Impact of CRB on Production of Surface Ozone in Punjab Plains of North-West Indo-Gangetic Plains: Development of Prediction Model using Statistical Approach

A

Thesis submitted in partial

Fulfillment of the Requirement

For the Degree of

Doctor of Philosophy

by

Madhvi Rana
(901409008)

Under the Supervision Of

Dr. Susheel Mittal CChem FRSC
Senior Professor

Dr. Gufran Beig
Scientist G, IITM Pune

School of Chemistry & Biochemistry
Thapar Institute of Engineering & Technology
(Deemed to be University)
Patiala-147004

August, 2019

Certificate

Certified that the thesis entitled "**Impact of CRB on production of surface ozone in Punjab plains of North-West Indo-Gangetic Plains: Development of prediction model using statistical approach**" which is submitted by Ms. Madhvi Rana in partial fulfillment of the requirement for the award of the Degree of Doctor of Philosophy in the School of Chemistry and Biochemistry, Thapar Institute of Engineering and Technology, Patiala, is a record of candidate's own independent and original research work carried out by her under our supervision and guidance. The material has not been submitted in part or full to any other University or Institute for the award of any degree.

(Supervisors)

Susheel Mittal 30.8.19
Dr. Susheel Mittal
CChem FRSC
Senior Professor
School of Chemistry and Biochemistry
Thapar Institute of Engineering
and Technology, Patiala-147004

Prof. (Dr.) Gufran Beig
Scientist G & Project Director
Indian Institute of Tropical
Meteorology
Pune - 411008

Prof. (Dr.) Gufran Beig
Scientist - G
Project Director, SAFAR
Indian Institute of Tropical Meteorology., Pune-8
MoES (Govt. of India)

(Head)

Dr. Amjad Ali
Professor
School of Chemistry and Biochemistry
Thapar Institute of Engineering
and Technology, Patiala-147004

Dedicated to my Parents and elder Brother Pankaj for blessings and encouragement

Candidate's Declaration

I, hereby declare that the work presented in the thesis entitled "**Impact of CRB on production of surface ozone in Punjab plains of North-West Indo-Gangetic Plains: Development of prediction model using statistical approach**" in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy in the School of Chemistry and Biochemistry, Thapar Institute of Engineering and Technology, Patiala, is an authentic record of my own work carried out under the joint supervision of Dr. Susheel Mittal, Senior Professor, School of Chemistry and Biochemistry, Thapar Institute of Engineering and Technology, Patiala and Dr. Gufran Beig, Scientist G, Indian Institute of Tropical Meteorology, Pune. The matter embodied in this thesis has not been submitted in part or full to any other university or institute for the award of any degree in India or Abroad.

Madhvi Rana

Place: Patiala
Date: 30/08/2019

Madhvi Rana
901409008

Acknowledgements

This thesis is the culmination of my journey of Ph.D which was just like climbing a high peak step by step accompanied with guidance, hardship, and frustration. The only way to success and knowledge is constantly doing efforts with zeal and passion. When I found myself at top experiencing the feeling of fulfilment, I realized though only my name appears on the cover of this dissertation, a great many people including my family members, well-wishers, my friends, colleagues and various institutions have contributed to accomplish this huge task.

First of all, I would like to express my warm and sincere gratitude to my thesis supervisors Dr. Susheel Mittal and Dr. Gufran Bieg. They always supported me through the work by motivating me from time to time. It would have been very difficult for me to come out with results without there invaluable suggestion.

My earnest thanks to Dr. Susheel Mittal for his scientific advice and knowledge and many insightful discussions and suggestions throughout the work. He is like an ocean of knowledge and ideas from where I got inspired to work and attain the clues for my assignments. I hope that I could be as enthusiastic, and energetic as my Sir and to someday be able to command an audience as well as he can.

I would like to express my deep gratitude to Dr. Prashant Rana for his valuable guidance, timely suggestions and support throughout my research work. Much of this work has not been completed without his vision and encouragement.

I am thankful to Indian Institute of Tropical Meteorology (IITM), Pune for the data collected on the Ambient Air Quality Monitoring System (AAQMS) established at TIET, Patiala, under the MAPAN project. I am also thankful to the Director and Head of School of Chemistry & Biochemistry, Thapar Institute of Engineering and Technology, Patiala, for providing the research facilities. I acknowledge the financial support under the INSPIRE-fellowship program (IF-140931) from the Department of Science and Technology (DST), Government of India.

I would like to thank my family for standing by me through all the joys and sorrows that research scholar life had to offer. My heartfelt thanks and life-long gratitude go to my mother and father for all their love, constant support, encouragement and sacrifices and affections that they have showed upon me. I would like to express my heartily thanks to younger brother Mr. Aman Rana for always being supportive. My heartfelt regard goes to my mother in law and sisters in law for their love and moral support.

I would like to thank my friends and lab mates Mr. Rahul Shukla, Mr. Suneel, Mr.Deepak Batish, Mr. Sanjeev Kumar, Ms. Sonia Rana and Ms. Ravinder Kaur Sandhu for their help in every phase

of the journey. It is impossible to sustain my patience and mood at every stage of research work without their company. They always managed to make me feel special and with them, I had the best tea breaks in my life.

Special thanks to my thunder buddy "Gurpreet Saggu", my friend cum colleague. You are always beside me during the happy and hard moments to push me and motivate me. Thank you so much, for the generous care and unconditional support throughout the research tenure.

I owe thanks to a very special person, my husband, Mr. Rituraj Singh Pundeer for his continued and unfailing love, support and understanding during my pursuit of Ph.D degree that made the completion of thesis possible. I greatly value his contribution and deeply appreciate his belief in me.

I thank the Almighty for giving me the strength and patience to work through all these years so that today I can stand proudly with my head held high.

Madhvi Rana

Madhvi Rana

Abstract

Air quality is a complex function of emissions, meteorology and topography and a sound framework for the connection of these variables can be provided by statistical approaches. Among the most significant contributors of air pollution like automobiles, industries and domestic emissions, the post-harvest crop residue burning is an episodic contributor to the pollutants level in the IGP region. Rice and wheat are the major crop in the North-Western part of the country, which generate a large amount of leftover crop residue, intensified with the usage of mechanical combine harvester technology which is further, subjected to open field crop residue burning. Natural and anthropogenic activities emit aerosols and trace gases such as carbon monoxide (CO) and oxide of nitrogen (NO_x), methane (CH_4), total non-methane (TNMHCs), which impart significant role in the lower atmosphere's reactivity. Meteorology, photochemistry, emissions, and deposition are the major factors that lead to production and accumulation of secondary pollutants like peroxy-acetyl nitrate (PAN) and surface ozone. This research provides a continuous in situ measurements of gaseous pollutants (surface O_3, CO, NO_x, MHC, TNMHCs), and meteorological variables (AT, SR, RH, WS and WD) for a time span of four years between 1st January 2013 to 31st December 2016, at a semi-urban site (Patiala) of North-West Indo-Gangetic Plains (NW-IGP). The meteorological impact on gaseous pollutant has been studied on a daily, monthly, seasonal and diurnal scale. The daily concentrations of gaseous pollutants like O_3, CO, NO_x, MHC, and TNMHCs ranged from 5 to 83 ppb, 0.01 to 2.6 ppm, 2.3 to 113.5 ppb, 562 to 3166 ppb and 27 to 413 ppb, respectively. The meteorological variables RH, SR, AT and WS range between 33.5 to 95.0 %, 68.0 to 378 W m^{-2}, 8 to 38 °C and 0.05 to 7.85 m sec^{-1}, respectively.

In diurnal ozone profile, a steep minimum is observed during the early morning around 0600-0800 hours and rise in ozone concentration is observed after 0800 hour, and hit maxima around in the range of 1500-1600 hours and starts decreasing after late afternoon or evening, in each season with variations in amplitude. The precursor gases exhibited two maxima, during the morning (0700-1000 hours) and evening hours (1700-1900 hours), correspond to boundary layer and prominent anthropogenic emissions during evening hours, when the vehicular activity is maximum and lower mixing ratios in the afternoon hours (1200-1600 hours) influenced by rise in the boundary, provides favorable dispersion conditions coupled

with reduced vehicular emissions. Comparison of diurnal ozone profiles with those of precursor gases (CO, NO_x, MHC, and TNMHCs) showed that maximum O_3 concentration was coincident with the concentration minima of NO_x and CO. This reveals the photochemical production of ozone resulted from the reaction of CO, CH_4, and NMHCs with OH radicals in the presence of NO_x. On seasonal basis, precursor gas (NO_x, CO, MHC, TNMHC), delineated the higher levels in winter and post monsoon season followed by summer and monsoon season.

Ozone is negatively correlated with its precursor gases CO (-0.250), NO_x (-0.262), MHC (-0.151) and TNMHCs (-0.252), as they get consumed in the photochemical production of ozone. The precursor gases (CO, NO_x, MHC, and TNMHCs) showed a strong positive correlation with each other which is attributable to common anthropogenic emission sources such as road traffic and combustion sources. The impact of meteorology on ozone production was positively correlated with SR (0.565), AT (0.449), and WS (0.208) and negatively correlated with RH (-0.682), which highlights the importance of photochemistry in the formation of ozone. CO, NO_x, MHC and TNMHCs were negatively correlated with SR (-0.557, -0.477, -0.300, and -0.441), AT (-0.527, -0.485, -0.320, -0.355), and WS (- 0.401, -0.392, -0.268,-0.468) and positively correlated with RH (0.222, 0.130, 0.217, 0.123) respectively. The correlations observed in all cases are statisticaly significant (p<0.01). Hence, primary pollutants concentration gets declined as the photochemical reaction for the production of secondary pollutant (O_3) become more active with increasing solar radiation, ambient air temperature, and wind speed.

The contribution from regional pollution to long-distance aged pollutants was identified using seasonal backward trajectories and cluster analysis. The main contribution in the four clusters was from cluster 1 (56%), followed by cluster 3 (17%). The biggest influx of gaseous pollutants reached the receptor site from regional continental areas and was not substantially affected by long-range transport. The highest ozone concentrations were observed in the summer, followed by post-monsoon, monsoon and winter season. The rate of increase of ozone (dO_3/dt) is highest during post-monsoon area varies from 14.3 to 18.4 ppb/hour, followed by the summer season (10.2 to 13.5 ppb/hour), indicating the influence of increased emission of precursor gases, especially from the episodic crop residue burning event in

respective periods. Ozone exceedance of 24 hour national standard occurred on 50% and 8% of the sampling periods during wheat and rice crop residue burning, respectively. The daily $O_{3\ max}$ / $O_{3\ min}$ ratio is a pollution index, value of the order 15.60 shows that the location of the study has significant ozone pollution.

A wide range of statistical methodologies were applied to develop satisfactory predictions of daily ozone levels. The development of statistical models involved the detailed choice of both site-specific predictor variables (pollutant and meteorological variables) along with the interactions between predictor variables in order to better capture the target variable (ozone) behavior. Random forest and backward stepwise regression methods are used for the importance and selection of the variables using R packages. It is found that among meteorological variables relative humidity and in precursor gases NO_x has the highest impact on the production of ozone. The fourteen prediction algorithms and their possible combinations of ensemble models were employed in this study. Compared with individual models the ensemble model approach showed an index of agreement of 0.91, the accuracy of 95.5% and mean absolute error of -0.001 ppb between the predicted and observed diurnal cycle and daily averaged data of the year 2016 for benchmark analysis. Predictions on air quality are very coherent and efficient in controlling the measures and can be proposed as a preventive action for regulations and protect public health.

List of Tables

List of Figures

LIST OF CONTENTS

INTRODUCTION

CHAPTER- 1

1.1 Agricultural importance of Punjab plains of North West Indo-Gangetic Plains

Agriculture is the backbone of the Indian economy with an average rate of growth approximately 7% over the last two decade.[1] The Indo-Gangetic Plain (IGP) ranging from 21.75º N, 74.25 ºE to 31.0º N, 91.50º E, enclosed five States namely; Haryana, Punjab, West Bengal, Uttar Pradesh, and Bihar, which accounts for 21% of land area and 40% of the total population of the Indian subcontinent. Punjab in the North -West side of IGP is the second biggest food grain producing state after Uttar Pradesh in India. Rice and wheat are the main crops cultivated in a cyclic manner in the region. It adds 35-40 % of total rice grains and 40-75 % of wheat grains to the food basket of the country.[2] This region experienced extensive economic growth, urban expansion, and development in the past few decades which resulted in worst air quality in the region. Among the most significant contributors of air pollution like automobiles, industries and domestic emissions, the post-harvest crop residue burning is an episodic contributor to the pollutants level. WHO reported the data of critically polluted cities, stretching from west to east of Northern India.[3]

1.2 Agricultural crop residue and open field burning practices

Rice and wheat crops are cultivated in a cyclic manner due to their sensitive biological relations with season and weather constraints. According to Ministry of Agriculture and Farmers Welfare (Government of India) annual report (2017-18), the rice crop has the highest percentage of production followed by wheat in Punjab region.[2] These crops are very environmentally vulnerable. Any kind of lag between the sowing and threshing process of these crops may affect their health and productivity.[4] As per their required environmental conditions**,** rice crop is sown in the middle of June when the weather is hot and humid and harvesting is done in the last week of October when the weather is moderately hot and dry. The burning of rice crop residue from harvesting is done from mid of October till mid-November (in post-monsoon season). Wheat crop is sown in the month of December when weather is cool and dry and is ready to harvest in the month of April when the environment is moderately cool and dry. The burning of wheat crop residue from harvesting is done from mid-April till May (in summer season). The post-harvest crop residue burning events are the conspicuous features in the N.W IGP.

The main crops are rice and wheat, which generate a large amount of leftover crop residue, intensified with the usage of mechanical combine harvester technology, which is further, subjected to open field crop residue burning due to short duration between rice harvesting and wheat sowing, labour shortage, uncertainty of weather and no economical technologies are used for collecting leftover agricultural residues from the field.[2,5] In these months the levels of pollutants are higher than the permissible limits, especially when the weather conditions are unfavourable for the dispersion of air pollutants. As per NASA's satellite-based pictures, every year approximately 46 Metric tons of crops residue waste is burned in the fields of Punjab.[6] The combustion of crop residue owing to less than optimal combustion conditions generates a dense smoke cloud, which can be seen in the IGP area which contain micron-sized aerosols, trace gases (NOx, CO, and VOCs), and the secondary species like ozone, peroxy-acetyl nitrate (PAN) and acid rain impart negative impacts on the composition and reactivity of atmosphere, human health, environment, climate and agriculture crop yield to larger extent.[7-13]

1.3 Importance of trace gases in the atmosphere

The three main gaseous components typically occur in a steady ratio to the total (99.9%) underneath a height of about 86 kilometres: nitrogen accounts for 78% of air by volume, whereas oxygen accounts for 21% of air by volume, and argon accounts for another 0.9% of the complete atmosphere. The rest 0.1% of the atmosphere is made up of trace constituents including water vapor, carbon dioxide, ozone, methane, nitrogen oxides, neon, and helium. These are called trace gases because they exist in limited quantities. Furthermore, environmental variations in worldwide temperature, air pollution and chemical structure changes of troposphere and stratosphere, must be comprehended and quantified.

Crop residue biomass burning is ranked as the second-largest source of air pollution after industrial or production units.[14-17] The studies of the tropospheric chemistry of geographical areas (Africa, South America, and South Asia) are emphasized due to the global impact of biomass burning on air quality and climate change. Mainly China and India are responsible for burning of approximately 67 million crop waste in the open fields with a yearly growth of 19% due to modernization in the agriculture field.[18] It emits aerosols and trace gases such as methane (CH_4), total non-methane (TNMHCs), carbon monoxide (CO) and oxide of nitrogen (NO_x), which impart significant role in the lower

atmosphere's reactivity. Meteorology, photochemistry, emissions, and deposition are the major factors that lead to production and accumulation of secondary pollutants like peroxy-acetyl nitrate (PAN) and surface ozone. ACRB practice contaminates ecological system that directly affects global and regional ambient air, visibility, climate, soil, weather and seasonal cycles.[19]

Tropospheric ozone, its precursor gases, and associated meteorological variables have been broadly monitored and reported in various part of the world.[20-22] The rate of increase of tropospheric ozone per decade in northern India is more than double than in southern India and residence timescale also varied with altitude and season to season.[23] Tropospheric ozone, a greenhouse gas, a powerful oxidant contributes radiative forcing in the atmosphere and also imposes a big threat for crop pro[20]ductivity and respiratory problems of the human beings.[13,24]

1.4 Air quality monitoring program in India- MAPAN

The Central Pollution Control Board of India has established regulations and acceptable standards for air pollutants to protect human health and the environment. Modeling and prediction can be used to evaluate compliance or non-compliance of current pollutant levels with regard to recommendations. The first set of air quality standards were introduced in India in 1982, followed in 1994 and 2009 by revisions.

Modelling Atmospheric Pollution and Networking (MAPAN) project is presently operating in about approx. 20 cities across India, financed by Minister of Earth Science (MoES). Pollutants monitored under this program include PM_{10} and $PM_{2.5}$, O_3, CO, NOx, CO_2, hydrocarbons (MHC and TNMHCs), black carbon (BC), and organic carbon (OC). The monitoring network cover the entire India as Srinagar (Jammu and Kashmir), Patiala (Punjab), Udaipur (Rajasthan), Delhi (NCT of Delhi),Varanasi (Uttar Pradesh), Jabalpur (Madhya Pradesh), Bhubaneshwar (Odisha), Hyderabad (Telangana), Pune and Satara (Maharashtra), Chennai (Tamil Nadu), Thiruvananthapuram (Kerala), Guwahati and Sonitpur (Assam), Aizwal (Mizoram). The air quality monitoring stations are operated in most towns in partnership with local academics/research organizations. It is expected that data generated as part of the network will be used to evaluate the IITM quality forecasting model, but very limited information is available to the public at the moment.

1.5 Chemistry of surface ozone

1.5.1 Dependence of O_3 on precursor gases and meteorology

Surface ozone, a greenhouse gas, contributes to global warming, radiative forcing and index substance of photochemical smog, exerting prominent effect on human health, vegetation and environment.[12,25-31] Indian region has higher ozone production efficiency, followed by China and Japan.[32] Further, in Northern India, surface ozone increases per decade are twice as much as in South India and the residence timescale also varied with altitude and season to season.[23] The photochemical activity of the tropical regions is very high due to the availability of OH radicals, produced by intense solar radiation and a large abundance of higher water vapor content. This feature makes the tropical region as the most powerful photochemically largely active and vulnerable region.[33,34]

Surface ozone is a secondary pollutant and is not directly emitted from sources but is formed by photochemical oxidation processes of CH_4, CO, VOC and NMHCs in the existence of sufficient amount of NO_x concentration in the presence of solar radiation.[35-38] The dominant anthropogenic sources of these prominent precursor gases are motor vehicular exhaust activity, industrial emissions, wood and waste agricultural biomass and bio-fuel incomplete burning activities and chemical solvents.[39] The importance of O_3 production and destruction mechanisms by photochemistry was first suggested by.[40] Further, [41] [42] summarized the ozone formation and its relative chemistry.

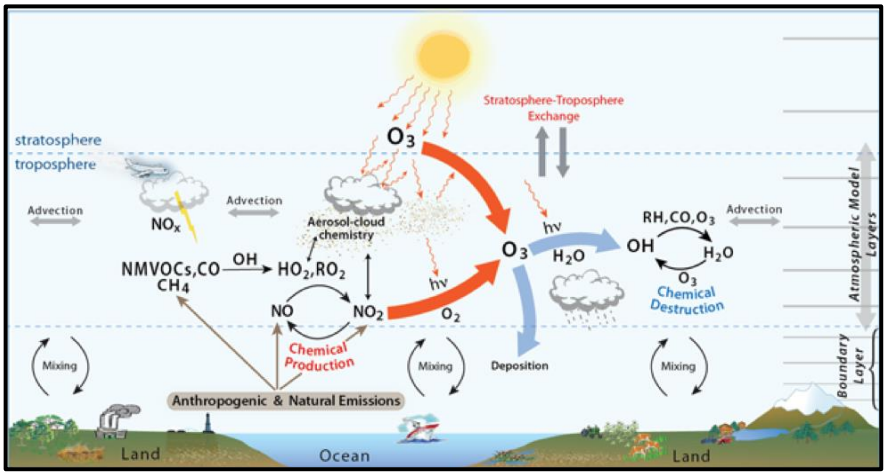


Figure 1.1: Schematic view of the sources and sinks of O_3 in the troposphere[43]

Tropospheric O_3 is controlled by a number of factors such as chemical destruction, photo-chemical formation, transportation of air, dry surface and exchanges between

stratosphere and troposphere. The relative importance of each of these processes depends on the existing meteorological conditions in a particular location. The stratospheric O_3 input into the troposphere is much smaller (10-20 %) than that is believed in the earlier decades, and the photochemical production of ozone is the main source in the troposphere (80- 90%) and an indirect and most important locally related source is advective ozone transport.[44,45]

1.5.2 Photochemical production of ozone

The photochemistry of ozone formation in the atmosphere is, in principle, rather well understood. Ozone formation in the free troposphere is determined by the reaction between methane (CH_4), carbon monoxide (CO) and NOx compounds, as indicated in Figure 1.1. Ozone formation is started by OH attack on hydrocarbons and CO, which leads to peroxy-radicals (HO_2, RO_2) formation. These react with nitric oxide to generate nitrogen dioxide, which photolysis yielding O-atoms, which in change, combine with molecular oxygen contributing to the production of ozone. The photochemistry of tropospheric O_3 formation in simplified form [46,47] is given below from equations (1.1 to 1.6). In photochemical reaction, 10-14 molecules of O_3 were produced by 1 molecule of CO, 3 molecules of CH_4 and 5 molecules of NMHCs on an average. The maximum concentration of ozone is observed during the 1500-1600 hours and starts decreasing after the late afternoon hours. The effectiveness of formation of ozone is restricted by accessible VOC and UV radiation (referred to as VOC-limited regime) while NO concentration limits the ozone production further away from source areas (NO_x limited regime). The ratio of $\frac{\mathrm{VOC}}{\mathrm{NO_x}}$ also varies from urban to rural regions. Stratospheric Tropospheric exchange (STE) and long-range transport from upwind are the nonchemical contributors in ozone levels. The removal process of ozone includes dry deposition (on plants and the Earth's surface), wet deposition (by precipitation), photo-chemical destruction and Stratospheric Tropospheric exchange (STE). The residence timescale of ozone in the atmosphere also varied with altitude and season to season. The chemical lives of ozone are highly variable from weeks to months to hours and less in the shallow nocturnal boundary layer where dry deposition and ' titration ' with fresh NO emissions quickly destroys ozone.

Ozone from oxides of nitrogen

$$NO_2 + h\nu\ (> 380nm) \rightarrow NO + O \quad (1.1)$$

$$O + O_2 + M \rightarrow O_3 + M \quad (1.2)$$

M (O_2 or N_2) - non-reactive species that can take up the energy released in reaction

$$HO_2 + NO \rightarrow OH^{\cdot} + NO_2$$

$$RO_2 + NO \rightarrow RO^{\cdot} + NO_2$$

Ozone from carbon monoxide

$$CO + 2O_2 + h\nu \rightarrow CO_2 + O_3 \quad (1.3)$$

Ozone from methane hydrocarbons

$$CH_4 + 4O_2 + 2h\nu \rightarrow HCHO + H_2O + 2O_3 \quad (1.4)$$

$$HCHO + h\nu \rightarrow H^{\cdot} + HCO^{\cdot}\ (h\nu < 330nm)$$

$$HCO^{\cdot} + h\nu \rightarrow H^{\cdot} + CO\ (h\nu < 360nm)$$

$$CO + 2O_2 + h\nu \rightarrow CO_2 + O_3 \quad (1.5)$$

Ozone from non-methane hydrocarbons

$$RH + 4O_2 + 2h\nu \rightarrow RCHO + H_2O + O_3 \quad (1.6)$$

1.5.3 Atmospheric processes and roles of boundary layer and meteorology

The physical state of the boundary layer is defined with factors such as temperature, relative humidity, atmospheric pressure, and wind, whereas it's chemical condition is determined by variables that measured the ample of a trace gas. The physical state variables are closely related to the boundary layer's dynamic aspects (turbulence level, diffusion efficiency, and air stability) while the chemical state variables are passive scalars without direct dynamic effects with the exception of aerosols. Changes in the state of boundary-layer are governed by means that transmit momentum, energy, and materials within and below the boundary layer. The levels of air pollutants in the atmospheric air are not only determined by their production rate but also by the nature and capacity of physical and chemical "sinks" such as chemical transition, wet and dry deposition on the surface of the earth. Generally, boundary layer height range 100 to 3000 m elevation from the surface of the earth. The remaining of the overlying troposphere is called the free troposphere. Atmosphere's boundary layer, agricultural meteorological, hydrologic, aeronautic, meso-scale meteorology and climate prediction play an important part in many areas of the air pollution. Quantification of the surface-atmosphere trace gas

exchange is acknowledged as a prerequisite to understanding the biosphere's role in the global climate. Surface-atmosphere quantification of trace gas exchange is recognized as a precondition to understanding the biosphere's position in the world climate. Atmospheric boundary layer, approximately 1 km in thickness is the interface of a free atmosphere and the surface of the earth that can be land ecosystems, lake, ice or ocean as shown in Figure 1.2.

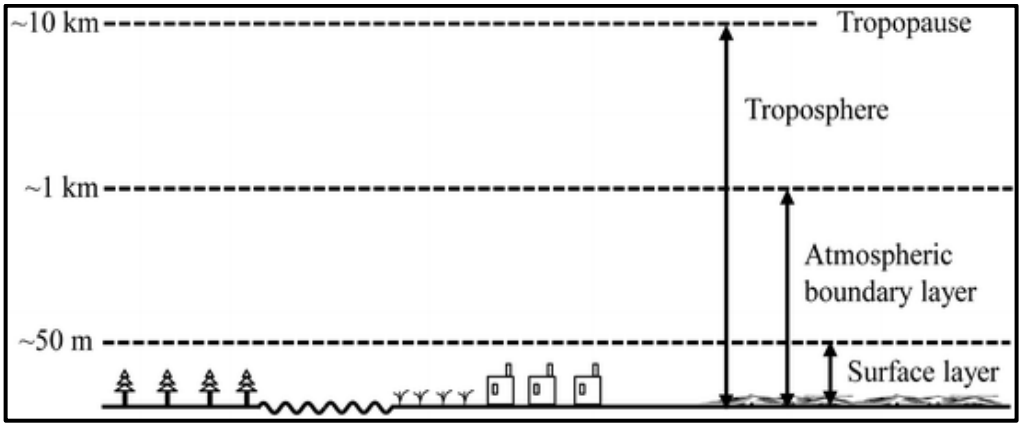


Figure 1.2: Division of the lower atmospheric layers.

When the Earth is hot during sunrise time, it evidently results in mixing of the nocturnal border and residual layer content above it, such modifications in the meteorological factor can have an important impact on the spatial dispersion of pollutants produced on the Earth, leading to spatially and hourly chemical variation, respectively.[48] At noon hours, the convective layer lies usually quite well mixed over the surface layer in which direct emissions are inducted. During the sunset hours, radiative cooling leads to a stable nocturnal boundary layer height, comparable to a radiation inversion. Beyond this lies a residual layer which regulates the thoroughly mixed species in the boundary layer during the day but that do not blend quickly with either the nocturnal boundary layer below or free troposphere above, during the night hours.

Ozone is the two phytotoxic and cytotoxic. It is an important air quality issue and causes major health problems and materials and ecosystems damage.[49] The Environmental Protection Agency (EPA) and the World Health Organization (WHO) consider it as a criteria pollutant.

1.5.4 Effects of surface ozone on vegetation

Plants have developed ozone-preventing mechanisms, ozone exposures above a limit can no longer be coped with by the plant's detoxification process. The critical ozone concentrations which are capable of substantially damaging crops are expressed with "Accumulated Ozone Exposure above the 40 ppbv limit" (AOT40 index).[48,50] This level of ozone concentration is already reached in many nations of the Northern hemisphere such as India, often cause significant losses in agricultural manufacturing.[51,52]

Ozone is primarily absorbed by the stomata when photosynthetic gas exchange is carried out in the mesophyll.[53] When it is inside the leaf, it responds quickly with the moisture of the cell surfaces. The resulting oxygen species are highly reactive and become a burden for the naturally occurring protective mechanisms of the plant. Effects include decreases in crop yield and quality, decreases in development, morphological and physiological alterations, visual and early senescent injuries.[54-56] Over and above physiological damage, surface ozone can result in decreased fungal, bacterial, viral and insect resistance and reproductive inhibition. This can lead to deteriorating quality and reduced biodiversity in natural habitats in crop production on delicate species.[48,57] The recorded highest overall surface ozone-related crop production losses (CCPs) in the U.S. (21-28 million tonnes), followed by China (18-27 million tonnes) and India (8-25 billion tonnes).[58] This delineates for 59 percent of the global economic damage (21, 21, and 17%, respectively). In view of projected losses in worldwide crop production, measurements of surface ozone through a network of locations are extremely necessary and must be controlled by appropriate ozone policies.

1.5.5 Effects of surface ozone on human health

Ozone is a highly potent oxidant, as it reacts with a broad range of biological components and cellular materials.[59] Besides, several factors including exposure levels and durations, intervals between short-term exposures and the averages volumes of breathed air per minute also influence the health impacts.[49,60,61]

1.5.6 Effects of surface ozone on materials

Ozone damages different types of materials more significantly organic materials, both functionally and aesthetically, independently of others and in conjunction with pollutants and environmental aspects.[62] It is known to damage textiles and some polymeric materials such as paints, plastics, rubber and surface coatings.[63] Ozone stimulates the

corrosion effect of sulphur dioxide on metals such as copper, zinc, silver, aluminium, nickel, and iron in conjunction with sulphur dioxide.

1.5.7 Effects of surface ozone on climate

Surface ozone can absorb both solar and infrared radiation[64], so it plays a significant role in the energy budget and earth's climate system.[64-66] A decrease or increase in surface ozone in the lower stratosphere cause decreases or increase in surface temperature.

Ozone is a triatomic molecule with a structure like water vapor has three radiatively active vibrational bands 9, 9.6, and 14.4 micrometers.[67] The 9 and 9.6 micrometers bands fall in the atmospheric window, which is otherwise transparent in terms of Earth's atmosphere large climatic impact. The 14.4-micrometer band matches the powerful carbon dioxide absorption band and therefore has little effect on the environment. Hence, the resultant effect on surface temperature and consequently climate of the earth critically dependent on the vertical distribution of the trend in ozone. Recent findings have revealed that most of the ozone in the troposphere is locally produced by human activities and the increasing trend leading to an overall decrease in the oxidation efficiency of the atmosphere.

1.6 Air quality forecasting

During airborne pollution episodes, short term forecast of air quality is required to take preventive and evasive actions. Air quality forecasting is crucial to determine the impacts of regional economic development and for the establishment of regional environmental management plans. It cut-down the requirement of stations, frequency of sampling and hence reduces cost on producing the data. It also provides data during the period of a technical error in the monitoring station. It provides pre-information to the general public about the status of the air environment in which they can take precautionary steps to take their health. Various prediction approaches (deterministic and statistical models) have been developed for the purpose of air quality forecasting.

1.6.1 Prediction approaches: Deterministic and Statistical models

To predict surface ozone concentration level, one must understand the linkages between precursor pollutants levels and meteorological variables, which can depict the processes involved in photochemical reactions (formation and destruction) of ozone.[68,69] However, predictions capability might also get affected due to the choice of community structure,

training techniques, and sample records. The numerous prediction strategies might be extensively labelled underneath deterministic models and statistical-empirical models as shown in Figure 1.3.[70]

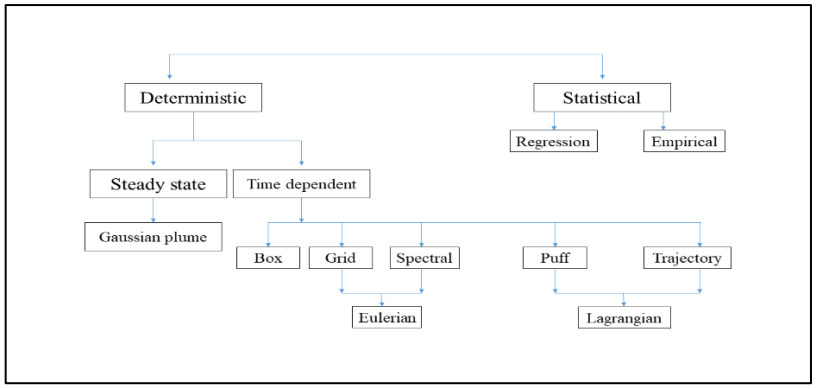


Figure 1.3: Types of forecasting approaches

Deterministic or numerical models, usually simulate atmospheric chemistry and dispersion models (cause/impact models) have been applied for a specific region having gridded air quality data.[71,72] Besides being useful, this approach cannot always be applicable, because of the very precise data requirement. For example, emission and transportation of pollutants and meteorological conditions.[73] Because of the lack of availability of rigorous data, statistical models becomes an outstanding approach for air quality forecasting.[74,75] They rely on the time series analysis between meteorological and air pollution data using historical trends, without dealing with the causes/impacts in the production of secondary pollutants.

Statistical approach received an increased attention for ozone prediction in last decade for different geographical regions, with a range of linear and nonlinear regression methods. [75-92] These models offer forecasts of higher accuracy and with higher computational efficiency than deterministic models.[93,94]

Advantages of a statistical approach:

- Relationships are established between input factors (predictors) and output variables (predicts and results) without detailing the causes and impacts of pollutant creation.
- Annual average air quality is primarily governed by source factors for a particular local region.
- There is no need to know the causes of pollutant dispersion.

- It doesn't consider physical and chemical methods and rather uses historical pollution records.
- The output of these models is similar and can be superior to deterministic modes many times without the need for elaborate output .[95]

Regression based models are the most commonly used statistical models in the literature of air pollution, with complexity varying extensively. In turn, these models can be categorized into three main groups: linear regression, regression tree, and non-linear regression. Linear regression models are the easiest type of regression models, but they have restricted flexibility to capture meteorological and pollutant interactions and non-linearities. Linear models were used by many researchers to predict ground ozone concentration as well as other pollutants.[96,97] Classification and Regression Trees can depict a non-linear connection between meteorological factors and levels of pollutants. Regression trees were used to model peak daily levels of ozone. The trend component in such models cannot, however, be recorded immediately. Finally, non-linear regression may be suitable if the nature of the physical issue requires a nonlinear connection. For exploratory analysis, the conditions in the model were created using graphical techniques and nonparametric regression to determine the outstanding relationships involving ozone and many candidate predictor variables.

1.6.2 Ensemble approach - The Necessity

Sometimes single linear or non-linear models might not be enough to capture all the relationships among time-related input variables. Hence, no single linear or non-linear approach is fitting suitably in every situation. Considering these limitations, further ensemble/ hybrid model was generated with the aim of retaining their advantages and avoids overfitting to get the best predictive combination. Ensemble strategies are meta-algorithms that integrate several machine learning strategies into one predictive model a good way to lower variance (bagging), bias (boosting), or enhance predictions (stacking) that lead to higher performance for forecasting if a single model fails to recognize the actual information generating process or while a single model may not be good enough to understand the traits of the time series data (Bates and ranger, 1969; Clemen, 1989, Makridaki, et al., 1982; Makridakis et al., 1993). The combination of linear and non-linear models will improve the opportunity to obtain distinct factual patterns over the individual model and improve forecast efficiency (Zhang, 2003).

1.6.3 Performance statistics

A variety of performance indicators such as correlation coefficient (r), fractional variance (FV), root mean square error (RMSE), accuracy, and index of agreement (d) are provided to mark the uncertainty of applying the suitable statistics to select the optimized model. In the literature, the performance of the models is evaluated using applying only one or more indices and noted that inconsistency existed in the aspects by which air quality models and their ensemble are tested and evaluated.[98,99] Hence, five performance indices were applied to this work. The model's behavior in both development and validation steps was assessed by C_p (modelled observation) and C_O (observed observation) by considering the following performance indices, given by equations (1-5).

$$\text{Accuracy} = [(\text{Co} - \text{Cp}) / \text{Co}\,] * 100 \qquad (1)$$

$$r = \sqrt{\frac{\sum_{i=1}^{n}(Co - \overline{Co})^2 - \sum_{i=1}^{n}(Co - Cp)^2}{\sum_{i=1}^{n}(Co - \overline{Co})^2}} \qquad (2)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Co - Cp)^2} \qquad (3)$$

$$d = 1 - \frac{[\sum_{i=1}^{n}|Cp - Co|^2]}{[\sum_{i=1}^{n}(|Cp - \overline{Co}| + |Co - \overline{Co}|)^2]} \qquad (4)$$

$$FV = 2\frac{\sigma_{Co} - \sigma_{Cp}}{\sigma_{Co} + \sigma_{Cp}} \qquad (5)$$

These statistics permit an evaluation of the performance of various models. The correlation coefficient (r) provides the measure of the variation of the reproduced data within the model. The level of statistical significance of the test is expressed as a *p*-value, traditionally 5% or 1% and denoted as α. If the *p*-value (5% or 1%) is less than the chosen significance level (α), that suggests that the results is statistically significant. RMSE measures residual error, which gives a general concept of the distinction between the observed and modelled value and is a prominent technique for measuring the error rate of a regression model. Index of agreement (d) measures the variation between the mean, predicted and observed values, which illustrates the degree of error-free predictions within limits of 0 (no agreement) to 1 (perfect agreement). Fractional variance (FV) measures the difference between the predicted and observed variance. Its value equal to zero means equal variance of the predicted and observed measurements. The accuracy measures the percentage variation of the observed and predicted values with acceptable error. The K-fold cross-validation justifies the robustness of the prediction method.

1.7 References

1. In *Global Economic Prospects, June **2018**: The Turning of the Tide?*, pp 111-148.
2. Jain, N.; Bhatia, A.; Pathak, H. *Aerosol and Air Quality Research* **2014**, *14*, 422-430.
3. Organization, W. H. *Geneva: World Health Organization* **2016**.
4. Awasthi, A.; Singh, N.; Mittal, S.; Gupta, P. K.; Agarwal, R. *Science of the total environment* **2010**, *408*, 4440-4445.
5. Kumar, V.; Sarkar, C.; Sinha, V. *Journal of Geophysical Research: Atmospheres* **2016**, *121*, 3619-3633.
6. Thumaty, K. C.; Rodda, S. R.; Singhal, J.; Gopalakrishnan, R.; Jha, C. S.; Parsi, G. D.; Dadhwal, V. K. *Curr. Sci* **2015**, *109*, 1850-1855.
7. Galanter, M.; Levy, H.; Carmichael, G. R. *Journal of Geophysical Research: Atmospheres* **2000**, *105*, 6633-6653.
8. Mittal, S. K.; Singh, N.; Agarwal, R.; Awasthi, A.; Gupta, P. K. *Atmospheric Environment* **2009**, *43*, 238-244.
9. Awasthi, A.; Agarwal, R.; Mittal, S. K.; Singh, N.; Singh, K.; Gupta, P. K. *Journal of Environmental Monitoring* **2011**, *13*, 1073-1081.
10. Agarwal, R.; Awasthi, A.; Singh, N.; Mittal, S. K.; Gupta, P. K. *International journal of environmental health research* **2013**, *23*, 281-295.
11. Kaskaoutis, D.; Kumar, S.; Sharma, D.; Singh, R. P.; Kharol, S.; Sharma, M.; Singh, A.; Singh, S.; Singh, A.; Singh, D. *Journal of Geophysical Research: Atmospheres* **2014**, *119*, 5424-5444.
12. Ashworth, K.; Wild, O.; Eller, A. S.; Hewitt, C. N. *Environmental science & technology* **2015**, *49*, 8566-8575.
13. Sinha, B.; Singh Sangwan, K.; Maurya, Y.; Kumar, V.; Sarkar, C.; Chandra, B.; Sinha, V. *Atmospheric Chemistry and Physics* **2015**, *15*, 9555-9576.
14. Badarinath, K.; Chand, T.; Prasad, V. K. *Current Science (00113891)* **2006**, *91*.
15. Badarinath, K.; Kharol, S. K.; Sharma, A. R.; Prasad, V. K. *Journal of Atmospheric and Solar-Terrestrial Physics* **2009**, *71*, 1267-1276.
16. Agarwal, R.; Awasthi, A.; Singh, N.; Gupta, P. K.; Mittal, S. K. *Science of the Total Environment* **2012**, *429*, 161-166.
17. Tiwari, S.; Srivastava, A.; Bisht, D.; Safai, P.; Parmita, P. *Natural hazards* **2013**, *65*, 1745-1764.
18. Kumar, P.; Khare, M.; Harrison, R. M.; Bloss, W. J.; Lewis, A.; Coe, H.; Morawska, L. *Atmospheric Environment* **2015**, *122*, 657-661.
19. Singh, R. P.; Kaskaoutis, D. G. *Eos, Transactions American Geophysical Union* **2014**, *95*, 333-334.
20. Akpinar, S.; Oztop, H. F.; Akpinar, E. K. *Environmental monitoring and assessment* **2008**, *146*, 211-224.
21. Xu, W.; Zhao, C.; Ran, L.; Deng, Z.; Liu, P.; Ma, N.; Lin, W.; Xu, X.; Yan, P.; He, X. *Atmospheric Chemistry and Physics* **2011**, *11*, 4353-4369.
22. Singla, V.; Satsangi, A.; Pachauri, T.; Lakhani, A.; Kumari, K. M. *Atmospheric research* **2011**, *101*, 373-385.
23. Lal, D.; Ghude, S. D.; Patil, S.; Kulkarni, S. H.; Jena, C.; Tiwari, S.; Srivastava, M. K. *Atmospheric research* **2012**, *116*, 82-92.
24. Jerrett, M.; Burnett, R. T.; Pope III, C. A.; Ito, K.; Thurston, G.; Krewski, D.; Shi, Y.; Calle, E.; Thun, M. *New England Journal of Medicine* **2009**, *360*, 1085-1095.

25. Hansen, J.; Nazarenko, L.; Ruedy, R.; Sato, M.; Willis, J.; Del Genio, A.; Koch, D.; Lacis, A.; Lo, K.; Menon, S. *science* **2005**, *308*, 1431-1435.
26. Schmale, J.; Shindell, D.; von Schneidemesser, E.; Chabay, I.; Lawrence, M. *Nature News* **2014**, *515*, 335.
27. Chameides, W. L.; Kasibhatla, P.; Yienger, J.; Levy, H. *Science* **1994**, *264*, 74-77.
28. Kampa, M.; Castanas, E. *Environmental pollution* **2008**, *151*, 362-367.
29. Cooper, O.; Parrish, D.; Stohl, A.; Trainer, M.; Nédélec, P.; Thouret, V.; Cammas, J.-P.; Oltmans, S.; Johnson, B.; Tarasick, D. *Nature* **2010**, *463*, 344.
30. Debaje, S.; Kakade, A. *Aerosol Air Qual. Res* **2006**, *6*, 322-333.
31. Jung, S.-W.; Lee, K.; Cho, Y.-S.; Choi, J.-H.; Yang, W.; Kang, T.-S.; Park, C.; Kim, G.-B.; Yu, S.-D.; Son, B.-S. *International journal of environmental research and public health* **2016**, *13*, 728.
32. Berntsen, T.; Isaksen, I. S.; Wang, W.-C.; Liang, X.-Z. *Tellus B: Chemical and Physical Meteorology* **1996**, *48*, 13-32.
33. Kuhlbusch, T.; Crutzen, P. *Global Biogeochemical Cycles* **1995**, *9*, 491-501.
34. Mohanakumar, K. *Stratosphere troposphere interactions: an introduction*; Springer Science & Business Media, **2008**.
35. Fishman, J.; Crutzen, P. J. *Nature* **1978**, *274*, 855.
36. Crutzen, P. J. *Annual review of earth and planetary sciences* **1979**, *7*, 443-472.
37. Lelieveld, J.; Crutzen, P. *Journal of Atmospheric Chemistry* **1991**, *12*, 229-267.
38. Horowitz, L. W.; Walters, S.; Mauzerall, D. L.; Emmons, L. K.; Rasch, P. J.; Granier, C.; Tie, X.; Lamarque, J. F.; Schultz, M. G.; Tyndall, G. S. *Journal of geophysical research: Atmospheres* **2003**, *108*.
39. Safai, P.; Raju, M.; Rao, P.; Pandithurai, G. *Atmospheric environment* **2014**, *92*, 493-500.
40. Crutzen, P. *Pure and Applied Geophysics* **1973**, *106*, 1385-1399.
41. Pitts, B. F.; Pitts, J. N. *Academic, US* **2000**.
42. Seinfeld, J. H.; Pandis, S. N.; John Wiley & Sons Inc, New York, **2006.**
43. Young, P. J.; Naik, V.; Fiore, A. M.; Gaudel, A.; Guo, J.; Lin, M.; Neu, J.; Parrish, D.; Reider, H.; Schnell, J. *Elementa: Science of the Anthropocene* **2018**, *6*.
44. Moller, D. *Arhiv za higijenu rada i toksikologiju* **2004**, *55*, 11-23.
45. Zhang, R.; Lei, W.; Tie, X.; Hess, P. *Proceedings of the National Academy of Sciences* **2004**, *101*, 6346-6350.
46. Chameides, W.; Walker, J. C. *Journal of Geophysical Research* **1973**, *78*, 8751-8760.
47. Liu, S.; Trainer, M.; Fehsenfeld, F.; Parrish, D.; Williams, E.; Fahey, D. W.; Hübler, G.; Murphy, P. C. *Journal of Geophysical Research: Atmospheres* **1987**, *92*, 4191-4207.
48. Mills, N. L.; Törnqvist, H.; Robinson, S. D.; Gonzalez, M. C.; Söderberg, S.; Sandström, T.; Blomberg, A.; Newby, D. E.; Donaldson, K. *Inhalation toxicology* **2007**, *19*, 81-89.
49. Nuvolone, D.; Petri, D.; Voller, F. *Environmental Science and Pollution Research* **2018**, *25*, 8074-8088.
50. Fuhrer, J.; Skärby, L.; Ashmore, M. R. *Environmental pollution* **1997**, *97*, 91-106.
51. Kärenlampi, L.; Skärby, L. In *UN-ECE Convention on Long-Range Transboundary Air Pollution Workshop (1996: Kuopio, Finland)*; University of Kuopio, Dept. of Ecology and Environmental Science, **1996**.
52. Fowler, D.; Amann, M.; Anderson, F.; Ashmore, M.; Cox, P.; Depledge, M.; Derwent, D.; Grennfelt, P.; Hewitt, N.; Hov, O. *Royal Society Science Policy Report* **2008**, *15*.
53. Fiscus, E. L.; Booker, F. L.; Burkey, K. O. *Plant, Cell & Environment* **2005**, *28*, 997-1011.
54. Finnan, J.; Jones, M.; Burke, J. *Agriculture, ecosystems & environment* **1996**, *57*, 159-167.
55. Gimeno, B.; Bermejo, V.; Sanz, J.; De La Torre, D.; Elvira, S. *Environmental Pollution* **2004**, *132*, 297-306.

56. Debaje, S.; Kakade, A.; Jeyakumar, S. J. *Journal of hazardous materials* **2010**, *183*, 773-779.
57. Morgan, P. B.; Mies, T. A.; Bollero, G. A.; Nelson, R. L.; Long, S. P. *New Phytologist* **2006**, *170*, 333-343.
58. Avnery, S.; Mauzerall, D. L.; Liu, J.; Horowitz, L. W. *Atmospheric Environment* **2011**, *45*, 2284-2296.
59. Avol, E. L.; Navidi, W. C.; Rappaport, E. B.; Peters, J. *Research report (Health Effects Institute)* **1998**, iii, 1-18; discussion 19-30.
60. Ghorani-Azam, A.; Riahi-Zanjani, B.; Balali-Mood, M. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences* **2016**, *21*.
61. Puri, P.; Nandar, S. K.; Kathuria, S.; Ramesh, V. *Indian Journal of Dermatology, Venereology, and Leprology* **2017**, *83*, 415.
62. Kucera, V.; Fitz, S. *Water, Air, and Soil Pollution* **1995**, *85*, 153-165.
63. Hanning-Lee, M.; Brady, B.; Martin, L.; Syage, J. *Geophysical research letters* **1996**, *23*, 1961-1964.
64. Brasseur, G. P.; Muller, J.; Tie, X.; Horowitz, L. *Present and Future of Modeling Global Environmental Change: Toward Integrated Modeling. TERRAPUB* **2001**, 63-75.
65. Schnell, J. L.; Prather, M. J.; Josse, B.; Naik, V.; Horowitz, L. W.; Zeng, G.; Shindell, D. T.; Faluvegi, G. *Geophysical Research Letters* **2016**, *43*, 3509-3518.
66. Tian, H.; Ren, W.; Tao, B.; Sun, G.; Chappelka, A.; Wang, X.; Pan, S.; Yang, J.; Liu, J.; S. felzer, B. *Ecosystem Health and Sustainability* **2016**, *2*, e01203.
67. Petty, G. W. *A first course in atmospheric radiation*; Sundog Pub, **2006.**
68. Abdul-Wahab, S. A.; Al-Alawi, S. M. *Environmental Modelling & Software* **2002**, *17*, 219-228.
69. Abdul-Wahab, S. A.; Bakheit, C. S.; Al-Alawi, S. M. *Environmental Modelling & Software* **2005**, *20*, 1263-1271.
70. Kalenderski, S. *Stochastic modeling of space-time processes: an air pollution problem*. University of British Columbia **2009**.
71. Jacob, D. J. *Introduction to atmospheric chemistry*; Princeton University Press, **1999**.
72. Zannetti, P. *Southampton and Van Nostrand Reinhold, New York* **1990**.
73. Denby, B.; Cassiani, M.; de Smet, P.; de Leeuw, F.; Horálek, J. *Atmospheric environment* **2011**, *45*, 4220-4229.
74. Kalenderski, S.; Steyn, D. G. *Environmetrics* **2011**, *22*, 572-586.
75. Pires, J.; Martins, F. *Atmospheric Environment* **2011**, *45*, 2413-2417.
76. Gardner, M.; Dorling, S. In *Proceedings 1st International Conference on GeoComputation*; University of Leeds, **1996**, pp 359-370.
77. Quintela-del-Río, A.; Francisco-Fernández, M. *Science of The Total Environment* **2011**, *409*, 1123-1133.
78. Kumar, A.; Goyal, P. *Pure and Applied Geophysics* **2013**, *170*, 711-722.
79. Flaum, J. B.; Rao, S. T.; Zurbenko, I. G. *Journal of the Air & Waste Management Association* **1996**, *46*, 35-46.
80. Gardner, M. W.; Dorling, S. *Atmospheric environment* **1998**, *32*, 2627-2636.
81. Thompson, M. L.; Reynolds, J.; Cox, L. H.; Guttorp, P.; Sampson, P. D. *Atmospheric environment* **2001**, *35*, 617-630.
82. Elkamel, A.; Abdul-Wahab, S.; Bouhamra, W.; Alper, E. *Advances in environmental research* **2001**, *5*, 47-59.
83. Sousa, S.; Martins, F. G.; Alvim-Ferraz, M.; Pereira, M. C. *Environmental Modelling & Software* **2007**, *22*, 97-103.
84. Lin, Y.; Cobourn, W. G. *Atmospheric Environment* **2007**, *41*, 3502-3513.

85. Al-Alawi, S. M.; Abdul-Wahab, S. A.; Bakheit, C. S. *Environmental Modelling & Software* **2008**, *23*, 396-403.
86. Cheng, C.-H.; Huang, S.-F.; Teoh, H.-J. *Computers & Mathematics with Applications* **2011**, *62*, 2016-2028.
87. Feng, Y.; Zhang, W.; Sun, D.; Zhang, L. *Atmospheric Environment* **2011**, *45*, 1979-1985.
88. Petelin, D.; Grancharova, A.; Kocijan, J. *Simulation modelling practice and theory* **2013**, *33*, 68-80.
89. Sun, W.; Zhang, H.; Palazoglu, A. *Atmospheric environment* **2013**, *81*, 199-208.
90. Baawain, M. S.; Al-Serihi, A. S. *Aerosol and air quality research* **2014**, *14*, 124-134.
91. Fontes, T.; Silva, L.; Silva, M.; Barros, N.; Carvalho, A. *Science of the Total Environment* **2014**, *488*, 197-207.
92. Kumar, N.; Middey, A.; Rao, P. S. *Urban Climate* **2017**, *20*, 148-167.
93. Draxler, R. R. *Journal of the Air & Waste Management Association* **2000**, *50*, 259-271.
94. Zhang, Y.; Bocquet, M.; Mallet, V.; Seigneur, C.; Baklanov, A. *Atmospheric Environment* **2012**, *60*, 632-655.
95. Kandya, A.; Mohan, M. In *Proceedings of the 7th international conference on urban climate*, **2009**.
96. McKendry, P. *Bioresource technology* **2002**, *83*, 37-46.
97. Robeson, S.; Steyn, D. *Atmospheric Environment (1967)* **1989**, *23*, 689-692.
98. Fox, D. G. *Bulletin of the American Meteorological Society* **1981**, *62*, 599-609.
99. Willmott, C. J. *Bulletin of the American Meteorological Society* **1982**, *63*, 1309-1313.

REVIEW OF LITERATURE

CHAPTER- 2

2.1 Influence of precursor gases and meteorology on surface ozone production

Ozone pollution is major threat to health and exerts a wide range of impacts on biological and economic systems. The aim of this literature review is to justify this study's research goals, taking into account prior studies by investigating past air quality monitoring and prediction studies and determining where future research is needed. The aim of this literature review is to justify this study's research goals, taking into account prior study studies and surveillance of air quality and prediction and to determine where future research is required. This section contains review of the already documented work by many authors related to the relationship between ozone, its precursor gases, meteorological variables and application of a wide range of statistical methodologies. They were applied to develop satisfactory predictions of ozone levels for regional air quality that had taken place in different corners of the world including India. This section provides traditional and recent modelling and forecasting approaches and methods for air quality parameters and indices. It has been noted that various methods have been used for environmental modeling. The problem's complexity determines the approach preferred **(Gardner and Dorling, 1998).**[1]

2.1.1 National scenario

Several research organizations have actively involved the measurements of surface ozone, its precursors, and meteorology for long and short-term in urban, semi-urban and rural regions on the Indian continent area. It is requisite to estimate the regional air quality and to collect the necessary trace gases inventories over the Indian subcontinent. The Himalayan Indo-Gangetic plains and foothills are a significant source of trace gases and particulate matter pollution impact on both regional and global climate.

One seventh of the world's population lives in Indo-Gangetic Plains (IGPs) and this fertile region supports agricultural production of food crops for many of South Asia. However, in terms of atmospheric composition and chemistry, it is still one of the most under studied areas of the globe. Surface ozone, photochemical reaction includes the oxidation, depending on levels of NO_x and Hydrogen oxide radicals, which act as catalysts in the response, of CO, CH_4, NMHCs and other volatile organic compounds.

Pancholi *et al.* (2018)[2] described continuous measurement of O_3, NO_x and meteorological parameters in the urban areas of Jodhpur District in Rajasthan, India. The diurnal cycles of O_3 and NO_x were inverse, with surface O_3 showing a peak in mid-day and lower concentration at night-time. For O_3, the significant seasonal difference during pre-monsoon / monsoon was noted to be the highest / lowest, respectively. In contrast, NO and NO_2 marked the highest/lowest levels during post-monsoon/ monsoon, respectively. Significant positive correlations of O_3 with temperature and solar radiation were found during the studied seasons and negatively correlated with relative humidity. In the entire monitoring period there was a noticeable difference of < 10 mg / m^3 in the O_3 levels between weekends and weekdays. **Kumari *et al.* (2018)[3]** have reported the downwind effect of the post-harvested crop residue fires on two sites Agra and Delhi. In the post-harvest period, there has been an increase of 17.3 and 31.7 ppb respectively in the hourly average O_3 mixing ratios of Agra and Delhi, under comparable weather conditions.

Sharma *et al.* (2017)[4] recorded the emission inventories of ozone and its precursors in NCR-Delhi. In places where NOx emissions are greater, low levels of ozone were observed. Higher levels of ozone were reported close to the boundary of study relative to the center of Delhi. Emission inventories shows that the transportation industry accounts for 55% of total NOx emissions and the power plants (23%) and diesel generators (7%). The transport sector again accounts for 33% in NMVOC inventories, followed by evaporative emissions resulting from solvent and fuel handling (30%) and agricultural residue burning (28%). The refuse burning makes up 73% of CO emissions primarily as a result of combustion incompleteness, followed by the combustion of agricultural residue (14%). Regional scale checks are needed to reduce ozone in NCR. **Verma *et al.* (2017)[5]** reported the influence of atmospheric conditions on ozone episodes over a semi-urban site, Agra, India. In June, the highest (48.2 ± 31.0 ppb) average monthly ozone concentration was almost three times that of December (16.7 ± 10.5 ppb). For elevated episodes of ozone, the hourly ozone limit of 78 days for India exceeds 90 ppb (NAAQS, CPCB, 2009), while the highest daily ozone limit of 70 ppb or greater exceeds 75 days. Trajstat was used during the episodic days for air mass cluster analysis, the maximum frequency (35.4%) of these air clusters were northwest-directed clusters. Local photochemistry and/or transportation were responsible to the episodes.

Peshin *et al.* (2017)[6] reported air quality data of network having 8 stations, located across Delhi NCR, comprised over 2,000 square kilometres of the area, monitored the gaseous and particulate air pollutants (O_3, NO_2, CO, $PM_{2.5}$ and PM_{10}) on various temporal and spatial scales. The analysis of backward air mass trajectory and residence time (RT) helped to measure the quantity of photochemical O_3 at the research site. The marking of O_3 values with RT disclosed important positive correlation that the build-up of O_3 starts when fresh air masses are accumulated in the polluted domain for 1-3 days and produces O3 by photooxidation. After this, the less pronounced photochemical effects are seen by the saturated and aged air masses. Due to the suppression of the boundary layer with O3 built up at a rate of 25.8 ppb^{-1} during the season, Maximum air mass stagnation was observed during winter. The research showed that the greatest impact on O_3 manufacturing in the study site was the large-scale anthropogenic emissions in the surrounding areas, however long range transportation was a major probability, particularly during a pre-monsoon season. In urban site Delhi, India, **Sharma *et al.* (2016)[7]** reported the ground concentrations of O_3 and its precursors gases (NO, NO_2, CO, CH_4 and NMHCs). The maximum/minimum average levels of O_3, NO and NO_2 were observed summer and monsoon season. The O_3, NO and NO_2 levels shown to have significant diurnal changes over the seasons. The findings reveal a negative correlation between ground O_3 and NO_x and CO during the research.

Gaur *et al.* (2014)[8] reported air quality of an urban site in Kanpur, by monitoring long-term near-surface measured of trace gases, (SO_2, NOx, CO and O_3) from June 2009 to May 2013. NO_x and CO exhibit peaks in morning and evening and minima in the afternoon, regardless of the season, that was clearly associated with the boundary layer height evolution. Ozone depicted with maximum afternoon levels and lowest in the morning, whereas levels of O_3 peaked in the summer owing to increased O_3 chemical production in the summer season. For all seasons except in the post-monsoon, O_3 has noted a favorable linear correlation with temperature that has a strong adverse correlation with relative humidity. During the winter season, SO_2, NOx and CO levels were attributed to near surface anthropological sources (e.g. cars, residential kitchen, brick ovens, coal-fired energy stations, land clearance and biomass burning) and lower boundary layer. In the monsoon season, the lowest concentration of all trace gases were reported due to be efficient wet scavenging. **Saini *et al.* (2014)**[9] described the meteorological dependence of air contaminants on Agra's kerbside. The average

maximum of daily ozone ranged between 44-51 ppb during 10:00 to 16:00 hours, and in night and early morning the average daily minimum varied from 20-26 ppb attributed to photo-chemical reaction of precursor gasses from anthropogenic sources. CO and NO_2 exhibits anti-correlation on diurnal scale with ozone. The peak SO_2 and O_3 levels were appeared at humidity ≤ 30 percent and below 40 percent for CO, NO_2 and $PM_{2.5}$ for strong vertical mixing height. Winds appear to have a notable impact on pollutants.

Diurnal and seasonal variability for trace gases at Udaipur, India, have been recorded by **Yadav *et al.* (2014)**.[10] The daily average data for O_3, CO and NO_x ranges from 5 to 51 ppbv, from 140 to 795 ppbv and from 3 to 25 ppbv. During the afternoon and early morning the O_3 diurnal distribution showed a highest level and lower values. The CO-NOx trends shows inverse diurnal profile, as peaks in the morning and evening but lower levels in afternoon. The diurnal profile of O_3 is primarily controlled by changes in the depth of planetary boundary layer and photo-chemistry of precursor gases. The highest mixing ratio of O_3, CO and NO_x were observed in pre-monsoon and winter season, respectively. Whereas lowest mixing ratio of O_3, CO and NO_x were observed in monsoon season. The long-range transport linked primarily to the summer and winter monsoon circulations throughout the Indian subcontinent, on the other side governing the seasonality of O_3, CO and NO_x. **Sarangi *et al.* (2014)[11]** reported that the diurnal variations are mainly attributed to the dynamic processes, including the vertical winds and boundary layer evolution at Nainital. The highest levels of O_3, CO and NO_y in the spring have been related primarily to local pollution supplemented by northern India biomass burning. The higher CO/NO_y level also confirm that the fresh emissions at the site have a minimal impact. **Sinha *et al.* (2014)[12]** detailed the volatile organic compounds (VOC) and NO_x emission levels and photo-chemistry that form the surface ozone and secondary organic aerosol at Mohali, suburban site (northwest IGP). The burning of wheat residues induced huge rises in all measured VOCs and primary pollutants (> 3-fold baseline values). Daily surface ozone exceeded the national air quality limit of 8 hours of 100 μg O_3 m^{-3}, that was marked by long-range transportation from the west.

Bhuyan *et al*. (2014)[13] reported the variation in surface ozone and precursor gases (NOx and CO) on temporal and diel scale in the Brahmaputra basin at Dibrugarh, a sub-Himalayan location in Assam. After sunrise, the concentration in O_3 gradually increases and then decreases by about 1500 LT hours, from night to sunrise the next day. The peak

monthly concentrations of O_3 was the highest and lowest of 42.9±10.3 ppb and 17.3±7.0 ppb in March and July month, respectively. NO_x and CO concentrations precede the highest concentration of O_3. Concentration weighted trajectory (CWT) analysis were performed to determine the possible airmass path and to identify the possible pollution source region. The majority of the airmasses are confined locally in post-monsoon and winter while in pre-monsoon and monsoon, arises at the Indo-Gangetic plains, Bangladesh, and Bay of Bengal. A state-wide inventory of crop residues burned in India was presented by **Jain *et al.* (2014),**[14] which was drafted using the (2008-2009) National Inventory Guidelines of the Inter-Governmental Panel on Climate Change. The total residue produced was 620 million tons and 15.9% of residual was burnt in the fields. Rice straw was responsible for 40% of the complete residue burnt, followed by wheat straw (22%) and sugar cane wastes (20%). **Ghude *et al.* (2014)**[15] reported the crop damage caused by ozone and loses of 3.5 ± 0.8 and 2.1 ± 0.8 (Mt) for wheat and rice, respectively in central and northern India. As the variation in inventories of NOx creates up to 36 percent uncertainty in crop loss, NOx should be the main objective of decreasing pollution effects on food safety.

The analysis now shows that one NOx or NO_2 molecule is responsible for seven to nine ozone molecules in the surface. Similarly, **Swamy *et al.* (2012)**[16] discussed the impact of precursors on surface ozone at a semi-arid urban location in Hyderabad. They emphasized the role of local pollutant and meteorological emissions and the strength of other air pollutants on surface ozone levels on diurnal scale. Furthermore, they reported that greater levels of ozone were observed on weekends than on weekdays in the winter season. **Singla *et al.* (2012)**[17] observed that the daytime peak ozone concentration was recorded at Dayalbagh, Agra, approximately after noon hours corresponding to on-site photochemical manufacturing, between (12:00–14:00) hours. In addition, seasonal fluctuations in the ozone surface are indicative of the pronounced summer, winter, and minima in the monsoon/post-monsoon seasons. The inverse relationship of O_3 with the precursor gases was underlined by **Singla *et al.* (2011)**[18] at the same site**.** The days of the high ozone incident were immediately related to the hot weather and low humidity during peak summers. Based on the EPA classification, 8-hour ozone concentration stayed fine 90% of days, tolerable 8.1% of days, unhealthy 1.9% of the days during the research period. The ground-based observations of air pollutant at a semi-urban location Pantnagar in the Indo-Gangetic Plain (IGP) region were reported by **Ojha *et al.* (2012).**[19]

Because of photochemistry with local pollutants or regional dynamics, seasonal ozone levels of 32-41 ppbv and 9–14 ppbv were observed in spring-autumn and August-September, respectively. In spring season, northern India biomass burning activity act as an additional source of ozone precursor source. The correlation analyses between Pantnagar and Nainital ozone concentrations detailed the impact of IGP emissions and photochemical procedures on the air quality, especially during midday spring hours, of the pristine Himalayan region. **Mahapatra *et al.* (2012)[20]** outlined the daily (0700–1745) ozone measurements at Bhubaneswar, Odisha and their relationship with meteorological parameters in the atmosphere. This seasonal variation indicates a well-defined daily ozone peak in January (approx. 85 ppbv), followed by secondary peaks with lower amplitude in June (38 ppbv), and August (20 ppbv). During winter months, the backward trajectory analysis reveals that long distance transport of air masses were predominately from the Indo-Gangetic Plains (IGP) and the western part of the Indian peninsula, a large industrial hub. Winds from less polluted land and the Bay of Bengal reaches the point of observation in other seasons, which substantially reduces the air pollution burden. In winter, ozone anti-weekend effect (approx. 5 ppbv) has been observed.

David *et al.* (2011)[21] recorded the ozone and NO_2 measurements, with mean values of 61 ± 7 ppb and 53 ± 6 ppb respectively, during winter months in the Bay of Bangladesh. The ozone blending percentage decreased in most diurnal patterns during afternoon followed by a rise in the night time and a peak in the morning. The transport from the neighbouring land mass was strongly linked with reported levels. Anthropogenic activities, forest fires and biomass burning in the Gangetic plains and other eastern Asia led to the ozone and its precursors throughout the site. In regulating ozone variations, water vapor played a significant role, instead of NO_2 levels. **David and Nair, 2011[22]** examined the temporal variations of surface ozone levels to study the regional air quality of Kerala, Trivandruam a tropical coastal site. The ozone variation was strongly linked to the mesoscale circulation and the accessibility of NOx. Ozone peaks differ in amplitude and duration with season. Daytime ozone correlated positively to maximum temperature and negatively to absolute water vapour. **Elampari and Chithambarathanu, 2011[23]** reported surface ozone measurements at the southernmost tropical semi-urban site Nagercoil, Kanyakumari in Tamil Nadu. The measurement demonstrated that ozone concentration is time-dependent, vary from minima in morning to peak in afternoon. The annual mean daily ground ozone concentration lies between

18.91 ppb to 20.08 ppb. Ozone show significant positive correlation with ambient temperature and negative correlation with relative moisture. The highest seasonal level was reported in summer and the lowest in north-eastern monsoon. **Reddy *et al*. (2010)[24]** reported distinct variability in surface ozone levels and their precursors in Anantapur on the diurnal and seasonal period. Surface ozone concentration is high in daytime (25 to 77 ppbv) and low in late night and early morning hours on diurnal scale, because of the lower ozone titration by nitrogen oxides. The ozone levels at weekend's shows that the greater O_3 levels on weekends compared to weekdays suggest that the photochemical ozone production on this site is non-linear. The average daily levels of ozone, correlated positively with temperature (r=0.76) and negativley with relative humidity (r=0.62) among meteorological parameters. Higher O_3 levels (25-77ppbv) in the summer and lower in the monsoon season have been recorded. At a specified site, oxidant concentration ([OX] = NO_2 + O_3) is the sum of a "regional contribution" (background O_3) that is NOx-independent and a "local contribution" that is linearly dependent on NO_x (Reddy et al. 2012).

The surface ozone measurement at Nainital, a high-altitude observational has identified diurnal differences in ozone levels that do not demonstrate daytime photochemical development like urban or rural areas by **Kumar *et al.* (2010).[25]** Seasonally variation indicates a unique peak ground ozone level (above 100 ppbv) late in the spring and a minimum surface ozone level in the summer and monsoon seasons. Furthermore, the interpretation of wind trajectories and the simulation of the WRF model reveal that the stratospheric input is more important than that of regional pollutant components at the site.

2.1.2 International scenario

Regional scale, human health and crop exposure of ozone is substantial in China, as compared to other developed regions, such as Japan, South Korea, Europe, the United States and Europe (JKEU), according to **Lu *et al.* (2018)[26]**. Fourth maximum daily 8-hour average (MDA8) and the number of days with MDA8 values > 70 ppb in warm season are 6.3-30 percent and 93-575 percent higher than the JKEU regional averages. Health exposure metrics such as MDA8 and annual SOMO35 (sum of ozone means over 35 ppb) are 6.3-16 and 25-95 percent higher, respectively in China. **Nuvolone *et al.* (2018)[27]** reported the short and long term health impact assessment studies of ozone

exposure on human respiratory and cardiovascular systems. In the 2013 EU-28, 16 000 early fatalities, or 192,000 years of loss of life, are attributable to ozone exposure, according to the most latest estimates. The impacts of ozone exposure on health in EU countries have been estimated very high in Italy with 3,380 premature fatalities and loss of lives for 61 years (per 100,000 residents). **Li *et al.* (2017)[28]** reported the measurements of O_3, CO, NO_x and NMHCs, along with meteorological data at urban (ZH), semi-urban (XS) and rural (QDH) sites in Hangzhou, China. Both ZH and XS had a greater amount of ozone than QDH during the observation period, reaching 41.3 and 47.8 percent, respectively. High levels of O_3 were observed at night in QDH, explained by well-known effect of NO titration in rural areas. The increasing upwind automobiles and VOC emissions could results in ozone episodes of pollution. in urban area (ZH) along with more efficient synergistic decrease in VOCs and NOx levels. **Wang *et al.* (2017)[29]** reported a review on the characteristics of urban and rural levels of ozone, its precursor gases source, processes and their effects on crops and human health. The tremendous increase in the levels of ozone has significantly exceed that of Northern America and Europe. Ozone levels in China's major urban centres were 100-200 % higher than environmental quality standards. **Ainsworth, 2017[30]** indicate the variability of air pollutants over temporal and spatial scale and further quantify the effects on global crop productivity. The plant species have been recognized with quantitative characteristics loci (QTL) for O_3 tolerance, and while none have been cloned so far to increase the tolerance of ozone pollution in crops. Future attempts are necessary to screen population mapping on the ground and recognizing more promising phenotypes of O_3 tolerance.

Awang *et al.* (2016)[31] recorded the ozone concentrations on diurnal scale for three major port cities (Klang, Perai, and Pasir Gudang) in Malaysia. On diurnal scale, a unimodal peak appeared between 1-3 pm in the ozone levels. The higher levels of O_3 were noted during the first quarter of the study period due to an rise in local photochemical production. The results of the principal component analysis revealed that O_3 levels vary by meteorological variables at 47.7, 39.5 and 18.2 percent higher than primary air pollutants at 28.9, 32.6 and 45 percent, respectively, for all three sites, respectively. The concentrations of ozone were still below those recommended by the MAAQG, in the three ports and development was feasible in the transport industry, particularly in Klang, in case of ozone exaggerating issues of ozone.

Cooper *et al.* (2014)[32] monitored the fluctuations in the hourly O_3 level reported in seven stations in Malaysia's Borneo, rich rainforests area affected from growing urbanisation and business-industrial operations, results in changes in pollutant profiles. The highest levels of hourly average and daily maxima of O_3 were observed at sampling stations of Bintulu (S_3) were 31 and 92 ppbv. The sampling locations close to local petrochemical industries, like S_3 and S_4, reported highest NO_2/NO ratio between 3.21 and 5.67. The O_3 levels recorded in all stations during the weekend were greater relative to the weekdays, which indicates that O_3 titration by NO during weekdays has a strong influence. This research also demonstrates the effect of meteorology on distinct seasonal variations in the surface O_3 concentration. **Luna *et al.* (2014)[33]** analysed that concentration of ozone and its precursor gases (CO, NO, NO_x). Ozone is correlated negatively with its precursor gases, as expected. Similarly, negatively correlated with humidity due to the reaction between water and ozone, leading to the oxidative specie OH.

Wang *et al.* (2013)[34] noted the highest levels of NO_x in winter due to increased pollutant emissions and unfavourable dispersion conditions, while the concentration of O3 levels were higher during the summer. A positive correlation between O_3 to CO was found at the background site where air mass was largely anthropogenic emissions, which exhibit photochemical processing, while negative slope both rural and urban were shown to have a major impact on the local emission of precursors. Positive correlations between NO_x to O_3 in rural and background areas indicated that NOx helped to create ozone, while a negative correlation at urban site involved titration of ozone due to high level of NO. **Alenezi *et al.* (2012)[35]** reported the rising levels of criteria air pollutants ((NM-HC, CH_4, CO, CO_2, O_3, SO_2, NO_2 and PM_{10}) along with influence of meteorological variables in Al Jahra, Kuwait. In order to identify the most likely sources of use of future mitigation methods for pollution control, the seasonal fluctuations in pollutants concentrate from winter and summer was analyzed. With the exception of O_3, CO_2 and PM_{10}, the pollutant concentrations are greater in winter rather than summer due to limited dispersion and shallow boundary layer. All pollutant levels, except NMHC, are below the permissible norms. **Xu *et al.* (2011)[36]** showed that trace gas concentrations are highly dependent on wind and provides information regarding regional pollution features. There were also clear temperature and relative humidity dependency on O_3 levels.

2.2 Air quality prediction approaches

The increasing diversity of statistical literature over the last decade, together with the use of very different data sets and performance indices. The statistical method has been improved with a range of linear and nonlinear regression techniques for various geographical areas. The lack of availability and uncertainty of most of the inputs of deterministic techniques, advanced statistics with excellent potential for predicting air pollution **(Mohan *et al.* 2007).**[37] Statistical modelling approaches were developed in order to better record the target variable's performance. The selection of the two site particular predictor factors (pollutant and meteorological variables), together with relationships between the predictor factors, was comprehensive. Predictions can also be influenced by the selection of the structure of the model, training methods and sample records. Sometimes single linear or non-linear models might not be enough to capture all the relationships among time-related input variables. Hence, no single linear or non-linear approach is fitting suitably in every situation. Considering these limitations, further ensemble/ hybrid model was generated with the aim of retaining their advantages and avoids overfitting to get the best predictive combination. Combining linear and non-linear models will increase the chance to procure different patterns in the facts over the individual model and enhance forecasting performance. Air quality forecasting is very consistent and efficiently controlling the measures and can be proposed as a preventive action for regulations and protect public health.

2.2.1 National scenario

Garaga *et al.* (2018)[38] reported the status of modelling approaches in India. The examination of the local and regional numerical models showed that the emission outputs need to be improved, whereas the statistics showed the need to select main tracers for the distribution of sources for better predictions. The significant sources of particulate matter in India include dust and traffic emissions. **Naveen and Anu, 2017**[39] evaluated the the air quality assessment and the air quality index prediction of the basis of database supervised by three monitoring stations of Kerala State Pollution Control Board (KSPCB) in Thiruvananthapuram District, Kerala, India. The Auto Regressive Integrated Moving Average (ARIMA) and Seasonal Auto Regressive Integrated Moving Average (SARIMA) method methods have been employed for air quality predictions. As

compared to the SARIMA models, the ARIMA models produced satisfying results that can be combined with other models in order to produce more exact results.

Kumar *et al.* (2017)[40] presented a review of the studies related to air quality monitoring and modelling of surface ozone and aerosol optical depth, in South Asia. Surface ozone production is largely limited by emissions of NOx. The air quality is deteriorated by dust storms by the increase of the load of particulate matter, but for the trace gasses like ozone act as a sink. Since 1990, ozone concentrations have risen by approximately 5–30 % and will continue to raise. Exposure to greater ozone concentrations considerably reduces production of wheat and rice and reduces the life expectancy in South Asian. The results are quantified using empirical relation based on based on studies carried out by **Sinha *et al.* (2015)[41]** in North America and Europe. **Niharika and Venkatadri, 2014[42]** stated that soft computing is regarded an outstanding technology for air quality forecast and data analysis. The study was carried by collecting air pollutant information and built an optimized model for prediction of air quality, for Nagpur, Maharashtra. Artificial Neural Networks (ANN) has demonstrated to be the most optimized information analyser of air quality among different soft computing methods.

Sharma *et al.* (2013)[43] reported the prediction of peak surface ozone at Suchindrum, Tamilnadu, delineating the peak values on diurnal scale at 1430 hours and minimum around 0530 hrs. The neural network models are suggested to estimate the maximum concentration of surface ozone (1430 hrs) according to the multiple input parameters, such as temperature, relative humidity and nitrogen dioxide, corresponding to the morning hours. The network produces excellent forecasts with a correlation coefficient ($r = 0.80$) between observed and forecasted concentrations. **Singh *et al.* (2013)[44]** conducted the principal components analysis to identify sources of air pollution and tree-based ensemble learning models to forecast Lucknow's (India's), urban air quality using databases over a five-year period. The results indicated that the significant source of air pollution are fuel combustion and vehicular emissions. The decision tree and supported vector machines model were built and several statistical parameters assessed for their generalization and predictive efficiency. The seasonal air quality contributing the misclassification rate of 8.32% (single decision tree); 4.12% (decision tree boost); 5.62% (decision tree boost), and 6.18% (support vector machine), respectively in complete data. The ensemble models have been designed to distinguish between seasonal air quality,

discriminatory variables, and air quality indices. In summer and winter, the air quality was unhealthy indicated by air quality indices. In classification and regression, both the DTF and DTB model were higher than the SVM, which could be attributed to the integration of bagging and boosting algorithms within such models. The suggested models of the ensemble effectively forecast the quality of urban ambient air. **Singh *et al.* (2012)[45]** noted that non-linear (MPR, ANNs) models were comparatively more efficient than linear partial least squares regression models, whereas ANN models performed more satisfactory then the low-order non-linear multivariate polynomial regression models. Among the three (multilayer perceptron, radial-basis function, and generalized regression neural network ANN models, GRNN over-performed with high correlation between measurements and model predicted values for the training, validation and test sets. In the course to enhance air quality index prediction outcomes one day in advance by the use of Delhi's prior day's AQI and meteorological variables, have introduced a PCA-Multilinear Regressions (MLRs) and PCA-neural network models by **Kumar and Goyal, (2011).[46]**

Kandya *et al.* (2009)[47] reported the statistical approach for the prediction of urban air quality over Delhi, using five models (Single Exponential Smoothing (SES), Adaptive Response Rate Single Exponential Smoothing (ARRSES), Holt's Linear Method (HLM) ARX (Auto Regressive eXogenous) Model and Auto Regressive Integrated Moving Averages (ARIMA).The findings indicate that there is no single modeling strategy that produces optimal outcomes in terms of performance indicators. In this research, ARIMA method ranks well above other methods among the five statistical methods.

2.2.2 International scenario

Pucer *et al.* (2018)[48] reported that air quality monitoring sites across Slovenia, show that both PM_{10} and O_3 predictions, Bayesian models outperform classic models. The proposed models perform better than PM_{10} experts and are on par with O_3 experts, already base their forecasts on statistical model predictions. **Zhai *et al.* (2018)[49]** developed a stacked ensemble model, for forecasting and analyzing the daily average concentrations of fine particulate matter (PM2.5) in Beijing, China. In order to select key variables and assess feature significance levels, stability selection and tree-based techniques of choice of features are implemented. Special feature extraction procedures, including those of simplification, polynomial, transformation and combination, are conducted before

modeling to identify potentially significant features based on an exploratory data analysis. Single models including LASSO, Adaboost, XGBoost and multi-layer perceptron optimized by the genetic algorithm (GA-MLP) are developed in the level 0 space and then integrated by support vector regression (SVR) in the level 1 space via stacked generalization. The ensemble model assessment indicates, that, the coefficient of determination (R^2) of 0.90 and root-mean squared error (RMSA) of 23.69 μg / m^3, usually is better than a single non-linear model predictor. The general precision of classification is 73.93%, with most misclassifications among adjacent categories. These findings show that the stacked ensemble model can be interpreted and generalized.

Mok *et al.* (2018)[50] implemented the efficacy of the adaptive Bayesian Model Averaging scheme in a forecasting system, which is used to provide a weighted average forecast, to assess probabilities for all ozone prediction models (ozone season, non-ozone season and transitional period). While, considering the strong seasonal cycles of the season, seasonal time variables are typically intended for predicting the daily ozone level. The sudden changes of models into their specified time span can influence their efficiency, particularly during the transitional season. The system for the daily peak 8-hour average ozone levels is implemented, with satisfactory outcomes. A linear regression model for the predictive daily maximum 8-hour (MDA8) levels of previous spring, on the basis of large-scale climate patterns was developed by **Shen and Mickley (2017)**.[51] The high level of ozone was correlated with springtime patterns of warm tropical Atlantic and cold northeast Pacific sea surface temperatures (SSTs), as well as positive sea level pressure (SLP) anomalies over Hawaii and negative SLP anomalies over the Atlantic and North America. This model indicated that the MDA8 concentration of ozone and amount of ozone episodes (> 70 ppbv) have a variation of 45 and 30 percent over East America.

Biancofiore *et al.* (2015)[52] analysed the four statistical criteria (correlation coefficient, fractional preference, normalized mean squared error and a factor of two) for the performance of neural network model, in a seaside town in central Italy. All criteria have been shown that in all model scenarios the neural network produces better outcomes than the regression model for ozone. In the town of Nova Gorica, Slovenia, the regressor selection techniques for ozone forecasting were used by **Kocijan *et al.* (2015).[53]** For various models distinct regressors have been chosen with various techniques depending on the cost functions of the validation procedure. Ten regressors have been selected,

namely, for the model to estimate maximum ozone daily concentration. For the average ozone concentrations between 8 a.m. and 20 p.m., fifteen regressors are selected and ten regressors have been selected for the average daily concentration.

In comparison to traditional PCA, **Zhang *et al.* (2014)[54]** proposes the PCA models with non-parametric T_2 control limit to predict ozone exceedance days. Indicators of process variation also contribute considerably to the identification of days of ozone exceedance. Multiple regressions combining PCA were used by **Rajab *et al.* (2013)[55]** and it showed that O_3 was correlated negatively with CH_4, H_2O vapour, RH, and positively with CO, AST (air surface temperature), SSKT (skin surface temperature), and AT (atmosphere temperature) during both the north east and south west monsoon periods. A variable selection method based on varimax rotating principal components was applied, to acquire the subset of predictor variables for linear regression model. The combined method based on both multiple regression and principal component was applied to enhance the prediction accuracy. For both north east and south west monsoon, the result of 8 independent variables for the best models for the columnar ozone, indicated about the same values as the r (0.93) and r^2 (0.86). **Arhami *et al.* (2013)[56]** reported the high correlations between the model and observation of hourly concentrations of pollutants of CO, NO_x, NO_2, NO and PM_{10}. ANN models were then designed and validated. Combining the ANNs with the MCSs based on Latin Hypercube Sampling (LHS) was then developed into the calculation of prediction interval (PI) methodologies and probability of exceeding the air quality thresholds.

Saithanu and Mekparyup (2013)[57] assessed two distinct methods for evaluating and forecasting the ozone level exceeds the air quality standard in the east of Thailand for the period 2006-2010. Firstly, the traditional procedure, the discriminant analysis model and secondly, neural network the model. The daily maximum ozone concentration and other atmospheric factors were used for the training and validation of both predictive models and the model performance were evaluated by correct classification rate. The outcome of results shows that the neural network model is used to overcome both the training and validation information set over discriminant analysis model. The average CCR of 87.22 % , 86.58% for the training and validation dataset for the neural network model while for the discriminant analysis model, 79.77% and 78.98% for training and testing, respectively. **Pires and Martins, (2011)[58]** suggested two techniques for improving the

performance of statistical models in the forecast of levels of tropospheric ozone. The first approach corrected the statistic model based on the average daily profile of the model error in the training set and results in improved the results with the ANN model. The second technique estimated the model error by analogizing three fundamental feedback control modes: proportional, integral and derivative, and gathered the highest hourly average O_3 concentration forecast with the correction of MLR model. O_3 levels have adverse relation to NO_2 reported by **Ghazali *et al.* (2010)[59]** using a MLR model, and this is typical of metropolitan regions where traffic that discharges NO_2 has the effect on the diurnal cycle of O_3. **Konstantinos *et al.* (2010)[60]** forecasted the maximum daily value (24 to 72 hours ahead) of the European Regional Pollution Index as well as the number of consecutive hours in a day, with at least one of the pollutants above a threshold concentration, using Artificial Neural Networks. The results were in agreement at the statistical rate of $p < 0.01$ in accordance with monitored data.

2.3 Research Gaps

Air quality data is generated on real time basis at various stations in the country, still it is not covering the entire land-area. It is a costly and time-consuming process. To overcome this issue, the concept of forecasting using statistical approach comes in light, simultaneously it cut-down the requirement of stations and cost on producing the data. This approach aims to protect the population from peak concentrations and fosters severe episode prevention also through the implementation of air quality. In this process, there is a need to study the variable relationships in large databases such as air pollution and meteorological parameters that can provide important information concerning the nature of dependencies and trends in the data set. During recent decades, many efforts have been devoted in exploring the relationships between secondary pollutant (ozone), precursor gases and meteorological parameters. Therefore, several statistical methodologies have been proposed. These methodologies are often based on linear or nonlinear regression models where air pollution concentration, in a specific location, is related to meteorological parameters. Sometimes single linear or non-linear models might not be enough to capture all the relationships among time-related input variables.

Keeping in mind the following research gaps identified from the survey of literature the research work has been done under the PhD programme:

- *No single linear or non-linear approach is fitting suitably in every situation. Considering these limitations, further ensemble/ hybrid model could be generated with the aim of retaining their advantages and avoid overfitting to get the best predictive combination.*
- *Application of Multivariate statistical methods have not been applied extensively in predicting ozone (O_3) concentrations at the ground level of the troposphere as a function of precursor gases and meteorological parameters. These models acting as a useful and effective tool, have not been applied successfully to air pollution investigation and management in quite a number of populated cities and regions over the world.*
- *Regression trees were used to model peak daily levels of ozone. The trend component in such models cannot, however, be recorded immediately. Finally, non-linear regression may be suitable if the nature of the physical issue requires a nonlinear connection. For exploratory analysis, the conditions in the model were created using graphical techniques and nonparametric regression to determine the outstanding relationships involving ozone and many candidate predictor variables.*

2.4 References

1. Gardner, M.; Dorling, S. *Atmospheric Environment* **2000**, *34*, 21-34.
2. Pancholi, P.; Kumar, A.; Bikundia, D. S.; Chourasiya, S. *Sustainable Environment Research* **2018**, *28*, 79-89.
3. Kumari, S.; Verma, N.; Lakhani, A.; Tiwari, S.; Kandikonda, M. K. *Environmental Science and Pollution Research* **2018**, *25*, 18879-18893.
4. Sharma, S.; Khare, M. *Atmospheric Environment* **2017**, *151*, 117-132.
5. Verma, N.; Lakhani, A.; Kumari, K. M. *Atmospheric Research* **2017**, *197*, 232-243.
6. Peshin, S.; Sharma, A.; Sharma, S.; Naja, M.; Mandal, T. *Sustainable Cities and Society* **2017**, *35*, 740-751.
7. Sharma, A.; Sharma, S. K.; Mandal, T. K. *Sustainable Environment Research* **2016**, *26*, 76-83.
8. Gaur, A.; Tripathi, S.; Kanawade, V.; Tare, V.; Shukla, S. *Journal of Atmospheric Chemistry* **2014**, *71*, 283-301.
9. Saini, R.; Singh, P.; Awasthi, B. B.; Kumar, K.; Taneja, A. *Atmospheric Pollution Research* **2014**, *5*, 796-804.
10. Yadav, R.; Beig, G.; Jaaffrey, S. *Atmospheric Environment* **2014**, *85*, 147-151.
11. Sarangi, T.; Naja, M.; Ojha, N.; Kumar, R.; Lal, S.; Venkataramani, S.; Kumar, A.; Sagar, R.; Chandola, H. *Journal of Geophysical Research: Atmospheres* **2014**, *119*, 1592-1611.
12. Sinha, V.; Kumar, V.; Sarkar, C. *Atmospheric Chemistry and Physics* **2014**, *14*, 5921-5941.
13. Bhuyan, P. K.; Bharali, C.; Pathak, B.; Kalita, G. *Environmental Science and Pollution Research* **2014**, *21*, 6696-6713.
14. Jain, N.; Bhatia, A.; Pathak, H. *Aerosol and Air Quality Research* **2014**, *14*, 422-430.
15. Ghude, S. D.; Jena, C.; Chate, D.; Beig, G.; Pfister, G.; Kumar, R.; Ramanathan, V. *Geophysical Research Letters* **2014**, *41*, 5685-5691.
16. Swamy, Y.; Venkanna, R.; Nikhil, G.; Chitanya, D.; Sinha, P.; Ramakrishna, M.; Rao, A. *Aerosol and Air Quality Research* **2012**, *12*, 662-671.
17. Singla, V.; Pachauri, T.; Satsangi, A.; Kumari, K. M.; Lakhani, A. *The Scientific World Journal* **2012**, *2012*.
18. Singla, V.; Satsangi, A.; Pachauri, T.; Lakhani, A.; Kumari, K. M. *Atmospheric Research* **2011**, *101*, 373-385.
19. Ojha, N.; Naja, M.; Singh, K.; Sarangi, T.; Kumar, R.; Lal, S.; Lawrence, M.; Butler, T.; Chandola, H. *Journal of Geophysical Research: Atmospheres* **2012**, *117*.
20. Mahapatra, P.; Jena, J.; Moharana, S.; Srichandan, H.; Das, T.; Chaudhury, G. R.; Das, S. *Journal of Earth System Science* **2012**, *121*, 1163-1175.
21. David, L.; Girach, I.; Nair, P. *Annales Geophysicae (09927689)* **2011**, *29*.
22. David, L. M.; Nair, P. R. *Journal of Geophysical Research: Atmospheres* **2011**, *116*.
23. Elampari, K.; Chithambarathanu, T. *International Journal of Science and Technology* **2011**, *1*, 80-88.
24. Reddy, B. S. K.; Kumar, K. R.; Balakrishnaiah, G.; Gopal, K. R.; Reddy, R.; Ahammed, Y. N.; Narasimhulu, K.; Reddy, L. S. S.; Lal, S. *Atmospheric Research* **2010**, *98*, 125-139.
25. Kumar, R.; Naja, M.; Venkataramani, S.; Wild, O. *Journal of Geophysical Research: Atmospheres* **2010**, *115*.
26. Lu, X.; Hong, J.; Zhang, L.; Cooper, O. R.; Schultz, M. G.; Xu, X.; Wang, T.; Gao, M.; Zhao, Y.; Zhang, Y. *Environmental Science & Technology Letters* **2018**, *5*, 487-494.

27. Nuvolone, D.; Petri, D.; Voller, F. *Environmental Science and Pollution Research* **2018**, *25*, 8074-8088.
28. Li, K.; Chen, L.; Ying, F.; White, S. J.; Jang, C.; Wu, X.; Gao, X.; Hong, S.; Shen, J.; Azzi, M. *Atmospheric Research* **2017**, *196*, 40-52.
29. Wang, T.; Xue, L.; Brimblecombe, P.; Lam, Y. F.; Li, L.; Zhang, L. *Science of the Total Environment* **2017**, *575*, 1582-1596.
30. Ainsworth, E. A. *The Plant Journal* **2017**, *90*, 886-897.
31. Awang, N. R.; Elbayoumi, M.; Ramli, N. A.; Yahaya, A. S. *Air Quality, Atmosphere & Health* **2016**, *9*, 25-39.
32. Cooper, O. R.; Parrish, D.; Ziemke, J.; Cupeiro, M.; Galbally, I.; Gilge, S.; Horowitz, L.; Jensen, N.; Lamarque, J.-F.; Naik, V. **2014**.
33. Luna, A.; Paredes, M.; De Oliveira, G.; Corrêa, S. *Atmospheric Environment* **2014**, *98*, 98-104.
34. Wang, Y.; Shen, L.; Wu, S.; Mickley, L.; He, J.; Hao, J. *Atmospheric Environment* **2013**, *75*, 374-382.
35. Alenezi, R.; Al-Anzi, B.; Abusam, A.; Ashfaque, A. *Journal of Environmental Protection* **2012**, *3*, 1711.
36. Xu, W.; Zhao, C.; Ran, L.; Deng, Z.; Liu, P.; Ma, N.; Lin, W.; Xu, X.; Yan, P.; He, X. *Atmospheric Chemistry and Physics* **2011**, *11*, 4353-4369.
37. Mohanakumar, K. *Stratosphere Troposphere Interactions: An Introduction*; Springer Science & Business Media, 2008.
38. Garaga, R.; Sahu, S. K.; Kota, S. H. *Current Pollution Reports* **2018**, *4*, 59-73.
39. Naveen, V.; Anu, N. *J Eng Res Appl* **2017**, *7*, 66-84.
40. Kumar, N.; Middey, A.; Rao, P. S. *Urban Climate* **2017**, *20*, 148-167.
41. Sinha, B.; Singh Sangwan, K.; Maurya, Y.; Kumar, V.; Sarkar, C.; Chandra, B.; Sinha, V. *Atmospheric Chemistry and Physics* **2015**, *15*, 9555-9576.
42. Niharika, V. M.; Rao, P. S. *International Journal of Computer Science and Information Technologies* **2014**, *5*, 103-107.
43. Sharma, R. K.; Thanu, T. C.; Elampari, K. *Universal Journal of Environmental Research & Technology* **2013**, *3*.
44. Singh, K. P.; Gupta, S.; Rai, P. *Atmospheric Environment* **2013**, *80*, 426-437.
45. Singh, K. P.; Gupta, S.; Kumar, A.; Shukla, S. P. *Science of the Total Environment* **2012**, *426*, 244-255.
46. Kumar, A.; Goyal, P. *Science of the Total Environment* **2011**, *409*, 5517-5523.
47. Kandya, A.; Mohan, M. In *Proceedings of the 7th international conference on urban climate*, 2009.
48. Pucer, J. F.; Pirš, G.; Štrumbelj, E. *Knowledge and Information Systems* **2018**, *57*, 635-654.
49. Zhai, B.; Chen, J. *Science of the Total Environment* **2018**, *635*, 644-658.
50. Mok, K.; Yuen, K.; Hoi, K.; Chao, K.; Lopes, D. *Stochastic Environmental Research and risk assessment* **2018**, *32*, 1283-1297.
51. Achakulwisut, P.; Shen, L.; Mickley, L. J. *Journal of Geophysical Research: Atmospheres* **2017**, *122*, 12,449-412,467.
52. Biancofiore, F.; Verdecchia, M.; Di Carlo, P.; Tomassetti, B.; Aruffo, E.; Busilacchio, M.; Bianco, S.; Di Tommaso, S.; Colangeli, C. *Science of the Total Environment* **2015**, *514*, 379-387.
53. Kocijan, J.; Hančič, M.; Petelin, D.; Božnar, M. Z.; Mlakar, P. *Simulation Modelling Practice and Theory* **2015**, *54*, 101-115.
54. Zhang, H.; Palazoglu, A.; Zhang, X.; Zhang, W.; Zhao, Z.; Sun, W.; Liu, S. *Chemometrics and Intelligent Laboratory Systems* **2014**, *133*, 42-48.

55. Rajab, J. M.; MatJafri, M.; Lim, H. *Atmospheric Environment* **2013**, *71*, 36-43.
56. Arhami, M.; Kamali, N.; Rajabi, M. M. *Environmental Science and Pollution Research* **2013**, *20*, 4777-4789.
57. Saithanu, K.; Mekparyup, J. *International Journal of Pure and Applied Mathematics* **2013**, *84*, 109-121.
58. Pires, J.; Martins, F. *Atmospheric Environment* **2011**, *45*, 2413-2417.
59. Ghazali, N. A.; Ramli, N. A.; Yahaya, A. S.; Yusof, N. F. F. M.; Sansuddin, N.; Al Madhoun, W. A. *Environmental Monitoring and Assessment* **2010**, *165*, 475-489.
60. Moustris, K. P.; Ziomas, I. C.; Paliatsos, A. G. *Water, Air, & Soil Pollution* **2010**, *209*, 29-43.

MATERIAL & METHODOLOGY

CHAPTER-3

This chapter describes techniques, the principle of operation and calibration of instruments employed for the measurements of atmospheric trace gases and the local meteorology which formed the basis of a primary database of this study. However, the secondary database sets include active fire counts and air mass backward trajectories that show the more realistic motion of air parcels and their transport pathways. Further, a wide range of statistical methodologies was applied to develop satisfactory air quality forecasting of trace gases. Enhanced trace gas concentrations can intimidate life and make air pollution one of the main problems of the North West Indo Gangetic Plains region.[1] It is a subject of profound research concern and limited research work was done in the tropical regions. Air quality predictions are extremely coherent and efficient in the management of interventions and can be suggested as a preventive measure for legislation and public health protection.[2,3] The description of the observational site and its meteorology are briefly elaborated in section 3.1 and 3.2, respectively. The information of different measurable parameters and their instrumentation is described in section 3.3. The realistic motion of air parcels and active fire counts in the proximity of the study area during the research period are described in section 3.4 and 3.5. Standard prescribed methods were used and all precautions were taken during different type of data collection and analysis.

3.1 Site description

The Patiala area covers a surface area of 3430 km^2 in Punjab State, India, between 29° 49′ to 30° 47′ north and 75° 58′ to 76° 54′ east. It is located in the northwestern section of the Indo-Ganges plain, in the eastern section of Shivalik Hills and in the middle of the Thar Desert. The observations were carried out at the Thapar Institute of Engineering and Technology at a station (Latitude- 30° 21′ North, Longitude- 76° 21′ East) located in Patiala city. Figure 3.1 shows a schematic design of the site of observation. The Continuous Ambient Air Quality Monitoring Station (CAAQMS) was established with the expertise of the Indian Institute of Tropical Meteorology (IITM), Pune by Envirotech Online Equipment's (P) Ltd. Hyderabad on a continuous basis under MAPAN (Modelling Atmospheric Pollution and Networking) project.

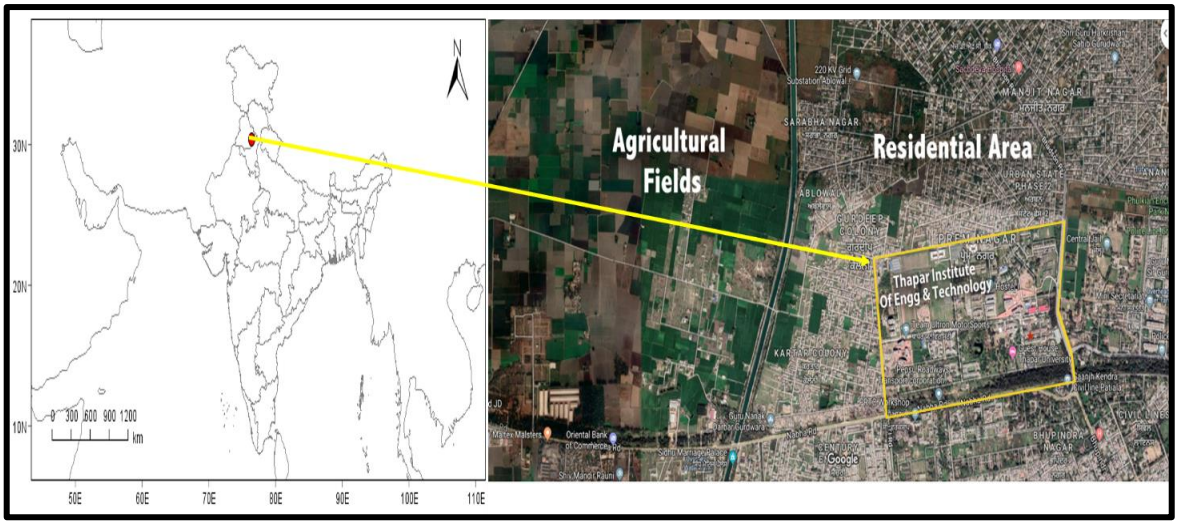


Figure 3.1: The geographical location of Continuous Ambient Air Quality Monitoring Station situated at Thapar Institute of Engineering and Technology, Patiala (India)

The station is installed on the second floor and the inlet system of analyzers and samplers is installed at the rooftop (12 m height) from the ground level. As according to records of 2011 census Patiala city, the 4th most populated district of Punjab had a population of 4,06,192 and population density of 10950 /sq. km which has grown at a rate of 25%. The study site is encircled by the inhabited colonies, inter-linked roads (NH-64 in north-east, NH-1in north and State Highway no. 8) and agricultural land. A bulk of the district population depends on agriculture for their livelihood either directly or indirectly. Punjab state is ranked as second-biggest food grain producing source of the country and adds 37.8% of total rice grains and 57% of wheat grains to the food basket of the country.[4] Patiala is a significant industrial development center on the national industrial plan, and the industrial areas in Rajpura, Patiala, Samana, and Nabha are dispersed throughout the district. Anthropogenic emissions of both trace gases and aerosols from domestic work, vehicular exhaust and industries are present throughout the year near the observational site along with a seasonal practice of agricultural crop residue burning for crop rotation.[5,6]

3.2 Basic meteorology of the observation site

Winter (December-February), summer (March-May), monsoon (June-August), and post-monsoon (September-November) are the four main seasons for the area. In summer, the city's climate is very warm and in winter, very cold. The region is generally dry and warm with maximum temperature of 42-46 °C during May. January

is the coldest month at a minimum average temperature of 1-5 °C. This region is likewise laid low with the western disturbances, a well-known synoptic meteorological phenomenon that is universal in winter and summer seasons. Annual rainfall accounts for 600 mm in the monsoon season, which lasts for three months. The wind rose for the corresponding study period has been plotted in Figure 3.2 and taken into consideration while analysing the data. These wind plots show the wind direction with a frequency of wind blowing from along with the wind speed (m/sec) based on hourly wind data. The length of each "spoke" is associated with the frequency of time of the blown wind from a given direction. The y-axis represents the increasing percentage frequencies of each bin, emanating from zero at the centre. The fetch region ranges from North-West to North- East. Most common direction of the wind at the site can be seen from the South-East followed by North-West, South-South West, and East North East.

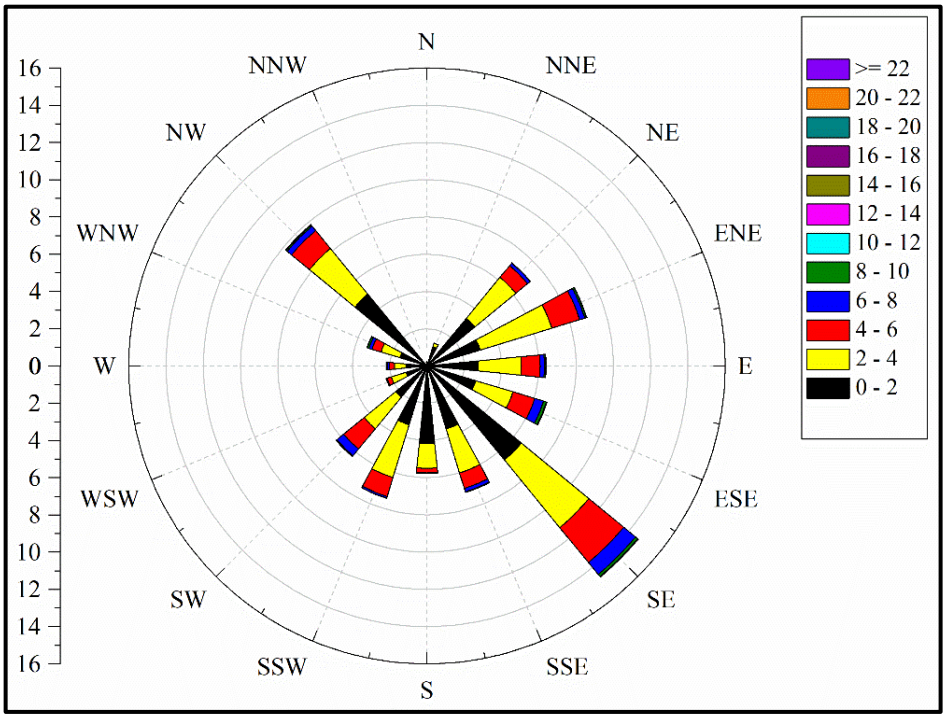


Figure 3.2: Wind rose plots showing wind direction, speed and frequency in the study area for the period in January 2013 to December 2015

3.3 Instrument description

The data used was collected from Continuous Ambient Air Quality Monitoring Station established at TIET, as shown in Figure 3.3. Analyzers and samplers used in the system are based on USEPA equivalent referenced methods. A dynamic gas calibrator (Ecotech Gas Cal 1100) in combining with the zero air supply and a set of standard gas mixtures was used for the daily zero / span checks.

The one hourly averaged data of surface meteorological variables (ambient temperature, solar radiation, relative humidity, wind speed, and wind direction) retrieved for the given duration.

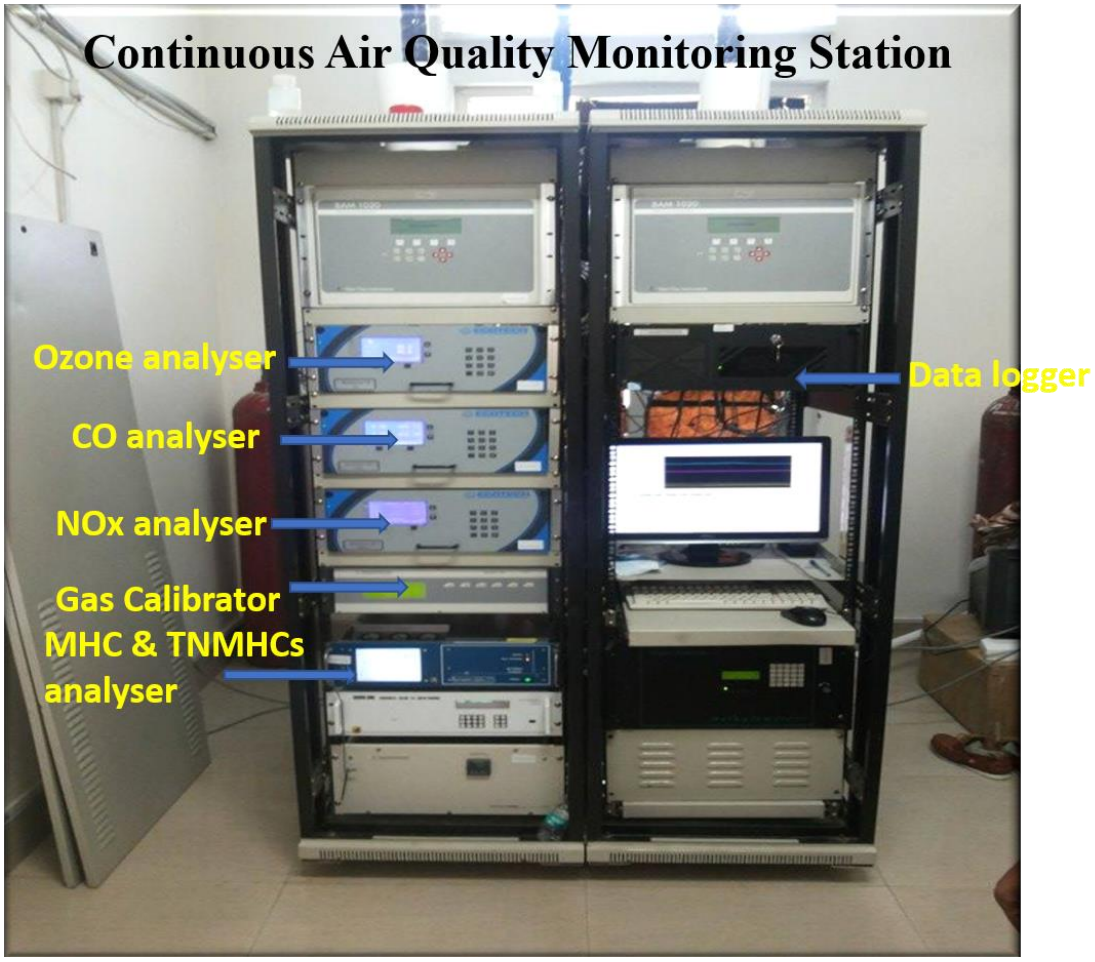


Figure 3.3: Continuous Ambient Air Quality Monitoring Station established at TIET, Patiala ((Latitude- 30° 21′ North, Longitude- 76° 21′ East)) by Envirotech Online Equipment (P) Ltd.

3.3.1 Measurement of surface ozone

Measurement theory

In the present study, surface O_3 is measured by an analyzer (Serinus 10, Ecotech), which works on the UV light absorption principle (UV photometry). The ozone level is measured in a sample gas by detecting UV absorption at 253.7 nm by O_3 molecule at the ambient pressure.

The sample air to be analyzed is passed through a Teflon 5 µm filter, before it is directed to the measurement cell to remove all particles bigger than 5 µm, susceptible to sample measurements. The two components of the sampling are the reference cycle (O_3

free air) and the sample cycle conducted by the ozone analyzer. The sample gas and the O_3-free reference gas are passed through UV light absorption cell and the intensity is measured by the light detector. The detected UV signal strength is proportional to the quantity of UV light absorbed by ozone, and the Beer-Lambert relationship is used for ozone concentration calculations. Ozone is not the only gas which absorbs UV (254 nm) but is also absorbed by SO_2 and aromatic compounds. The sample air passes through an ozone scrubber, MnO_2, to avoid interference with all gasses and thus removes ozone from the samples measurement signal, which is the precise measurement of the interfering gasses. The O_3 analyzer is streamlined in Figure 3.4. The major components include mercury lamp (UV source), photodiode detector and ozone scrubber (Manganese dioxide, MnO_2).

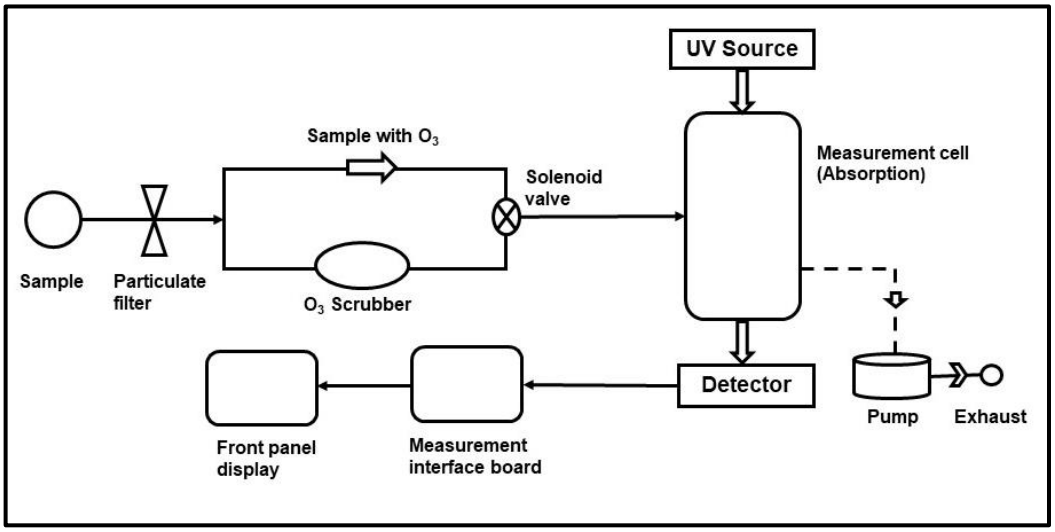


Figure 3.4: A simplified layout of the ozone analyzer showing the working principle of the analyzer

Instrument description

The analyzer series Serinus utilizes the sophisticated Kalman digital filter. This filter offers the best possible compromise for the type of signal and noise current in ambient air monitoring systems, between response time and noise decrease. The here used analyzer also includes correctness's due to temperature modifications, pressure/flow rate as well as lamp intensity drift. The device can work over a range of 0-45° C. The analyzer has zero drift below 2 ppbv and a drift span below 1 percent per month. For gas temperature and pressure changes, the ozone concentration is automatically corrected to 0°, 20°, or 25° C at 1 atmosphere. The specimen sample rates are approximately 1-3 litters per minute and the instrument's Lower Detectable

Limit (LDL) is approximately 1 ppbv. The analyzers response time is approximately 20 seconds. It is recorded that the absolute precision of this equipment is about 5%.[7]
Five primary assemblies consisting of ozone analyzer are:

- The pneumatics to transfer sample and exhaust gas.
- The ozone (optical cell) measurement devices and other parameters relevant.
- All circuit boards are controlling the control scheme for all sensors and pneumatic elements.
- The power supply that provides all device processors with energy.
- Data access communications module.

Calibration

Periodic multi-point calibration and subsequent zero / span checks are required for the analyzer. They were performed using Ecotech Gas Cal 1100, a gas calibrator.[8] The maintenance and calibration protocol as recommended by the supplier for ozone analyzer was followed. The analyzer's zero point was set with zero calibration. Span calibrations calibrate the device to the maximum cap of normal monitoring. Ecotech proposes that 80% of the full-scale calibration in environmental surveillance circumstances be adequate (i.e. 1ppm).

Precision check

In this, the instrument is run through a known concentration of span gas or zero air. The observation is made with no adjustment and it is called level two calibration. Manual or automatic mode can be used for precision check. In case of failure of precision check by the instrument, again perform the span calibration or zero calibration.

3.3.2 Measurement of NO_x

Measurement theory

The Serinus 40 oxides of nitrogen analyzer comply continuous analysis of nitric oxide (NO), total nitrogen oxides (NO_X), and nitrogen dioxide (NO_2) using gas-phase chemiluminescence. The primary source of NO_X is from automobile exhaust in their corresponding forms and in the formation of ozone (O_3) or acid rain, which are hazardous to human health. As an asthma contributor and smog observed in many industrialized towns, NO_x has a high prevalence in our societies.

It works on the principle of NO oxidation by ozone, with a peak at 630 nm and produces chemiluminescence effect. NO_2 is converted into NO by molybdenum converter of higher sensitivity with an efficiency of ~100% conversion. The two modes of operation exist i.e. NO mode and NO_X mode. In NO mode, NO in the sample of ambient air is measured by the light intensity emitted from chemiluminescence reaction and detected by Photo Multiplier Tube (PMT). Ambient air directly enters into the reaction chamber and reacts with ozone. In NO_X mode, NO_2 is converted into NO which goes into the reaction chamber and ambient air flows through molybdenum converter. It is used to suck the ambient air sample which is heated to 625° C temperature. NO_2 gets converted into NO in the reaction chamber. Thus, NO_X (NO_2 + NO) is measured in this mode. NO_2 is measured by the difference between NO and NO_x. 5 groups of NO and 5 groups of NO_x are measured by the instrument in a group of 10 (5 of each), and per cycle, it swaps 3 times between the channels. Background noise is the sample per cycle is filtered from the signal. Figure 3.5 represents the basic layout of nitrogen analyzer. Photomultiplier tube (PMT), Ozone source with reaction chamber and Delay coil constitutes the major components.

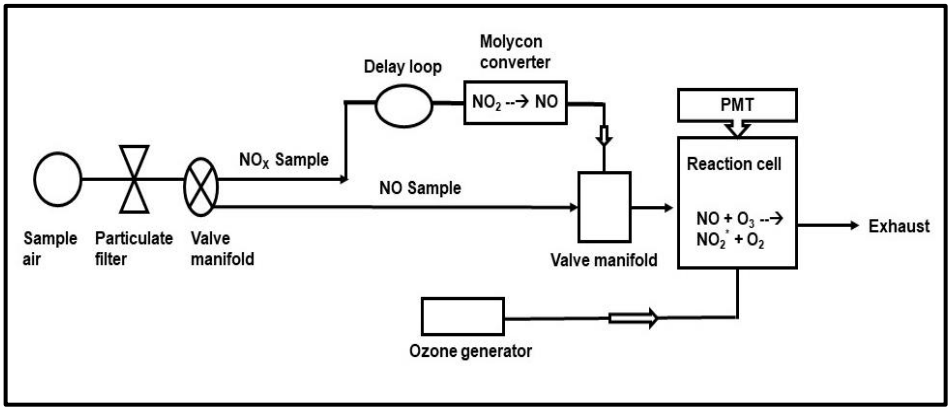


Figure 3.5: A simplified layout of the NO_X analyzer showing the working principle of the Analyser

Instrument description

For NO_x measurements of 0-20ppm, the Serinus 40 analyzer measures with the sensitivity of 0.4 ppb. For gas temperature and pressure change, the concentrations are automatically corrected and referenced at 1 atmosphere to 0°, 20°, and 25°C. The sophisticated Kalman Digital Filter offers the best possible compromise for the type of signal and noise in ambient air analyzers between response time and noise reduction. The nitrogen oxides consist of five primary assemblies:

- The sample and exhaust gas transmission pneumatics.
- Nitrogen (optical cell) oxide measurement sensors and other appropriate parameters.
- The control scheme that includes the control of sensors and pneumatics on all circuit boards.
- The energy supply that powers all the pneumatic and sensor control circuit panels.
- The power that supplies power to all procedures of the tool.
- Data access communications module.

Calibration

Zero calibrating utilizes zero air (NO_x scrubbed ambient air) to calibrate/check the instrument's reduced range. To set the analyzer's zero point, zero calibrations are used. Span calibrate the device to the upper boundaries of the regular monitoring. Ecotech recommended the calibration in environmental surveillance scenarios 400 ppb should be adequate for 80 % of the entire scale. The multipoint comprises a span gas instrument at several recognized levels, which records the instrument output. The linearity of the concentration curve against concentration voltage is calibrated using multi-point calibrations; an instrument gain should not be adapted to each individual point. The calibration of the pressure is performed at two points, one under vacuum and another under ambient pressure.

Precision check

A level 2 calibration is a precision check. This implies that the instrument has a known concentration or zero air, and no provision is made for the measurement of the instrument concentration. An accuracy check can be done manually or automatically.

3.3.3 Measurements of CO

Measurement theory

A non-diffuse infrared (NDIR) measurement of CO in ambient air is done with the Serinus 30 carbon monoxide analyzer. Carbon Monoxide is incomplete combustion of hydrocarbon fuels that is colorless, odious and tasteless gas. When inhaled, it connects to hemoglobin, myoglobin and mitochondrial cytochrome oxidase which reduce oxygen storage, transport and breathing in these organelles. It is a toxic gas to the human body. Serinus 30 allows continuous analysis based on the capacity of CO to absorb infrared radiation. The maximum CO absorption range of 4.67μm is the range of

the optical filter. The maximum is 4.67μm. As is shown in the Beer-Lambert Law, the power of the received signal is proportional to the sample quantity of CO. The signal detector has a band-pass filter to guarantee that only light is detected near 4.7 microns of the wavelength. This scheme combines a gas filter correlation wheel in the light route. This wheel has three components to improve the precision of measurement, i.e. CO window, N_2, and mask window. A CO saturation functioning as a beam is included in the CO window. The N_2 window is not IR-absorbing at 4.7 microns and is applied for ordinary CO-measuring. The mask blocks the light source completely and determines, in relation to each other and the background, the background signals and other signals. Figure 3.6 shows a streamlined design of the carbon monoxide analyzer. A Teflon filter of diameter 47 mm and a filter size of 5 microns is the particulate filter. This filter excludes all particles greater than 5 μm which reduce sample measurement interferences. The following major sub-components are used to measure: particle filter; CO-CO_2 converter; optical cell; primary controller PCB; pressure sensor.

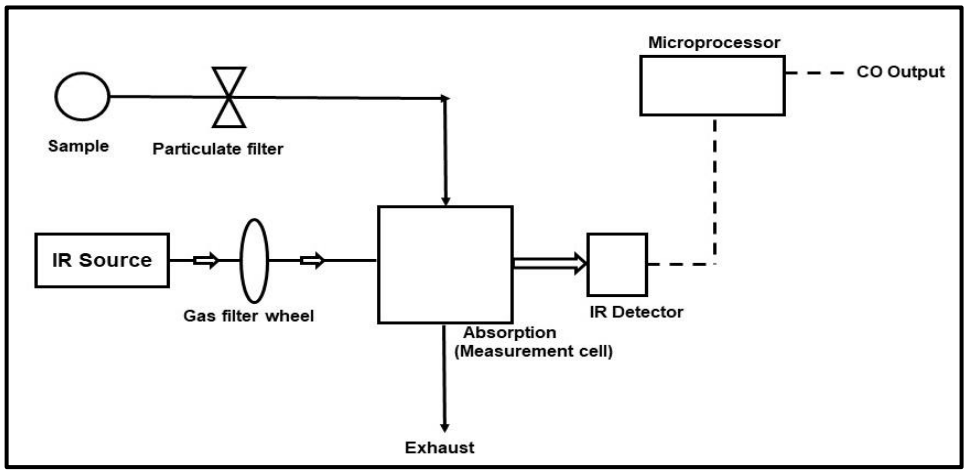


Figure 3.6: A simplified layout of the CO analyzer showing the working principle of the analyzer

Instrument description

The Serinus series of analyzers use the sophisticated Kalman digital filter. For the type of signals and noises in environmental air analyzers, this filter offers the highest potential compromise between reaction times and noise decrease. It has a sensitivity of 0.05ppm in the range of 0-200ppm. For gas temperature and pressure changes, the concentration of CO is automatically corrected and is based on 0 ° C, 20 ° C or 25 ° C at 1 atmosphere. The carbon monoxide analyzer comprises the five main assemblies:

- Sample and exhaust gas transfer pneumatics.

- Carbon monoxide (optical cell) measurement devices and other pertinent parameters.
- The control system comprising all the pneumatic and sensor control circuit panels.
- Power supply for all device processors that provides energy.
- Data access communications module.

3.3.4 Methane hydrocarbons and Total non-methane hydrocarbons

Measurement theory

The Synspec Alpha 115 analyzer uses true gas chromatographic separation to measure methane and total non-methane hydrocarbons in ambient air. Methane and total non-methane hydrocarbons are separated by column in a compact oven. FID is the detector used in the instrument. A specific layered packed column moves through the gas sample. Methane (CH4) is injected first in the detector. A second later, the column is "back-flushed" and the detector is filled with all other hydrocarbons. The FID generates two peaks: one for methane and a peak of TNMHC. Due to its high sensitivity and broad linearity, FID is the perfect detector for measuring hydrocarbons. Oxygen is burned with hydrogen in a flame. When other molecules pass through a hydrogen flame, they burn and boost the ion level. The flame-generated ions are drawn to the positive side of an electrical field in the detector cell. This change is assessed and has a direct correlation to the concentration. The electrical current within the detector cell changes.

Figure 3.7 shows a streamlined design of the analyzer for MHC and TNMHC. A compact oven (to heat up a column), a column (for separating methane from non-methanal complete hydrocarbons) and the FID (Flame ionization detector) are measuring principal components. This is done by a loop (to retrieve a sample).

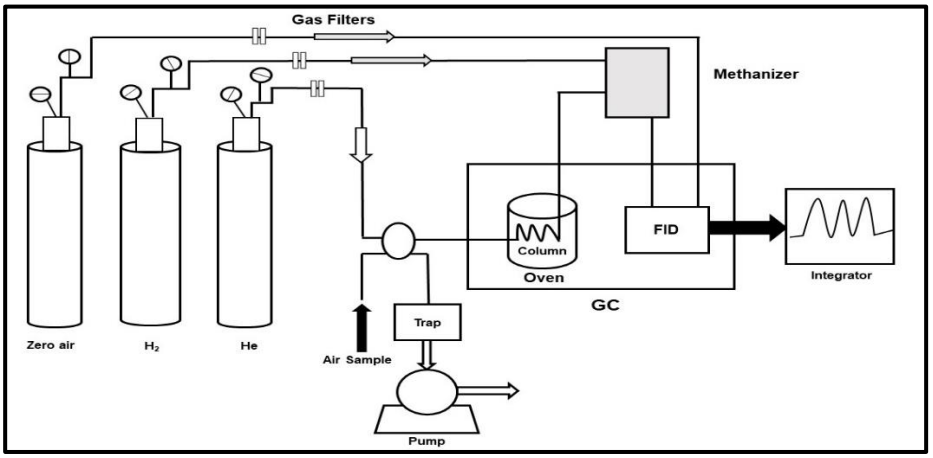


Figure 3.7: A simplified layout of the MHC and TNMHCs analyzer showing the

working principle of the analyzer

Instrument description

Alpha 115 can detect methane from 0.1 ppm to 1000 ppm and total non-methane hydrocarbons from 0.05 ppm to 100 ppm with its packed column and FID detector. The standard time for analyses is 3 min (180 sec) with repeatability and linearity <1% of full scale.

Calibration

Gas stream at 25 ml ambient pressure is required for internal calibration switch for zero calibration and span gas. It determines the conversion factor for (automatic) calculation of concentrations from area values. It is also used for checking the correct functioning of the instrument. Validation is normally used for checking the validity of an existing calibration. However, it can also be used to calibrate, if i.e. span gas deviation is higher than 10 %. Span gas at various known concentrations is involved in multipoint calibration and output of the instrument is recorded. A well-calibrated Alpha 115 will most certainly read the precision of +/- 1%.

3.3.5 Meteorological Parameter Measuring Sensor

Relative humidity, temperature, solar radiation, wind speed and direction along with rainfall data are various meteorological parameters that were recorded at the rooftop (12 m height) from the ground level using Automatic Weather Station (tower-based measurements). The measurements of meteorological parameters have been carried out well equipped with online sophisticated sensors including thermo-hygrometer, thermometer, electronic rain gauge, solar radiation pyranometer and wind speed and wind direction sensors.

Relative humidity and temperature sensors provide a compact, economical & reliable system for monitoring ambient air. The humidity sensor is based upon the capacitance change of a polymer solid-state thin-film capacitor. In-built signal conditioning electronics provides an amplified and linear output proportional to the relative humidity. A dielectric polymer absorbs water molecules from the air by a thin metal electrode, resulting in a moisture shift. The sensor assembly including the sensor for temperature is a PT1000 Platinum Resistance element having a linear response over the

0 to 60 degree C operating temperature range. Envirotech supplies the humidity/temperature sensor complete with a Radiation shield and 20 m cable for interface to data loggers. The white powder-coated radiation shield is made from seasoned wood. The concentric plates allow free movement of ambient air and provide an effective shield minimizing the effect of solar and terrestrial radiation on the temperature measurement. The range lies between 0-60 °C within accuracy and resolution of +/- 0.2 °C and 0.1 °C, respectively. For relative humidity sensor range lies between 0 -100 %, with an accuracy and resolution of +/- 3 % and 0.1%., respectively.

In agricultural, meteorological and solar power research, the LI-200SA Pyranometer is intended to measure global solar radiation in fields. LI-200SA has a photovoltaic silicon detector installed in a completely cosine-corrected miniature head. Current output, that is directly proportional to solar radiation, is calibrated in a unit of watts per square meter against an Eppley Precision Spectral Pyranometer (PSP). In most natural daylight circumstances the error is < 5%.

Wind Speed Sensor (Model WS10) is an accurate and durable three-cup anemometer with non-contact optical sensing, optimized for the study of wind speed for environmental applications where low speed or calm conditions are especially important. For minimum friction and mechanical load, the sensor uses a non-contact photo chopper based system for sensing rotation of the anemometer shaft. The sensor output is a pulse train with a frequency proportional to the rate of the shaft rotation or horizontal wind speed. The TTL level output from the sensor can be directly connected to a data-logger via the 20m cable provided with the sensor. The working range lies between 0.5 to 100 km/hr with an accuracy of +/- 0.5 m/sec.

Wind vane sensor (Envirotech WD10) provides azimuth data for use in micrometeorological studies. The sensor is especially useful when the measurement of wind direction is critical in near calm conditions. The WD10 sensor is based on the principle of potentiometer type with SS shaft with low starting threshold and a high damping ratio. The range of the sensor lies between 0-360 degrees, resolution of 1 degree and accuracy of + / - 3 degrees.

Precise rain/precipitation measurement is one of the fundamental meteorological criteria. The Envirotech RF10 rainfall sensor incorporates a tipping bucket design

providing an accurate measurement of rainfall with minimum operator effort or maintenance. The system is economical and field-proven with several hundred units in operation all over the country. Underneath a collecting funnel is a dual-chambered tipping bucket mount. If an accurate amount of precipitation is gathered on one side of the bucket gravity tips it over, it is indicative that the reed interrupter sensor is temporarily shut down. The rain sample gathered will be discharged via the system base. Use of inert materials such as fiber-glass and stainless steel allow prolonged trouble-free field use. It measures with an accuracy and resolution of 1mm and 0.5 mm of rainfall, respectively.

3.4 Quality Control and Quality Check of data

Real-time online raw data get stored in WinAQMS server. A schematic of data management and WinAQMS data acquisition and control have been shown in Figure 3.8. Before perforating any data analysis it is very important to check the quality of data so it will truly represent the existing conditions. Before using the data for further analysis rigorous quality checks have been performed on raw data. The Continuous Data Analyzers (by Ecotech) was used in the system are based on the United States Environmental Protection Agency (USEPA) equivalent referenced methods and are linked digitally to Data Acquisition System (DAS) having features like automated information validation software with records capture percentage. The data is recorded continuously and reported for every 5 minutes for all the analyzers during the period of measurements.

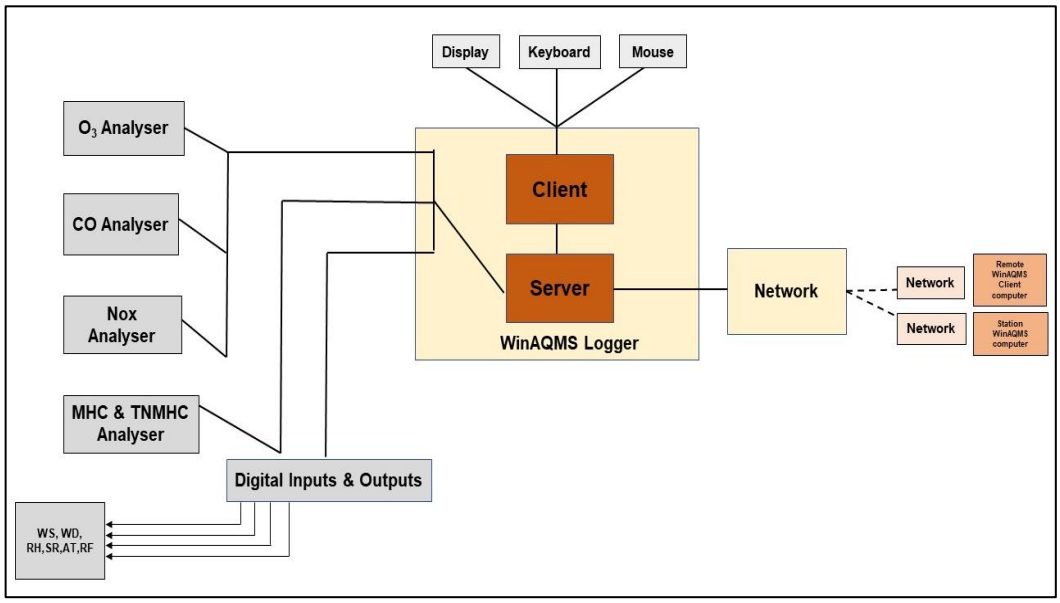


Figure 3.8: A diagrammatic representation of data management and WinAQMS data acquisition control

After collecting the raw data, a quality check has been done on the data to identify for the outlier data points. To ensure the quality of data, guidelines for monitoring and calibration, repair of instruments and evaluation of ambient air quality monitoring station was performed on regular basis as per protocol is given in material and methodology. This correct accounting is taken of the standard instrumental error and the standard deviation from the mean in a month for a specific hour. Average daily concentrations were calculated only if at least 18 of each day's hourly concentration values were available. This obtained data is used as a baseline/input to the programs to get the data in the required form of analysis.

3.4.1 Calibration

To ensure the good health of instruments in terms of their stability, functioning and analyzer response to pollutants concentration, calibration has been carried at appropriate time intervals (daily, weekly, 6 monthly and yearly) in calibration protocol for ozone, CO, NO_x, MHC, and TNMHC. To establish the relationship between analyzer response and pollutants concentration over the analyzer's full-scale range, the multiple calibrations have been performed. In order to confirm for an indication of analyzer stability and function zero and span checks have been performed. In general calibration process consists of the establishment of a reliable and stable calibration source, provide a successful link between the analyzer and the calibration source and calibrate the analyzer against the calibration source. Figure 3.9 shows the overall overview of the calibration scheme for CO, NOx, and NMHC. Ozone calibration was performed against a well-known ozone source that was produced and tested with a UV photometer.

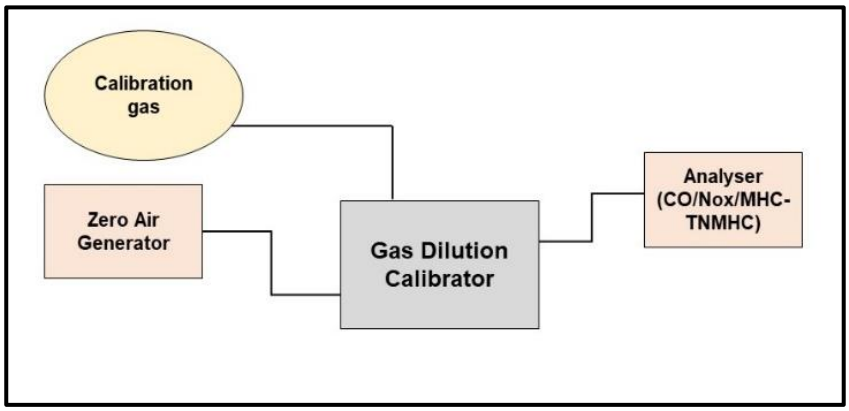


Figure 3.9: General outline of calibration system used for CO, NO_x, MHC, and TNMHCs

3.5 Satellite Data: Moderate Resolution Imaging Spectroradiometer (MODIS)

Pre, during and the post biomass burning periods for wheat and rice crop residues have been outlined over the study period using Terra-MODIS (Moderate Resolution Imaging Spectroradiometer) true color composites. Wheat crop residue burning (WCRB), designated as; Pre- Wheat crop residue burning (15th March to 31st March), During- Wheat crop residue burning (20th April to 31st May) and Post- Wheat crop residue burning (15th June to 30th June). Rice crop residue burning (RCRB), designated as; Pre- Rice crop residue burning (1st September to 15th September), During- Rice crop residue burning (15th October to 20th November) and Post- Rice crop residue burning (1st December to 15th December). Terra-MODIS (Moderate Resolution Imaging Spectroradiometer) satellite observations clearly indicate a large number of fire spots (red dots) during wheat and rice crop residue burning period and related emissions in the form of smoke over the IGP region shown in Appendix-II.

3.6 Back Air Trajectory Simulations using HYSPLIT and Cluster analysis

The 5-days backward trajectory at 500 m level over Patiala for four different seasons: winter (December-February), summer (March-May), monsoon (June-August), and post-monsoon (September-November) have been plotted in order to show the more real air parcel movement and their transport pathways, as shown in Figure 3.10. It was done using trajectories statistics software, Traj-Stat (version 1.2.2.6). The substantial sources of emissions at this site are local vehicular exhaust and agricultural crop residue burning in the surrounding regions. In summer season, long-range transport of air masses from desert regions of Iran, Afghanistan, Pakistan, and Arabian Peninsula contributed to aged pollutants and regional emissions of wheat crop residue burning (WCRB) event contributed to the increased emission levels of precursor gases (like NO_x and CO) for ozone production, under favorable meteorological conditions. In the monsoon season, trajectories indicate the transport of cleaner air from the Indian Ocean and Arabian Sea due to SW winds. The levels of pollutants show a significant impact of wash-out due to dilution with the transport of cleaner air, frequent rainfalls, decreased solar radiation and negligible contribution from the agricultural crop residue burning events during monsoon compared to other seasons. Post-monsoon, trajectories show the

mixed air masses (continental and marine) and the transition mainly from SW (during monsoon) to NW. High levels of emissions were reported due to regional rice crop residue burning (RCRB) event, festival period and the advent of low temperature and calm wind weather conditions. In the winter season, regional air masses reported high levels of CO and NO_x and low levels of ozone, due to decreased solar radiation intensity and sunshine hours, which inhibits the photochemical processes. The low boundary layer and calm meteorological conditions, further limits the dispersion of regional emissions of biomass burning and vehicular exhaust.[9]

These groups of trajectories were divided into different groups (clusters) depending on the velocity and direction of transportation. Their length and curvature ultimately produce clusters representing the main directions of air mass that offer a better understanding of the source areas (Li et al. 2012).[10] Out of these four clusters, the major percentage contribution was from cluster 1 (56 %) followed by cluster 3 (17 %) shown in Figure 3.11. Hence, the major influx of gaseous and particulate pollutants originated from regional continental landmasses to reach the receptor site and was not influenced significantly by long-range transport.

Further, potential source contribution function (PSCF) and concentration weighted trajectory (CWT) analysis was performed to highlight the source potentials of the studied geographical region and also to allocate their discrete contribution to the measured concentrations at the receptor site.[11-13] It was done using trajectories statistics software, TrajStat (version 1.2.2.6).

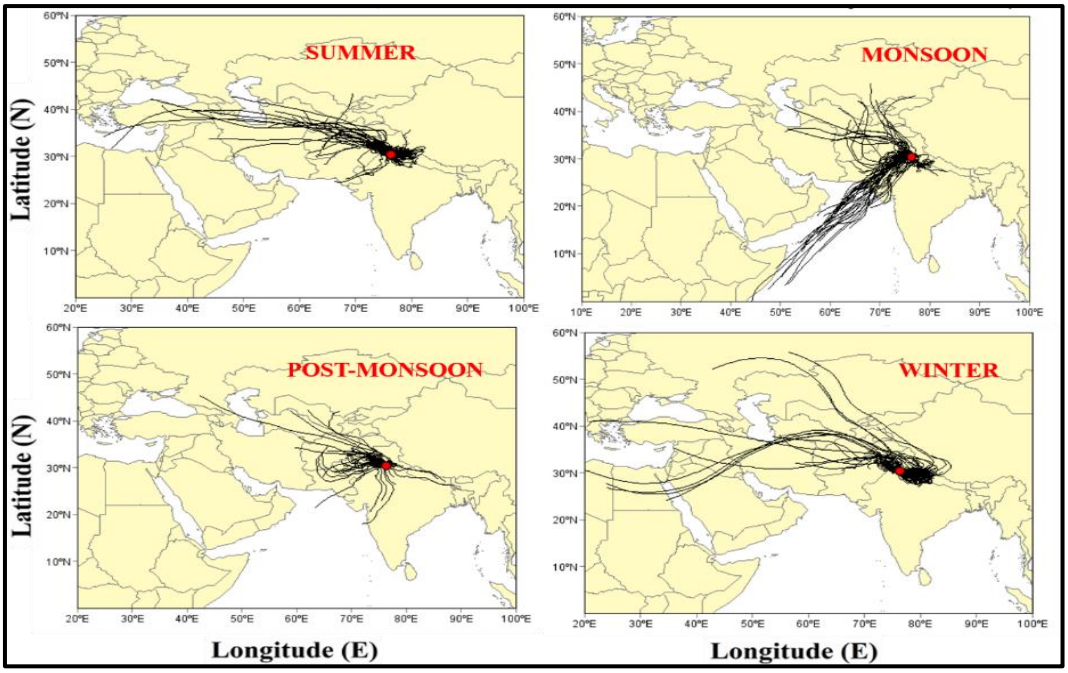


Figure 3.10: Backward trajectories at 500m level over Patiala in different seasons (Summer, Monsoon, Post-monsoon, and Winter) during the study period (2013-2016)

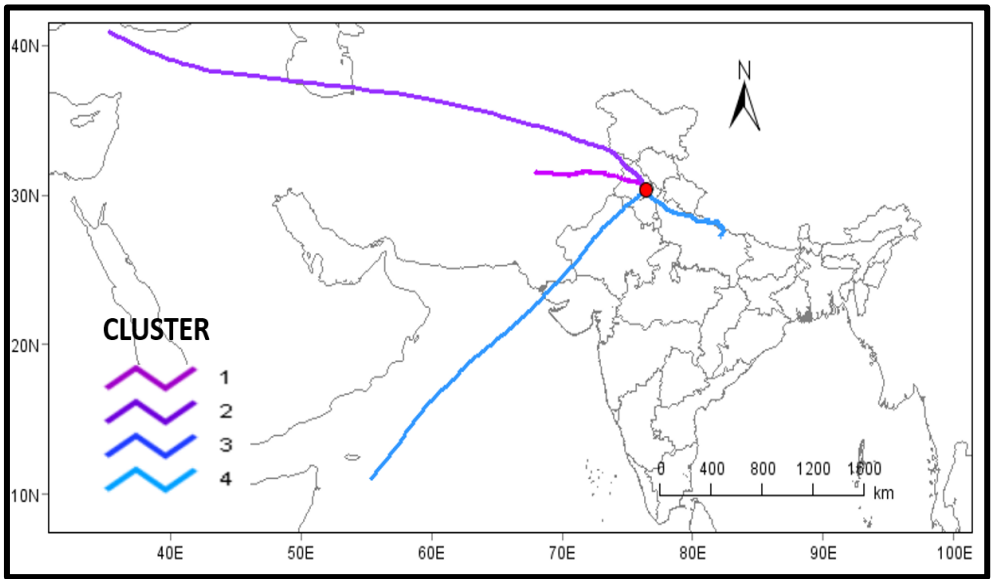


Figure 3.11: Clustered trajectory analysis of air masses during the study period (2013-2016)

3.7 Statistical modeling approach

Statistical approaches with a range of linear and non-linear regression techniques have gained enhanced attention in the last ten years for ozone prediction in distinct geographical areas. These models provide greater precision and greater calculation efficiency forecasts. Air Quality prediction can be carried out using various traditional

and soft computing techniques. The prediction algorithms used in this research work include fourteen techniques (random forest, tree bag, neural network, linear regression, linear regression with backward selection, independent component regression, least angle regression, generalized linear model, k- nearest neighbors, kernel support vector machine, partial least square, projection pursuit regression, bayesian generalized linear model, least angle regression) from different families. Table 3.1 contains the detailed information of the family of models along with its tuning parameters.

Table 3.1: Prediction models with their family to which a method belongs and its tuning parameters.

Model	Method	Tuning parameters and value(s)	Family
Random Forest	rf	No. of trees=500; No. of variable tried at each split =2	Classification, Regression
Bagged CART	tree bag	None	Classification, Regression
Neural network	nnet	size= 10, linout= TRUE MaxNWts= 10000, trace= FALSE	Classification, Regression
k-Nearest neighbor	kknn	k=9	Classification, Regression
Support vector machines	ksvm	kernel= "rbfdot" prob.model= TRUE	Classification, Regression
Partial least squares	pls	ncomp= 3	Classification, Regression
Linear Regression	lm	intercept=89.65	Regression
Linear Regression With backward selections	leapbakwacd	nvmax=4, leaps	Regression

Generalized linear model	glm	None	Classification, Regression
Independent Component regression	icr	ncomp=3,fast	Regression
Bayesian Generalized Linear model	bayesglm	None	Classification, Regression
Relaxed Lasso	relaxo	lambda=6386708 and phi=0.1	Regression
Projection pursuit regression	ppr	interms=3	Regression
Least angle	lars	fraction=1	Regression

3.8 References

1. Kumar, V.; Sarkar, C.; Sinha, V. *Journal of Geophysical Research: Atmospheres* **2016**, *121*, 3619-3633.

2. Kolehmainen, M.; Martikainen, H.; Ruuskanen, J. *Atmospheric Environment* **2001**, *35*, 815-825.

3. Sansuddin, N.; Ramli, N. A.; Yahaya, A. S.; Yusof, N. F. F. M.; Ghazali, N. A.; Al Madhoun, W. A. *Environmental Monitoring and Assessment* **2011**, *180*, 573-588.

4. Jain, N.; Bhatia, A.; Pathak, H. *Aerosol and Air Quality Research* **2014**, *14*, 422-430.

5. Gupta, P. K.; Sahai, S.; Singh, N.; Dixit, C.; Singh, D.; Sharma, C.; Tiwari, M.; Gupta, R. K.; Garg, S. *Current science* **2004**, 1713-1717.

6. Punia, M.; Nautiyal, V. P.; Kant, Y. *Current Science (00113891)* **2008**, *94*.

7. Kleinman, L. I. *Journal of Geophysical Research: Atmospheres* **1994**, *99*, 16831-16838.

8. Yadav, R.; Beig, G.; Jaaffrey, S. *Atmospheric Environment* **2014**, *85*, 147-151.

9. Nair, V. S.; Moorthy, K. K.; Alappattu, D. P.; Kunhikrishnan, P.; George, S.; Nair, P. R.; Babu, S. S.; Abish, B.; Satheesh, S.; Tripathi, S. N. *Journal of Geophysical Research: Atmospheres* **2007**, *112*.

10. Li, Y.; Lau, A. H.; Fung, J. H.; Zheng, J.; Zhong, L.; Louie, P. *Journal of Geophysical Research: Atmospheres* **2012**, *117*.

11. Hopke, P. K. *Journal of Chemometrics: A Journal of the Chemometrics Society* **2003**, *17*, 255-265.

12. Wang, Y.; Zhang, X.; Draxler, R. R. *Environmental Modelling and Software* **2009**, *24*, 938-939.

13. Cheng, I.; Zhang, L.; Blanchard, P.; Dalziel, J.; Tordon, R. *Atmospheric Chemistry and Physics* **2013**, *13*, 6031-6048.

RESULTS & DISCUSSION

CHAPTER -4

The severity of the impacts of air pollution depends upon two variable factors i.e., the ambient concentration of pollutants and its exposure time. The measurement of primary and secondary gaseous pollutants helps to study the changes in atmospheric chemical composition and to understand the ambient air quality. The data used was collected from Continuous Ambient Air Quality Monitoring Station having analyzers (by Ecotech), based on the United States Environmental Protection Agency (USEPA) equivalent referenced methods and reported for every 5 minutes for all the analyzers during the period of measurements. Average daily concentrations were calculated only if at least 18 of each day's hourly concentration values were available.

4.1 Day to day variation in trace gases (surface O_3, CO, NO_x, MHC, TNMHC) and meteorological variables

This study presents four year period of 1st January 2013 to 31st December 2016 of continuous measurement of O_3, CO, NO_x, MHC, TNMHCs and meteorological variables (AT, SR, RH, WS, and WD). The data coverage for the entire study period was very high (about 95%) and hence statistically robust. Figures (4.1 to 4.5) show the daily mean, maxima and minimal levels of gaseous pollutants (O_3, CO, NO_x, MHC, and TNMHCs). The daily mean variation of meteorological variables (RH, SR, AT, T_{max} and WS) are shown in Figures (4.6, 4.7) for the study period.

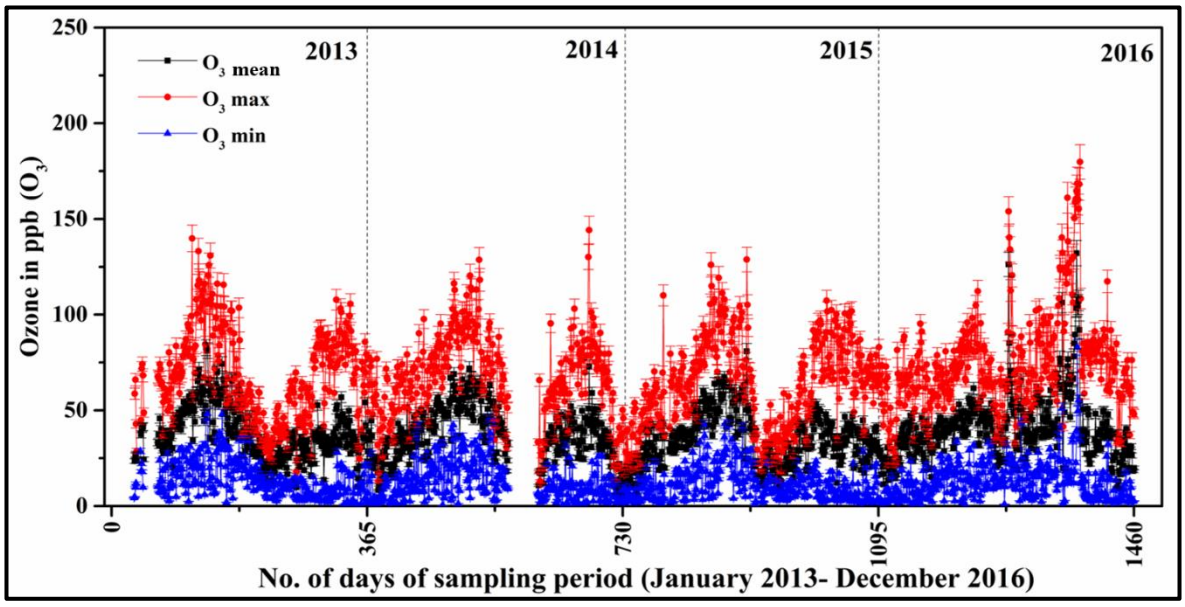


Figure 4.1: Daily maximum and daily minimum levels (24 hours averaged) of ozone observed at Patiala during the study period (2013-2016)

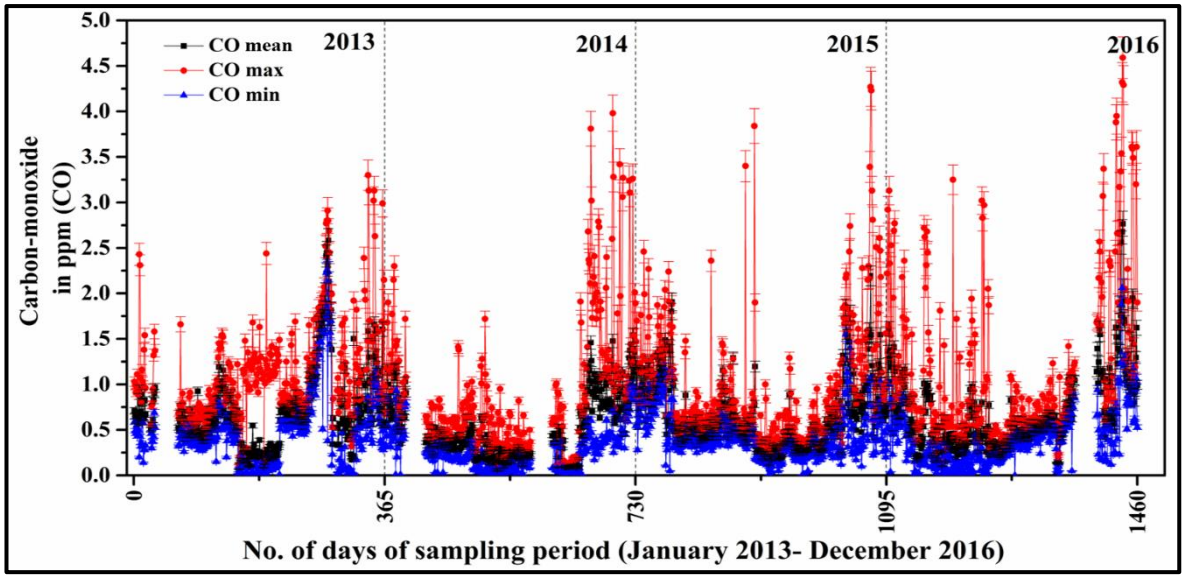


Figure 4.2: Daily maximum and daily minimum levels (24 hours averaged) of carbon-monoxide during the study period (2013- 2016)

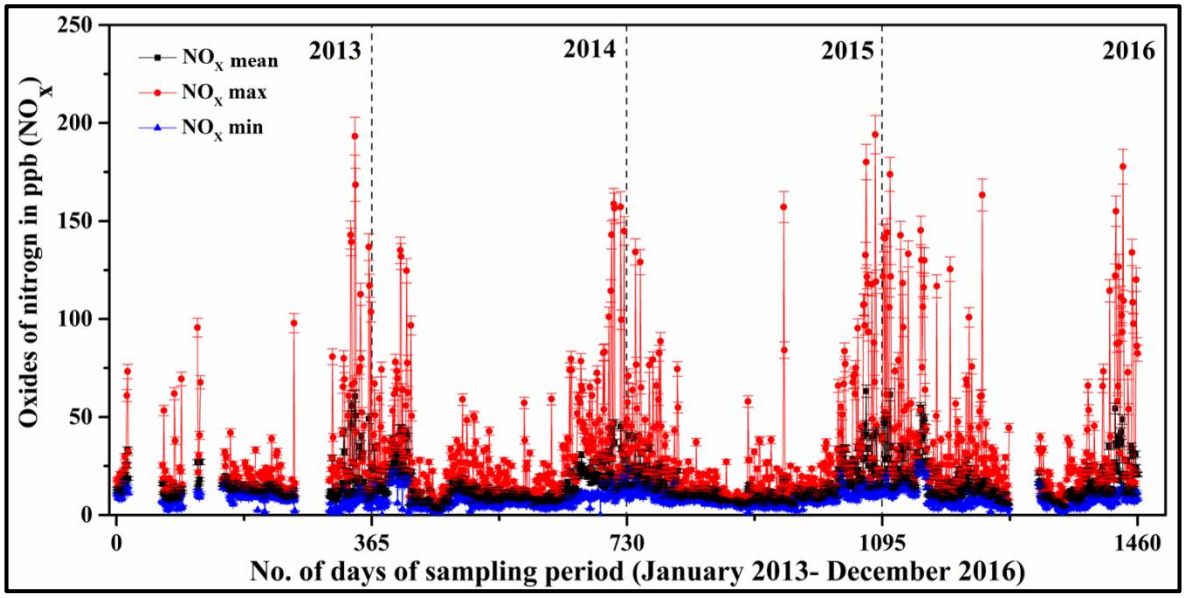


Figure 4.3: Daily maximum and daily minimum levels (24 hours averaged) of Oxides of nitrogen during the study period (2013- 2016)

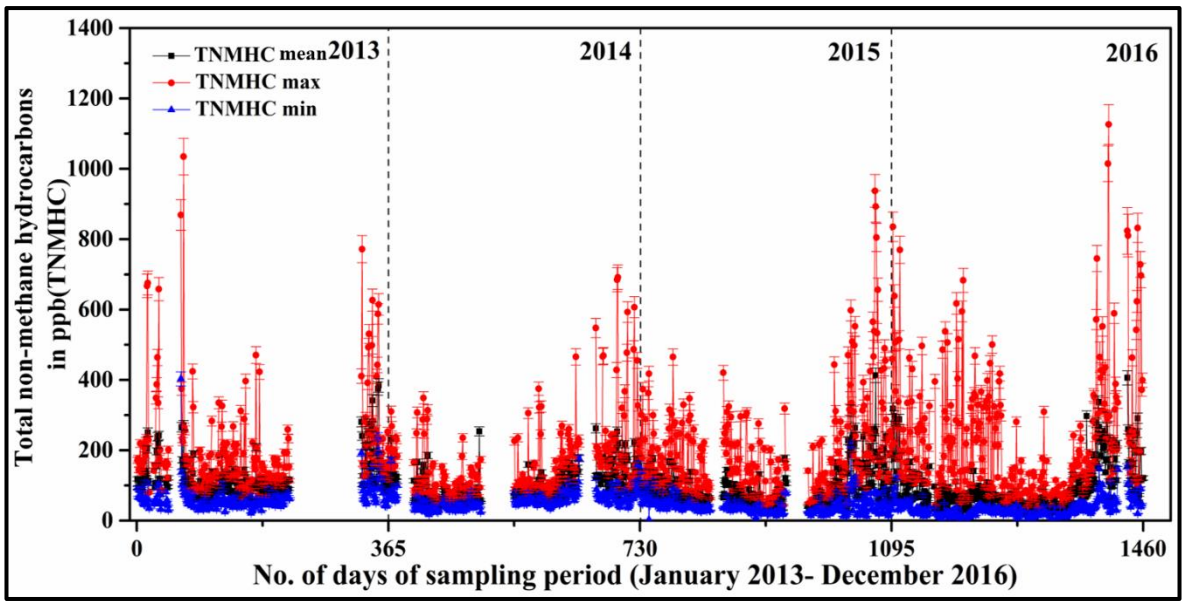


Figure 4.4: Daily maximum and daily minimum levels (24 hours averaged) of total non- methane hydrocarbon during the study period (2013-2016)

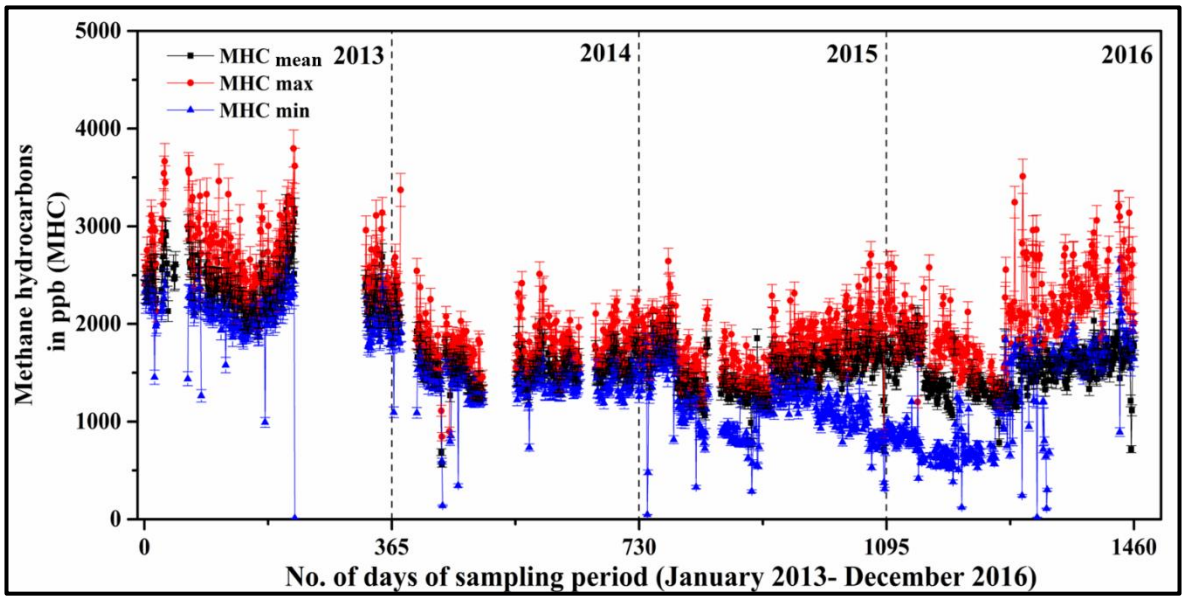


Figure 4.5: Daily maximum and daily minimum levels (24-hours averaged) of methane hydrocarbons during the study period (2013- 2016)

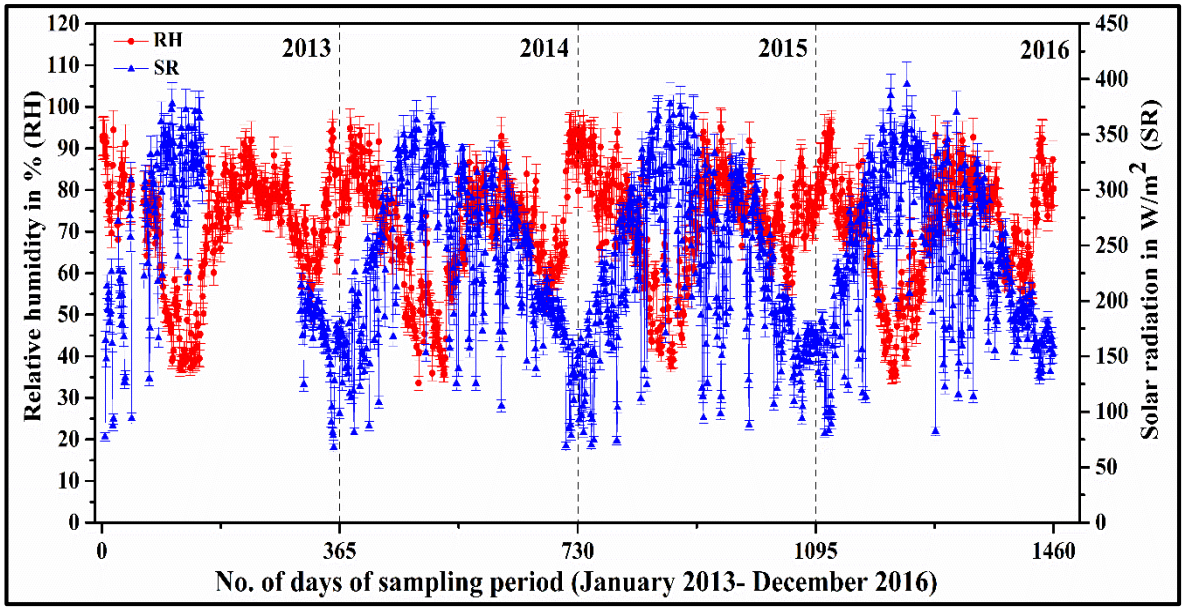


Figure 4.6: Levels (24 hours averaged) of relative humidity and solar radiation during the study period (2013- 2016)

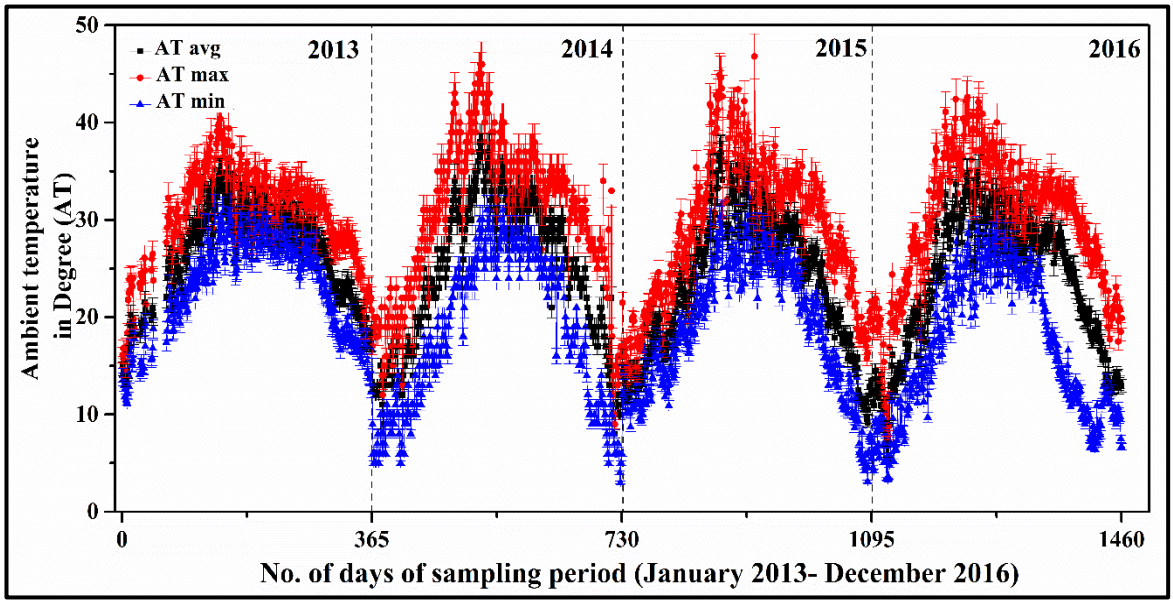


Figure 4.7: Daily maximum and daily minimum levels (24 hours averaged) of ambient temperature during the study period (2013 – 2016)

The daily concentrations of gaseous pollutants like O_3, CO, NO_x, MHC, and TNMHCs ranged from 5 to 83 ppb, 0.01 to 2.6 ppm, 2.3 to 113.5 ppb, 562 to 3166 ppb and 27 to 413 ppb, respectively. The meteorological variables RH, SR, AT and WS varied from 33.5 to 95.0 %, 68.0 to 378 W m^{-2}, 8 to 38 °C and 0.05 to 7.85 m sec^{-1}, respectively.

Anthropogenic emissions from sources such as vehicles, domestic activities, road-dust, construction activity, and industrial processes are released throughout the year, but the episodic events of post-harvest crop residue burning of wheat and rice occur during April–May (summer) and October–November (post-monsoon), respectively. The intermediate period from December to February (winter season) experiences fog formation, a manifestation of shallower planetary boundary layer and emissions from biofuel burning and fossil-fuel combustion sources and moisture from western disturbances.

4.2 Diurnal variations in surface O_3, CO, NO_x, MHC, TNMHC during different seasons

Diurnal cyclic variations of the gaseous mixture of ambient air are important to get information about local pollution conditions. The hourly averaged values recorded for 5 minutes further averaged for each month and for a particular season (± standard deviation) to get the diurnal variations, reported using Envirotech online equipment's discussed in details in section 2.4. Diurnal seasonal variations of ozone and its precursor gases are shown in Figures (4.8 to 4.12) for the study period (January 2013 to December 2016).

In diurnal ozone profile in Figure 4.8, a steep minimum is observed during the early morning around 0600- 0800 hours in all seasons of the year. After 0800 hour, a rise in ozone concentration is observed and hit maxima around in the range of 1500-1600 hours and starts decreasing after late afternoon or evening. However, ozone concentration showed significant strong seasonal variability and a similar diurnal pattern with varied amplitude, from season to season. A pronounced and sharper increment in daytime maximum ozone is observed during the post-monsoon season; whereas, the daytime broader area of ozone is observed in the summer season. During the daytime, the photochemical reaction of precursor gases, in the presence of favourable meteorological conditions and vertical mixing from ozone rich upper layer contribute to surface ozone levels.[1] During early morning and night hour, low boundary layer height reduces the vertical mixing process and the higher concentration of NO and NO_2 leads to rapid titration of ozone.

In the summer season, the monthly daytime maxima of O_3 were measured to vary in the range of (62-74 ppb) at 1500-1700 hours. In this season, the lowest values of

around (20-22 ppb) were measured in the morning hours (0700-0800 hours). In the monsoon season, the diurnal variation shows less variability with maxima of (50-82 ppb) at 1500-1700 hours and minima of (17-22 ppb) in the early morning hours. In the post-monsoon season, the daytime maxima of O_3 in the range of (69-91 ppb) were measured during 1500- 1700 hours, while low values (11-17 ppb) in the early morning hours. The strong diurnal variation with daytime peak values of (44-59 ppb) and the lowest mixing ratios of (10- 11 ppb) in the early morning hours were observed during the winter season. However, the distribution of O_3 in the monsoon season shows weaker diurnal dependency.

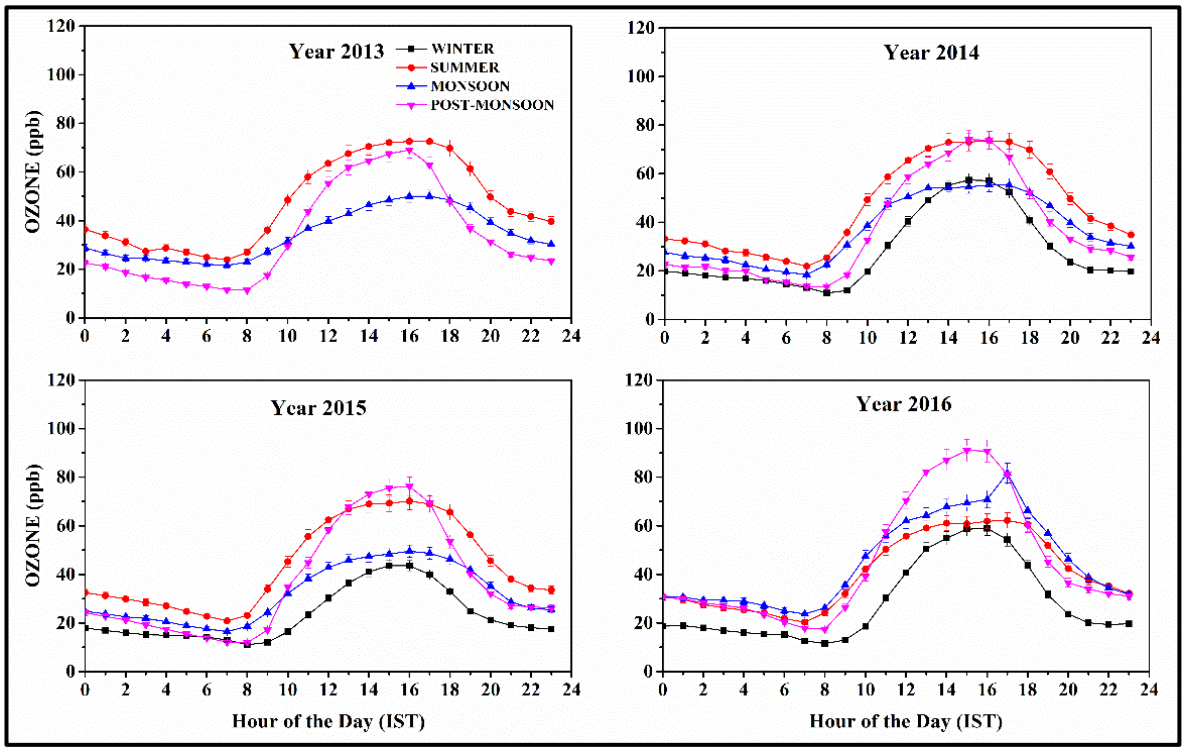


Figure 4.8: Diurnal variations in levels of ozone in four seasons (winter, summer, monsoon, and post-monsoon) at Patiala during the study period (2013-2016)

The comparison of levels of ozone in the ambient air of the Patiala site with that of some other locations in India is reported in Table 4.1. Moderate to high values of the gaseous pollutants were reported for other rural, semi-urban, urban, coastal and mountain sites in India.[1-7]

Table 4.1: Comparison of ozone levels of ambient air at Patiala site with those at some other locations in India

Loation	Latitude (deg)	Longitude (deg)	Site type	Period of study (mm/year)	Ozone (ppb)	Reference
Patiala	30.35 ºN	76.37 ºE	Semi-Urban	January-2013 to December-2016	37.52±10.58	(This study)
Agra	27.18 ºN	78.02 ºE	Urban	January-2002 to December-2002	20.0± *	[8]
Ahmedabad	23.0 ºN	72.60 ºE	Urban	January-2002 to December-2002	26.27± 10	[9]
Pune	18.53 ºN	73.86 ºE	Urban	June-2003 to May-2004	30.9±14	[10]
Anantapur	14.68 ºN	77.60 ºE	Rural	January-2010 to December-2010	40.7±3.7	[11]
Kanpur	26.46 ºN	80.33 ºE	Urban	June-2009 to May-2013	27.9±17.8	[5]
Jabalpur	23.17 ºN	79.95 ºE	Urban	January-2014 to December-2014	34.0± *	[12]
Udaipur	26.58 ºN	73.68 ºE	Urban	April-2011 to March-2012	22.420± *	[7]

*Data not reported

Figures 4.9 to 4.12 show diurnal variations of precursor gases of surface ozone i.e., NO_x, CO, MHC and TNMHCs during different seasons of the study period and peaks were observed in the early morning and late evening hours. On a diurnal scale, precursor gases exhibited two maxima, at around 0700-0900 hours and 1800-2200 hours, corresponding to low boundary layer and prominent anthropogenic emissions during evening hours, when the vehicular activity is maximum. The rise in PBL depth during 1400-1800 hours provides a larger mixing region along with consumption of precursor gases in photochemical production of surface ozone.

Comparison of daily ozone profiles with those of precursor gases (NO_x and CO) showed that maximum O_3 concentration was coincident with the concentration minima of NO_x and CO. Diurnal profile of each season of precursor gases showed a similar pattern with variations in amplitude. The diurnal seasonal average of precursor gases demonstrated significant seasonal deviation with a clear order of winter > post-monsoon > summer > monsoon.

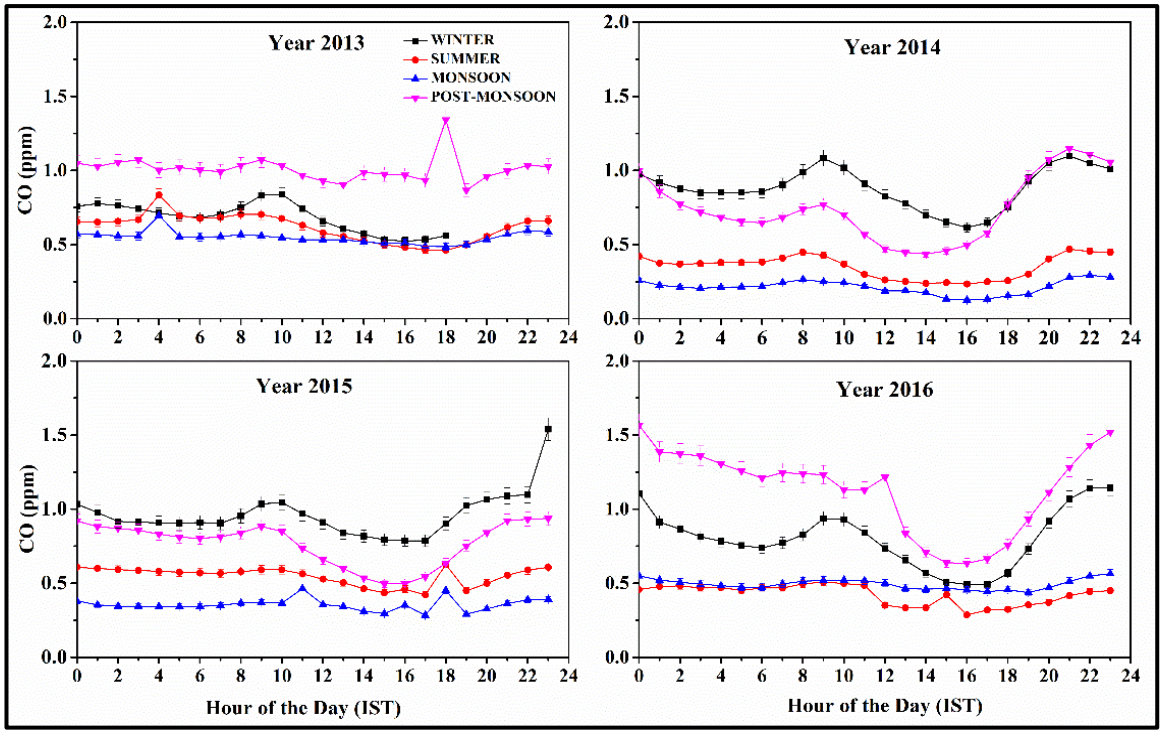


Figure 4.9: Diurnal variations in levels of CO four seasons (winter, summer, monsoon and post-monsoon) at Patiala during the study period (2013-2016)

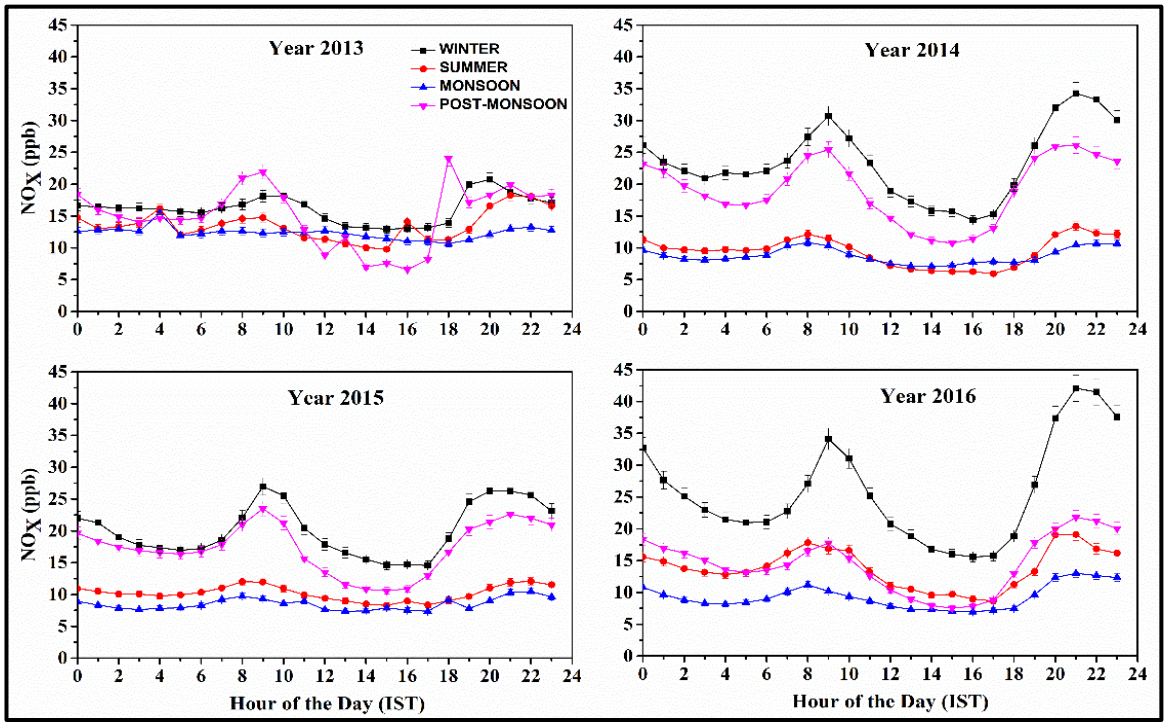


Figure 4.10: Diurnal variations in levels of NO_x in four seasons (winter, summer, monsoon and post-monsoon) at Patiala during the study period (2013-2016)

The mixing ratios of CO, NO_x, MHC, and TNMHCs show largest diurnal peaks in post-monsoon and winter seasons followed by summer and monsoon seasons. Besides, the diurnal profile of each season of precursor gases showed a similar pattern with variations in amplitude.

In the winter and post-monsoon seasons, the morning and evening peak amplitudes of NO_x varied in the ranges of 20.73 - 42.10 ppb and 21.78-26.12 ppb, respectively. The afternoon values of NO_x also show seasonality with higher values 12.93-15.59 ppb in winter and 6.58-10.76 ppb in post-monsoon season, respectively. During these seasons, the peaks were most prominent as compared to other seasons. In the summer season, the morning and evening maxima varied in the range of 12.08 -19.14 ppb and the lowest values of around 5.93-9.78 ppb were measured in the afternoon hours. In the monsoon months, weak diurnal dependence was observed in the levels of NOx, with almost absent morning and evening peaks in the lowest range of 5.93- 9.78 ppb, as compared to other seasons. In winter and post-monsoon seasons, the most prominent morning and evening diurnal peaks of CO varied in the ranges of 0.63-1.14 ppm and 0.94-1.57 ppm, respectively. The summer and monsoon, peak hours levels were in the ranges of 0.47-0.83 ppm and 0.29-1.70 ppm, respectively. The afternoon mixing ratios of CO were 0.49-0.79 ppm, 0.44-0.87 ppm, 0.23-0.46 ppm, and 0.13-0.49 ppm during the winter, post-monsoon, summer and monsoon seasons, respectively.

Similar seasonal pattern with varied amplitude was observed for MHC and TNMHCs levels. In the winter and post-monsoon seasons, the diurnal maxima and minima of MHC and TNMHCs were measured to vary in the range of 1824.72-2822.39 ppb, 1739.94- 2464.89 ppb, 142.64-169.28 ppb, 191.29- 247.20 ppb and 916.63- 2281.71 ppb, 1086.80- 1881.14 ppb, 49.17-84.46 ppb, 60.70- 94.94 ppb, respectively. In the monsoon season, the diurnal variation shows less variability with increased levels of MHC and TNMHC with maxima of 1598.90-3761.54 ppb and 73.69-138.16 ppb and minima of 1270.59 -2158.98 ppb and 35.67-85.06 ppb in the afternoon hours. In the winter and post-monsoon season, emissions from several natural and anthropogenic sources get trapped below the temperature inversion layer, while the pollutants get mixed well in the summer and monsoon seasons, causing dilution of pollutants resulting in minima in diurnal amplitudes.

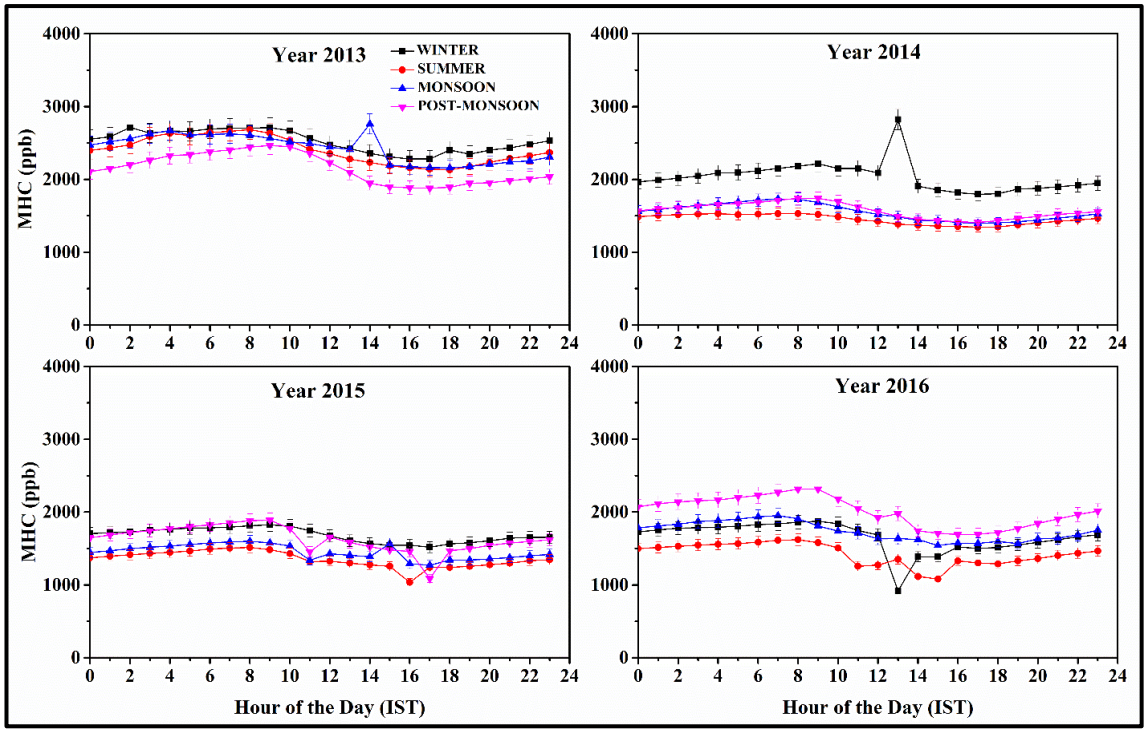


Figure 4.11: Diurnal variations in levels of MHC in four seasons (winter, summer, monsoon and post-monsoon) at Patiala during the study period (2013-2016)

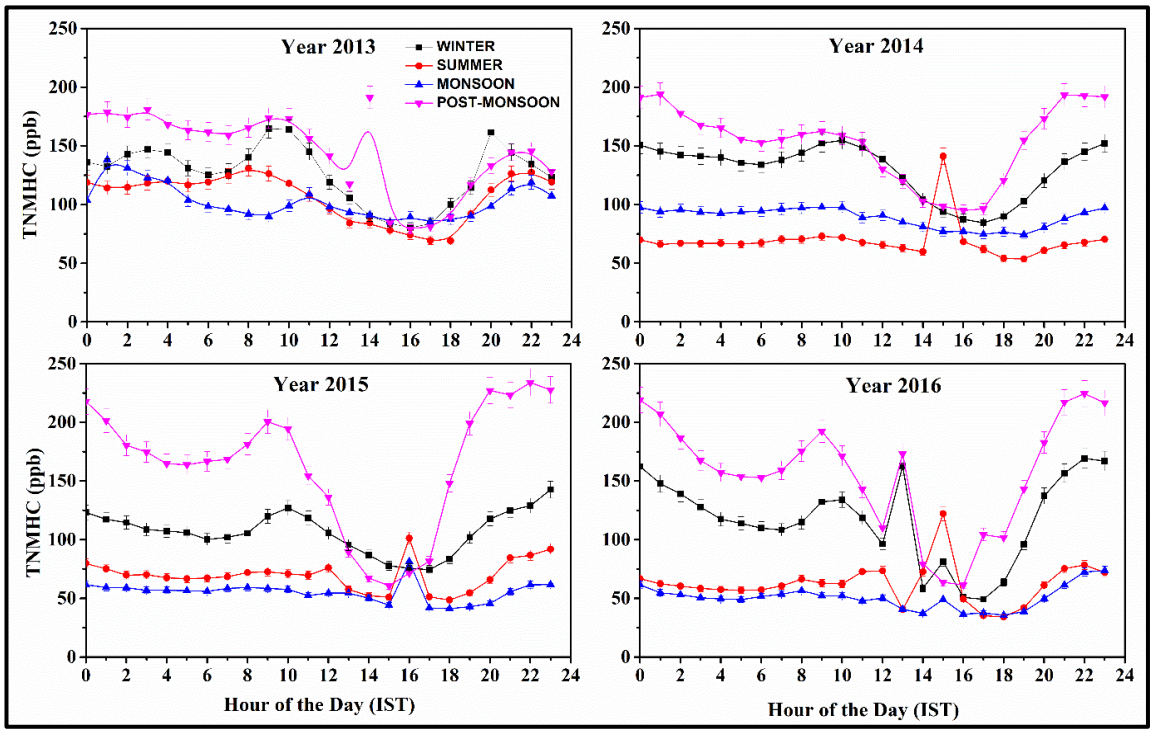


Figure 4.12: Diurnal variations in levels of TNMHC in four seasons (winter, summer, monsoon and post-monsoon) at Patiala during the study period (2013-2016)

4.3 Seasonal variations in surface O_3, CO, NO_x, MHC, TNMHC levels during the study period

Box and Whisker plots in Figures 4.13 (a-e) and Figures (4.14 - 4.18) show the seasonal variation of O_3, CO, and NO_x, MHC and TNMHC levels on an average for the study period (2013-2016) and on yearly basis, respectively. The vertical lines, stretched-out from each end of the box called whiskers, representing 1% (bottom cross) and 99% (top cross) of the data. The vertical box encloses the minimum (bottom line), maximum (top line) and middle line 50% of data. The mean value represented by the cross and horizontal line inside the box stands for the median value. There is very distinct and systematic seasonal variation observed in the levels of trace gases.

Seasonal fluctuations in ozone concentrations can be put in decreasing order as: summer > post-monsoon > monsoon > winter. The observed changes in ozone concentrations are regulated by various interlinked processes of atmospheric chemistry, daily variations in boundary layer height, dynamics and transport of pollutants. Among all seasons, ozone production is intense in summer, because of the increased intensity of solar radiations during the day as well as strong and persistent inversions at nights. In monsoon, low levels of ozone were observed because of wet surface deposition of precursor pollutants by monsoonal rains and limited solar radiation intensity.

NO_x, CO, MHC, and TNMHC emission sources are anthropogenic in nature and followed similar seasonal variability in concentration levels during the study period with higher levels in winter and post monsoon season followed by summer and monsoon season. Minimum ozone is produced in winter due to the low solar intensity and low temperature while there is not much fall in anthropogenic emissions of precursor gases in winter followed by the post-monsoon season. This illustrates that the photochemical production of ozone resulted from the reaction of CO, CH_4, and NMHCs with OH radicals in the presence of NO_x.

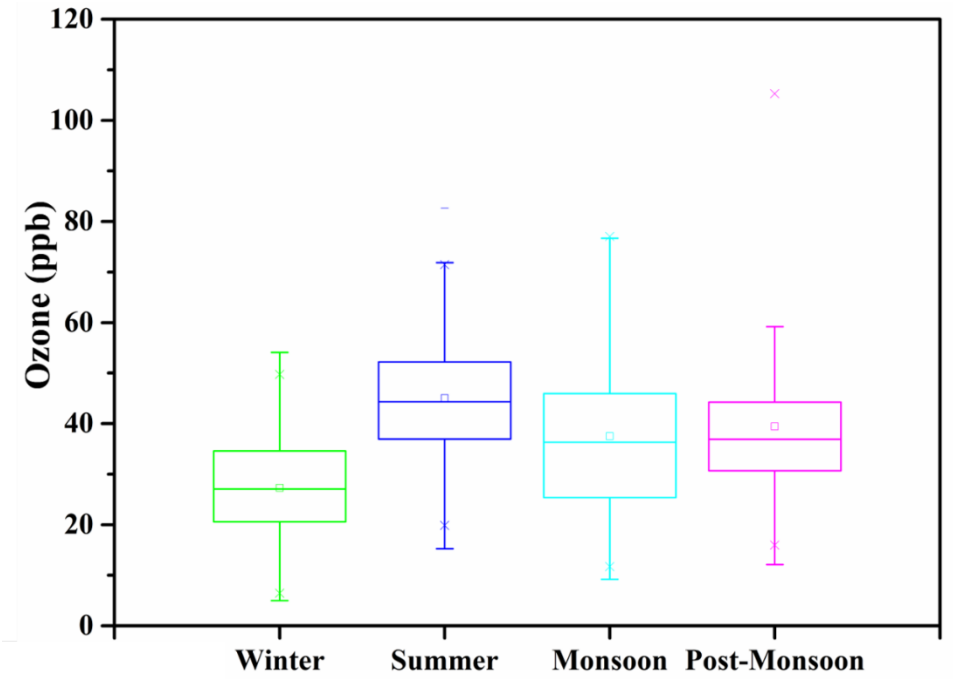


Figure 4.13(a): Box plots of ozone concentrations (seasonal average) at Patiala during the study period (2013–2016)

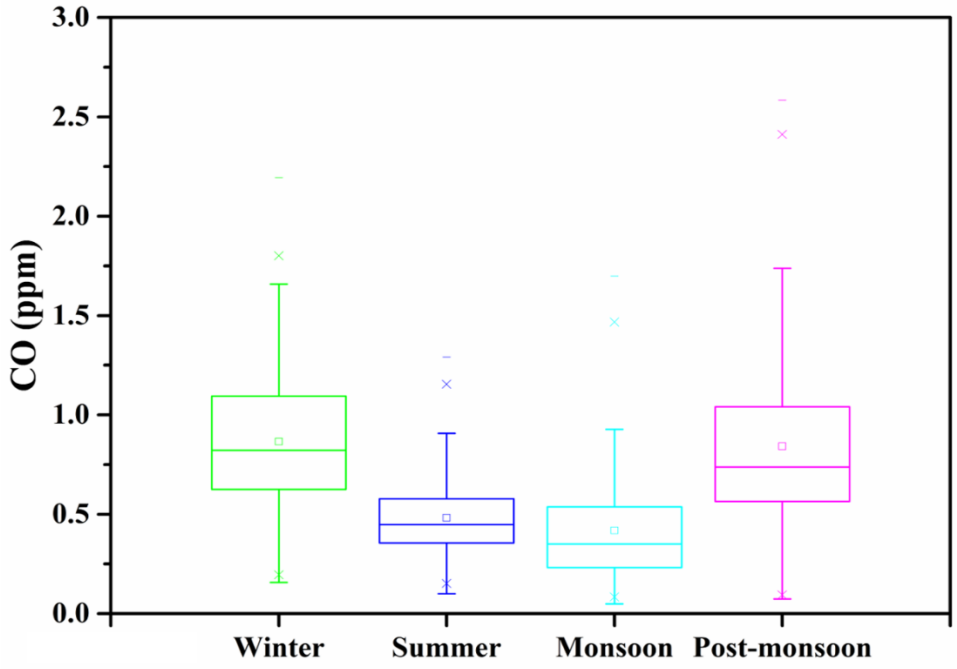


Figure 4.13(b): Box plots of CO concentrations (seasonal average) at Patiala during the study period (2013–2016)

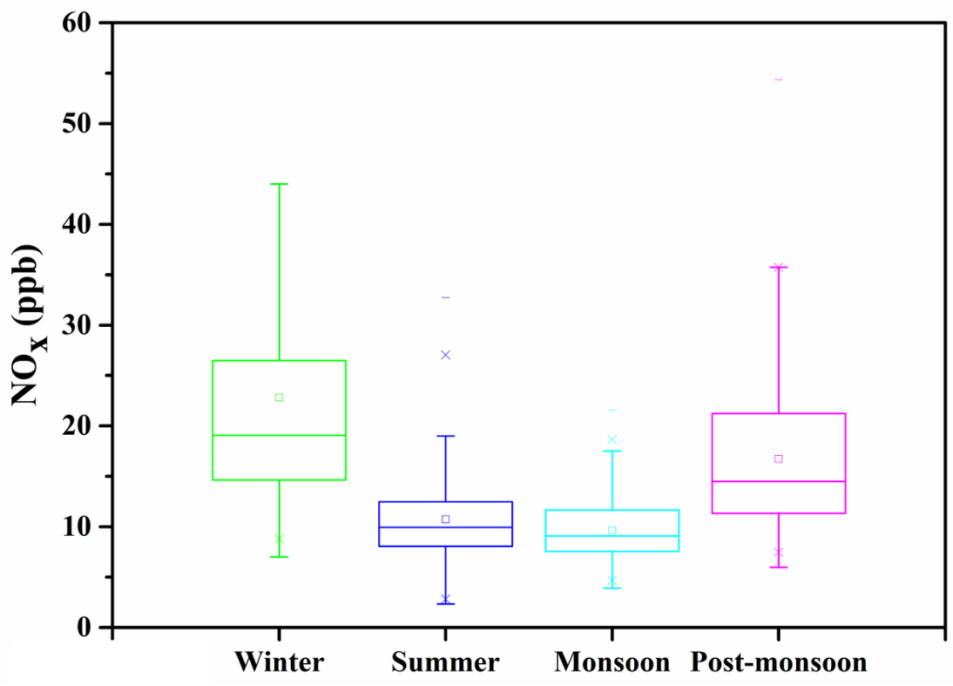


Figure 4.13(c): Box plots of NO_x concentrations (seasonal average) at Patiala during the study period (2013–2016)

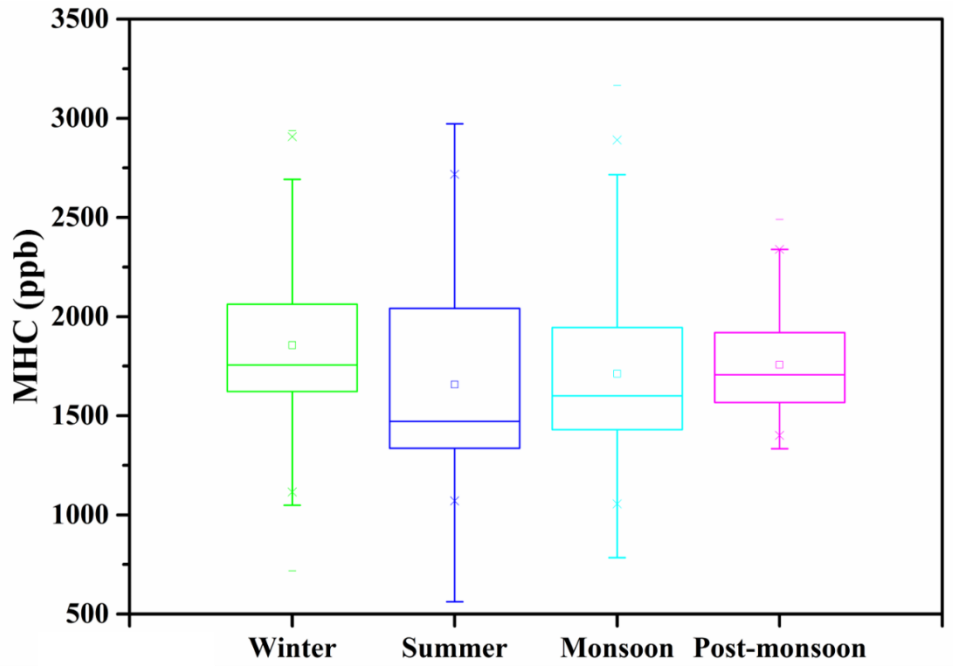


Figure 4.13(d): Box plots of MHC concentrations (seasonal average) at Patiala during the study period (2013–2016)

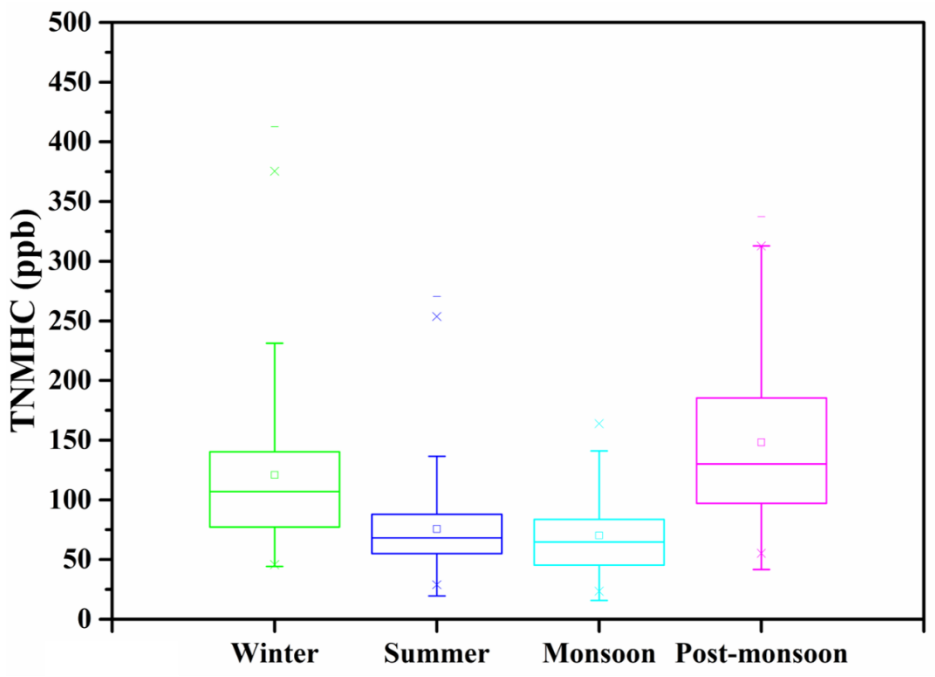


Figure 4.13(e): Box plots of TNMHC concentrations (seasonal average) at Patiala during the study period (2013–2016)

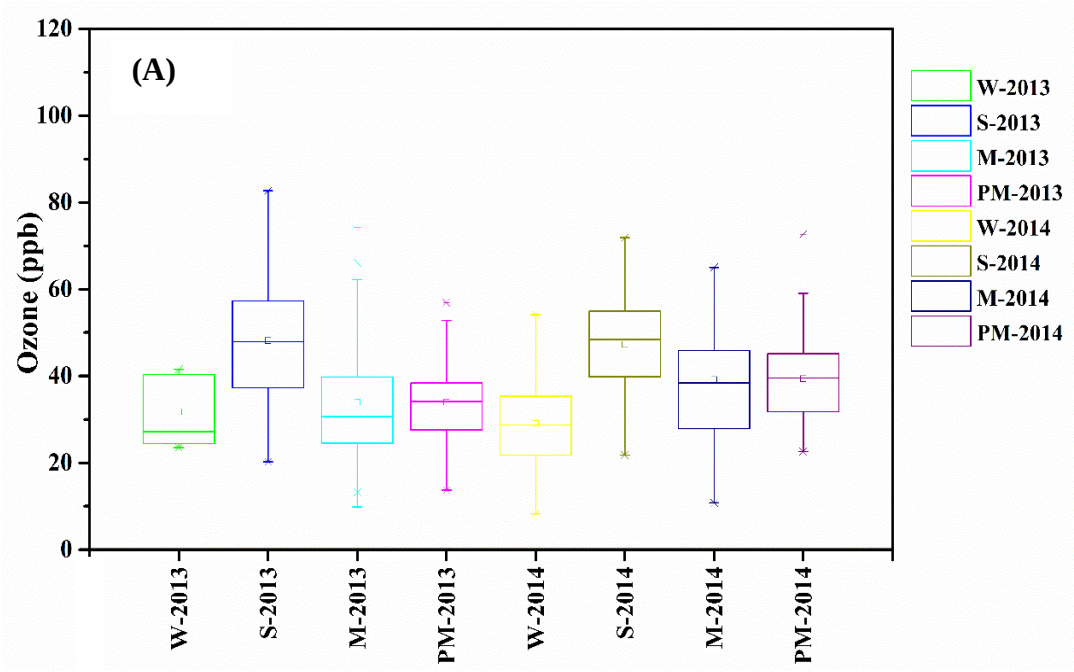


Figure 4.14(a): Box plots of ozone concentrations (seasonal average) at Patiala during the study period (2013–2014)

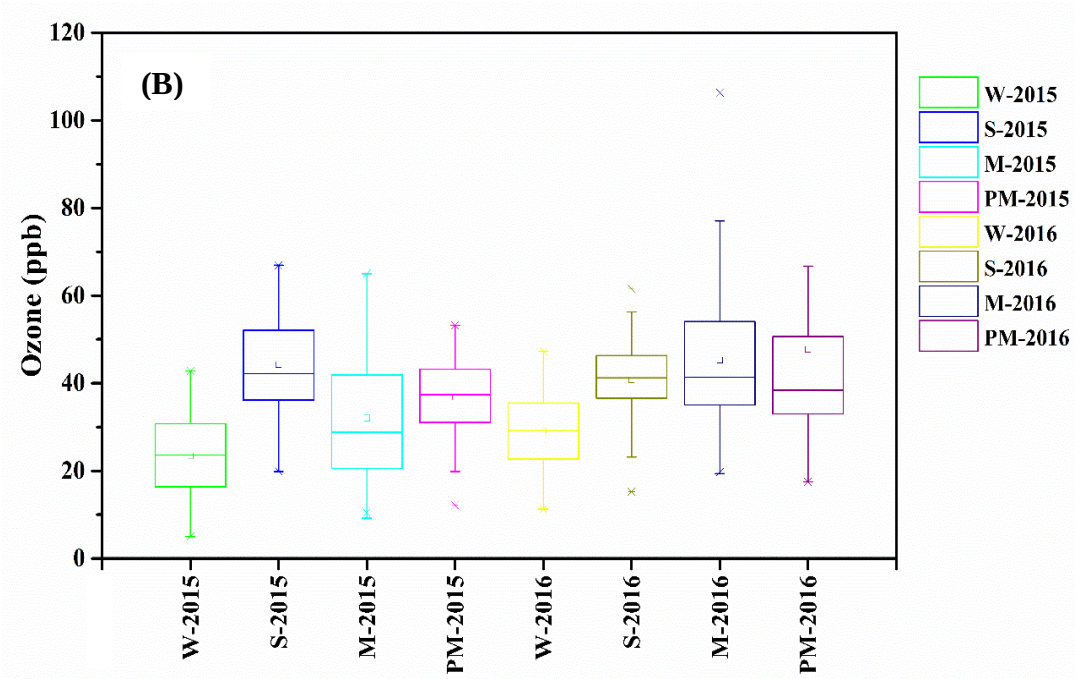


Figure 4.14(b): Box plots of ozone concentrations (seasonal average) at Patiala during the study period (2015–2016)

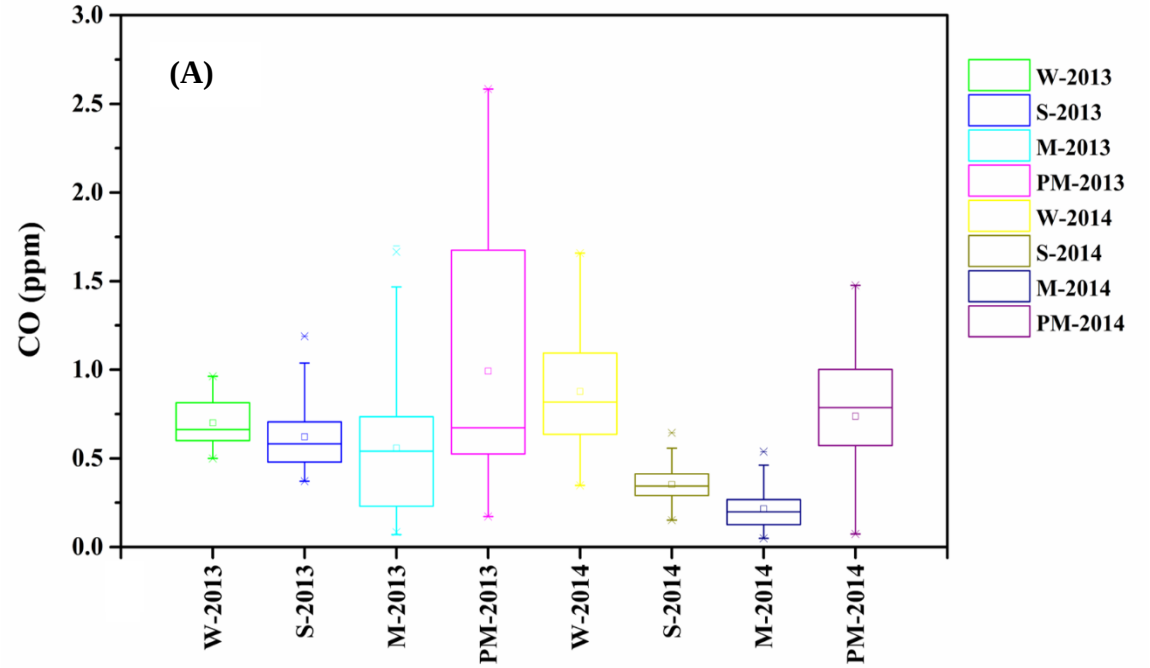


Figure 4.15 (a): Box plots of CO concentrations (seasonal average) at Patiala during the study period (2013–2014)

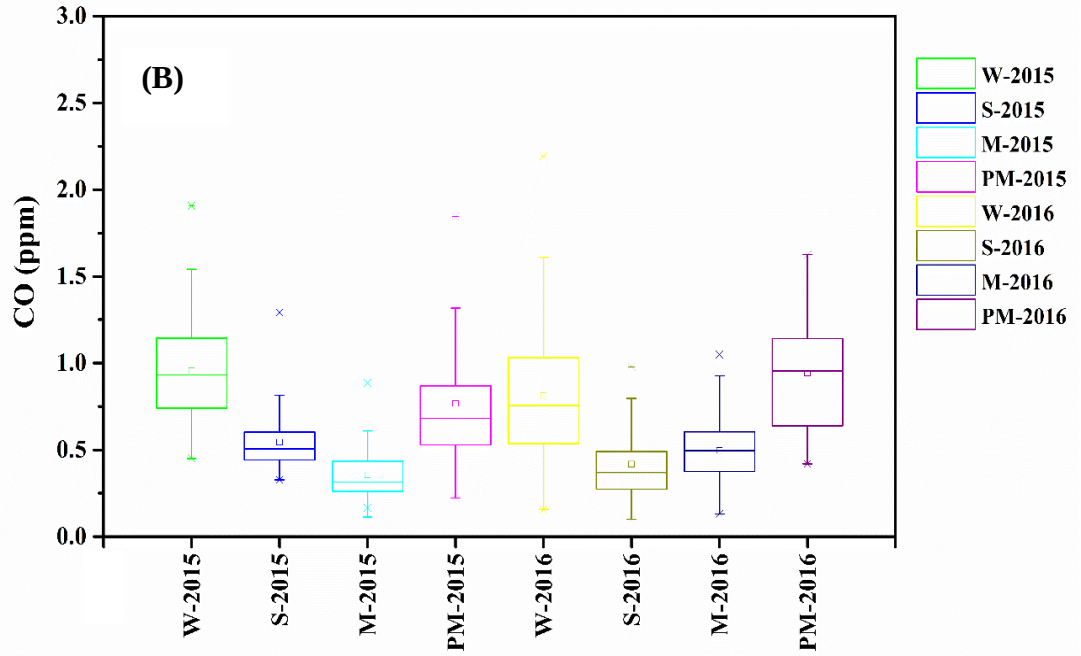


Figure 4.15 (b): Box plots of CO concentrations (seasonal average) at Patiala during the study period (2015–2016)

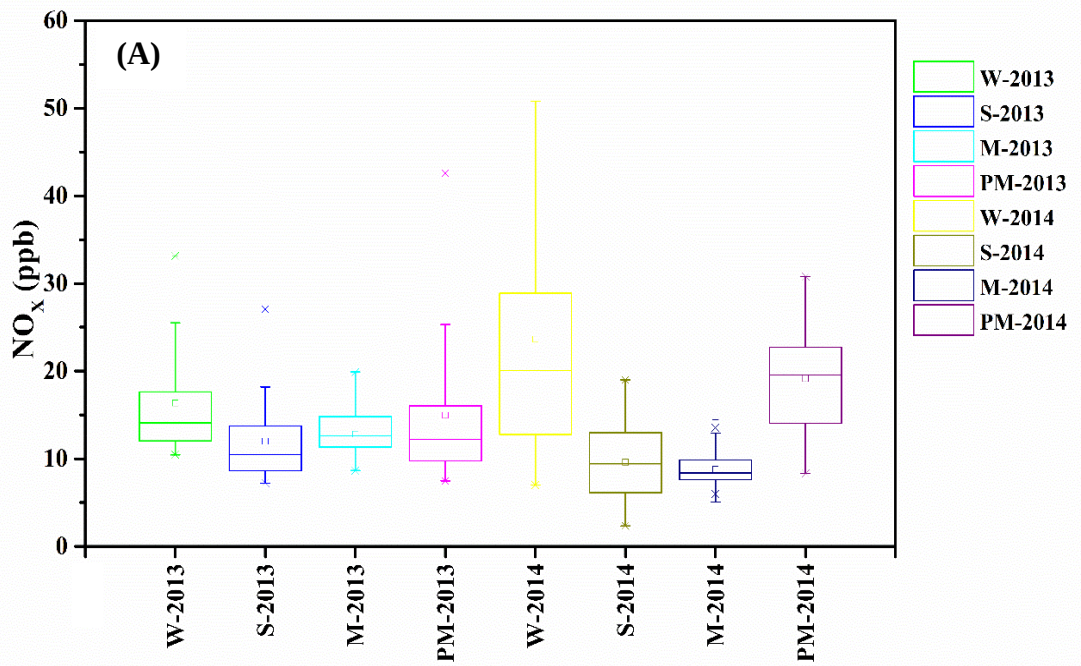


Figure 4.16 (a): Box plots of NO_x concentrations (seasonal average) at Patiala during the study period (2013–2014)

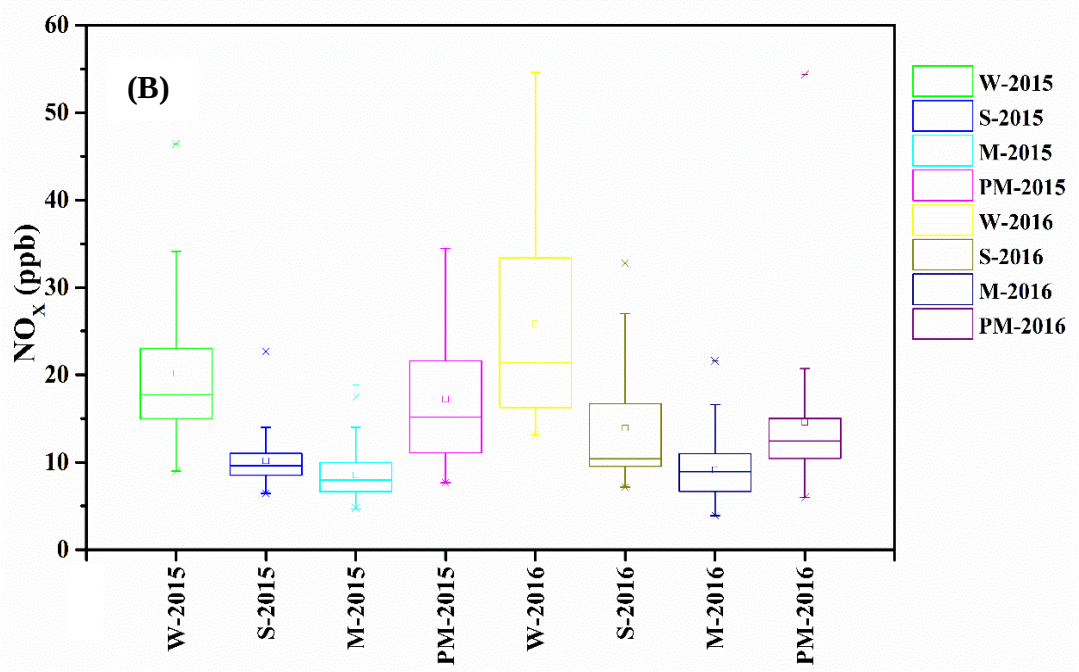


Figure 4.16 (b): Box plots of NO_x concentrations (seasonal average) at Patiala during the study period (2015–2016)

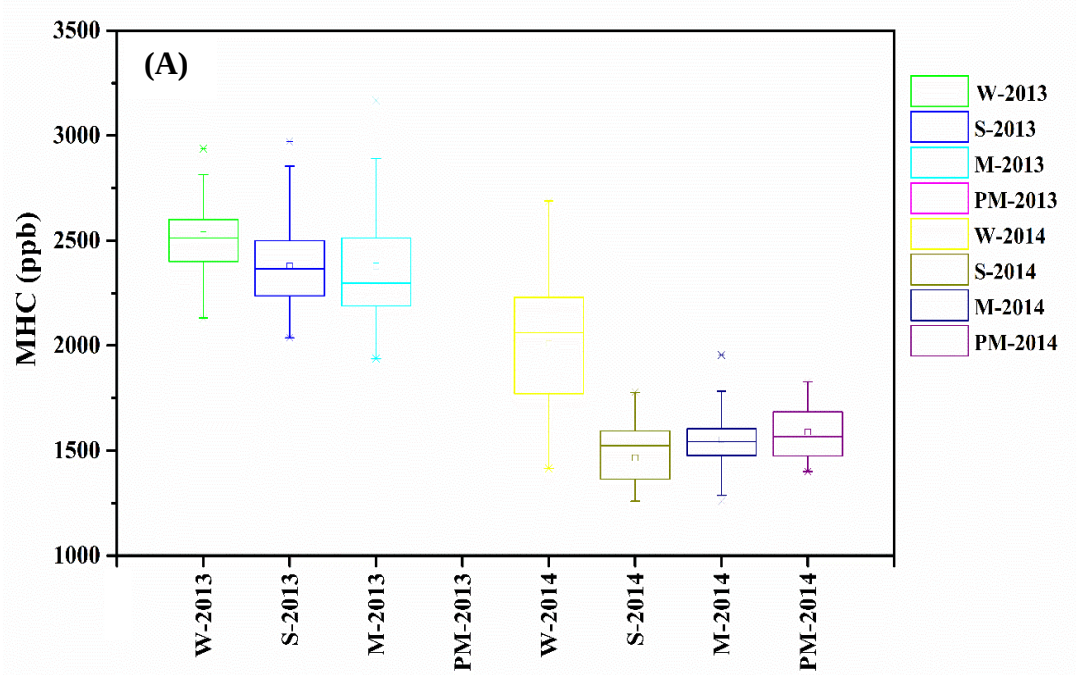


Figure 4.17 (a): Box plots of MHC concentrations (seasonal average) at Patiala during the study period (2013–2014)

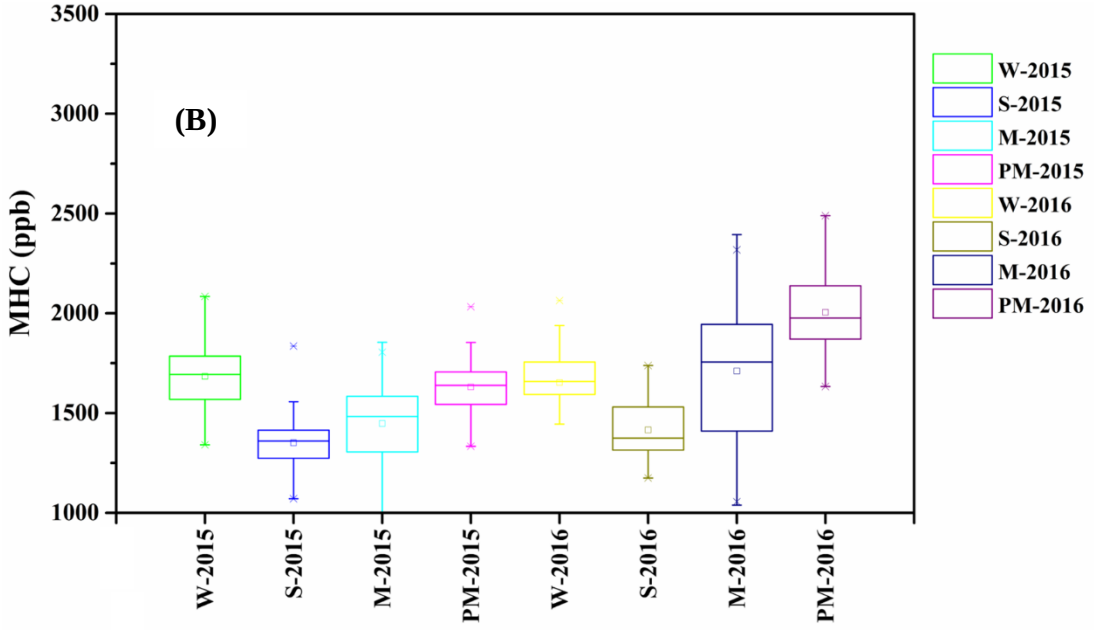


Figure 4.17 (b): Box plots of MHC concentrations (seasonal average) at Patiala during the study period (2015–2016)

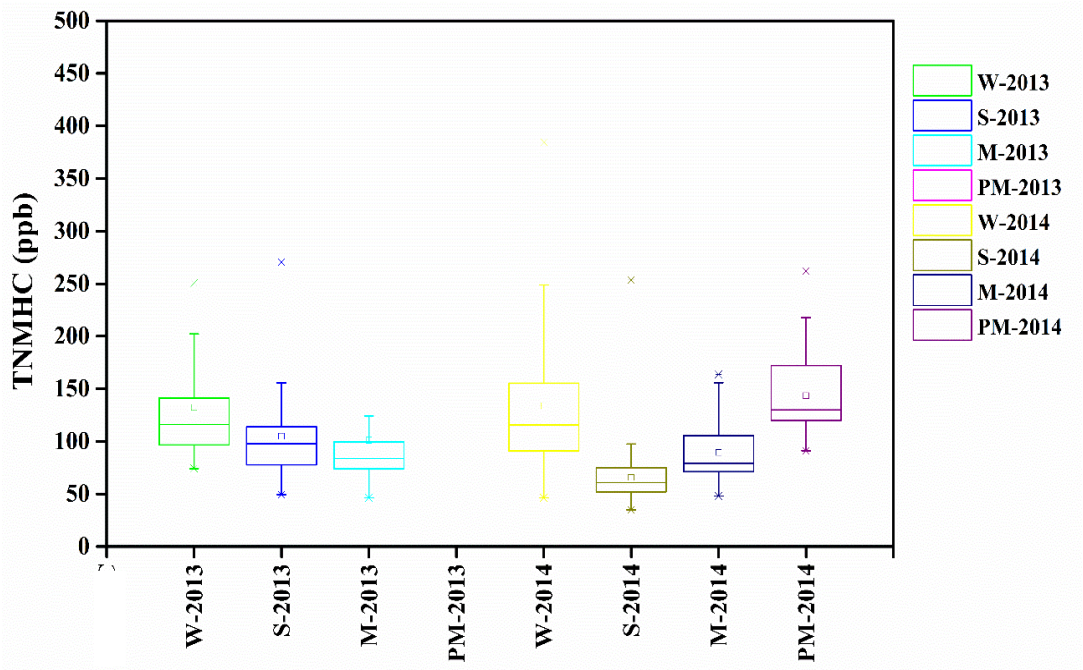


Figure 4.18 (a): Box plots of TNMHCs concentrations (seasonal average) at Patiala during the study period (2013–2014)

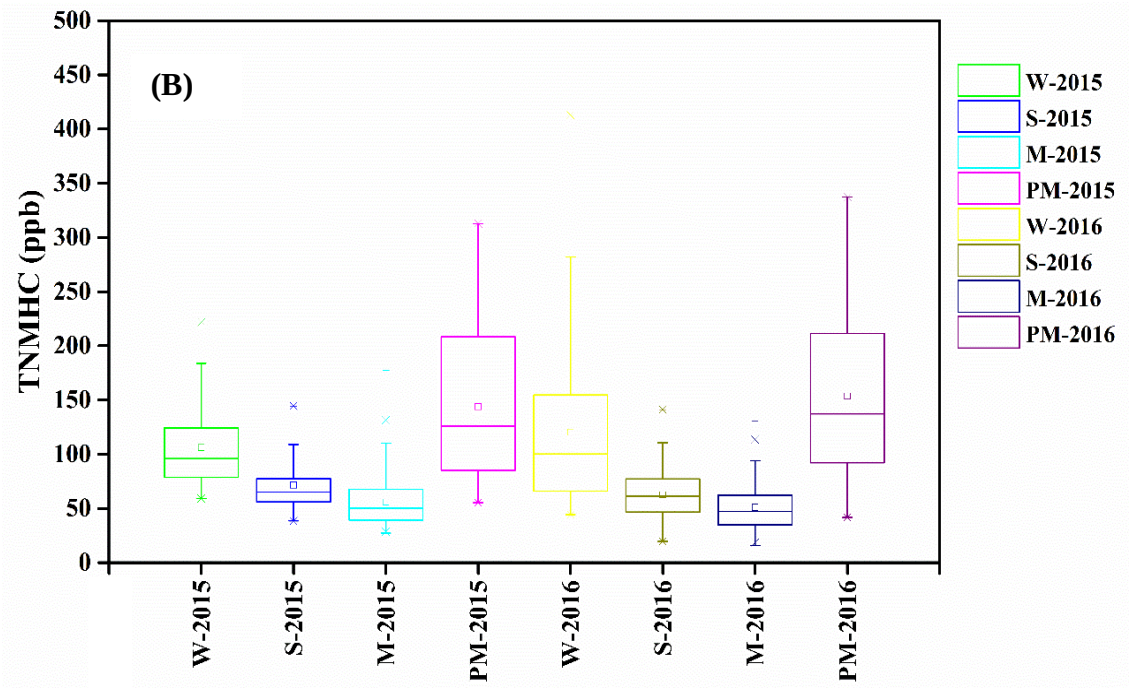


Figure 4.18 (b): Box plots of TNMHCs concentrations (seasonal average) at Patiala during the study period (2015–2016)

4.4 Frequency distribution in levels of surface O_3, CO, NO_x, MHC, TNMHC levels during the study period

The frequency distribution of the concentrations of surface O_3, CO, NO_x, MHC, TNMHC for the study period (2013-2016) on an annual and seasonal basis are shown in Figures 4.19 (a-e) and 4.20 (a-e). The daily levels (24 hours averaged) are divided into classes with appropriate bin size on the basis of the upper and lower limit of the data. On an annual basis as shown in Figures 4.19 (a-e), the concentration range of surface O_3, CO, NO_x, MHC, TNMHC corresponding to percentage maximum frequency are 31-60 ppb (60%), 0-0.05 ppm (45%), 0-30 ppb (92%), 1001-2000 ppb (80%) and 0-100 ppb (67%), respectively.

In Figure 4.20 (a), the daily levels of ozone on a seasonal basis are divided into 4 classes with a bin size of 30 ppb ranges from 0-120 ppb. The maximum percentage frequency of occurrence were 80 % and 70 % in the range of 31-60 ppb in summer and post-monsoon season, followed by 60 % (in 0-30 ppb range) and 55% (in 31-60 ppb range) in winter and monsoon season, respectively. In summer and post-monsoon season, the maximum frequencies were towards higher concentration ranges (31- 60 ppb), this feature is caused by a significant contribution from post-harvest wheat and rice crop residue burning period in summer and post-monsoon season, respectively. Other sources include vehicular emissions, domestic burning and construction

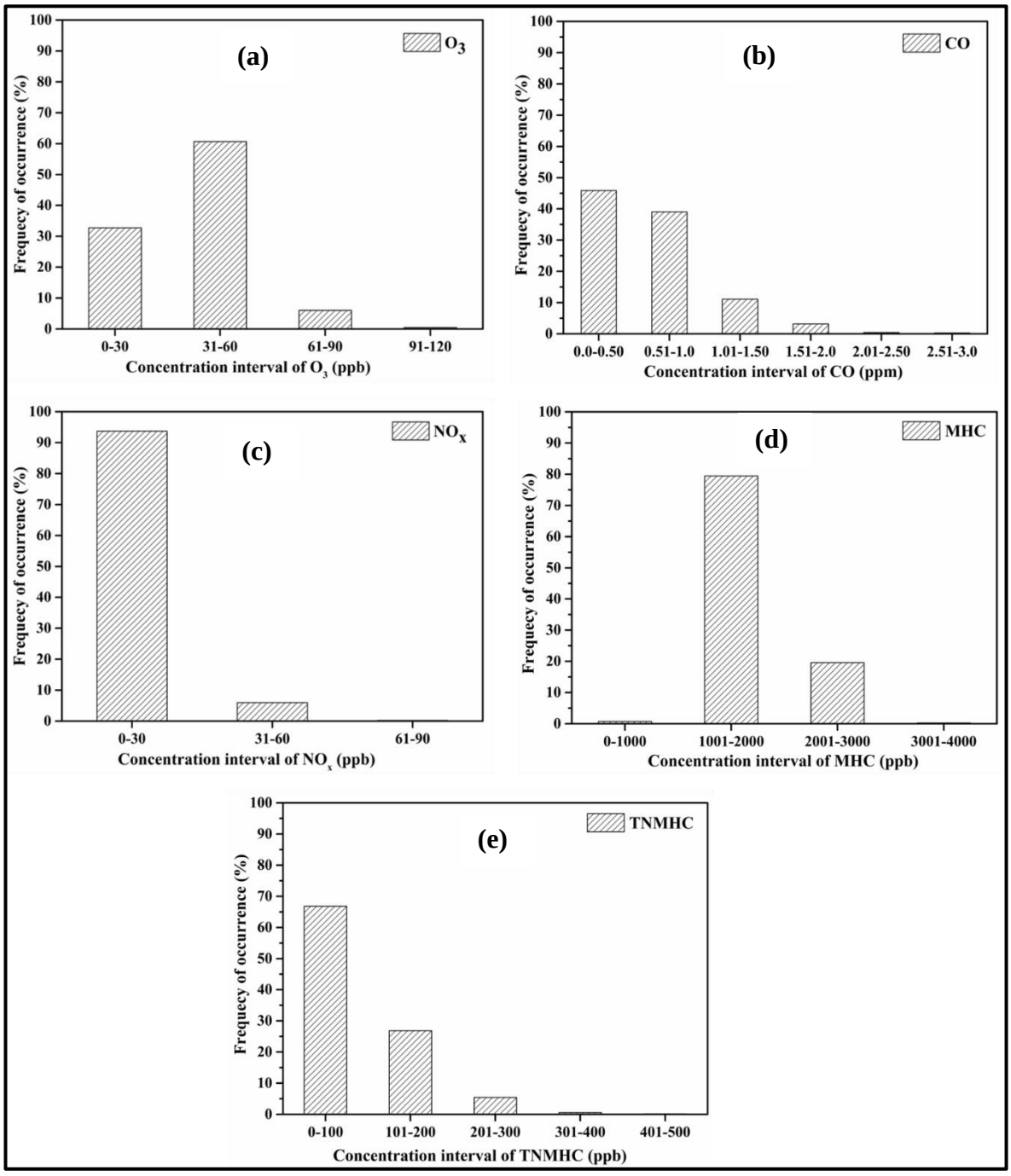


Figure 4.19: Annual percentage frequency distribution of ozone (a), CO (b), NO_x (c), MHC (d), TNMHCs (e) levels on a daily basis during the study period (2013-2016)

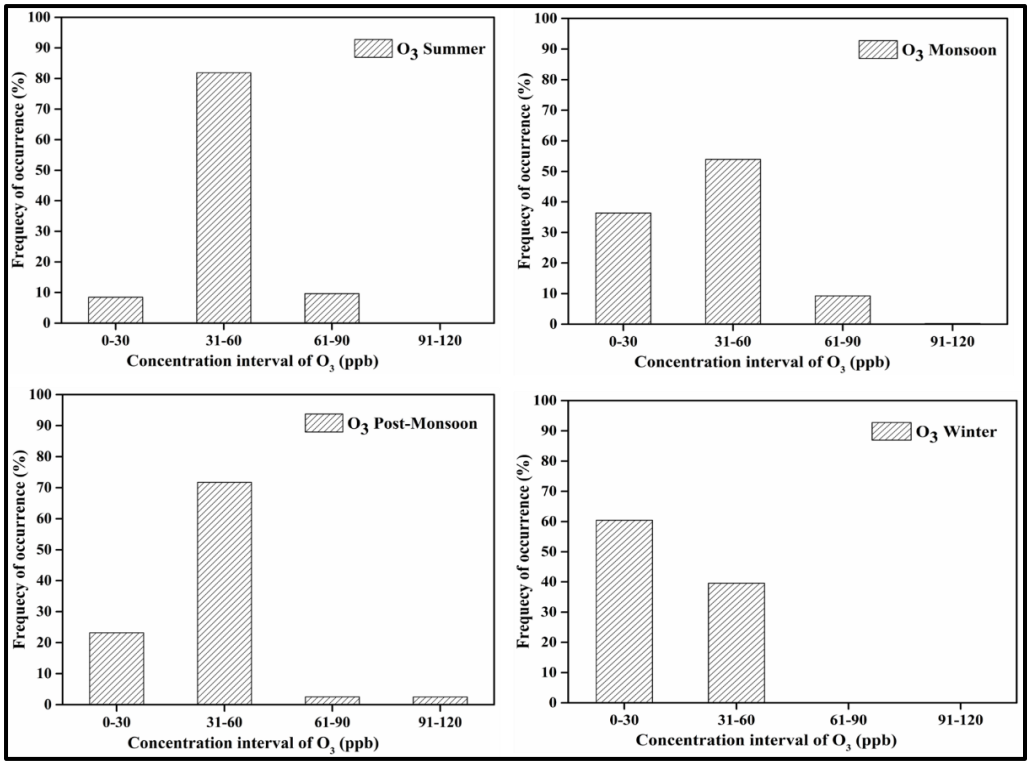


Figure 4.20 (a): Frequency distribution of ozone levels on a seasonal basis during the study period (2013-2016)

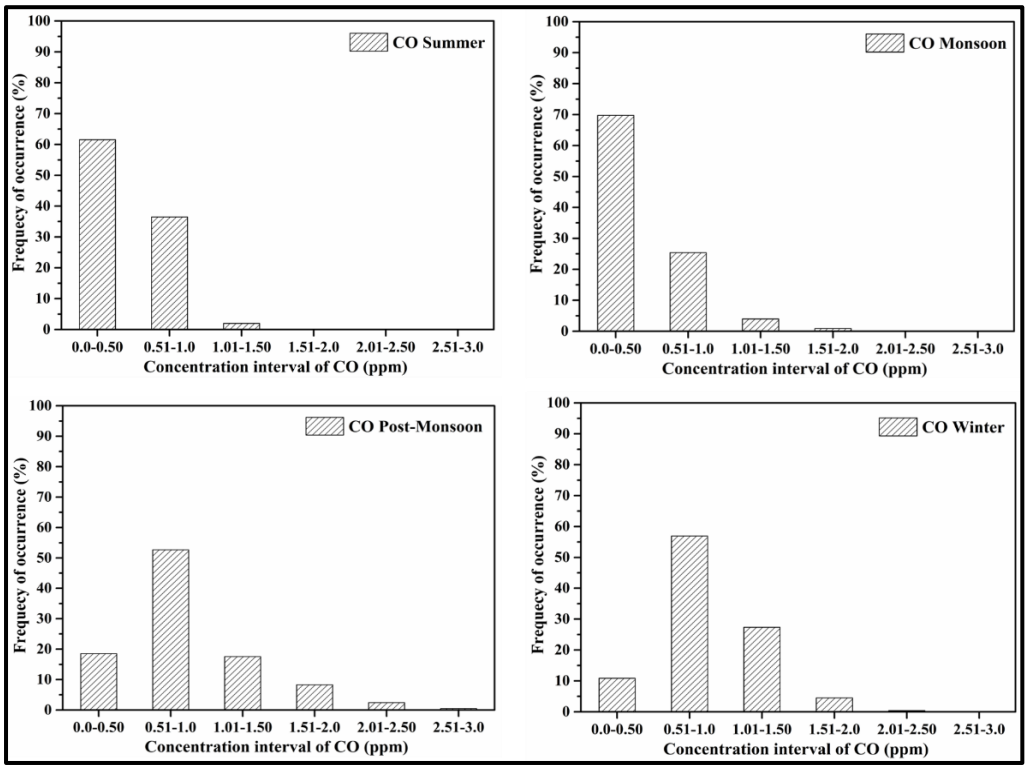


Figure 4.20 (b): Frequency distribution of CO levels on a seasonal basis during the study period (2013-2016)

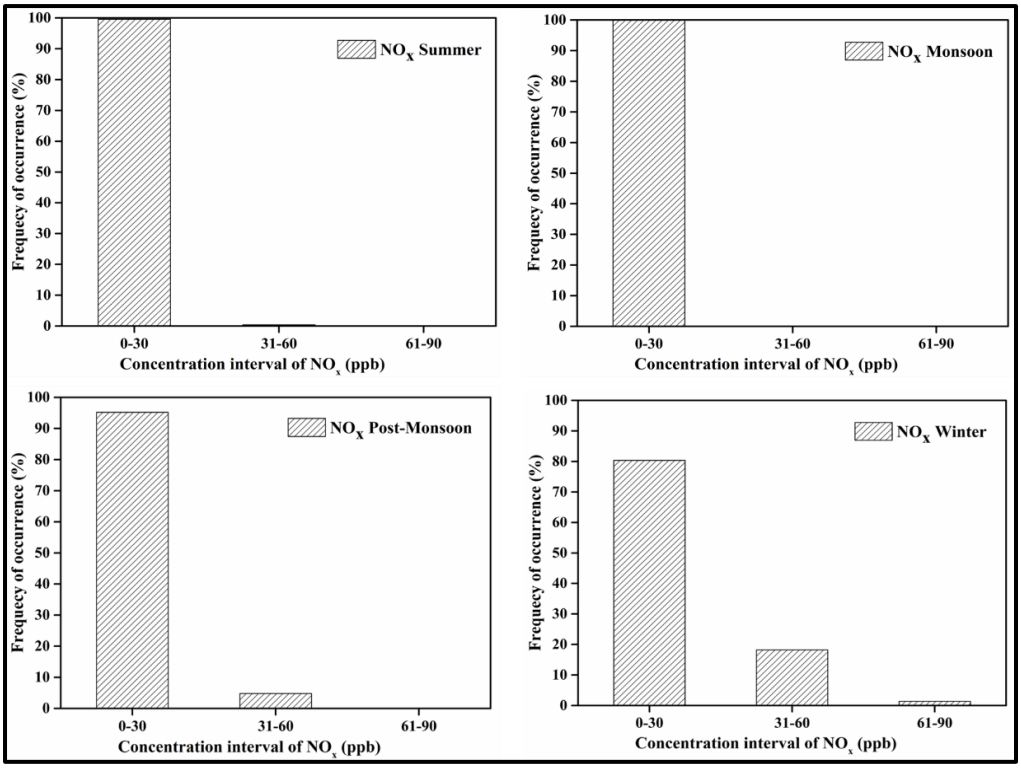


Figure 4.20 (c): Frequency distribution of NO_X levels on a seasonal basis during the study period (2013-2016)

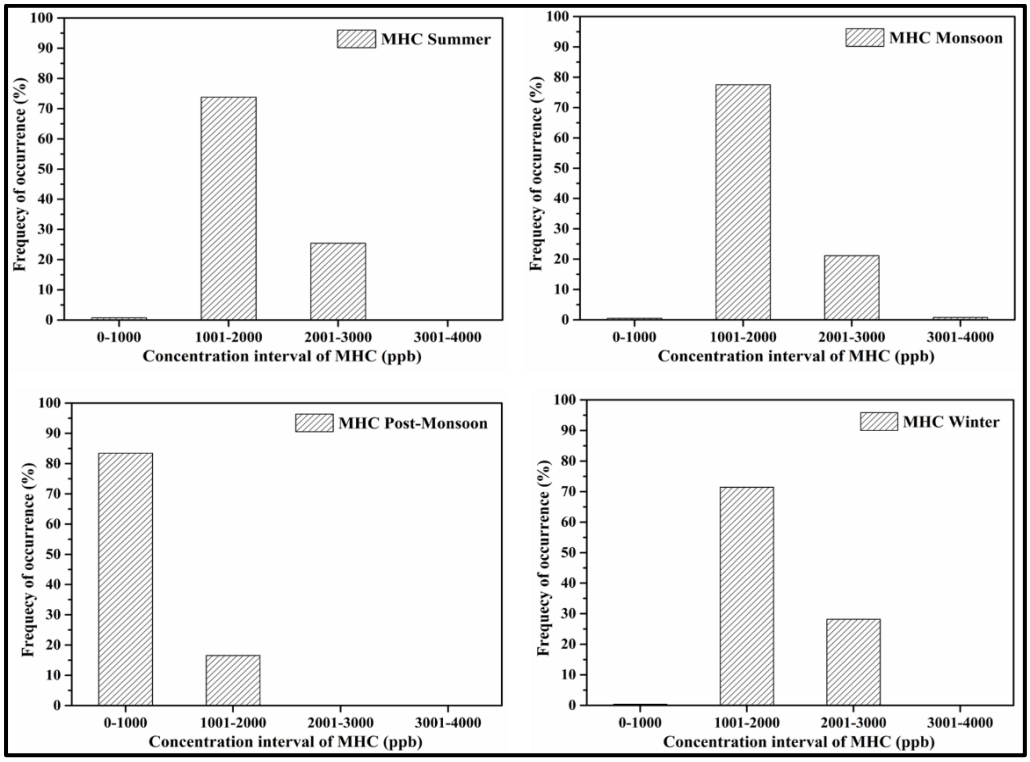


Figure 4.20 (d): Frequency distribution of MHC levels on a seasonal basis during the study period (2013-2016)

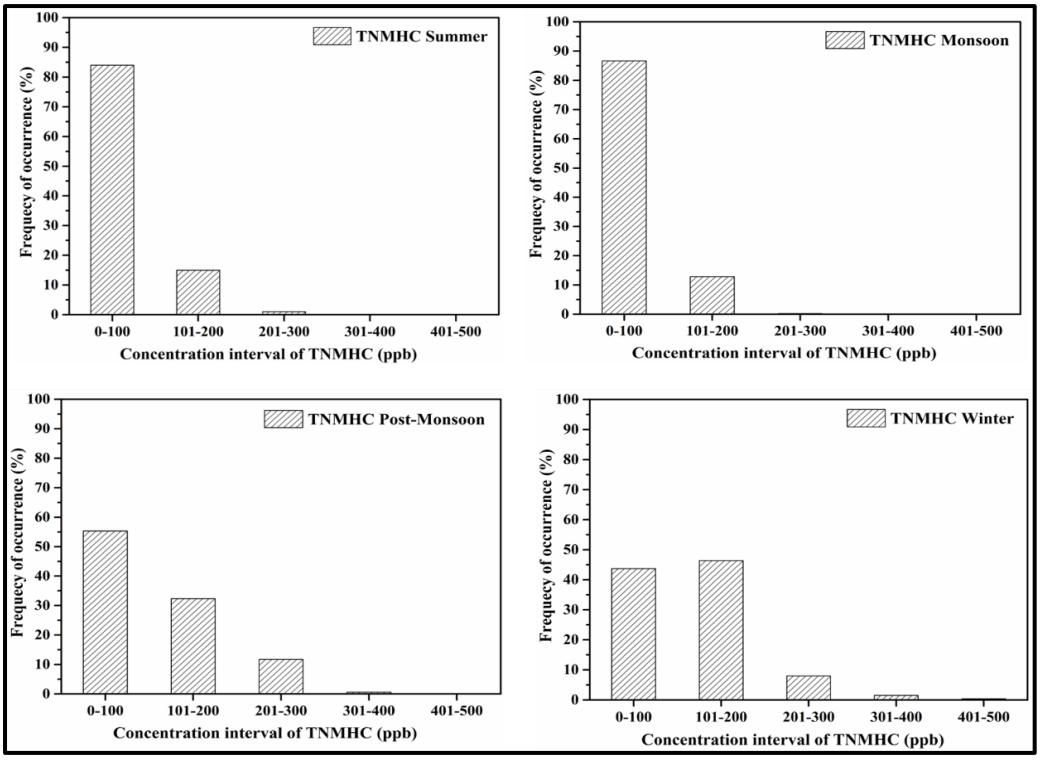


Figure 4.20 (e): Frequency distribution of TNMHCs levels on a seasonal basis during the study period (2013-2016)

activities. But during winter and monsoon season, the peak frequencies shifted towards lower concentration ranges due to limited dispersion of pollutants due to the calm weather conditions in the winter season and wet deposition due to rainfall in monsoon season, respectively.

For all precursor gases (CO, NO_x, MHC, and TMHCs) the percentage frequency distribution peaks shifted towards higher concentration intervals in post-monsoon due to WCRB episodes (burning at non-ideal combustion conditions) followed by winter season due to entrapment of emissions in lower temperature and calm wind conditions. In Figure 4.20 (b-e), the daily levels of CO, NO_x, MHC, and TNMHCs on a seasonal basis are divided into 6, 3, 4 and 5 classes with a bin size of 0.50 ppm, 0-30 ppb, 0-1000 ppb and 0-100 ppb ranges from 0-3 ppm, 0-90 ppb, 0-4000 ppb, and 0-500 ppb, respectively.

As in Figure 4.20 (b), the maximum percentage frequency of occurrence of CO was 70 % and 60 % in the range of 0-0.50 ppm monsoon and summer season, followed by 55 % (in 0-0.50 ppb interval) and 52% (in 0.51-1.0 ppm interval) in winter and post-monsoon season, respectively. For NO_x, the percentage of a frequency distribution is

maximum (80-99%) in the interval of 0–30 ppb for all the four seasons as shown in Figure 4.20 (c).

As in Figure 4.20 (d), the maximum percentage frequency of occurrence of methane hydrocarbons (MHC) was 85 % in the range of 0-1000 ppb in post-monsoon season, followed by 70-75 % (in 1001-2000 ppb interval) in summer, monsoon and winter season, respectively. The maximum percentage frequency of occurrence of total non-methane hydrocarbons (TNMHCs) was 85 % in the range of 0-100 ppb in summer and monsoon season, followed by 55 % in post-monsoon and 45% (in 101-200 ppb interval) in the winter season, shown in Figure 4.20 (e).

4.5 References

1. Reddy, K. K.; Naja, M.; Ojha, N.; Mahesh, P.; Lal, S. *Journal of Earth System Science* **2012**, *121*, 911-922.
2. Ahammed, Y. N.; Reddy, R.; Gopal, K. R.; Narasimhulu, K.; Basha, D. B.; Reddy, L. S. S.; Rao, T. *Atmospheric Research* **2006**, *80*, 151-164.
3. Mittal, S. K.; Singh, N.; Agarwal, R.; Awasthi, A.; Gupta, P. K. *Atmospheric Environment* **2009**, *43*, 238-244.
4. Singh, N.; Mittal, S. K.; Agarwal, R.; Awasthi, A.; Gupta, P. K. *International Journal of Environmental and Analytical Chemistry* **2010**, *90*, 829-843.
5. Gaur, A.; Tripathi, S.; Kanawade, V.; Tare, V.; Shukla, S. *Journal of Atmospheric Chemistry* **2014**, *71*, 283-301.
6. Kumar, V.; Sarkar, C.; Sinha, V. *Journal of Geophysical Research: Atmospheres* **2016**, *121*, 3619-3633.
7. Yadav, R.; Sahu, L.; Beig, G.; Jaaffrey, S. *Atmospheric Research* **2016**, *176*, 96-107.
8. Satsangi, G. S.; Lakhani, A.; Kulshrestha, P.; Taneja, A. *Journal of Atmospheric Chemistry* **2004**, *47*, 271-286.
9. Sahu, L.; Lal, S. *Atmospheric Environment* **2006**, *40*, 880-891.
10. Beig, G.; Gunthe, S.; Jadhav, D. *Journal of Atmospheric Chemistry* **2007**, *57*, 239-253.
11. Reddy, B. S. K.; Kumar, K. R.; Balakrishnaiah, G.; Gopal, K. R.; Reddy, R.; Sivakumar, V.; Lingaswamy, A.; Arafath, S. M.; Umadevi, K.; Kumari, S. P. *Aerosol and Air Quality Research* **2012**, *12*, 1081-1094.
12. Sagar, K.; Srivastava, R.; Sarkar, S. *Br. Journal of Environmental Science* **2015**, *3*, 62-70.

CHAPTER -5

5.1 Dependence of Ozone on precursor gases

5.1.1 Monthly averaged levels of ozone and precursor gases (CO, NOx, MHC, and TNMHCs)

Figures 5.1 and 5.2 show concentration levels of monthly averaged data for gaseous pollutants (O_3, CO, and NO_x) and (O_3, MHC, and TNMHCs) for the study period, calculated from 24 h hours data sets. The increase/decrease in monthly concentration levels of each gas indicates the significant influence of human activities, seasonal variations, and related chemical transformations. A maximum of ozone concentration level is corresponding to the minimum levels in precursor gases (CO, NO_x, MHC, and TNMHCs) that lead to the formation of ozone, with favorable meteorological conditions. The daily $O_{3\ max}$ / $O_{3\ min}$ ratio is a pollution index, for polluted sites, has values in excess of 10 while low values smaller than 1.40 for clean areas.[1] The index value of the order of 15.60 indicates the study location has remarkable ozone pollution.

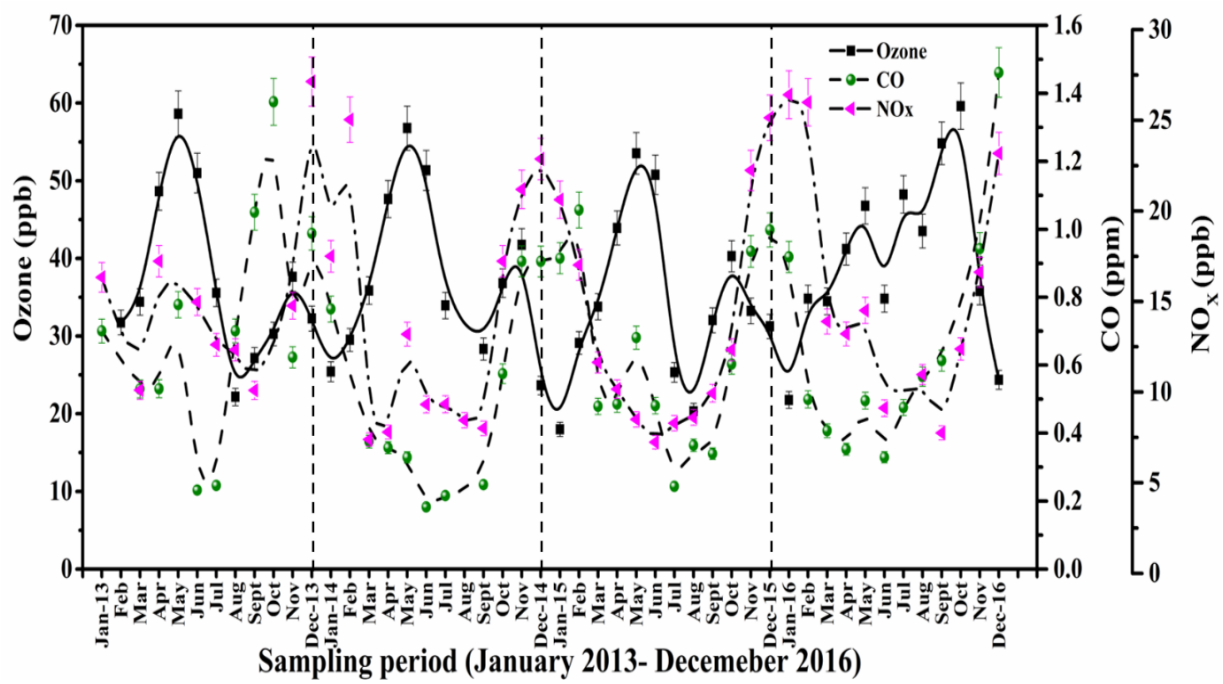


Figure 5.1: Concentration levels (monthly average) gaseous pollutants (O_3, CO and NO_x) at Patiala during the study period (2013–2016)

It can be seen from Figures 5.1 and 5.2 that ozone level starts increasing from January and the first maxima appeared in May followed by decreases till September with a minimum in monsoon season. Second maxima appeared in November followed again by a decrease till January. The same trend of variations in the level of ozone was

through-out the observation period with a little change in the magnitude of the average concentration levels.

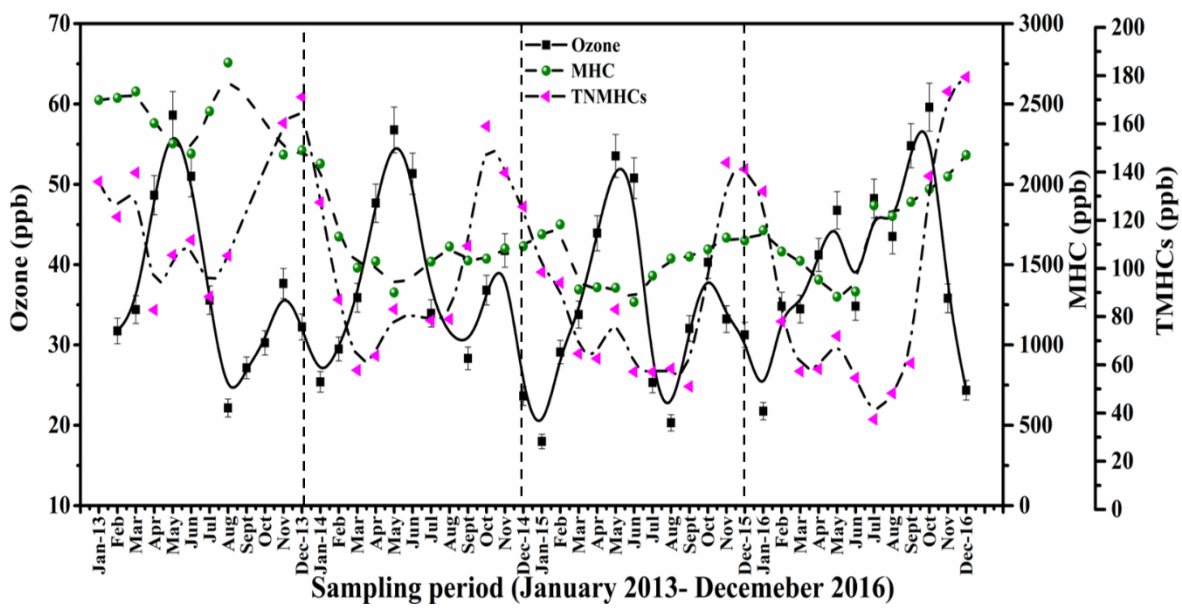


Figure 5.2: Concentration levels (monthly average) gaseous pollutants (O_3, MHC and TNMHCs) at Patiala during the study period (2013–2016)

5.1.2 Contour plots of Ozone with CO and NO_x

Figures 5.3 (a-d) shows the dependence of O_3 on the levels of CO and NO_x during daytime (0800-1700 hours) for different seasons. During pre-monsoon/summer and post-monsoon seasons, the enhancements in levels of O_3 (as indicated by the intensity of red color) were observed in the low-CO and low-NO_x regimes, which is a result of the rise in the photochemical production of O_3. The large production of ozone is also indicated by a large number of data points in this region. In the winter season, the low levels of ozone are spread over the whole contour plot indicating continuous production of ozone as reflected from yellow to blue color. For low NO_x regime, an increase in NO_x leads to an increase in OH and to corresponding increases in the oxidation rate of hydrocarbons and in the levels of ozone the rate is independent of CO. In the high-NO_x limit (NO-saturated), the rate of ozone formation increases with CO but decreases as NO increases. [2]

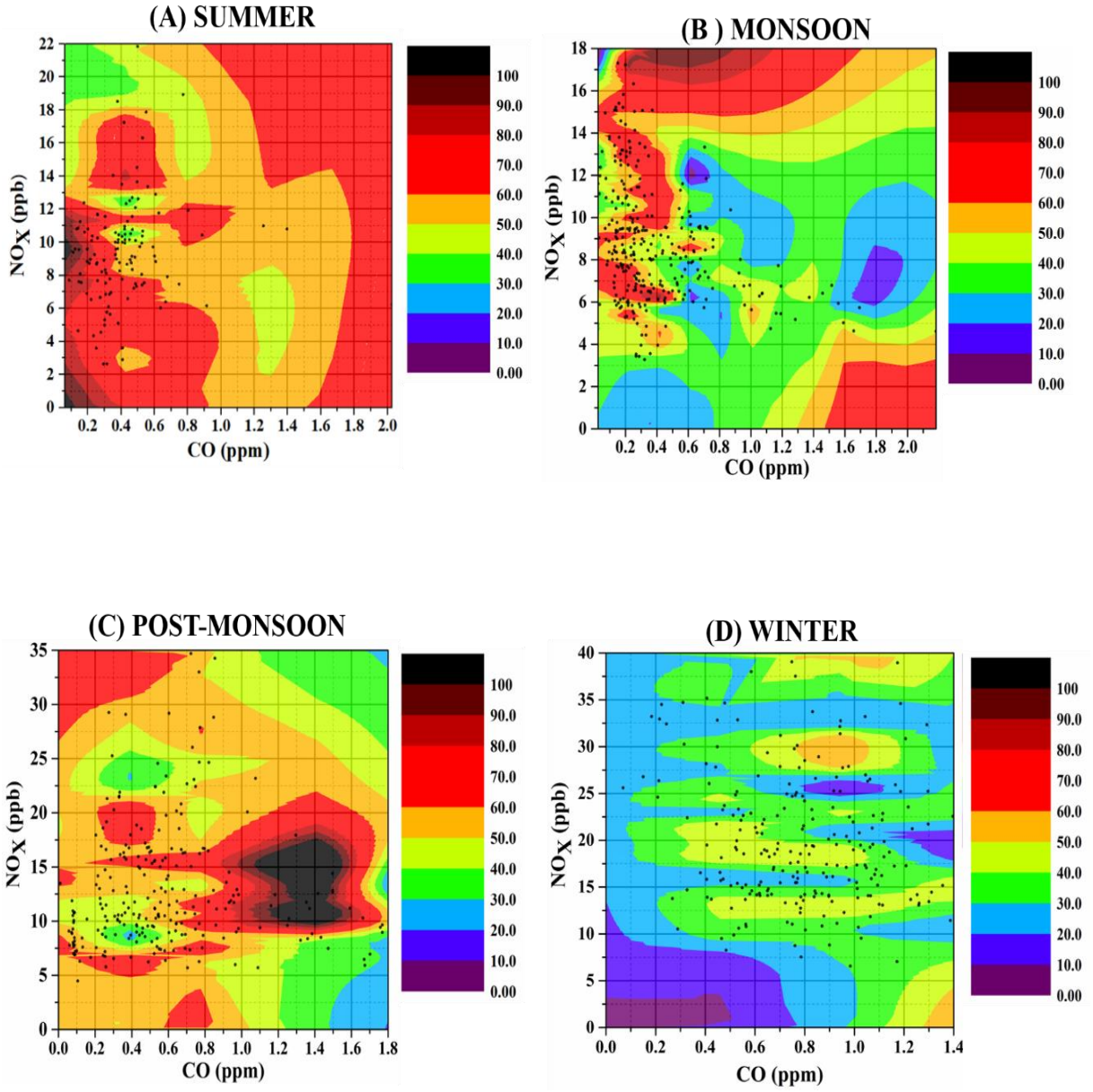


Figure 5.3: Contour plots showing the dependence of ozone concentrations on NOx and CO during daytime (0600–1700 hours) in four seasons (a) summer, (b) monsoon, (c) post- monsoon and (d) winter at Patiala during 2013–2016. Colour codes indicate ozone levels in ppb.

5.2 Dependence of ozone on meteorology

Figures 5.4 (a and b) shows the monthly variation in ozone levels, the increase/decrease indicates the significant influence of meteorological variations and related chemical transformations. The rise and fall in ozone levels are directly corresponding to the solar radiation and ambient temperature exposure and inverse to the relative humidity. The ozone level starts increasing from January and the first

maxima appeared in May followed by decreases till September with minima in monsoon season. The maximum monthly average of ozone concentration is corresponding to minimum relative humidity (34-70%) and high solar radiation (153-366w/m^2) and daily ambient temperature (24-34°C) along with the emissions of precursor gases from the wheat crop residue (with moisture content less than 5%) burning. The second-highest monthly average of ozone concentration occurred in November corresponding to moderate relative humidity (54-71%), solar radiation (159-219w/m^2) and daily ambient temperature (17-25°C) along with the emissions of precursor gases from the rice crop residue (with moisture content 40-50%) burning. There is a noteworthy difference in the amplitude of two maxima (May and November) observed during post-harvest crop residue burning of wheat and rice crop, respectively, pinpoints the favourable meteorological conditions i.e., high ambient temperature, solar radiation, and low relative humidity for the photochemical production of ozone in summer than the post-monsoon season, during rice crop residue burning (RCRB) period.

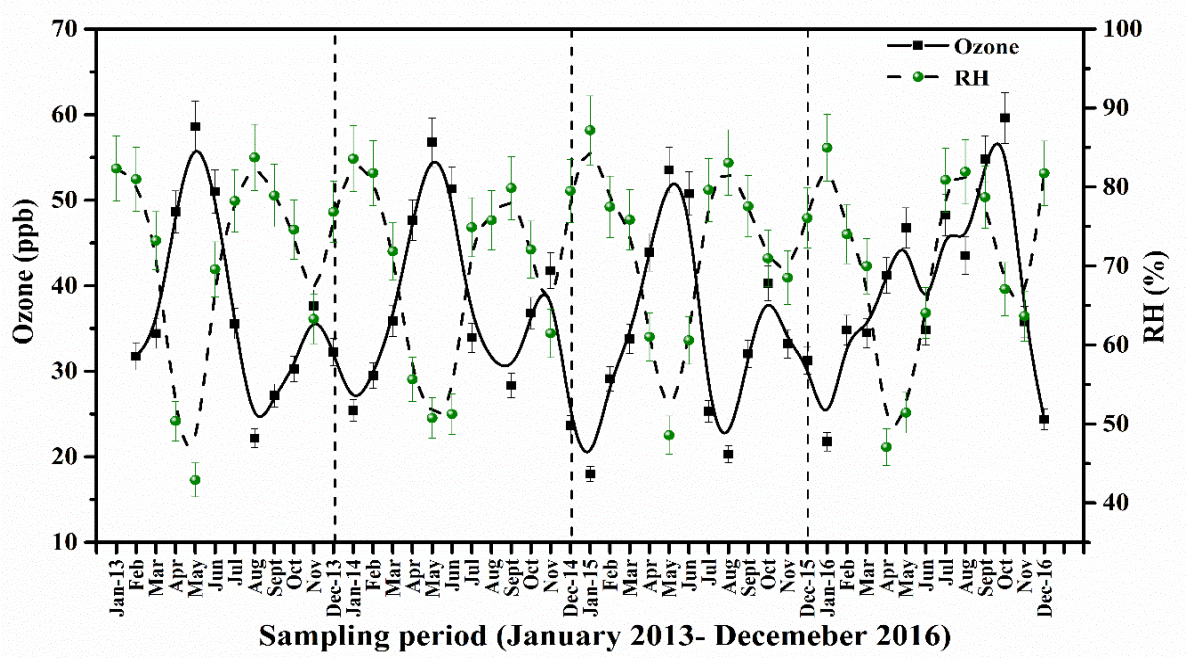


Figure 5.4 (a): Concentration levels (monthly average) of ozone and meteorological variable (RH) at Patiala during (2013–2016)

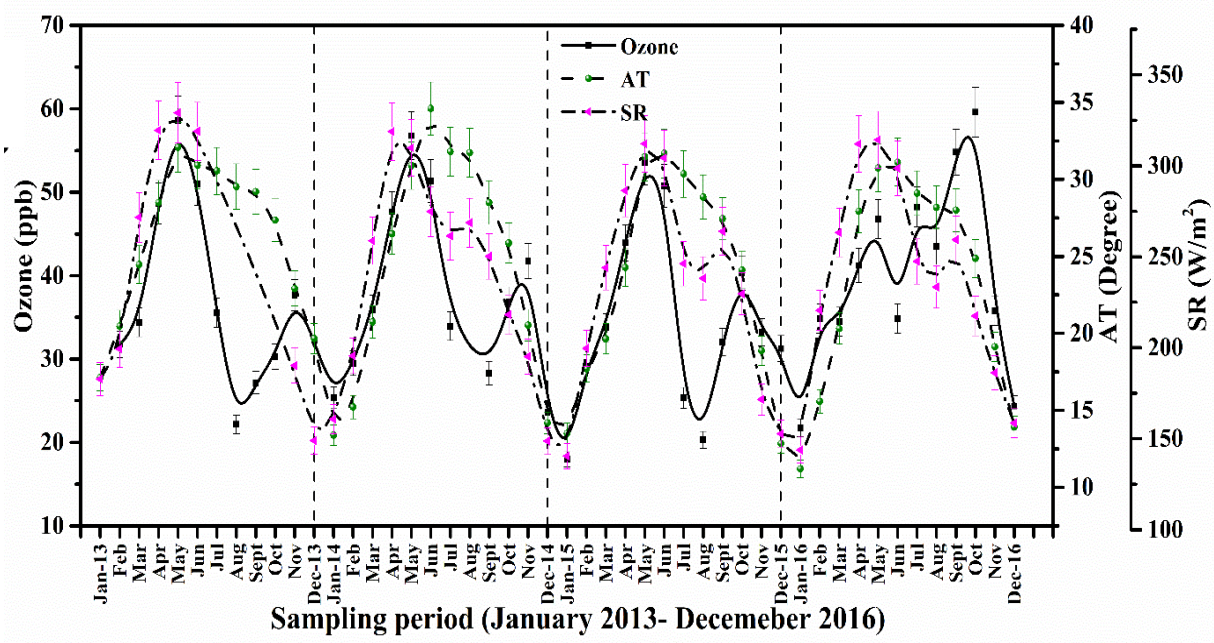


Figure 5.4 (b): Concentration levels (monthly average) of ozone and meteorological variable (AT and SR) at Patiala during (2013–2016)

5.3 Rate of change of ozone concentration on the diurnal scale during different seasons

The seasonal variations of the rate of increase/decrease of ozone have been plotted as a function of diurnal scale for each season of the study period (2013 - 2016) as shown in Figure 5.5 (a). The change in the rate of ozone (dO_3/dt) increase/decrease for all the seasons varies almost on the same pattern with the difference in magnitude. The rate of increase (in the morning) and decrease (in the evening) were highest during 0900-1000 hour and 1700-1900 hours, respectively. On the other hand, the values of dO_3/dt show little variations from night till early morning hours. The highest rate of increase of dO_3/dt during post-monsoon ranges between 14.3 to 18.4 ppb/hr followed by summer season (10.2 to 13.5 ppb/hr) indicated the influence of increased emission of precursor gases, especially from the episodic crop residue burning event in respective periods. The average rates of increase/decrease [d(O3)/dt] was as high +6.3 ppb/hr as compared to different cities of India, e.g., Delhi [3], Agra [4], Kanpur [5], Udaipur [6] lies between (2.5 to 5.9) ppb/hr during 0800-1100 hours and -8.1 ppb/hr lies between (-2.4 to -6.4) during 1700-1900 hours. The rate of change of ozone of surface ozone at any site is dependent on its net chemical production (production loss), surface deposition, and advection.

In the seasonal index (SI) of ozone, the well-marked month-to-month cycle was shown in Figure 5.5(b) with the highest and lowest ozone means occurring in May and August, respectively. This indicates that there is a fluctuation in the seasonal index (SI) from a minimum of about 60 % (in January and August) of average to 159 % during the course of one complete cycle of one year. The peak in May is followed by a secondary peak in November, corresponding to the emission levels contributed by wheat and rice crop residue burning events, respectively.

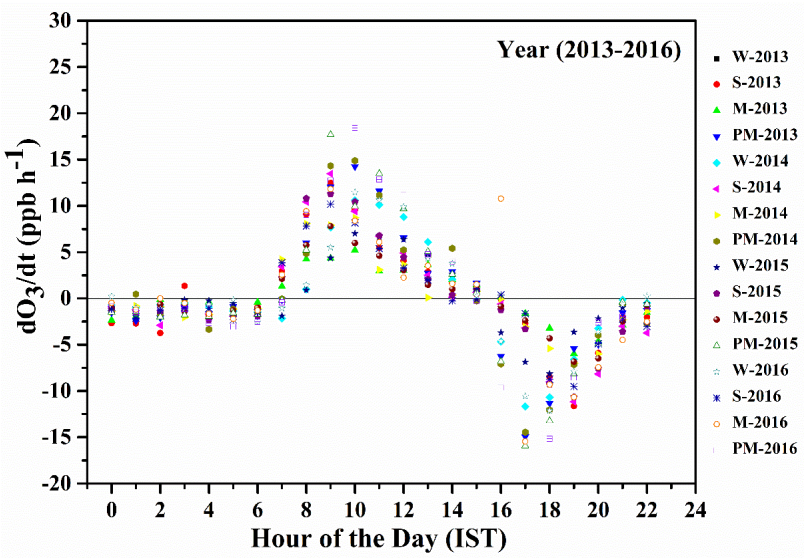


Figure 5.5 (a): Diurnal dependence of the rate of change of ozone (dO_3/dt) for four seasons (winter, summer, monsoon, and post-monsoon) at Patiala during years 2013–2016

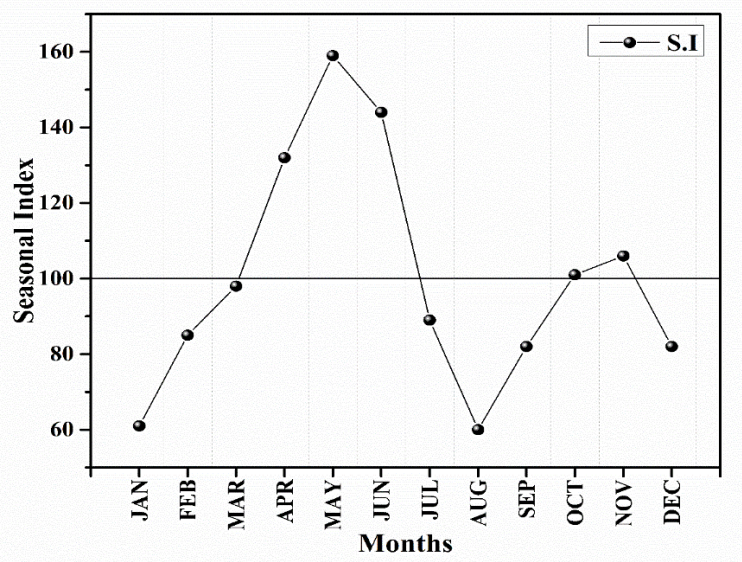


Figure 5.5 (b): Month-to-month variability of the seasonal index of ozone

5.4 Dependence of O_3 on Precursors and Meteorology: Correlation Analysis

The significant dependence of ozone with its precursor gases (CO, NO_x, MHC, TNMHCs) and meteorological variables (AT, SR, RH, and WS) is highlighted in this section. The complex relationship between ozone, precursor gases and meteorological variables that may induce or impede its production was analyzed using correlation analysis computed using SPSS® IBM® Version 20 (Statistical Package for the Social Sciences). A parametric, Pearson correlation test is advantageous for the analysis of quantitative variables with different units. The significant correlations ($p < 0.01$) were observed in all cases as shown in Table 5.1.

Ozone exhibits strong day to day, seasonal and long-term timescale variations which are determined by anthropogenic processes and meteorological conditions. Since, CO, NO_x, MHC, and TNMHCs being the precursor gases for the production of ozone under photochemical conditions. The ozone is negatively correlated with the levels of CO (-0.250), NO_x (-0.262), MHC (-0.151) and TNMHCs (-0.252) in the tropospheric region. The precursor gases (CO, NO_x, MHC, and TNMHCs) showed a strong positive correlation with each other is attributable to common anthropogenic emission sources such as road traffic and combustion sources. The high positive correlation coefficient of 0.565 and 0.449 was observed between SR, ambient temperature and ozone which highlights the importance of photochemistry in the formation of ozone. The sun radiations control the temperature, and hence, the photolysis efficiency will be high, leading to the generation of ozone from its precursors such as hydrocarbons, NO*x*, etc. The negative correlation ($r = -0.682$) between the surface ozone and humidity lies in the fact that with an increase in humidity levels, the precursor gases are washed away from the environment, resulting in a decrease in the production of tropospheric ozone concentration, the major photochemical path for the removal of ozone. Higher levels of humidity slow down the photochemical process due to its association with greater cloud abundance, atmospheric instability and low incoming SR.[7] Also, the surface ozone is depleted through the deposition of its molecules on water droplets.[8] The positive correlation ($r = 0.208$) between ozone and wind seed indicates ozone transport. The increase of WS implies the increasing transport of air, thus the influence of WS on primary and secondary pollutants. In the case of primary pollutants, it acts as a diluting agent, while ozone concentration increases due to its transport from other regions.

The precursor gases (CO, NO_x, MHC and TNMHCs) were negatively correlated with SR (-0.557, -0.477, -0.300, and -0.441), AT (-0.527, -0.485, -0.320, -0.355), and WS (- 0.401, -0.392, -0.268,-0.468) and positively correlated with RH (0.222, 0.130, 0.217, 0.123) respectively. Hence, primary pollutants concentration gets declined as the photochemical reaction for the production of secondary pollutant (O_3) become more active with increasing solar radiation, ambient air temperature, and wind speed. The negative correlation between precursor gases pollutants and wind velocity, which serves as a diluting agent. Similar results were shown in other studies, and correlation values are highly sited specific.[9]

Table 5.1: Pearson correlation coefficient matrix between O_3, precursor gases (CO, NO_x, MHC, TNMHCs) and meteorological variables (RH, SR, WS, and AT) for the study period

	O_3	**CO**	**NOx**	**MHC**	**TNMHCs**	**RH**	**SR**	**WS**	**AT**
O_3	1	-0.250	-0.262	-0.151	-0.252	-0.682	0.565	0.208	0.449
CO		1	0.639	0.351	0.656	0.222	-0.557	-0.401	-0.527
NOx			1	0.335	0.642	0.130	-0.477	-0.392	-0.485
MHC				1	0.474	0.217	-0.300	-0.268	-0.320
TNMHC					1	0.123	-0.441	-0.468	-0.355
RH						1	-0.608	-0.226	-0.488
SR							1	0.311	0.724
WS								1	0.240
AT									1

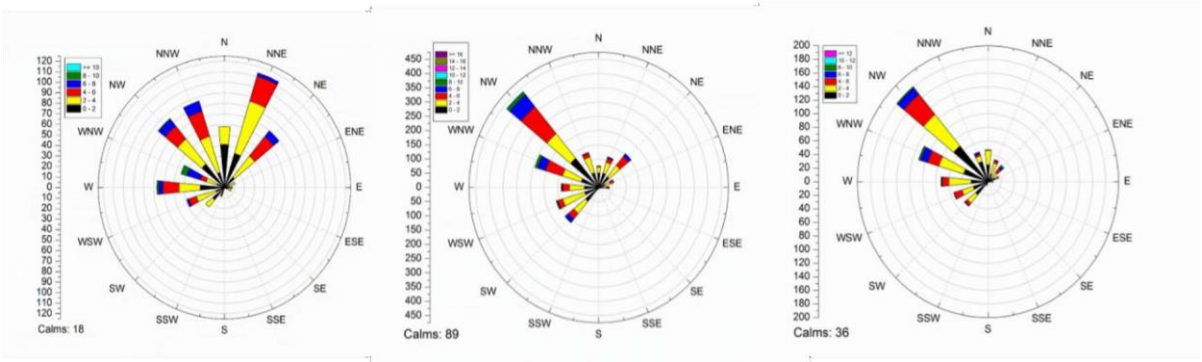
*O_3 (Ozone), CO (Carbon-monoxide), NO_X (Oxides of nitrogen), MHC (Methane hydrocarbon), TNMHCs (Total non-methane hydrocarbons), RH (Relative humidity), SR (Solar radiation), WS (Wind speed) and AT (Ambient air temperature)
**Correlation is significant at the p-value < 0.01 level

5.5 Case study: Ozone levels during crop residue burning (CRB) episodes at a semi-urban site

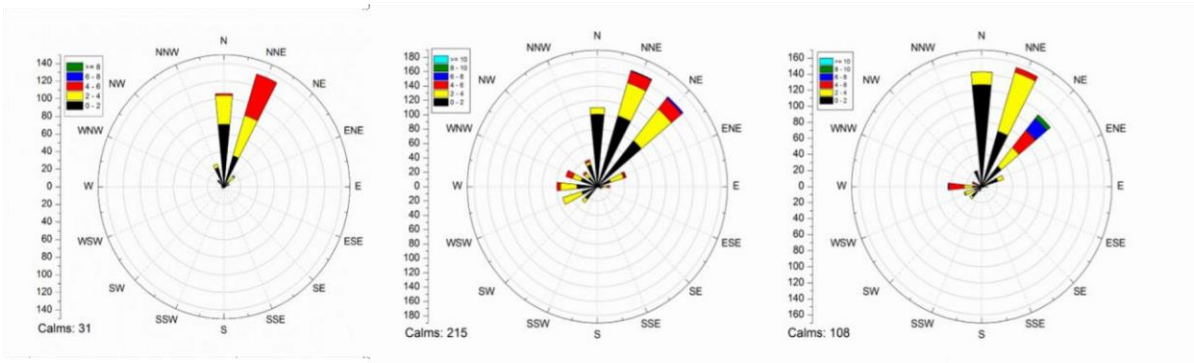
A preliminary analysis of the plots (Figure 5.1 and Figure 5.2) indicate that ozone levels are a reflection of the trends of CO, NO_x, MHC, and TNMHCs besides being more susceptible to crop residue burning (CRB) activities. Hence, further discussion is confined to the ozone levels during crop residue burning (CRB) episodes at a semi-urban site for the period of January 2013 to December 2016. With a view to understanding variations in ozone levels, especially during the episodic CRB event. Pre, during and post biomass burning period have been depicted over the study period using Terra MODIS true-color composites shown in Appendix-II.

The wind rose diagrams for the corresponding periods that is, pre-, during- and post-CRB events have also been plotted (Figure 5.6) and taken into consideration while

analyzing the data. These wind plot diagrams represent the wind direction with a frequency of wind blowing from along with the wind speed based on hourly wind data. The frequency of the time for which wind blows from a particular direction is represented by the length of each "spoke". Each concentric circle represents a different frequency, starting from zero at the center to increasing frequencies at the outer circles. It is observed that most prevalent wind direction at the site is from the NW in wheat and NNE and NE in rice period. Almost similar fetch regions were observed in pre, during and post-harvesting periods of wheat and rice, respectively. Calm conditions were observed more often in during and post period of rice crop residue burning (RCRB) whereas high wind speeds (2-8 m/sec) were observed in the wheat period. The pre-harvesting periods can be considered as "baseline periods" for evaluating the impacts of crop residue burning on ambient ozone concentration levels.



(a) PRE-WCRB **(b) DURING-WCRB** **(c) POST-WCRB**



(d) PRE-RCRB **(e) DURING-RCRB** **(f) POST-RCRB**

Figure 5.6: Wind rose plots (a-f) showing wind direction, wind speed and wind frequency in the study area for pre, during and post-crop residue burning period in 2014 and 2015.

As shown in Figures 5.7 and 5.8, the levels of primary gaseous pollutants (CO, NO_x, MHC, and TNMHCs) are almost double during rice crop residue burning (RCRB) period than in the wheat crop residue burning period (WCRB). The probable reasons are, complete rice stubble as it is not used for animal feed while wheat stubble is stocked for animal feed, combustion process occurred at low intensity of solar radiation and calm wind conditions as compared to WCRB which influence the chemical reactions of pollutants and their dispersion. [10-12]

In the WCRB period, the primary gaseous pollutants (CO, NO_x, MHC, and TNMHCs) got consumed in photo-oxidation reaction for the production of a secondary pollutant, ozone under favorable meteorological conditions. Figures 5.7 (a-e) and 5.8 (a-e) shows pre-, during- and post-CRB events for wheat and rice, respectively for the study period. The trends in variations in pre-, during- and post-CRB events are more or less the same for wheat and rice crops. The concentration levels of gaseous pollutants were are significantly higher during harvesting season than the pre and post CRB period.

Ozone exceedance of 24 hour national standard occurred on 50% and 8% of the sampling periods during- wheat and rice crop residue burning, respectively. This episodic event influences the gaseous pollutants levels in a sustained manner and influence positively for a period much beyond one month as the levels don't come back to the pre CRB period levels. In case of rice period, low levels of ozone were monitored during-RCRB period as compared to wheat, despite profound emissions of precursors gases from CRB, due to low temperature and solar radiation in the post-monsoon season. As seen from the wind rose diagram, the fetch region was distributed from SW to NW in WCRB and N to NE in RCRB. In conclusion, we can safely interpret that the remarkable increase in pollution levels was done due to episodic CRB event and fetch regions were remained almost same throughout the study of pre-, during- and post- CRB, neutralizing the long-range transport of pollutants.

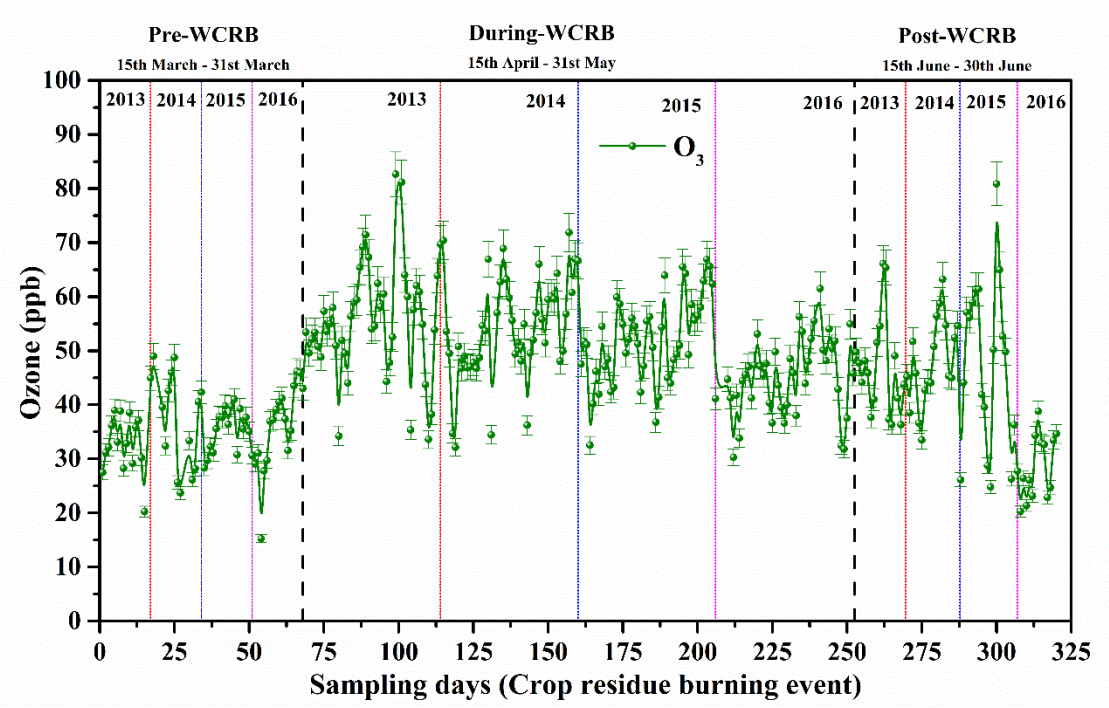


Figure 5.7 (a): Ozone levels (24-hour average) during pre-, during- and post-crop residue burning of wheat during the study period

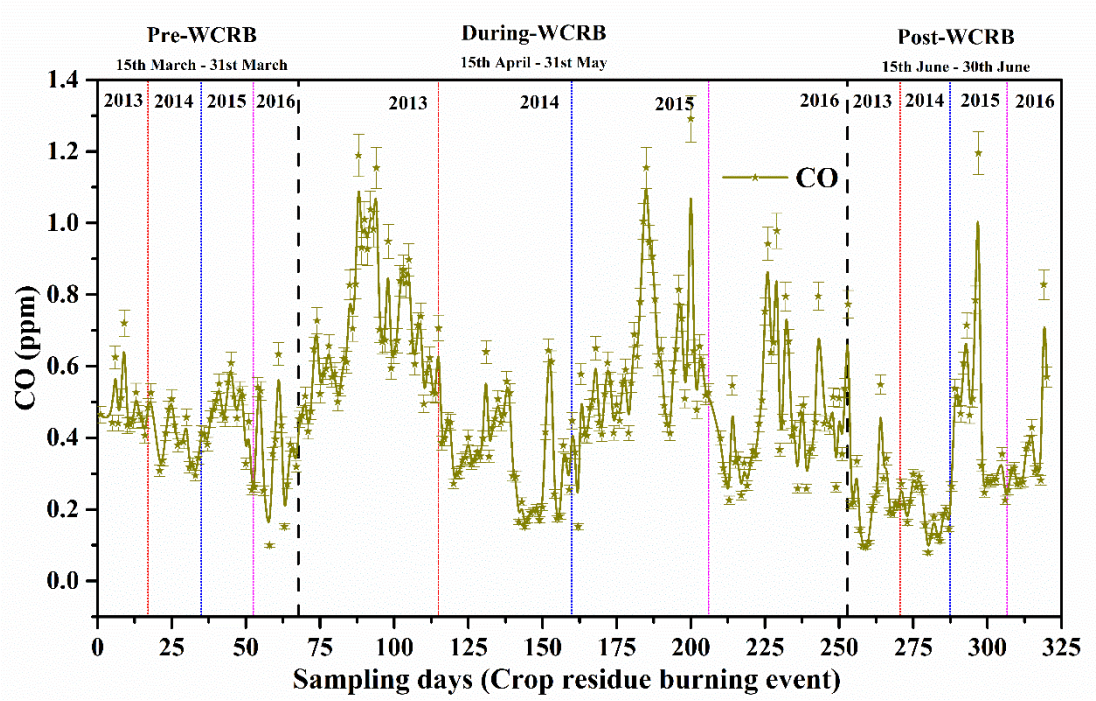


Figure 5.7 (b): CO levels (24-hour average) during pre-, during- and post-crop residue burning of wheat during the study period

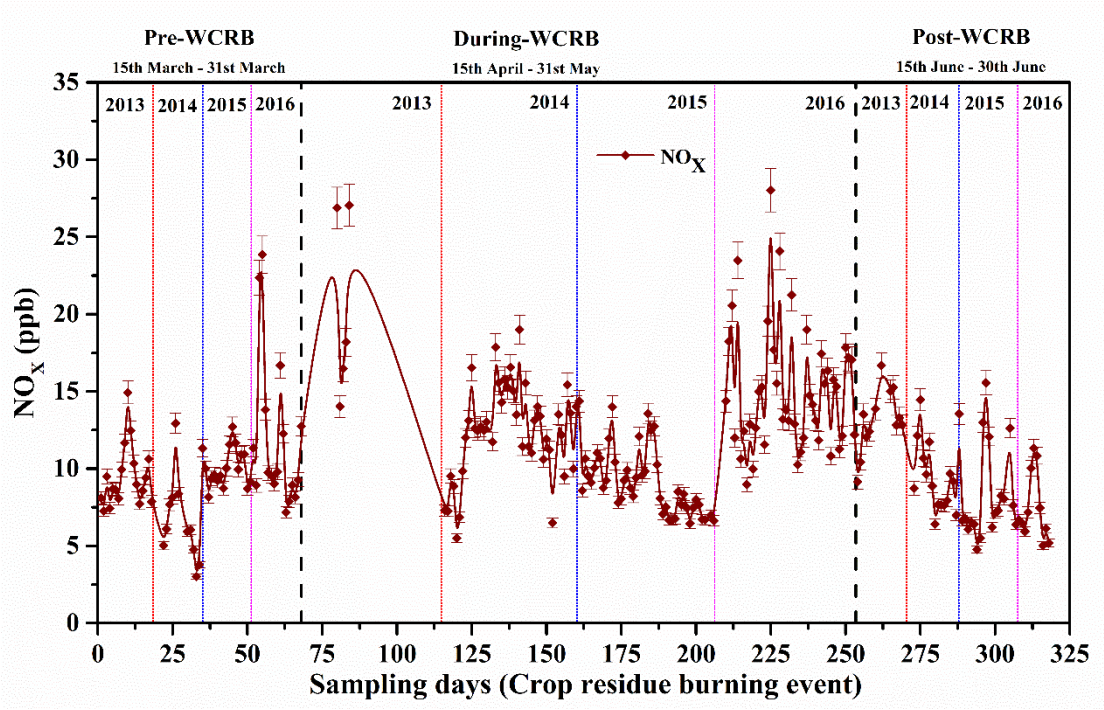


Figure 5.7 (c): NO_x levels (24-hour average) during pre-, during- and post-crop residue burning of wheat during the study period

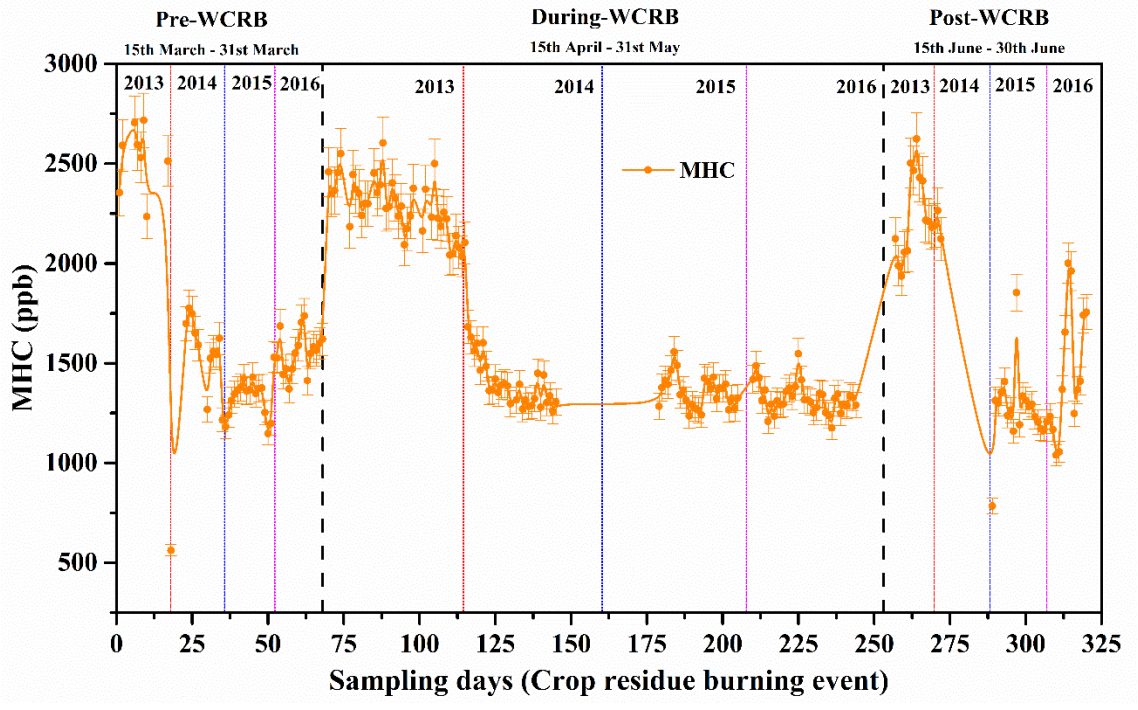


Figure 5.7 (d): MHC levels (24-hour average) during pre-, during- and post-crop residue burning of wheat during the study period

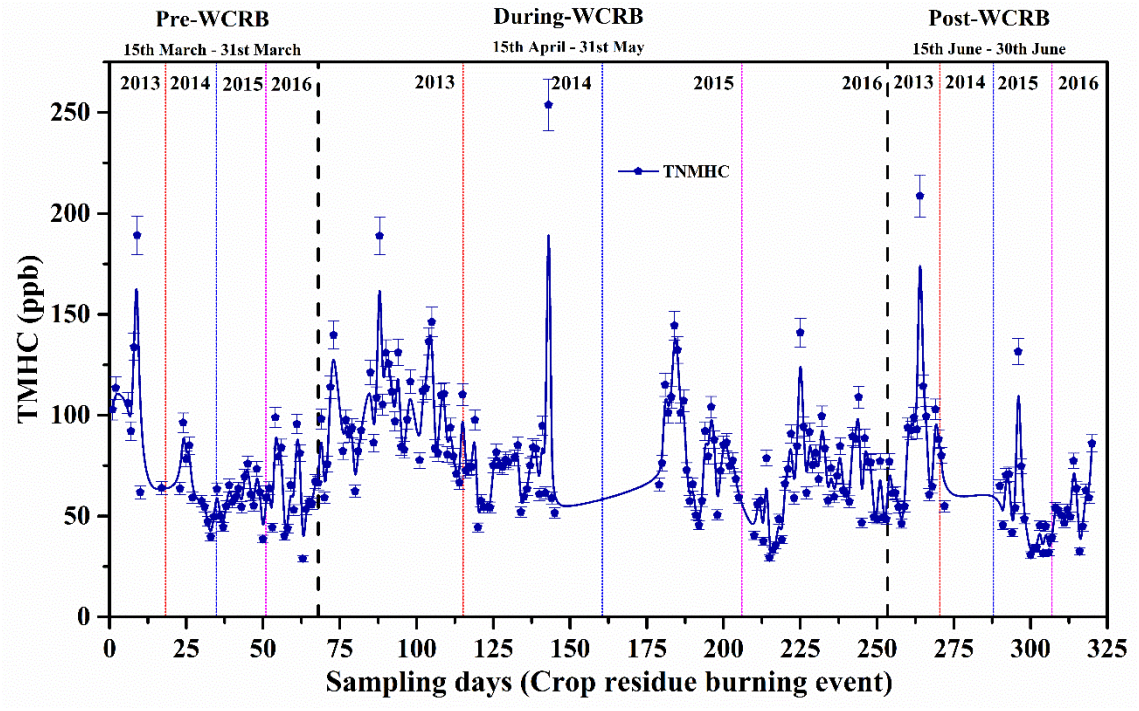


Figure 5.7 (e): TNMHCs levels (24-hour average) during pre-, during- and post-crop residue burning of wheat during the study period

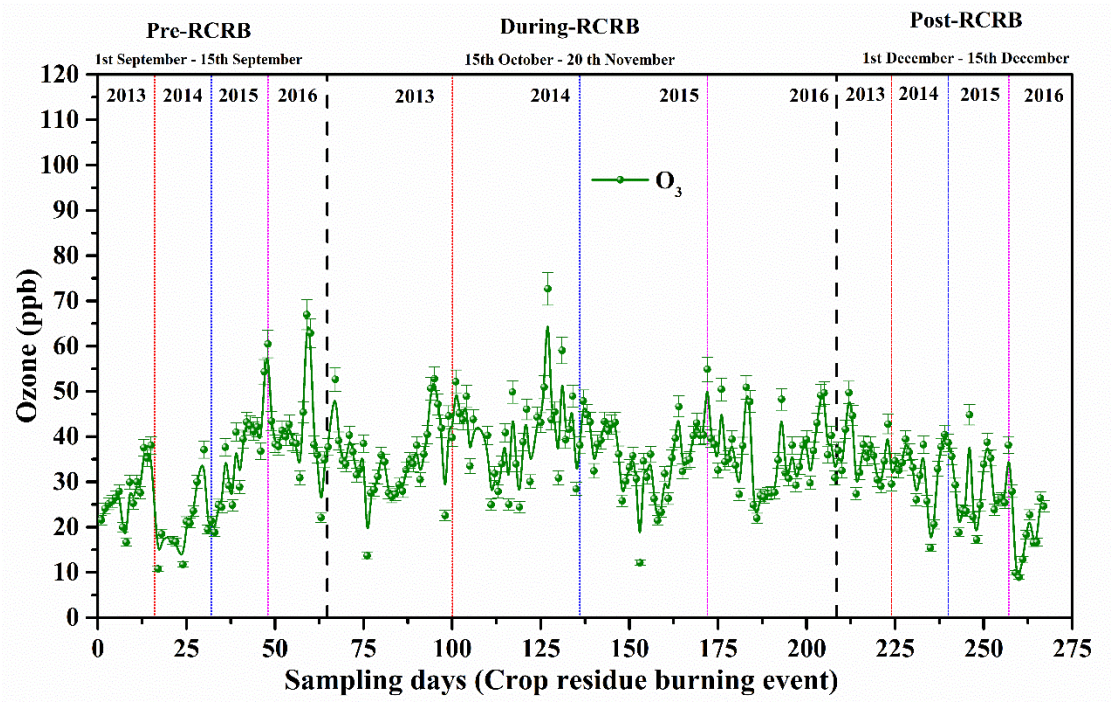


Figure 5.8 (a): Ozone levels (24-hour average) during pre-, during- and post-crop residue burning of rice during the study period

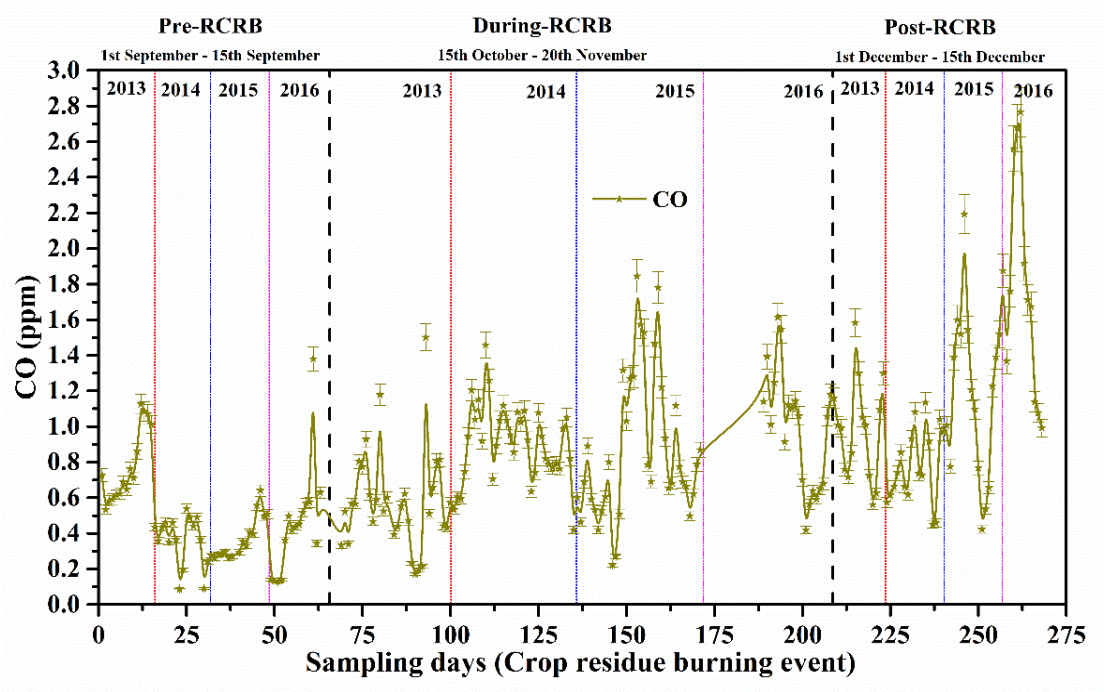


Figure 5.8 (b): CO levels (24-hour average) during pre-, during- and post-crop residue burning of rice during the study period

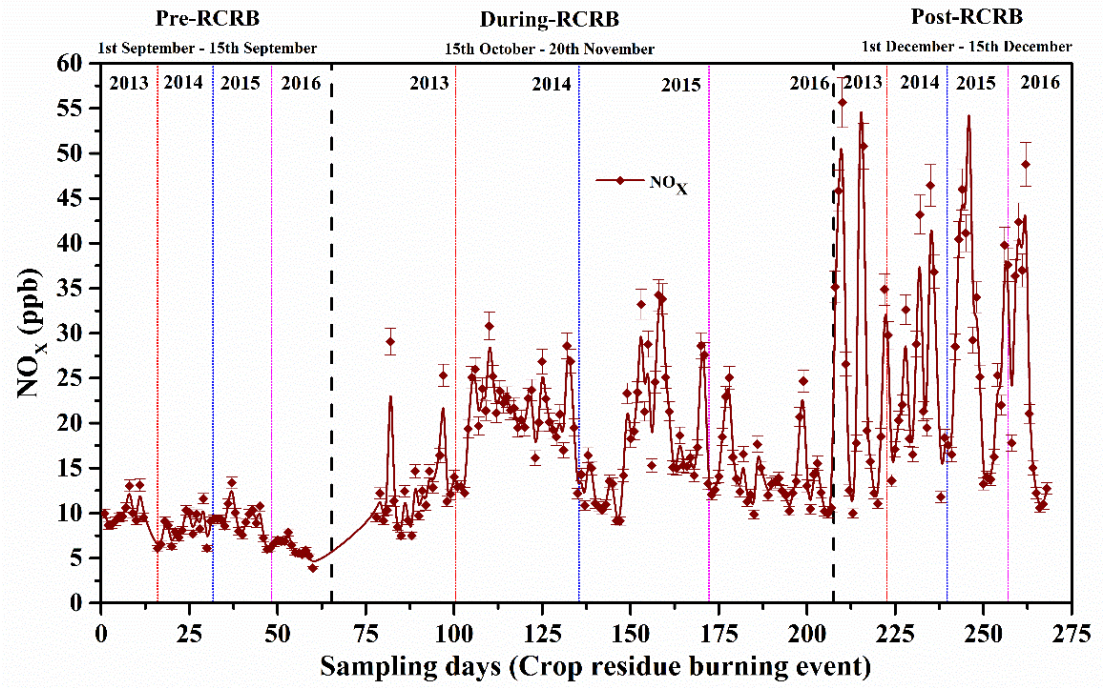


Figure 5.8 (c): NO_x levels (24-hour average) during pre-, during- and post-crop residue burning of rice during the study period

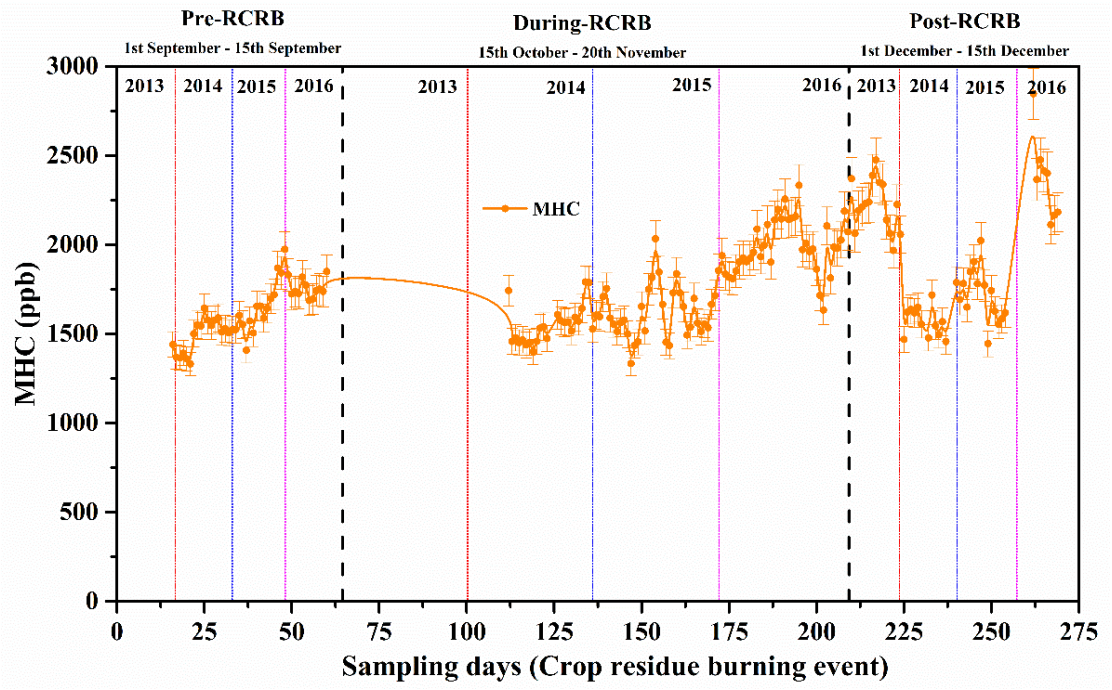


Figure 5.8 (d): MHC levels (24-hour average) during pre-, during- and post-crop residue burning of rice during the study period

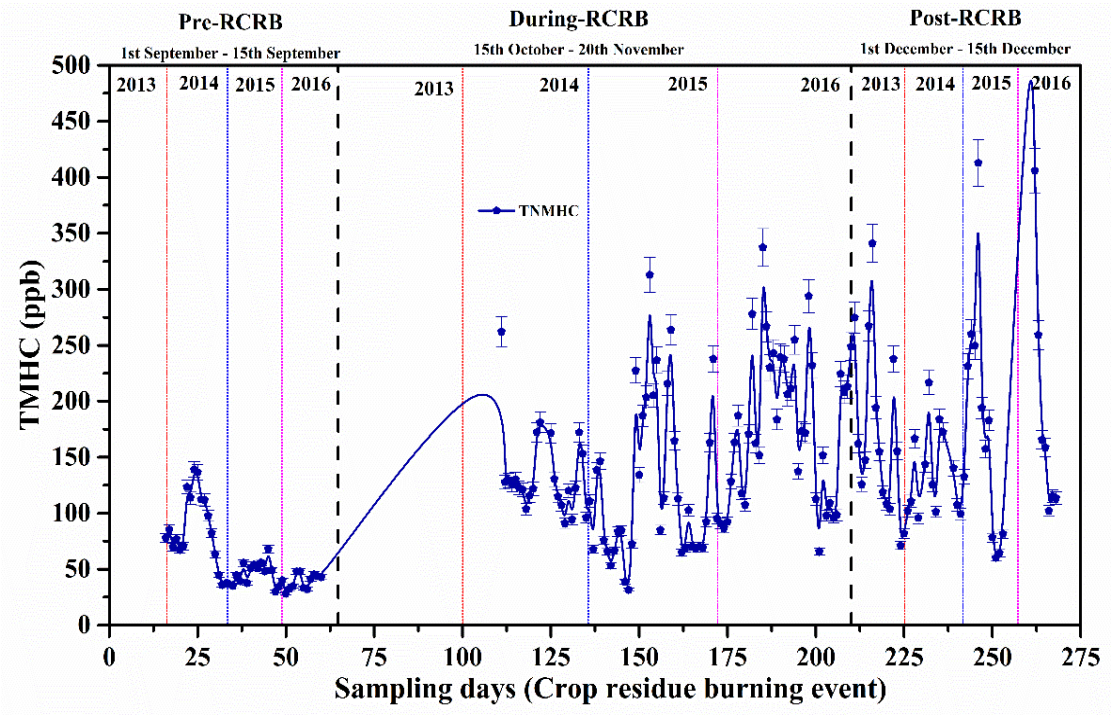


Figure 5.8 (e): TNMHCs levels (24-hour average) during pre-, during- and post-crop residue burning of rice during the study period

5.6 Backward trajectory, Potential source contribution function (PSCF) and Concentration weighted trajectory (CWT) analysis

The backward trajectory plot showed that the realistic motion of air parcels and their transport pathways was calculated at 500 m above ground level (AGL) for the sampling duration and have been presented in Figure 5.9 (a). These groups of trajectories have been clustered together into distinct groups (clusters) based on the transport speed and direction, eventually yielding clusters with similar length and curvature to represent the major directions of the air mass offering a better conception of the source regions.[13] Along with air clusters, moderate resolution imaging spectroradiometer (MODIS) satellite observations of fire counts in the proximity of the study area during the study period are shown in Figure 5.9 (b). Out of these four clusters, the major percentage contribution was from cluster 1 (56%), followed by cluster 3 (17%), as shown in Figure 5.9 (b). Hence, the major influx of gaseous was from regional continental landmasses to reach the receptor site and was not significantly influenced by long-range transport.

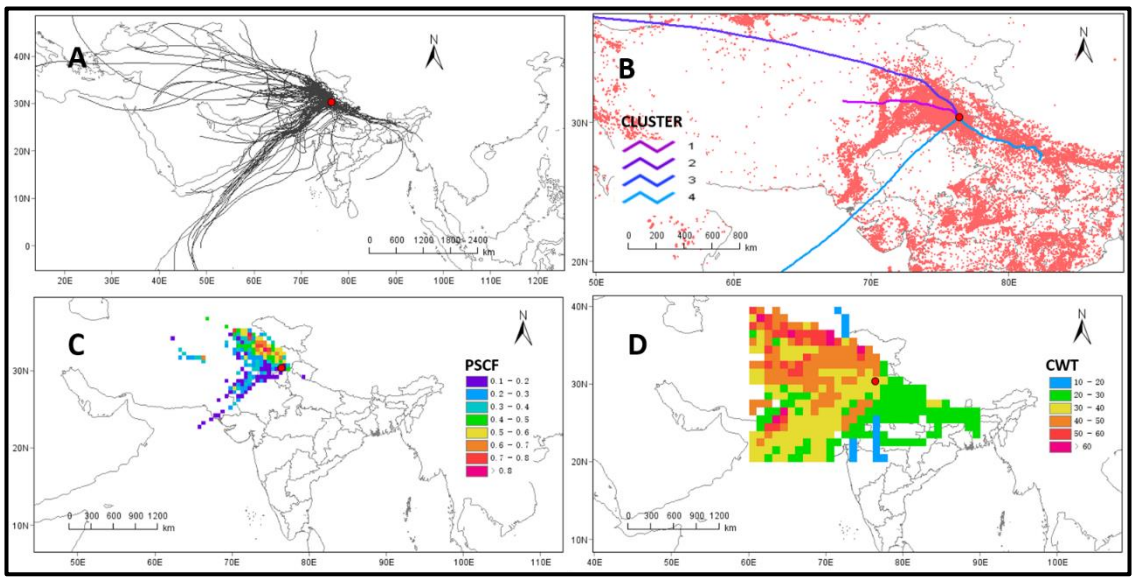


Figure 5.9 (a): Backward trajectories at 500m level over Patiala of the study period (January 2013 - December 2016)
(b): Clustered Trajectory Analysis and MODIS fire counts
(c): PSCF analysis of ozone during the study period
(d): CWT analysis of surface ozone during the study period

Furthermore, potential source contribution function (PSCF) and concentration weighted trajectory (CWT) analysis was performed to highlight the source potentials of the studied geographical region and also to allocate their discrete contribution to the measured concentrations at the receptor site.[14] It was done using trajectories

statistics software, TrajStat (version 1.2.2.6). Figures 5.9 (c) and 5.9 (d) show the PSCF and CWT analyses for ozone. The probable source grids span the densely populated and polluted north-west regions, including parts of India as Punjab, Haryana and Rajasthan and parts of Pakistan which contain key oil fields, refineries and major thermal power plants.[9] Moreover, the enormous amounts of the burning of agricultural crop residues in these fertile farmlands is a critical issue [12] that has also been adjudged a contributing factor to the high values of gaseous and particulate emissions.[15] Thus, the air masses flowing from these regions are bound to bring pollutant-laden gases to the receptor site and in due course generate secondary pollutants such as ozone.

5.7 References

1. Saliba, M.; Ellul, R.; Camilleri, L.; Güsten, H. *Journal of Atmospheric Chemistry* **2008**, *60*, 117-135.
2. Yadav, R.; Sahu, L.; Jaaffrey, S.; Beig, G. *Journal of Atmospheric Chemistry* **2014**, *71*, 125-144.
3. Ahammed, Y. N.; Reddy, R.; Gopal, K. R.; Narasimhulu, K.; Basha, D. B.; Reddy, L. S. S.; Rao, T. *Atmospheric Research* **2006**, *80*, 151-164.
4. Singla, V.; Satsangi, A.; Pachauri, T.; Lakhani, A.; Kumari, K. M. *Atmospheric Research* **2011**, *101*, 373-385.
5. Gaur, A.; Tripathi, S.; Kanawade, V.; Tare, V.; Shukla, S. *Journal of Atmospheric Chemistry* **2014**, *71*, 283-301.
6. Yadav, R.; Sahu, L.; Beig, G.; Jaaffrey, S. *Atmospheric Research* **2016**, *176*, 96-107.
7. Saini, R.; Singh, P.; Awasthi, B. B.; Kumar, K.; Taneja, A. *Atmospheric Pollution Research* **2014**, *5*, 796-804.
8. Londhe, A.; Jadhav, D.; Buchunde, P.; Kartha, M. *Current Science* **2008**, 1724-1729.
9. Sharma, A.; Sharma, S. K.; Mandal, T. K. *Sustainable Environment Research* **2016**, *26*, 76-83.
10. Mittal, S. K.; Singh, N.; Agarwal, R.; Awasthi, A.; Gupta, P. K. *Atmospheric Environment* **2009**, *43*, 238-244.
11. Badarinath, K.; Chand, T.; Prasad, V. K. *Current Science (00113891)* **2006**, *91*.
12. Jain, N.; Bhatia, A.; Pathak, H. *Aerosol and Air Quality Research* **2014**, *14*, 422-430.
13. Li, K.; Chen, L.; Ying, F.; White, S. J.; Jang, C.; Wu, X.; Gao, X.; Hong, S.; Shen, J.; Azzi, M. *Atmospheric Research* **2017**, *196*, 40-52.
14. Bhuyan, P. K.; Bharali, C.; Pathak, B.; Kalita, G. *Environmental Science and Pollution Research* **2014**, *21*, 6696-6713.
15. Badarinath, K.; Kharol, S. K.; Sharma, A. R.; Prasad, V. K. *Journal of Atmospheric and Solar-Terrestrial Physics* **2009**, *71*, 1267-1276.

CHAPTER -6

In this study, the prediction model has been developed for daily ozone levels as a function of precursor gases and meteorological parameters as predictor variables. The prevailing work is probably the primary work within the literature that deals with the day by day ozone levels and their predictions within the Patiala area of N.W IGP, in which continuous ambient air quality monitoring has currently started out. The relationship between ozone, its precursor gases, and meteorological variables has been documented by many authors, and application of a wide range of statistical methodologies were applied to develop satisfactory predictions of ozone levels. The growing diversity of statistical literature in the last decade, along with the use of widely differing input dataset and performance indices were reported in Table 6.1 provides a summary and critical evaluation of the literature. The development of statistical models involved the detailed choice of both site-specific predictor variables (pollutant and meteorological variables) along with the interactions between predictor variables in order to better capture the target variable (ozone) behavior.

The main objective of this work is to develop a prediction model for daily ozone for the improvement of progressed statistical methods using an ensemble approach. The assessment of the significant interactions among predictor variables was observed before developing linear and non-linear and ensemble models. Air quality forecasting is very consistent and effective in controlling the measures and can be proposed as a preventive action for regulations and protect public health[1,2]

6.1 Feature importance and selection of variables for the determination of input variables and training sets for target variable (ozone)

In machine learning and statistics, the process of selection of relevant features or predictors for use in model construction is very important. When using feature selection methods, the main assumption is that the data includes some characteristics that are either redundant or meaningless and can, therefore, be deleted without much information loss. In this study, the predictor-selection of input variables was in details explored deliberating the predictability of input variables on the overall performance of the prediction model. First, based on reported literature and secondly model-based.

Table 6.1: List of statistical models for prediction of tropospheric levels of ozone at a single site

Location	Latitude-Longitude (degree)	Statistical methods	Model Input Variables	Performance indices	Reference
Patiala, India	30.35 ºN, 76.37 ºE	RF, Treebag, Nnet, LM, LeapBkwd, GLM, ICR, Kknn, Ksvm, BayesGLM, Relaxo, Ppr, Pls, Lars, EM-1, EM-2, EM-3, EM-4	CO, NO_X, MHC, TNMHCs, RH, SR, AT, WS and WD	r, RMSE, FV, Accuracy, and d	(This study)
Oporto, Northern Portugal	41.10 °N, 8.40 °W	MLR, Feed-forward NN	NO, NO_2, T, RH, and WV	r, MBE, MAE, RMSE, d.	[3]
Kuwait	28-30 °N, 46- 48 °E	PCR, ANN, Combined PCR-ANN	CH_4, TNMHCs, CO, CO_2, NO, NO_2, SO_2, WS WD, AT, RH, SR	R^2, RMSE	[4]
Isfahan, Iran	32.37 °N, 51.40 °E	MLR	RH, T_{wet}, T_{dry}, T_{max}, T_{min}, T_{ground}, TSSH, and O3	R^2	[5]

Bahadurshah Zafar Marg (ITO), Delhi, India	28.62 °N, 77.24 °E	SVM, MLP, MR	Previous day O_3, T_{max}, Daily Predominant WS, RH, and TSSH	MAPE, NRMSE, RMSE, d	6
Pantnagar, Tarai belt of Himalayas	29.01 °N, 79.28 °E	MLR	SO_2, NO_2, TSPM, WS, RH, AT, and P	R^2	7
Annaba, Algeria	36.54 °N, 7.46°E	RBF-NN, MLP-NN ANN-based model, linear model, hybrid model	NO, CO, O_3, PM_{10}, NOx, NO_2 and SO_2, WS, AT, and RH	d, MSE	8
Tehran, Iran	35.68 °N, 51.38°E	NN	NOx, NO_2, NO, O_3, CO, PM_{10}, WS, T, RH, and WD	MAE, RMSE, d, r	9
Thamrin, Jakarta	6.11°S, 106.49 °E	MLR	CH_4, CO, NMHC, NO, NO_2, THC, WD, WS, T, SR, RH	R^2, r	10

*1 RF(Random Forest), Treebag(Bagged CART), Nnet (Neural network), LM (Linear Regression), LeapBkwd (Linear Regression with backward selection), GLM (Generalized Linear Model), ICR (Independent component Regression), Kknn(k-Nearest neighbor), Ksvm (Support Vector Machines), BayesGLM (Bayesian Generalized Linear Model), Relaxo (Relaxed Lasso), Ppr (Projection pursuit regression), Pls (Partial least squares), Lars (Least angle regression), EM-1(Ensemble Model-1), EM-2 (Ensemble Model-2), EM-3 (Ensemble Model-3), EM-4 (Ensemble Model-4), CO (Carbon-monoxide), NO_X (Oxides of nitrogen), MHC (Methane hydrocarbon), TNMHCs (Total non-methane hydrocarbons), RH (Relative humidity), SR (Solar radiation),AT (Air temperature), WS (Wind speed) and WD (Wind direction), r (Correlation coefficient), RMSE (Root mean square error), FV (Fractional variance), Accuracy and d (Index of agreement)

*3 PCR (Principle component regression) , ANN (Artificial neural network), CH_4 (Methane hydrocarbon), CO_2 (Carbon-dioxide), NO (Nitrogen monoxide), NO_2 (Nitrogen dioxide), SO_2 (Sulphur dioxide)

*4 T_{wet} (wet temperature), T_{dry} (dry temperature), T_{max} (maximum temperature),T_{min} (minimum temperature), T_{ground} (ground temperature)

*5 SVM(Support vector machines), MLP (Multilayer perceptron- neural network),MR (Multiple regression), TSSH (Total sunshine hour), MAPE (mean absolute percentage error), NRMSE (normalized root mean square error);

*7 RBF-NN (Radial basis function- neural network), MLP-NN (Multilayer perceptron- neural network), O_3 (Ozone), PM_{10} (Particulate matter with diameter $\leq$10), (MSE (Mean square error)
*8 MAE (Mean absolute error)
*9 MLR (Multiple linear regression), THC (Total hydrocarbon carbon)

*8 MAE (Mean absolute error)

techniques that rely on training with different combinations of input variables.[11,12] Random forest and backward stepwise regression methods are used for the importance and selection of the variables using R packages. Random forest model results highlighted the importance of the variables according to their weightage in term of their rank as shown in Table 6.2. It is found that among meteorological variables relative humidity and in precursor gases NO_x has the highest impact on the production of ozone.

Table 6.2: Feature importance of the precursor gases and meteorological variables for the development of the prediction model for ozone

RANK	Variable Importance
1	RH
2	NO_x
4	MHC
5	WD
6	WS
7	AT
8	CO
9	SR
10	TNMHCs

*O_3 (Ozone), CO (Carbon-monoxide), NO_X (Oxides of nitrogen), MHC (Methane hydrocarbon), TNMHCs (Total non-methane hydrocarbons), RH (Relative humidity), SR (Solar radiation), WS (Wind speed), WD (Wind direction) and AT (Ambient air temperature)

Backward stepwise regression methods begin with all available features for the model along with performance indices such as correlation coefficient and root mean square error (RMSE). The variable that is the least important is discarded at each step. This method goes on until there are no non-significant variables. The user establishes the level of meaning at which variables from the model can be deleted to obtain the significant dependence between input variables prior to developing the statistical models.

Variance inflation factor (VIF) is an index that measures the increased variance of estimated regression coefficients due to collinearity with the value of 2.94. The value of VIF 10 and above, then the multicollinearity is problematic. They all provide the significant dependence of ozone with CO, NO_x, MHC, TNMHCs, AT, SR, RH, and WS which highlight the significant relevance of each variable on the model performance. So, the input data set comprise the daily average value of above-mentioned pollution and meteorological variables for the development of prediction models to minimize the prediction uncertainty.

Table 6.3: Backward stepwise regression for the selection of features (or variables) for the development of the prediction model for surface ozone

Features									Performance parameters	
									r	**RMSE**
RH	NOx	MHC	WD	WS	AT	CO	SR	NMHC	0.81	10.09
RH	NOx	MHC	WD	WS	AT	CO	SR		0.81	10.07
RH	NOx	MHC	WD	WS	AT	CO			0.80	10.33
RH	NOx	MHC	WD	WS	AT				0.79	10.5
RH	NOx	MHC	WD	WS					0.79	10.42
RH	NOx	MHC	WD						0.78	10.76
RH	NOx	MHC							0.78	10.85
RH	NOx								0.76	11.02
RH	NOx								0.76	11
RH									0.76	11.35

6.2 Techniques applied in modelling and optimization of models – The Ensemble Approach

Air quality prediction can be carried out using various traditional and soft computing techniques. The prediction algorithms used in this research work include fourteen techniques (Random forest, tree bag, neural network, linear regression, linear regression with backward selection, independent component regression, least angle regression, generalized linear model, k- nearest neighbors, kernel support vector machine, partial least square, projection pursuit regression, Bayesian generalized linear model, least angle regression) from different families. Table 3.1 contains names of the family to which a method belongs and its tuning parameters detailed in material and methodology in section 3.7. The flow chart of the methodology of work is shown in Figure 6.1. The specific descriptions of the applied models are to be had on caret R package.[13] The input data set of pollutant and meteorological variables from (2013-2015) was used to build the prediction models. The distribution of input data set into the training-testing experiment is optimized to 70% and 30%, respectively for all the strategies to evaluate the forecasting performance of models. For model training, 10,000 iterations were used to decide model parameters and the test set is used to guarantee the generalization capability of the model.

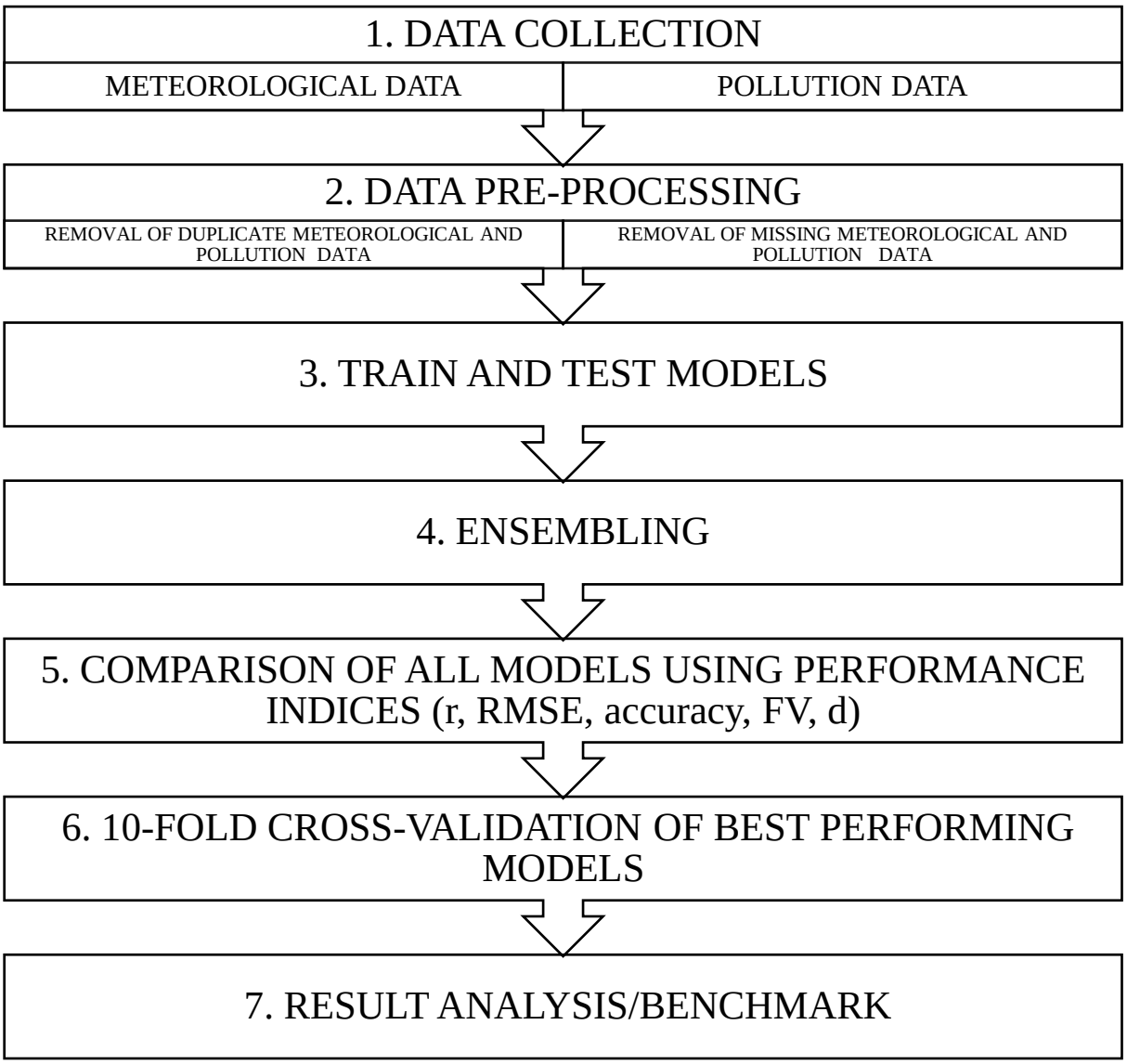


Figure 6.1: The flow chart of the methodology of the work

In this work, ensemble/combination approach was proposed, in which linear and non-linear models are used at the same time, aiming to acquire the dynamic physical and chemical processes involved in photochemical reactions between ozone, meteorological variables and/or precursor emissions. It takes advantage of the individual strength of both linear and nonlinear models and a good strategy for practical use to improve forecasting performance on real data set over the individual models.[14] Further, ensemble models were generated with the possible combination of fourteen models. Four combinations of ensemble models (EM-1, EM-2, and EM-3, EM-4) were selected on the basis of performance indices from the group of ensemble models. In order to have a comprehensive idea of the potential of each individual and four ensemble models, their performances were

assessed by calculating five performance indices (r, RMSE, accuracy, FV, d). Results of performance indices are shown in Table 6.4 and also presented in Figures 6.2 (a) and 6.2 (b).

Table 6.4: Performance comparison of machine learning methods and their ensembles in the prediction of daily ozone levels on training/testing data set with performance indices (correlation coefficient (r), fractional variance (FV), index of agreement (d); root mean square error (RMSE)

S.No.	Models	R	FV	D	Accuracy	RMSE
1	RF	0.82	0.59	0.82	90.64	9.55
2	Treebag	0.80	0.52	0.82	90.08	9.92
3	Nnet	0.78	0.44	0.77	87.75	10.41
4	LM	0.79	0.46	0.78	87.94	10.28
5	LeapBkwd	0.70	0.41	0.65	79.35	9.93
6	GLM	0.79	0.46	0.78	87.94	10.28
7	ICR	0.70	0.41	0.65	79.35	13.38
8	Kknn	0.75	0.25	0.78	84.95	11.19
9	Ksvm	0.74	0.33	0.78	84.00	11.53
10	BayesGLM	0.79	0.46	0.78	87.94	10.28
11	Relaxo	0.79	0.51	0.78	88.23	10.14
12	Ppr	0.79	0.46	0.78	87.94	10.28
13	Pls	0.54	1.18	0.51	74.75	15.18
14	Lars	0.79	0.46	0.78	87.94	10.28
15	EM-1	0.85	0.50	0.86	95.85	10.00
16	EM-2	0.84	0.48	0.85	96.0	10.00
17	EM-3	0.86	0.54	0.86	96.33	10.00
18	EM-4	0.86	0.57	0.86	96.25	10.00

*RF(Random Forest), Treebag (Bagged CART), Nnet (Neural network), LM (Linear Regression), LeapBkwd (Linear Regression with backward selection), GLM (Generalized Linear Model), ICR (Independent component Regression), Kknn (k-Nearest neighbor), Ksvm (Support Vector Machines), BayesGLM (Bayesian Generalized Linear Model), Relaxo (Relaxed Lasso), Ppr (Projection pursuit regression), Pls (Partial least squares), Lars (Least angle regression), EM-1(Ensemble Model-1), EM-2 (Ensemble Model-2), EM-3 (Ensemble Model-3), EM-4 (Ensemble Model-4)

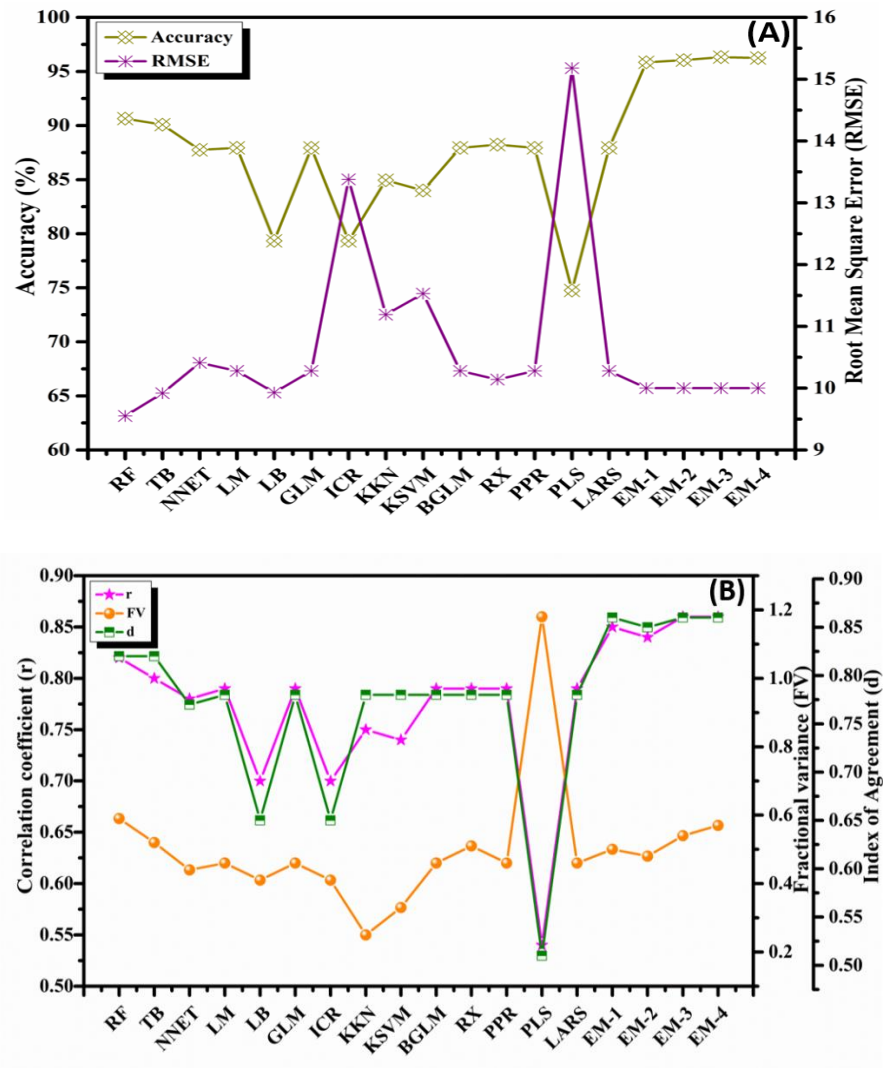


Figure 6.2: Comparison of individual 14 models and four best ensemble models via performance indices (r, RMSE, accuracy, FV, d)

6.3 Performance evaluation of the proposed Ensemble Approach

In terms of accuracy, the ensemble models (EM-1, 96%) provide at least 5% improvement over best performing an individual model (RF, 91%), clearly shown in Table 6.4. The corresponding index of agreement was from 0.85 - 0.86 for ensemble models and 0.51- 0.82 for the individual linear and non-linear models. The correlation coefficient (r) for ensemble models varies between 0.84 - 0.86, whereas for individual models it is limited to 0.54 - 0.82.

The results of the ensemble model revealed that via combining linear and non-linear models inclusively, the overall forecasting variance or error was reduced significantly appreciably as fractional variance and RMSE for ensemble varies between 0.48-0.57 and 10, respectively and for individual models 0.25-1.18 and 9.92-15.18, respectively. The

results clearly indicate that ensemble models outperform significantly over the individual linear or non-linear models. 10-fold cross-validation method shows the robustness of ensemble models on all of the model evaluation parameters shown in Figure 6.3 (a, b, c, and d). The model ensemble (EM-1) outperforms over the three ensemble models (EM-2, EM-3, and EM-4).

Table 6.5: Performance indices for training/testing and benchmark data set for the optimized EM-1ensemble model

Performance parameters (EM-1)	Training and Testing	Benchmark
1. Accuracy	95.85	95.52
2. r	0.85	0.91
3. RMSE	10	6.9
4. FV	0.50	0.33
5. d	0.86	0.92

*Correlation coefficient (r); Root mean square error (RMSE); Fractional variance (FV); Index of agreement (d)

Further, benchmark analysis was done on the optimized EM-1 model using 2016 data set. The ozone data for the months of July and August 2016 is not reported due to a technical error in the analyzer. The results reveal a close index of agreement of 0.91, with an accuracy of 95.5% among the predicted and observed data for tropospheric ozone done on diurnal monthly and daily average plots, as shown in Figures 6.4 and 6.5. The observed and the predicted mean values are 39.1 ppb and 35.1 ppb, respectively with a mean absolute error of -0.001 ppb. The observed and predicted measurement curves undulations are within 95% confidence interval using model ensemble (EM-1) throughout the year (2016). The higher levels of ozone during September and October was observed due to rice crop harvesting season. The emission of ozone precursor was increased by the burning of left crop residue along with slow wind speeds leads to less dispersion of pollutants.

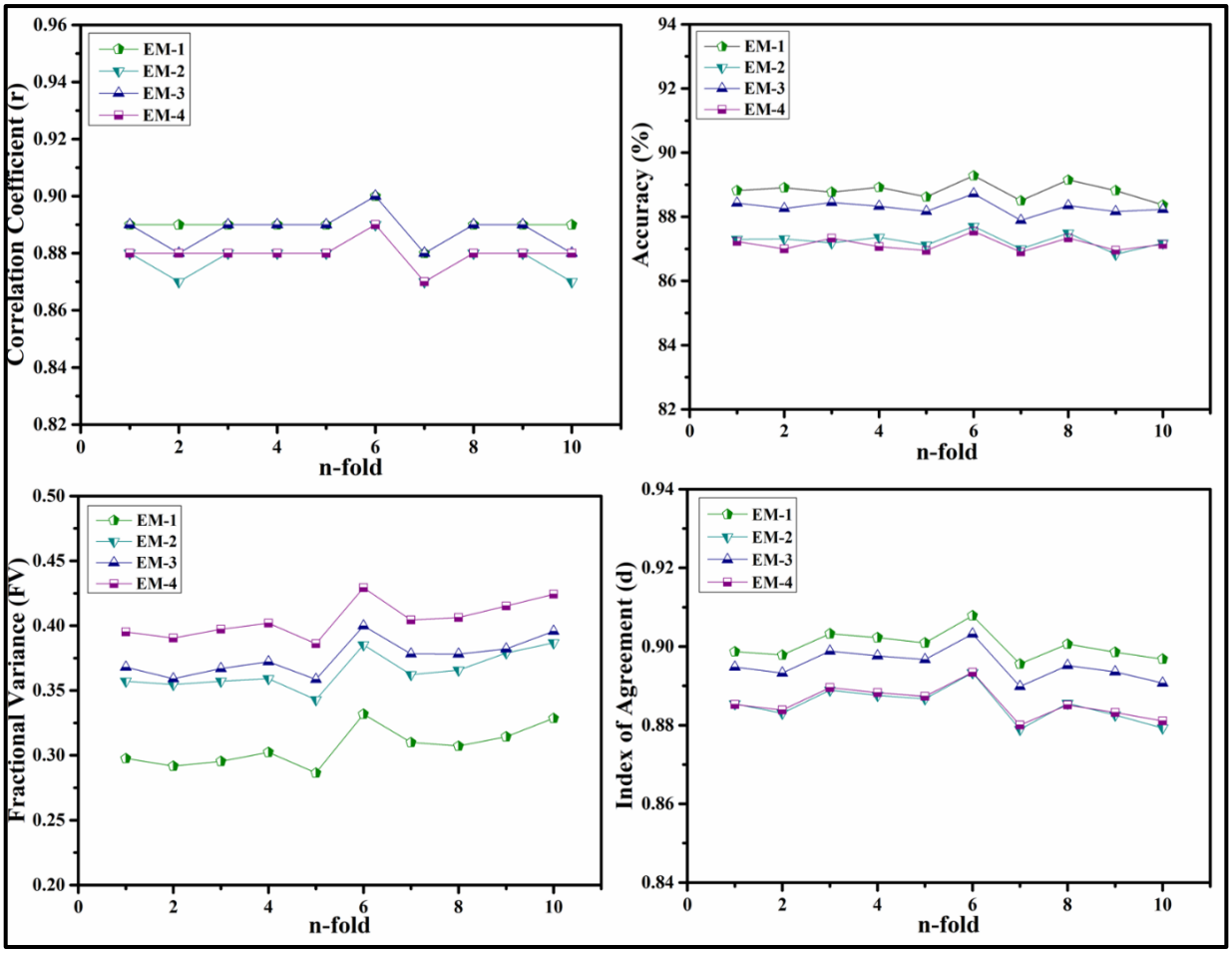
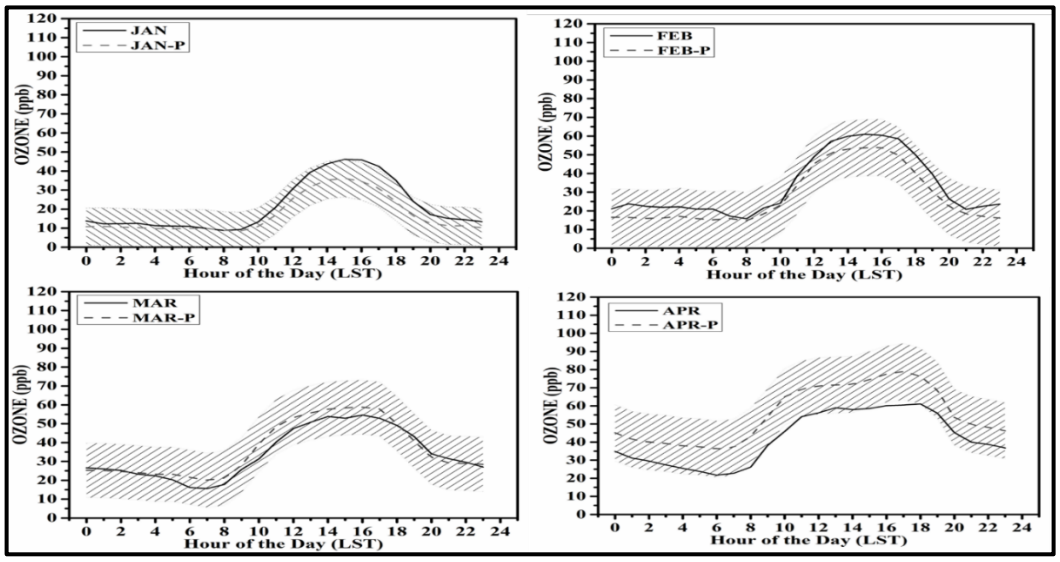
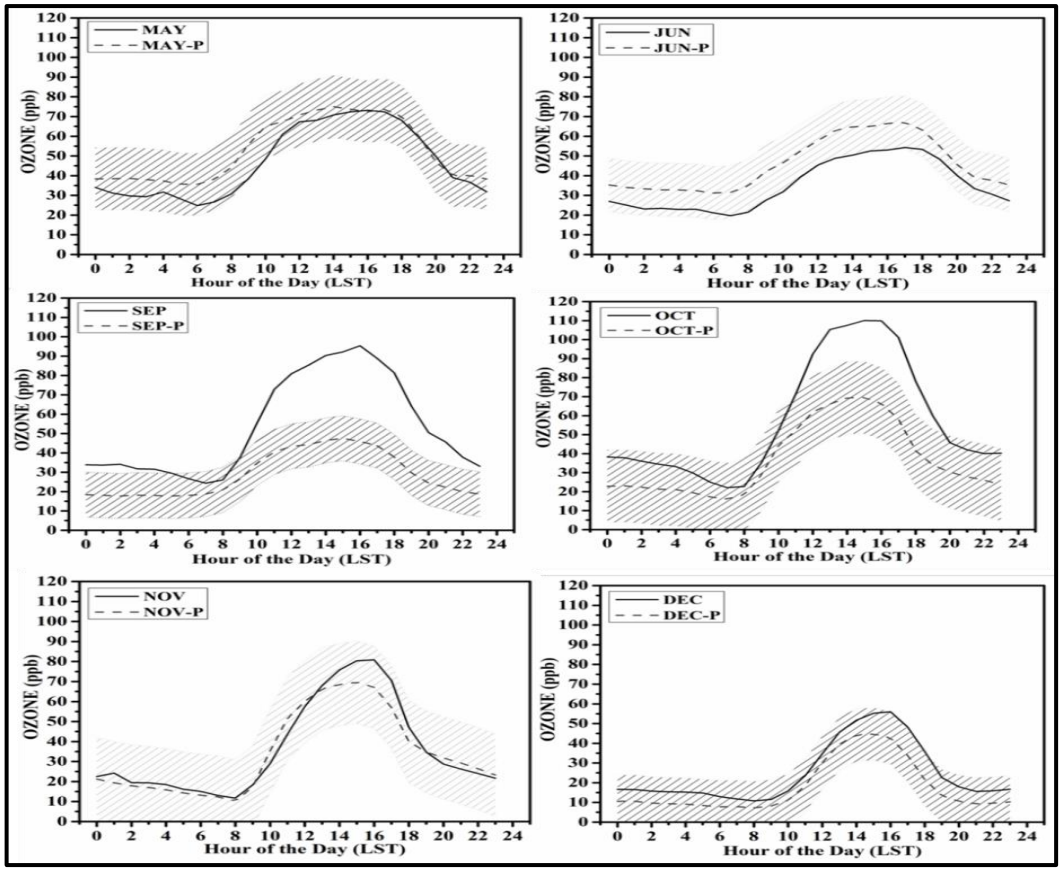


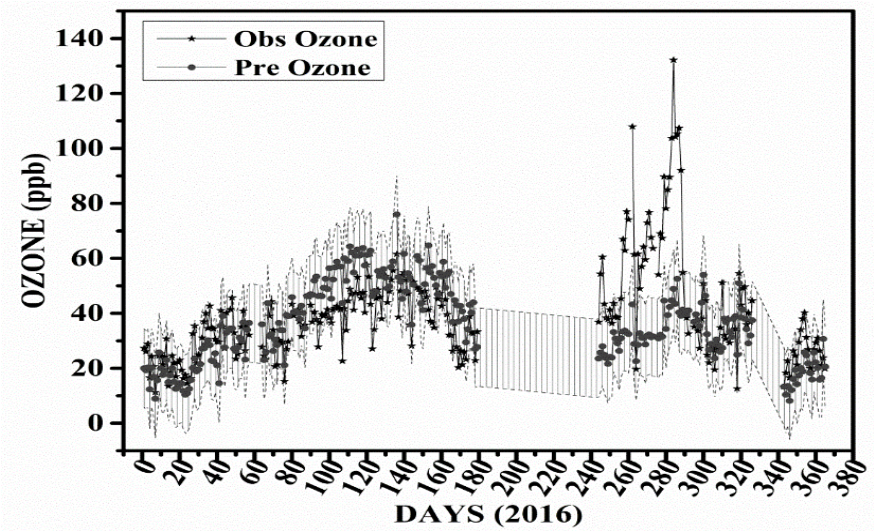
Figure 6.3: 10-fold cross-validation of performance indices (r, accuracy, FV and d) for four best-performing ensemble models (EM-1, EM-2, EM-3, and EM-4)





*****July and August data is not reported due to instrument failure

Figure 6.4: Diurnal cycle month-to-month variation of observed (2016) and predicted concentration (Month-P) using EM-1 model along with shaded area (predicted mean ± 1 standard deviation)



*****Data not reported for the period of (180 – 240 days) due to instrument failure

Figure 6.5: Daily variations of observed (2016) and predicted concentration using EM-1 model along with the shaded area of (predicted mean ± 1 standard deviation)

The results of the optimal ensemble model (EM-1) for the training and benchmark data sets are shown in Table 6.5. The analysis shows that the ensemble models under realistic conditions provide good fits between measured and modeled ozone levels. This is supported by the high value of correlation coefficient (0.91) and index of agreement (0.92), indicating that the EM-1 model is capable of predicting the variability of ozone's concentration. It is clear that predictions via ensemble approach capture the more accurate integral relationship between meteorological variables and precursor pollutants and resulted in improved model prediction performance and reduced complexity. This work is supposed to improve the usage of scientific knowledge via policymakers and regulatory bodies throughout member states by way of integrating assessment for air best plans at the regional and neighborhood scales to arrest the pollutants.

6.4 References

1. Kolehmainen, M.; Martikainen, H.; Ruuskanen, J. *Atmospheric Environment* **2001**, *35*, 815-825.
2. Sansuddin, N.; Ramli, N. A.; Yahaya, A. S.; Yusof, N. F. F. M.; Ghazali, N. A.; Al Madhoun, W. A. *Environmental Monitoring and Assessment* **2011**, *180*, 573-588.
3. Sousa, S.; Martins, F. G.; Alvim-Ferraz, M.; Pereira, M. C. *Environmental Modelling & Software* **2007**, *22*, 97-103.
4. Al-Alawi, S. M.; Abdul-Wahab, S. A.; Bakheit, C. S. *Environmental Modelling & Software* **2008**, *23*, 396-403.
5. Hassanzadeh, S.; Hosseinibalam, F.; Omidvari, M. *Physica A: Statistical Mechanics and its Applications* **2008**, *387*, 2317-2327.
6. Chelani, A. B. *Environmental Monitoring and Assessment* **2010**, *162*, 169-176.
7. Banerjee, T.; Singh, S.; Srivastava, R. *Atmospheric Research* **2011**, *99*, 505-517.
8. Sabri, G.; Tarek, K. M. *International Journal of Environmental Studies* **2012**, *69*, 79-89.
9. Arhami, M.; Kamali, N.; Rajabi, M. M. *Environmental Science and Pollution Research* **2013**, *20*, 4777-4789.
10. Wasi'ah, N. R.; Driejana, D. *IPTEK Journal of Proceedings Series* **2017**, *3*.
11. Glavan, M.; Gradišar, D.; Atanasijević-Kunc, M.; Strmčnik, S.; Mušič, G. *The International Journal of Advanced Manufacturing Technology* **2013**, *68*, 2743-2759.
12. Kocijan, J.; Hančič, M.; Petelin, D.; Božnar, M. Z.; Mlakar, P. *Simulation Modelling Practice and Theory* **2015**, *54*, 101-115.
13. Kuhn, M. *URL\url {https://topepo. github. io/caret/}.[Online* **2018**.
14. Zhang, G. P. *Neurocomputing* **2003**, *50*, 159-175.

Conclusion and Summary

Indo-Gangetic Plain (IGP) is one the world's largest and agriculturally productive and densely populated area has experienced extensive economic growth, urban expansion and development in various sectors, i.e., industry, automobile, infrastructure, agriculture, etc. which resulted in pollution problems in the past few decades. Natural and anthropogenic activities emit aerosols and trace gases such as methane (CH_4), total non-methane (TNMHCs), carbon monoxide (CO) and oxide of nitrogen (NO_x), which impart important role in the production and accumulation of secondary pollutants like peroxy-acetyl nitrate (PAN) and surface ozone. Among the most significant contributors of air pollution like automobiles, industries and domestic emissions, the post-harvest crop residue burning is an episodic contributor to the pollutants level in the region. Under MAPAN (Modelling Atmospheric Pollution and Networking) project, the continuous ambient air quality of trace (O_3, CO, NO_x, MHC, TNMHCs) and meteorological variables (AT, SR, RH, WS, and WD) was monitored for the four year period of 1st January 2013 to 31st December 2016, at a semi-urban site, Patiala in the north-western region of IGP.

On diurnal scale, the peak in ozone levels were coincident with the concentration minima of precursor gases (CO, NO_x, MHC, and TNMHCs). A steep minimum is observed during the early morning around 0600- 0800 hours in all seasons. After 0800 hour, a rise in ozone concentration is observed and hit maxima around in the range of 1500-1600 hours and starts decreasing after late afternoon or evening. For precursor gases (CO, NO_x, MHC, and TNMHCs), the elevated levels during the morning (0700-1000 hours) and evening hours (1700-1900 hours) and lower mixing ratios in the afternoon hours (1200-1600 hours). The elevated levels during morning and evening hours at Patiala coincide with peak traffic hours indicating the roles of the vehicular exhaust. On the other hand, the observations from night till early morning hours show high values due to shallow nocturnal boundary layer depth resisting mixing of local emissions with the free tropospheric air. On the other hand, during after noontime, the higher PBL depth provides a larger mixing region along with consumption of precursor gases in photochemical production of surface ozone. Seasonally averaged ozone concentrations delineated the highest levels in summer season followed by post-monsoon, monsoon and winter season. The mixing ratios of CO, NO_x, MHC, and TNMHCs show largest diurnal peaks in post-monsoon and winter seasons followed by summer and monsoon seasons. Besides, the diurnal profile of each season of gases showed a similar pattern with variations in

amplitude. There is significant negative correlation between precursor gases CO (-0.250), NO_x (-0.262), MHC (-0.151) and TNMHCs (-0.252) and high positive correlation coefficient between SR (0.565), ambient temperature (0.449) and wind speed (0.208). The negative correlation (r = -0.682) between the surface ozone and humidity, as the precursor gases are washed away from the environment, resulting in a decrease in the production of tropospheric ozone concentration.

The pollution index value ($O_{3\ max}$ / $O_{3\ min}$ ratio) of 15.60 indicates the study location has remarkable ozone pollution. The highest rate of increase of dO_3/dt during post-monsoon ranges between 14.3 to 18.4 ppb/hr followed by summer season (10.2 to 13.5 ppb/hr), highlights the clear contribution of emissions from post-harvest crop residue burning periods of wheat and rice, respectively. The National Ambient Air Quality Standards for ozone (60ppb for 24hour) were violated 50% and 8% of the duration during wheat and rice crop residue burning, respectively. Crop residue burning events influenced the ozone levels positively in a sustained manner for a period much beyond the post-harvesting season. Regulation to curb crop residue burning would surely extenuate ozone pollution and bring co-benefit in productivity of crops. The contribution from regional pollution to long-distance aged pollutants was identified using seasonal backward trajectories. Furthermore, potential source contribution function (PSCF) and concentration weighted trajectory (CWT) analysis was performed to highlight the source potentials of the studied geographical region.

The proposed ensemble approach for the selection/grouping of various statistical models has successfully been applied to predict concentration levels of ozone in the surface layer of the atmosphere. The relationship of ozone levels with meteorological and air pollution variables were examined in detail for the study period, aiming to capture the dynamic physical and chemical processes involved in photochemical reactions. Models selected for ensemble were carefully chosen on the basis of five performance indices like r, accuracy, FV, d and RMSE. Taking five parameters as selection criteria was the clue to success of the ensemble approach. Moreover, the ensemble models were a combination of linear and non-linear models, as any single approach was not giving good results.

A comparison between fourteen (linear and non-linear) prediction approaches and their ensemble models for daily average ozone concentrations prediction showed an improved performance of the later and were consistent with observed values of the daily average

and monthly diurnal values. Four combinations of ensemble models (EM-1, EM-2, and EM-3, EM-4) were constituted on the basis of performance indices from the group of ensemble models. EM-1 model comprising of the selected models (random forest and kernel-support vector machine) outperformed significantly over the individual and ensemble models. Robustness of the model was established with 2016 data set. This analysis made on the basis of high value of correlation coefficient (0.91) and index of agreement (0.92) along with low values of root mean square error (10) and fractional variance (0.50) shows that the ensemble models under realistic conditions provide good fits between measured led ozone levels. Since prediction of ozone in the surface layer of the atmosphere is a challenging task due to frequently changing dynamics of atmosphere, yet based on huge amount of data, some degree of reliability of prediction can always be achieved.

It is clear that combining predictions generated by different methods capture more accurate underlying relationship between both the meteorological and precursor pollutants predictor variables and results in improved model prediction performance and reduced complexity. Better prediction models can be established with data from the network of monitoring stations. Moreover, inclusion of data on gaseous monitoring would further improve the prediction quality.

List of Publictions

- Madhvi Rana, Susheel K. Mittal*, Gufran Beig, 2019. "**Enhanced Ozone Production in Ambient Air at Patiala Semi-Urban Site During Crop Residue Burning Events",** MAPAN-Journal of Metrology Society of India, Springer 34(2):273–288. DOI: 10.1007/s12647-019-00315-x (Sci IF= 1.25)
- Madhvi Rana, Susheel K. Mittal*, Gufran Beig, Prashant Rana, 2019. **"The impact of crop residue burning (CRB) on the diurnal and seasonal variability of the ozone and PM levels at a semi-urban site in the north-western Indo-Gangetic plain"** Journal of Earth System Science, Springer 128:166. DOI: 10.1007/s12040-019-1164-z (Sci IF= 1.42)
- Madhvi Rana, Susheel K. Mittal*, Gufran Beig, 2019. "**Assessment and prediction of surface ozone in Northwest Indo-Gangetic Plains using ensemble approach**" in Environment, Development and Sustainability, Springer. DOI: 10.1007/s10668-020-00841-8 (Sci I.F 2.19).

Papers/Posters in Conferences

• Secured 2nd position in Conference on "International Conference on Advances in Science, Engineering & Technology" held in Haridwar, Gurukul Kangari on 29-30 January, 2016.

• Attended Pre-AdMet Workshop on "Best Practices in Measurements and Evaluation of Uncertainty" from February, 22-23, 2016 organized at CSIR-National Physical Laboratory, New Delhi -110012.

• Participated in 9th International Conference on "Advances in Metrology" (AdMet-2016) from February, 24 – 26, 2016 organized at CSIR-National Physical Laboratory, New Delhi- 110012.

• Presented poster in International Conference on Emerging Areas of Environmental Science and Engineering (EAESE-2017) Organized by Department of Environmental Science and Engineering, Guru Jambheshwar University of Science and Technology, Hisar on February, 16-18, 2017.

• Presented Poster presentation Conference award category at the 2nd Indian International Conference on Air Quality Management: Health and Exposure (IICAQM

2017), jointly organized by the Department of Civil Engineering, IIT Madras, Department of Civil Engineering, IIT Delhi and Newcastle University, UK on June 1-2, 2017, at IIT Delhi.

APPENDIX-I

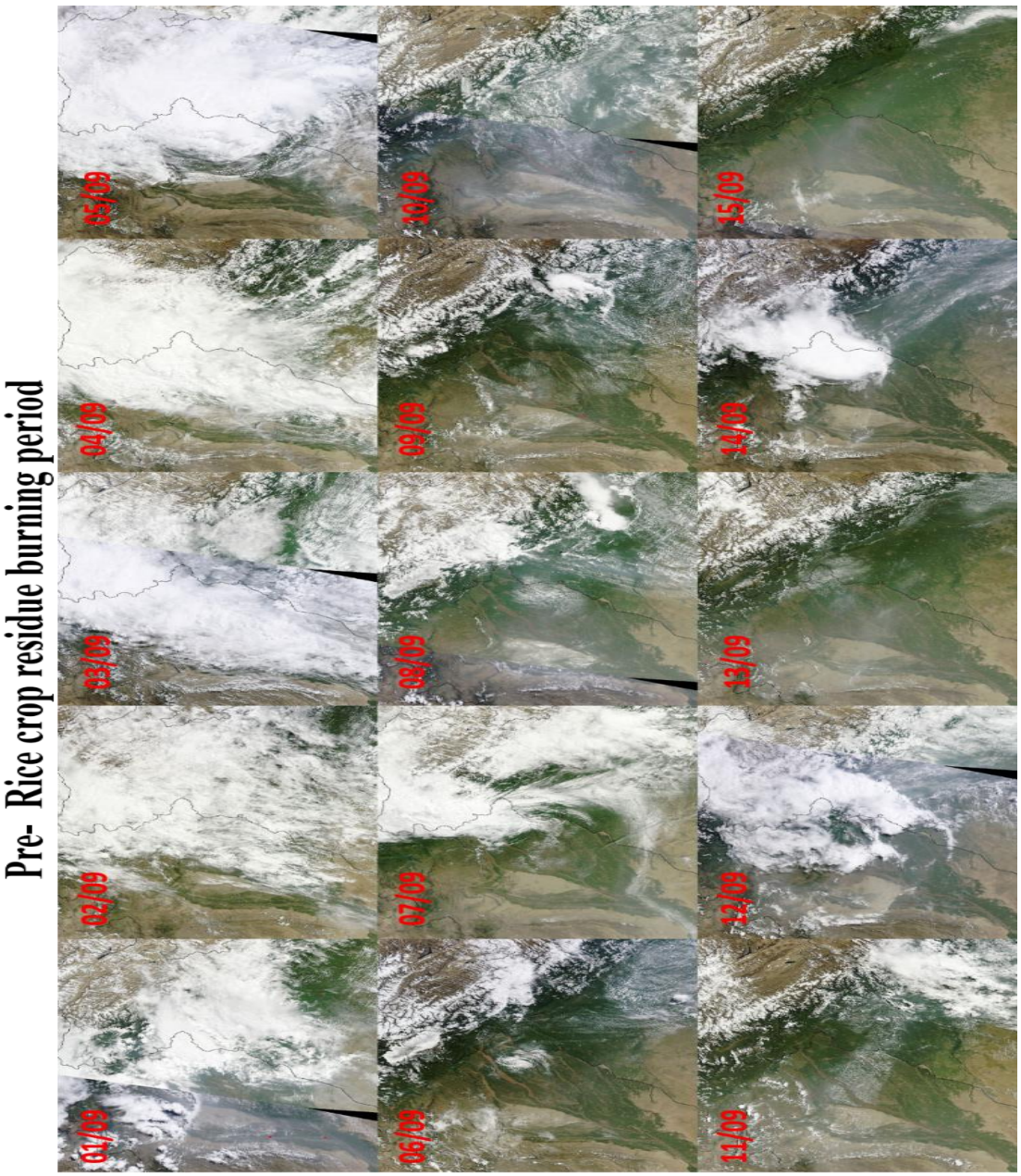
Tittle: Ambient air quality standard values for important pollutants as prescribed by Central Pollution Control Board

S.No.	Pollutant	Time weighted average	Concentration in Ambient Air		
			Industrial, Rural and other area	Ecologically Sensitive Area	Methods of measurement
1.	Sulphur dioxide (SO_2), µg/m^3	Annual 24 hours	50 80	20 80	-Improved west and Gaeke. -Ultraviolet fluorescence
2.	Nitrogen Dioxide (NO_2), µg/m^3	Annual 24 hours	40 80	30 80	Modified Jacob and Hochheiser (Na- Arsenite) -Chemiluminescence
3.	PM_{10}, µg/m^3	Annual 24 hours	60 100	60 100	-Gravimetric -TOEM -Beta attenuation
4.	$PM_{2.5}$, µg/m^3	Annual 24 hours	40 60	40 60	-Gravimetric -TOEM -Beta attenuation
5.	Ozone (O_3), µg/m^3	24 hours 8 hours 1 hours	60 100 180	60 100 180	-UV photometric -Chemiluminescence -Chemical method
6.	Lead (Pb), µg/m^3	Annual 24 hours	0.50 1	0.50 1	-AAS/ ICP method after sampling on EPM 2000 or equiv. filter paper -ED-XRF using Teflon filter
7.	Carbon monoxide (CO), mg/m^3	8 hours 1 hour	02 04	02 04	-NDIR (Non Dispersive Infra-red) spectroscopy.
8.	Ammonia (NH_3), µg/m^3	Annual 24 hours	100 400	100 400	-Chemiluminescence -Indophenol blue method.
9.	Benzene(C_6H_6), µg/m^3	Annual	05	05	-GC -Adsorption and desorption followed by GC.

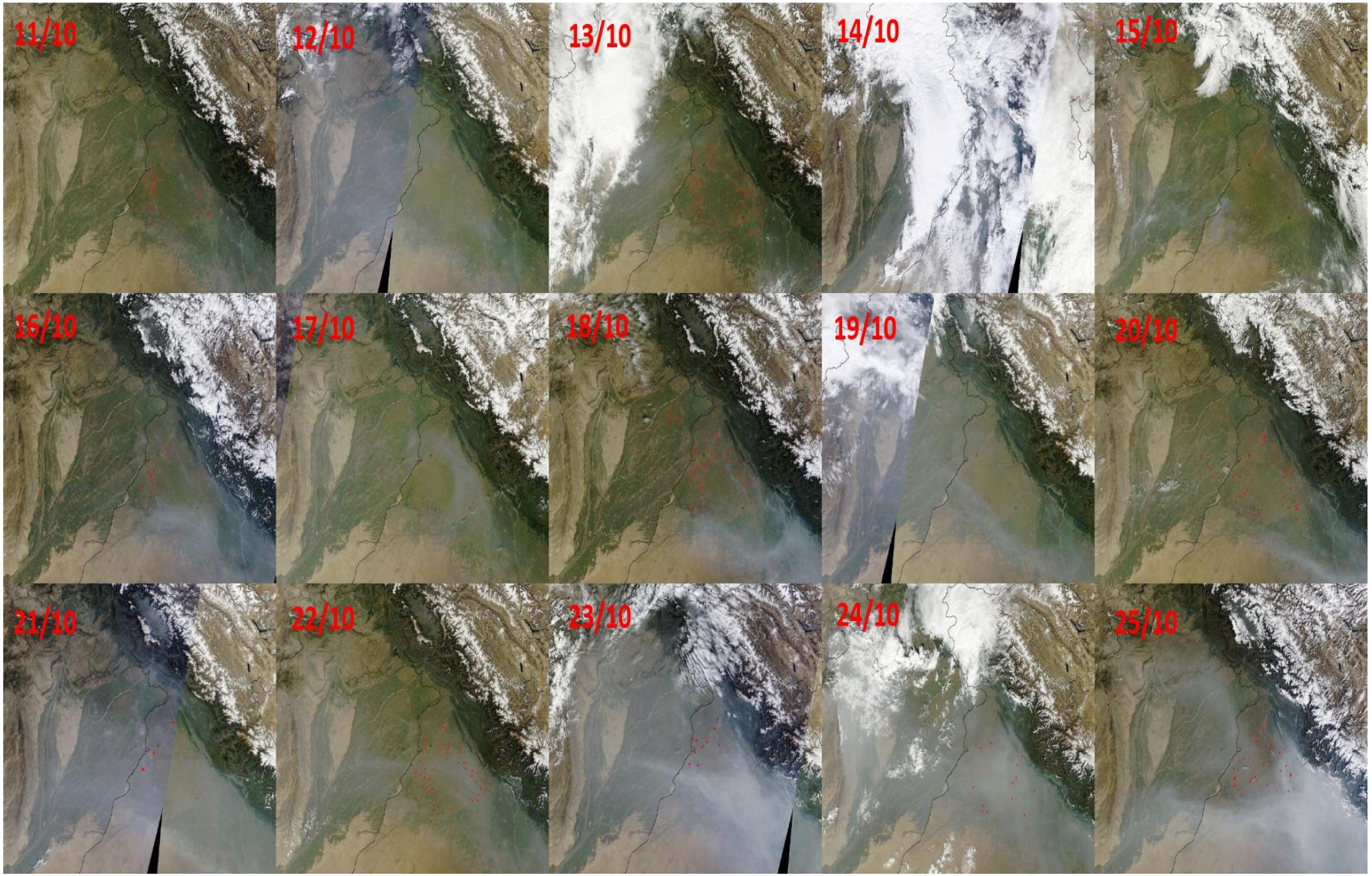
10.	Benzo(a)pyrene (Bap), µg/m^3	Annual	01	01	-Solvent extraction followed by HPLC/GC analysis
11.	Arsenic(As), ng/m^3	Annual	06	06	-AAS/ICP method after sampling on EPM 2000
12.	Nickel (Ni), ng/m^3	Annual	20	20	-AAS/ICP method after sampling on EPM 2000.

APPENDIX-II

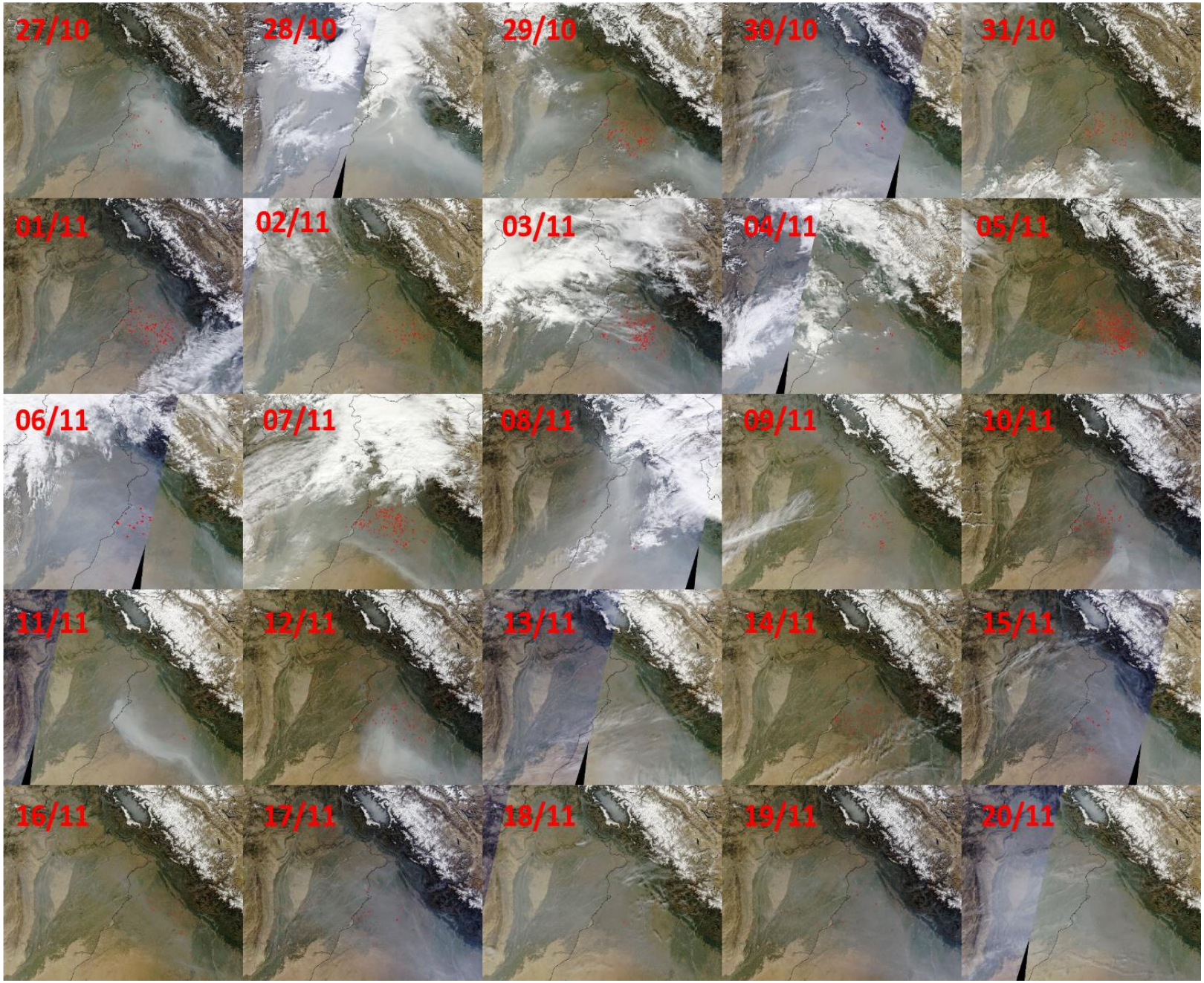
Tittle: Pre, during and post biomass burning periods of rice and wheat have been delineated over the study period using Terra-MODIS (Moderate Resolution Imaging Spectroradiometer) true color composites



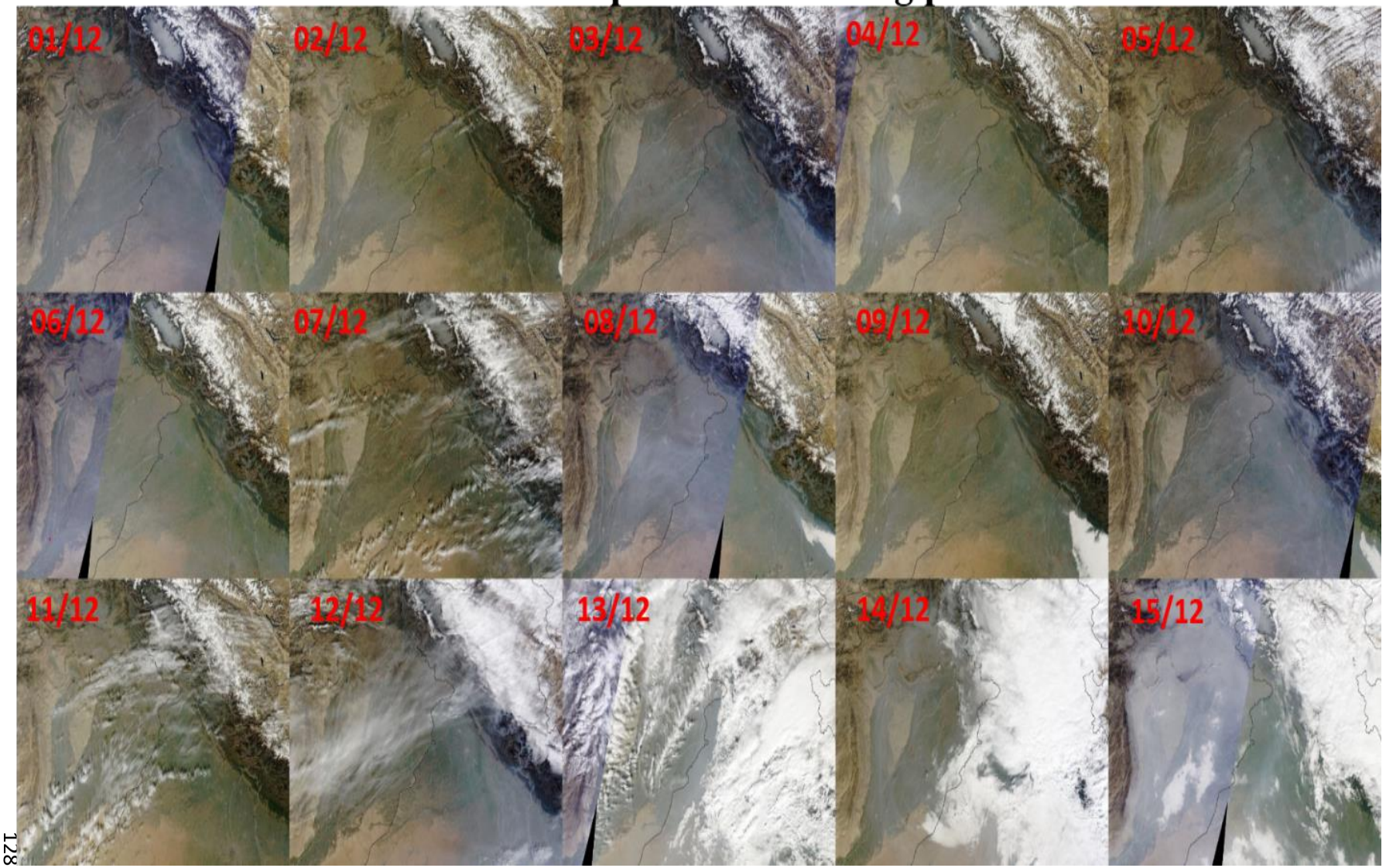
During- Rice crop residue burning



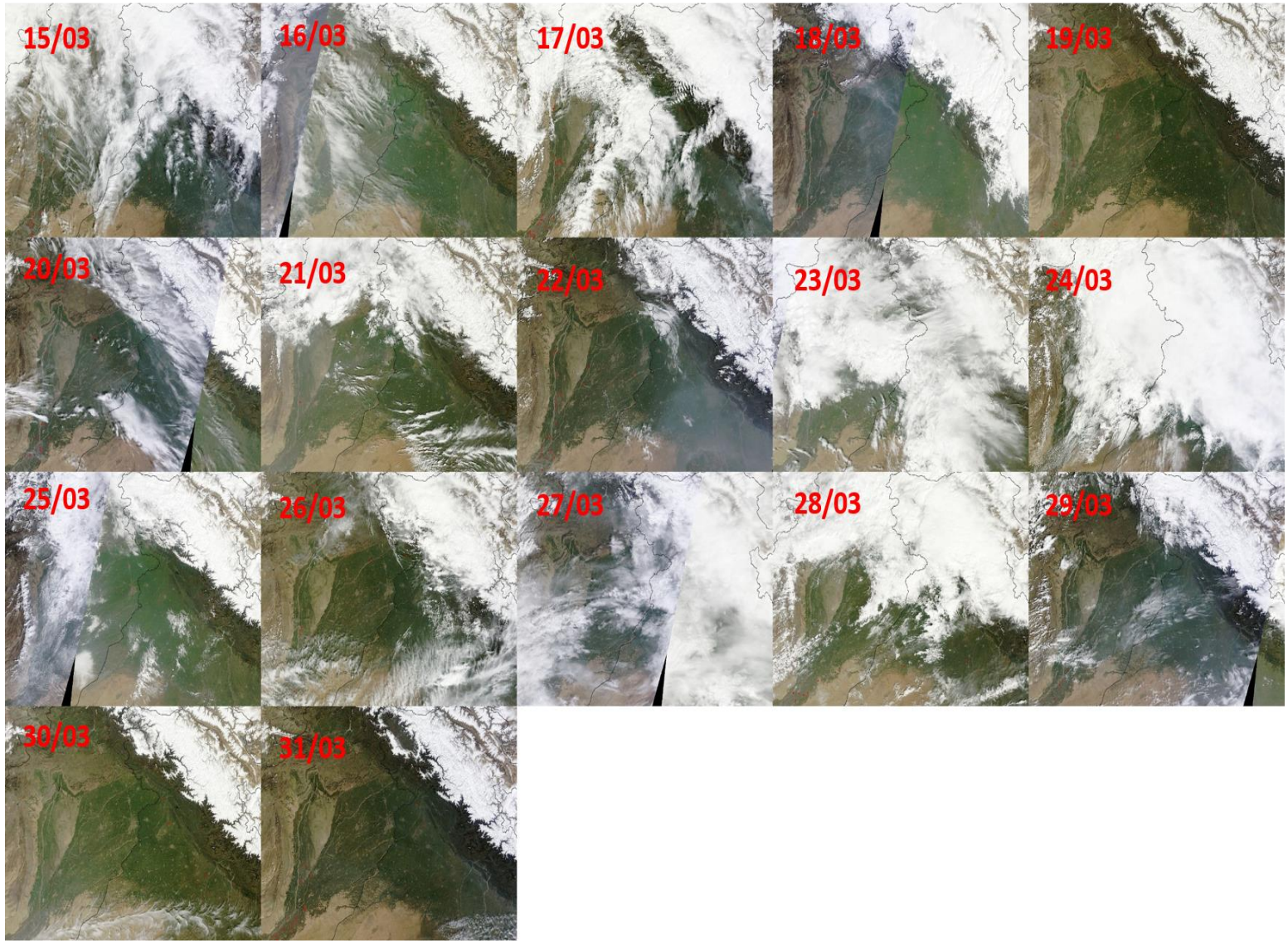
During- Rice crop residue burning



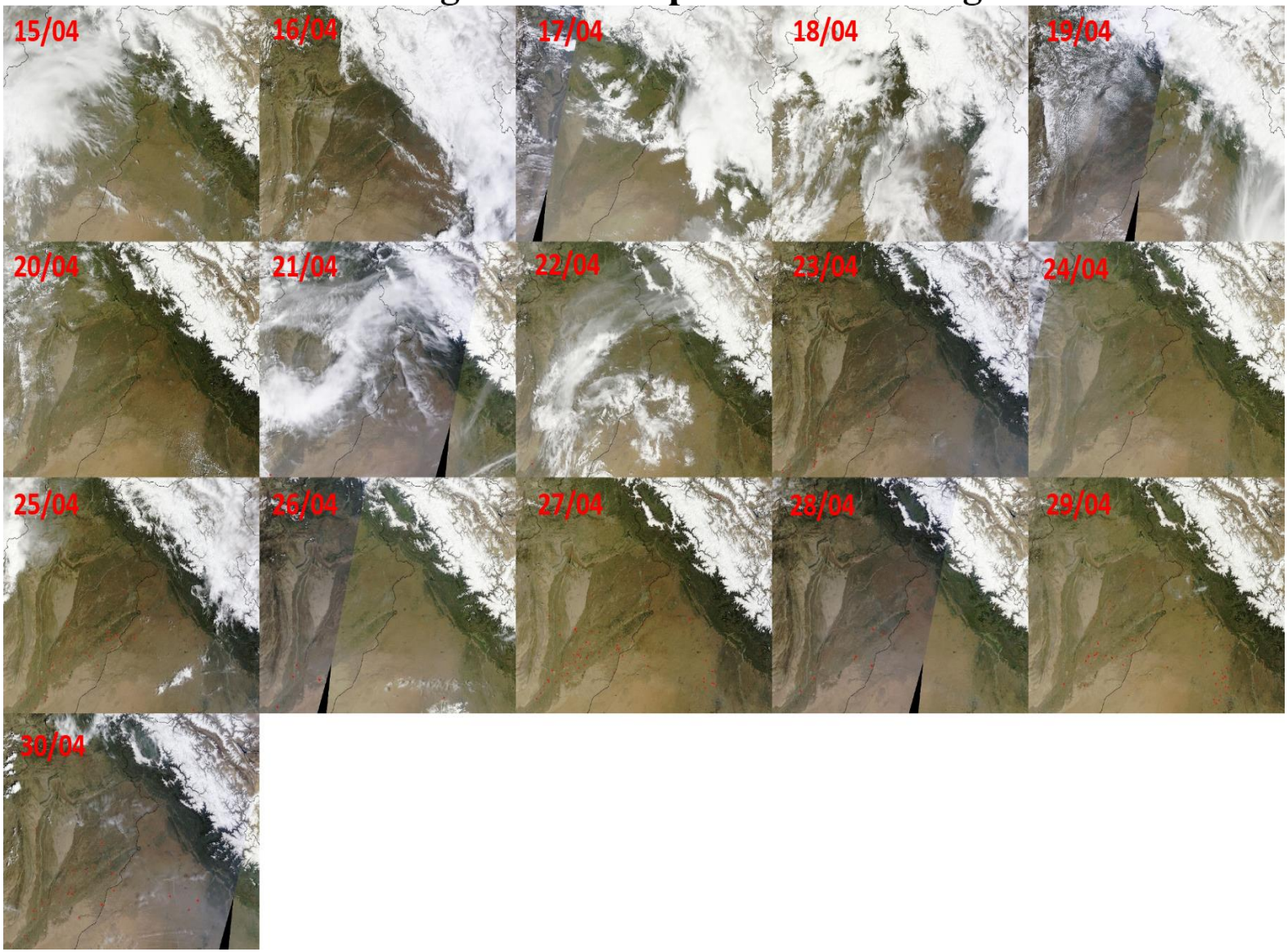
Post-Rice crop residue burning period



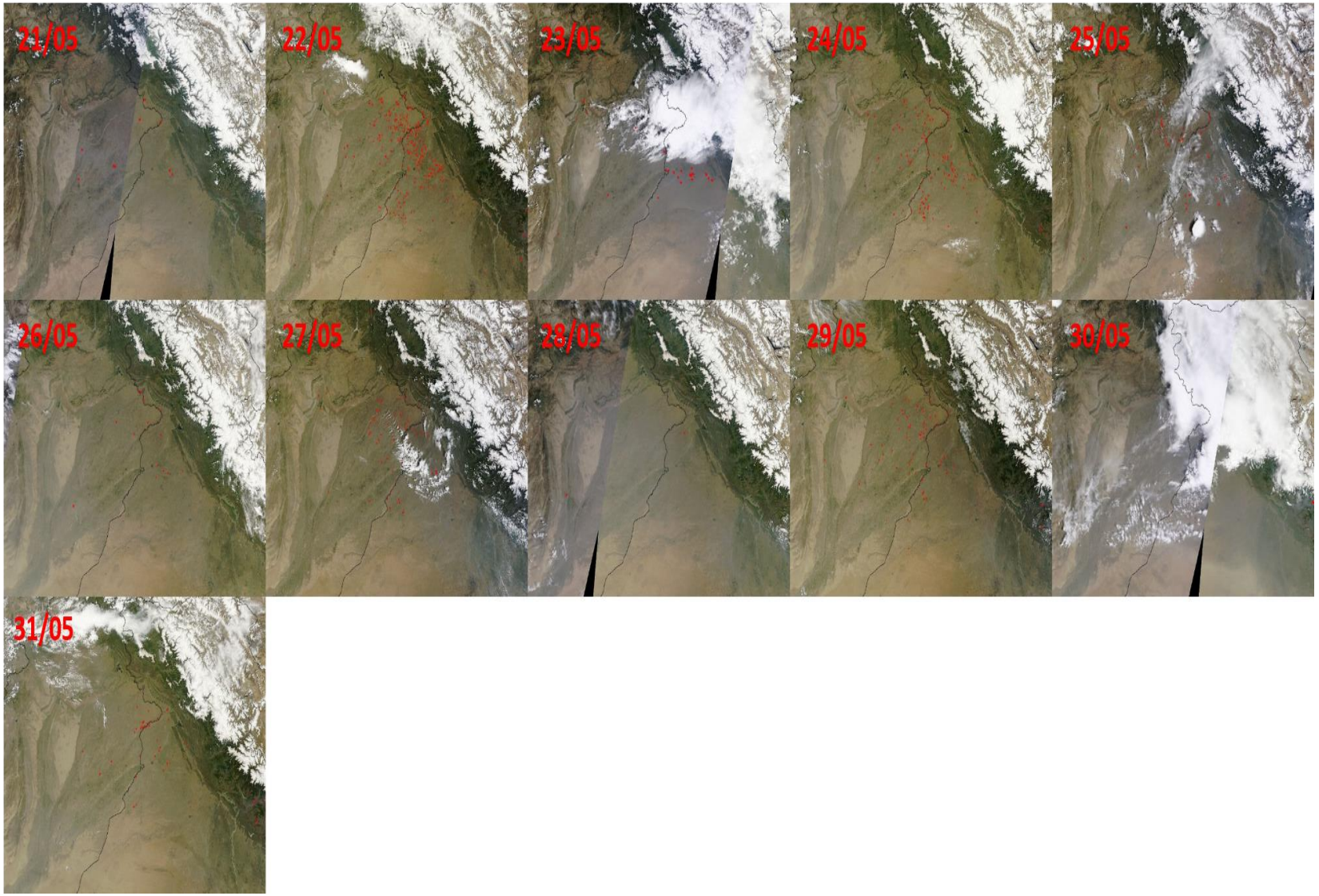
Pre- Wheat crop residue burning period



During- Wheat crop residue burning



During- Wheat crop residue burning



Post- Wheat crop residue burning