SUMMER INTERNSHIP PROGRAM

at

National Remote Sensing Centre, ISRO

From 12/05/2025 to 25/07/2025

A REPORT BY

Samaresh Ghorami

Master of Business Administration

Hospital and Health Management

Roll no - 2067

2024-26

under the Mentorship of

Dr. Varsha Tanu

भारत सरकार
अन्तरिक्ष विभाग
राष्ट्रीय सुदूर संवेदन केन्द्र
बालानगर, हैदराबाद - 500 037, तेलंगाना, भारत
टेलिफोन : +040-23879572-76
+040-23879261-65
फैक्स : +040-23878648

Government of India
Department of Space (DOS)
National Remote Sensing Centre
Balanagar, Hyderabad - 500 037, Telangana, India
Telephone : +040-23879572-76
: +040-23879261-65
Fax : +040-23878648

CERTIFICATE

This is to certify that this project report entitled "**Identification and Predictive Analysis of Air Pollution Hotspots and Correlation with Health Index using QGIS & Remote Sensing**" is a record of the Bonafide work carried out by Mr. Samaresh Ghorami under my guidance and supervision at the National Remote Sensing Centre (NRSC), ISRO, Hyderabad during the period 12th May 2025 to 25th July 2025.

Mr. Samaresh Ghorami has demonstrated sincerity, dedication, and a high degree of reliability throughout his project work. His consistent efforts and commitment to excellence were evident in all aspects of the project. I wish him continued success in his future endeavors and a rewarding career in the field of research.

(Signature of the Guide)

DR. MAHESH PATHAKOTI

डॉ. महेश पाताकोटी / Dr. Mahesh Pathakoti
वैज्ञानिक-एस एफ / Scientist-SF
पृथ्वी और जलवायु विज्ञान क्षेत्र (ईसीएसए)
Earth and Climate Sciences Area (ECSA)
राष्ट्रीय सुदूर संवेदन केंद्र (एनआरएससी)
National Remote Sensing Centre (NRSC)
भारतीय अंतरिक्ष अनुसंधान संगठन (इसरो)
Indian Space Research Organisation (ISRO)
अंतरिक्ष विभाग Dept. of Space / भारत सरकार Govt. of India
बालानगर - Balanagar / हैदराबाद - Hyderabad-500037

IIHMR UNIVERSITY

IIHMR University Faculty Mentor Approval

This is to certify that **Samaresh Ghorami** of academic year 2024-26 of Institute of Health Management Research has prepared a poster with respect to his summer internship held during May-July 2025.

He has carried out work as per the guidelines. His poster is approved for presentation.

29/7/2025

Date: 29/07/2025

Dr. Varsha Tanu
Associate Professor

1, Prabhu Dayal Marg, Near Sanganer Airport, Jaipur- 302029, Rajasthan, India.
Ph.: 0141-3924700, 2791431- 32, E-mail: iihmr@iihmr.edu.in, Web.: www.iihmr.edu.in

Identification and Predictive Analysis of Air Pollution Hotspots and Correlation with Health Index using QGIS & Remote Sensing

Faculty mentor: Dr. Varsha Tanu, Associate Professor at IIHMR University
Industry Mentor: Dr. Mahesh Pathakoti, Scientist/Engineer 'SF', NRSC/ISRO

Government of India
National Remote Sensing Centre
Indian Space Research Organisation
ISO 9001:2015

Name: Samaresh Ghorami
Roll no.: 2067
Stream: MBA HM29

About Organisation

NRSC is part of the ISRO, Department of Space, and is located in Hyderabad. NRSC collects remote sensing data from various domestic and international Earth observation satellites. This data, characterized by diverse geographical, spectral, and temporal resolutions, supports applications in urban planning, forestry, agriculture, water management, and disaster monitoring. Advanced data processing techniques extract valuable information, aiding decision-making across multiple sectors. NRSC operates and maintains satellite ground stations to process, store, and retrieve data from ISRO's satellites, ensuring smooth satellite mission operations and timely data delivery to end users.

NRSC Vision Statement

To continue to be in the forefront of developing remote sensing applications in the country and be the technology expert in establishing ground stations and generating high quality satellite and aerial data products.

Organisation Mandate

National Remote Sensing Centre (NRSC) has the mandate for establishment of ground stations for receiving satellite data, generation of data products, dissemination to the users, development of techniques for remote sensing applications, including disaster management support, geospatial services for good governance, and capacity building for professionals, faculty and students.


Disaster Management Support


Geospatial Services


Satellite Data Reception

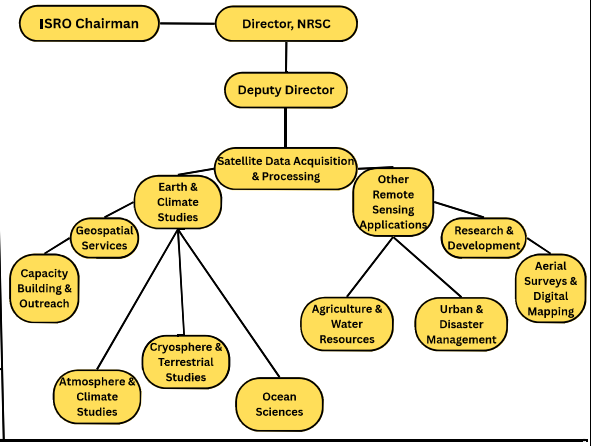

Data Processing & Dissemination


Remote Sensing Applications


Capacity Building


Aerial Services & Digital Mapping

National Remote Sensing Centre (NRSC), ISRO Structure



Objectives

- To establish and empirically quantify the correlations between various ground-based air pollutants and mortality rates across different Indian states, thereby illuminating regional specificities in these relationships.
- To leverage the developed predictive model to project prospective mortality rates and identify future mortality hotspots across India for the year 2030, offering vital foresight for strategic public health planning and resource allocation.
- To develop a robust machine learning model capable of predicting state-level standardized mortality rates in India by synthesizing a comprehensive dataset primarily integrating satellite-derived air quality data and mortality data.

Study Area

This study focuses on the entirety of India, a South Asian nation that, as of recent estimates, stands as the most populous country globally. Comprising 28 states and 8 Union Territories.

Data & Methodology

Datasets Used-

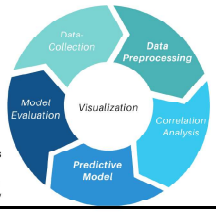
- Standardized Mortality rates and Population Data: IHME GBD database (Indian states and UTs, 2017 - 2021).
- Ground Station Air Quality Data: Ground-level air quality observations for PM10, PM2.5, SO2, and NO2 (National Air Quality Monitoring Program 9 & Continuous Ambient Air Quality Monitoring network, managed by CPCB, 2017-2021).
- Aerosol Optical Depth: Annual average AOD from the MODIS (Moderate Resolution Imaging Spectroradiometer) sensor, 2010-2023. AOD serves as a proxy for atmospheric particulate matter loading and is highly correlated with ground-level PM2.5 and PM10 concentrations.
- SO2: Annual average SO2 concentrations from the TROPOMI (Tropospheric Monitoring Instrument) sensor on the Sentinel-5 Precursor satellite, 2018-2023. TROPOMI offers high spatial resolution and daily global coverage, making it ideal for tracking SO2 emissions from industrial sources.
- NO2: Annual average NO2 concentrations also from the TROPOMI sensor, 2018-2023. NO2 is a key indicator of combustion processes, primarily from vehicular traffic and fossil fuel power generation.

DATA ANALYSIS WORKFLOW

Stage-1 Data Collection
Mortality Data, Ground Station Pollutant Data (PM2.5, PM10, SO2, NO2), Satellite Data (AOD, NO2, SO2)

Stage-2 Data Preprocessing
Feature Engineering (Lagged Pollutants), Missing Value Handling.

Stage-3 Correlation Analysis
Correlation between Ground Station Pollutants & Mortality



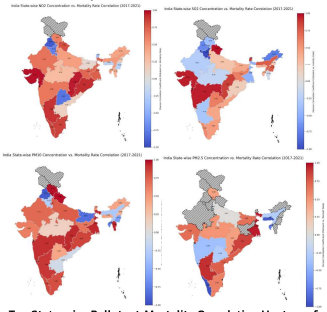
Stage-4 Predictive Model
Random Forest Regressor

Stage-5 Model Evaluation
Performance Metrics (R-squared, MAE, RMSE), Feature Importance Analysis.

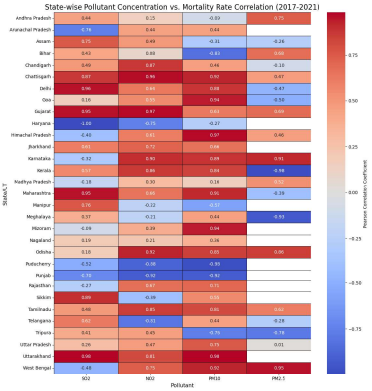
Stage-6 Future Projection
2030 Mortality Rate Predictions

Results & Discussion

Pollutant-Mortality Correlations

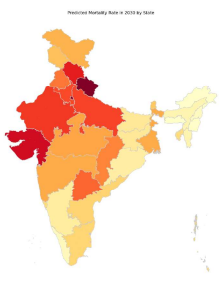


Top State-wise Pollutant-Mortality Correlation Heatmap for Specific Pollutants



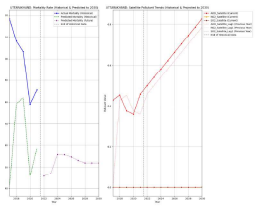
States with strong positive SO2 correlations include Uttarakhand(highest 0.97), Delhi, Maharashtra, Gujarat, Sikkim, Chattisgarh, showing strong links to industrial activity. This finding is consistent with the known health effects of SO2 from industrial emissions.
States with strong positive NO2 correlations include Gujarat(highest 0.97), Chattisgarh, Odisha, Karnataka, Chandigarh, and Kerala. highlighting the contribution of traffic-related NO2 to mortality.
States with strong positive PM2.5 correlations include West Bengal, Karnataka, Odisha, Andhra Pradesh, Gujarat, and Bihar. This finding is consistent with expectations for the influence of factors such as vehicular traffic, urbanization, and industrialization
States with strong positive PM10 correlations include Uttarakhand (highest 0.98), Himachal Pradesh, Goa, Mizoram, West Bengal, and Chattisgarh. Rapid urbanization and industrial growth suggest that coarse particulate matter from construction and vehicular emissions is a significant health factor.

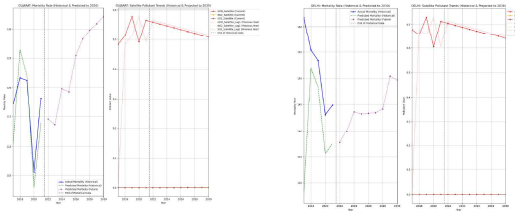
Predicted Mortality Hotspots (2030 Projection)



State/Union Territory	Predicted Mortality Rate (2030)
Uttarakhand	150.05
Delhi	150
Gujarat	132
Himachal Pradesh	125
Rajasthan	117
Haryana	112
Uttar Pradesh	111
Telangana	106
Punjab	103
Chhattisgarh	98.5
Puducherry	97
Maharashtra	95
Madhya Pradesh	88.5
West Bengal	87
Ladakh	82.5
Jammu and Kashmir	80.5
Dadra and Nagar Haveli and Daman and Diu	80
Andhra Pradesh	78
Sikkim	77
Lakshadweep	73.5
Andaman & Nicobar Islands	72
Chandigarh	71.5
Tamil Nadu	70.2
Goa	69.5
Tripura	69.5
Bihar	61
Jharkhand	61
Karnataka	57
Odisha	56
Mizoram	53
Assam	52.5
Kerala	49
Nagaland	48
Manipur	46
Arunachal Pradesh	45.5
Meghalaya	44.5

Top State-wise Predicted Mortality Rate Graphs (2030)





Uttarakhand is predicted to have the highest mortality rate in 2030, with a value around 151. This high rate is not directly tied to a corresponding spike in the AOD pollutant value, as its pollutant levels are also projected to increase, reaching close to 0.8 by 2030, which is high but not the highest among all states. The high mortality rate in Uttarakhand, as seen in the scatter plots, is an outlier compared to other states with similar pollutant concentrations. This suggests that while air pollution may be a contributing factor, other variables not detailed in the provided graphs, such as healthcare access or specific health crises, could be driving the high and rapidly increasing mortality rate in the state. Delhi is projected to have the second-highest mortality rate in 2030, at approximately 150. This trend is accompanied by a slight decrease in the AOD pollutant value, which is expected to be below 0.65 by 2030. The scatter plots show Delhi has a very high mortality rate relative to its high PM10 concentrations, indicating a strong correlation between pollutants and mortality in this region. Gujarat is expected to come as the 3rd with 132 Mortality rates.

Practical Exposure	Theoretical Learnings
Directly experienced the complexities of working with multiple, heterogeneous datasets (mortality, satellite AOD, SO2, NO2, population, geographical boundaries). This involved significant data cleaning, standardization (e.g., state names), and merging.	This project applied Biostatistical concepts like correlation to examine the relationship between environmental factors (air pollution) and health outcomes (mortality rates). Using Demographic tools, it analyzed population dynamics and mortality trends, aligning with Health and Development theories that emphasize the impact of social and environmental determinants on public health.

Challenges faced during internship

Deals with lots of different pieces of information (like mortality records, satellite data, and population numbers) that didn't easily fit together and had missing parts.

Major learnings from internship

Learned how to take a project from start to finish, from getting all the raw information to showing clear results and advice.

Suggestions for improvement

Thinking about how to make our system work for much larger amounts of data or even in real-time, so it can be used more widely and quickly.

SWOT Analysis

Strength
- Our project tackles a major health and environmental issue in a very important region (Indo-Pacific), especially in India.
- We can forecast future pollution and its health impacts, which helps policymakers to make a real difference.
- Our project brings Unique Data Combination.

Weakness
- Getting all the different kinds of data ready and clean can be a very time-consuming and manual process.
- Reliance on specific satellite data (MODIS, TROPOMI) means we depend on their availability and consistency.

SWOT

Opportunity
- The methods we developed can be used in other countries facing similar air pollution issues.
- Growing public interest in air quality can create more demand for our insights.
- There's a strong need for data-driven policies to combat air pollution, and our project provides exactly that.

Threats
- Changes in access to satellite data or ground monitoring networks could impact our ability to continue the project.
- Unexpected, rapid changes in pollution sources or weather patterns might affect the accuracy of our long-term predictions.

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 1 **From:** 13th May to 17th May

Activity Done	• Joined the organization on 13th may. • Received the project named "Identification of Air pollution Hotspot and correlation with Health Index using GIS & Remote Sensing". • Had an initial meeting with my mentor, Mahesh P. sir, to discuss the project. • Reviewed introductory documents and data on the project background (Health index, Pollutants data) and objectives. • Was introduced to the internal task management software, "Q-GIS"(which will be used for GIS analysis) and completed account setup.
Linking theory (Classroom learning) with practice at your organisation	The principles of Research Methodology, particularly regarding systematic data review and planning for analysis, are proving relevant in understanding the project's approach to identifying air pollution hotspots and their correlation with health indices.
Gaps identified on the working area.	The organization utilizes specific internal software(Q-GIS) for task management. This week involved initial exposure, and there is a gap in my proficiency with these tools.
Suggestions for improvement	Suggesting a brief training session or providing access to comprehensive user manuals or video tutorials for the internal software and processes.

Plan for the next week	• I will conduct a detailed examination of the provided datasets on health indices and pollutants. This will involve understanding the specific variables included, the geographical coverage (identifying the cities in India for which data is available). • I will focus on learning the basics of Python programming and the Pandas library. Pandas is crucial for working with and analyzing data, which will be essential for this project. • Alongside Python learning, I will also dedicate some time to reviewing introductory materials or articles on the fundamental GIS concepts related to identifying spatial hotspots. • Schedule a follow-up meeting with my mentor, Mahesh P. sir, to discuss my initial findings from the data exploration and my progress in Q-GIS.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 2 **From:** 19th May to 24th May

Activity Done	•Sourcing state-wise mortality data related to air pollution (e.g., from the Global Burden of Disease study, for causes like Ischemic Heart Disease, Stroke, COPD, Lung Cancer, and Lower Respiratory Infections attributable to air pollution). •Joining this data to state administrative boundaries in QGIS. •Configuring the QGIS map to reflect these cause-specific mortality rates for each year (2017-2021). •Generating individual map frames for each year (2017-2021), accurately reflecting these mortality changes and including a dynamic year label. •Demonstrating the change in cause-specific mortality values for specific states, providing a click-and-inspect view of the data for any given year(2017-2021)
Linking theory (Classroom learning) with practice at your organisation	This week's work, involving the sourcing, integration, and visualization of cause-specific mortality data, directly applied "Principles from Demography" to understand population health dynamics.by mapping the spatial and patterns of air pollution-related deaths, these activities contribute to the broader field of "Health and Development", offering insights crucial for public health planning and policy.

Gaps identified on the working area.	A notable gap identified this week is the challenge in sourcing publicly available, consistently updated, and high-resolution (city/district-wise) cause-specific mortality data, particularly for air pollution-related illnesses. This limits the depth of sub-state level spatial analysis and detailed impact assessment at a finer administrative scale.
Suggestions for improvement	To enhance the analytical depth, focus on calculating and utilizing population-adjusted mortality rates (e.g., per 100,000 population) for cause-specific deaths. This normalization is crucial for valid spatial and temporal comparisons across states, providing more accurate public health insights.
Plan for the next week	•Calculate population-adjusted mortality rates for cause-specific deaths (2017-2021). •Begin sourcing raw air pollutant data (e.g., PM2.5, PM10). •Analyze district-wise air quality data and compute state averages. •Strategize with mentor about the refined data approach, including the planned use of population-adjusted mortality rates and air pollutant data integration.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 3 **From:** 26th May to 31st May

Activity Done	•Obtained and processed age-standardized mortality rates attributable to Ambient Particulate Matter Pollution (APMP) for Indian states from the Global Burden of Disease (GBD) database for the relevant years. •This normalization is crucial for valid spatial and temporal comparisons across states, providing more accurate public health insights. •Gathered and sorted air pollutant data (PM2.5, PM10, SO2, NO2, O3) for the years 2015 to 2025, organized year-wise. •Successfully integrated prepared health data into QGIS by joining it to the India states boundary shapefile. •Generated comprehensive choropleth maps visualizing the Total APMP Mortality Rate across Indian states for each year. •Developed temporal animations in QGIS to illustrate the changes in APMP mortality rates over time across Indian states.
Linking theory (Classroom learning) with practice at your organisation	My approach to processing the health data was directly informed by the principles of Epidemiology, a subject where I learned about concepts such as age-standardized mortality rates and their crucial role in comparative public health analysis. This foundational understanding was gained under the guidance of Prof. Sanjay Gupta Sir. The normalization of mortality rates, was emphasized in class as essential for valid spatial and temporal comparisons

	across diverse populations like those in Indian states, ultimately providing more accurate public health insights.
Gaps identified on the working area.	•The current analysis does not yet include spatial visualization or direct correlation of the gathered pollutant data. •The current analysis does not account for potential lag effects in health outcomes from air pollution exposure.
Suggestions for improvement	•Integrate and visualize the gathered air pollutant data in QGIS to directly compare its spatial and temporal patterns with health outcomes. •Perform comprehensive statistical correlation analysis between APMP mortality rates and individual air pollutant levels for each year. •Investigate the potential for lag effects by analyzing relationships between past pollutant exposures and current health outcomes.
Plan for the next week	•Generate comprehensive choropleth maps for each key air pollutant (PM2.5, PM10, SO2, NO2, O3) for each relevant year. •Develop animations in QGIS for individual pollutants to illustrate their changes over time across Indian states. •Conduct statistical correlation analysis between the Total APMP Mortality Rate and each individual air pollutant at the state level for each year

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 4 **From:** 2nd May to 7th May

Activity Done	•I gathered all the raw air quality data from the CPCB ground station spanning 2015 through 2023. •I directly adjusted and standardized column headers in several source Excel files for better consistency. •I used Python to automatically read, combine, and clean all the raw data, including fixing missing state and city information. •I created numerous graphs to show national pollution trends, state-level pollutant trends, top polluted states/cities, year-by-year comparisons across all states, and analyzed pollution changes from 2015 to 2023.
Linking theory (Classroom learning) with practice at your organisation	**Excel application** classes were crucial for understanding the initial data structure and performing necessary manual adjustments, while insights from **Digital Health** helped contextualize the pollution data for public health implications, guiding the visualization and analysis towards meaningful outcomes.
Gaps identified on the working area.	The existing dataset, while good for historical analysis and comparison, currently lacks enough raw data points or specific influencing factors needed for robust future pollution prediction models.
Suggestions for improvement	To address the data limitations for future prediction, our next step will be to integrate satellite data. This will provide

	a broader and potentially more consistent source of air quality information.
Plan for the next week	•We will analyze the relationship between the collected air quality data and mortality rates to understand potential health impacts. •We will begin the process of gathering satellite-derived air quality data to enhance our dataset for future predictions and broader analysis

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 5 **From:** 9th June to 14th June

Activity Done	•**Data Acquisition & Preprocessing (Ground-Based Data):** Successfully loaded, cleaned, and merged air quality (SO2, NO2, PM10, PM2.5) and all-cause mortality rate data for Indian states from 2017-2021. •**Exploratory Data Analysis (EDA) & Initial Correlation Analysis:** Performed EDA, conducted detailed state-wise Pearson correlation analysis between pollutants and mortality rates (2017-2021), and generated visualizations including heatmaps, scatter plots, and initiated choropleth maps. •**Initiation of Satellite Data Collection for Predictive Analysis:** Set up Google Earth Engine (GEE)and uploading the Indian state boundaries. Developed Python code to extract annual average AOD (2010-2023), NO2 (2018-2023), and SO2 (2018-2023) concentrations per state from satellite data.
Linking theory (Classroom learning) with practice at your organisation	The analysis of health impacts, correlation studies, and understanding mortality rates directly apply principles from **Biostatistics**, particularly in data interpretation and drawing statistically sound conclusions.
Gaps identified on the working area.	While the existing dataset is good for historical analysis and comparison, it currently lacks enough raw data points or specific influencing factors needed for robust future pollution prediction models or predictive analysis.

Suggestions for improvement	To address the data limitations for future prediction, our next step will be to integrate satellite data. This will provide a broader and potentially more consistent source of air quality information.
Plan for the next week	•**Finalize Satellite Data Acquisition:** Monitor and confirm the completion of GEE export tasks. Download the exported AOD, NO2, and SO2 CSV files from Google Drive to the local machine. •**Integrate Satellite Data:** Load the newly acquired satellite data into the Jupyter Notebook and merge it with the existing ground-based air quality and mortality datasets. This will create a comprehensive dataset for predictive modeling. •**Search for Extended Mortality Data:** Actively investigate and acquire additional years of mortality data (e.g., before 2017 or after 2021) to match the extended satellite dataset timeframe (2010-2023), if available. •**Begin Predictive Modeling Preparation:** Start exploring and preparing additional features (e.g., demographic data, meteorological variables) and outline the methodology for building the predictive model.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 6 **From:** 16th June to 21st June

Activity Done	• Finalized Satellite Data Acquisition: Successfully monitored and confirmed the completion of GEE export tasks. • Integrated Satellite Data: Loaded the newly acquired satellite data and successfully merged it with the existing ground-based air quality and mortality datasets. This has created a comprehensive dataset for predictive modeling. • Searched for Extended Mortality Data: Initiated active investigation and acquisition of additional years of mortality data (e.g., before 2017 or after 2021) to match the extended satellite dataset timeframe (2010-2023), if available. • Began Predictive Modeling Preparation: Started exploring and preparing additional features and outlined the methodology for building the predictive model.
Linking theory (Classroom learning) with practice at your organisation	Concepts from Biostatistics, Demography, and Digital Health were crucial in guiding the integration of diverse datasets, understanding population-level health impacts, and framing the predictive modeling approach. Geospatial analysis skills were directly applied in handling and merging satellite data.
Gaps identified on the working area.	While significant data integration has occurred, refining feature selection for the predictive model and ensuring robust data consistency across all merged sources remain

	areas for further attention. The availability of extended historical mortality data is still a potential limitation.
Suggestions for improvement	To enhance predictive accuracy, the next steps will involve rigorous feature engineering and selection, potentially exploring more advanced imputation techniques for any remaining data gaps. Continued efforts to acquire comprehensive historical mortality data are also vital.
Plan for the next week	• To conduct in-depth feature engineering, including the creation of lagged pollutant features and other relevant indicators. • To begin training and validating initial predictive models (e.g., SVR, Random Forest) using the integrated dataset. • To perform preliminary model evaluation using relevant metrics (R-squared, MAE, RMSE). • To refine the methodology for nationwide forecasting and hotspot identification based on model outputs.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 7 **From:** 23rd June to 28th June

Activity Done	• Conducted in-depth feature engineering, including the creation of lagged pollutant features and other relevant indicators. • Began training and validating initial predictive models (e.g., SVR, Random Forest) using the integrated dataset. • Performed preliminary model evaluation using relevant metrics (R-squared, MAE, RMSE). • Refined the methodology for nationwide forecasting and hotspot identification based on model outputs.
Linking theory (Classroom learning) with practice at your organisation	The practical application of statistical modeling and validation directly reinforced concepts from Biostatistics, particularly in model evaluation and interpreting metrics like R-squared, MAE, and RMSE for health outcomes. The integration of complex datasets and the development of predictive tools demonstrated the real-world application of Digital Health and Demography principles in public health forecasting.
Gaps identified on the working area.	While significant progress has been made in model development, further fine-tuning of hyperparameters and exploring ensemble methods could potentially enhance predictive performance. Ensuring the interpretability of complex model outputs for policy translation remains an ongoing area of focus.
Suggestions for improvement	To further optimize model performance, continued experimentation with different model architectures and hyperparameter tuning is suggested. Additionally,

	developing more intuitive visualizations or dashboards for policymakers would aid in actionable insights from the complex predictive outputs.
Plan for the next week	• Conduct an in-depth analysis of model results, including feature importance. • Develop visualization tools to effectively present forecasting and hotspot identification. • Prepare a comprehensive summary of the project's findings and methodologies. • Begin drafting the final project report, incorporating all data analysis and model outcomes.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 8 **From:** 30th June to 5th July

Activity Done	• Conducted an in-depth analysis of model results, including feature importance. • Developed visualization tools to effectively present forecasting and hotspot identification. • Prepared a comprehensive summary of the project's findings and methodologies. • Began drafting the final project report, incorporating all data analysis and model outcomes.
Linking theory (Classroom learning) with practice at your organisation	The detailed analysis of model results and feature importance directly applied advanced statistical concepts from Biostatistics, allowing for a deeper understanding of the relationship between air pollution and health outcomes. The development of visualization tools effectively translated complex data into actionable insights, reflecting principles of Digital Health for effective communication and public health interventions. The comprehensive report drafting solidified skills in presenting interdisciplinary research.
Gaps identified on the working area.	While the model results and visualizations are well underway, further refinement of the interactive elements within the visualization tools could enhance user engagement and policy impact. Ensuring the narrative of the final report seamlessly integrates all technical aspects with broader health and development implications remains a focus.

Suggestions for improvement	To maximize the project's impact, consider developing interactive dashboards for the visualization tools to allow for dynamic exploration by policymakers. A dedicated review process for the final report, involving experts from both technical and health policy domains, would also be beneficial to ensure clarity and relevance.
Plan for the next week	Finalize all visualization tools and ensure their accuracy and user-friendliness. Complete the drafting of the final project report, including conclusions and recommendations. Prepare for the writing and submission of a conference research paper publication.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 9 **From:** 7th July to 12th July

Activity Done	• Finalized all visualization tools and ensured their accuracy and user-friendliness. • Completed the drafting of the final project report, including conclusions and recommendations. • Prepared for the writing and submission of a conference research paper publication.
Linking theory (Classroom learning) with practice at your organisation	The meticulous process of finalizing visualization tools and drafting the final report deeply leveraged principles from Research Methodology, ensuring the project's findings were accurately represented and conclusions were well-supported by evidence. The preparation for conference paper submission also directly applied theoretical knowledge of academic writing, ethical considerations, and dissemination of research, as taught in Research Methodology.
Gaps identified on the working area.	While the project's core deliverables are complete, the ongoing process of academic publication requires continuous engagement with peer review feedback and potential revisions. Optimizing the visual presentation of complex statistical outputs for a broad audience within publication constraints can always be refined.
Suggestions for improvement	To further enhance the impact of the research, dedicate time to actively seeking feedback on the draft conference paper from mentor. Explore opportunities for interactive elements in digital versions of the report or supplementary materials for wider accessibility.

Plan for the next week	• Incorporate feedback received on the final project report and prepare for submission. • Continue refining the conference research paper draft based on early reviews. • Start writing a full journal research paper.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Reporting Format (Weekly)

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Week no: 10 **From:** 14th July to 25th July

Activity Done	• Incorporated feedback received on the final project report and prepared it for submission. • Continued refining the conference research paper draft based on early reviews. • Started writing a full journal research paper. • Successfully submitted both the conference and journal research papers.
Linking theory (Classroom learning) with practice at your organisation	This week's activities, particularly the refinement and submission of research papers, deeply applied principles from Research Methodology, encompassing rigorous data presentation, adherence to academic standards, and the ethical dissemination of findings. The overall project's journey, from data integration to predictive modeling and public health implications, continuously drew upon concepts from Biostatistics, Digital Health, and Demography, bridging theoretical understanding with practical application in addressing complex societal challenges.
Gaps identified on the working area.	While the research papers have been submitted, the ongoing peer-review process may highlight areas for further clarification or additional analysis. The transition from project completion to sustained research dissemination requires continued focus on academic engagement and feedback incorporation.
Suggestions for improvement	To maximize the impact of the research, proactively monitor the status of submitted papers and be prepared for

	potential revisions based on reviewer comments. Begin exploring opportunities for presenting findings at wider forums, such as webinars or industry workshops, beyond academic conferences.
Plan for the next week	Prepare a comprehensive poster and report for the college poster presentation.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor

Compiled Report

Name of the Organisation: National Remote Sensing Centre, ISRO

Area/ Department/ Project of Summer Internship Program: Identification and Predictive Analysis of Air Pollution Hotspots and correlation with Health Index using GIS & Remote Sensing.

Name of the Student: SAMARESH GHORAMI

Roll no: 2067 **Stream:** MBA(Hospital and Health Management)

Activity Done	• Joined the National Remote Sensing Centre, ISRO, and was assigned the project. Held an initial meeting with mentor Mahesh P. Sir, reviewed project background documents (Health Index, Pollutants data), and set up an account on Q-GIS for GIS analysis. • Sourced state-wise mortality data (2017-2021) from the Global Burden of Disease study for air pollution-related causes, integrated it with state boundaries in QGIS, and created yearly map frames with dynamic labels for visualization. • Processed age-standardized mortality rates for Ambient Particulate Matter Pollution (APMP) from the GBD database, gathered air pollutant data (PM2.5, PM10, SO2, NO2, O3) for 2015-2025, integrated health data into QGIS, and developed choropleth maps and temporal animations for APMP mortality rates. • Collected raw air quality data (2015-2023) from CPCB ground stations, standardized Excel files, used Python for data cleaning and merging, and created graphs for national/state-level pollution trends and comparisons. • Loaded, cleaned, and merged ground-based air quality (SO2, NO2, PM10, PM2.5) and mortality data (2017-2021), performed exploratory data analysis (EDA) with Pearson correlation, and initiated satellite data collection using Google Earth Engine (GEE) for AOD, NO2, and SO2 (2010-2023).

	• Finalized GEE satellite data acquisition, merged it with ground-based datasets, searched for extended mortality data (pre-2017, post-2021), and began preparing features and methodology for predictive modeling. • Conducted feature engineering (e.g., lagged pollutant features), trained and validated predictive models (SVR, Random Forest), performed preliminary model evaluation (R-squared, MAE, RMSE), and refined methodology for forecasting and hotspot identification. • Analyzed model results for feature importance, developed visualization tools for forecasting and hotspot identification, summarized project findings, and began drafting the final project report. • Finalized visualization tools for accuracy and user-friendliness, completed the final project report with conclusions, and prepared a conference research paper for submission. • Incorporated feedback on the final project report, refined and submitted conference and journal research papers, and prepared for a college poster presentation.
Linking theory (Classroom learning) with practice at your organisation	The project integrated principles from Research Methodology, Biostatistics, Epidemiology, Demography, and Digital Health. Systematic data review, statistical modeling (e.g., correlation analysis, predictive modeling), and visualization techniques were applied to analyze air pollution and health impacts. Academic writing and ethical dissemination principles guided the preparation and submission of research papers, bridging theoretical knowledge with practical public health applications.
Gaps identified on the working area.	• Initial lack of proficiency in Q-GIS software. • Limited availability of high-resolution, city/district-wise mortality data. • Absence of spatial visualization and lag effect analysis for pollutants.

	• Insufficient raw data points for robust predictive modeling. • Need for refined feature selection and data consistency in predictive modeling. • Requirement for hyperparameter tuning and model interpretability. • Need for more interactive visualization tools and comprehensive report narratives. • Ongoing peer-review feedback and optimization of visual outputs for publications.
Suggestions for improvement	• Provide training or manuals for Q-GIS to enhance proficiency. • Normalize mortality data with population-adjusted rates for better comparisons. • Integrate and visualize pollutant data with lag effect analysis. • Incorporate satellite data for broader air quality insights. • Employ advanced feature engineering and imputation techniques. • Experiment with model architectures and develop intuitive dashboards. • Create interactive dashboards and involve interdisciplinary experts in report reviews. • Seek mentor feedback on papers and explore interactive report elements. • Monitor peer-review feedback and explore broader dissemination platforms.

Key Learnings	• Mastered Q-GIS for geospatial analysis and data visualization. • Developed skills in sourcing, cleaning, and integrating diverse datasets (mortality, air quality, satellite). • Gained expertise in Python for data processing and predictive modeling (SVR, Random Forest). • Learned statistical techniques (correlation, model evaluation with R-squared, MAE, RMSE). • Understood the importance of feature engineering, lag effects, and data normalization in public health analysis. • Created effective visualization tools (choropleth maps, animations, graphs) for policy communication. • Gained experience in drafting and submitting academic research papers, engaging with peer-review processes. • Recognized the value of bridging theoretical concepts (Biostatistics, Epidemiology, Digital Health) with practical public health applications.

Samaresh Ghorami

Signature of the Student

Approved by the IIHMR University's Mentor