8 \mu\text{g m}^{-3}$ ),  $84.2 \pm 24.6 \mu\text{g m}^{-3}$  ( $34.1 \pm 7.3 \mu\text{g m}^{-3}$ ),  $38.3 \pm 7.6 \mu\text{g m}^{-3}$  ( $22.5 \pm 0.3 \mu\text{g m}^{-3}$ ), and  $30.7 \pm 0.4 \mu\text{g m}^{-3}$  ( $22.3 \pm 17.3 \mu\text{g m}^{-3}$ ), respectively, from 20 November to 19 December 2019 (Figures 5.14 and 5.15). The contribution of heating demand-induced residential heating emissions was the lowest in the urban cities, Mohali, and Patiala, out of all four sites. Therefore, the promotion of cleaner technology for water and space heating can play a substantial role in improving air quality in periods with low ambient temperatures, particularly in rural areas.

Compared to residential heating emissions, the combined average contribution of agricultural activities (paddy harvest, local and regional paddy residue burning) to the PM mass concentration was significantly smaller ( $< 10 \mu\text{g m}^{-3}$ ) at all four receptor sites from 20 November 2019 onwards. Therefore, while the impact of paddy residue burning is limited to just one month, heating emissions triggered by low ambient temperatures would continue to deteriorate the air quality throughout the three months of the winter season (December, January, and February, monthly mean  $< 20 \text{ }^\circ\text{C}$ , Awasthi et al. (2011)).

The effect of daily fluctuations in the heating demand on the overall PM mass loading could be included in emission inventories and air quality forecasting using normalized emission maps and interactive meteorological data as is currently done for biogenic emissions in the Model for

5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM Emission of Gases and Aerosols from Nature (MEGAN) (Guenther et al., 2012). This could vastly improve wintertime air quality forecasts.

## **5.5. Conclusions**

This chapter is the first to report ground-based PM measurements over Nadampur, a rural village in the Sangrur district of Punjab. The trends in PM levels were compared across four sites in Punjab from 1 October to 19 December 2019. The major results are:

Satellite detection of fire activity over northwest India was constrained by estimating the Ce (a top-down proxy based on AOT) and PM<sub>10</sub> and PM<sub>2.5</sub> EFs (a laboratory-based proxy). Agreement of the estimated Ce with the reported Ce, while disagreement between the EFs suggests that the fire counts detected by satellites are vastly underestimated over this region.

The contribution of agricultural activities (paddy harvest, local and regional paddy residue burning from 20 October to 19 November 2019) and residential heating emissions (from 20 November to 19 December 2019) to the PM mass loadings was more significant in rural sites than the urban sites. Agricultural activities increased the PM<sub>10</sub> (PM<sub>2.5</sub>) at Nadampur, a rural site, and Mohali, an urban site by  $141.7 \pm 38.3 \mu\text{g m}^{-3}$  ( $73.5 \pm 11.6 \mu\text{g m}^{-3}$ ) and  $61.3 \pm 5.9 \mu\text{g m}^{-3}$  ( $34.5 \pm 9.0 \mu\text{g m}^{-3}$ ). Residential heating, triggered by the dip in ambient temperature, increased the PM<sub>10</sub> (PM<sub>2.5</sub>) at Nadampur and Mohali by  $151.2 \pm 47.2 \mu\text{g m}^{-3}$  ( $120.1 \pm 8.8 \mu\text{g m}^{-3}$ ) and  $30.7 \pm 0.4 \mu\text{g m}^{-3}$  ( $22.3 \pm 17.3 \mu\text{g m}^{-3}$ ), respectively. Therefore, mitigating PM emissions requires different policies for rural and urban sites.

Overall, the impact of paddy residue burning on the ambient air quality was confined to roughly one month in the year. The air quality continued to be poor even after the abatement of paddy residue fires due to the increased residential solid fuel burning for space heating, triggered by a dip in the ambient temperatures. Unlike paddy residue burning, which receives massive attention from all stakeholders, this overlooked source profoundly impacts air quality throughout the winter. The promotion of cleaner technology for water and space heating can substantially improve the air quality, particularly in rural areas.

A network of ground-based low-cost PM sensors can help validate emissions inventories. A dense monitoring network with spatial interpolation techniques that account for

## 5. Contribution of paddy residue burning and domestic biofuel burning for heating purposes to PM

dispersion could potentially be used to locate PM sources with a disproportionate impact on air quality in urban environments and industrial areas (Garg and Sinha, 2017).

A key limitation of this chapter was the unavailability of PM measurements at Nadampur from 7 November to 13 November 2019 due to the spider infestation of the inlet. Future studies should deploy multiple low-cost sensors to ensure data completeness. Securing the inlet with a net fabric after confirming that the same does not affect the sensor accuracy is also a cost-effective solution. Co-located temperature and RH measurements, crucial for predicting corrected PM measurements using the RF machine learning algorithm, were not available for Nadampur. Therefore, a dedicated co-located meteorological station can further improve the prediction accuracy.

This is the first study to highlight the significant contribution of harvesting activities to the PM mass loading in rural areas. However, the reported emission factors of soil preparation methods such as tilling (plowing, disking) are considerably larger than harvesting (Chen et al., 2017; Maffia et al., 2020). Apart from monitoring the PM levels, future studies should measure wind speed, soil moisture content, and soil composition in agricultural fields that significantly impact the dust flux (Foroutan et al., 2017) emitted by soil preparation activities.

# Chapter 6

---

## **Conclusions: Major findings and outlook**

This chapter highlights the major results compiled in this thesis and presents a discussion on their real-world implications. The Indo-Gangetic Plains in India have extremely high particulate matter levels that pose a severe health risk to the residents. Unprecedented urbanization and industrialization have further led to worsening of the air quality. While the Central Pollution Control Board has mandated the National Ambient Air Quality Standards, most sites in this region fail to comply with them almost throughout the year.

No city or region is an island. Since air is a shared resource, the pollutants measured in any region's air may have been emitted elsewhere and gotten advected to the receptor site via winds. To gauge the success or failure of any mitigation policy or intervention, one needs to know the relative contribution of air mass transport to the regional background of a pollutant. Chapter 3 of this thesis quantifies the contribution of long-range transport to the particulate matter loadings at Mohali, a suburban site in the north-western Indo Gangetic Plain. The major conclusions are:

1. Long-range transport from the arid regions of Arabia, the Thar desert, the Middle East, and Afghanistan significantly increases the loading of fine- and coarse-mode particles in north-west India, particularly in the summer season characterized by frequent dust storms.
2. Further, most policies assume particulate matter as an isolated entity that does not undergo physical and/or chemical modifications in the atmosphere. Chapter 3 demonstrates that a significant fraction of secondary particulate matter belonging to coarse mode can originate from the aqueous-phase processing of gas-phase precursors, especially in winter and monsoon season. A shallow boundary layer in winter coupled with the additional pollutant (both primary emissions of particulate matter and gas-phase precursors) burden from burning waste/biomass to keep warm further exacerbates the health impacts. Therefore, standalone policies that aim to curb only the primary aerosol fraction may not be successful in reducing PM levels in north-west India.
3. While long-range transport results in significant enhancements in the average PM loading in north-west India, its impact on the percentage of exceedance events (days not in compliance with the daily average  $PM_{10}$  and  $PM_{2.5}$  air quality standards) is at best minimal. This shows that residents of the north-western Indo-Gangetic Plains suffer from chronic exposure to particulate matter loadings above the air quality standards due to the presence

of regional sources. The average regional background of fine mode PM was found to be  $80 \mu\text{g m}^{-3}$  and  $110 \mu\text{g m}^{-3}$  in the post-monsoon and winter season, respectively, which is higher than the daily average  $\text{PM}_{2.5}$  NAAQS of  $60 \mu\text{g m}^{-3}$ .

4. Combustion drives the levels of  $\text{PM}_{2.5}$  in the atmosphere during all seasons. Regional combustion sources must be targeted to ensure more frequent compliance with the air quality standards,

Effective mitigation policies are marred by several challenges, including a lack of spatially and temporally dense air quality data. The expansion of ground-based PM measurement networks is limited by the high cost and complicated troubleshooting of research-grade instruments. Simultaneously, there has been a surge in demand for open-source air quality data as several people are eager to track their personal exposure to particulate matter. In response, several low-cost personal optical sensors have directly permeated the market without undergoing a comprehensive field evaluation. I have addressed the issues dealing with the data quality of such sensors in Chapters 4 and 5 of my thesis. The primary results are:

1. The accuracy of low-cost sensors is affected by variations in meteorological conditions. Although the positive bias induced due to water absorption at high relative humidity was well known, the effect of inlet losses at high wind speeds has been recognized for the first time. This has important implications for sites in the Indo-Gangetic Plains that witness frequent dust storms.
2. The Airveda sensors (calibrated in India) outperformed the Laser Egg sensors (calibrated in China) when their performance was compared with a US EPA-approved  $\beta$ -attenuation analyzer. Given the wide spatial variations in aerosol composition, low-cost sensors must be calibrated at the site of application (or at places with identical aerosol composition) rather than the country/site of origin. Further, if there is seasonal variation in the aerosol sources and meteorological conditions at the application site, a multi-season calibration is vital to ensure robust data quality.
3. Low-cost optical sensors assume a size distribution and aerosol density to convert the volume derived from scattering into mass concentration. A singular value of density is neither spatially nor temporally accurate. Besides, equating the optical diameter with the aerodynamic equivalent diameter can lead to large deviations in the particulate matter measurements from the sensor, particularly when the loading of high-density mineral dust is high.

4. I have proposed issue-specific empirical corrections and machine learning-based corrections in this thesis to correct the inaccurate low-cost sensor PM measurements. Although the Random Forest method outperformed the empirical correction in terms of improving the sensor accuracy, its performance was limited by its inability to extrapolate beyond the training dataset. The predictive performance of Random Forests needs to be improved for further applications in the field of air quality. Zhang et al. (2019) suggested coupling Random Forests, a non-parametric algorithm, with a parametric regression, termed as Regression Enhanced Random Forests (RERFs). RERFs significantly improved the predictive performance of Random Forests. Improved Random Forest models have been used in several applications such as predicting travel decisions, retweet behaviour, and the academic performance of students. However, the use of this improved model remains unexplored in the field of air quality.
5. While the proposed corrections may be beneficial for the integration of low-cost sensors with existing measurement networks, they may not be feasible for an end-user looking to monitor personal exposure. Therefore, the sensor design and data processing algorithm must be improved based on the recommendations in Chapter 4 to limit the need for complicated post-processing.
6. Out of all the proposed corrections, correction for hygroscopic growth can be performed relatively conveniently using accurate measurements of ambient relative humidity from an integrated humidity sensor. However, it is critical to place the sensor outside the particulate matter sensor rather than inside to avoid artifacts from thermal absorption.

Another major challenge in the path of effective mitigation policies is the sheer diversity in sources that contribute to the pollutant loading. For instance, construction, traffic, resuspended dust, waste burning, leaf litter burning, biomass burning, fossil fuel burning, etc., are some of the dominant sources. People of different socio-economic backgrounds have different fuel choices, which causes further variations in the fuel types used for cooking, space heating, and transportation. Although most sites in the NW-IGP do not comply with the PM NAAQS throughout the year, from November onwards, the poor air quality in this region is blamed on the large-scale agricultural residue burning in Punjab and Haryana. However, the emission strength remains poorly constrained due to the lack of ground-based measurements over rural areas within the source region.

I used the Random Forest corrected PM measurements from the Airveda sensors to quantify the contribution of paddy residue fires and heating demand induced domestic biofuel burning to the overall PM mass loading in north-west India (Chapter 5). The primary results are:

1. The coefficient of emissions ( $C_e$ ), which directly relates the FRE to the mass of aerosol emitted, was estimated to be  $0.038 \text{ kg MJ}^{-1}$  and  $0.024 \text{ kg MJ}^{-1}$  for  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$ , respectively, which is comparable to the satellite-derived  $C_e$  of  $0.04 \text{ kg MJ}^{-1}$  for  $\text{PM}_{10}$  over this region.
2. Punjab and northern Haryana were responsible for more than 80% of the total  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  emissions over the NW-IGP from paddy residue fires between 1 October and 19 December.
3. Heating demand-induced PM emissions contribute significantly to the PM mass loadings in rural areas. Promoting cleaner technology for water and space heating can play a key role in improving air quality during the cold season.
4. Harvest activities, currently not accounted for in emission inventories, were found to impact the air quality in rural areas and contributed roughly 14% of the  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  mass loading at Nadampur.
5. Overall, the impact of paddy residue burning on the ambient air quality was confined to roughly one month in the year. The air quality continued to be poor even after the abatement of paddy residue fires due to the increased residential solid fuel burning for space heating, triggered by a dip in the ambient temperatures. Unlike paddy residue burning, which receives massive attention from all stakeholders, this overlooked source profoundly impacts air quality throughout the winter. The promotion of cleaner technology for water and space heating can substantially improve the air quality, particularly in rural areas.

This thesis unravels several facets that make PM mitigation challenging in north-west India. This region witnesses poor air quality primarily in the winter, summer, and post-monsoon seasons as wet scavenging in the monsoon season leads to a significant reduction in the PM levels. North-west remains the dominant fetch for most sites in this region in winter, summer, and post-monsoon seasons. Synoptic scale transport from distant regions in the north-west was found to significantly contribute to the average PM levels. This fraction of the PM (varying between 4% to 57% of the average  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  mass concentration in different seasons), representing the transboundary pollution sources, can hardly be mitigated. However, by themselves, regional combustion sources present in north-west India are responsible for frequent non-compliance with the air quality



standards, particularly in the winter and the post-monsoon seasons. The blanket term, “regional combustion sources,” however, conceals an enormous diversity in sources including but not limited to the transport sector, powerplants, brick kilns, industries, burning firecrackers on festivals, waste burning of different scales and composition, agricultural residue burning, leaf litter burning, and residential fuel usage for cooking, space, and water heating.

By and large, it has been a common tendency to blame more “visible” sources such as dust storms, festivals, and agricultural residue burning in April to May and October to November for the poor air-quality that persists for roughly three quarters of the year. This issue picks up the maximum momentum each year in October and November, the time of paddy residue burning. It should be noted that while this discussion dwindles after November, the air quality continues to be poor. Additionally, the current air pollution discourse in north-west India is limited to cities. Research studies have selectively focused on evaluating the impact of distant crop-residue fires on New Delhi and the National Capital Region. In contrast, a similar investigation scale is conspicuously absent in rural areas where the much-hyped residue is burnt.

Therefore, policymakers looking to address poor air quality over north-west India need to adopt a data-driven multi-sectoral approach. The network of ground-based measurement needs to be expanded, particularly over rural areas. Low-cost sensors have a promising potential as they can provide an accuracy between 80 and 90% with comprehensive correction algorithms.

Improving rural and urban air quality are two separate problems. Therefore, separate source apportionment studies should be carried out for rural and urban locations using appropriate chemical tracers. Access to clean fuel for space and water heating can ameliorate roughly 50% of the PM<sub>2.5</sub> mass concentration in rural areas. Besides, a life cycle assessment of the contribution of agriculture activities (including harvest, plowing, disking) to the PM mass loadings should be carried out. The reported emission factors of soil preparation methods such as tilling (plowing, disking) are considerably larger than harvesting (Chen et al., 2017; Maffia et al., 2020). Apart from monitoring the PM levels, wind speed, soil moisture content, and soil composition should be monitored in agricultural fields that significantly impact the dust flux (Foroutan et al., 2017) emitted by soil preparation activities. In contrast, urban areas need to look inwards and address sources as more than 70% of the variability in PM<sub>10</sub> and PM<sub>2.5</sub> in Mohali could not be explained by paddy residue burning.

# List of Figures

---

Figure 1.1: Approximate emission fluxes of the primary aerosols and the precursors to secondary aerosols classified based on the source [The estimates in this figure have been compiled from (Boucher, 2015) and (Tomasi and Lupi, 2016)] .....	5
Figure 1.2: Scanning electron micrographs highlighting the diversity in aerosol particles collected at IISER Mohali showing a) Inorganic aerosol particle, b) dust aerosol, c) inorganic aerosol, d) soot particle, e) inorganic aerosol, f) Internally mixed SOA-inorganic particle, g) SOA droplet.....	6
Figure 2.1: Broad classification of analytical techniques for the characterization of ambient particulate matter .....	16
Figure 2.2: Schematic of the 5014i $\beta$ -attenuation analyzer showing the major internal parts .....	21
Figure 2.3: Figure showing different regimes of light scattering as a function of particle size and the wavelength of the incident light. [Adopted from (Brasseur and Jacob, 2017)] .....	27
Figure 2.4: Basic scattering geometry depicting the incident wave, the object, and the scattered wave ...	28
Figure 2.5: Figure showing the major internal components of the modified Laser Egg (left) sensor and the Airveda sensor (right) .....	32
Figure 3.1: (Left) Location of Mohali on a land classification map (courtesy ESA GlobCover 2009 project). The site is located in the north-western Indo-Gangetic Plain, close to the forested slopes of the foothills of the Himalayan mountain range. (Right) The exact location of the measurement facility and its spatial relationship with respect to the nearby cities, the mountain range, and potential local point sources of particulate matter. ....	50
Figure 3.2: Wind rose plot for the measurement site for the winter (December–February), summer (March–June), monsoon (July– September), and post-monsoon (October and November) seasons. Wind speed and wind direction were measured at the height of 20 m a.g.l. ....	53
Figure 3.3: Seasonal composites of sea level pressure, surface winds, and 700hPa winds. The values on the top and bottom of the legend in each figure show the maximum and minimum, respectively. (Images provided by the NOAA-SRL Physical Sciences Division, Boulder Colorado from their Web site at <a href="http://www.esrl.noaa.gov/psd/">http://www.esrl.noaa.gov/psd/</a> ).....	54
Figure 3.4: Diel box and whisker plots for fine-mode (top four panels) and coarse-mode (bottom four panels) particulate matter for the winter, summer, monsoon, and post-monsoon seasons for the period November 2011 to August 2013 respectively. The box indicates the upper and lower quarter values; the cross indicates the median, and the dots connected by lines provide the mean. The whiskers indicate the 5 <sup>th</sup> and 95 <sup>th</sup> percentiles, respectively. Periods of calm ( $< 1 \text{ m s}^{-1}$ ) have been excluded while preparing the graph. The interval highlighted in red shows the daytime low (12:00 to 16:00LT) and nighttime low (03:00 to 06:00LT). ....	57
Figure 3.5: Percentage change in t-RMSD with increase in cluster number.....	60

Figure 3.6: Mean height (a.g.l.) of all trajectories in an individual cluster as a function of trajectory running time (72h) for trajectories arriving at 09:00 and 23:00 UTC.....	60
Figure 3.7: (a–f) All individual trajectories that contributed to each of the clusters and the cluster mean, superimposed on a land classification map (courtesy of the ESA GlobCover 2009 project). The length of each mean trajectory is 3 days, and the distance between two successive data points represents a 1h interval. The average trajectory of each cluster has been superimposed using circles with a black outline. ....	62
Figure 3.8: Contribution of individual clusters to air-mass flow for all four seasons. Magenta: calm; red: south-easterly cluster; orange: south-westerly cluster; green: local cluster; light blue: slow westerly cluster; dark blue: medium westerly cluster; purple: fast westerly cluster. ....	63
Figure 3.9: Mean coarse-mode (PM <sub>10-2.5</sub> ) and fine-mode (PM <sub>2.5</sub> ) mass loading for each air-mass cluster and season at the IISER Mohali air quality station. Hatched bars indicate coarse-mode and fine-mode PM mass loadings observed during rain events.....	67
Figure 3.10: Dependence of coarse-mode PM mass loadings on the emission of gas-phase precursors and meteorological parameters for the different seasons. The marker shape distinguishes PM mass loadings measured during rain events (circles) and under dry conditions (diamonds); data points obtained while the rain gauge was not working are marked with crosses. Marker size is proportional to wind speed. The smallest markers indicate $WS \leq 1 \text{ m s}^{-1}$ , the largest markers $WS \geq 15 \text{ m s}^{-1}$ . Markers are colour coded with relative humidity.....	70
Figure 3.11: Scatter plots of fine-mode PM with CO, acetonitrile, benzene and NO <sub>2</sub> for the winter, summer, monsoon, and post-monsoon seasons. ‘a’ stands for intercept and ‘b’ stands for slope in the linear regression equation. ....	71
Figure 3.12: Percentage of days where the 24h average PM <sub>10</sub> and PM <sub>2.5</sub> mass loading (in $\mu\text{g m}^{-3}$ ) exceed the national ambient air quality standard for each air-mass cluster and season. *For calm conditions, PM <sub>2.5</sub> or PM <sub>10</sub> were averaged for all times during a 24h interval that had $WS < 1 \text{ m s}^{-1}$ only. The fraction of exceedance days is calculated based on this average and is not based on a genuine 24h average, as wind speeds do not remain low continuously. ....	78
Figure 4.1: a) Location of IISER Mohali Atmospheric Chemistry Facility (red marker) inside the campus of Indian Institute of Science Education and Research (IISER) (30.667 N, 76.729 E, 310m a.s.l.) shown in a Google Earth imagery acquired on August 10, 2016 (Map image: ©2016 DigitalGlobe, Google Earth). b) Location of the two modified Laser Egg PM sensors denoted as LE Sensor 1 and LE Sensor 2 next to the inlets of reference PM <sub>10</sub> and PM <sub>2.5</sub> $\beta$ -attenuation analyzers at the IISER Mohali Atmospheric Chemistry Facility.....	86
Figure 4.2: Figure showing the external casing, accessory inlet fan, and internal circuitry of the modified Laser Egg (LE) PM sensor.....	89
Figure 4.3: Hourly averaged values of locally measured meteorological parameters: wind speed, ambient temperature, solar radiation, relative humidity and absolute humidity, the hourly median value of wind	

direction and the hourly sum of rainfall recorded at IISER Mohali Atmospheric Chemistry Facility during the three-month calibration period from April 27, 2016, till July 25, 2016. The shaded portion represents the dry summer season (April 27, 2016 – June 15, 2016), while the unshaded portion represents the monsoon season (June 16, 2016 – July 25, 2016)..... 92

Figure 4.4: Time series plot of hourly averaged  $PM_{10}$  mass concentration from reference analyzer (circular markers in orange) and two Laser Egg (LE) sensors (green triangular and blue square markers respectively) for the period from April 27, 2016, to July 25, 2016. The shaded portion represents the dry summer period (April 27, 2016 – June 15, 2016) with average RH varying from 9% to 58%, while the unshaded portion represents the wet monsoon period (June 16, 2016 – July 25, 2016) with average RH varying from 61% to 85%. (a, bottom) The daily average value of the coefficient of divergence (COD) for raw  $PM_{10}$  measurements from the LE sensors. Vertical bars represent the daily variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles of the COD. b) Reduced Major Axis (RMA) regression of  $PM_{10}$  from LE Sensor 2 versus Sensor 1. Marginal rugs have been added to show the distribution of data c) RMA regression of  $PM_{10}$  from the LE sensors versus the reference analyzer. Markers are color-coded according to ambient RH. .... 93

Figure 4.5 (a, top) Time series plot of hourly averaged  $PM_{2.5}$  mass concentration from the reference analyzer (circular markers in orange) and two Laser Egg (LE) sensors (green triangular and blue square markers respectively) for the period from April 27, 2016, to July 25, 2016. The shaded portion represents the dry summer period (April 27, 2016 – June 15, 2016) with average RH varying from 9% to 58%, while the unshaded portion represents the wet monsoon period (June 16, 2016 – July 25, 2016) with average RH varying from 61% to 85%. (a, bottom) The daily average value of the coefficient of divergence (COD) for raw  $PM_{2.5}$  measurements from the LE sensors. Vertical bars represent the daily variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles of the COD. b) Reduced Major Axis (RMA) regression of  $PM_{2.5}$  from LE Sensor 2 versus Sensor 1. Marginal rugs have been added to show the distribution of data. c) RMA regression of  $PM_{2.5}$  from the LE sensors versus the reference analyzer. Markers are color-coded according to ambient RH. .... 98

Figure 4.6: Aspiration efficiency of the Laser Egg (LE) inlet as a function of external wind speed ( $U_o$ ) for particles with varying aerodynamic diameters ( $D_a$ ) and a density of  $2 \text{ g cm}^{-3}$ ..... 101

Figure 4.7: (Top)  $PM_{10}$  (solid markers) and  $PM_{2.5}$  (hollow markers) measurements from the reference and the two Laser Egg (LE) sensors during two dust storms observed from May 3, 2016, 17:00:00 to May 4, 2016, 08:00:00 and May 4, 2016, 17:00:00 to May 5, 2016, 02:00:00 respectively (Bottom) Relative Humidity (RH) and wind speed during the two dust storm periods. .... 102

Figure 4.8: (a and b) Markers (color-coded according to ambient relative humidity) showing the observed ratio of Laser Egg (LE) PM (after density and RH correction) to reference PM versus wind speed. The maroon, red, blue, and yellow lines (in a and b) represent the theoretical ratio of LE PM to reference PM derived from the assumed aerosol volume size distribution for summer (April 27, 2016, to June 15, 2016) shown in (c), two observed size distributions from Kanpur (Kaskaoutis et al., 2012) (d) and

Gual Pahari (Hyvärinen et al., 2011a) (e) and another assumed “low dust” distribution (f). A density of  $2 \text{ g cm}^{-3}$  has been assumed for the ambient aerosol..... 107

Figure 4.9: (a and b) Markers (color-coded according to ambient relative humidity) showing the observed ratio of Laser Egg (LE) PM (after density and RH correction) to reference PM versus wind speed. The maroon, red, blue, and yellow lines (in a and b) represent the theoretical ratio of LE PM to reference PM derived from the assumed aerosol volume size distribution for summer (April 27, 2016, to June 15, 2016) shown in (c), two observed size distributions from Kanpur (Kaskaoutis et al., 2012) (d) and Gual Pahari (Hyvärinen et al., 2011a) (e) and another assumed “low dust” distribution (f). A density of  $1.2 \text{ g cm}^{-3}$  has been assumed for the ambient aerosol..... 109

Figure 4.10: (a and b) Markers (color-coded according to ambient relative humidity) showing the observed ratio of Laser Egg (LE) PM (after density and RH correction) to reference PM versus wind speed. The maroon, red, blue, and yellow lines (in a and b) represent the theoretical ratio of LE PM to reference PM derived from the assumed aerosol volume size distribution for summer (April 27, 2016, to June 15, 2016) shown in (c), two observed size distributions from Kanpur (Kaskaoutis et al., 2012) (d) and Gual Pahari (Hyvärinen et al., 2011a) (e) and another assumed “low dust” distribution (f). ..... 111

Figure 4.11: Correlation plots of the Laser Egg (LE) PM versus the reference PM obtained after performing reduced major axis (RMA) regression for raw LE measurements (a and d), after RH and density correction (b and e), and after correcting for aspiration losses and cut-off diameter (c and f). The solid (red) line indicates the 1:1 line. In each of the above plots, “S” refers to the slope of the best-fit line, and “r” refers to the Pearson correlation coefficient. The markers are color-coded according to ambient RH..... 113

Figure 4.12: Time series of hourly averaged  $\text{PM}_{10}$  and  $\text{PM}_{2.5}$  mass concentration measured by the reference and Laser Egg (LE) sensors, pre- and post-correction for the calibration period from April 27, 2016, to July 25, 2016. The shaded portion represents the dry summer season (April 27, 2016 – June 15, 2016), while the unshaded portion represents the monsoon season (June 16, 2016 – July 25, 2016). ..... 114

Figure 4.13: a) Diel plot of ambient temperature from the reference analyzer (circular markers in orange) and two Laser Egg (LE) sensors (green triangular and blue square markers, respectively) for the calibration period. The markers connected by lines indicate the hourly averages, and the shaded portion represents the ambient variability as the 75<sup>th</sup> and 25<sup>th</sup> percentiles. b) Time series plot of hourly averaged ambient temperature from the reference and the LE sensor plotted along with solar radiation measurements (hollow circular markers) from July 13, 2016, to July 15, 2017..... 116

Figure 4.14: The deviation in Laser Egg (LE) temperature from the reference (Ref.) ambient temperature sensor explained as a function of (a) Solar Radiation (SR) during the day [06:00 to 18:59 LT] and as a function of (b) Relative Humidity (RH) during the night [19:00 to 05:59 LT]. Markers, color-coded according to the hour of the day, represent the averaged values of hourly bins derived using the measurements from the three-month calibration period (April 27, 2016, till July 25, 2016). In each of

the above plots, “I” refers to the intercept, “S” refers to the slope of the best – fit line, and “r” refers to the Pearson correlation coefficient obtained from the ordinary linear regression of the deviation in temperature versus SR (Figure 4.14a) and RH (Figure 4.14b). ..... 117

Figure 4.15: Time series plot of hourly averaged (a) relative humidity (RH) and (c) absolute humidity reported by reference RH sensor (orange circular markers) and two Laser Egg (LE) sensors (green triangular and blue square markers, respectively) for the calibration period from April 27, 2016, until July 25, 2016. A diel plot of (b) relative humidity and (d) absolute humidity for the calibration period. The markers connected by lines indicate the hourly averages, while the shaded portion represents ambient variability as 75<sup>th</sup> and 25<sup>th</sup> percentiles. .... 118

Figure 5.1: a) Map showing the locations of the Nadampur, Mohali, Patiala, and Kalal Majra in the state of Punjab, India. (b) Google Earth imagery (captured on 3 October 2020, © Google Earth and Maxar Technologies) of Nadampur village showing agricultural fields in all directions and the Airveda AV\_0027 sensor installed at the Government Senior Secondary School, Nadampur. .... 124

Figure 5.2: The inside and the inlet of the low-cost Airveda PM sensor ..... 126

Figure 5.3: (a and c) Scatter plot of the daily averaged raw and corrected Airveda AV\_0013 PM (using “RF Model I”) versus the reference PM obtained in the field calibrations performed between 20 September 2019 and 18 February 2019. (b and d) Diel profile of the Airveda (raw and corrected) and reference PM. The solid markers connected by dots indicate hourly averaged mass concentration, while the hollow markers indicate the hourly median accuracy..... 132

Figure 5.4: Spider webs outside a) and in the inlet b) of the Airveda low-cost sensor (AV\_0027) installed at Nadampur. The images were captured on 20 December 2019 at the end of the field deployment at Nadampur. .... 133

Figure 5.5: (a and c) Scatter plot of daily averaged raw and corrected 30-min averaged Airveda AV\_0027 PM versus the reference PM obtained in the field calibration performed from 31 December 2019 to 6 January 2020. (b and d) Diel profile (b and d) of the Airveda (raw and corrected) and reference PM. The solid markers connected by dots indicate hourly averaged mass concentration, while the hollow markers indicate the hourly median accuracy..... 133

Figure 5.6: Daily average measurements of PM<sub>10</sub> (a) and PM<sub>2.5</sub> (b) at four sites in Punjab: Nadampur, Mohali, Patiala, and Kalal Majra from 1 October 2019 to 19 December 2019. For Nadampur and Mohali, the line connected by markers indicates the Random Forest-corrected PM measured using the Airveda low-cost PM sensors AV\_0027 and AV\_0013, respectively. The Punjab Pollution Control Board (PPCB) measured PM at Kalal Majra and Patiala, and the measurements (in the public domain) were obtained from <https://app.cpcbcr.com/ccr/#/caaqm-dashboard-all/caaqm-landing>..... 134

Figure 5.7: The correlation coefficient between the daily average PM at Nadampur with the daily sum of VIIRS detected Fire Radiative Power (FRP) in fetch regions with edge length varying from 10 km to 100 km. .... 135

Figure 5.8: Active fire detections (yellow dots) from the Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) captured between 1 October 2019 and 19 December 2019 in a 50 km × 50 km grid box surrounding Nadampur (n = 2899), Mohali (n = 602), Kalal Majra (n = 1146), and Patiala (n = 2140)..... 136

Figure 5.9: a) Daily sum of active fires detected by the VIIRS in a 50km × 50 km grid box surrounding Mohali and Nadampur from 1 October 2019 to 19 December 2019. The orange markers show the days with dense cloud cover over the study area. [b) to f)] Satellite images (The Visible Infrared Imaging Radiometer Suite (VIIRS) Corrected Reflectance imagery) showing the cloud cover over the study area on 2, 7, and 13-15 November 2019..... 136

Figure 5.10: Scatter plot of the daily sum of paddy residue burnt versus the mass of a) PM<sub>10</sub> and c) PM<sub>2.5</sub>. Scatter plot of the daily sum of fire radiative energy versus the mass of b) PM<sub>10</sub> and d) PM<sub>2.5</sub>..... 142

Figure 5.11: PM<sub>10</sub> and PM<sub>2.5</sub> emissions in the NW-IGP at a spatial resolution of 0.5° by 0.5° from agricultural residue fires detected by the VIIRS..... 143

Figure 5.12: Scatter plot between the daily average heating demand versus the PM mass concentration (a and c) and absolute mass of PM (b and d) observed at Nadampur, Mohali, Kalal Majra, and Patiala. .... 145

Figure 5.13: Factor fingerprints obtained from the constrained 5-factor PMF run for the four sites. HV refers to paddy harvest, LF refers to local paddy residue fires, RF refers to regional paddy residue fires, HD refers to heating demand, SR refers to solar radiation, and RH refers to relative humidity. The five factors are local paddy residue fires (red), paddy harvest (green), regional paddy residue fires (brown), heating demand (blue), and others (orange). .... 147

Figure 5.14: Measured daily averaged PM mass concentration and the daily contribution of different sources to the overall PM mass loading at Nadampur (a and b) and Kalal Majra (c and d) from 1 October to 19 December 2019, evaluated using Multiple Linear Regression (MLR) and Positive Matrix Factorization (PMF). The predicted source-contribution on days with cloud cover/haze and data gaps, which were not included while evaluating the MLR regression, are shown in lighter shades in a) and c) ..... 148

Figure 5.15: Measured daily averaged PM mass concentration and the daily contribution of different sources to the overall PM mass loading at Mohali (a and b) and Patiala (c and d) from 1 October to 19 December 2019, evaluated using Multiple Linear Regression (MLR) and Positive Matrix Factorization (PMF). The predicted source-contribution on days with cloud cover/haze and data gaps, which were not included while evaluating the MLR regression, are shown in lighter shades in a) and c)..... 150

# List of Tables

---

Table 3.1: Average of the locally measured meteorological parameters for daytime/nighttime for the different clusters and seasons. For solar radiation, I provided the daytime average only. For rain, I calculated the sum of the rainfall instead of the average, and the numbers in brackets represent the number of rain events. .... 64

Table 3.2: Lower limit for the contribution of long-range transport and local pollution events to PM mass loadings in  $\% \pm 1\sigma$  of the total PM. “Negative” indicates that the PM mass loadings are not enhanced compared to the local cluster, which represents the regional background levels. .... 67

Table 3.3: Statistical significance of the difference of the mean for each pair of clusters. Values to the right of the principal diagonal denote significance among PM<sub>10-2.5</sub> pairs. In contrast, values to the left of the principal diagonal denote significance among PM<sub>2.5</sub> pairs. Pairwise comparison based on Tukey’s studentised HSD (honestly significant differences) test was used to assess the statistical significance of the difference of the mean for each pair of clusters and each season. Values in brackets indicate the statistical significance after all rain events were removed from the data set. .... 69

Table 4.1: Characteristics of the low-cost Laser Egg (LE) air quality monitor and the reference Thermo Fisher 5014i  $\beta$ -attenuation PM analyzer. .... 87

Table 4.2: Average ( $1\sigma$  ambient variability) of PM<sub>10</sub>, PM<sub>2.5</sub>, ambient temperature, relative humidity, and absolute humidity from the reference analyzers and the two Laser Egg (LE) sensors for the calibration period from April 27, 2016, until July 25, 2016. The number in the bracket denotes one sigma ambient variability. .... 94

Table 4.3: Table showing the Root Mean Square Error (RMSE), normalized Root Mean Square Error (nRMSE), and Mean Bias Error (MBE) computed for raw and post-correction PM<sub>10</sub> and PM<sub>2.5</sub> measurements obtained from Laser Egg sensors denoted as LES1 and LES2. .... 96

Table 4.4: Value of the slope (S), intercept (I), and Pearson correlation coefficient (r) obtained from the regression of Laser Egg sensors (LE Sensor 1 and 2) with Reference (Ref) analyzers and inter-comparison of two LE sensors. % CV (coefficient of variation) depicts the relative precision between two LE sensors. RMA denotes Reduced Major axis regression, and OLS represents Ordinary Least Squares regression. .... 97

Table 4.5: Impact of using various size distributions and a density of  $2 \text{ g cm}^{-3}$  shown in Figure 4.8c-f for correction on the slope, correlation coefficient (r), normalized Root Mean Square Error (nRMSE), and Mean Bias Error (MBE) of the PM<sub>10</sub> and PM<sub>2.5</sub> measurements from the two Laser Egg (LE) sensors. .... 108



Table 4.6: Impact of using various size distributions and a density of 1.2 g cm <sup>-3</sup> shown in Figure 4.9c-f for corrections on the slope, correlation coefficient (r), normalized Root Mean Square Error (nRMSE), and Mean Bias Error (MBE) of the PM <sub>10</sub> and PM <sub>2.5</sub> measurements from the two LE sensors.....	110
Table 4.7: Impact of using various size distributions and a density of 2.7 g cm <sup>-3</sup> shown in Figure 4.10c-f for corrections on the slope, correlation coefficient (r), normalized Root Mean Square Error (nRMSE), and Mean Bias Error of the PM <sub>10</sub> and PM <sub>2.5</sub> measurements from the two LE sensors.....	112
Table 5.1: Major differences in the Airveda PM sensor and the reference Thermo Fisher 5014 β-attenuation PM analyzer.....	127
Table 5.2: Location of the Airveda low-cost sensors and the duration of the experiment. The sensors with names underlined were used as part of the training dataset for the Random Forest machine learning algorithm.....	130
Table 5.3: Slope/Intercept and correlation coefficient (r) obtained from the RMA regression between the Airveda sensor and reference PM between 20 September 2019 to 18 February 2020 when they were co-located. AV_0027* indicates the measurements obtained during the post-Nadampur deployment field calibration of the sensor when its inlet was infested with spiders (performed between 31 December 2019 and 6 January 2020). A separate RF model was used to correct the measurements of this period. RMSE, MBE, and COE stand for root mean square error, mean bias error, and coefficient of efficiency, respectively. ‘n’ refers to the number of available measurement pairs used for regression evaluation.....	131
Table 5.4: Estimates of emissions of PM <sub>10</sub> and PM <sub>2.5</sub> from agricultural residue burning compiled from different studies. ....	144

## List of Abbreviations

---

- NW-IGP North west Indo-Gangetic Plain
- PM Particulate Matter
- NAAQS National Ambient Air Quality Standard
- CPCB Central Pollution Control Board
- DMS Dimethyl sulfate
- DRH Deliquescence relative humidity
- UFP Ultrafine Particles
- ERF Effective Radiative Forcing
- CCN Cloud condensation nuclei
- SOA Secondary organic aerosol
- ERF<sub>ari</sub> Effective Radiative Forcing due to aerosol-radiation interaction
- ERF<sub>aci</sub> Effective Radiative Forcing due to aerosol-cloud interaction
- PAH Polycyclic aromatic hydrocarbons
- LULC Land use land cover
- LE Laser Egg
- RH Relative Humidity
- RH<sup>Ref</sup> RH measured from the reference sensor
- RH<sup>LE</sup> RH reported by Laser Egg
- OPC Optical Particle Counter
- BAM  $\beta$ -attenuation monitor
- TEOM Tapered Elemental Oscillating Microbalance
- CPC Condensation particle counter
- DMPS Differential mobility particle sizer
- DMA Differential mobility analyzer
- FEM Federal Equivalent Method
- FRM Federal Reference Method
- CPF Conditional probability function
- NWP Numerical weather prediction
- LPDM Lagrangian particle dispersion model
- CTM Chemical transport model
- PSCF Potential Source Contribution System
- CF Concentration field
- SR Solar Radiation
- AQI Air Quality Index

- OLS Ordinary Least Squares
- RMA Reduced Major Axis
- nRMSE Normalized Root Mean Square Error
- MBE Mean Bias Error
- COD Coefficient of divergence
- AH Absolute Humidity
- AH<sup>Ref</sup> Absolute humidity calculated using reference temperature and relative humidity
- r Pearson correlation coefficient
- r<sup>2</sup><sub>adj</sub> adjusted coefficient of determination
- LT Local time
- COE Coefficient of efficiency
- AOD Aerosol Optical Depth
- VIIRS Visible Infrared Imaging Radiometer Suite
- MODIS Moderate Resolution Imaging Spectroradiometer
- HYSPLIT Hybrid Single-Particle Lagrangian Integrated Trajectory
- GDAS Global data acquisition system
- FRP Fire Radiative Power
- FRE Fire Radiative Energy
- Ce Coefficient of Emissions
- EF Emission factor
- BLH Boundary Layer Height
- VC Ventilation Coefficient
- STILT Stochastic Time-Inverted Lagrangian Transport

## Bibliography

---

Abdalmogith, S. S., and Harrison, R. M.: The use of trajectory cluster analysis to examine the long-range transport of secondary inorganic aerosol in the UK, *Atmospheric Environment*, 39, 6686-6695, <https://doi.org/10.1016/j.atmosenv.2005.07.059>, 2005.

Adar, S. D., Filigrana, P. A., Clements, N., and Peel, J. L.: Ambient coarse particulate matter and human health: A systematic review and meta-analysis, *Curr Environ Health Rep*, 1, 258-274, <https://doi.org/10.1007/s40572-014-0022-z>, 2014.

Agnihotri, C. L., and Singh, M. S.: Satellite study of western disturbances, *Mausam*, 33, 249-254, 1982.

Akagi, S. K., Yokelson, R. J., Wiedinmyer, C., Alvarado, M. J., Reid, J. S., Karl, T., Crounse, J. D., and Wennberg, P. O.: Emission factors for open and domestic biomass burning for use in atmospheric models, *Atmos. Chem. Phys.*, 11, 4039-4072, 10.5194/acp-11-4039-2011, 2011.

Al-Kindi, S. G., Brook, R. D., Biswal, S., and Rajagopalan, S.: Environmental determinants of cardiovascular disease: lessons learned from air pollution, *Nature Reviews Cardiology*, 10.1038/s41569-020-0371-2, 2020.

Alduchov, O. A., and Eskridge, R. E.: Improved magnus form approximation of saturation vapor pressure, *Journal of Applied Meteorology*, 35, 601-609, [https://doi.org/10.1175/1520-0450\(1996\)035<0601:IMFAOS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1996)035<0601:IMFAOS>2.0.CO;2), 1996.

Allen, R. J., Hassan, T., Randles, C. A., and Su, H.: Enhanced land–sea warming contrast elevates aerosol pollution in a warmer world, *Nature Climate Change*, 9, 300-305, 10.1038/s41558-019-0401-4, 2019.

Anderson, J. O., Thundiyil, J. G., and Stolbach, A.: Clearing the air: a review of the effects of particulate matter air pollution on human health, *J Med Toxicol*, 8, 166-175, 10.1007/s13181-011-0203-1, 2012.

Andreae, M. O., Jones, C. D., and Cox, P. M.: Strong present-day aerosol cooling implies a hot future, *Nature*, 435, 1187-1190, 10.1038/nature03671, 2005.

Andreae, M. O.: Atmospheric aerosols versus greenhouse gases in the twenty-first century, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365, 1915-1923, 10.1098/rsta.2007.2051, 2007.

Aneja, V. P., Agarwal, A., Roelle, P. A., Phillips, S. B., Tong, Q., Watkins, N., and Yablonsky, R.: Measurements and analysis of criteria pollutants in New Delhi, India, *Environment International*, 27, 35-42, 2001.

Asadi, S., Bouvier, N., Wexler, A. S., and Ristenpart, W. D.: The coronavirus pandemic and aerosols: Does COVID-19 transmit via expiratory particles?, *Aerosol Science and Technology*, 54, 635-638, 10.1080/02786826.2020.1749229, 2020.

Ashbaugh, L. L., Malm, W. C., and Sadeh, W. Z.: A residence time probability analysis of sulfur concentrations at grand Canyon National Park, *Atmospheric Environment (1967)*, 19, 1263-1270, [https://doi.org/10.1016/0004-6981\(85\)90256-2](https://doi.org/10.1016/0004-6981(85)90256-2), 1985.

Asselman, J., Van Acker, E., De Rijke, M., Tilleman, L., Van Nieuwerburgh, F., Mees, J., De Schamphelaere, K. A. C., and Janssen, C. R.: Marine biogenics in sea spray aerosols interact with the mTOR signaling pathway, *Scientific Reports*, 9, 675, 10.1038/s41598-018-36866-3, 2019.

Austin, E., Novosselov, I., Seto, E., and Yost, M. G.: Laboratory evaluation of the Shinyei PPD42NS low-cost particulate matter sensor, *PLOS ONE*, 10, e0137789, <https://doi.org/10.1371/journal.pone.0141928>, 2015.

Awasthi, A., Agarwal, R., Mittal, S. K., Singh, N., Singh, K., and Gupta, P. K.: Study of size and mass distribution of particulate matter due to crop residue burning with seasonal variation in rural area of Punjab, India, *Journal of Environmental Monitoring*, 13, 1073-1081, 10.1039/C1EM10019J, 2011.

Ayers, G. P.: Comment on regression analysis of air quality data, *Atmospheric Environment*, 35, 2423-2425, [https://doi.org/10.1016/S1352-2310\(00\)00527-6](https://doi.org/10.1016/S1352-2310(00)00527-6), 2001.

Badarinath, K. V. S., Kharol, S. K., Sharma, A. R., and Krishna Prasad, V.: Analysis of aerosol and carbon monoxide characteristics over Arabian Sea during crop residue burning period in the Indo-Gangetic Plains using multi-satellite remote sensing datasets, *Journal of Atmospheric and Solar-Terrestrial Physics*, 71, 1267-1276, <https://doi.org/10.1016/j.jastp.2009.04.004>, 2009.

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., and Modzel, P.: Evaluation of low-cost sensors for ambient PM<sub>2.5</sub> monitoring, *Journal of Sensors*, 2018, 16, <https://doi.org/10.1155/2018/5096540>, 2018.

Badura, M., Batog, P., Drzeniecka-Osiadacz, A., and Modzel, P.: Regression methods in the calibration of low-cost sensors for ambient particulate matter measurements, *SN Applied Sciences*, 1, 622, 10.1007/s42452-019-0630-1, 2019.

Barkjohn, K. K., Bergin, M. H., Norris, C., Schauer, J. J., Zhang, Y., Black, M., Hu, M., and Zhang, J.: Using low-cost sensors to quantify the effects of air filtration on indoor and personal exposure relevant PM<sub>2.5</sub> concentrations in Beijing, China, *Aerosol and Air Quality Research*, 10.4209/aaqr.2018.11.0394, 2018.

Bates, T. S., Huebert, B. J., Gras, J. L., Griffiths, F. B., and Durkee, P. A.: International Global Atmospheric Chemistry (IGAC) Project's First Aerosol Characterization Experiment (ACE 1): Overview, *Journal of Geophysical Research: Atmospheres*, 103, 16297-16318, 10.1029/97JD03741, 1998.

Baumann, K., and Stohl, A.: Validation of a Long-Range Trajectory Model Using Gas Balloon Tracks from the Gordon Bennett Cup 95, *Journal of Applied Meteorology*, 36, 711-720, 10.1175/1520-0450-36.6.711, 1997.

Begum, B. A., Biswas, S. K., Markwitz, A., and Hopke, P. K.: Identification of sources of fine and coarse particulate matter in Dhaka, Bangladesh, *Aerosol and Air Quality Research*, 10, 345-353, 2010.

Beig, G., Sahu, S. K., Singh, V., Tikle, S., Sobhana, S. B., Gargeva, P., Ramakrishna, K., Rathod, A., and Murthy, B. S.: Objective evaluation of stubble emission of North India and quantifying its impact on air quality of Delhi, *Science of The Total Environment*, 709, 136126, <https://doi.org/10.1016/j.scitotenv.2019.136126>, 2020.

Belis, C. A., Larsen, B. R., Amato, F., El Haddad, I., Favez, O., Harrison, R. M., Hopke, P. K., Nava, S., Paatero, P., and Prevot, A.: European guide on air pollution source apportionment with receptor models, in, European Union, 2014.

Bell, M. L., and Davis, D. L.: Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution, *Environmental health perspectives*, 109 Suppl 3, 389-394, 10.1289/ehp.01109s3389, 2001.

Bhat, G. S.: The Indian drought of 2002—a sub-seasonal phenomenon?, *Quarterly Journal of the Royal Meteorological Society*, 132, 2583-2602, 10.1256/qj.05.13, 2006.

Bikkina, S., Andersson, A., Kirillova, E. N., Holmstrand, H., Tiwari, S., Srivastava, A. K., Bisht, D. S., and Gustafsson, Ö.: Air quality in megacity Delhi affected by countryside biomass burning, *Nature Sustainability*, 2, 200-205, 10.1038/s41893-019-0219-0, 2019.

Bohren, C. F., and Huffman, D. R.: Absorption and scattering by a sphere, Absorption and scattering of light by small particles, Wiley New York, 82-129 pp., 1983.

Borge, R., Lumberras, J., Vardoulakis, S., Kassomenos, P., and Rodríguez, E.: Analysis of long-range transport influences on urban PM<sub>10</sub> using two-stage atmospheric trajectory clusters, *Atmospheric Environment*, 41, 4434-4450, <https://doi.org/10.1016/j.atmosenv.2007.01.053>, 2007.

Boucher, O.: *Atmospheric aerosols*, *Atmospheric Aerosols*, Springer, 9-24 pp., 2015.

Bové, H., Bongaerts, E., Slenders, E., Bijnens, E. M., Saenen, N. D., Gyselaers, W., Van Eyken, P., Plusquin, M., Roeffaers, M. B. J., Ameloot, M., et al.: Ambient black carbon particles reach the fetal side of human placenta, *Nature Communications*, 10, 3866, [10.1038/s41467-019-11654-3](https://doi.org/10.1038/s41467-019-11654-3), 2019.

Bow, S.: *Pattern recognition, application to large data-set problems*, 1984, in, Marcel Dekker, Inc., New York, 1984.

Brankov, E., Rao, S. T., and Porter, P. S.: A trajectory-clustering-correlation methodology for examining the long-range transport of air pollutants, *Atmospheric Environment*, 32, 1525-1534, [https://doi.org/10.1016/S1352-2310\(97\)00388-9](https://doi.org/10.1016/S1352-2310(97)00388-9), 1998.

Brasseur, G. P., and Jacob, D. J.: *Modeling of atmospheric chemistry*, Cambridge University Press, 2017.

Brook Robert, D., Rajagopalan, S., Pope, C. A., Brook Jeffrey, R., Bhatnagar, A., Diez-Roux Ana, V., Holguin, F., Hong, Y., Luepker Russell, V., Mittleman Murray, A., et al.: Particulate Matter Air Pollution and Cardiovascular Disease, *Circulation*, 121, 2331-2378, [10.1161/CIR.0b013e3181dbee1](https://doi.org/10.1161/CIR.0b013e3181dbee1), 2010.

Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A.: Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, *Science of The Total Environment*, 518-519, 626-635, <https://doi.org/10.1016/j.scitotenv.2015.01.022>, 2015.



Brunekreef, B., and Forsberg, B.: Epidemiological evidence of effects of coarse airborne particles on health, *European Respiratory Journal*, 26, 309, 10.1183/09031936.05.00001805, 2005.

Bryson, R. A., and Swain, A. M.: Holocene variations of monsoon rainfall in Rajasthan, *Quaternary Research*, 16, 135-145, [https://doi.org/10.1016/0033-5894\(81\)90041-7](https://doi.org/10.1016/0033-5894(81)90041-7), 1981.

Buchanan, C. M., Beverland, I. J., and Heal, M. R.: The influence of weather-type and long-range transport on airborne particle concentrations in Edinburgh, UK, *Atmospheric Environment*, 36, 5343-5354, [https://doi.org/10.1016/S1352-2310\(02\)00579-4](https://doi.org/10.1016/S1352-2310(02)00579-4), 2002.

C3S: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS), in, Copernicus Climate Change Service Climate Data Store (CDS) Reading, United States 2017.

Cabello, M., Orza, J. A. G., Galiano, V., and Ruiz, G.: Influence of meteorological input data on backtrajectory cluster analysis? a seven-year study for southeastern Spain, 2008.

Calderón-Garcidueñas, L., González-Maciel, A., Mukherjee, P. S., Reynoso-Robles, R., Pérez-Guillé, B., Gayosso-Chávez, C., Torres-Jardón, R., Cross, J. V., Ahmed, I. A. M., Karloukovski, V. V., et al.: Combustion- and friction-derived magnetic air pollution nanoparticles in human hearts, *Environmental Research*, 176, 108567, <https://doi.org/10.1016/j.envres.2019.108567>, 2019.

Cappa, C. D., Onasch, T. B., Massoli, P., Worsnop, D. R., Bates, T. S., Cross, E. S., Davidovits, P., Hakala, J., Hayden, K. L., Jobson, B. T., et al.: Radiative Absorption Enhancements Due to the Mixing State of Atmospheric Black Carbon, *Science*, 337, 1078, 10.1126/science.1223447, 2012.

Carslaw, D. C., Bevers, S. D., Ropkins, K., and Bell, M. C.: Detecting and quantifying aircraft and other on-airport contributions to ambient nitrogen oxides in the vicinity of a large international airport, *Atmospheric Environment*, 40, 5424-5434, <https://doi.org/10.1016/j.atmosenv.2006.04.062>, 2006.

Carslaw, D. C., and Ropkins, K.: Openair—an R package for air quality data analysis, *Environmental Modelling & Software*, 27, 52-61, 2012.

Carslaw, D. C., and Beever, S. D.: Characterising and understanding emission sources using bivariate polar plots and k-means clustering, *Environmental Modelling & Software*, 40, 325-329, <https://doi.org/10.1016/j.envsoft.2012.09.005>, 2013.

Carter, A. W., Zaitchik, B. F., Gohlke, J. M., Wang, S., and Richardson, M. B.: Methods for Estimating Wet Bulb Globe Temperature From Remote and Low-Cost Data: A Comparative Study in Central Alabama, 4, e2019GH000231, <https://doi.org/10.1029/2019GH000231>, 2020.

Chan, Y. C., Simpson, R. W., McTainsh, G. H., Vowles, P. D., Cohen, D. D., and Bailey, G. M.: Source apportionment of PM<sub>2.5</sub> and PM<sub>10</sub> aerosols in Brisbane (Australia) by receptor modelling, *Atmospheric Environment*, 33, 3251-3268, [https://doi.org/10.1016/S1352-2310\(99\)00090-4](https://doi.org/10.1016/S1352-2310(99)00090-4), 1999.

Chandra, B. P., and Sinha, V.: Contribution of post-harvest agricultural paddy residue fires in the N.W. Indo-Gangetic Plain to ambient carcinogenic benzenoids, toxic isocyanic acid and carbon monoxide, *Environment International*, 88, 187-197, <https://doi.org/10.1016/j.envint.2015.12.025>, 2016.

Charlson, R. J.: Giants' footprints in the greenhouse, *Nature*, 401, 741-742, 10.1038/44456, 1999.

Chen, W., Tong, D. Q., Zhang, S., Zhang, X., and Zhao, H.: Local PM<sub>10</sub> and PM<sub>2.5</sub> emission inventories from agricultural tillage and harvest in northeastern China, *Journal of Environmental Sciences*, 57, 15-23, <https://doi.org/10.1016/j.jes.2016.02.024>, 2017.

Chien, C.-H., Theodore, A., Wu, C.-Y., Hsu, Y.-M., and Birky, B.: Upon correlating diameters measured by optical particle counters and aerodynamic particle sizers, *Journal of Aerosol Science*, 101, 77-85, <https://doi.org/10.1016/j.jaerosci.2016.05.011>, 2016.

---

Guidelines for ambient air quality monitoring [http://www.indiaairquality.info/wp-content/uploads/docs/2003\\_CPCB\\_Guidelines\\_for\\_Air\\_Monitoring.pdf](http://www.indiaairquality.info/wp-content/uploads/docs/2003_CPCB_Guidelines_for_Air_Monitoring.pdf), access: 18 January 2020, 2003.

List of stations with Continuous Automatic Air Quality Monitoring Stations (CAAQMS): <https://app.cpcbcr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/caaqm-data-availability>, access: January 18, 2020, 2019a.

Operating stations under the National Air Monitoring Programme (NAMP): [https://cpcb.nic.in/uploads/Stations\\_NAMP.pdf](https://cpcb.nic.in/uploads/Stations_NAMP.pdf), access: January 18, 2020, 2019b.

Crilley, L. R., Shaw, M., Pound, R., Kramer, L. J., Price, R., Young, S., Lewis, A. C., and Pope, F. D.: Evaluation of a low-cost optical particle counter (Alphasense OPC-N2) for ambient air monitoring, *Atmospheric Measurement Techniques*, 11, 709-720, <https://doi.org/10.5194/amt-11-709-2018>, 2018.

Cross, E. S., Williams, L. R., Lewis, D. K., Magoon, G. R., Onasch, T. B., Kaminsky, M. L., Worsnop, D. R., and Jayne, J. T.: Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements, *Atmospheric Measurement Techniques*, 10, 3575-3588, <https://doi.org/10.5194/amt-10-3575-2017>, 2017.

Crutzen, P. J.: The “anthropocene”, in: *Earth system science in the anthropocene*, Springer, 13-18, 2006.

Cusworth, D. H., Mickley, L. J., Sulprizio, M. P., Liu, T., Marlier, M. E., DeFries, R. S., Guttikunda, S. K., and Gupta, P.: Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India, *Environmental Research Letters*, 13, 044018, [10.1088/1748-9326/aab303](https://doi.org/10.1088/1748-9326/aab303), 2018.

Das, P. K.: Mean vertical motion and non-adiabatic heat sources over India during the monsoon, *Tellus*, 14, 212-220, [10.1111/j.2153-3490.1962.tb00132.x](https://doi.org/10.1111/j.2153-3490.1962.tb00132.x), 1962.

---

De Gouw, J., and Jimenez, J. L.: Organic aerosols in the Earth's atmosphere, in, ACS Publications, 2009.

Dey, S., and Tripathi, S. N.: Estimation of aerosol optical properties and radiative effects in the Ganga basin, northern India, during the wintertime, *Journal of Geophysical Research: Atmospheres*, 112, 10.1029/2006JD007267, 2007.

Dimri, A. P.: Impact of horizontal model resolution and orography on the simulation of a western disturbance and its associated precipitation, *Meteorological Applications*, 11, 115-127, 10.1017/S1350482704001227, 2004.

Dockery, D. W., Speizer, F. E., Stram, D. O., Ware, J. H., Spengler, J. D., and Ferris, B. G.: Effects of Inhalable Particles on Respiratory Health of Children, *American Review of Respiratory Disease*, 139, 587-594, 10.1164/ajrccm/139.3.587, 1989.

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E.: An Association between Air Pollution and Mortality in Six U.S. Cities, *New England Journal of Medicine*, 329, 1753-1759, 10.1056/NEJM199312093292401, 1993.

Dorling, S. R., Davies, T. D., and Pierce, C. E.: Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry—Method and applications, *Atmospheric Environment. Part A. General Topics*, 26, 2575-2581, [https://doi.org/10.1016/0960-1686\(92\)90110-7](https://doi.org/10.1016/0960-1686(92)90110-7), 1992.

Draxler, R., and Rolph, G.: HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory), NOAA Air Resources Laboratory, College Park, MD, Model access via NOAA ARL READY Website, in, 2013.

Draxler, R. R., and Hess, G.: An overview of the HYSPLIT\_4 modelling system for trajectories, *Australian meteorological magazine*, 47, 295-308, 1998.

---

Draxler, R. R.: Evaluation of an ensemble dispersion calculation, *Journal of Applied Meteorology*, 42, 308-317, 2003.

Duhanyan, N., and Roustan, Y.: Below-cloud scavenging by rain of atmospheric gases and particulates, *Atmospheric Environment*, 45, 7201-7217, <https://doi.org/10.1016/j.atmosenv.2011.09.002>, 2011.

Eatough, D. J., Long, R. W., Modey, W. K., and Eatough, N. L.: Semi-volatile secondary organic aerosol in urban atmospheres: meeting a measurement challenge, *Atmospheric Environment*, 37, 1277-1292, [https://doi.org/10.1016/S1352-2310\(02\)01020-8](https://doi.org/10.1016/S1352-2310(02)01020-8), 2003.

Englert, N.: Fine particles and human health—a review of epidemiological studies, *Toxicology Letters*, 149, 235-242, <https://doi.org/10.1016/j.toxlet.2003.12.035>, 2004.

EPA: Appendix A to Part 58 of 40 CFR - Ambient Air Quality Surveillance 2015a.

EPA: Appendix A to Part 58 of 40 CFR - Ambient Air Quality Surveillance (Subchapter C). 2015b.

EPA: List of designated reference and equivalent methods, December 15, 2019., 2019.

Ervens, B., Carlton, A. G., Turpin, B. J., Altieri, K. E., Kreidenweis, S. M., and Feingold, G.: Secondary organic aerosol yields from cloud-processing of isoprene oxidation products, *Geophysical Research Letters*, 35, 10.1029/2007GL031828, 2008.

ESOPB: District Domestic Product of Punjab, Economic and Statistical Organization, Government of Punjab, 2015.

Firket, J.: The cause of the symptoms found in the Meuse Valley during the fog of December, 1930, *Bull Acad R Med Belg*, 11, 683-741, 1931.

Fleming, Z. L., Monks, P. S., and Manning, A. J.: Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition, *Atmospheric Research*, 104-105, 1-39, <https://doi.org/10.1016/j.atmosres.2011.09.009>, 2012.

Foroutan, H., Young, J., Napelenok, S., Ran, L., Appel, K. W., Gilliam, R. C., and Pleim, J. E.: Development and evaluation of a physics-based windblown dust emission scheme implemented in the CMAQ modeling system, 9, 585-608, <https://doi.org/10.1002/2016MS000823>, 2017.

Ganguly, D., Jayaraman, A., Rajesh, T. A., and Gadhavi, H.: Wintertime aerosol properties during foggy and nonfoggy days over urban center Delhi and their implications for shortwave radiative forcing, *Journal of Geophysical Research: Atmospheres*, 111, 10.1029/2005JD007029, 2006.

Garg, S., and Sinha, B.: Determining the contribution of long-range transport, regional and local source areas, to PM10 mass loading in Hessen, Germany using a novel multi-receptor based statistical approach, *Atmospheric Environment*, 167, 566-575, <https://doi.org/10.1016/j.atmosenv.2017.08.029>, 2017.

Gautam, R., Hsu, N. C., Tsay, S. C., Lau, K. M., Holben, B., Bell, S., Smirnov, A., Li, C., Hansell, R., Ji, Q., et al.: Accumulation of aerosols over the Indo-Gangetic plains and southern slopes of the Himalayas: distribution, properties and radiative effects during the 2009 pre-monsoon season, *Atmos. Chem. Phys.*, 11, 12841-12863, 10.5194/acp-11-12841-2011, 2011.

Gebhart, K. A., Schichtel, B. A., and Barna, M. G.: Directional Biases in Back Trajectories Caused by Model and Input Data, *Journal of the Air & Waste Management Association*, 55, 1649-1662, 10.1080/10473289.2005.10464758, 2005.

George, I. J., and Abbatt, J. P. D.: Heterogeneous oxidation of atmospheric aerosol particles by gas-phase radicals, *Nature Chemistry*, 2, 713-722, 10.1038/nchem.806, 2010.

Gildemeister, A. E., Hopke, P. K., and Kim, E.: Sources of fine urban particulate matter in Detroit, MI, *Chemosphere*, 69, 1064-1074, <https://doi.org/10.1016/j.chemosphere.2007.04.027>, 2007.

Goswami, B. N.: Interannual Variations of Indian Summer Monsoon in a GCM: External Conditions versus Internal Feedbacks, *Journal of Climate*, 11, 501-522, 10.1175/1520-0442(1998)011<0501:IVOISM>2.0.CO;2, 1998.

Goswami, B. N., Venugopal, V., Sengupta, D., Madhusoodanan, M. S., and Xavier, P. K.: Increasing Trend of Extreme Rain Events Over India in a Warming Environment, 314, 1442-1445, 10.1126/science.1132027 %J *Science*, 2006a.

Goswami, B. N., Wu, G., and Yasunari, T.: The Annual Cycle, Intraseasonal Oscillations, and Roadblock to Seasonal Predictability of the Asian Summer Monsoon, 19, 5078-5099, 10.1175/jcli3901.1, 2006b.

Grivas, G., Chaloulakou, A., and Kassomenos, P.: An overview of the PM10 pollution problem, in the Metropolitan Area of Athens, Greece. Assessment of controlling factors and potential impact of long range transport, *Science of The Total Environment*, 389, 165-177, <https://doi.org/10.1016/j.scitotenv.2007.08.048>, 2008.

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions, *Geosci. Model Dev.*, 5, 1471-1492, 10.5194/gmd-5-1471-2012, 2012.

Guttikunda, S. K., and Gurjar, B. R.: Role of meteorology in seasonality of air pollution in megacity Delhi, India, *Environmental Monitoring and Assessment*, 184, 3199-3211, 10.1007/s10661-011-2182-8, 2012.

Haagen-Smit, A. J.: Chemistry and physiology of Los Angeles smog, *Industrial & Engineering Chemistry*, 44, 1342-1346, 1952.

Habib, G., Venkataraman, C., Shrivastava, M., Banerjee, R., Stehr, J. W., and Dickerson, R. R.: New methodology for estimating biofuel consumption for cooking: Atmospheric emissions of

black carbon and sulfur dioxide from India, *Global Biogeochemical Cycles*, 18, 10.1029/2003GB002157, 2004.

Hagan, D. H., and Kroll, J. H.: Assessing the accuracy of low-cost optical particle sensors using a physics-based approach, *Atmos. Meas. Tech.*, 13, 6343-6355, 10.5194/amt-13-6343-2020, 2020.

Hakkim, H., Sinha, V., Chandra, B. P., Kumar, A., Mishra, A. K., Sinha, B., Sharma, G., Pawar, H., Sohpaal, B., Ghude, S. D., et al.: Volatile organic compound measurements point to fog-induced biomass burning feedback to air quality in the megacity of Delhi, *Science of The Total Environment*, 689, 295-304, <https://doi.org/10.1016/j.scitotenv.2019.06.438>, 2019.

Hama, S. M. L., Kumar, P., Harrison, R. M., Bloss, W. J., Khare, M., Mishra, S., Namdeo, A., Sokhi, R., Goodman, P., and Sharma, C.: Four-year assessment of ambient particulate matter and trace gases in the Delhi-NCR region of India, *Sustainable Cities and Society*, 54, 102003, <https://doi.org/10.1016/j.scs.2019.102003>, 2020.

Hammer, Ø., Harper, D. A., and Ryan, P. D.: PAST: paleontological statistics software package for education and data analysis, *Palaeontologia Electronica* 4, 9, 2001a.

Hammer, Ø., Harper, D. A., and Ryan, P. D.: PAST: paleontological statistics software package for education and data analysis, *Palaeontologia Electronica* 4, art. 4: 9 pages, 2001b.

Hangal, S., and Willeke, K.: Aspiration efficiency: unified model for all forward sampling angles, *Environmental Science & Technology*, 24, 688-691, <https://doi.org/10.1021/es00075a012>, 1990.

Harris, E., Sinha, B., van Pinxteren, D., Schneider, J., Poulain, L., Collett, J., D'Anna, B., Fahlbusch, B., Foley, S., Fomba, K. W., et al.: In-cloud sulfate addition to single particles resolved with sulfur isotope analysis during HCCT-2010, *Atmos. Chem. Phys.*, 14, 4219-4235, 10.5194/acp-14-4219-2014, 2014.



---

Harrison, R. M., Laxen, D., Moorcroft, S., and Laxen, K.: Processes affecting concentrations of fine particulate matter (PM<sub>2.5</sub>) in the UK atmosphere, *Atmospheric Environment*, 46, 115-124, <https://doi.org/10.1016/j.atmosenv.2011.10.028>, 2012.

Holstius, D. M., Pillarisetti, A., Smith, K. R., and Seto, E.: Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California, *Atmospheric Measurement Techniques*, 7, 1121-1131, <https://doi.org/10.5194/amt-7-1121-2014>, 2014.

Hopke, P. K.: Review of receptor modeling methods for source apportionment, *Journal of the Air & Waste Management Association*, 66, 237-259, 10.1080/10962247.2016.1140693, 2016.

Hyvärinen, A. P., Raatikainen, T., Komppula, M., Mielonen, T., Sundström, A. M., Brus, D., Panwar, T. S., Hooda, R. K., Sharma, V. P., de Leeuw, G., et al.: Effect of the summer monsoon on aerosols at two measurement stations in Northern India – Part 2: Physical and optical properties, *Atmos. Chem. Phys.*, 11, 8283-8294, 10.5194/acp-11-8283-2011, 2011a.

Hyvärinen, A. P., Raatikainen, T., Komppula, M., Mielonen, T., Sundström, A. M., Brus, D., Panwar, T. S., Hooda, R. K., Sharma, V. P., de Leeuw, G., et al.: Effect of the summer monsoon on aerosols at two measurement stations in Northern India – Part 2: Physical and optical properties, *Atmospheric Chemistry and Physics*, 11, 8283-8294, <https://doi.org/10.5194/acp-11-8283-2011>, 2011b.

Ichoku, C., and Ellison, L.: Global top-down smoke-aerosol emissions estimation using satellite fire radiative power measurements, *Atmos. Chem. Phys.*, 14, 6643-6667, 10.5194/acp-14-6643-2014, 2014.

India, C. o.: District Census Handbook, Sangrur Directorate of Census Operations, Punjab Series-04, Part XII-B, 2011.

Jacobson, M. Z.: Strong radiative heating due to the mixing state of black carbon in atmospheric aerosols, *Nature*, 409, 695-697, 10.1038/35055518, 2001.

---

Jain, N., Bhatia, A., and Pathak, H.: Emission of Air Pollutants from Crop Residue Burning in India, *Aerosol and Air Quality Research*, 14, 422-430, 10.4209/aaqr.2013.01.0031, 2014.

Jandacka, D., and Durcanska, D.: Differentiation of Particulate Matter Sources Based on the Chemical Composition of PM10 in Functional Urban Areas, 10, 583, 2019.

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., and Morawska, L.: The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog, *Atmospheric Measurement Techniques*, 11, 4883-4890, <https://doi.org/10.5194/amt-11-4883-2018>, 2018.

Jethva, H., Satheesh, S. K., and Srinivasan, J.: Seasonal variability of aerosols over the Indo-Gangetic basin, *Journal of Geophysical Research: Atmospheres*, 110, 10.1029/2005JD005938, 2005.

Jethva, H., Torres, O., Field, R. D., Lyapustin, A., Gautam, R., and Kayetha, V.: Connecting Crop Productivity, Residue Fires, and Air Quality over Northern India, *Scientific Reports*, 9, 16594, 10.1038/s41598-019-52799-x, 2019.

Jiao, W., Hagler, G., Williams, R., Sharpe, R., Brown, R., Garver, D., Judge, R., Caudill, M., Rickard, J., Davis, M., et al.: Community Air Sensor Network (CAIRSENSE) project: evaluation of low-cost sensor performance in a suburban environment in the southeastern United States, *Atmospheric Measurement Techniques*, 9, 5281-5292, <https://doi.org/10.5194/amt-9-5281-2016>, 2016.

Johnson, K. K., Bergin, M. H., Russell, A. G., and Hagler, G. S. W.: Field test of several low-cost particulate matter sensors in high and low concentration urban environments, *Aerosol and Air Quality Research*, 18, 565-578, 10.4209/aaqr.2017.10.0418, 2018.

Jones, R. M., and Brosseau, L. M.: Aerosol transmission of infectious disease, *Journal of occupational and environmental medicine*, 57, 501-508, 2015.

Joseph, P. V., and Raman, P. L.: Existence of low level westerly jet stream over peninsular India during July, *Indian J Meteorol Geophys*, 17, 407-410, 1966.

Joseph, P. V.: A tentative model of Andhi, *Mausam*, 33, 417, 1982.

Joseph, P. V., and Sijikumar, S.: Intraseasonal variability of the low-level jet stream of the Asian summer monsoon, *Journal of Climate*, 17, 1449-1458, 2004.

Kahl, J. D., Harris, J. M., Herbert, G. A., and Olson, M. P.: Intercomparison of three long-range trajectory models applied to Arctic haze, *Tellus B*, 41B, 524-536, 10.1111/j.1600-0889.1989.tb00139.x, 1989.

Kahl, J. D.: A cautionary note on the use of air trajectories in interpreting atmospheric chemistry measurements, *Atmospheric Environment. Part A. General Topics*, 27, 3037-3038, [https://doi.org/10.1016/0960-1686\(93\)90336-W](https://doi.org/10.1016/0960-1686(93)90336-W), 1993.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., et al.: The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, 77, 437-472, 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kanamitsu, M.: Description of the NMC Global Data Assimilation and Forecast System, *Weather and Forecasting*, 4, 335-342, 10.1175/1520-0434(1989)004<0335:DOTNGD>2.0.CO;2, 1989.

Kandlikar, M.: Air pollution at a hotspot location in Delhi: detecting trends, seasonal cycles and oscillations, *Atmospheric Environment*, 41, 5934-5947, 2007.

Kappos, A. D., Bruckmann, P., Eikmann, T., Englert, N., Heinrich, U., Höpfe, P., Koch, E., Krause, G. H. M., Kreyling, W. G., and Rauchfuss, K.: Health effects of particles in ambient air, *International journal of hygiene and environmental health*, 207, 399-407, 2004.

Karagulian, F., Barbieri, M., Kotsev, A., Spinelle, L., Gerboles, M., Lagler, F., Redon, N., Crunaire, S., and Borowiak, A.: Review of the performance of low-cost sensors for air quality monitoring, *Atmosphere*, 10, <https://doi.org/10.3390/atmos10090506>, 2019.

Kaskaoutis, D. G., Singh, R. P., Gautam, R., Sharma, M., Kosmopoulos, P. G., and Tripathi, S. N.: Variability and trends of aerosol properties over Kanpur, northern India using AERONET data (2001–10), *Environmental Research Letters*, 7, 024003, 10.1088/1748-9326/7/2/024003, 2012.

Kaskaoutis, D. G., Sinha, P. R., Vinoj, V., Kosmopoulos, P. G., Tripathi, S. N., Misra, A., Sharma, M., and Singh, R. P.: Aerosol properties and radiative forcing over Kanpur during severe aerosol loading conditions, *Atmospheric environment*, 79, 7-19, 2013.

Kassomenos, P., Vardoulakis, S., Borge, R., Lumberras, J., Papaloukas, C., and Karakitsios, S.: Comparison of statistical clustering techniques for the classification of modelled atmospheric trajectories, *Theoretical and applied climatology*, 102, 1-12, 2010.

Kaushik, C. P., Ravindra, K., Yadav, K., Mehta, S., and Haritash, A. K.: Assessment of ambient air quality in urban centres of Haryana (India) in relation to different anthropogenic activities and health risks, *Environmental monitoring and assessment*, 122, 27-40, 2006.

Kelly, K. E., Whitaker, J., Petty, A., Widmer, C., Dybwad, A., Sleeth, D., Martin, R., and Butterfield, A.: Ambient and laboratory evaluation of a low-cost particulate matter sensor, *Environmental Pollution*, 221, 491-500, <https://doi.org/10.1016/j.envpol.2016.12.039>, 2017.

Kim, E., and Hopke, P. K.: Source Apportionment of Fine Particles in Washington, DC, Utilizing Temperature-Resolved Carbon Fractions, *Journal of the Air & Waste Management Association*, 54, 773-785, 10.1080/10473289.2004.10470948, 2004.

Kim, K.-H., Kabir, E., and Kabir, S.: A review on the human health impact of airborne particulate matter, *Environment International*, 74, 136-143, <https://doi.org/10.1016/j.envint.2014.10.005>, 2015.

Kim, M., Deshpande, S. R., and Crist, K. C.: Source apportionment of fine particulate matter (PM<sub>2.5</sub>) at a rural Ohio River Valley site, *Atmospheric Environment*, 41, 9231-9243, <https://doi.org/10.1016/j.atmosenv.2007.07.061>, 2007.

Kingma, B. R. M., Frijns, A. J. H., Schellen, L., and van Marken Lichtenbelt, W. D.: Beyond the classic thermoneutral zone, *Temperature*, 1, 142-149, 10.4161/temp.29702, 2014.

Kuht, J., and Farmery, A. D.: Body temperature and its regulation, *Anaesthesia & Intensive Care Medicine*, 15, 273-278, 2014.

Kulkarni, P., Baron, P. A., and Willeke, K.: *Aerosol measurement: principles, techniques, and applications*, John Wiley & Sons, 2011.

Kulkarni, S. H., Ghude, S. D., Jena, C., Karumuri, R. K., Sinha, B., Sinha, V., Kumar, R., Soni, V. K., and Khare, M.: How Much Does Large-Scale Crop Residue Burning Affect the Air Quality in Delhi?, *Environmental Science & Technology*, 10.1021/acs.est.0c00329, 2020a.

Kulkarni, S. H., Ghude, S. D., Jena, C., Karumuri, R. K., Sinha, B., Sinha, V., Kumar, R., Soni, V. K., and Khare, M.: How Much Does Large-Scale Crop Residue Burning Affect the Air Quality in Delhi?, *Environmental Science & Technology*, 54, 4790-4799, 10.1021/acs.est.0c00329, 2020b.

Kulshrestha, U. C., Saxena, A., Kumar, N., Kumari, K. M., and Srivastava, S. S.: Chemical composition and association of size-differentiated aerosols at a suburban site in a semi-arid tract of India, *Journal of Atmospheric Chemistry*, 29, 109-118, 1998.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., and Britter, R.: The rise of low-cost sensing for managing air pollution in cities, *Environment International*, 75, 199-205, <https://doi.org/10.1016/j.envint.2014.11.019>, 2015.

Kumar, R., Srivastava, S. S., and Kumari, K. M.: Characteristics of aerosols over suburban and urban site of semiarid region in India: seasonal and spatial variations, *Aerosol Air Qual. Res.*, 7, 531-549, 2007.

Kumar, R., Lee, J.-H., Shie, R.-H., Chio, C.-P., and Chan, C.-C.: Patterns and Sources of PM10 in the Ecologically Sensitive Himalayan Region in Himachal Pradesh, India, *Aerosol and Air Quality Research*, 20, 410-418, 2020.

Kumar, S., Sharma, D. K., Singh, D. R., Biswas, H., Praveen, K. V., and Sharma, V.: Estimating loss of ecosystem services due to paddy straw burning in North-west India, *International Journal of Agricultural Sustainability*, 17, 146-157, 10.1080/14735903.2019.1581474, 2019.

Kumar, V., Sarkar, C., and Sinha, V.: Influence of post-harvest crop residue fires on surface ozone mixing ratios in the N.W. IGP analyzed using 2 years of continuous in situ trace gas measurements, 121, 3619-3633, <https://doi.org/10.1002/2015JD024308>, 2016.

Lang, T. J., and Barros, A. P.: Winter storms in the central Himalayas, *Journal of the Meteorological Society of Japan. Ser. II*, 82, 829-844, 2004.

Larson, E. J. L., and Portmann, R. W.: Anthropogenic aerosol drives uncertainty in future climate mitigation efforts, *Scientific Reports*, 9, 16538, 10.1038/s41598-019-52901-3, 2019.

Lawrence, M. G., and Lelieveld, J.: Atmospheric pollutant outflow from southern Asia: a review, *Atmospheric Chemistry and Physics*, 10, 11017, 2010.

Lee, E., Chan, C. K., and Paatero, P.: Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong, *Atmospheric Environment*, 33, 3201-3212, 1999.

Legates, D. R., and McCabe Jr, G. J.: Evaluating the use of “goodness-of-fit” Measures in hydrologic and hydroclimatic model validation, *Water Resources Research*, 35, 233-241, 10.1029/1998WR900018, 1999.

Legates, D. R., and McCabe, G. J.: A refined index of model performance: a rejoinder, *International Journal of Climatology*, 33, 1053-1056, 10.1002/joc.3487, 2013.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, *Nature*, 525, 367-371, 10.1038/nature15371, 2015.

Levy Zamora, M., Xiong, F., Gentner, D., Kerkez, B., Kohrman-Glaser, J., and Koehler, K.: Field and laboratory evaluations of the low-cost plantower particulate matter sensor, *Environmental Science & Technology*, 53, 838-849, <https://doi.org/10.1021/acs.est.8b05174>, 2019.

Lewis, S. L., and Maslin, M. A.: Defining the Anthropocene, *Nature*, 519, 171-180, 10.1038/nature14258, 2015.

Li, C., Bosch, C., Kang, S., Andersson, A., Chen, P., Zhang, Q., Cong, Z., Chen, B., Qin, D., and Gustafsson, Ö.: Sources of black carbon to the Himalayan–Tibetan Plateau glaciers, *Nature Communications*, 7, 12574, 10.1038/ncomms12574, 2016.

Lim, H.-J., Carlton, A. G., and Turpin, B. J.: Isoprene Forms Secondary Organic Aerosol through Cloud Processing: Model Simulations, *Environmental Science & Technology*, 39, 4441-4446, 10.1021/es048039h, 2005.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M. A., Amann, M., Anderson, H. R., Andrews, K. G., et al.: A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, *The Lancet*, 380, 2224-2260, [https://doi.org/10.1016/S0140-6736\(12\)61766-8](https://doi.org/10.1016/S0140-6736(12)61766-8), 2012.

Lin, M., Chen, Y., Burnett, R. T., Villeneuve, P. J., and Krewski, D.: The influence of ambient coarse particulate matter on asthma hospitalization in children: case-crossover and time-series analyses, *Environmental Health Perspectives*, 110, 575-581, 10.1289/ehp.02110575, 2002.

Liu, H.-Y., Schneider, P., Haugen, R., and Vogt, M.: Performance assessment of a low-cost PM<sub>2.5</sub> sensor for a near four-month period in Oslo, Norway, *Atmosphere*, 10, <https://doi.org/10.3390/atmos10020041>, 2019a.

Liu, T., Marlier, M. E., DeFries, R. S., Westervelt, D. M., Xia, K. R., Fiore, A. M., Mickley, L. J., Cusworth, D. H., and Milly, G.: Seasonal impact of regional outdoor biomass burning on air pollution in three Indian cities: Delhi, Bengaluru, and Pune, *Atmospheric Environment*, 172, 83-92, <https://doi.org/10.1016/j.atmosenv.2017.10.024>, 2018.

Liu, T., Marlier, M. E., Karambelas, A., Jain, M., Singh, S., Singh, M. K., Gautam, R., and DeFries, R. S.: Missing emissions from post-monsoon agricultural fires in northwestern India: regional limitations of MODIS burned area and active fire products, *Environmental Research Communications*, 1, 011007, 10.1088/2515-7620/ab056c, 2019b.

Liu, T., Mickley, L. J., Singh, S., Jain, M., DeFries, R. S., and Marlier, M. E.: Crop residue burning practices across north India inferred from household survey data: Bridging gaps in satellite observations, *Atmospheric Environment: X*, 100091, <https://doi.org/10.1016/j.aeaoa.2020.100091>, 2020.

Logan, W. P. D.: Mortality in the London fog incident, 1952, *Lancet*, 336-338, 1953.

Lohan, S. K., Jat, H. S., Yadav, A. K., Sidhu, H. S., Jat, M. L., Choudhary, M., Peter, J. K., and Sharma, P. C.: Burning issues of paddy residue management in north-west states of India, *Renewable and Sustainable Energy Reviews*, 81, 693-706, <https://doi.org/10.1016/j.rser.2017.08.057>, 2018.

Lundborg, M., Johard, U., Låstbom, L., Gerde, P., and Camner, P.: Human Alveolar Macrophage Phagocytic Function is Impaired by Aggregates of Ultrafine Carbon Particles, *Environmental Research*, 86, 244-253, <https://doi.org/10.1006/enrs.2001.4269>, 2001.

Maffia, J., Dinuccio, E., Amon, B., and Balsari, P.: PM emissions from open field crop management: Emission factors, assessment methods and mitigation measures – A review, *Atmospheric Environment*, 226, 117381, <https://doi.org/10.1016/j.atmosenv.2020.117381>, 2020.



Mahal, J. S., Manes, G. S., Dixit, A., Verma, A., and Singh, A.: Development of a conveyor seeder for direct sowing of wheat in combine-harvested rice field, *Agricultural Research Journal*, 53, 421-424, 2016.

Maher, B. A., Ahmed, I. A. M., Karloukovski, V., MacLaren, D. A., Foulds, P. G., Allsop, D., Mann, D. M. A., Torres-Jardón, R., and Calderon-Garciduenas, L.: Magnetite pollution nanoparticles in the human brain, *Proceedings of the National Academy of Sciences*, 113, 10797, 10.1073/pnas.1605941113, 2016.

Makar, M., Antonelli, J., Di, Q., Cutler, D., Schwartz, J., and Dominici, F.: Estimating the Causal Effect of Low Levels of Fine Particulate Matter on Hospitalization, *Epidemiology*, 28, 627-634, 10.1097/EDE.0000000000000690, 2017.

Mandal, P., Sarkar, R., Mandal, A., and Saud, T.: Seasonal variation and sources of aerosol pollution in Delhi, India, *Environmental Chemistry Letters*, 12, 529-534, 10.1007/s10311-014-0479-x, 2014.

Manikonda, A., Zíková, N., Hopke, P. K., and Ferro, A. R.: Laboratory assessment of low-cost PM monitors, *Journal of Aerosol Science*, 102, 29-40, <https://doi.org/10.1016/j.jaerosci.2016.08.010>, 2016.

Massey, D., Masih, J., Kulshrestha, A., Habil, M., and Taneja, A.: Indoor/outdoor relationship of fine particles less than 2.5  $\mu\text{m}$  (PM<sub>2.5</sub>) in residential homes locations in central Indian region, *Building and Environment*, 44, 2037-2045, 2009.

Masson-Delmotte, V.: Global Warming of 1.5 oC: An IPCC Special Report on the impacts of global warming of 1.5 C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, World Meteorological Organization, 2018.

---

Massoud, R., Shihadeh, A. L., Roumié, M., Youness, M., Gerard, J., Saliba, N., Zaarour, R., Abboud, M., Farah, W., and Saliba, N. A.: Intraurban variability of PM10 and PM2.5 in an Eastern Mediterranean city, *Atmospheric Research*, 101, 893-901, <https://doi.org/10.1016/j.atmosres.2011.05.019>, 2011.

McConnell, J. R., Edwards, R., Kok, G. L., Flanner, M. G., Zender, C. S., Saltzman, E. S., Banta, J. R., Pasteris, D. R., Carter, M. M., and Kahl, J. D. W.: 20th-Century Industrial Black Carbon Emissions Altered Arctic Climate Forcing, *Science*, 317, 1381, 10.1126/science.1144856, 2007.

Miller, L., Farhana, S., and Xu, X.: Trans-boundary air pollution in Windsor, Ontario (Canada), *Procedia Environmental Sciences*, 2, 585-594, 2010.

Mishra, A. K., and Shibata, T.: Synergistic analyses of optical and microphysical properties of agricultural crop residue burning aerosols over the Indo-Gangetic Basin (IGB), *Atmospheric Environment*, 57, 205-218, 2012.

Mohanraj, R., and Azeez, P. A.: Health effects of airborne particulate matter and the Indian scenario, *Current science*, 87, 741-748, 2004.

Molina, M. J., and Rowland, F. S.: Stratospheric sink for chlorofluoromethanes: chlorine atom-catalysed destruction of ozone, *Nature*, 249, 810-812, 10.1038/249810a0, 1974.

Mooley, D. A.: The role of western disturbances in the production of weather over India during different seasons, *Indian J. Meteor. Geophys*, 8, 253-260, 1957.

Mukherjee, A., Stanton, G. L., Graham, R. A., and Roberts, T. P.: Assessing the utility of low-cost particulate matter sensors over a 12-Week period in the Cuyama valley of California, *Sensors*, 17, <https://doi.org/10.3390/s17081805>, 2017.

Murphy, D. M., Cziczo, D. J., Hudson, P. K., Schein, M. E., and Thomson, D. S.: Particle density inferred from simultaneous optical and aerodynamic diameters sorted by composition, *Journal of Aerosol Science*, 35, 135-139, [https://doi.org/10.1016/S0021-8502\(03\)00386-0](https://doi.org/10.1016/S0021-8502(03)00386-0), 2004.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestedt, J., Huang, J., Koch, D., Lamarque, J.-F., D. Lee, B. M., Nakajima, T., et al.: Anthropogenic and Natural Radiative Forcing, in: Climate Change 2013 – The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Intergovernmental Panel on Climate, C., Cambridge University Press, Cambridge, 659-740, 2014.

NAAQS: National Ambient Air Quality Standard 2009, The Gazette of India, Extraordinary, Kartika 27, 1931, Government of India, New Delhi, 2009.

Nag, S., Gupta, A. K., and Mukhopadhyay, U. K.: Size distribution of atmospheric aerosols in Kolkata, India and the assessment of pulmonary deposition of particle mass, *Indoor and Built Environment*, 14, 381-389, 2005.

Nagar, P. K., Singh, D., Sharma, M., Kumar, A., Aneja, V. P., George, M. P., Agarwal, N., and Shukla, S. P.: Characterization of PM<sub>2.5</sub> in Delhi: role and impact of secondary aerosol, burning of biomass, and municipal solid waste and crustal matter, *Environmental Science and Pollution Research*, 24, 25179-25189, 10.1007/s11356-017-0171-3, 2017.

Nakayama, T., Matsumi, Y., Kawahito, K., and Watabe, Y.: Development and evaluation of a palm-sized optical PM<sub>2.5</sub> sensor, *Aerosol Science and Technology*, 52, 2-12, <https://doi.org/10.1080/02786826.2017.1375078>, 2018.

Nemery, B., Hoet, P. H. M., and Nemmar, A.: The Meuse Valley fog of 1930: an air pollution disaster, *The Lancet*, 357, 704-708, 2001.

Nessler, R., Weingartner, E., and Baltensperger, U.: Effect of humidity on aerosol light absorption and its implications for extinction and the single scattering albedo illustrated for a site in the lower free troposphere, *Journal of Aerosol Science*, 36, 958-972, <https://doi.org/10.1016/j.jaerosci.2004.11.012>, 2005.

Norris, G., Duvall, R., Brown, S., and Bai, S.: Epa positive matrix factorization (pmf) 5.0 fundamentals and user guide prepared for the us environmental protection agency office of research and development, washington, dc, Washington, DC, 2014.

NSSO: Energy Sources of Indian Households for Cooking and Lighting, 2011-12, National Sample Survey Office, Ministry of Statistics and Programme Implementation, Government of India Report No. 567 (68/1.0/4), 2011.

Nyanganyura, D., Makarau, A., Mathuthu, M., and Meixner, F. X.: A five-day back trajectory climatology for Rukomechi research station (northern Zimbabwe) and the impact of large-scale atmospheric flows on concentrations of airborne coarse and fine particulate mass, *South African Journal of Science*, 104, 43-52, 2008.

Ofori, F. G., Hopke, P. K., Aboh, I. J. K., and Bamford, S. A.: Characterization of fine particulate sources at Ashaiman in Greater Accra, Ghana, *Atmospheric Pollution Research*, 3, 301-310, <https://doi.org/10.5094/APR.2012.033>, 2012.

Oh, M.-S., Lee, T.-J., and Kim, D.-S.: Quantitative source apportionment of size-segregated particulate matter at urbanized local site in Korea, *Aerosol and Air Quality Research*, 11, 247-264, 2011.

Oliva, P., and Schroeder, W.: Assessment of VIIRS 375m active fire detection product for direct burned area mapping, *Remote Sensing of Environment*, 160, 144-155, <https://doi.org/10.1016/j.rse.2015.01.010>, 2015.

Oliveira, A. V. M., Raimundo, A. M., Gaspar, A. R., and Quintela, D. A.: Globe Temperature and Its Measurement: Requirements and Limitations, *Annals of Work Exposures and Health*, 63, 743-758, [10.1093/annweh/wxz042](https://doi.org/10.1093/annweh/wxz042) *Annals of Work Exposures and Health*, 2019.

Ortiz, J. D., and Jackson, R.: Understanding Eunice Foote's 1856 experiments: heat absorption by atmospheric gases, 0, [doi:10.1098/rsnr.2020.0031](https://doi.org/10.1098/rsnr.2020.0031).

Paatero, P., and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5, 111-126, 1994.

Pallavi, Sinha, B., and Sinha, V.: Source apportionment of volatile organic compounds in the northwest Indo-Gangetic Plain using a positive matrix factorization model, *Atmos. Chem. Phys.*, 19, 15467-15482, 10.5194/acp-19-15467-2019, 2019.

Pandey, J. S., Kumar, R., and Devotta, S.: Health risks of NO<sub>2</sub>, SPM and SO<sub>2</sub> in Delhi (India), *Atmospheric Environment*, 39, 6868-6874, 2005.

Pandithurai, G., Dipu, S., Dani, K. K., Tiwari, S., Bisht, D. S., Devara, P. C. S., and Pinker, R. T.: Aerosol radiative forcing during dust events over New Delhi, India, *Journal of Geophysical Research: Atmospheres*, 113, 2008.

Pant, P., Shukla, A., Kohl, S. D., Chow, J. C., Watson, J. G., and Harrison, R. M.: Characterization of ambient PM<sub>2.5</sub> at a pollution hotspot in New Delhi, India and inference of sources, *Atmospheric Environment*, 109, 178-189, 2015.

Papastefanou, C.: Aerosol sampling and measurement techniques, *Radioactivity in the Environment*, 12, 113-159, 2008.

Park, M., Joo, H. S., Lee, K., Jang, M., Kim, S. D., Kim, I., Borlaza, L. J. S., Lim, H., Shin, H., Chung, K. H., et al.: Differential toxicities of fine particulate matters from various sources, *Scientific Reports*, 8, 17007, 10.1038/s41598-018-35398-0, 2018.

Pattey, E., and Qiu, G.: Trends in primary particulate matter emissions from Canadian agriculture, *Journal of the Air & Waste Management Association*, 62, 737-747, 10.1080/10962247.2012.672058, 2012.

Pawar, H., Garg, S., Kumar, V., Sachan, H., Arya, R., Sarkar, C., Chandra, B. P., and Sinha, B.: Quantifying the contribution of long-range transport to particulate matter (PM) mass loadings at a

---

suburban site in the north-western Indo-Gangetic Plain (NW-IGP), *Atmos. Chem. Phys.*, 15, 9501-9520, 10.5194/acp-15-9501-2015, 2015.

Pease, P. P., Tchakerian, V. P., and Tindale, N. W.: Aerosols over the Arabian Sea: geochemistry and source areas for aeolian desert dust, *Journal of Arid Environments*, 39, 477-496, 1998.

Pérez, I. A., Artuso, F., Mahmud, M., Kulshrestha, U., Sánchez, M. L., and García, M.: Applications of air mass trajectories, *Advances in Meteorology*, 2015, 2015.

Péterfalvi, N., Keller, B., and Magyar, M.: PM10 emission from crop production and agricultural soils, *Agrokémia és Talajtan*, 67, 143-159, 2018.

Petterssen, S.: *Weather analysis and forecasting*, 1956.

Piedra, P., Kalume, A., Zubko, E., Mackowski, D., Pan, Y.-L., and Videen, G.: Particle-shape classification using light scattering: An exercise in deep learning, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 231, 140-156, <https://doi.org/10.1016/j.jqsrt.2019.04.013>, 2019.

Pisharoty, P. R., and Desai, B. N.: Western disturbances and Indian weather, *Indian J. Meteorol. Geophys*, 7, 333-338, 1956.

Pöhlker, C., Wiedemann, K. T., Sinha, B., Shiraiwa, M., Gunthe, S. S., Smith, M., Su, H., Artaxo, P., Chen, Q., Cheng, Y., et al.: Biogenic Potassium Salt Particles as Seeds for Secondary Organic Aerosol in the Amazon, *Science*, 337, 1075, 10.1126/science.1223264, 2012.

Pope III, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., and Heath, C. W.: Particulate air pollution as a predictor of mortality in a prospective study of US adults, *American journal of respiratory and critical care medicine*, 151, 669-674, 1995.

Pope III, C. A.: Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk?, *Environmental health perspectives*, 108, 713-723, 2000a.

---

Pope III, C. A.: Epidemiology of fine particulate air pollution and human health: biologic mechanisms and who's at risk?, *Environmental Health Perspectives*, 108 Suppl 4, 713-723, 10.1289/ehp.108-1637679, 2000b.

Pope III, C. A., and Dockery, D. W.: Health effects of fine particulate air pollution: lines that connect, *Journal of the Air & Waste Management Association*, 56, 709-742, <https://doi.org/10.1080/10473289.2006.10464485>, 2006.

Pöschl, U.: Atmospheric Aerosols: Composition, Transformation, Climate and Health Effects, *Angewandte Chemie International Edition*, 44, 7520-7540, 10.1002/anie.200501122, 2005.

Pöschl, U., Martin, S. T., Sinha, B., Chen, Q., Gunthe, S. S., Huffman, J. A., Borrmann, S., Farmer, D. K., Garland, R. M., and Helas, G.: Rainforest aerosols as biogenic nuclei of clouds and precipitation in the Amazon, *science*, 329, 1513-1516, 2010.

Prata, J. C.: Airborne microplastics: consequences to human health?, *Environmental pollution*, 234, 115-126, 2018.

Ramachandran, S.: Aerosols and Climate Change: Present Understanding, Challenges, and Future Outlook, in: *Land-Atmospheric Research Applications in South and Southeast Asia*, edited by: Vadrevu, K. P., Ohara, T., and Justice, C., Springer International Publishing, Cham, 341-378, 2018.

Ramaswamy, C.: On the sub-tropical jet stream and its role in the development of large-scale convection, *Tellus*, 8, 26-60, 1956.

Rao, P. S., and Sikka, D. R.: Intraseasonal variability of the summer monsoon over the North Indian Ocean as revealed by the BOBMEX and ARMEX field programs, *pure and applied geophysics*, 162, 1481-1510, 2005.

Reff, A., Eberly, S. I., and Bhave, P. V.: Receptor Modeling of Ambient Particulate Matter Data Using Positive Matrix Factorization: Review of Existing Methods, *Journal of the Air & Waste Management Association*, 57, 146-154, 10.1080/10473289.2007.10465319, 2007.

Renwick, L. C., Donaldson, K., and Clouter, A.: Impairment of Alveolar Macrophage Phagocytosis by Ultrafine Particles, *Toxicology and Applied Pharmacology*, 172, 119-127, <https://doi.org/10.1006/taap.2001.9128>, 2001.

Riddle, E. E., Voss, P. B., Stohl, A., Holcomb, D., Maczka, D., Washburn, K., and Talbot, R. W.: Trajectory model validation using newly developed altitude-controlled balloons during the International Consortium for Atmospheric Research on Transport and Transformations 2004 campaign, *Journal of Geophysical Research: Atmospheres*, 111, 10.1029/2006JD007456, 2006.

Riemer, N., Ault, A. P., West, M., Craig, R. L., and Curtis, J. H.: Aerosol mixing state: measurements, modeling, and impacts, *Reviews of Geophysics*, 57, 187-249, 2019.

Rowland, F. S., and Molina, M. J.: Chlorofluoromethanes in the environment, 13, 1-35, 10.1029/RG013i001p00001, 1975.

Saha, D., Soni, K., Mohanan, M. N., and Singh, M.: Long-term trend of ventilation coefficient over Delhi and its potential impacts on air quality, *Remote Sensing Applications: Society and Environment*, 15, 100234, <https://doi.org/10.1016/j.rsase.2019.05.003>, 2019.

Sahai, S., Sharma, C., Singh, D. P., Dixit, C. K., Singh, N., Sharma, P., Singh, K., Bhatt, S., Ghude, S., Gupta, V., et al.: A study for development of emission factors for trace gases and carbonaceous particulate species from in situ burning of wheat straw in agricultural fields in india, *Atmospheric Environment*, 41, 9173-9186, <https://doi.org/10.1016/j.atmosenv.2007.07.054>, 2007.

Samset, B. H.: How cleaner air changes the climate, *Science*, 360, 148, 10.1126/science.aat1723, 2018.



Sanchis, E., Ferrer, M., Calvet, S., Coscollà, C., Yusà, V., and Cambra-López, M.: Gaseous and particulate emission profiles during controlled rice straw burning, *Atmospheric Environment*, 98, 25-31, <https://doi.org/10.1016/j.atmosenv.2014.07.062>, 2014.

Sarkar, C., Kumar, V., and Sinha, V.: Massive emissions of carcinogenic benzenoids from paddy residue burning in North India, *Curr. Sci*, 104, 1703-1706, 2013.

Sarkar, C., Sinha, V., Kumar, V., Rupakheti, M., Panday, A., Mahata, K. S., Rupakheti, D., Kathayat, B., and Lawrence, M. G.: Overview of VOC emissions and chemistry from PTR-TOF-MS measurements during the SusKat-ABC campaign: high acetaldehyde, isoprene and isocyanic acid in wintertime air of the Kathmandu Valley, *Atmos. Chem. Phys.*, 16, 3979-4003, 10.5194/acp-16-3979-2016, 2016.

Sarkar, C., Sinha, V., Sinha, B., Panday, A. K., Rupakheti, M., and Lawrence, M. G.: Source apportionment of NMVOCs in the Kathmandu Valley during the SusKat-ABC international field campaign using positive matrix factorization, *Atmos. Chem. Phys.*, 17, 8129-8156, 10.5194/acp-17-8129-2017, 2017.

Sayahi, T., Butterfield, A., and Kelly, K. E.: Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors, *Environmental Pollution*, 245, 932-940, <https://doi.org/10.1016/j.envpol.2018.11.065>, 2019a.

Sayahi, T., Kaufman, D., Becnel, T., Kaur, K., Butterfield, A. E., Collingwood, S., Zhang, Y., Gaillardon, P. E., and Kelly, K. E.: Development of a calibration chamber to evaluate the performance of low-cost particulate matter sensors, *Environmental Pollution*, 255, 113131, <https://doi.org/10.1016/j.envpol.2019.113131>, 2019b.

Field evaluation of Laser Egg PM sensor: <http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/laser-egg---field-evaluation.pdf>, access: 18 January 2020, 2016.

---

Field evaluation of Kaiterra Laser Egg 2+ sensor: <http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/kaiterra-laser-egg-2---field-evaluation.pdf?sfvrsn=6>, access: 18 January 2020, 2019.

Schraufnagel, D. E.: The health effects of ultrafine particles, *Experimental & Molecular Medicine*, 52, 311-317, 10.1038/s12276-020-0403-3, 2020.

Schrenk, H. H., Heimann, H., Clayton, G. D., Gafafer, W. M., and Wexler, H.: Air pollution in Donora, Pa: epidemiology of the unusual smog episode of October 1948. Preliminary Report, *Public health bulletin*, 306, 1949.

Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment, *Remote Sensing of Environment*, 143, 85-96, <https://doi.org/10.1016/j.rse.2013.12.008>, 2014.

Schwartz, J.: Air pollution and daily mortality: a review and meta analysis, *Environmental research*, 64, 36-52, 1994.

Seinfeld, J. H., and Pandis, S. N.: *Atmospheric chemistry and physics: from air pollution to climate change*, John Wiley & Sons, 2016.

Shah, J. J., Kneip, T. J., and Daisey, J. M.: Source apportionment of carbonaceous aerosol in New York City by multiple linear regression, *Journal of the Air Pollution Control Association*, 35, 541-544, 1985.

Sharma, D., Singh, D., and Kaskaoutis, D. G.: Impact of two intense dust storms on aerosol characteristics and radiative forcing over Patiala, northwestern India, *Advances in Meteorology*, 2012, 2012.

Shiraiwa, M., Kondo, Y., Iwamoto, T., and Kita, K.: Amplification of Light Absorption of Black Carbon by Organic Coating, *Aerosol Science and Technology*, 44, 46-54, 10.1080/02786820903357686, 2010.

Showman, A. P., and Dowling, T. E.: Chapter 20 - Earth as a Planet: Atmosphere and Oceans, in: Encyclopedia of the Solar System (Third Edition), edited by: Spohn, T., Breuer, D., and Johnson, T. V., Elsevier, Boston, 423-444, 2014.

Shy, C. M.: Epidemiologic evidence and the united states air quality standards, American Journal of Epidemiology, 110, 661-671, 10.1093/oxfordjournals.aje.a112847 %J American Journal of Epidemiology, 1979.

Sidhu, H. S., Singh, M., Singh, Y., Blackwell, J., Lohan, S. K., Humphreys, E., Jat, M. L., Singh, V., and Singh, S.: Development and evaluation of the Turbo Happy Seeder for sowing wheat into heavy rice residues in NW India, Field Crops Research, 184, 201-212, <https://doi.org/10.1016/j.fcr.2015.07.025>, 2015.

Sikka, D. R., and Gadgil, S.: On the maximum cloud zone and the ITCZ over Indian, longitudes during the southwest monsoon, Monthly Weather Review, 108, 1840-1853, 1980.

Sikka, D. R.: Evaluation of monitoring and forecasting of summer monsoon over India and a review of monsoon drought of 2002, Proceedings-Indian National Science Academy Part A, 69, 479-504, 2003.

Singh, N., Mittal, S. K., Agarwal, R., Awasthi, A., and Gupta, P. K.: Impact of rice crop residue burning on levels of SPM, SO<sub>2</sub> and NO<sub>2</sub> in the ambient air of Patiala (India), International Journal of Environmental and Analytical Chemistry, 90, 829-843, 2010.

Singh, R. P., Dey, S., Tripathi, S. N., Tare, V., and Holben, B.: Variability of aerosol parameters over Kanpur, northern India, Journal of Geophysical Research: Atmospheres, 109, <https://doi.org/10.1029/2004JD004966>, 2004.

Singh, T., Biswal, A., Mor, S., Ravindra, K., Singh, V., and Mor, S.: A high-resolution emission inventory of air pollutants from primary crop residue burning over Northern India based on VIIRS thermal anomalies, Environmental Pollution, 266, 115132, <https://doi.org/10.1016/j.envpol.2020.115132>, 2020.

Sinha, V., Kumar, V., and Sarkar, C.: Chemical composition of pre-monsoon air in the Indo-Gangetic Plain measured using a new air quality facility and PTR-MS: high surface ozone and strong influence of biomass burning, *Atmos. Chem. Phys.*, 14, 5921-5941, 10.5194/acp-14-5921-2014, 2014a.

Sinha, V., Kumar, V., and Sarkar, C.: Chemical composition of pre-monsoon air in the Indo-Gangetic Plain measured using a new air quality facility and PTR-MS: high surface ozone and strong influence of biomass burning, *Atmospheric Chemistry and Physics*, 14, 5921-5941, <https://doi.org/10.5194/acp-14-5921-2014>, 2014b.

Snyder, E. G., Watkins, T. H., Solomon, P. A., Thoma, E. D., Williams, R. W., Hagler, G. S. W., Shelow, D., Hindin, D. A., Kilaru, V. J., and Preuss, P. W.: The changing paradigm of air pollution monitoring, *Environmental Science & Technology*, 47, 11369-11377, <https://doi.org/10.1021/es4022602>, 2013.

Sorensen, C. M., Flagan, R. C., Baltensperger, U., and Pui, D. Y. H.: Grand challenges for aerosol science and technology, *Aerosol Science and Technology*, 53, 731-734, 10.1080/02786826.2019.1611333, 2019.

Sousan, S., Koehler, K., Thomas, G., Park, J. H., Hillman, M., Halterman, A., and Peters, T. M.: Inter-comparison of low-cost sensors for measuring the mass concentration of occupational aerosols, *Aerosol Science and Technology*, 50, 462-473, <https://doi.org/10.1080/02786826.2016.1162901>, 2016.

Sousan, S., Koehler, K., Hallett, L., and Peters, T. M.: Evaluation of consumer monitors to measure particulate matter, *Journal of Aerosol Science*, 107, 123-133, <https://doi.org/10.1016/j.jaerosci.2017.02.013>, 2017.

Speight, J. G.: Chapter 1 - Chemicals and the Environment, in: *Environmental Organic Chemistry for Engineers*, edited by: Speight, J. G., Butterworth-Heinemann, 1-41, 2017.

---

Stevens, B., and Feingold, G.: Untangling aerosol effects on clouds and precipitation in a buffered system, *Nature*, 461, 607-613, 10.1038/nature08281, 2009.

Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H.: Interpolation Errors in Wind Fields as a Function of Spatial and Temporal Resolution and Their Impact on Different Types of Kinematic Trajectories, *Journal of Applied Meteorology*, 34, 2149-2165, 10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2, 1995.

Stohl, A.: Trajectory statistics—a new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe, *Atmospheric Environment*, 30, 579-587, 1996.

Stohl, A.: Computation, accuracy and applications of trajectories—a review and bibliography, *Atmospheric Environment*, 32, 947-966, 1998.

Stohl, A., Eckhardt, S., Forster, C., James, P., Spichtinger, N., and Seibert, P.: A replacement for simple back trajectory calculations in the interpretation of atmospheric trace substance measurements, *Atmospheric Environment*, 36, 4635-4648, [https://doi.org/10.1016/S1352-2310\(02\)00416-8](https://doi.org/10.1016/S1352-2310(02)00416-8), 2002.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, *Atmos. Chem. Phys.*, 5, 2461-2474, 10.5194/acp-5-2461-2005, 2005.

Sujatha, P., Mahalakshmi, D. V., Ramiz, A., Rao, P. V. N., and Naidu, C. V.: Ventilation coefficient and boundary layer height impact on urban air quality, *Cogent Environmental Science*, 2, 1125284, 10.1080/23311843.2015.1125284, 2016.

Thomas, D., and Charvet, A.: 1 - An Introduction to Aerosols, in: *Aerosol Filtration*, edited by: Thomas, D., Charvet, A., Bardin-Monnier, N., and Appert-Collin, J.-C., Elsevier, 1-30, 2017.

---

Thumaty, K. C., Rodda, S. R., Singhal, J., Gopalakrishnan, R., Jha, C. S., Parsi, G. D., and Dadhwal, V. K.: Spatio-temporal characterization of agriculture residue burning in Punjab and Haryana, India, using MODIS and Suomi NPP VIIRS data, *Curr. Sci*, 109, 1850-1855, 2015.

Tiwari, S., Bisht, D. S., Srivastava, A. K., Pipal, A. S., Taneja, A., Srivastava, M. K., and Attri, S. D.: Variability in atmospheric particulates and meteorological effects on their mass concentrations over Delhi, India, *Atmospheric Research*, 145, 45-56, 2014.

Tomasi, C., and Lupi, A.: Primary and Secondary Sources of Atmospheric Aerosol, in: *Atmospheric Aerosols*, 1-86, 2016.

Triantafyllou, E., Diapouli, E., Tsilibari, E. M., Adamopoulos, A. D., Biskos, G., and Eleftheriadis, K.: Assessment of factors influencing PM mass concentration measured by gravimetric & beta attenuation techniques at a suburban site, *Atmospheric Environment*, 131, 409-417, <https://doi.org/10.1016/j.atmosenv.2016.02.010>, 2016.

Vadrevu, K., and Lasko, K.: Intercomparison of MODIS AQUA and VIIRS I-Band fires and emissions in an agricultural landscape—Implications for air pollution research, *Remote sensing*, 10, 978, 2018.

Vadrevu, K. P., Lasko, K., Giglio, L., Schroeder, W., Biswas, S., and Justice, C.: Trends in Vegetation fires in South and Southeast Asian Countries, *Scientific Reports*, 9, 7422, [10.1038/s41598-019-43940-x](https://doi.org/10.1038/s41598-019-43940-x), 2019.

Van Hung, N., Maguyon-Detras, M. C., Migo, M. V., Quilloy, R., Balingbing, C., Chivenge, P., and Gummert, M.: Rice Straw Overview: Availability, Properties, and Management Practices, in: *Sustainable Rice Straw Management*, edited by: Gummert, M., Hung, N. V., Chivenge, P., and Douthwaite, B., Springer International Publishing, Cham, 1-13, 2020.

Vanderpool, R. W., Krug, J. D., Kaushik, S., Gilberly, J., Dart, A., and Witherspoon, C. L.: Size-selective sampling performance of six low-volume “total” suspended particulate (TSP) inlets,

---

Aerosol Science and Technology, 52, 98-113, <https://doi.org/10.1080/02786826.2017.1386766>, 2018.

Varutbangkul, V., Brechtel, F. J., Bahreini, R., Ng, N. L., Keywood, M. D., Kroll, J. H., Flagan, R. C., Seinfeld, J. H., Lee, A., and Goldstein, A. H.: Hygroscopicity of secondary organic aerosols formed by oxidation of cycloalkenes, monoterpenes, sesquiterpenes, and related compounds, *Atmospheric Chemistry and Physics*, 6, 2367-2388, <https://doi.org/10.5194/acp-6-2367-2006>, 2006.

Venkataraman, C., Habib, G., Eiguren-Fernandez, A., Miguel, A. H., and Friedlander, S. K.: Residential biofuels in South Asia: carbonaceous aerosol emissions and climate impacts, *Science*, 307, 1454-1456, 2005.

Venkataraman, C., Habib, G., Kadamba, D., Shrivastava, M., Leon, J. F., Crouzille, B., Boucher, O., and Streets, D. G.: Emissions from open biomass burning in India: Integrating the inventory approach with high-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) active-fire and land cover data, *Global Biogeochemical Cycles*, 20, 10.1029/2005GB002547, 2006.

Viana, M., Kuhlbusch, T. A. J., Querol, X., Alastuey, A., Harrison, R. M., Hopke, P. K., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A. S. H., et al.: Source apportionment of particulate matter in Europe: A review of methods and results, *Journal of Aerosol Science*, 39, 827-849, <https://doi.org/10.1016/j.jaerosci.2008.05.007>, 2008.

Wang, K., Chen, F.-E., Au, W., Zhao, Z., and Xia, Z.-l.: Evaluating the feasibility of a personal particle exposure monitor in outdoor and indoor microenvironments in Shanghai, China, *International Journal of Environmental Health Research*, 29, 209-220, <https://doi.org/10.1080/09603123.2018.1533531>, 2019.

Wang, Y., Li, J., Jing, H., Zhang, Q., Jiang, J., and Biswas, P.: Laboratory evaluation and calibration of three low-cost particle sensors for particulate matter measurement, *Aerosol Science and Technology*, 49, 1063-1077, <https://doi.org/10.1080/02786826.2015.1100710>, 2015.

Ware, J. H., Thibodeau, L. A., Speizer, F. E., Colome, S., and Ferris, B. G., Jr.: Assessment of the health effects of atmospheric sulfur oxides and particulate matter: evidence from observational studies, *Environmental health perspectives*, 41, 255-276, [10.1289/ehp.8141255](https://doi.org/10.1289/ehp.8141255), 1981.

Warton, D. I., Wright, I. J., Falster, D. S., and Westoby, M.: Bivariate line-fitting methods for allometry, *Biological Reviews*, 81, 259-291, <https://doi.org/10.1017/S1464793106007007>, 2006.

Wesely, M. L., and Hicks, B. B.: A review of the current status of knowledge on dry deposition, *Atmospheric Environment*, 34, 2261-2282, [https://doi.org/10.1016/S1352-2310\(99\)00467-7](https://doi.org/10.1016/S1352-2310(99)00467-7), 2000.

West, J. J., Cohen, A., Dentener, F., Brunekreef, B., Zhu, T., Armstrong, B., Bell, M. L., Brauer, M., Carmichael, G., Costa, D. L., et al.: “What We Breathe Impacts Our Health: Improving Understanding of the Link between Air Pollution and Health”, *Environmental Science & Technology*, 50, 4895-4904, [10.1021/acs.est.5b03827](https://doi.org/10.1021/acs.est.5b03827), 2016.

Whitby, K. T., Husar, R. B., and Liu, B. Y. H.: The aerosol size distribution of Los Angeles smog, *Journal of Colloid and Interface Science*, 39, 177-204, 1972.

Wiedinmyer, C., Quayle, B., Geron, C., Belote, A., McKenzie, D., Zhang, X., O'Neill, S., and Wynne, K. K.: Estimating emissions from fires in North America for air quality modeling, *Atmospheric Environment*, 40, 3419-3432, <https://doi.org/10.1016/j.atmosenv.2006.02.010>, 2006.

Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, *Geosci. Model Dev.*, 4, 625-641, [10.5194/gmd-4-625-2011](https://doi.org/10.5194/gmd-4-625-2011), 2011.

Williams, R., Duvall, R., Kilaru, V., Hagler, G., Hassinger, L., Benedict, K., Rice, J., Kaufman, A., Judge, R., Pierce, G., et al.: Deliberating performance targets workshop: Potential paths for emerging PM<sub>2.5</sub> and O<sub>3</sub> air sensor progress, *Atmospheric Environment: X*, 2, 100031, <https://doi.org/10.1016/j.aeaoa.2019.100031>, 2019.



Wilson, J. G., Kingham, S., Pearce, J., and Sturman, A. P.: A review of intraurban variations in particulate air pollution: Implications for epidemiological research, *Atmospheric Environment*, 39, 6444-6462, <https://doi.org/10.1016/j.atmosenv.2005.07.030>, 2005.

Wilson, J. G., Kingham, S., and Sturman, A. P.: Intraurban variations of PM<sub>10</sub> air pollution in Christchurch, New Zealand: Implications for epidemiological studies, *Science of The Total Environment*, 367, 559-572, <https://doi.org/10.1016/j.scitotenv.2005.08.045>, 2006.

Woodward, S., Roberts, D. L., and Betts, R. A.: A simulation of the effect of climate change-induced desertification on mineral dust aerosol, 32, 10.1029/2005gl023482, 2005.

Zajac, D., Russjan, E., Kostrzon, M., and Kaczyńska, K.: Inhalations with Brine Solution from the 'Wieliczka' Salt Mine Diminish Airway Hyperreactivity and Inflammation in a Murine Model of Non-Atopic Asthma, 21, 4798, 2020.

Zalasiewicz, J., Waters, C. N., Williams, M., Barnosky, A. D., Cearreta, A., Crutzen, P., Ellis, E., Ellis, M. A., Fairchild, I. J., Grinevald, J., et al.: When did the Anthropocene begin? A mid-twentieth century boundary level is stratigraphically optimal, *Quaternary International*, 383, 196-203, <https://doi.org/10.1016/j.quaint.2014.11.045>, 2015.

Zhang, H., Nettleton, D., and Zhu, Z.: Regression-enhanced random forests, arXiv preprint arXiv:1904.10416, 2019.

Zhang, J., Marto, J. P., and Schwab, J. J.: Exploring the applicability and limitations of selected optical scattering instruments for PM mass measurement, *Atmospheric Measurement Techniques*, 11, 2995-3005, <https://doi.org/10.5194/amt-11-2995-2018>, 2018.

Zhang, R., Khalizov, A. F., Pagels, J., Zhang, D., Xue, H., and McMurry, P. H.: Variability in morphology, hygroscopicity, and optical properties of soot aerosols during atmospheric processing, *Proceedings of the National Academy of Sciences*, 105, 10291, 10.1073/pnas.0804860105, 2008.

Zheng, T., Bergin, M. H., Johnson, K. K., Tripathi, S. N., Shirodkar, S., Landis, M. S., Sutaria, R., and Carlson, D. E.: Field evaluation of low-cost particulate matter sensors in high- and low-concentration environments, *Atmospheric Measurement Techniques*, 11, 4823-4846, <https://doi.org/10.5194/amt-11-4823-2018>, 2018.

Zhou, S., Davy, P. K., Huang, M., Duan, J., Wang, X., Fan, Q., Chang, M., Liu, Y., Chen, W., Xie, S., et al.: High-resolution sampling and analysis of ambient particulate matter in the Pearl River Delta region of southern China: source apportionment and health risk implications, *Atmos. Chem. Phys.*, 18, 2049-2064, 10.5194/acp-18-2049-2018, 2018.

Zíková, N., Hopke, P. K., and Ferro, A. R.: Evaluation of new low-cost particle monitors for PM<sub>2.5</sub> concentrations measurements, *Journal of Aerosol Science*, 105, 24-34, <https://doi.org/10.1016/j.jaerosci.2016.11.010>, 2017.

Zimmerman, N., Presto, A. A., Kumar, S. P. N., Gu, J., Haurlyliuk, A., Robinson, E. S., Robinson, A. L., and Subramanian, R.: A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring, *Atmospheric Measurement Techniques*, 11, 291-313, <https://doi.org/10.5194/amt-11-291-2018>, 2018.

Zuberi, B., Johnson, K. S., Aleks, G. K., Molina, L. T., Molina, M. J., and Laskin, A.: Hydrophilic properties of aged soot, *Geophysical Research Letters*, 32, 10.1029/2004GL021496, 2005.

Zuo, J., Ji, W., Ben, Y., Hassan, M. A., Fan, W., Bates, L., and Dong, Z.: Using big data from air quality monitors to evaluate indoor PM<sub>2.5</sub> exposure in buildings: Case study in Beijing, *Environmental Pollution*, 240, 839-847, <https://doi.org/10.1016/j.envpol.2018.05.030>, 2018.