Science Journal of Public Health
2024; Vol. 12, No. 1, pp. 9-17
https://doi.org/10.11648/j.sjph.20241201.12

Research Article

Heterogenous Association Between Mortality and Environmental Factors

Andrew Ju*

Choate Rosemary Hall, Wallingford, United States

Abstract

The global environment has changed rapidly since the Industrial Revolution. Human emissions of heat-trapping greenhouse gases have warmed our earth, leading to more climate extremes. Human activities have also caused air pollution and, thus, worse air quality. Warmer climates and polluted air pose severe risks to human health. This paper focuses on temperature and air pollution as the primary environmental factors and studies their relationship with mortality across different ethnicity and age groups in the U.S. from 2001 to 2021. The main research methods employed in this paper are correlation analysis and least-square regressions. This paper finds that, on average, environmental factors are moderately and positively related to total deaths at a statistically significant level. Such a positive relationship still holds when I further investigate how the environment is associated with mortality by each individual cause. Moreover, heterogeneity in the relationship is identified among different races. In particular, temperature seems to have a larger impact on the Native Hawaiian/Pacific Islander and Black populations. At the same time, air pollution is observed to have a very strong association with the mortality of the Asian population, Pacific Islanders, and Hawaiian Natives. Finally, age disparities are not that significant. One finding worth pointing out is that children and teenagers (Age 1-14) appear to be more susceptible to air pollution than other age groups.

Keywords

Environmental Factors, Temperature, Air Quality, Mortality, Heterogeneity, Disparity

1. Introduction

Our living environment has changed dramatically since the Industrial Revolution. Human emissions of greenhouse gases have led to a warmer climate, with the average global temperature being 1.9 degrees Fahrenheit higher than the preindustrial world in the 1800s [1], and is expected to continue to rise for many decades [2]. Human activities have also produced vast amounts of pollutants, worsening air quality.

Numerous studies have shown that environmental changes like warmer climates and air pollution pose increasing risks to human health by affecting the morbidity and mortality rate of many diseases [3], such as cardiovascular diseases, respiratory illnesses, cancer, infectious diseases, psychological illnesses, etc. Research shows that exposure to hot temperatures reduces the birth weight of babies [4] and detriments the respiratory health of children [5]. Warm climates and severe air pollution would increase the chances of virus transmissions [6]. On top of that, the climate impacts people's psychological health [7], and extreme heat may even lead to higher suicidal rates [8]. Pollution would also lead to reduced happiness and higher rates of depression [9]. Moreover, the

*Corresponding author: aju25@choate.edu (Andrew Ju)

Received: 10 November 2023; **Accepted:** 28 November 2023; **Published:** 13 March 2024

health toll imposed by environmental changes varies across different socioeconomic groups [10]. The leading cause of such disparity is material inequality, such as poverty and lack of education and insurance. The evidence can be found in the racial disparities in cancer treatments [11] and mortality rates in pregnancy [12].

This paper focuses on temperature and air pollution as environmental factors and studies their impacts on mortality in the U.S. during the period of 2001-2021. In this paper, I first examine how environmental factors are correlated with total deaths. Then, I take a further step and look into the relationship between the environment and the mortality of different causes. Finally, this paper investigates the possible heterogeneity in the relationship between environment and mortality across different races and age cohorts. The findings of this project would inspire better policymaking regarding environment protection, public health, and the allocation of resources among various ethnicities and age groups.

The remainder of this paper is organized as follows: Part 2 introduces the data and methods used in this paper. Part 3 reports the research results. Part 4 concludes.

2. Data and Methods

2.1. Data

This paper uses three datasets: CDC WONDER, a national mortality database [13]; a state-level monthly temperature database from NOAA's National Centers for Environmental Information [14]; and air quality data from the U.S. Environmental Protection Agency (EPA) [15].

2.1.1. CDC WONDER Dataset

The CDC WONDER database provides detailed national mortality information, where mortality is defined as the number of deaths. Specifically, the CDC WONDER dataset records mortality by state, death year, gender, age range, race, and cause of death. The data sample used in this paper ranges from 2001 to 2021.

2.1.2. NOAA Temperature Dataset

This paper obtains statewide average and minimum temperatures in a particular year from the NOAA's National Centers for Environmental Information. Temperature data is then merged with the CDC WONDER dataset by state and year to analyze the relationship between temperature and mortality.

2.1.3. EPA AQI Dataset

The EPA provides daily county-level Air Quality Index (AQI) as a measure of air quality. Every day, the EPA calculates an index for each of the five major air pollutants regulated by the Clean Air Act: ground-level ozone, particle pollution, carbon monoxide, sulfur dioxide, and nitrogen dioxide. Whichever pollutant has the highest index is identified as the primary pollutant of that day, and its index value is recorded as the AQI value of that day. The AQI typically runs from 0 to 500. In more detail, the EPA divides AQI values into six categories corresponding to different levels of health concern. A "Good" day has an AQI at or below 50, which means the pollution has little risk to human health. "Moderate" days have AQI values running from 51 to 100, indicating that the pollution is acceptable but may pose a moderate risk to susceptible people. "Unhealthy for Sensitive Groups" days have AQI values varying from 101 to 150 when old and young people and people with respiratory diseases face greater pollution hazards. An "Unhealthy" Day has an AQI from 151 to 200, and most people are beginning to experience adverse health effects. A "Very Unhealthy" day has an AQI between 201 and 300, and a "Hazardous" day higher than 300.

Besides reporting daily AQI values, the EPA also summarizes annual max and median AQI values for each county. In addition, the EPA AQI dataset reports the number of days a county has air quality data in one particular year and the number of days each of the five major pollutants is the most responsible for air pollution in that year. Based on the EPA AQI data, I construct annual AQI values at the state level, which can be merged with the CDC WONDER dataset and enable me to study the relationship between air quality and mortality.

2.2. Methods

This paper first provides descriptive statistics of environmental factors, such as temperature and air quality (AQI), throughout 2001-2021. In particular, six representative states are selected to show how the environment has improved or worsened since 2001.

Then, this paper performs correlation analysis to establish the direction and strength of the relationship between environmental factors and mortality. More specifically, I calculate the correlation coefficients between mortality and various proxy variables for environmental conditions, like temperature, AQI values, probability of unhealthy days, and probability of ozone days (i.e., those days ozone is identified as the primary pollutant). Furthermore, I calculate the correlation coefficients for each race/age group in the CDC WONDER dataset to explore the possible heterogeneity in the relationship between environmental factors and mortality.

Finally, this paper conducts least-square regression analysis to further study the relationship between the environment and the mortality by different causes of death. The regression model is represented as follows:

$$y = a + bx + \varepsilon,$$

Where x denotes environmental factors like temperature and AQI, y represents the mortality resulting from one of the

15 leading causes as identified in the CDC WONDER database, and ε is the error term. The regression estimator of the coefficient b indicates how many more/fewer deaths are associated with one unit increase/decrease in temperature or AQI values. Regression results are presented in figures for the easiness of comparison. All the data analysis in this paper is conducted by using R.

3. Results

3.1. Descriptive Statistics

Figure 1 shows the trend of average temperature in Fahrenheit from 2001 to 2021 on a national level. Each dot represents the national average temperature of that year, which is calculated by taking the average of the 12-month average temperature of all states as provided by the NOAA Temperature database. The graph shows that even though the average temperature only increased slightly from 2001 to 2021, it seems to become more volatile since 2011, which coincided with global warming and more frequent climate extremes.

Figure 2 presents the annual air quality trend in the U.S. from 2001 to 2021, with each dot representing the national average of annual max AQI of all counties in that year as reported in the EPA AQI database. The graph shows an apparent downward tendency in the max AQI, which implies a gradual improvement in air quality over 2001-2021.

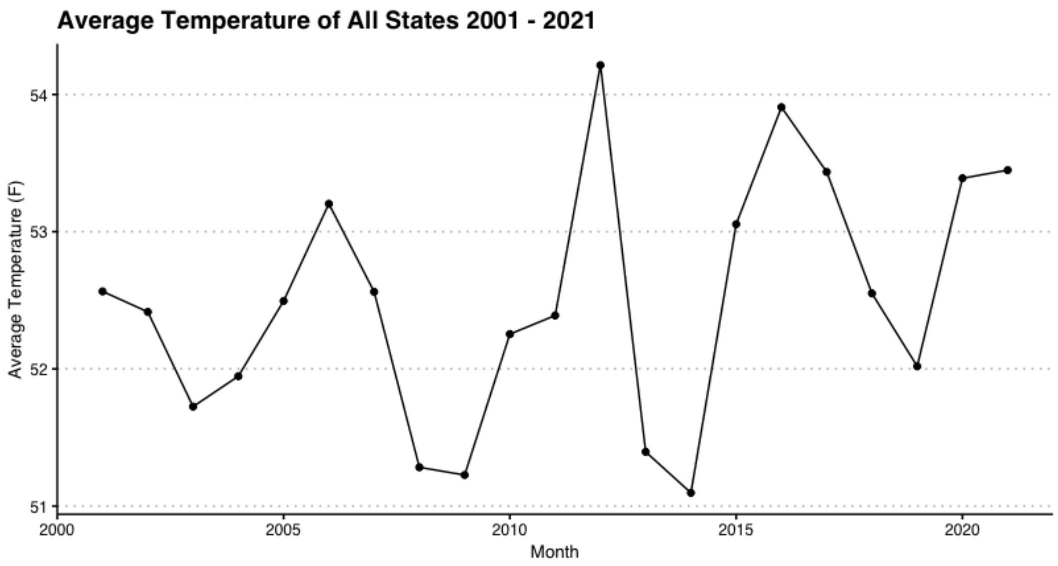

Figure 1. *The Trend of Average Temperature.*

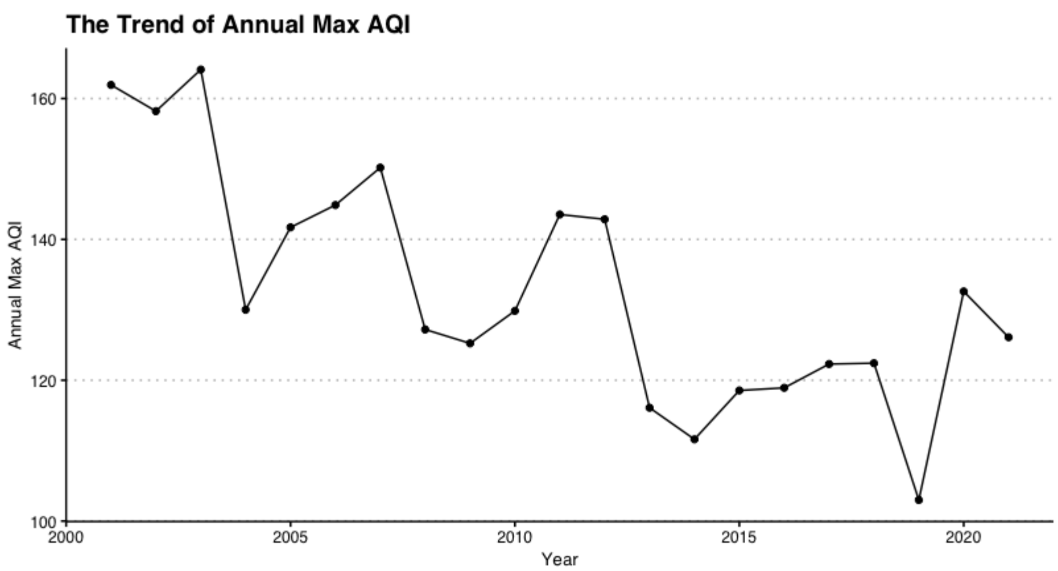

Figure 2. *The Trend of Annual Max AQI.*

Figure 3 shows the development of air quality in six representative states that experienced the most or the slightest improvement in air quality over the period 2001-2021 in terms of two criteria: the average chance of having a "Good" day (i.e., AQI is no larger than 50) and the average chance of having an “Unhealthy for Sensitive Groups” day (i.e., AQI is between 101 and 150). At the county level, the annual chance of having a "Good"/"Unhealthy for Sensitive Groups" day is calculated by dividing the number of days of which the AQI is no larger than 50/between 101 and 150 by the total number of days with AQI data in that year, and is presented in percentage points. By taking the average of these county-level data, I obtain the annual chance of having a "Good"/"Unhealthy for Sensitive Groups" day on a state level. Then, I select the three states that saw the most remarkable/minor improvement in air quality by both criteria from 2001 to 2021. As shown in Figure 3, Delaware, Pennsylvania, and Texas all saw significant improvements in their air quality over the past two decades, which may be attributed to the fact that these states were under worrisome and unhealthy air quality conditions at the beginning of this period. North Dakota, Montana, and California experienced the most minor improvement in air quality over 2001-2021. Among them, however, California remained relatively stable and has more room to improve, which is understandable considering its massive scale. Surprisingly, North Dakota and Montana have been witnessing their air quality getting worse over the past twenty years, although they have a relatively small population and were in a much better condition in 2001.

Among the five major air pollutants the Clean Air Act regulates, ozone is the chief pollutant responsible for unsatisfactory air quality. Figure 4 shows the overall trend of the annual average chance of having an “ozone day” (i.e., ozone is most responsible for air pollution on that day) in percentage points from 2001 to 2021. Each dot represents the average chance of ozone days occurring in the nation that year. More specifically, the county-level chance of having ozone days is calculated by dividing the number of ozone days by the number of days with AQI data in that county. The national average chance is obtained by taking the average of the county values. Figure 4 indicates an upward drift in the probability of having ozone days throughout 2001-2021, except for a slight dip during the COVID-19 pandemic. Moreover, ozone has been responsible for most of the air pollution in the U.S. since 2010.

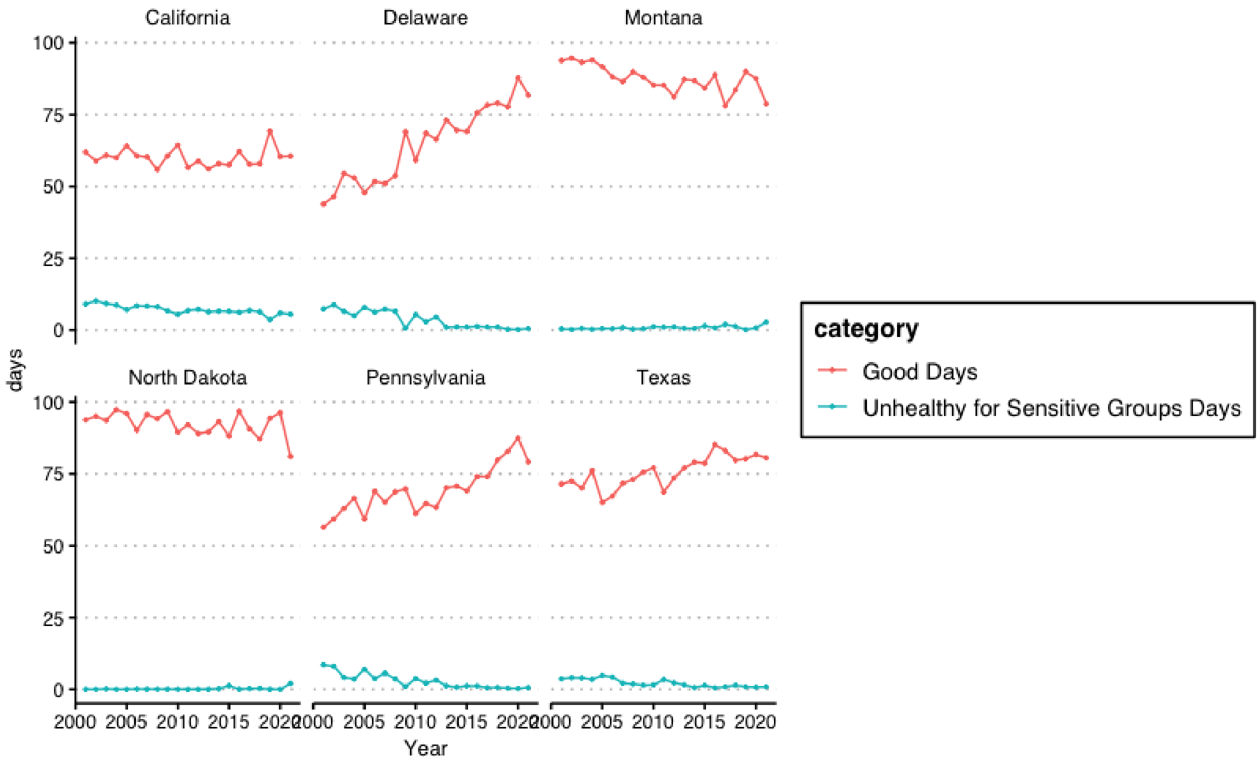

Figure 3. *Good Days and Unhealthy Days in Six Representative States.*

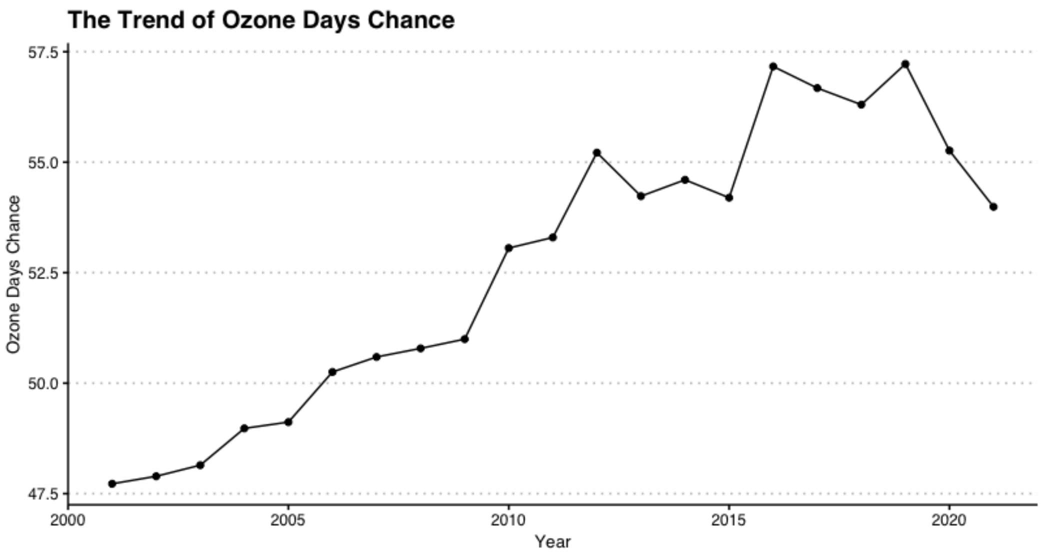

Figure 4. The Chance of Ozone Days.

3.2. The Relationship Between Environment and Mortality

Table 1 reports the correlation coefficients between various measures of environmental conditions and mortality, where the environment is proxied by temperature and air pollution, and mortality is the total number of deaths. Annual data at the state level is used to conduct the correlation analysis. According to Table 1, there is a positive relationship between environment and mortality, which is moderate in magnitude (except for ozone days) and significant at statistical levels. Moreover, such a positive relationship is strongly consistent across all environmental variables. For example, the correlation coefficients between mortality and the two measures of temperature are very similar in magnitude. This also holds for the correlation between mortality and air pollution variables represented by AQI. The probability of having an ozone day is weakly correlated with mortality, which is reasonable given ozone is only one of the five major pollutants tracked by AQI. Overall, Table 1 reveals a significantly positive correlation between mortality and environmental conditions. A worse environment is typically associated with more deaths. Further actions need to be taken to protect the environment so that it is more livable for human beings.

Table 1. Correlation between Environment and Mortality.

Measures of Environmental Condition		Correlation Coefficients
Temperature	Average temperature	0.409 (14.38)
	Minimum temperature	0.403 (14.10)
Air Pollution	Highest value of median AQI	0.462 (16.69)
	Highest value of AQI	0.413 (14.52)
	Highest probability of having unhealthy days for sensitive groups	0.449 (16.09)
	Average probability of having an ozone day	0.270 (9.000)

Note: T-statistics are indicated under the correlation coefficients in parenthesis.

Figures 5-7 describe the results of least-square regressions of deaths resulting from one of the 15 leading causes on various environmental factors. For the sake of space, this paper only reports the results for the three environmental factors: the average temperature, the highest value of AQI, and the highest value of max AQI; and the ten death causes that have the largest coefficients on environment among the 15 leading causes.

Figure 5 presents the regression results of mortality by different causes, where each dot indicates the coefficient on average temperature. On the whole, temperature is found to be positively associated with mortality of various causes, which corroborates the findings of correlation analysis. For example, "Diseases of the Heart" has the largest regression coefficient of 621.6, meaning that on average one Fahrenheit increase in temperature is associated with an additional 621.6 deaths due to heart diseases. The second and third significant coefficients in terms of magnitude are from the regressions of "Malignant Neoplasms" and "Ischemic Heart Diseases", respectively. The regression results are consistent with the fact that these three types of diseases are identified as the top three causes of death in the U.S.

Figure 6 and Figure 7 depict the regression results when air quality measures like the highest value of AQI and the chance of having ozone days are used as environmental factors. The results are very similar to and consistent with those in Figure 5. One thing to note is that, compared to Figure 5, "Chronic Lower Respiratory Diseases" appears among the top ten death causes that have the most significant regression coefficients on AQI or probability of ozone days. This finding may imply that air pollution does pose a higher death risk to people suffering from respiratory diseases.

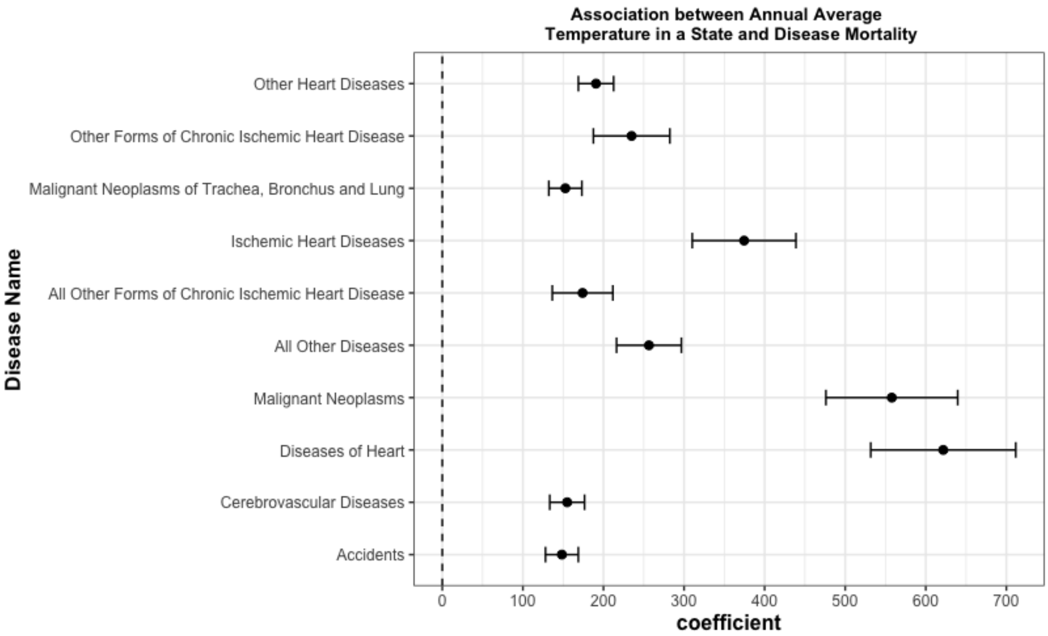

Figure 5. *Linear Regression Results on the Average Temperature.*

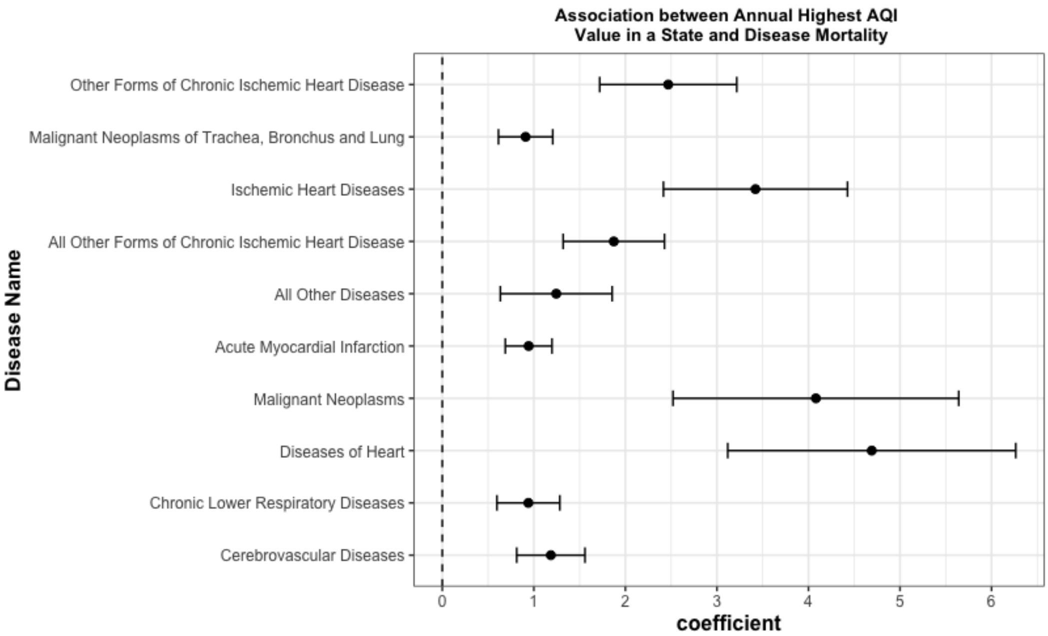

Figure 6. *Linear Regression Results on the Max AQI.*

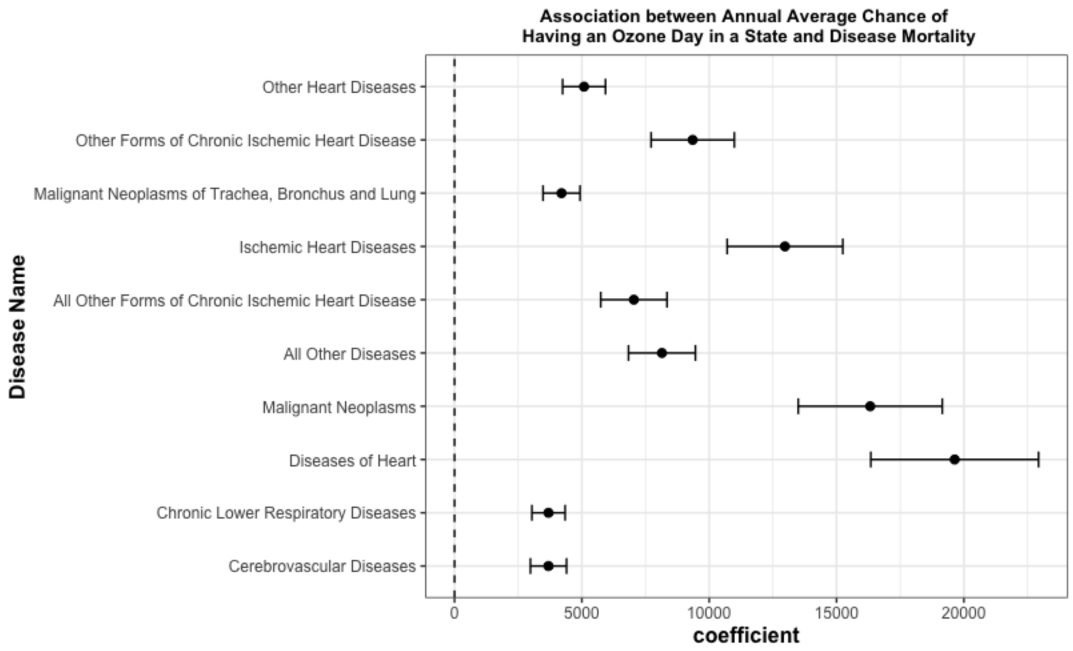

Figure 7. *Linear Regression Results on Ozone Days.*

3.3. Heterogeneity in the Relationship Between Environment and Mortality

This subsection conducts correlation analysis for different ethnicity/age groups to explore the possible heterogeneity in the relationship between environment and mortality.

Table 2 reports the correlation coefficients between various environmental factors and mortality for different races. In line with the results in Table 1, the number of deaths in each ethnicity group is positively associated with the environment. However, heterogeneity exists among different races, as reflected by the difference in the magnitude of coefficients. For example, temperature seems significantly related to mortality for Native Hawaiian/Pacific Islander and Black populations while exhibiting a much weaker correlation with mortality for either the Asian or the Asian/Pacific Islander group. Air pollution, as measured by AQI, on the other hand, very strongly correlates with the total deaths of the Asian population, Pacific Islanders, and Hawaiian Natives. However, such a relationship appears milder for the Black group. Ozone overall has a weak correlation with the total deaths of all races. One exception might be that the correlation coefficient for the Black, which is statistically significant, is greater in magnitude compared to other races.

Table 3 presents the correlation coefficients between environment and mortality for different age groups. In general, environmental factors, except for ozone days, are moderately positively related to the number of deaths across all age groups, which does not exhibit sizable age disparities. One thing to note is that children and teenagers (Age 1 - 14) seem most affected by air pollution: this age group's mortality is strongly associated with the highest AQI value and the highest probability of unhealthy days for sensitive people. Therefore, reducing air pollution might be particularly beneficial for children and teenagers.

Table 2. *Correlation Analysis by Ethnicity Groups.*

Measures of Environmental Condition		**Correlation Coefficients by Ethnicity Groups**					
		Asian	**American Indian/Alaska Native**	**Black**	**White**	**Asian / Pacific Islander**	**Native Hawaiia / Pacific Islander**
Temperature	Average temperature	0.106 (1.187)	0.378 (7.767)	0.439 (13.96)	0.388 (13.51)	0.106 (2.535)	0.501 (2.005)
	Minimum temperature	0.179 (2.015)	0.367 (7.494)	0.400 (12.47)	0.382 (13.24)	0.179 (4.139)	0.584 (2.495)

Measures of Environmental Condition		Correlation Coefficients by Ethnicity Groups					
		Asian	American Indian/Alaska Native	Black	White	Asian / Pacific Islander	Native Hawaiia / Pacific Islander
Air Pollution	Highest value of median AQI	0.513 (6.628)	0.443 (9.391)	0.167 (4.856)	0.458 (16.50)	0.513 (15.75)	0.889 (6.733)
	Highest value of AQI	0.805 (15.04)	0.187 (3.625)	0.181 (5.263)	0.403 (14.12)	0.805 (21.86)	0.834 (5.243)
	Highest probability of having unhealthy days for sensitive groups	0.692 (10.64)	0.287 (5.728)	0.205 (6.004)	0.441 (15.73)	0.692 (19.98)	0.934 (9.059)
	Average probability of having an ozone day	0.054 (0.600)	0.164 (3.156)	0.341 (10.39)	0.265 (8.824)	0.112 (2.363)	0.367 (1.367)

Note: T-statistics are indicated under the correlation coefficients in parenthesis.

***Table 3**. Correlation Analysis by Age Groups.*

Measures of Environmental Condition		Correlation Coefficients by Age Groups				
		Age 1	Age 1-14	Age 15-34	Age 35-64	Age 65+
Temperature	Average temperature	0.445 (15.48)	0.349 (9.740)	0.485 (17.78)	0.472 (17.16)	0.385 (13.37)
	Minimum temperature	0.427 (14.70)	0.371 (10.45)	0.483 (17.69)	0.463 (16.76)	0.379 (13.11)
Air Pollution	Highest value of median AQI	0.503 (18.13)	0.489 (14.70)	0.485 (17.77)	0.470 (17.06)	0.450 (16.13)
	Highest value of max AQI	0.442 (18.13)	0.549 (14.70)	0.418 (17.77)	0.407 (17.06)	0.406 (16.13)
	Highest probability of having unhealthy days for sensitive groups	0.487 (17.35)	0.577 (18.50)	0.460 (16.60)	0.436 (15.53)	0.441 (15.78)
	Average probability of having an ozone day	0.214 (6.850)	0.209 (5.587)	0.252 (8.360)	0.264 (8.786)	0.271 (9.048)

Note: T-statistics are indicated under the correlation coefficients in parenthesis.

4. Conclusion

In summary, the U.S.'s climate is gradually getting warmer and more volatile from 2001 to 2021. Air pollution has been reduced, and air quality is in general improving. Temperature and air quality, as the main environmental measures employed in this paper, are found to be positively correlated with total mortality, and such a correlation is moderate in magnitude and statistically significant. In particular, cardiovascular diseases, respiratory diseases, and cancer deaths are all positively associated with environmental impacts. Furthermore, racial disparities exist in the relationship between environment and mortality, where Hawaiian Natives, Pacific Islanders, and Black people seem more susceptible to temperature changes. At the same time, Asians, Pacific Islanders, and Hawaiian Natives are more affected by air pollution. Finally, for each age group, moderate correlations between environment and mortality are generally observed, and there is no pronounced heterogeneity across age cohorts. However, air pollution does appear to be particularly harmful to non-infant children (Age 1-14).

A few limitations of this research project are worth mentioning to avoid ambiguity and provide directions for future research. First, the data used in this paper was only aggregate information. Individual characteristics such as socioeconomic status, educational background, gender, etc., are not considered and may confound the relationship between environment and mortality. Second, this research focuses on mortality as the health impacts that temperature and air pollution may have on people. However, death is an extreme and rare outcome compared to morbidity, and therefore, the association between environment and mortality reported in this paper might not be able to capture the overall effects of environmental changes on human health. Finally, this paper does not try to establish a causal relationship between environment and mortality, and caution should be taken when interpreting the findings of this research.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] NASA. (2023a). *World of change: Global temperatures*. NASA. https://earthobservatory.nasa.gov/world-of-change/global-temperatures#:~:text=According%20to%20an%20ongoing%20temperature,1.9%C2%B0%20Fahrenheit)%20since%201880

[2] NASA. (2023b). *The effects of climate change*. NASA. https://climate.nasa.gov/effects/

[3] Patz, J. A., Campbell-Lendrum, D., Holloway, T., & Foley, J. A. (2005). Impact of regional climate change on human health. *Nature*, *438* (7066), 310–317. https://doi.org/10.1038/nature04188

[4] Deschenes, O., Greenstone, M., & Guryan, J. (2009). Climate change and birth weight. *American Economic Review*, *99* (2), 211–217. https://doi.org/10.1257/aer.99.2.211

[5] Beatty, T. K. M., & Shimshack, J. P. (2014). Air pollution and children's respiratory health: A cohort analysis. *Journal of Environmental Economics and Management*, *67* (1), 39–57. https://doi.org/10.1016/j.jeem.2013.10.002

[6] Manoj, M. G., Satheesh Kumar, M. K., Valsaraj, K. T., Sivan, C., & Vijayan, S. K. (2020). Potential link between compromised air quality and transmission of the novel corona virus (SARS-COV-2) in affected areas. *Environmental Research*, *190*, 110001. https://doi.org/10.1016/j.envres.2020.110001

[7] Obradovich, N., Migliorini, R., Paulus, M. P., & Rahwan, I. (2018). Empirical evidence of mental health risks posed by climate change. *Proceedings of the National Academy of Sciences*, *115* (43), 10953–10958. https://doi.org/10.1073/pnas.1801528115

[8] Carleton, T. A. (2017). Crop-damaging temperatures increase suicide rates in India. *Proceedings of the National Academy of Sciences*, *114* (33), 8746–8751. https://doi.org/10.1073/pnas.1701354114

[9] Zhang, X., Zhang, X., & Chen, X. (2015). Happiness in the air: How does a dirty sky affect subjective well-being? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2655352

[10] Braveman, P. A., Egerter, S. A., & Mockenhaupt, R. E. (2011). Broadening the focusthe need to address the social determinants of health. *American Journal of Preventive Medicine*, *40* (1). https://doi.org/10.1016/j.amepre.2010.10.002

[11] Ward, E., Jemal, A., Cokkinides, V., Singh, G. K., Cardinez, C., Ghafoor, A., & Thun, M. (2004). Cancer disparities by race/ethnicity and socioeconomic status. *CA: A Cancer Journal for Clinicians*, *54* (2), 78–93. https://doi.org/10.3322/canjclin.54.2.78

[12] Harper, M. A., Espeland, M. A., Dugan, E., Meyer, R., Lane, K., & Williams, S. (2004). Racial disparity in pregnancy-related mortality following a live birth outcome. *Annals of Epidemiology*, *14* (4), 274–279. https://doi.org/10.1016/s1047-2797(03)00128-5

[13] CDC WONDER. (2023). *Multiple cause of death data on CDC Wonder*. Centers for Disease Control and Prevention. https://wonder.cdc.gov/mcd.html

[14] NOAA National Centers for Environmental information. (2023). *Statewide time series: Climate at a glance*. Statewide Time Series | Climate at a Glance | National Centers for Environmental Information (NOAA). https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/statewide/time-series

[15] EPA. (n.d.). Air Data: Air Quality Data Collected at Outdoor Monitors Across the US. https://www.epa.gov/outdoor-air-quality-data