

# CalEnviroScreen

Mia Forsline

2021-12-02

## Research Question

In California from 2013 - 2021, does air quality (as measured by annual mean PM2.5 concentrations per census tract) vary with poverty rates (as measured by the percent of the population living below two times the federal poverty level per census tract)?

## Background

*Why is this important? Is there existing evidence on this question? If so, why is it inconclusive? If not, why not?*

In California, events like wildfires can significantly decrease air quality by releasing fine particles such as particulate matter, or PM2.5 (Shi et al., 2019). PM2.5 refers to particles with diameters  $\leq 2.5 \mu\text{m}$ , which are known to be hazardous for human health. They are especially detrimental for our respiratory and cardiovascular health (Cleland et al., 2021). As California's wildfires continue to worsen over time, it is becoming increasingly important to monitor air quality, PM2.5 concentrations, and their impacts on populations (Gupta et al., 2018).

The environmental burden of poor air quality is not shared equally across the state. For example, the San Joaquin Valley's economically disadvantaged and ethnically diverse communities breathe some of the most polluted air in the nation (Cisneros et al., 2017). As a result, vulnerable communities such as Mexican American immigrant farm workers and their families experience disproportionately high rates of asthma attacks, hospital admissions, and other medical issues (Schwartz & Pepper, 2009). This inequitable pattern is well recognized in other states (Qian & Wu, 2019), the United States overall (Tessum et al., 2021), and even other countries (Li et al., 2018).

While there are many possible ways to explore the inequity of air pollution in California, I specifically use annual mean PM2.5 to measure of air quality and poverty rate to quantify socioeconomic disparities.

## Hypotheses

My null hypothesis ( $H_0$ ) is that, in California, there is no relationship between annual mean PM2.5 concentrations per census tract and percent of the population living below twice the federal poverty line per census tract.

My alternative hypothesis ( $H_A$ ) is that, in California, there is a relationship between annual mean PM2.5 concentrations per census tract and percent of the population living below twice the federal poverty line per census tract

## Data Description and Collection

*Where did you access it? What are its spatial and temporal features? What are its limitations? What do you know about the sampling strategy and what biases that may introduce? If helpful, you can use a histogram, scatterplot, or summary statistics table to describe your data.*

I downloaded 2013 - 2021 CalEnviroScreen data from the California Office of Environmental Health Hazard Assessment (OEHHA) and the California Open Data Portal :

- Versions 1.1 (Sept 2013)
- Versions 2.0 (Oct 2014)
- Versions 3.0 (June 2018)
- Versions 4.0 (Oct 2021)

Each CalEnviroScreen dataset contains columns of:

- environmental pollution burden indicators
  - exposure indicators such as PM2.5
  - environmental effects indicators such as groundwater threats
- population characteristics:
  - sensitive population indicators such as rates of asthma
  - socioeconomic factor indicators such as rates of poverty

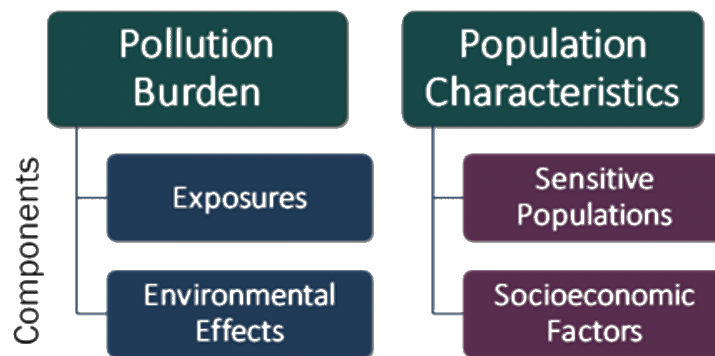


Figure 1: CalEnviroScreen Components (Source: <https://oehha.ca.gov/media/downloads/calenviroscreen/report/calenviroscreen40reportf2021.pdf>)

Each dataset also contains one row per census tract, meaning each census tract in California is assigned a value per environmental indicator or population characteristic.

In particular, I am interested in:

## 1. PM2.5

The annual mean concentration of PM2.5, which CalEnviroScreen (CES) calculates based on a weighted average of measured monitor concentrations and satellite observations ( $\mu\text{g}/\text{m}^3$ ) over 3 years. For example, the CES 1.1 report used data from 2007 - 2009 while the CES 4.0 report used 2015 - 2017. All reports used data from the California Air Resources Board's Air Monitoring Network. In addition to the Air Monitoring Network, CES 3.0 and CES 4.0 incorporated Satellite Remote Sensing Data.

Data were more likely to be high resolution around certain cities or localized areas, but not all cities have air monitoring stations. Depending on the year, locales with little to no data were either omitted or estimated using nearby locations' data. For example, in CES 1.1, census tracts with centers  $> 50\text{km}$  away from the nearest air monitor were omitted from the analysis. In CES 4.0, missing data was estimated using regression relationships with nearby sites.

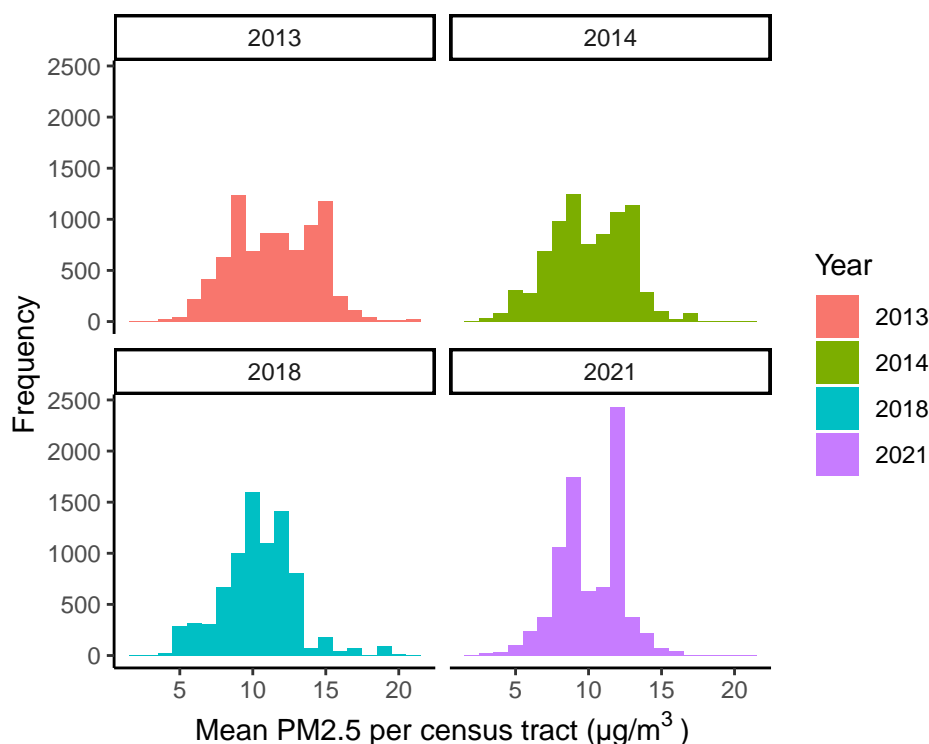


Figure 2: 2013 - 2021 mean PM2.5 in California was not normally distributed. In 2013, 2014, 2018, and 2021, the annual mean concentrations of PM2.5 ( $\mu\text{g}/\text{m}^3$ ) per census tract were 11.52, 10.01, 10.38, and 10.15 respectively. Data was sourced from CalEnviroScreen 1.1 - 4.0 (<https://oehha.ca.gov/>).

PM2.5 annual mean monitoring data were extracted from all monitoring sites where possible. For CES 4.0, PM2.5 annual mean concentrations were also calculated using Aerosol Optical Depth measurements as well as land use and meteorology data via regression on ground monitor data.

For CES 1.1 - 3.0, the geographic center of each census tract, quarterly mean PM2.5 concentrations were estimated using ordinary kriging. For CES 4.0, overall PM2.5 annual mean concentrations were estimated for each 1km x 1km grid cell using both the monitoring and satellite data in a weighted average. An inverse-distance weighting method was used, so grid cells close to monitors relied more heavily on monitor estimates while grid cells further from monitors relied more heavily on satellite data. Grid cells with monitors > 50km away relied solely on satellite data. Concentrations were estimated at the center of each 1km x 1km grid cell.

The quarterly estimates were then averaged to calculate annual means. The annual means were estimated over 3 years to avoid account for uneven sampling frequency.

## 2. Poverty

The percent of the population living below two times the federal poverty level, which CalEnviroScreen calculates using a 5-year estimate. For example, the CES 1.1 report used a 5-year estimate from 2007 - 2011 data while the CES 4.0 report used a 5-year estimate from 2015 - 2019 data. Poverty data came from the American Community Survey. Multiple years of data are used to calculate more reliable results for geographic areas with small populations.

CalEnviroScreen defined poverty as twice below the federal poverty line to account for California's high cost of living relative to other states and because the federal poverty threshold has not changed since the 1980s despite the cost of living increasing over time.

The percent per census tract was calculated by individuals living below 200% the poverty level per census tract / total individuals living below 200% of the poverty level.

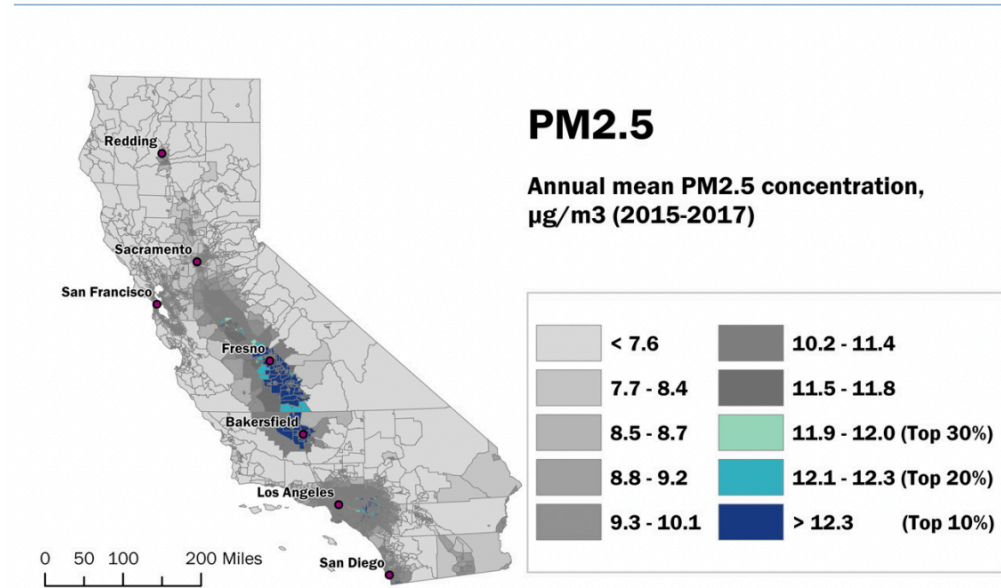


Figure 3: CalEnviroScreen 4.0 PM2.5 Map (Source: <https://oehha.ca.gov/media/downloads/calenviroscreen/report/calenviroscreen40reportf2021.pdf>)

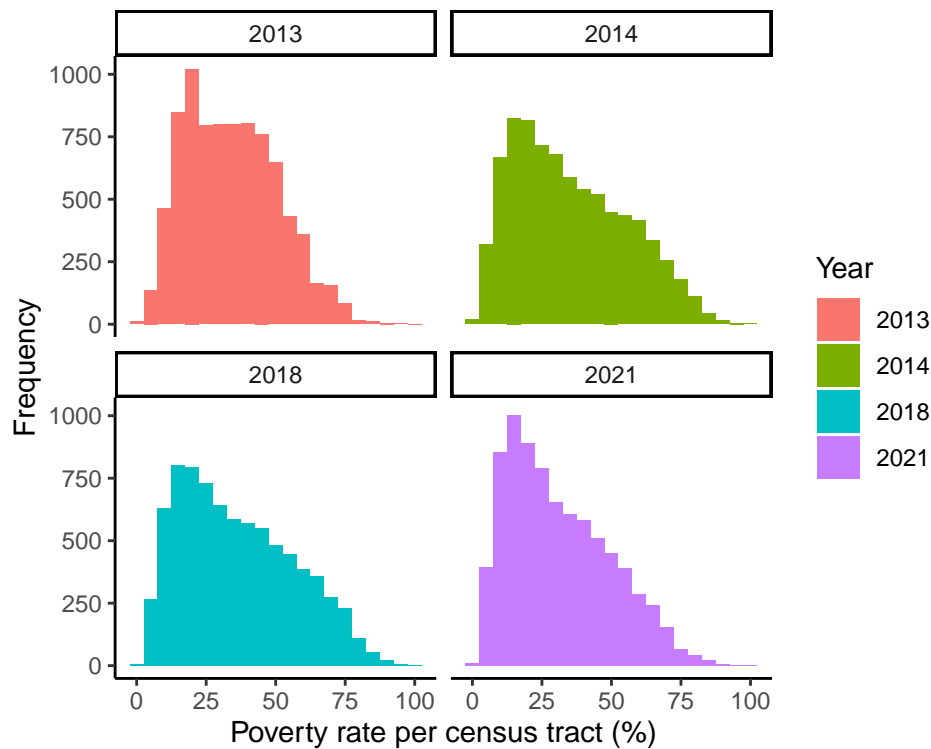


Figure 4: 2013 - 2021 poverty rates in California were not normally distributed. In 2013, 2014, 2018, and 2021, the mean percentages of the population per census tract living below two times the federal poverty level were 34.24%, 35.28%, 36.39%, and 31.34% respectively. Data was sourced from CalEnviroScreen 1.1 - 4.0 (<https://oehha.ca.gov/>).

Standard error and relative standard error were calculated to determine the reliability of the calculated poverty rate. Census tracts with unreliable estimates were assigned no value for poverty rate (NULL).

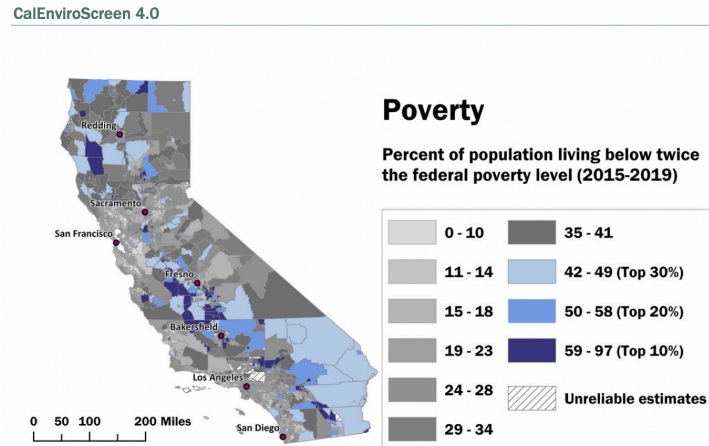


Figure 5: CalEnviroScreen 4.0 Poverty Map (Source: <https://oehha.ca.gov/media/downloads/calenviroscreen/report/calenviroscreen40reportf2021.pdf>)

## Methods - Statistical Analysis Plan

*What is your analysis plan? Why did you choose this analysis, given your data and question? What are the limitations?*

To assess if, in California from 2013 - 2021, air quality varies with poverty rates, I plan to run linear regressions of PM2.5 ~ Poverty for each year (2013, 2014, 2018, 2021). This analysis is appropriate to describe how air quality might be changing with respect to poverty rates. Running multiple regressions over the different years can help us understand how this relationship might be changing over time.

This method is limited by the fact that I am only including one independent variable (Poverty) in the model. It is likely that there are many different factors in addition to poverty influencing air quality, but this analysis is a strong starting point for detangling those complex relationships.

## Results

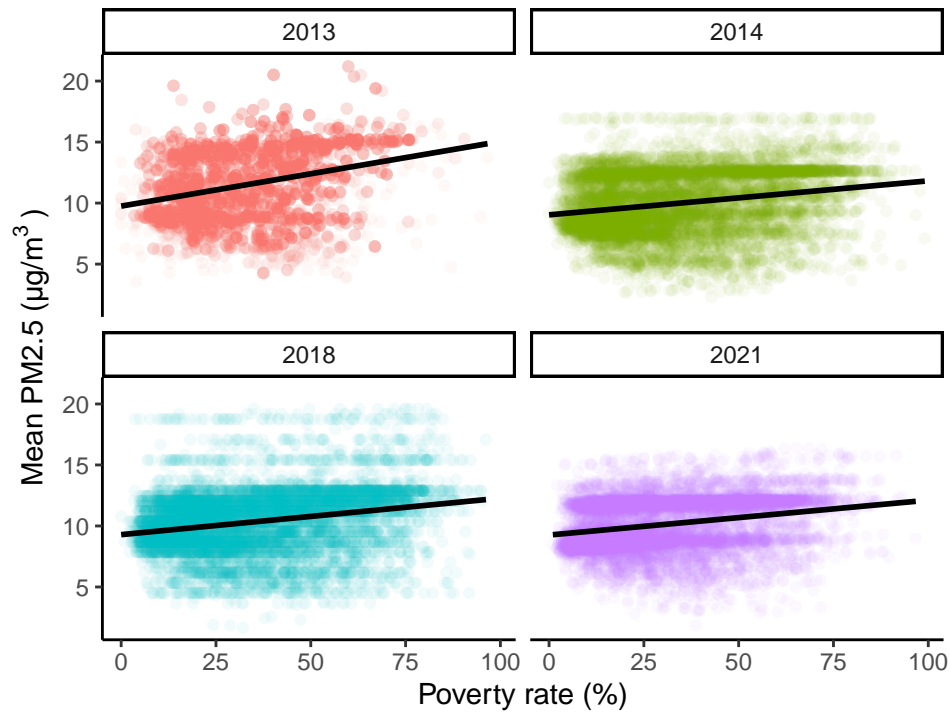
*Show us your results in figure(s) and/or table(s) that are carefully labeled and captioned. Describe in the text (and orally when presenting) what you found, and how these results either do or do not help you answer your question.*

For all time periods, annual mean PM2.5 concentrations were significantly influenced by the percent of people living below twice the federal poverty level. In 2013, PM2.5 increased by 0.035  $\mu\text{g}/\text{m}^3$  as the poverty rate increased by 1% (p-value =  $2.574341 \times 10^{-15}$ , sd = 0.0044083). In 2014, PM2.5 increased by 0.028  $\mu\text{g}/\text{m}^3$  as the poverty rate increased by 1% (p-value =  $1.2707682 \times 10^{-83}$ , sd = 0.0014217). In 2018, PM2.5 increased by 0.030  $\mu\text{g}/\text{m}^3$  as the poverty rate increased by 1% (p-value =  $3.9602597 \times 10^{-99}$ , sd = 0.0013935). In 2021, PM2.5 increased by 0.029  $\mu\text{g}/\text{m}^3$  as the poverty rate increased by 1% (p-value =  $1.6134962 \times 10^{-105}$ , sd = 0.0012904).

Every year's results supported my hypothesis that mean PM2.5 and poverty in California are significantly related (Figure @ref(fig:final)).

## Conclusions

As expected, I found—in California during 2013, 2014, 2018, and 2021—a statistically significant relationship between air quality and poverty rates. Specifically, for all four years, annual mean concentrations of PM2.5 ( $\mu\text{g}/\text{m}^3$ )



Caption

Figure 6: Air quality significantly associates poverty in California. For each sampled time period, as poverty rates increase in California, mean PM2.5 increases and air quality deteriorates (p-value  $\ll 0.05$ ).

increased as the percent of people living below twice the federal poverty level increased (Figure @ref(fig:final)). In other words, air quality was lower in census tracts with higher poverty rates.

These findings supported my hypothesis and corroborated prior research that has identified PM2.5 disparities based on socioeconomic factors in California (Mousavi et al. 2021). This further emphasizes the importance of an environmental justice lens when investigating issues such as air quality.

## Future Directions

*One short analysis cannot fully answer an interesting scientific question. If you had time to collect more data or conduct more analysis, what would help you answer this question better?*

While my analysis focused on four specific years of comprehensive data based on years that the CalEnviroScreen was carried out, it could be interesting to expand that time frame to before 2013. Exploring air quality and poverty before and after 2013 is especially interesting because 2013 was the year that California's cap-and-trade program was initiated. During this time, there is evidence that while greenhouse gases were overall reduced in California, socioeconomically disadvantaged communities actually experienced emission increases (Cushing et al., 2018).

## GitHub

Full code can be found here.

Cisneros, R., Brown, P., Cameron, L., Gaab, E., Gonzalez, M., Ramondt, S., Veloz, D., Song, A., & Schweizer, D. (2017). Understanding Public Views about Air Quality and Air Pollution Sources in the San Joaquin Valley, California. *Journal of Environmental and Public Health*, 2017, e4535142. <https://doi.org/10.1155/2017/4535142>

- Cleland, S. E., Serre, M. L., Rappold, A. G., & West, J. J. (2021). Estimating the Acute Health Impacts of Fire-Originated PM<sub>2.5</sub> Exposure During the 2017 California Wildfires: Sensitivity to Choices of Inputs. *GeoHealth*, 5(7), e2021GH000414. <https://doi.org/10.1029/2021GH000414>
- Cushing, L., Blaustein-Rejto, D., Wander, M., Pastor, M., Sadd, J., Zhu, A., & Morello-Frosch, R. (2018). Carbon trading, co-pollutants, and environmental equity: Evidence from California’s cap-and-trade program (20112015). *PLOS Medicine*, 15(7), e1002604. <https://doi.org/10.1371/journal.pmed.1002604>
- Gupta, P., Doraiswamy, P., Levy, R., Pikelnaya, O., Maibach, J., Feenstra, B., Polidori, A., Kiros, F., & Mills, K. C. (2018). Impact of California Fires on Local and Regional Air Quality: The Role of a Low-Cost Sensor Network and Satellite Observations. *GeoHealth*, 2(6), 172–181. <https://doi.org/10.1029/2018GH000136>
- Li, V. O., Han, Y., Lam, J. C., Zhu, Y., & Bacon-Shone, J. (2018). Air pollution and environmental injustice: Are the socially deprived exposed to more PM<sub>2.5</sub> pollution in Hong Kong? *Environmental Science & Policy*, 80, 53–61. <https://doi.org/10.1016/j.envsci.2017.10.014>
- Qian, X., & Wu, Y. (2019). Assessment for health equity of PM<sub>2.5</sub> exposure in bikeshare systems: The case of Divvy in Chicago. *Journal of Transport & Health*, 14, 100596. <https://doi.org/10.1016/j.jth.2019.100596>
- Schwartz, N. A., & Pepper, D. (2009). Childhood Asthma, Air Quality, and Social Suffering Among Mexican Americans in California’s San Joaquin Valley: “Nobody Talks to Us Here”. *Medical Anthropology*, 28(4), 336–367. <https://doi.org/10.1080/01459740903303944>
- Shi, H., Jiang, Z., Zhao, B., Li, Z., Chen, Y., Gu, Y., Jiang, J. H., Lee, M., Liou, K.-N., Neu, J. L., Payne, V. H., Su, H., Wang, Y., Witek, M., & Worden, J. (2019). Modeling Study of the Air Quality Impact of Record-Breaking Southern California Wildfires in December 2017. *Journal of Geophysical Research: Atmospheres*, 124(12), 6554–6570. <https://doi.org/10.1029/2019JD030472>
- Tessum, C. W., Paoletta, D. A., Chambliss, S. E., Apte, J. S., Hill, J. D., & Marshall, J. D. (2021). PM<sub>2.5</sub> pollutants disproportionately and systemically affect people of color in the United States. *Science Advances*, 7(18), eabf4491. <https://doi.org/10.1126/sciadv.abf4491>