EJScreen

Environmental Justice Mapping and Screening Tool

EJScreen Technical Documentation for Version 2.3

July 31, 2024

U.S. Environmental Protection Agency
Office of Environmental Justice and External Civil Rights
Washington, D.C. 20460

Suggested citation:

U.S. Environmental Protection Agency (EPA), 2024. EJScreen Technical Documentation.

For more information:
www.epa.gov/EJScreen

Contents

1 INTRODUCTION 4

Geographic Framework 4

2 OVERVIEW OF DATA IN EJSCREEN 6

Socioeconomic Indicators in EJScreen 6

Summary Overview of Socioeconomic Indicators Featured in EJScreen 6

Detailed Descriptions of Socioeconomic Indicators 7

Environmental Burden Indicators in EJScreen 14

Summary Overview of Environmental Burden Indicators Featured in EJScreen 15

Detailed Descriptions of Environmental Burden Indicators 16

Considerations for Selection of Environmental Burden Indicators in EJScreen 29

3 OVERVIEW OF INDEXES IN EJSCREEN 31

Demographic Indexes Included in EJScreen 31

Demographic Index 31

Supplemental Demographic Index 33

Indexes Included in EJScreen 36

EJ Indexes 36

Supplemental Indexes 38

4 TECHNICAL DETAILS ON PERCENTILES 40

What a Percentile Means 40

Color-coded High Percentile Bins 40

Use of Percentiles 41

How Percentiles are Calculated 41

5 THRESHOLDS 43

Indexes Threshold Map Widget 43

Initial Filter Approach for Screening 43

6 BUFFER REPORTS 45

7 TECHNICAL DETAILS ON PROXIMITY INDICATORS 47

Calculating Proximity to Facilities 48

Calculating Proximity to Hazardous Waste & RMP Sites 49

Calculating Proximity to Superfund Sites 49

Calculating Proximity to Traffic 50

Calculating Proximity to Toxic Weighted Wastewater Dischargers 50

Calculating Proximity – Additional Details 51

Extremely Small Block/Site Distance Values 51

Caveats and Observations 52

8 DETAILS ON U.S. TERRITORIES 53

Data Sources 53
Processing Steps 53
Index Availability for the Territories 53
Notes about Puerto Rico 54
9 OTHER DATA ELEMENT DESCRIPTIONS 55
Health Disparities Data 55
Climate Change Data 56
Critical Service Gaps 57
EJ Indexes and Supplemental Indexes Availability by Area 58

1 Introduction

EJScreen is EPA's environmental justice (EJ) screening and mapping tool that utilizes standard and nationally-consistent data to highlight places that may have higher environmental burdens and vulnerable populations. The tool offers EJ (2-factor) and supplemental (5-factor) indexes by combining environmental and socioeconomic indicators, using Census block groups as the basic geographic unit. The tool also provides a variety of powerful data and mapping capabilities that enable users to access environmental and demographic information across the entire country, at high spatial resolution, displayed in color-coded maps and standard data reports.

This technical document provides details on the data and methods used to select the indicators and create the indexes in EJScreen. EPA annually updates EJScreen with the newest datasets available and improvements to the interface. For more information on the updates to EJScreen over time, visit the EJScreen Change Log.

Geographic Framework

- The geographic framework for EJScreen was built from 2022 Census TIGER/Line data for all 50 states, the District of Columbia, and Puerto Rico.
- The socioeconomic data source is U.S. Census Bureau's American Community Survey (ACS) 2018-2022 5-Year Estimates (ACS 2022).
- There are a total of 242,336 block groups.
- For 2022, the Census Bureau implemented changes in Connecticut due to new county-equivalent geographic units. The number of counties went from 8 to 9 with new county codes and county names. Also, block group and tract levels needed for EJScreen are impacted. Block group and tract codes are changed only for the county portion of the codes, except there are now one more water block group and one more water tract. Associated ACS tables have the new codes. TIGER/Line boundary layers for tract, block group, and county are also changed. Census also published relationship files to assist with moving pre-2022 codes to 2022. Details can be found here: https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2022/geography-changes.html.
- The application defines the spatial reference as "WGS 1984 Web Mercator (Auxiliary Sphere)."
- The U.S. Territories of American Samoa, Commonwealth of the Northern Mariana Islands (CNMI), Guam, Puerto Rico, and the U.S. Virgin Islands are included in the application. However, Puerto Rico is included with all 50 states and the District of Columbia. The term "territories" in this document refers to the four territories excluding Puerto Rico.
- Puerto Rico uses ACS 2022. American Samoa, CNMI, and Guam use 2022 Census Place boundaries. The U.S. Virgin Islands use 2022 Census Estate boundaries. All spatial boundaries are extracted from the 1:500,000 Cartographic Boundary files. The territories data provides an additional 686 records. Note that socioeconomic data is built from Census 2020 Island Areas Demographic Profile data released in October 2022.

- The Census Place boundaries and Census Estate boundaries for the territories are appended to the block groups for all 50 states and the District of Columbia and Puerto Rico (see Section 8 Details on U.S. Territories for details).
- Census block centroids with population from 2020 Decennial Census P.L. 94-171 Redistricting data are used to provide population-weights for the application.

2 Overview of Data in EJScreen

This section describes the environmental and socioeconomic indicators featured in the application, why they are included, and the data used to derive them.

Socioeconomic Indicators in EJScreen

For Version 2.3:

- All socioeconomic indicators are from Census Bureau's ACS 2018-2022 5-year Summary.
- Two health disparities indicators are also used for the supplemental demographic index:
    - Persons with Disabilities is also from Census Bureau's ACS 2018-2022 5-year Summary.
    - Low life expectancy is from Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS).
- All territories' socioeconomic data are from the Census 2020 Demographic Profile Summary File for each territory, using the Place summary level for American Samoa, CNMI, and Guam, and Estates summary level for U.S. Virgin Islands. The Demographic Profiles were published in October 2022 by the U.S. Census Bureau.

Summary Overview of Socioeconomic Indicators Featured in EJScreen

EJScreen uses socioeconomics indicators as very general indicators of a community's potential susceptibility to the types of environmental factors included in EJScreen. There are seven socioeconomic indicators and five health indicators featured in EJScreen. These indicators form the basis for both the demographic index and the supplemental demographic index.

Socioeconomic Indicators:

1. **People of color:**
    - The percent of individuals in a block group who list their racial status as a race other than white alone and/or list their ethnicity as Hispanic or Latino. That is, all people other than non-Hispanic white-alone individuals. The word "alone" in this case indicates that the person is of a single race, not multiracial.
2. **Low income:**
    - The percent of a block group's population in households where the household income is less than or equal to twice the federal "poverty level."
3. **Unemployment rate:**
    - The percent of a block group's population that did not have a job at all during the reporting period, made at least one specific active effort to find a job during the prior four weeks, and were available for work (unless temporarily ill).
4. **Limited English speaking household**:
    - A "limited English speaking household" is one in which no member 14 years old and over (1) speaks only English or (2) speaks a non-English language and speaks English "very well." In other words, all members 14 years old and over have at least some difficulty with English.

5. **Less than high school education:**
    - Percent of people age 25 or older in a block group whose education is short of a high school diploma.
6. **Under age 5:**
    - Percent of people in a block group under the age of 5.
7. **Over age 64:**
    - Percent of people in a block group over the age of 64.

Detailed Descriptions of Socioeconomic Indicators

People of Color

People of color are individuals who list their racial status as a race other than white alone and/or list their ethnicity as Hispanic or Latino. That is, all people other than non-Hispanic white-alone individuals. The word "alone" in this case indicates that the person is of a single race, not multiracial.

The U.S. Census Bureau classifies races based on the Office of Management and Budget (OMB) standards on race and ethnicity and it is determined based on individuals' self-identification with one or more groups. These groups are:

- White
- Black or African American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Other Pacific Islander

Furthermore, ethnicity is defined as to whether an individual is of Hispanic origin or not. Individuals of any race can report as Hispanic.

How does EJScreen determine percent people of color?

- The ACS information on people of color is captured in the table *Hispanic or Latino Origin by Race* (ACS Table ID: B03002). The ACS divides race and Hispanic status into 21 categories, four of which are shown in Table 1.

Table 1: ACS Hispanic or Latino Origin by Race Table Element Reference

Table Element	Hispanic Status/Race Population
B03002.001	Total Population: All races/ethnicities
B03002.002	Total Population: Non-Hispanic
B03002.003	Total Population: Non-Hispanic, White Alone
B03002.012	Total Population: Hispanic

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate percent people of color, two elements from Table 1 are used in the following equation:

$$\%\ People\ of\ Color = \frac{B03002.001 - B03002.003}{B03002.001}$$

- In EJScreen, the raw values for the percent people of color indicator range from 0% to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the percent people of color indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 31%. If a raw value is higher than 31%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the percent people of color indicator raw value for a block group is 72%, which is higher than 31%, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Low Income

Low income is defined as a household whose income is less than or equal to twice the poverty level. For example, a household of four with a reported $40,000 total annual income is lower than twice the poverty threshold of $59,900 ($29,950 is the poverty threshold defined by the U.S. Census Bureau for 2022). This household will fall into the category of "low income" in EJScreen. The poverty level is updated by the U.S. Census Bureau annually and varies by family size and composition.

The poverty level is a national number and the same across all geographic regions. To accommodate differences in the varying costs of living across the United States and other factors, EJScreen uses twice the poverty level to capture low-income households, especially in high-cost areas. The rationale for using twice the poverty threshold rather than just the poverty threshold includes the following considerations:

- The effects of income on baseline health and probably on other aspects of susceptibility are not limited to those below the poverty thresholds.
- Many studies in various fields use 2x poverty.

How does EJScreen determine percent low income?

- The ACS low income information is captured in the table *Ratio of Income to Poverty Level in the Past 12 Months* (ACS Table ID: C17002). The ACS divides the ratio of income to poverty into seven categories, as shown in Table 2.

Table 2. ACS Income/Poverty Level Table Element Reference

Table Element	Income/Poverty Level
C17002.001	Total Population Whose Poverty Status is Known
C17002.002	People with Ratio of Income to Poverty under .50
C17002.003	People with Ratio of Income to Poverty from .50 to .99
C17002.004	People with Ratio of Income to Poverty from 1.00 to 1.24
C17002.005	People with Ratio of Income to Poverty from 1.25 to 1.49
C17002.006	People with Ratio of Income to Poverty from 1.50 to 1.84
C17002.007	People with Ratio of Income to Poverty from 1.85 to 1.99

C17002.008	People with Ratio of Income to Poverty from 2.00 and over

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate percent low income, two elements from Table 2 are used in the following equation:

$$\%\ Low\ Income = \frac{C17002.001 - C17002.008}{C17002.001}$$

- In EJScreen, the raw values for the percent low income indicator range from 0% to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the percent low income indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 26%. If a raw value is higher than 26%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the percent low income indicator raw value for a block group is 47%, which is higher than 26%, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Unemployment

Unemployment refers to all those who did not have a job at all during the reporting period, made at least one specific active effort to find a job during the prior four weeks, and were available for work (unless temporarily ill). The U.S. Census Bureau determines unemployment through a series of questions in the ACS that establish several labor force elements.

The first parameter is whether or not an individual is age 16 or older. This is necessary because, under federal law, individuals under the age of 16 cannot fully participate in the labor force.

The second parameter is labor force participation. Individuals over the age of 16 who are not classified as members of the labor force include students, stay-at-home parents, retired workers, and others who are not actively searching for employment. An individual can only be considered unemployed if they are part of the labor force.

The last parameter is employment status. An individual is considered unemployed if they were neither "at work" nor "with a job but not at work" during the reference week (the week the survey was conducted), were actively looking for work during the last four weeks, and were available to accept a job. This also includes individuals who did not work at all during the reference week, were waiting to be called back to a job from which they had been laid off, and were available for work except for temporary illness.

How does EJScreen determine unemployment?

- The ACS unemployment information is captured in the table *Employment Status for the Population 16 Years and Over* (Table ID: B23025). The ACS divides the population age 16 or older into seven categories, which are shown in Table 3. EJScreen uses the Civilian Labor Force numbers.

Table 3: ACS Employment Status Table Element Reference

Table Element	Employment Status
B23025.001	Total Population Age 16 Years and Over
B23025.002	Total Labor Force
B23025.003	Total Civilian Labor Force
B23025.004	Civilian Labor Force: Employed
B23025.005	Civilian Labor Force: Unemployed
B23025.006	In Armed Forces
B23025.007	Total Not In Labor Force

Note: These data can be downloaded from the US Census Bureau's FTP Server.

- To calculate the percentage of unemployment in the civilian labor force, two elements from Table 3 are used in the following equation:

$$\%\ Unemployment = \frac{B23025.005}{B23025.003}$$

- In EJScreen, the raw values for the percent unemployment indicator range from 0% to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the percent unemployment indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 3.5%. If a raw value is higher than 3.5%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the percent unemployment indicator raw value for a block group is 10.6%, which is higher than 3.5%, that block group will be placed in the 85th percentile (which is higher than the 50th percentile).

Limited English Speaking Household

A limited English speaking household is defined as a household in which no one age 14 or over speaks English at least "very well" as reported in the U.S. Census Bureau's ACS. The percent of limited English speaking households is used instead of the actual number of limited English speaking households because percentages account for possible differences among population sizes and make block groups comparable.

How does EJScreen determine Limited English Speaking Households?

- The ACS limited English speaking household information is captured in the table *Household Language by Household Limited English Speaking Status* (ACS Table ID: C16002). The ACS divides limited English speaking households into four language groups as shown in Table 4.

Table 4. ACS Limited English Speaking Table Element Reference

Table Element	Language & Ability Status
C16002.001	Total Households
C16002.004	Limited English Speaking – Spanish

C16002.007	Limited English Speaking – Other Indo-European Languages
C16002.010	Limited English Speaking – Asian/Pacific Island Languages
C16002.013	Limited English Speaking – Other Languages

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate percent limited English speaking, the elements from Table 4 are used in the following equation:

$$\%\ Limited\ English\ Speaking = \frac{C16002.004 + C16002.007 + C16002.010 + C16002.013}{C16002.001}$$

- In EJScreen, the raw values for the percent limited English speaking household indicator range from 0% to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the percent limited English speaking household indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 0%. If a raw value is higher than 0%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the percent limited English speaking household indicator raw value for a block group is 5.9%, which is higher than 0%, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Less than High School Education

Percent of people age 25 or older in a block group whose education is less than a high school diploma.

How does EJScreen determine less than high school education?

- The ACS education information is captured in the table *Sex by Educational Attainment for the Population 25 Years and Over* (ACS Table ID: B15002). The ACS divides the education information by grade as shown in Table 5.

Table 5. ACS Educational Attainment Table Element Reference

Table Element	Sex	Educational Attainment
B15002.001		Total Population Age ≥ 25
B15002.003	Male	No Schooling Completed
B15002.004		Nursery – 4th Grade
B15002.005		5th & 6th Grade
B15002.006		7th & 8th Grade
B15002.007		9th Grade
B15002.008		10th Grade
B15002.009		11th Grade
B15002.010		12th Grade, No Diploma
B15002.020		No Schooling Completed
B15002.021		Nursery – 4th Grade
B15002.022		5th & 6th Grade

B15002.023	Female	7th & 8th Grade
B15002.024		9th Grade
B15002.025		10th Grade
B15002.026		11th Grade
B15002.027		12th Grade, No Diploma

Note: These data can be downloaded from the US Census Bureau's FTP Server.

- To calculate the percentage of the 25 and older population with less than high school education, the elements from Table 5 are used in the following equation:

$$\% \text{ Less Than High School Education} = \frac{B15002.003 + .004 + .005 + .006 + .007 + .008 + .009 + .010 + .020 + .021 + .022 + .023 + .024 + .025 + .026}{B15002.001}$$

- In EJScreen, the raw values for the percent less than high school education indicator range from 0% to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the less than percent high school education indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 7.5%. If a raw value is higher than 7.5%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the less than percent high school education indicator raw value for a block group is 18%, which is higher than 7.5%, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Individuals under Age 5

Percent of people in a block group under the age of 5.

How does EJScreen determine individuals under age 5?

- The ACS information on individuals under age 5 is captured in the table *Sex by Age* (ACS Table ID: B01001). The ACS divides the population into 23 age groups for each sex. The elements that were used are shown in Table 6.

Table 6. ACS Age Table Element Reference

Table Element	Age
B01001.001	Total Population
B01001.003	Male, Under 5 Years
B01001.027	Female, Under 5 Years

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate the percentage individuals under age of 5 for a block group, the elements from Table 6 are used in the following equation:

$$\% \, Individuals \; Under \; Age \; 5 = \frac{B01001.003 + B01001.027}{B01001.001}$$

- In EJScreen, the raw values for the percent individuals under age 5 indicator range from 0% to 50%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the percent individuals under age 5 indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 0%. If a raw value is higher than 0%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the percent individuals under age 5 indicator raw value for a block group is 6%, which is higher than 0%, that block group will be placed in the 85th percentile (which is higher than the 50th percentile).

Individuals over Age 64

Percent of people in a block group over the age of 64.

How does EJScreen determine individuals over age 64?

- The ACS information on individuals over age 64 is captured in the table *Sex by Age* (ACS Table ID: B01001). The ACS divides the population into 23 age groups for each sex. The elements that were used are shown in Table 7.

Table 7. ACS Age Table Element Reference

Table Element	Sex	Age
B01001.001		Total Population
B01001.020	Male	65 and 66 years
B01001.021		67 to 69 years
B01001.022		70 to 74 years
B01001.023		75 to 79 years
B01001.024		80 to 84 years
B01001.025		85 years and over
B01001.044	Female	65 and 66 years
B01001.045		67 to 69 years
B01001.046		70 to 74 years
B01001.047		75 to 79 years
B01001.048		80 to 84 years
B01001.049		85 years and over

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate the percentage individuals over the age of 64 for a block group, the elements from Table 7 are used in the following equation:

$$\% \text{ Individuals over Age 64} = \frac{B01001.020+.021+.022+.023+.024+.025+.044+.045+.046+.047+.048+.049}{B01001.001}$$

- In EJScreen, the raw values for the percent individuals over age 64 indicator range from 0% to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for

the percent individuals over age 64 indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 15.7%. If a raw value is higher than 15.7%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the percent individuals over age 64 indicator raw value for a block group is 23%, which is higher than 15.7%, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Environmental Burden Indicators in EJScreen

For Version 2.3:

- Particulate Matter (PM) 2.5 and Ozone values are derived from 2020 tract-level source data from EPA's Office of Air Quality Planning and Standards (OAQPS), Non-attainment areas (NAA) provided in March 2024.
- Nitrogen Dioxide (NO_2) is from 2020 block group-level source data from NASA's Health and Air Quality Applied Sciences Team (HAQAST).
- Air Toxics data, Diesel PM is derived from 2020 block group-level source data from EPA's OAQPS provided in March 2024.
- Toxic Releases to Air source provided by EPA's Office of Pollution Prevention and Toxics (OPPT) on May 16, 2023 from 2021 Risk-Screening Environmental Indicators (RSEI) modeled results.
- Traffic proximity source is derived from the 2020 Highway Performance Monitoring System (HPMS).
- Lead paint has been updated to Census Bureau's ACS 2018-2022 5-year Summary.
- Superfund proximity source is derived from Superfund Enterprise Management System (SEMS) database on February 26, 2024; and NPL Superfund Site Boundaries from EPA Superfund Public map service on February 26, 2024. Superfund boundaries contain National Priorities List (NPL)Final and Proposed plus Superfund Alternative Approach (SAA) sites.
- Risk Management Plan (RMP) facility proximity source is derived from EPA's Facility Registry Service (FRS) by selecting active facilities included in the RMP National Program System on February 5, 2024.
- Hazardous waste proximity sources are derived from operating Treatment, Storage, and Disposal Facilities (TSDFs) from RCRAInfo and Large Quantity Generators (LQGs) from the 2021 Biennial Reports (BR) on February 5, 2024.
- Underground Storage Tanks (UST) source provided by EPA's Office of Underground Storage Tanks on March 5, 2024 at the Census 2022 block group level.
- Wastewater discharge source provided by EPA's OPPT in February 2024 from 2021 RSEI modeled results.
- The Drinking Water Non-Compliance indicator is a weighted measure of violations by community water systems (CWSs) aggregated to block groups. The data used for this indicator is from the Safe Drinking Water Information System (SDWIS) Federal Reporting Services System submission year 2023 quarter 4.

Summary Overview of Environmental Burden Indicators Featured in EJScreen

This section describes the environmental burden indicator data used in EJScreen. EJScreen contains these 13 environmental burden indicators:

1. **Particulate Matter 2.5**
    - PM 2.5 levels in air measured using an annual average.
2. **Ozone**
    - Ozone annual mean top 10 of daily maximum 8-hour concentration in air.
3. **Nitrogen Dioxide (NO_2)**
    - Average annual nitrogen dioxide (NO_2) levels expressed as part per billion (by volume).
4. **Diesel Particulate Matter**
    - Diesel PM level in air.
5. **Toxic releases to air**
    - RSEI modeled toxicity-weighted concentrations in air of TRI listed chemicals.
6. **Traffic proximity and volume**
    - Count of vehicles (annual average daily traffic [AADT]) at major roads within 10 kilometers (km), divided by distance in km.
7. **Lead paint**
    - Percent of housing units built before 1960.
8. **Superfund proximity**
    - Count of Superfund sites (proposed and final NPL sites and SAA sites) within 10 km, each divided by distance in km.
9. **RMP facility proximity**
    - Count of RMP (potential chemical accident management plan) facilities within 10 km, each divided by distance in km.
10. **Hazardous waste proximity**
    - Count of hazardous waste management facilities (TSDFs and LQGs) within 10 km, each divided by distance in km.
11. **Underground storage tanks (UST) and leaking UST (LUST)**
    - Count of LUSTs (multiplied by a factor of 7.7) and the number of USTs within a 1,500-foot buffered block group.
12. **Wastewater discharge**
    - RSEI modeled toxicity-weighted concentrations of TRI listed chemicals in water stream segments within 500 meters, divided by distance in km.
13. **Drinking Water Non-Compliance**
    - Score based on number of Safe Drinking Water Act violations not returned to compliance that community water systems have received over the past five years; violations are weighted based age and severity.

Detailed Descriptions of Environmental Burden Indicators

Environmental Burden Indicator—Particulate Matter 2.5

What is the PM 2.5 indicator?

The PM 2.5 indicator is a measure of potential exposure to inhalable particles that are 2.5 micrometers or smaller. This is measured in terms of annual average concentration in air measured in micrograms per cubic meter. PM 2.5 information included in EJScreen highlights areas across the U.S. that have the highest PM 2.5 levels when compared to the nation or state. It does not indicate whether a selected area is or is not meeting the national ambient air quality standard for PM 2.5.

The PM 2.5 indicator in EJScreen is a measure of potential exposure but not a measure of risk. The raw PM 2.5 data is compiled by 2020 Census tract, which is publicly available for use in the tool by EPA's OAQPS. For air toxics risk measures (as opposed to exposure) users can turn to EJScreen's other indicator: diesel PM.

How does EJScreen determine the Particulate Matter 2.5 indicator?

The PM 2.5 indicator data was provided by EPA's OAQPS using a fusion of monitor data and Community Multiscale Air Quality (CMAQ) air quality modeling. For more information about the methods used, see EPA Report EPA-454/S-15-001. This is provided to EJScreen as a spreadsheet compiled by 2020 Census tracts. The tract values are re-assigned to each child block group, so all block groups within each tract have the same PM 2.5 value as for the tract. 2020 block groups are remapped to 2022 block groups.

In EJScreen, the raw values for the PM 2.5 indicator range from 3.28858 to 25.6121. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the PM 2.5 indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 8.209934. If a raw value is higher than 8.209934, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the PM 2.5 indicator raw value for a block group is 10.07, which is higher than 8.209934, that block group will be placed in the 89th percentile (which is higher than the 50th percentile).

Where to find more information on PM 2.5?

For more information about PM, please visit EPA's website on PM. For real-time and forecasted information about all standard EPA air quality criteria, please visit EPA's AirNow website.

Environmental Burden Indicator—Ozone

What is the ozone indicator?

The ozone indicator in EJScreen v2.3 reflects the annual mean of the 10 highest MDA8 O_3 concentrations (henceforth referred to as the "peak concentration metric"). EJScreen is replacing the seasonal mean

metric with the peak concentration metric so that it is more consistent with the form of the O_3 NAAQS[1]. This better represents the NAAQS attainment status of areas across the nation and exposure to peak ozone concentrations tied to respiratory effects in controlled human exposure studies (U.S. EPA, 2020a; U.S. EPA, 2020b). The ozone indicator in EJScreen is a measure of potential exposure but not a measure of risk.

How does EJScreen determine the ozone indicator?

The ozone indicator data was provided as a spreadsheet by EPA's OAQPS using a fusion of monitor data and CMAQ air quality modeling. For more information about the methods used, see EPA Report EPA-454/S-15-001. This is provided to EJScreen as a spreadsheet compiled by 2020 Census tracts. The tract values are assigned to each child block group, so all block groups within each tract have the same ozone value as for the tract. 2020 block groups are remapped to 2022 block groups.

In EJScreen, the raw values for the ozone indicator range from 6.95112to 77.5368. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the ozone indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 40.489515. If a raw value is higher than 40.489515, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the ozone indicator raw value for a block group is 45.7, which is higher than 40.489515, that block group will be placed in the 86th percentile (which is higher than the 50th percentile).

Where to find more information on ozone?

To learn more, visit EPA's website on ozone. For real-time and forecasted information about all standard EPA air quality criteria, please visit EPA's AirNow website.

References:

Berrocal, V. J.; Gelfand, A. E.; and Holland, D. M. (2011). Space-Time Data fusion Under Error in Computer Model Output: An Application to Modeling Air Quality. *Biometrics*, DOI: 10.1111/j.1541-0420.2011.01725.x.

U.S. EPA. (2015). Regulatory Impact Analysis of the Final Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone. U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA-452/R-15-007. Available from: https://www.epa.gov/sites/default/files/2020-07/documents/naaqs-o3_ria_final_2015-09.pdf

[1] While the form of the O_3 NAAQS is based on the annual 4th highest MDA8 O_3 value, here we look at an average across the top 10 days which will span days above and below the value of the 4th high. By looking at an average across multiple days rather than a single day, this metric provides more year-to-year stability while still representing concentrations that correspond to peak ozone exposure.

U.S. EPA. (2020a). Integrated Science Assessment (ISA) for Ozone and Related Photochemical Oxidants (Final Report, Apr 2020). U.S. Environmental Protection Agency, Washington, DC, EPA-600/R-20-012. Available from: https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=348522

U.S. EPA. (2020b). Policy Assessment (PA) for the Review of the Ozone National Ambient Air Quality Standards (Final Report, May 2020). U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA-452/R-20-001. Available from: https://www.epa.gov/sites/default/files/2020-05/documents/o3-final_pa-05-29-20compressed.pdf

U.S. EPA. (2021). Regulatory Impact Analysis for the Final Revised Cross-State Air Pollution Rule (CSAPR) Update for the 2008 Ozone NAAQS. U.S. Environmental Protection Agency, Research Triangle Park, NC. EPA-452/R-21-002. Available from: https://www.epa.gov/sites/default/files/2021-03/documents/revised_csapr_update_ria_final.pdf

Environmental Burden Indicator—Nitrogen Dioxide

What is the nitrogen dioxide indicator?

This EJScreen indicator provides screening level information, on a block group by block group basis, of average annual nitrogen dioxide (NO_2) levels expressed as part per billion (by volume). Breathing air with a high concentration of NO_2 can irritate airways in the human respiratory system. Such exposures over short periods can aggravate respiratory diseases, particularly asthma, leading to respiratory symptoms (such as coughing, wheezing or difficulty breathing), hospital admissions, and visits to emergency rooms. Longer exposures to elevated concentrations of NO_2 may contribute to the development of asthma and potentially increase susceptibility to respiratory infections. People with asthma, as well as children and the elderly, are generally at greater risk for the health effects of NO_2. The health-based National Ambient Air Quality Standards for NO_2 are a 1-hour standard set at a level of 100 parts per billion (ppb) (based on the 3-year average of 98th percentile of the yearly distribution of 1-hour daily maximum concentrations) and an annual standard set at a level of 53 ppb. EPA concludes that these standards protect the public health with an adequate margin of safety. As of the most recent period of valid monitoring data (2020-2022), all block groups in the U.S. are reporting NO_2 concentrations well below the standards.

How does EJScreen determine the nitrogen dioxide indicator?

Source data for the NO_2 indicator are developed by NASA's Health and Air Quality Applied Sciences Team (HAQAST). Unlike other EJScreen air indicators, the source data does not come from EPA modeling fused with on-the-ground air monitor data. The NO_2 indicator is developed from a regression model, satellite data, and large-scale models. The resulting product is an approximately one square kilometer grid of NO_2 concentrations. These grids are then projected to the 2020 Census block groups using a high-resolution NASA satellite product. The gridded data is then aggregated to census block group geographies using mean pixel values. Please note this data is not used for EPA standards or regulations.

In EJScreen, the raw values for the NO_2 indicator range from 0.0368339 to 25.2692. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the NO_2 indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 7.472829. If a raw value is higher than 7.472829, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the NO_2 indicator raw value for a block group is 10.6, which is higher than 7.472829, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Environmental Burden Indicator—Diesel Particulate Matter

What is the Diesel PM indicator?

The Diesel PM indicator is the estimated concentration of Diesel PM as provided by the 2020 AirToxScreen. The raw indicator is expressed in units of micrograms per cubic meter ($\mu g/m^3$), and reported at the Census block group-level. 2020 block groups are remapped to 2022 block groups. EJScreen converts Diesel PM concentrations into percentiles.

How does EJScreen determine the Diesel PM indicator?

Diesel PM concentrations are provided by a spreadsheet from EPA OAQPS. The source data is compiled by Census tract. The tract values are repeated for each Census block group.

In EJScreen, statistical percentiles are calculated from the raw Diesel PM indicator to make it more comparable. The raw values for the Diesel PM indicator range from 5.559 X10-5 to 2.979 µg/m3. The percentiles for the Diesel PM indicator range from 0 to 100% with a median at the 50th percentile corresponding to a raw value of 0.156 µg/m3, and with the 95th percentile corresponding to a raw value of 0.4472 µg/m3. The 95th percentile and greater mapping bin corresponds to Diesel PM concentrations ranging from 0.4472 to 2.979 µg/m3.

Where to find more information on Diesel PM?

Diesel PM comes from EPA's AirToxScreen. The AirToxScreen is the Agency's ongoing, comprehensive evaluation of air toxics in the United States. This effort aims to prioritize air toxics, emission sources, and locations of interest for further study. It is important to remember that the air toxics data presented here provide broad estimates of health risks over geographic areas of the country, not definitive risks to specific individuals or locations. More information on the AirToxScreen can be found at: https://www.epa.gov/AirToxScreen.

Environmental Burden Indicator—Toxic Releases to Air

What is the Toxic Releases to Air indicator?

The Toxic Releases to Air indicator quantifies relative potential human health impacts of certain chemicals included on the list of toxic chemicals established under section 313 of the Emergency Planning and Community Right-to-Know Act (EPCRA) based on the quantities released to air by facilities.

The indicator is derived by considering the quantities of the chemical(s) released to air, modeling the fate of the chemical(s) in air based on an approximate point of release, determining chemical concentrations in air, and weighting the chemical concentrations by the relative toxicity(ities) of the chemical(s). Average annual toxicity-weighted concentration results from all modeled air releases are aggregated at the U.S. Census block group and other geographies in EJScreen.

How does EJScreen determine the Toxic Releases to Air indicator?

- The Toxic Releases to Air indicator utilizes chemical release and transfer data reported by facilities to EPA's Toxics Release Inventory (TRI) Program along with EPA's RSEI model to estimate average annual chemical concentrations in air weighted by relative toxicity.
- Average annual toxicity-weighted concentrations for the air environmental pathway from the RSEI Geographic Microdata were calculated using the most recent RSEI model version and TRI reported data.
- OPPT provided average annual toxicity-weighted concentration results, aggregated and assigned based on the intersections of the RSEI model grid cells with the EJScreen boundaries for U.S. Census block groups and island areas (i.e., U.S. Territories). The data were provided as a polygon feature class in a file geodatabase.
- To put each indicator value in perspective, Office of Environmental Justice and External Civil Rights (OEJECR) generated statistical percentiles from the supplied toxicity-weighted air concentration data using standard EJScreen algorithms to produce indicator results for different geographic scale comparisons (e.g., national, state).
- In EJScreen, the raw values for the Toxic Releases to Air indicator range from 0 to 4,512,172.1. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the Toxic Releases to Air indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 603.032. If a raw value is higher than 603.032, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the Toxic Releases to Air indicator raw value for a block group is 3122, which is higher than 603.032, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).

Where to find more information on Toxic Releases to Air?

Source data for the Toxic Releases to Air indicator are derived from data used in and results from EPA's Risk-Screening Environmental Indicators (RSEI) model. RSEI model results offer a screening-level and geographic perspective for relative comparisons of potential health impacts resulting from certain waste management activities of chemicals included on the TRI list, established and maintained under section 313 of the EPCRA. RSEI primarily relies on data submitted by regulated facilities to EPA's Toxics Release Inventory (TRI) Program, which compiles and makes available the information reported by facilities that are subject to the TRI reporting requirements. More details and information on the RSEI Geographic Microdata results for air, as well as other geographic-based RSEI modeled results, can be found and further analyzed by visiting the RSEI Geographic Microdata website.

Environmental Burden Indicator—Traffic Proximity

What is the traffic proximity indicator?

The traffic proximity indicator is based on AADT count divided by distance in meters from the Census block centroid. The proximity score is based on the traffic within a search radius of 10 km. This distance was selected to be large enough to capture the great majority of road segments (with traffic data) that could have a significant impact on the local residents, balanced against the need to limit the scope due to computational constraints. The closest traffic is given more weight, and the distant traffic is given less weight, through inverse distance weighting. For example, traffic 500 meters away is given only one tenth as much weight as traffic 50 meters away.

For Version 2.3, the 2020 version of the HPMS data contained overlapping and/or coincident line segments, which would have resulted in overcounting. The following steps were taken to ensure each AADT is used only once: created a new text column by concatenating Route_ID and AADT, then performed a spatial dissolve operation based on the concatenated string, which resulted in a final set of non-overlapping segments. The process was applied to all states.

Why is traffic proximity an indicator in EJScreen?

- Proximity to roads can provide access to jobs, health care, food, recreational opportunities, and other benefits. However, in EJScreen, the indicator is designed to screen for the negative aspects of very close proximity to very high volumes of traffic, which include asthma and cardiovascular and heart disease, among others.
- Residential proximity to traffic has been associated with various health impacts, particularly asthma exacerbation and possibly onset of asthma, as well as mortality rates. Proximity to traffic has also been associated with subclinical atherosclerosis (a key pathology underlying cardiovascular disease [CVD]), prevalence of CVD and coronary heart disease (CHD), incidence of myocardial infarction, and CVD mortality.

How does EJScreen determine the traffic proximity indicator?

- Highway segments are from the HPMS lines and AADT counts are from the 2020 HPMS release, Federal Highway Administration (FWHA), U.S. Department of Transportation (DOT). Proximity scores were calculated by adding a 10km buffer to all non-zero population block groups; performing a one-to-many spatial join to identify all block group-segment pairs; assigning inverse distance scores for all identified 2020 Census blocks within each targeted block group with AADT multiplier (max score of 10); multiplying the scores by AADT and block population weights; and then aggregating the block scores to create a final block group score.

- Information about HPMS can be found on this FHWA website. The 2020 spatial HPMS data is available by state from this DOT-hosted ArcGIS server: https://geo.dot.gov/server/services. EJScreen processing uses a subset of highways that includes:
    - Interstates
    - Principal Arterial—Other Freeways and Expressways

- Principal Arterial—Other
- Minor Arterial in urban areas

- In EJScreen, the raw values for traffic proximity range from 0 to 30,181,802.6. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for traffic proximity range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 814,996.3. If a raw value is higher than 814,996.3, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the traffic proximity raw value for a block group is 3,934,280, which is higher than 814,996.3, that block group will be placed in the 88th percentile (which is higher than the 50th percentile).

Environmental Burden Indicator—Lead Paint Indicator

What is the lead paint indicator?

The lead paint indicator is the percentage of occupied housing units built before 1960, calculated from the U.S. Census Bureau's ACS 5-year summary estimates on age of housing stock. EJScreen uses age of housing stock as a surrogate for potential lead exposure as regulations banning lead-based residential paint in 1978 led to the reduction and finally an end to the use of such paint in housing. The percentage of older housing units is a proxy for potential exposure to lead paint and lead-containing dust that accumulates indoors, in homes, or in other buildings where lead paint was used. EJScreen uses housing units built before 1960.

How does EJScreen determine the lead paint indicator?

- The data are derived from the ACS Summary, block group-level estimates.
- The ACS information on lead paint is captured in the table *Year Structure Built* (ACS Table ID: B25034). The elements that were used are shown in Table 8.

Table 8. ACS Year Structure Built Table

Table Element	Year Structure Built
B25034.001	Total Housing Units
B25034.002	Housing Units Built 2014 or later
B25034.003	Housing Units Built 2010 to 2013
B25034.004	Housing Units Built 2000 to 2009
B25034.005	Housing Units Built 1990 to 1999
B25034.006	Housing Units Built 1980 to 1989
B25034.007	Housing Units Built 1970 to 1979
B25034.008	Housing Units Built 1960 to 1969
B25034.009	Housing Units Built 1950 to 1959
B25034.010	Housing Units Built 1940 to 1949

Table Element	Year Structure Built
B25034.011	Housing Units 1939 or earlier

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate the lead paint indicator, which is an indicator for percent pre-1960 housing units, four elements from Table 8 are used in the following equation:

$$\%\ Lead\ Paint = \frac{(B25034.009 + B25034.010 + B25034.011)}{B25034.001}$$

- In EJScreen, the raw values for the lead paint indicator range from 0 to 100%. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the lead paint indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 21.1%. If a raw value is higher than 21.1%, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the lead paint indicator raw value for a block group is 29.2%, which is higher than 21.1%, that block group will be placed in the 59th percentile (which is higher than the 50th percentile).

Environmental Burden Indicator—Superfund Proximity

What is the Superfund proximity indicator?

The Superfund proximity indicator is reflective of the total count of sites proposed and listed (final or proposed) on the NPL as well as SAA sites in each block group within 10 km of the average resident in a block group, divided by distance, calculated as the population-weighted average of blocks in each block group.

How does EJScreen determine the Superfund proximity indicator?

- NPL Superfund Site Boundaries from Superfund Public map service (FeatureServer) were downloaded from the Geoplatform. Final and proposed NPL sites and SAA sites were downloaded from the SEMS website. Superfund polygons were dissolved by EPA ID to produce 1,401 polygons (Final + Proposed + SAA). 37 points from SEMS were used to finish the universe of sites. These points were buffered by 1 meter distance and appended to polygons for 1,438 input records. Proximity scores were calculated by adding a 10km buffer to all non-zero population block groups; performing a one-to-many spatial join to identify all block group--facility boundary pairs; assigning inverse distance scores for all identified 2020 Census blocks within a targeted block group (max score of 11); multiplying the scores by block population weights; and aggregating scores to block group.
- In EJScreen, the raw values for the Superfund proximity indicator range from 0 to 46.0823. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the Superfund proximity indicator range from 0 to 100% with a median at the 50th

percentile, which corresponds to the raw value of 0.0. If a raw value is higher than 0.0, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the Superfund proximity indicator raw value for a block group is 0.1249, which is higher than 0.0, that block group will be placed in the 64th percentile (which is higher than the 50th percentile).

- More information about NPL sites can be found on this SEMS website.

Environmental Burden Indicator—RMP Facility Proximity

What is the RMP facility proximity indicator?

The RMP facility proximity is reflective of the total count of active RMP facilities in each block group within 10 km of the average resident in a block group, divided by distance, calculated as the population-weighted average of blocks in each block group. ForEJScreen Version 2.3, only active facilities were counted.

How does EJScreen determine the RMP facility proximity indicator?

- RMP facilities are pulled from EPA's FRS by selecting active facilities included in the RMP National Program System.
- Proximity scores are calculated by adding a 10km buffer to all non-zero population block groups; performing a one-to-many spatial join to identify all block group-facility pairs; assigning inverse distance scores for all identified 2020 Census blocks within each block group (max score of 10); multiplying the scores by block population weights; and aggregating scores to block group.
- RMP facilities are queried by using the FRS Query website and selecting RMP in the program system list.
- In EJScreen, the raw values for the RMP facility proximity indicator range from 0.0 to 18.05973. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the RMP facility proximity indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 0.269274. If a raw value is higher than 0.269274, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the RMP facility proximity indicator raw value for a block group is 0.922583, which is higher than 0.269274, that block group will be placed in the 79th percentile (which is higher than the 50th percentile).

Environmental Burden Indicator—Hazardous Waste Proximity

What is the hazardous waste proximity indicator?

The hazardous waste proximity indicator is reflective of the total count of hazardous waste facilities in each block group within 10 km of the average resident in a block group, divided by distance, calculated as the population-weighted average of blocks in each block group. Hazardous waste facilities are defined as Resource Conservation and Recovery Act (RCRA) handlers that are either operating TSDFs from RCRA or reporting LQGs from the 2021 BR.

How does EJScreen determine the hazardous waste proximity?

- TSDFs are collected by using the RCRAInfo Search website and selecting TSDF Handler Universe.
- 2021 BR LQGs are collected by using the BR Search website.
- Proximity scores are calculated by adding a 10km buffer to all non-zero population block groups; performing a one-to-many spatial join to identify all block group-facility pairs; assigning inverse distance scores for all identified 2020 Census blocks within each block group (max score of 10); multiplying the scores by block population weights; and aggregating scores to block group.
- In EJScreen, the raw values for the hazardous waste proximity indicator range is from 0.0 to 82.2821. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the hazardous waste proximity indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 1.274127. If a raw value is higher than 1.274127, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the hazardous waste proximity indicator raw value for a block group is 3.011323, which is higher than 1.274127, that block group will be placed in the 68th percentile (which is higher than the 50th percentile).

Environmental Burden Indicator—Underground Storage Tanks

What is the UST indicator?

The UST indicator quantifies the relative risk of being affected by a LUST for a block group. The indicator is derived by the weighted sum of active LUSTs and sum of active and temporarily out-of-service USTs within a certain distance from a block group.

How does EJScreen determine the UST indicator?

- The UST indicator scores are provided by the EPA's Office of Research and Development.
- EJScreen's UST indicator is calculated using the sum of LUSTs (multiplied by a factor of 7.7) and the number of USTs within a 1,500-foot buffered block group. That number is then divided by the area of the buffered block group in km^2. The 7.7 multiplier is derived from the average number of active USTs divided by the average number of LUSTs in the U.S. backlog (cleanups remaining) from 2011-2020. A 1,500-foot buffer is used as a radius of influence for the Benzene plume migration to encompass USTs/LUSTs near block groups that could potentially be affected by a release. See EPA's website on underground and leaking underground storage tanks for more information.
- In EJScreen, the raw values for the UST indicator range from 0.0 to 137.8717. In order to make the indicator more comparable, percentile ranks are used. The percentiles for the UST indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 0.824333. If a raw value is higher than 0.824333, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the UST indicator raw value for a block group is 5.17, which is higher than 0.824333, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).
- To calculate the UST indicator for a block group, the following equation is used:

$$UST\ Indicator = \frac{(\#\ of\ LUSTs\ \times 7.7) + (\#\ of\ Active\ USTs)}{Area\ of\ buffered\ block\ group_{1500ft-buffer}}$$

Environmental Burden Indicator—Wastewater Discharge

What is the wastewater discharge indicator?

The wastewater discharge indicator quantifies relative potential human health impacts of certain chemicals included on the list of toxic chemicals established under section 313 of the EPCRA based on the quantities discharged to water by facilities and proximity measures. The indicator is derived by considering the quantities of the chemical(s) discharged to water, modeling the fate of the chemical(s) in water based on an approximate point of release, determining chemical concentrations in water, and weighting the chemical concentrations by the relative toxicity(ities) of the chemical(s). Average annual toxicity-weighted concentration results in downstream stream segments within 10 km of a block centroid, divided by distance in meters, are presented as the population-weighted average of blocks in each block group.

How does EJScreen determine the wastewater discharge indicator?

- The wastewater discharge indicator utilizes discharge monitoring report (DMR) and TRI data reported by facilities along with EPA's RSEI model to estimate average annual chemical concentrations in water weighted by relative toxicity.
- Average annual toxicity-weighted concentrations for the water environmental pathway were calculated using the most recent RSEI model version and reported data.
- OPPT in collaboration with the Office of Water provided average annual toxicity-weighted concentration results, aggregated and indexed to National Hydrography Dataset (NHD) water stream segments (flowlines). The data were provided as a polyline feature class in a file geodatabase.
- To place higher emphasis on stream segments with higher toxicity-weighted chemical concentrations, proximity scores were calculated by adding a 10km buffer to all non-zero population block groups; performing a one-to-many spatial join to identify all block group--segment pairs; assigning inverse distance scores for all identified 2020 Census blocks within each targeted block group with toxic concentrations multiplier (max score of 10); multiplying the weighted scores by toxic concentrations and block population weight to produce the final block scores; then aggregating to block group. To put each indicator value in perspective, OEJECR generated statistical percentiles from the supplied toxicity-weighted water concentration data using standard EJScreen algorithms to produce indicator results for different geographic scale comparisons (e.g., national, state).
- In EJScreen, the raw values for the wastewater discharge indicator range from 0.0 to 7,709,094,779. In order to make the indicator more comparable, statistical percentiles are used. The percentiles for the wastewater discharge indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 50.475064. If a raw value is higher than 50.475064, it would be placed above the 50th percentile; if the value is lower, it would be

placed below the 50th percentile. For example, if the wastewater discharge indicator raw value for a block group is 1757.464869, which is higher than 50.475064, that block group will be placed in the 78th percentile (which is higher than the 50th percentile).

Where to find more information about wastewater discharges and potential impacts?

- The NPDES permit program established by the Clean Water Act (CWA) addresses water pollution by regulating point sources that discharge pollutants to waters of the United States. EPA's Enforcement and Compliance History Online tool and TRI Toxics Tracker can be used to search for facilities and wastewater discharges in a community. Users can also find out more about a local watershed through EPA's How's My Waterway resource. To find more details and information on how EPA's RSEI model calculates results involving wastewater discharges of TRI-listed chemicals, please visit https://www.epa.gov/rsei/modeling-water-releases-rsei.

Environmental Burden Indicator—Drinking Water Non-Compliance

What is the drinking water non-compliance indicator?

The goal of this indicator is to highlight populations served by community water systems (CWSs) that have challenges complying with National Primary Drinking Water Regulations (NPDWRs). The higher the Drinking Water Non-Compliance Indicator score, the more compliance challenges faced by the CWS(s) serving that census block group population. The drinking water non-compliance indicator distills hundreds of different violation types into a single value that attempts to capture Safe Drinking Water Act (SDWA) compliance performance of a CWS. It is important to note that the compliance data used by this indicator does not directly measure drinking water quality as this depends on violation type and other factors unique to each water system.

How does EJScreen determine the drinking water non-compliance indicator?

The drinking water non-compliance indicator is only applicable to households that get their drinking water from a CWS, which are subject to NPDWRs under the SDWA. The indicator excludes households on private wells, for example, which are not regulated by the EPA. The indicator reflects compliance performance for only the last 5-years (2019 Q4- 2023 Q4). Scores are only given to water systems for violations that have not been returned to compliance.

Drinking Water Non-Compliance Score Description:

- Violations are divided into 3 categories with associated point values weighted by severity (based on the three SDWA public notice violation tiers):
    1. Health Based Violations (immediate impacts), 10 points:
        a. Nitrate and/or Nitrite Maximum Contaminant Levels (MCLs), Acute Maximum Residual Disinfectant Level (MRDL), Revised Total Coliform Rule (R/TCR) Acute, Turbidity Treatment Technique (TT), Surface Water Treatment Rule Treatment Technique (SWTR TT).
    2. Other Health Based Violations, 5 points:

a. Any other health-based violation (violation type MCL, MRDL or TT), plus Total Coliform Rule TCR Monitoring/Reporting (MR) repeat violations (violation code 25 and 26) and Nitrate MR.

3. Non-Health Based Violations, 1 point:
    a. Any other Monitoring Violation (MON), Reporting Violation (RPT), Monitoring and Reporting (MR), or Other Violations categories.

- For all 10 and 5 point violation types (see 1a. and 2a. above), the time (in years) a violation has been in non-compliance is added to the violation score. 1-point is added for every year of non-compliance, up to 5-years (5-points).
- The violation data that makes up the drinking water non-compliance indicator was taken from EPA's internal Enforcement and Targeting Tool (ETT) which draws upon the SDWA system of record known as the Safe Drinking Water Information System, or SDWIS. SDWIS is updated on a quarterly basis and the violation data used to create the Drinking Water Non-Compliance Indicator score is from 2023 Q4.
- Not all community water system boundaries are mapped. The systems not mapped do not get represented in the EJScreen indicator. However, 99% of all public water consumers that receive their water from a community water system are represented in the service area dataset.
- This service area dataset is published by EPA's Office of Research and Development (ORD). Link to data application and relevant webpages:
    - https://www.epa.gov/ground-water-and-drinking-water/community-water-system-service-area-boundaries?tab=map The 5-year service area score of all violation types is first appended into the community water system service area boundaries:
- The 5-year service area score of all violation types is first appended into utility service area boundaries:

$$Service\ Area\ Score = \sum(Violation\ severity + Years\ in\ non-compliance_{health-based})$$

- The indicator score is then calculated at the block level and weighted by block population as a percent of its parent block groups' population (see equation below). If any part of the service area boundary is within a block, that block's score gets associated with that service area boundary(ies). If two or more service areas intersect the same block the scores are averaged. This method takes the most conservative approach: theoretically, if only one person is served by a public water system in that block, the score is applied to ***everyone*** living in that census block.

$$DWInScore_{blk} = \sum \frac{(DWInScore_{blk})}{n} * \%\ of\ block\ group\ population$$

- The block group score is the sum of the scores calculated at the associated census block level (see equation below). If only one system serves a block group, and the system encompasses all associated block group blocks, the block group score will be identical to the service area score for that system.

$$DWInScore_{blkGrp} = \sum(DWInScore_{blk})$$

- If no service areas intersect any part of a block group, that block group gets a null score. These areas are typically supplied by private well water.
- In EJScreen, the raw values for the drinking water non-compliance indicator range from 0 to 899. To make the indicator more comparable, statistical percentiles are used. The percentiles for the drinking water non-compliance indicator range from 0 to 100% with a median at the 50th percentile, which corresponds to the raw value of 0. If a raw value is higher than 0, it would be placed above the 50th percentile; if the value is lower, it would be placed below the 50th percentile. For example, if the drinking water non-compliance indicator raw value for a block group is 1, which is higher than 0, that block group will be placed in the 80th percentile (which is higher than the 50th percentile).
- For more information on how this indicator was created, click here.

Considerations for Selection of Environmental Burden Indicators in EJScreen

A variety of considerations has informed the selection of these environmental burden indicators; in general, the selected indicators exhibit the following characteristics:

- Resolution: Screening level data are available (or could be readily developed) at the block group level (or at least close to this resolution).
- Coverage: Screening level data are available (or could be readily developed) for the entire United States (or with nearly complete coverage).
- Relevance to EJ: Pollutants or impacts are relevant to environmental justice (e.g., differences between groups have been indicated in exposures, susceptibility, or health endpoints associated with the exposures).
- Public health significance: Pollutants or impacts are potentially important in the United States (e.g., notable impacts estimated or significant concerns have been expressed, at least locally, or exposure has been linked to health endpoints with substantial impacts nationwide).

Additionally, EPA triages the inclusion of various environmental datasets into EJScreen by several factors and criteria:

Description	Prioritize	De-prioritize
Geographic consistency	Nationally available datasets	Datasets only available regionally/by states/by communities
Temporal availability	Datasets representing current environmental burdens	Datasets representing historic environmental burdens
Data completeness	Data representing the entire population of the environmental burden	Only a small sample of pollution sources from a larger population

Homogeneousness of the environmental burden	The pollutant effects all or most geographic areas and people (i.e, traffic proximity effects almost all communities)	Pollution datasets that only effect certain communities, states, regions (i.e., uranium or coal mining is local to only a few states)
Impacts of the burden	Environmental burdens effects large populations	Environmental burdens only effecting a few people
The burden is federally regulated	Federally regulated pollution	Not regulated or state regulated pollution
The environmental burden's effect on health outcomes is not minimal	Pollution that can lead to seriously public health outcomes	Pollution that has negligible public health outcomes
Spatial accuracy	Spatially accurate datasets	Non-spatially accurate datasets
Spatial consistency	Environmental pollution samples that are spatially homogeneous (i.e., satellite data on NO_2 levels)	Samples of pollution are spatially heterogeneous. Example: USGS groundwater monitoring wells in Nebraska are in the hundreds; no wells are installed in Utah
The burden can be mapped	Mappable environmental burdens	Non-mappable environmental burdens
Spatial autocorrelations between pollution source and exposure	There is a strong relationship between the geography of the pollutant and the geography of those effected by the pollutant	No known relationship between the geography of the pollutant and the geography of those effected by the pollutant
The environmental burden can be expressed into block group geographies	Can be logically mapped to the block group (i.e., density of oil refineries)	Difficult or non-logical to map to block groups (i.e., impaired river-ways. Rivers being linear features flowing through block groups effecting a very small proportion of a given block group)

3 Overview of Indexes in EJScreen

This section describes the environmental and socioeconomic indexes featured in the tool, why they are included, and how they are derived.

Demographic Indexes Included in EJScreen

EJScreen includes two demographic indexes based on different combinations of the socioeconomic and health indicators. The two demographic indexes featured in EJScreen are:

- **Demographic Index** is based on the average of the Z-score values of the two socioeconomic indicators: percent low income and percent people of color.
- **Supplemental Demographic Index** is based on the average of the Z-score values of the five socioeconomic and health indicators: percent low income, percent limited English speaking, percent less than high school education, percent persons with disabilities, and low life expectancy.

Z-score is a statistical measurement that tells how far a value is from the mean of a group of values. Negative values are smaller than the mean, and positive values are greater than the mean. Using μ for the mean and σ for the standard deviation, a Z-score value can be calculated as follows:

$$\mathbf{z = (x - \mu) / \sigma}$$

For example, given a particular block group with percent low income value (x) of 45%, a mean value (μ) of 30.4% and the standard deviation (σ) of 21.3%, the Z-score value for percent low income can be calculated as follows:

$$z = \frac{45 - 30.4}{21.3} = 0.685$$

A Z-score value of percent low income, percent people of color, percent limited English speaking, percent less than high school education, percent persons of disabilities, or low life expectancy was computed before a demographic index or supplemental demographic index was calculated for each block group.

Demographic Index

What is the Demographic Index in EJScreen?

The Demographic Index in EJScreen is an average measure of two demographic variables of percent low income and percent people of color. These two demographic factors were explicitly named in Executive Order 12898 on Environmental Justice.

How does EJScreen determine the Demographic Index?

EJScreen uses these two socioeconomic indicators:

1. The ACS low income information is captured in the table *Ratio of Income to Poverty Level in the Past 12 Months* (ACS Table ID: C17002). The ACS divides the ratio of income to poverty into seven categories, as shown in Table 9.

Table 9. ACS Income/Poverty Level Table Element Reference

Table Element	Income/Poverty Level
C17002.001	Total Population Whose Poverty Status is Known
C17002.002	People with Ratio of Income to Poverty under .50
C17002.003	People with Ratio of Income to Poverty from .50 to .99
C17002.004	People with Ratio of Income to Poverty from 1.00 to 1.24
C17002.005	People with Ratio of Income to Poverty from 1.25 to 1.49
C17002.006	People with Ratio of Income to Poverty from 1.50 to 1.84
C17002.007	People with Ratio of Income to Poverty from 1.85 to 1.99
C17002.008	People with Ratio of Income to Poverty from 2.00 and over

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate percent low income, two elements from Table 9 are used in the following equation:

$$\%\ Low\ Income = \frac{C17002.001 - C17002.008}{C17002.001}$$

2. The ACS people of color information is captured in the table *Hispanic or Latino Origin by Race* (ACS Table ID: B03002). The ACS divides race and Hispanic status into 21 categories, four of which are shown in Table 10.

Table 10. ACS Hispanic or Latino Origin by Race Table Element Reference

Table Element	Hispanic Status/Race Population
B03002.001	Total Population: All races/ethnicities
B03002.002	Total Population: Non-Hispanic
B03002.003	Total Population: Non-Hispanic, White Alone
B03002.012	Total Population: Hispanic

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate percent people of color, two elements from Table 10 are used in the following equation:

$$\%\ People\ of\ Color = \frac{B03002.001 - B03002.003}{B03002.001}$$

3. To calculate the Demographic Index, the Z-score values for the percent low income and percent people of color are first calculated using the formula mentioned previously and then two Z-score values are averaged to produce the demographic index as follows:

$$Demographic\ Index = \frac{\%\ LowIncome_z + \%\ People\ of\ Color_z}{2}$$

For example, if a Census block group has a low income Z-score value of 0.25 and a people of color Z-score value of 0.75, the Demographic Index value would be 0.5. After calculation, the absolute value of the smallest demographic index value was added to individual demographic indexes to make them non-negative. This shift was applied to national and state levels.

Why is the Demographic Index in EJScreen?

The Demographic Index offers a perspective on community-level vulnerability across the country based on income level and people of color, which are the two demographic factors explicitly named in Executive Order 12898 on Environmental Justice. Low income and people of color populations often experience greater exposure to environmental burdens than the general population as a whole. Many studies have established that sources of environmental hazards are often located and concentrated in areas that are dominated by low income and people of color populations. The Demographic Index also provides additional functionality for use in decision-making considerations, such as the allocation of government resources when awarding grants.

Supplemental Demographic Index

What is the Supplemental Demographic Index in EJScreen?

The Supplemental Demographic Index uses the same updated methodology and calculation as the one used for the Demographic Indexes but replaces the current Demographic Index (the average of percent low income and percent people of color) with a supplemental five-factor demographic index. The five socioeconomic indicators considered are low life expectancy, percent low income, percent persons with disabilities, percent limited English speaking, and percent less than high school education.

How does EJScreen determine the Supplemental Demographic Index?

The Supplemental Demographic Index is the average of the following five indicators:

1. Low Life Expectancy—To highlight areas where the life expectancy is lower than national norms, EJScreen uses an inverse measure showing higher values for lower years and lower scores for higher years. Low Life Expectancy is an inverse of the normalized life expectancy (as defined below) derived from the Life Expectancy at Birth from CDC NCHS:

   % Low Life Expectancy is defined as "1 – (Life Expectancy / Max Life Expectancy)"

   Note: *This is derived from the CDC life expectancy at birth data using the formula above.*

2. The ACS low income information is captured in the table *Ratio of Income to Poverty Level in the Past 12 Months* (ACS Table ID: C17002). The ACS divides the ratio of income to poverty into seven categories, as shown in Table 11.

Table 11. ACS Income/Poverty Level Table Element Reference

Table Element	Income/Poverty Level
C17002.001	Total Population Whose Poverty Status is Known
C17002.002	People with Ratio of Income to Poverty under .50
C17002.003	People with Ratio of Income to Poverty from .50 to .99
C17002.004	People with Ratio of Income to Poverty from 1.00 to 1.24
C17002.005	People with Ratio of Income to Poverty from 1.25 to 1.49
C17002.006	People with Ratio of Income to Poverty from 1.50 to 1.84
C17002.007	People with Ratio of Income to Poverty from 1.85 to 1.99
C17002.008	People with Ratio of Income to Poverty from 2.00 and over

Note: The data can be downloaded from the US Census Bureau's FTP Server.

- To calculate percent low income, two elements from Table 11 are used in the following equation:

$$\%\ Low\ Income = \frac{C17002.001 - C17002.008}{C17002.001}$$

3. Persons with Disabilities—Percent of all persons with disabilities. This data is derived from 2022 ACS "Sex by Age by Disability Status" table (B18101) for Census tracts. Block group values are calculated by multiplying the tract value by the block group population weights. The weights are derived from the same Census source used by the EJScreen buffer reports and analysis—2020 Decennial Census P.L. 94-171 Redistricting data. Source: Download ACS Data

4. The ACS limited English speaking household information is captured in the table *Household Language by Household Limited English Speaking Status* (ACS Table ID: C16002). The ACS divides limited English speaking households into four language groups as shown in Table 12.

Table 12. ACS Limited English Speaking Table Element Reference

Table Element	Language & Ability Status
C16002.001	Total Households
C16002.004	Limited English Speaking – Spanish
C16002.007	Limited English Speaking – Other Indo-European Languages
C16002.010	Limited English Speaking – Asian/Pacific Island Languages
C16002.013	Limited English Speaking – Other Languages

Note: The data can be downloaded from the US Census Bureau's FTP Server.

5. The ACS education information is captured in the table *Sex by Educational Attainment for the Population 25 Years and Over* (ACS Table ID: B15002). The ACS divides the education information by grade as shown in Table 13.

Table 13. ACS Educational Attainment Table Element Reference

Table Element	Sex	Educational Attainment
B15002.001		Total Population Age ≥ 25
B15002.003	Male	No Schooling Completed
B15002.004		Nursery – 4th Grade
B15002.005		5th & 6th Grade
B15002.006		7th & 8th Grade
B15002.007		9th Grade
B15002.008		10th Grade
B15002.009		11th Grade
B15002.010		12th Grade, No Diploma
B15002.020	Female	No Schooling Completed
B15002.021		Nursery – 4th Grade
B15002.022		5th & 6th Grade
B15002.023		7th & 8th Grade
B15002.024		9th Grade
B15002.025		10th Grade
B15002.026		11th Grade
B15002.027		12th Grade, No Diploma

Note: These data can be downloaded from the US Census Bureau's FTP Server.

6. To calculate the Supplemental Demographic Index, the Z-score values for the low life expectancy, percent low income, percent persons with disabilities, percent limited English speaking, and percent less than high school education are first calculated using the formula mentioned previously and then five Z-score values are averaged to produce the supplemental demographic index as follows:

$$Supplemental\ Demographic\ Index = \frac{\%LowIncome_z + \%LimitedEnglish_z + \%LessthanHS_z + \%Disabilities_z + LowLifeExp_z}{5}$$

Note that the low life expectancy data are available for about 90% of the country because of availability of the CDC Life Expectancy data. For the areas where the low life expectancy data are not available, the Supplemental Demographic Index becomes the average of four indicator Z-scores instead of five, as follows:

$$Supplemental\ Demographic\ Index = \frac{\%LowIncome_z + \%LimitedEnglish_z + \%LessthanHS_z + \%Disabilities_z}{4}$$

After calculation, the absolute value of the smallest supplemental demographic index value was added to individual supplemental demographic indexes to make them non-negative. This shift was applied to national and state levels.

Why is the Supplemental Demographic Index in EJScreen?

The Supplemental Demographic Index offers a different perspective on community-level vulnerability across the country. The Supplemental Demographic Index also provides additional functionality for use in decision-making considerations, such as the allocation of government resources when awarding grants. For questions on the appropriate use of the EJ and Supplemental Demographic Indexes, please contact your servicing legal office.

Indexes Included in EJScreen

EJScreen features two sets of indexes, 13 EJ Indexes and 13 Supplemental Indexes, which are described in detail below. The indexes are a combination of the environmental burden indicators described above and the Demographic Index or the Supplemental Demographic Index, as described above.

EJ Indexes

The EJ Indexes are a combination of environmental indicators and the Demographic Index described previously. EJScreen features a single EJ Index for each of the 13 environmental burden indicators.

EJ Index Calculations:

To calculate a single EJ Index, EJScreen combines a single environmental burden indicator with the Demographic Index.

The equation for EJ Index calculations is as follows:

$$EJ\ Index\ =\ Demographic\ Index\ \times\ Normalized\ Environmental\ Indicator$$

where the Demographic Index is the average of two Z-score values of percent low income and percent of people of color, and the Normalized Environmental Indicator is the percentile of an environmental indicator.

Percentiles Methodology Notes:

- Percentiles calculations are unweighted.
- Percentile ties use a floor method. This produces the lowest value; for example, for an element with more than 50 percent zeros, like “Limited English Speaking”, a value of 0 will yield 0 percentile and not 50th percentile.

EJ Index Example:

For a block group (ID = 410510068022) in Lincoln County, OR, the percent of people of color is 0.252511 while the national mean is 0.395843 and the national standard deviation is 0.315152. The percent low income is 0.1873 while the corresponding national average is 0.30415 and the national standard deviation is 0.213371. All values here are in decimal. Based on the Z-score formula introduced

previously, a Z-score value for the percent people of color or percent low income can be computed as follows:

$$\%\ people\ of\ color_z = \frac{0.252511\ -\ 0.395843}{0.315152} = -0.454803$$

and

$$\%\ low\ income_z = \frac{0.252511\ -\ 0.395843}{0.315152} = -0.547638$$

So the demographic index is:

$$Demographic\ Index = \frac{\%\ People\ of\ Color_z + \%\ LowIncome_z}{2}$$

$$= \frac{-0.454803\ +\ (-0.547638)}{2}$$

$$= -0.501221$$

To make the demographic index non-negative, the absolute value of the smallest national demographic index is added. Therefore, the shifted demographic index is:

$$Demographic\ Index = -0.501221 + |-1.34074|$$

$$= 0.8395$$

The EJ Index for Superfund Proximity is a product of the demographic index and the national percentile for Superfund Proximity. For this block group, the national percentile for Superfund Proximity is 59. So the EJ Index is:

$$EJ\ Index = 59 \times 0.8395 = 49.532$$

After comparing the EJ Index lookup table for the Superfund Proximity using the EJ Index value above, the national EJ Index percentile for Superfund Proximity for this block group is 66.

Supplemental Indexes

The supplemental indexes are a combination of environmental indicators described above and the Supplemental Demographic Index. Similar to the EJ Indexes, EJScreen features a single supplemental index for each of the 13 environmental indicators.

Supplemental Index Calculations:

To calculate a single supplemental index, EJScreen combines a single environmental indicator with the Supplemental Demographic Index.

The equation for Supplemental Index calculations is as follows:

$$Supplemental\ Index = Supplemental\ Demogaphic\ Index\ \times Normalized\ Environmental\ Indicator$$

where the Normalized Environmental Indicator is Percentile of Environmental Indicator Source Data.

Supplemental Index Example:

For the same block group (ID = 410510068022) in Lincoln County, OR, the necessary data (in decimal) for five components used to compute the supplemental demographic index are included in the following table:

	%Low Income	%Limited English Speaking	%Less Than High School Education	%Persons with Disabilities	%Low Life Expectancy
Value	0.1873	0.023454	0.026035	0.102395	0.152821
Mean	0.30415	0.0481	0.11307	0.136065	0.19532
Stand Deviation	0.213371	0.108136	0.1169	0.062492	0.040118
Z-score	-0.547638	-0.227917	-0.744525	-0.538789	-1.05935

The Z-score value for each component is also computed and included in the table. The supplemental demographic index is the average of the Z-score values of all five components:

$$Supplemental\ Demographic\ Index = \frac{\%LowIncome_z + \%LimitedEnglish_z + \%LessthanHS_z + \%Disabilities_z + LowLifeExp_z}{5}$$

So

$$Supplemental\ Demographic\ Index = \frac{(-0.547638) + (-0.227917) + (-0.744525) + (-0.538789) + (-1.05935)}{5} = -0.6236438$$

To make the supplemental demographic index non-negative, the absolute value of the smallest national supplemental demographic index is added. Therefore, the shifted supplemental demographic index is:

$$Supplemental\ Demographic\ Index = -0.6236438 + |-1.6392| = 1.0156$$

The Supplemental Index for Superfund Proximity is a product of the supplemental demographic index and the national percentile for Superfund Proximity. For this block group, the national percentile for Superfund Proximity is 59. So the Supplemental Index is:

$$Supplemental\ Index = 59 \times 1.0156$$
$$= 59.92$$

After comparing the Supplemental Index lookup table for the Superfund Proximity using the Supplemental Index value above, the national Supplemental Index percentile for Superfund Proximity for this block group is 60.

4 Technical Details on Percentiles

What a Percentile Means

EJScreen puts each indicator or index value in perspective by reporting the value as a percentile. For example, an area may show 60% of housing was built prior to 1960. It may not be obvious whether this is a relatively high or low value, compared to the rest of the nation or in the state. Therefore, EJScreen also reports that 60% pre-1960 puts this area at the 80th percentile nationwide. For a place at the 80th percentile nationwide, that means 20% of the U.S. population has a higher value.

A percentile in EJScreen tells us roughly what percent of the U.S. block groups have a lower value (or in some cases, a tied value). This means that 100 minus the percentile tells us roughly what percent of U.S. block groups have a higher value. This is generally a reasonable interpretation because for most indicators there are not many exact ties between places and not many places with missing data.

Note that when there are ties, a "floor" method is used to make the assignment. For example, if an indicator value of 0 is assigned percentiles from 0 to 10, the reported percentile will be 0 instead of 10.

One note is that block group indicator null values are not included in percentile calculation. For example, Ozone or PM 2.5 are not available for Alaska, Hawaii, and Puerto Rico, so the block groups from those areas are excluded from percentile calculation. In cases like that, only block groups with valid values are included in percentile calculation.

Color-coded High Percentile Bins

All locations above the 50th percentile receive a color designation in EJScreen, going from light yellow to dark red as the percentile increases. The color bins are four categories of percentiles of 50-80 percentile, 80-90 percentile, 90-95 percentile, and 95-100 percentile. These colors call attention to certain locations as a very simple way to communicate relative screening results. There is no official policy significance assigned to each individual color on the maps, but the choice of these categories or "bins" is noteworthy because it signifies that certain ranges of percentiles may merit closer attention or can be easily mapped.

A dark color or relatively high percentile means the value is relatively uncommon. However, a high percentile is not necessarily a real concern from a health or legal perspective. To understand the actual health or other implications of any screening results requires looking at the actual data the indicator represents, and also looking at other relevant data if available. Besides the percentile, other important considerations in interpreting any screening results include the following:

1. whether and to what extent the environmental data shows values above any relevant health-based or legal threshold,
2. the significance of any such thresholds, or the magnitude and severity of the health or other impacts of the given environmental concern, nationally or locally, and

3. the degree of any disparity between various groups, in exposures to the relevant environmental pollutants.

Use of Percentiles

The U.S. percentiles are computed based on the block groups nationwide. The state percentiles are calculated within each individual state. EJScreen focuses on the U.S. percentiles as a way to visualize all results in common units. The U.S. percentile uses the U.S. block groups as the basis of comparison. The state percentile was also calculated based on the block groups of a given state (or the District of Columbia or Puerto Rico). The national or state mean value was calculated as the average of the block groups with data for that indicator, within the respective geographic scope.

Note that the U.S. and state percentiles both will rank block groups in exactly the same rank order within the given state. If the goal is just to rank or compare locations within a single state, it does not matter whether the U.S. or state percentile is used. The difference between state and U.S. percentiles becomes apparent mainly in two situations: when comparing places across states, or when comparing results to some pre-determined, specific reference percentile (e.g., 80th percentile).

The advantage of U.S. percentiles for an EJ Index, for example, is that a higher percentile in place A versus place B clearly indicates that the combination of the environmental burden indicator and Demographic Index is greater in place A than in place B. In a sense, the U.S. percentile indicates how uncommon it is to have such a high level for an indicator or index.

State percentiles cannot be compared across states as easily. If two places A and B, in two different states, happen to both be at the 80th percentile for the traffic proximity EJ Index, for example, it is not clear which actually has the higher index value. It just means that A's index is just as uncommon within that state as B's is in B's state. However, this may be useful information because an EJScreen user may want to know how high the indicator is relative to the rest of that state.

For a given location or community, the state and U.S. percentiles may not be the same or similar but instead would be different. The state and U.S. percentiles will typically be very similar if the state and U.S. average indicator values are very similar. However, if the state average is very low compared to the U.S., the state percentile shown will be higher than U.S. percentile shown, for a given raw value of an indicator. If the state average is much higher than the U.S. average, for an indicator like the traffic proximity indicator, then a traffic score that would normally be considered fairly high nationwide, such as the 90th percentile in the U.S., would not be considered very unusual within that state, so the state percentile would be lower, and might be only 78th percentile, for example. The state percentile being lower than the U.S. percentile does not mean the indicator value is lower in the given place, it just means the state average is higher than the U.S. average.

How Percentiles are Calculated

A percentile lookup table was built for the U.S. and each individual state, including the District of Columbia and Puerto Rico. The corresponding raw values to each of 0 to 100 percentile with incremental

one percentile were stored in the table. To determine the percentile for a given value, it compares against 101 values from 0 to 100 percentile. Once a range is detected, the percentile for the lower range value will be assigned. When there are ties, a “floor” method is used to make the assignment. For example, if an indicator value of 0 is assigned percentiles from 0 to 10, the reported percentile will be 0 instead of 10.

5 Thresholds

Indexes Threshold Map Widget

The threshold maps are a compilation of the various EJ or supplemental indexes on a single map. The threshold map widget allows EJScreen users to look across all 13 indexes at once, providing a cumulative outlook on vulnerable populations facing higher pollution burdens. Users select a percentile range, for example the 80th to 100th percentile, and the widget then maps places where one or more of the indexes is within that range. The widget also allows users to select the number of indexes within the user-defined range or to select specific indicators of interest. Threshold maps are available for both the EJ indexes and the supplemental indexes and are available for comparison at the national and state level.

The tool uses the data built from the following EJScreen datasets:

- National EJScreen Index Percentiles built based on the two-factor Demographic Index
- State EJScreen Index Percentiles built based on the two-factor Demographic Index
- National Supplemental Index Percentiles built based on the Supplemental Demographic Index
- State Supplemental Percentiles built based on the Supplemental Demographic Index

Datasets include these elements:

- 13 Calculated EJ Index Percentiles
- 101 Percentile bin counters (0 to 100)

Initial Filter Approach for Screening

In past screening experience, EPA has found it helpful to establish a suggested Agency starting point for the purpose of identifying geographic areas that may warrant further consideration, analysis, or outreach. The use of an initial filter promotes consistency and provides a pragmatic first step for EPA programs and regions when interpreting screening results. For early applications of EJScreen, EPA identified the 80th percentile filter as that initial starting point. In other words, an area with any of the 13 EJ Indexes at or above the 80th percentile nationally should be considered as a potential candidate for further review. Further review may include considering other factors and other sources of information such as health-based information, local knowledge, proximity and exposure to environmental hazards, susceptible populations, unique exposure pathways, and other federal, regional, state, and local data. This filter is simply a starting point, and program offices and regions should perform additional analysis before making any decisions about potential environmental justice issues. As EPA gains further experience and insight into the performance of the tool and its applicability for different uses, program offices and regions may opt to designate starting points that are more inclusive or specifically tailored to meet programmatic needs more effectively. For example, EPA's Office of Enforcement and Compliance Assurance uses at least one supplemental index above the 90th percentile

for the US or state, in the facility's Census block group or a 1-mile average around it, to identify facilities that may need additional considerations for EJ concerns.

The 80th percentile filter in EJScreen is not intended to designate an area as an “EJ community.” EJScreen provides screening level indicators, not a determination of the existence or absence of environmental justice concerns. Nor does the use of the 80th percentile filter suggest that all of the 13 environmental burden indicators are equal in terms of their impact on human health and the environment. Instead, the 80th percentile filter encourages programs to consider environmental burden indicators outside of their areas of concentration. The Agency may revise this approach in the future based on experience. This 80th percentile filter is for internal EPA use and is not intended to apply to states or other organizations.

6 Buffer Reports

EJScreen allows a user to define a buffer, such as the circle that includes everything within 1 mile of a specific point. Non-circular, user-defined shapes also can be defined to represent buffers of any shape. A report summarizes the demographics of residents within this buffer, as well as the environmental burden indicators and EJ Index values within the buffer.

The summary within a buffer is designed to represent the average resident within the buffer, and also provides an estimate of the total population residing in the buffer. For example, the traffic proximity indicator for a buffer is the population-weighted average of all the traffic indicator values in the buffer. Similarly, the percent people of color would be a weighted average, which is the same as the overall percent people of color for all residents in the buffer.

Some block groups will be partly inside and partly outside a buffer, and any buffer analysis must estimate how much of each block group's population is inside the buffer. Areal apportionment of block groups is one standard method, but it assumes that population is evenly spread throughout a block group, which may be far from the actual distribution of residents. Areal apportionment of blocks would be even more accurate but extremely computationally intensive.

To provide the most accurate counts that are currently feasible for a screening tool, EJScreen uses an approach based on decennial Census block internal points. EJScreen estimates the fraction of the Census block group population that is inside the buffer by using *block*-level population counts from the decennial Census. These blocks provide data about where residents are at a higher resolution than block groups. Each block has an internal point defined by the Census Bureau, and the entire block population is counted as inside or outside the buffer depending on whether the block internal point is inside or outside. This assumption typically introduces relatively little error because blocks are so small relative to a typical buffer, so a small fraction of the total buffer population is in blocks that span an edge of the buffer. Also, any blocks along the edge of a buffer whose populations are close to 0 or 100% inside the buffer will be well represented by this assumption.

As long as users draw buffers much larger than a local block group, this method should represent the average person inside the buffer reasonably well.

The calculation of a value for the buffer is essentially the population-weighted average of the indicator values in the blocks included in the buffer, where each block uses the indicator values of the block group containing it. A block group is weighted based on the fraction of the current ACS block group population that is considered in the buffer. That fraction is estimated as the decennial Census block population divided by the decennial Census block group population. The formula below is used to estimate the population average of a raw indicator value in a buffer. This formula is simply a population-weighted average – it sums the population-weighted raw values, and then divides that sum by the total population in the buffer.

$$Value(A) = \sum_{\forall Blk, Blk \cap A} \frac{\frac{BlockPop}{BGPop} * BGACSPop * BG_{RawValue}}{\sum_{\forall Blk, Blk \cap A} \frac{BlockPop}{BGPop} * BGACSPop}$$

"BlockPop" refers to the decennial Census block level population total (used here because the ACS does not provide block resolution), "BG" indicates block group, and "BGPop" denotes block group population. "BGACSPop" is the block group estimated population count from the current ACS, which is often different than the decennial Census total for all blocks in the block group, because the ACS data used here is a composite estimate based on survey samples spanning five years, while the decennial Census is a full count at one point in time.

7 Technical Details on Proximity Indicators

Several of EJScreen's environmental burden indicators are direct or indirect estimates of potential exposure or health risks, such as the Air Toxics cancer risk estimates and the ozone and PM 2.5 concentration estimates. There are other aspects of an individual's or a community's environmental concerns that are less readily quantified in terms of emissions, concentrations, or risk estimates.

People may be concerned about living near facilities that handle hazardous substances, and other potential sources of pollution, such as highways or abandoned waste sites. Concern over "locally undesirable land uses" (LULUs) is in some cases founded on the potential for routine or episodic releases of pollutants to the air, land, or water, and the potential for such releases to cause human health or environmental adverse effects or other societal disamenities.

The purpose of the proximity measures in EJScreen is to systematically and consistently quantify different degrees of potential for these effects. We have developed a method to calculate a score that represents the relative magnitude of the proximity of the population within a block group to facilities, waste sites, or traffic surrounding it. A block group with more facilities closer to the block group's residential population will have a higher score than a block group where facilities are further away. We have applied this method to these facility or site types:

- Superfund sites.
- Hazardous waste facilities, subject to regulations under the RCRA.
- RMP facilities, which are facilities that maintain greater than certain quantities of extremely hazardous substances, and are required to take certain actions, including filing risk management plans, under section 112 (r) of the Clean Air Act.

The Superfund polygonal data were used in conjunction with Superfund sites for which no polygons were available. In order to convert the point locations to polygons, a one-meter circle centered at each site was used to represent the site, and the small circles were appended to Superfund polygonal data to form the Superfund dataset. We have developed an approach to calculate proximity to polygonal Superfund sites.

We have also developed a similar approach to compute proximity to traffic volume on nearby highways and proximity to toxic concentrations on nearby wastewater segments.

In the sections below, we will describe the general approach, in terms of facility proximity. We will then describe how it differs for traffic and wastewater proximity. Then we will discuss certain adjustments we have made, mostly to make the approach computationally efficient, and summarize the data sources and computational routine that we applied to implement this approach. We conclude with caveats and other observations.

Calculating Proximity to Facilities

Each block group for the U.S. states, District of Columbia, and Puerto Rico is made up of between one block and several hundred blocks. Most block groups nationwide are smaller than approximately 0.5 square miles, an area that if circular would have a radius of about 640 meters. In block groups of this median size, the average residence generally would be about 430 to 720 meters (or less than half a mile) away from a given point within the block group, such as a facility, as explained at the end of this section. About 20-25% of block groups covered an area smaller than a circle of radius 300-350 meters (almost one quarter of a mile). Also, a very small number of block groups are extremely large in area, in very rural locations.

Blocks and block groups vary greatly in geographic area and in population. All blocks within a block group may or may not have residential population estimated by the 5-year ACS. Only non-zero population block groups are included in proximity calculation. The approach used here starts measuring block group level proximity to the facilities in or near the block groups. Then the distance to those facilities from each block (rather than block group) centroid within the block group is computed. The inverse distance scores are aggregated and then weighted by the number of people in each block. Thus, while population is considered in aggregating the block scores, the measure does not increase or decrease for block groups with higher or lower populations.

In EJScreen, a proximity score is considered as an inverse function of distance:

$$proximity\ score = \frac{1}{\text{distance from block centroid to site (km)}}$$

That is, a facility 1 km from a block's population contributes twice the score as a facility 2 km from the same block. We note that we have made a choice in using inverse distance for this function. Air dispersion modeling for pollutants following Gaussian plume assumptions would show a generally greater drop-off in concentration, roughly with the second power to 2.5 power of one over distance. But actual concentrations around individual plants follow often-complex patterns that depend on the particular mix of stack versus fugitive emissions, characteristics of stack height, exit velocity and temperature, the presence of buildings or other land surface characteristics, and meteorology. Some substances react readily with other substances in the atmosphere, or precipitate out readily. It is not uncommon for concentrations to rise for some distance from the emitting source, and then to fall from that peak concentration. The Gaussian plume model applies to gases, and emissions of particulates can drop off more quickly than gases.

Releases to land may follow extremely complex patterns of dispersion. Added to that are the very site-specific characteristics of potential human exposure via drinking water, vapor intrusion, contact with contaminated soils, etc. For water pollution, similar complexities exist, most notably that an effluent is carried away downstream of a running body of water, and dilution can be complicated by the presence of other water entering stream segments, by volatilization, by biological and chemical interactions, by deposition to sediments, and finally by the treatment and removal of a water pollutant sent to a publicly-owned treatment works.

We also note that researchers and others have taken varied approaches to representing the proximity of facilities to populations. The Environmental Justice Screening Method (EJSM) model of environmental justice concerns, developed for the state of California, scored facility proximity in concentric rings around a population centroid (Pastor Jr., Morello-Frosch, & Sadd, 2010; Sadd, Pastor, Morello-Frosch, Scoggins, & Jesdale, 2011). All facilities within 1 mile received a score of 3. All within the 1- to 3-mile band received a score of 2, and those between 3 and 5 miles received a score of 1. Anything beyond 5 miles received a score of zero. This step-wise scoring represents the judgment of the model developers, influenced by interactions with various stakeholders.

Finally, we note that EJScreen's measure of proximity is intended to represent more than simply real or potential human health adverse effects coming from exposure. Some parts of the environmental justice literature reflect semi-quantitative factors, such as increased psychological stress, fear, and other reactions to the presence of LULUs. This is not the forum for sorting through those factors.

However, we have made a judgment call: For the purposes of this EJScreen tool, we represent a facility's measure of proximity by the inverse of its distance from the estimated location of the average person. A block's proximity score is the sum of the inverse distances of all the facilities of a particular type.

Note that for the minority of block groups in the United States with no residential population, we assign a null value for proximity analysis indicators.

The units for these measures are facilities per km. Any block that is less than .1 km away from a facility gets a 1/d value of 10 (11 for Superfund sites).

Calculating Proximity to Hazardous Waste & RMP Sites

The hazardous waste (TSDF) and RMP proximity indicators were calculated in four steps. The first step was to overlay block groups with TSDF/RMP sites that fall within a 10km radius of each block group polygon (point to polygon). This step relates each block group with TSDF/RMP sites. The second step was to compute the distance between each block centroid within each targeted block group from the first step and the TSDF/RMP sites within the same block group (so it is a distance between point to point). Due to the parent-child relationship between block group and block, all blocks can be determined easily for a given block group. The third step was to calculate the score for each block/site combination using inverse distance. A maximum score of 10 was used for all distances under 0.1 km. The next step was to multiply each block/site combination by its block group population weight and to then aggregate the results to get the total block group score. A non-zero population block group was assigned a score of zero when no TSDF/RMP sites were found within a 10km buffer.

Calculating Proximity to Superfund Sites

The Superfund indicator proximity was calculated in four steps. The first step was to overlay block groups with Superfund polygons that fall within a 10km radius of each block group polygon (polygon to polygon). This step relates each block group with Superfund sites. The second step was to compute the distance between each block centroid within each targeted block group from the first step and the

Superfund sites within the same block group (so it is a distance between point to polygon). Due to the parent-child relationship between block group and block, all blocks can be determined easily for a given block group. The third step was to calculate the score for each block/site combination using inverse distance. A maximum score of 11 was used for all distances under 0.1 km. The next step was to multiply each block/site combination by its block population weight and to then aggregate the results to get the total block group score. A non-zero population block group was assigned a score of zero when no Superfund sites were found within a 10km buffer.

Calculating Proximity to Traffic

We have adopted essentially the same approach described above for representing proximity to highway segments—an inverse distance-weighted sum of highway segments surrounding each block, and a population-weighted sum of the individual blocks' contributions to the block group.

The highway segment database that we have used is described in section 2. These segments differ from a facility database in that they are lines on a geographic area, rather than points that represent the facilities. In our approach, we find the distance from the block centroids to each highway segment.

The traffic proximity indicator was calculated in five steps. The first step was to overlay block groups with traffic lines that fall within a 10km radius of each block group polygon (line to polygon). This step relates each block group with traffic. The second step was to compute the distance between each block centroid within each targeted block group from the first step and the traffic lines within the same block group (so it is a distance between point to polyline). Due to the parent-child relationship between block group and block, all blocks can be determined easily for a given block group. The third step was to calculate the score for each block/line combination using inverse distance. A maximum score of 10 was used for all distances under 0.1 km. Next, each block/line combination score was multiplied by the AADT estimate associated with each highway segment. This is meant to reflect the traffic intensity, and this differs from the facility approach, where we have taken each facility within each group as having equal importance. The final step was to multiply each block/line combination by its block group population weight and to then aggregate the results to get the total block group score. A non-zero population block group was assigned a score of zero when no traffic lines were found within a 10km buffer.

Calculating Proximity to Toxic Weighted Wastewater Dischargers

The proximity to toxicity-weighted wastewater discharges was calculated using the same methodology as the proximity to traffic—an inverse distance-weighted sum of water stream segments surrounding each block, and a population-weighted sum of the individual blocks' contributions to the block group. As described above in the Environmental Burden Indicators in EJScreen section, the source data is RSEI modeled output of toxicity-weighted concentrations mapped to water stream segments. So essentially EPA replaced highways with NHD stream geographies and traffic counts with toxicity-weighted concentrations.

The wastewater discharge indicator proximity was calculated in five steps. The first step was to overlay block groups with wastewater discharge lines that fall within a 10km radius of each block group polygon (line to polygon). This step relates each block group with wastewater discharge lines. The second step

was to compute the distance between each block centroid within each targeted block group from the first step and the wastewater discharge lines within the same block group (so it is a distance between point to polyline). Due to the parent-child relationship between block group and block, all blocks can be determined easily for a given block group. The third step was to calculate the score for each block/line combination using inverse distance. A maximum score of 10 was used for all distances under 0.1 km. Next, each block/line combination score was multiplied by the toxicity-weighted chemical concentration estimate associated with each wastewater segment. This is meant to reflect the toxicity concentration contribution, which differs from the facility approach where we have taken each facility within each group as having equal importance. The final step was to multiply each block/line combination by its block population weight and to then aggregate the results to get the total block group score. A non-zero population block group was assigned a score of zero when no wastewater discharger lines were found within a 10km buffer.

Calculating Proximity – Additional Details

Proximity Calculations—block group proximity scores for Superfund (polygon), RMP (point), Hazardous Waste (point), and Wastewater Discharges (line) were calculated with new source data. The scores for traffic used the same source data as those used for previous year. Census block centroids with population-weights are unchanged. These block points and associated populations and block group weights were derived from 2020 Decennial Census P.L. 94-171 Redistricting data.

We address two modifications to the general method described above. The first deals with instances where a facility or highway segment location is very close to the centroid of the block. The second is an accommodation to the computational intensity of the general method.

Extremely Small Block/Site Distance Values

Our intention is to represent the proximity of facilities, wastewater dischargers, or highway segments to the population within each block. All facilities and each part of all highway segments fall within one block. By chance, some portion of those points fall very close to the block centroids.

We do not know how the population is geographically distributed within any block, but we assume that people are more likely to be distributed across the blocks' expanses than to be concentrated at one point, such as the centroid. In fact, for rural, suburban, and many non-high rise urban areas, people's residences are more likely to be closer to the blocks' peripheries (bounded by roads) than clustered at the centroids. Thus, when a facility location happens to be very close to the block centroid, it would result in an artificially high contribution to the block's score. This is not a hypothetical problem: We have observed block/site distance values well below 100 meters, and some below 10 meters.

In looking for solutions to the problem, we conducted analyses and arrived at the approach we have adopted. Blocks vary widely in their total area and in their shapes. Both can be found in the Census Bureau's TIGER shape files. Dealing explicitly with the individual block shapes would be computationally

very intensive because there are over 11 million blocks. Since we cannot easily find out how the residents are actually distributed in those areas, we made two simplifying assumptions:

- residents are evenly distributed across the surface area of each block, and
- each block can be represented by a circle whose radius is a square root of block area divided by the circumference ratio.

We call this latter value the Block Area Equivalent Radius.

Our investigations indicate that for any block/site distance less than the Block Area Equivalent Radius, 0.9 times that value is a reasonable representation of the average distance from the facility for all residents in the block. We call this the block/site distance corrected.

Our computational scheme determines the block/site distance values as described above, tests for the comparison with Block Area Equivalent Radius, and substitutes individual corrected values. We found that we needed to make that correction for less than 1% of all facility/block combinations in an early testing dataset that used 2005-2009 ACS data.

Caveats and Observations

Several aspects of the proximity analysis approach have been mentioned above but deserve summary here.

- We recognize that our selection of the inverse of distance is a design choice that represents our judgment of a balance among competing factors.
- We recognize that one could potentially attempt to distinguish among facilities within each facility category by quantitative or qualitative measures of importance. These could include total pounds released or toxicity-weighted releases for NPDES facilities; the number of accidental releases and/or their apparent severity for RMP facilities; some classification of the likelihood of releases for Superfund sites or TSDFs; and general indications of scale for all of them. We note that CalEnviroScreen has addressed this issue to some extent, and that the RSEI tool based on TRI data may be relevant to future work on this issue. At this point, we have chosen not to develop any such potential scaling adjustments.
- We recognize that all location data are subject to potential error. While we have high confidence in the block centroid locations, we know that the facility or site or roadway location data may contain larger or smaller errors, and that for large facilities or sites, one point may not be an entirely adequate representation of the location of its releases or of neighbors' perceptions.

8 Details on U.S. Territories

EJScreen features select environmental and socioeconomic indicators for the following U.S. Territories: American Samoa, CNMI, Guam, Puerto Rico, and the U.S. Virgin Islands. American Samoa, CNMI, and Guam use 2022 Census Place boundaries. The U.S. Virgin Islands use 2022 Census Estate boundaries. However, Puerto Rico is included with all 50 states and the District of Columbia. The term "territories" in this document refers to the four territories excluding Puerto Rico. The territories data for American Samoa, CNMI, Guam, and the U.S. Virgin Islands provides an additional 686 records. These territories have state percentiles data and do not have national percentiles.

Data Sources

- Census boundaries (Places, Estates, and Counties) are from Cartographic Boundary files, 2022, 1:500K scale.
- Demographic data—2020 Demographic Profiles of Island Areas published in 2022.

Processing Steps

- Generated the block population weight table based on distributing 2020 population for Estate or Place boundaries
- Built socioeconomic indicator tables by extracting the Place/Estate summary level data from the Demographic Profiles. Indicators include: % people of color; % low income; % limited English speaking; % less than high school education; % under age 5; and % over age 64
- Calculated Demographic Index: $\%\text{Low Income}_{\text{z-score}} + \%\text{People of Color}_{\text{z-score}} / 2$
- Collected all the available environmental data sources
- Calculated EJ Indexes.

Index Availability for the Territories

	American Samoa	NMI	Guam	U.S. Virgin Islands
PM 2.5	No	No	No	No
Ozone	No	No	No	No
Nitrogen Dioxide (NO_2)	No	No	No	No
Diesel Particulate Matter	No	No	No	No
Toxic Releases to Air	Yes	Yes	Yes	Yes
Traffic Proximity	No	No	No	No
Lead Paint	Yes	Yes	Yes	Yes
Superfund Proximity	No	No	Yes	Yes
RMP Facility Proximity	No	No	Yes	Yes
Hazardous Waste Proximity	No	Yes	Yes	Yes
Underground Storage Tanks	Yes	Yes	Yes	Yes
Wastewater Discharge	No	No	No	No
Drinking Water Non-Compliance	No	No	No	No

Notes about Puerto Rico

- While Puerto Rico is a U.S. Territory, it is included in the Census ACS 5-year Summary datasets, so EJScreen treats it like a state and features both national and state-level percentiles and indexes.

9 Other Data Element Descriptions

This section describes additional datasets now available in EJScreen. They include health disparities, climate change data, and critical service gaps. These data are from a variety of data sources, including CDC, Census, First Street, EPA, National Oceanic and Atmospheric Administration (NOAA), Climate and Economic Justice Screening Tool (CEJST), and U.S. Department of Agriculture (USDA).

Health Disparities Data

- **Low Life Expectancy**—Average life expectancy data developed as a collaboration between NCHS, the National Association for Public Health Statistics and Information Systems (NAPHSIS), and the Robert Wood Johnson Foundation. This data is available at the tract level. Source: U.S. Small-area Life Expectancy Estimates Project (USALEEP)
  It is derived from Life Expectancy at Birth from CDC, National Center for Health Statistics using the formula of % Low Life Expectancy is defined as "1 – (Life Expectancy / Max Life Expectancy)". The following processing steps were used to bring the data into EJScreen:
  - The source is Census 2015 tract-level data, so it was first converted to 2010 tracts using the Census 2010 to 2015 relationship table.
  - 2010 tracts were then converted to 2020 tracts using the Census 2010 to 2020 relationship table, and were then converted to Census 2022 tracts.
  - Low Life Expectancy values were assigned to each child 2022 block group within the same Census tract.
- **Heart Disease**—Heart disease prevalence among adults aged 18 years or older. The term "heart disease" refers to several types of heart conditions. This data is available at the Census tract level; each value is then assigned to all child block groups. Source: Updated with January 2024 release of CDC Places Data
- **Asthma**—Asthma prevalence among adults aged 18 or older. This data is available at the Census tract level; each value is then assigned to all child block groups. Source: Updated with January 2024 release of CDC Places Data
- **Cancer**—Cancer (excluding skin cancer) prevalence among adults aged 18 or older. This data is available at the Census tract level; each value is then assigned to all child block groups. Source: Updated with January 2024 release of CDC Places Data
- **Persons with Disabilities**—Percent of all persons with disabilities. This data is derived from 2022 ACS "Sex by Age by Disability Status" table (B18101) for Census tracts. Block group values are calculated by multiplying the tract value by the block group population weights. The weights are derived from the same Census source used by the EJScreen buffer reports and analysis—2020 Decennial Census P.L. 94-171 Redistricting data. Source: Download ACS Data

Climate Change Data

- **Wildfire and Flood Risk Data from First Street**

First Street is partnering with EPA to provide climate risk data to EJScreen. First Street provided Census block group-level data, including percent of properties at risk of being affected by wildfires or flooding based on current climate conditions as well as conditions projected to exist in 30 years. EJScreen merged the data with current EJScreen spatial data and generated national and state percentiles. EJScreen includes:

- **Wildfire Risk**—The household risk of wildfire exposure under 2022 weather conditions as modeled by First Street. Source: https://firststreet.org/
- **Flood Risk**—The household risk of flooding under 2022 weather conditions as modeled by First Street. Source: https://firststreet.org/

The First Street Wildfire Model (FS-WFM) is a 30-meter resolution model representing the wildfire exposure for any specific location in the contiguous U.S., today and with the future climate change. The risk of wildfire is derived from a series of inputs associated with fire fuels, weather, human influence, and fire movement. Bringing all of these inputs together, at a national scale, in a high-resolution, climate-adjusted model represents a first-of-its-kind property-level wildfire risk model.

The First Street Flood Model (FS-FM) is a nationwide probabilistic flood model that shows the risk of flooding at any location in all 50 states and Puerto Rico due to rainfall (pluvial), riverine flooding (fluvial), and coastal surge flooding. While other hydraulic and hydrologic models show refined risks of flooding in certain areas, this model provides complete coverage across the United States at 3-meter resolution.

- **Estimated 100-year floodplains**

This map depicts the estimated 100-year floodplains for the Conterminous U.S. for 2016. These data can be used to look for socially and economically vulnerable communities located within the estimated 100-year floodplains. Source: EnviroAtlas

- **Sea Level Rise (NOAA)**

This dataset depicts land at risk of permanent flooding when sea level rises. There are six map layers showing the area inundated at one, two, three, four, five, and six feet of sea level rise. Over the next 30 years, scientists estimate a one- to three-foot rise along most of the U.S. coastline. These data can be used to look for areas of potential permanent inundation based on chosen sea level rise estimates (i.e., 1 to 6 feet) in relation to locations of social and economic vulnerability. Source: National Oceanic and Atmospheric Administration

- **Extreme Heat**

This layer maps maximum days above 90 degrees. Dates are from 01/01/2019 to 12/31/2023. Map data is derived by EPA from PRISM Climate Group.

Critical Service Gaps

- **Broadband Gaps**—Maps areas of households without a broadband internet subscription. This data is derived from 2022 ACS "Presence and Types of Internet Subscriptions in Household" table (B28002) for Census block groups. Source: Download ACS Data
- **Lack of Health Insurance**—Percent of all persons without Health Insurance Coverage. This data is derived from 2022 ACS "Health Insurance Coverage Status by Sex by Age" table (B27001) for Census tracts. Block group values are calculated by multiplying the tract value by the block group population weights. The weights are derived from the same Census source used by the EJScreen buffer reports and analysis—2020 Decennial Census P.L. 94-171 Redistricting data. Source: Download ACS Data
- **Housing Burden**—The data were downloaded from the Climate and Economic Justice Screening Tool (CEJST) website. This dataset contains census tract-level percentiles for housing cost, which is the share of households that are both earning less than 80% of Housing and Urban Development's Area Median Family Income and are spending more than 30% of their income on housing costs. The housing cost percentiles were adopted as Housing Burden for EJScreen.
- **Transportation Access Burden** —The data were downloaded from DOT's Transportation Disadvantaged Census Tracts website. The Average of Transportation Indicator was adopted as Transportation Access for EJScreen. According to DOT, it is an average of four transportation-related indicator percentiles, including Transportation Cost Burden, National Walkability Index, Percentage of Households with No Vehicle Available, and Mean Commute Time to Work.
- **Food Desert**—Low income and low access tract measured at 1 mile for urban areas and 10 miles for rural areas. This data is available at the tract level. Source: USDA

EJ Indexes and Supplemental Indexes Availability by Area

EJScreen features 13 EJ Indexes and 13 Supplemental Indexes that have varying levels of spatial coverage because of the differences in the source datasets. The following table provides an overview for EJ Indexes:

EJ Indexes	Lower 48	Alaska	Hawaii	Puerto Rico	American Samoa	CNMI	Guam	U.S. Virgin Islands
PM 2.5	Yes	No	No	No	No	No	No	No
Ozone	Yes	No	No	No	No	No	No	No
Nitrogen Dioxide (NO_2)	Yes	Yes	Yes	Yes	No	No	No	No
Diesel Particulate Matter	Yes	No	Yes	Yes	No	No	No	No
Toxic Releases to Air	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Traffic Proximity	Yes	Yes	Yes	Yes	No	No	No	No
Lead Paint	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Superfund Proximity	Yes	Yes	Yes	Yes	No	No	Yes	Yes
RMP Facility Proximity	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Hazardous Waste Proximity	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Underground Storage Tanks	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wastewater Discharge	Yes	No	Yes	Yes	No	No	No	No
Drinking Water Non-Compliance	Yes	Yes	Yes	No	No	No	No	No

The following table provides an overview for Supplemental Indexes:

Supplemental Indexes	Lower 48	Alaska	Hawaii	Puerto Rico	American Samoa	CNMI	Guam	U.S. Virgin Islands
PM 2.5	Yes	No	No	No	No	No	No	No
Ozone	Yes	No	No	No	No	No	No	No
Nitrogen Dioxide (NO_2)	Yes	Yes	Yes	Yes	No	No	No	No
Diesel Particulate Matter	Yes	No	Yes	Yes	No	No	No	No
Toxic Releases to Air	Yes	Yes	Yes	Yes	No	No	No	No
Traffic Proximity	Yes	Yes	Yes	Yes	No	No	No	No
Lead Paint	Yes	Yes	Yes	Yes	No	No	No	No
Superfund Proximity	Yes	Yes	Yes	Yes	No	No	No	No
RMP Facility Proximity	Yes	Yes	Yes	Yes	No	No	No	No
Hazardous Waste Proximity	Yes	Yes	Yes	Yes	No	No	No	No
Underground Storage Tanks	Yes	Yes	Yes	Yes	No	No	No	No
Wastewater Discharge	Yes	No	Yes	Yes	No	No	No	No
Drinking Water Non-Compliance	Yes	Yes	Yes	No	No	No	No	No