MODELLING AMBIENT $PM_{2.5}$ EXPOSURE AT HYPERLOCAL SCALE FOR DELHI AND SETTING EXPOSURE REDUCTION TARGET TO MINIMIZE HEALTH BURDEN

SHASHI TIWARI

CENTRE FOR ATMOSPHERIC SCIENCES

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JULY 2024

MODELLING AMBIENT $PM_{2.5}$ EXPOSURE AT HYPERLOCAL SCALE FOR DELHI AND SETTING EXPOSURE REDUCTION TARGET TO MINIMIZE HEALTH BURDEN

by

SHASHI TIWARI

Centre for Atmospheric Sciences

Submitted

In fulfilment of the requirements of the degree of doctor of philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

JULY 2024

DEDICATED TO MY PARENT

Certificate

This is to certify that the thesis entitled **"Modelling ambient $PM_{2.5}$ exposure at hyperlocal scale for Delhi and setting exposure reduction target to minimize health burden"** being submitted by Ms. Shashi Tiwari for the award of the degree of **Doctor of Philosophy**, is a record of the original bonafide research work carried out by her. She has worked under my guidance and supervision and has fulfilled the requirements for the submission of this thesis. The results presented in this thesis have not been submitted in part or whole to any University or Institution for award of any degree/diploma.

Prof. Sagnik Dey
Centre for Atmospheric Sciences,
Indian Institute of Technology Delhi,
Hauz Khas, New Delhi-110016, India

Acknowledgements

I sincerely thank my supervisor, Prof. Sagnik Dey, for his insightful advice and ongoing encouragement. He assisted me in becoming a capable researcher and problem-solver. I was inspired by his serenity at my most trying circumstances.

I would also like to thank Prof. S.N. Tripathi of IIT Kanpur for giving me the opportunity to work with him as a CO-PI on a project at the start of my PhD career. He trusted me, which boosted my self-confidence and advanced my professional career. Prof. Guy of the Max Planck Institute for Meteorology (MPI-M), Hamburg, Germany, provided me the chance to work on his project and extended an offer to me to join the organization as a research associate before I had even earned my doctorate. For this, I would like to convey my sincere appreciation. I thank Prof Somnath Baidya Roy, Prof. Manju Mohan all the Centre of Atmospheric Sciences faculty members for helpful suggestions during my seminar presentation.

I thank my friends Rohit Chowdhary and Anushree for always being the friends in need. I thank Mr. Alok Kumar for helping with data. I also thank Vijay for helping me with the logistics. Finally, I thank my husband he helped me become a woman to face difficult situations. My patience level improved during my PhD, which benefitted me both personally and professionally. And my mentor Prof. Sagnik Dey was instrumental in helping me happily reach my goal. Special thanks to my supporting parents, always being there for me, and their constant encouragement and support. I dedicate my thesis to my parents and Guide.

Date: Shashi Tiwari

Abstract

Exposure to ambient fine particulate matter ($PM_{2.5}$) has been established as a causal determinant of morbidity and premature mortality in children and adults. Delhi, as the capital city of India, has persistently garnered attention due to its elevated concentrations of $PM_{2.5}$, prompting multifaceted investigations into source identification, policy implementation, mitigation efforts, and their consequential health impacts. In this work, an analysis of ground monitoring stations spanning recent years (2017-2022) conducted. Delhi has achieved significant success in reducing annual $PM_{2.5}$ levels from 2017 to 2019, with a commendable rate of 4.7% per year. The substantial decline in $PM_{2.5}$ levels observed in 2020 was facilitated by the sharp reduction in emissions during the COVID-19 lockdown from March to May. If the current annual decline of 4.7% is sustained, the annual $PM_{2.5}$ level in 2024 will reach 85 µg m^{-3}, representing a 32.6% reduction that meets NCAP target. However, the reduction percentage decreased in 2022, posing a challenge to achieving the NCAP target by 2024. By 2022, there had been an 18% reduction in $PM_{2.5}$ levels in Delhi when considering all existing monitoring stations and their yearly updates from the NCAP base year. However, focusing solely on the common monitoring stations operational throughout the 2017-2022 period reveals a less impressive reduction of only 10%.

Deriving hyperlocal information is critical for quantifying exposure disparities and managing air quality at neighbourhood scales in cities. Delhi is one of the most polluted megacities in the world, where ground-based monitoring was limited before 2017.Here ambient $PM_{2.5}$ exposure at 100 m × 100 m spatial resolution for the period 2002–2019 using the random forest model is estimated. The model-predicted daily and annual $PM_{2.5}$ show a ten-

fold cross-validation R^2 of 0.91 and 0.95 and root mean square error of 19.3 and 9.7 µg m^{-3}, respectively, against coincident ground measurements from CPCB ground network. Annual mean $PM_{2.5}$ exposure varied in the range of 90–160 µg m^{-3} in Delhi, with shifts in local hotspots and a reduction in spatial heterogeneity over the years.

Additionally, a deep learning model is employed to test the forecasting of $PM_{2.5}$ concentrations at an hourly scale. The primary challenge encountered involves filling in-situ data gaps, for which four methods were adopted. Among that, the KNN Spatial iterator demonstrated superior accuracy at 0.97%. For short-term forecasting over the next 72 hours, the CNN-LSTM Encoder-Decoder model with Multivariate Input exhibited optimal performance, yielding an R value of 0.81 and an RMSE of 19.67 µg/m^3.

The rising health burden in Delhi, considering changes in exposure levels, baseline mortality of non-communicable diseases (NCDs), and population growth, is delved into, in the work. The changes in associated health burden during this period are estimated, and the contributing factors are isolated. The mortality burden attributable to ambient $PM_{2.5}$ in Delhi increased by 49% between 2002 and 2019, though its relative share of the total mortality burden has decreased from 10% in 2002 to 8% in 2019. Improvement baseline mortality helped to decrease in burden, otherwise 56% would have climbed. Hence, the changes in health burden attributable to air pollution needs to be interpreted in terms of socio-demographic factors in policy framing. . IHD had the highest contribution to mortality and DALY in 2002 and 2019; its relative contributions increased from 43% to 50% and from 36% to 43%, respectively. COPD has the second-largest contribution to mortality (decreased from 26% in 2002 to 24% in 2019) and DALY (increased from 215 in 2002 to 24% in 2019), and the highest contribution to YLD burden where its contribution decreased from 60% in 2002 to 56% in 2019. The third highest contribution to the disease burden is stroke, whose contribution to death, DALY, and YLD remained mostly unchanged. For death and DALY, the contributions of T2D and LC also remained unchanged, while that of LRI decreased in this period.

While doing the sensitivity analysis for the exposure resolution, it is found that, the mortality burden in 2002 and 2019 higher by 10% and 3.1%, respectively, for exposure assessment at 100 m scale relative to the estimates with 1 km scale. Finally, the exposure reduction target for Delhi to meet the UN SDG3.4 goal, which requires efforts beyond the current NCAP, is set. Delhi can meet the United Nations Sustainable Development Goal 3.4 target of reducing the non-communicable disease burden attributable to $PM_{2.5}$ by one-third in 2030 relative to 2015 by reducing ambient $PM_{2.5}$ exposure below the World Health Organization’s first interim target of 35 μg m^{-3}.

The results presented as part of this thesis is expected to help the Indian policymakers in identifying the areas where mitigation policies should be directed and formulate more efficient air quality management plans so that maximum health benefit can be obtaine

सारांश

परिवेशीय सूक्ष्म कण पदार्थ (पीएम2.5) के संपर्क को बच्चों और वयस्कों में रुग्णता और समय से पहले मृत्यु दर के एक कारक के रूप में स्थापित किया गया है। भारत की राजधानी के रूप में दिल्ली ने पीएम2.5 की उच्च सांद्रता के कारण लगातार ध्यान आकर्षित किया है, जिससे स्रोत की पहचान, नीति कार्यान्वयन, शमन प्रयासों और उनके परिणामी स्वास्थ्य प्रभावों में बहुआयामी जांच को बढ़ावा मिला है। इस कार्य में, हाल के वर्षों (2017-2022) में फैले ग्राउंड मॉनिटरिंग स्टेशनों का विश्लेषण किया गया। दिल्ली ने 2017 से 2019 तक प्रति वर्ष 4.7% की सराहनीय दर के साथ वार्षिक PM2.5 स्तर को कम करने में महत्वपूर्ण सफलता हासिल की है। 2020 में PM2.5 के स्तर में जो उल्लेखनीय गिरावट देखी गई, वह मार्च से मई तक COVID-19 लॉकडाउन के दौरान उत्सर्जन में तेज कमी के कारण हुई। यदि 4.7% की मौजूदा वार्षिक गिरावट बरकरार रहती है, तो 2024 में वार्षिक पीएम2.5 स्तर 85 $\mu g\ m^{-3}$ तक पहुंच जाएगा, जो एनसीएपी लक्ष्य को पूरा करने वाली 32.6% की कमी का प्रतिनिधित्व करता है। हालाँकि, 2022 में कमी का प्रतिशत कम हो गया, जिससे 2024 तक NCAP लक्ष्य प्राप्त करने में चुनौती पैदा हो गई। 2022 तक, सभी मौजूदा निगरानी स्टेशनों और NCAP से उनके वार्षिक अपडेट पर विचार करने पर दिल्ली में PM2.5 के स्तर में 18% की कमी आई थी। आधार वर्ष। हालाँकि, 2017-2022 की अवधि के दौरान केवल सामान्य निगरानी स्टेशनों पर ध्यान केंद्रित करने से केवल 10% की कम प्रभावशाली कमी का पता चलता है।

शहरों में पड़ोस के पैमाने पर जोखिम असमानताओं को मापने और वायु गुणवत्ता के प्रबंधन के लिए हाइपरलोकल जानकारी प्राप्त करना महत्वपूर्ण है। दिल्ली दुनिया के सबसे प्रदूषित मेगासिटीज में से एक है, जहां 2017 से पहले जमीन आधारित निगरानी सीमित थी। यहां यादृच्छिक वन मॉडल का

उपयोग करके 2002-2019 की अवधि के लिए 100 मीटर × 100 मीटर स्थानिक रिज़ॉल्यूशन पर परिवेशी पीएम 2.5 जोखिम का अनुमान लगाया गया है। मॉडल-अनुमानित दैनिक और वार्षिक PM2.5, सीपीसीबी ग्राउंड नेटवर्क से संयोगित ग्राउंड माप के मुकाबले क्रमशः 0.91 और 0.95 का दस गुना क्रॉस-वैलिडेशन R2 और 19.3 और 9.7 μg m-3 की मूल माध्य वर्ग त्रुटि दिखाता है। दिल्ली में वार्षिक औसत PM2.5 एक्सपोज़र 90-160 μg m−3 की रेंज में भिन्न-भिन्न है, स्थानीय हॉटस्पॉट में बदलाव और पिछले कुछ वर्षों में स्थानिक विविधता में कमी आई है।

इसके अतिरिक्त, प्रति घंटे के पैमाने पर PM2.5 सांद्रता के पूर्वानुमान का परीक्षण करने के लिए एक गहन शिक्षण मॉडल नियोजित किया जाता है। सामने आई प्राथमिक चुनौती में इन-सीटू डेटा अंतराल को भरना शामिल है, जिसके लिए चार तरीकों को अपनाया गया था। उनमें से, केएनएन स्पैटियल इटरेटर ने 0.97% पर बेहतर सटीकता का प्रदर्शन किया। अगले 72 घंटों में अल्पकालिक पूर्वानुमान के लिए, मल्टीवेरिएट इनपुट के साथ सीएनएन-एलएसटीएम एनकोडर-डिकोडर मॉडल ने इष्टतम प्रदर्शन प्रदर्शित किया, जिससे 0.81 का आर मान और 19.67 का आरएमएसई प्राप्त हुआ।μजी/एम3.

कार्य में दिल्ली में बढ़ते स्वास्थ्य बोझ, जोखिम के स्तर में बदलाव, गैर-संचारी रोगों (एनसीडी) की आधारभूत मृत्यु दर और जनसंख्या वृद्धि पर विचार किया गया है। इस अवधि के दौरान संबंधित स्वास्थ्य बोझ में परिवर्तन का अनुमान लगाया गया है, और योगदान करने वाले कारकों को अलग किया गया है। टीदिल्ली में परिवेशी PM2.5 के कारण होने वाली मृत्यु दर में 2002 और 2019 के बीच 49% की वृद्धि हुई है, हालांकि कुल मृत्यु दर में इसकी सापेक्ष हिस्सेदारी 2002 में 10% से घटकर 2019 में 8% हो गई है।आधारभूत मृत्यु दर में सुधार से बोझ कम करने में मदद मिली, अन्यथा 56% चढ़ जाता। इसलिए, वायु प्रदूषण के कारण स्वास्थ्य पर पड़ने वाले बोझ में बदलाव की व्याख्या नीति निर्धारण में सामाजिक-

जनसांख्यिकीय कारकों के संदर्भ में की जानी चाहिए।. 2002 और 2019 में मृत्यु दर और DALY में IHD का योगदान सबसे अधिक था; इसका सापेक्ष योगदान क्रमशः 43% से बढ़कर 50% और 36% से 43% हो गया। सीओपीडी का मृत्यु दर में दूसरा सबसे बड़ा योगदान है (2002 में 26% से घटकर 2019 में 24% हो गया) और DALY (2002 में 215 से बढ़कर 2019 में 24% हो गया), और YLD बोझ में सबसे अधिक योगदान है जहां इसका योगदान 60 से कम हो गया 2002 में % से 2019 में 56%। बीमारी के बोझ में तीसरा सबसे बड़ा योगदान स्ट्रोक है, जिसका मृत्यु, डीएएलवाई और वाईएलडी में योगदान ज्यादातर अपरिवर्तित रहा। मृत्यु और DALY के लिए, T2D और LC का योगदान भी अपरिवर्तित रहा, जबकि इस अवधि में LRI का योगदान कम हो गया।

एक्सपोज़र रिज़ॉल्यूशन के लिए संवेदनशीलता विश्लेषण करते समय, यह पाया गया कि, 1 किमी स्केल के अनुमान के सापेक्ष 100 मीटर स्केल पर एक्सपोज़र मूल्यांकन के लिए 2002 और 2019 में मृत्यु दर क्रमशः 10% और 3.1% अधिक है। अंत में, संयुक्त राष्ट्र एसडीजी3.4 लक्ष्य को पूरा करने के लिए दिल्ली के लिए जोखिम में कमी का लक्ष्य निर्धारित किया गया है, जिसके लिए वर्तमान एनसीएपी से परे प्रयासों की आवश्यकता है। दिल्ली विश्व स्वास्थ्य संगठन के पहले अंतरिम लक्ष्य 35 से नीचे परिवेशी पीएम2.5 के जोखिम को कम करके 2015 के सापेक्ष 2030 में पीएम2.5 के कारण होने वाले गैर-संचारी रोग के बोझ को एक तिहाई तक कम करने के संयुक्त राष्ट्र सतत विकास लक्ष्य 3.4 के लक्ष्य को पूरा कर सकती है। $\mu g \cdot m-3$.

इस थीसिस के हिस्से के रूप में प्रस्तुत परिणामों से भारतीय नीति निर्माताओं को उन क्षेत्रों की पहचान करने में मदद मिलने की उम्मीद है जहां शमन नीतियों को निर्देशित किया जाना चाहिए और अधिक कुशल वायु गुणवत्ता प्रबंधन योजनाएं तैयार करनी चाहिए ताकि अधिकतम स्वास्थ्य लाभ प्राप्त किया जा सके।

Table of contents

List of figures

List of tables

List of abbreviations

AIC: Akaike Information Criteria

AOD- Aerosol Optical Depth

AQI: Air Quality Index

AR: Autoregression

ARIMA: Autoregressive Integrated Moving Average Model

BM: Baseline Mortality

BP: Back Propagation

BPNN: Back Propagation Neural Network

CAAQMS- Continuous Ambient Air Quality Monitoring Stations

CI: Confidence Interval

CNN BP: Convolutional Neural Network Back Propagation

CNN-LSTM: Convolutional Neural Network - Long Short-Term Memory

COPD: Chronic Obstructive Pulmonary Disease

CPCB- Central Pollution Control Board

CTM- Chemical Transport Model

CV: Cross Validation

DALYs: Disability-Adjusted Life Years

EPCA: Environmental Pollution (Prevention and Control) Authority

FB: Fractional Bias

GAM- Generalized Additive Models

GBD- Global Burden Disease

GBM: Gradient Boosting Machine

GOI: Government of India

GRAP: Graded Response Action Plan

h: Hours

HDV: Heavy Duty Vehicles

I: Differencing

IHD: Ischemic Heart Disease

IITM: Indian Institute of Tropical Meteorology

IMD: India Meteorology Department

KNN: K-Nearest Neighbors

LC: Lung Cancer

LRI: Lower Respiratory Infection

LSTM: Long Short-Term Memory

MA: Moving Average

MAE: Mean Absolute Error

MAIAC : Multiangle Implementation of Atmospheric Correction

MAPE: Mean Absolute Percentage Error

MERRA-2: Modern-Era Retrospective analysis for Research and Applications version 2

ML-Machine Learning

MLR: Multiple Linear Regression

MOEF&CC- Ministry of Environment, Forest and Climate Change

MoES: Ministry of Earth Science

MRDs: Mean Relative Deviations

MSE- Mean Square Error

NAAQS: National Ambient Air Quality standards

NCAP- National Clean Air Program

NCAR: National Centre for Atmospheric Research

NCD- non-communicable diseases

NCR- National Capital Region

NH: National Highway

NMSE: Normalized Mean Square Error

NN: Neural Network

PBL: planetary boundary layer

PCA: principal component analysis

PM- Particulate matter

PM_{10} - Particulate matter with aerodynamical diameter less than 10µm

$PM_{2.5}$ - Particulate matter with aerodynamical diameter less than 2.5µm MAE

PMUY: Pradhan Mantri Ujjwala Yojana

R: correlation

R^2- coefficient of determination

RF: Random Forest

RMSE- root mean square error

RMSE: root mean square error

RNN: Recurrent Neural Network

RR: Relative Risks

SAFAR: system of air quality and weather forecasting and research

SARIMA: Seasonal Autoregressive Integrated Moving Average

SDG- Sustainable Development Goal

SVM: Support Vector Machine

T2D: Type-2-Diabetes

UIs: Uncertainty Intervals

UN: United Nation

VC- Ventilation coefficient

VC: Ventilation coefficient

VIIRS: Visible Infrared Imaging Radiometer Suite

WHO: World Health Organization

w.r.t.: With Respect to

XGBoost: Extreme Gradient Boosting

YLD: Years Of Life Lived with Disability