Analyzing Air Pollution Exposure in Affordable Housing Programs: A Data-Driven Approach to Environmental Risk

The City College of New York

Sajida Malik

Mentor : Yana Kucheva

Introduction

- Public Housing and the Low-Income Housing Tax Credit (LIHTC) program provide millions of households with stable and affordable living conditions, underscoring their importance in the U.S. housing landscape.
- PM2.5(Particulate Matter 2.5) refers to fine particulate matter that is small enough to penetrate deep into the lungs, causing serious health issues like respiratory and cardiovascular diseases.
- Ozone, a major component of smog, forms when pollutants chemically react in sunlight and is linked to worsening conditions like asthma and lung disease.
- Both pollutants continue to exceed national air quality standards in many areas.

Problem Statement

- Federally assisted affordable housing, including LIHTC and Public Housing, is vital for supporting low-income populations across the U.S.
- Many of these properties are located in areas with high air pollution, exposing residents to significant health risks.
- Current housing policies may inadequately address environmental risks associated with air quality.
- This project evaluates exposure to air pollutants, specifically PM2.5 and Ozone, in federally assisted housing.
- The study analyzes the relationship between housing locations, pollution levels, and residents' socioeconomic demographics.
- Findings aim to assess the effectiveness of current policies and provide insights for improving environmental health equity.

Design

Data

- National Housing Preservation Database (NHPD)
  - Geographic Variables
    - Housing Properties
    - Longitude and Latitude
- U.S. Environmental Protection Agency (EPA) EJScreen Dataset
  - Air Pollution Variables
  - Socioeconomic Variables
  - Demographic Variables

Features of the Datasets

Geographic Variables:

- Census Tract (Description: State 2 digits, County 3 digits and Census Tract 6 digits)
- LIHTC: Lihtc_status1 and 2
- Public Housing : ph_status1 and 2
- Longitude
- Latitude
- City
- State
- County

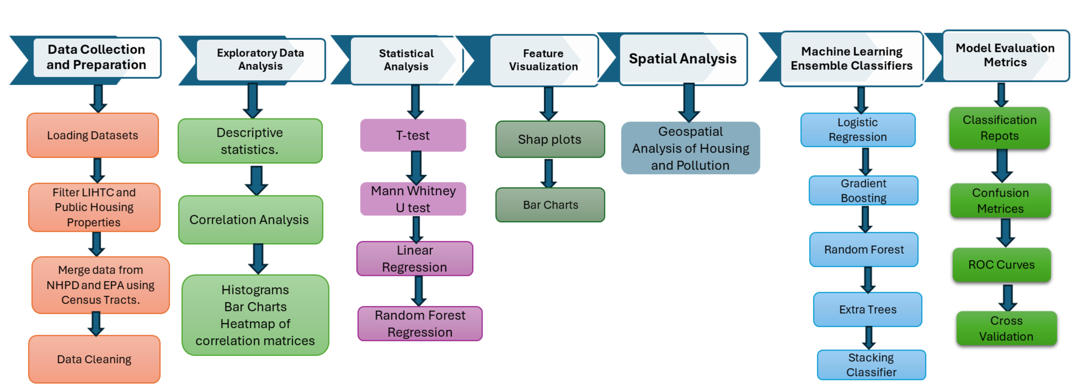
Target Variables: Air Pollution Levels:

- PM25: PM2.5
- OZONE: Ozone
- Classification of PM2.5 levels
  - Good: 0-9
  - Moderate: 9.1- 35.4
  - Unhealthy for Sensitive Group: 55.5-125.4
  - Unhealthy: 125.5-225.4
  - Very unhealthy: 125.4-225.4
  - Hazardous: 225.5+
- Classification of Ozone Levels
  - Good: 0- 54
  - Moderate: 55- 70
  - Unhealthy for Sensitive Group: 71-85
  - Unhealthy: 86-105
  - Very unhealthy: 106-200
  - Hazardous: 201

Socioeconomic and Demographic Variables:

- PEOPCOLORPCT: People of color percentage
- LOWINCPCT: Low Income percentage
- UNEMPPCT: Unemployment percentage
- UNDER5PCT: Under 5 years of age percentage
- OVER64PCT: Over 64 years of age percentage
- LIFEEXPPCT: Life expectancy percentage.

Analytical Framework



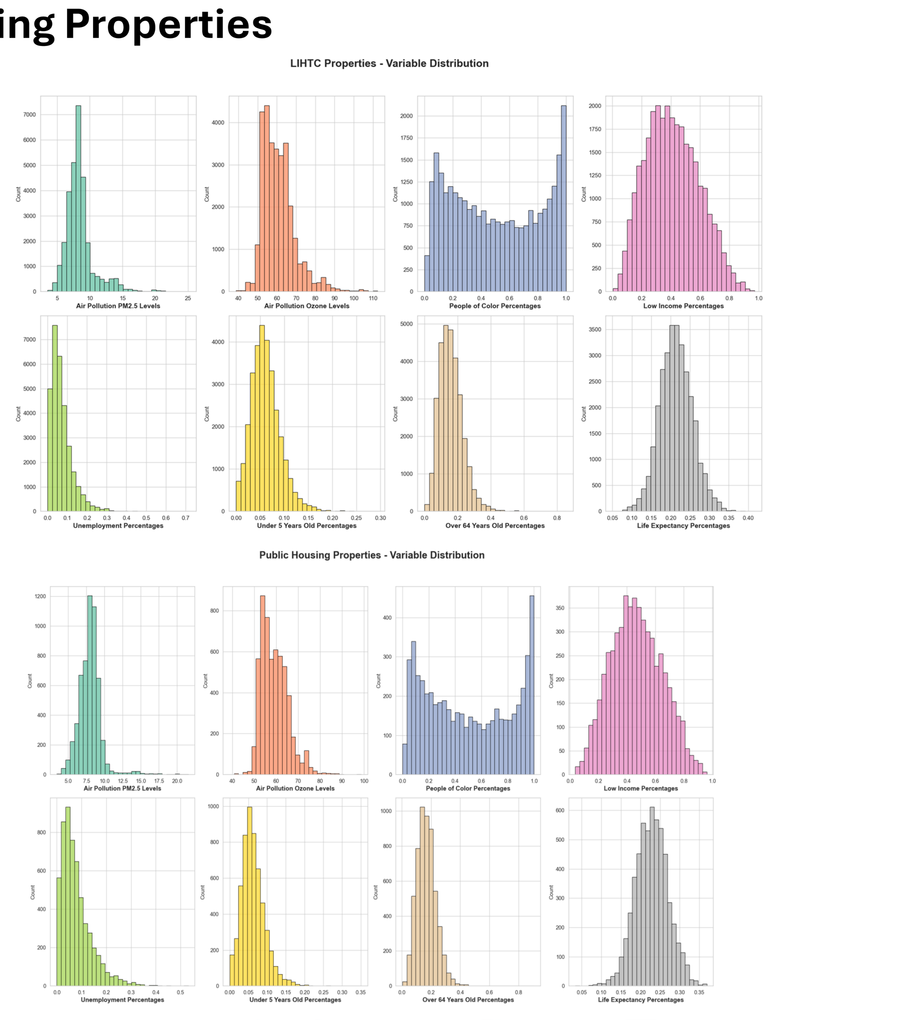
Results

Data Points:

LIHTC : 30321 buildings

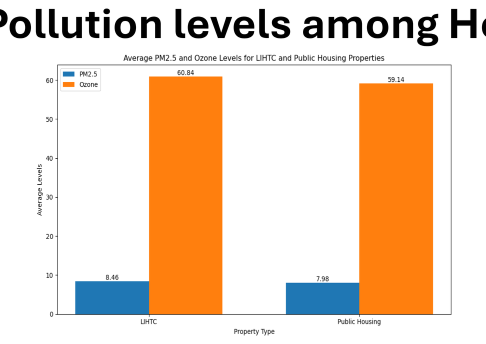
Public Housing: 5601 buildings

Fig. I. Data Distribution of LIHTC and Public Housing Properties



Results

Fig. II. Mean of air Pollution levels among Housing Properties



The P-value is less than 0.001 confirms that the average (mean) exposure is higher for LIHTC properties.

Fig. III. Correlation Matrices of all Features

LIHTC Public Housing

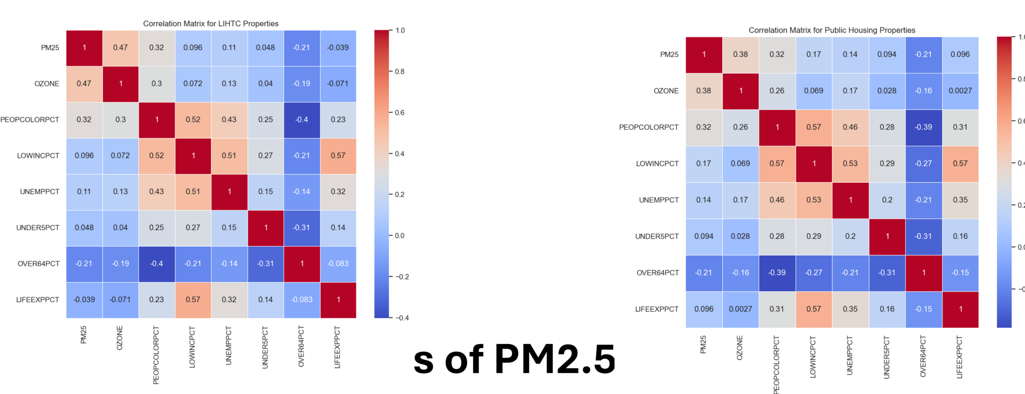


Table. I s of PM2.5 n for LIHTC and Public Housing Programs:

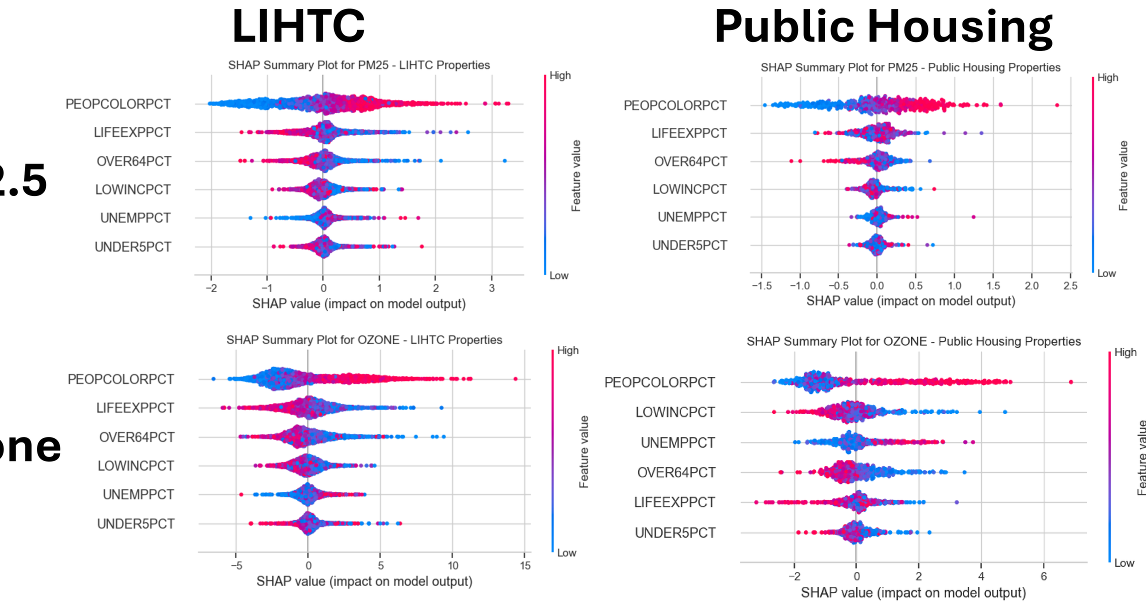
	LIHTC				Public Housing			
	PM 2.5		Ozone		PM2.5		Ozone	
	Linear Regression	Random Forest	Linear Regression	Random Forest	Linear Regression	Random Forest	Linear Regression	Random Forest
Testing R-squared:	0.12	0.54	0.11	0.52	0.10	0.19	0.08	0.19
Testing MSE:	4.26	2.24	64.90	34.96	2.07	1.87	35.86	31.50

Fig. IV. Feature Importances in predicting air pollution levels among affordable housing programs:

LIHTC **Public Housing**

PM2.5

Ozone



Percentages of people of color consistently shows the highest impact on model output, indicating it is a significant predictor for pollution levels. High values contribute positively to the output, suggesting areas with higher percentages of people of color are associated with higher predicted pollution levels.

Fig. V. Classification of air pollution levels among affordable housing Programs:

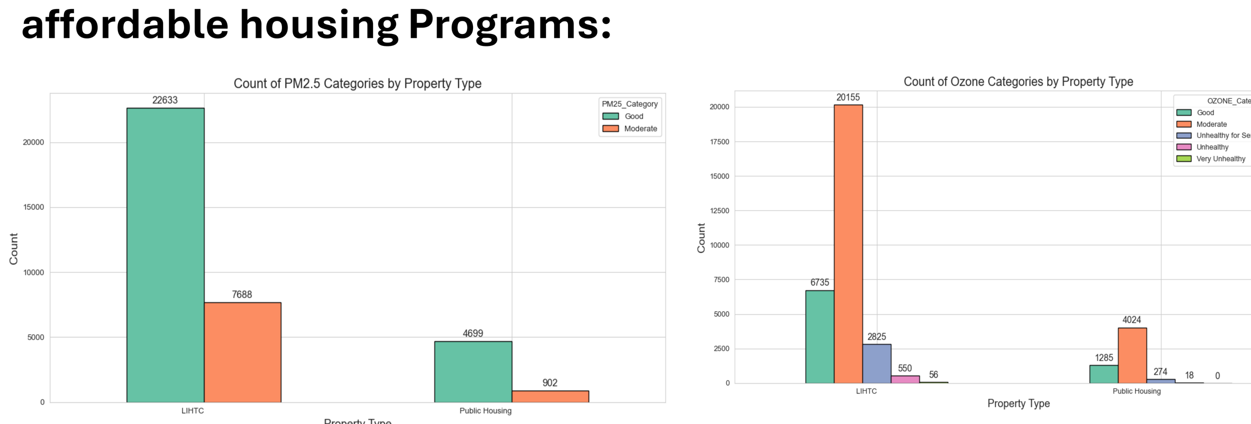
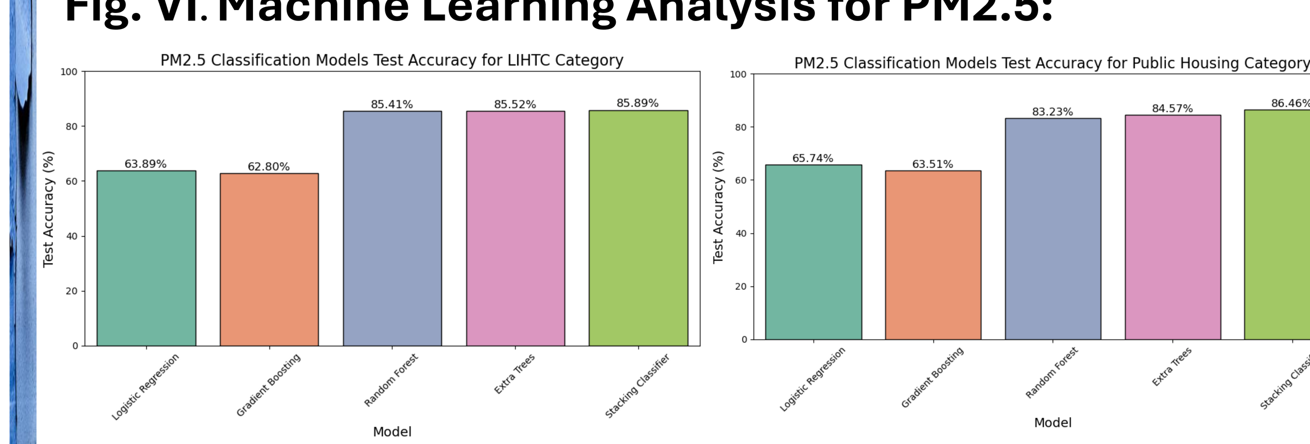


Fig. VI. Machine Learning Analysis for PM2.5:



Results

Fig VII: Machine Learning Analysis for Ozone:

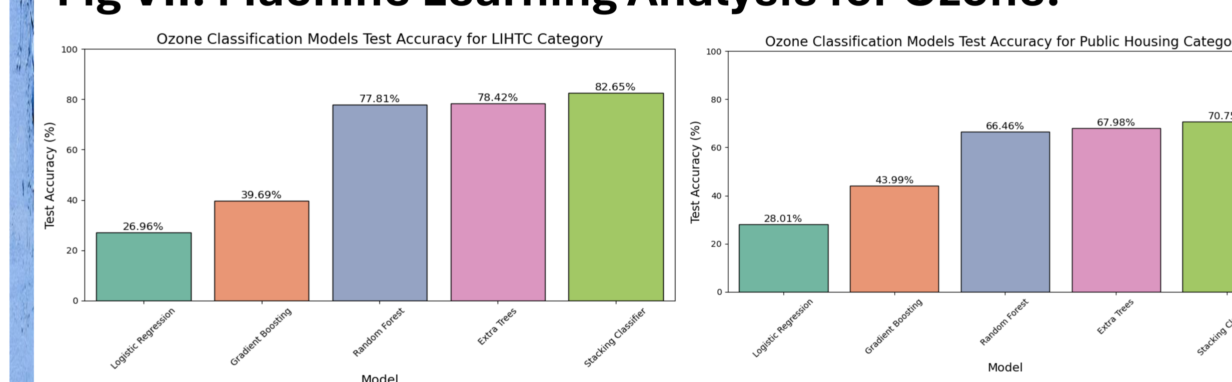


Fig .VII. Spatial Analysis of PM2.5 Exposure in United States

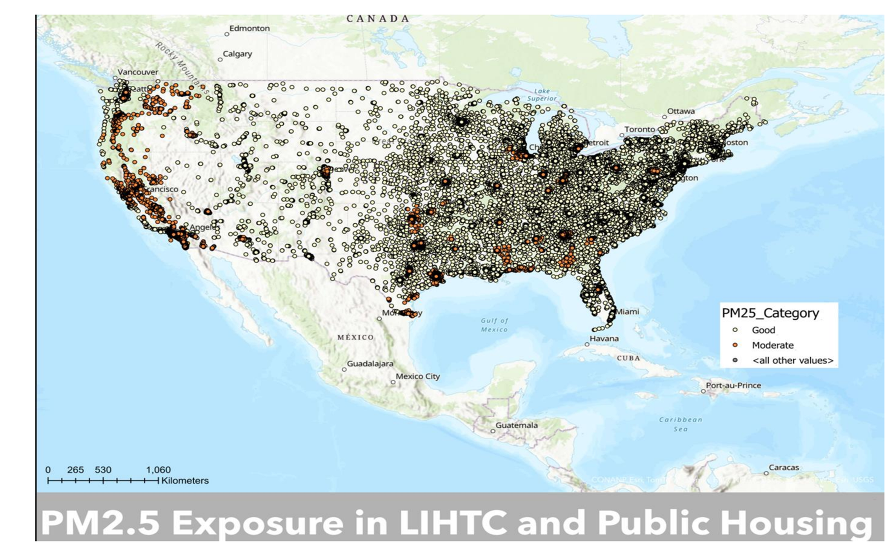
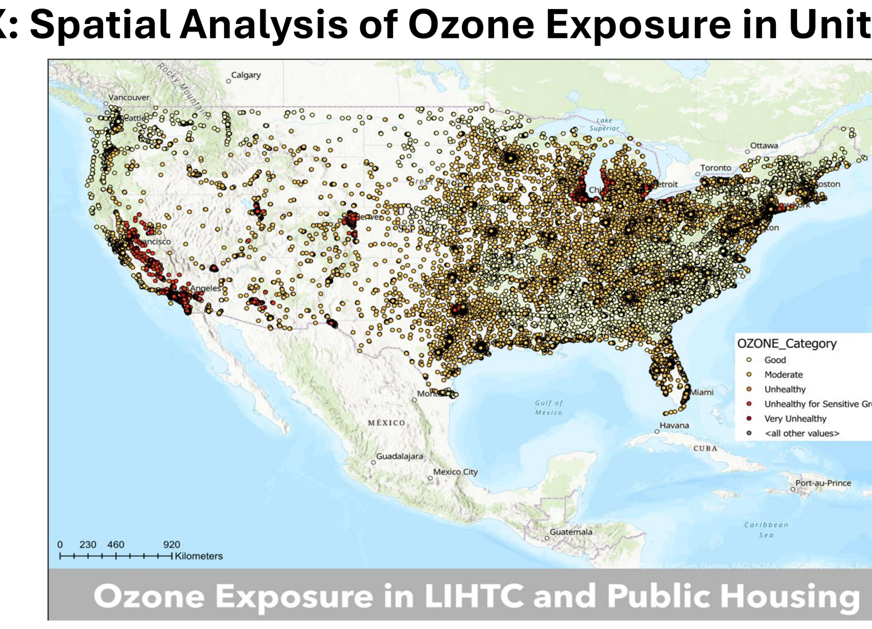


Fig IX: Spatial Analysis of Ozone Exposure in United States



Discussion

- Significant disparities in air pollution exposure were identified between LIHTC and Public Housing properties. LIHTC properties exhibited higher pollution levels on average, highlighting the urgent need for targeted interventions to address these environmental inequalities and protect vulnerable communities.
- SHAP analysis identified the percentage of people of color as a critical predictor of air pollution exposure, emphasizing the role of systemic environmental inequities that disproportionately affect marginalized groups.
- Random Forest regression demonstrated strong predictive capability for air pollution levels, validating its utility in understanding complex relationships between socioeconomic and environmental factors.
- Stacking classifiers demonstrated high accuracy for PM2.5 predictions, indicating robust model performance for particulate matter exposure. However, ozone classification accuracy was comparatively lower, which may reflect the complex dynamics of ozone formation and the inherent challenges in capturing these processes using available features. This highlights the need for additional data and refined features to better model ozone exposure and improve classification performance.
- Spatial analysis pinpointed geographic areas with higher pollution levels where affordable housing is disproportionately affected. These findings highlight significant geographic disparities and provide actionable insights for regional policy focus.
- These findings underscore the critical need for housing policies that integrate environmental equity considerations. By addressing air pollution exposure disparities, policymakers can better safeguard the health and well-being of vulnerable populations residing in federally assisted affordable housing.

Future Work

- Incorporate additional air pollution variables, such as nitrogen dioxide (NO_2) or sulfur dioxide (SO_2), for a more comprehensive understanding of environmental risks.
- Analyze time-series data to track trends in pollution exposure and assess the long-term impacts of housing policies and environmental regulations, such as the Clean Air Act.
- Improve ozone exposure modeling by integrating additional data and refining features to address its unique complexities.
- Evaluate the effectiveness of policies like the Clean Air Act in improving air quality around federally assisted housing and recommend adjustments based on findings.
- Collaborate with local communities and stakeholders to validate findings and incorporate lived experiences for actionable and equitable insights.

References

- Goplerud, D.K., Gensheimer, S.G., Schneider, B.K., Eisenberg, M.D., Smith, G.S. The Spatial Relationship Between the Low-Income Housing Tax Credit Program and Industrial Air Pollution. Cityscape: A Journal of Policy Development and Research, 24(3), 2022. U.S. Department of Housing and Urban Development, Office of Policy Development and Research. https://www.huduser.gov/portal/periodicals/cityscape/vol24num3/ch7.pdf
- U.S. Environmental Protection Agency (EPA). "Air Pollution: Current and Future Challenges." Accessed September 1, 2024. https://www.epa.gov/clean-air-act-overview/air-pollution-current-and-future-challenges
- Anyanwu, C., & Beyer, K. M. M. (2024). Intersections among housing, environmental conditions, and health equity: A conceptual model for environmental justice policy. Social Sciences & Humanities Open, 9, 100845. https://doi.org/10.1016/j.ssaho.2024.100845