Methodology for the [Petrochemical Air Pollution Map](#) (version 1.0 released January 2025)

Contents

Petrochemical industry risk

The default map view visualizes the modeled risk to human health caused by inhalation of hazardous air pollution released by petrochemical facilities into ambient air. Risk accounts for the combined air releases from multiple pollutants and facilities that have modeled data available. The data displayed are toxicity-weighted concentrations, which combine the modeled amount of chemicals in the air with the relative toxicity of the chemicals. Users can also filter the map for a set of priority chemicals to see the risk patterns for individual chemicals across the country.

Data source: [EPA Risk Screening Environmental Indicators (RSEI) Model Geographic Microdata](#)

Year: 2020 (most recent year with data from both EPA models, RSEI and AirToxScreen)

Resolution: The RSEI model produces estimates at 810-meter grid cell resolution. We used results averaged over census block group geographies.

Facilities included: We selected facilities specific to the petrochemical industry and supply chain if any of their [North American Industry Classification System (NAICS)](#) codes (primary or non-primary) matched the industries specified below.
- Chemical Manufacturing: NAICS 325
- Plastics and Rubber Products Manufacturing: NAICS 326
- Petroleum Sector: NAICS 324
- Other Chemical and Allied Products Merchant Wholesalers: NAICS 424690
- Petroleum Bulk Stations and Terminals: NAICS 424710
- Other Warehousing and Storage: NAICS 493190

Types of releases modeled: Stack and fugitive air releases

Toxicity weighting: We used cancer and noncancer toxicity weights (tox weights) provided in the [RSEI Chemical Data](#) for the inhalation exposure pathway. While the RSEI toxicity-weighted concentrations calculated by EPA use the greater of the cancer or non-cancer tox weights for each chemical, we instead calculated and displayed cancer and non-cancer toxicity-weighted concentrations separately in the map, allowing the user to select which health effect type to view. Chemicals with no inhalation toxicity weights available are excluded from the risk estimates which are therefore expected to be a lower bound. Of the 366 chemicals with modeled releases from the industries specified above, 71 chemicals had inhalation toxicity weights available for cancer and 146 had inhalation tox weights for noncancer.

Data aggregation: For each chemical, we calculated the cancer and non-cancer toxicity-weighted concentration by multiplying the tox weight by the 2020 modeled concentration in each block group. To calculate the multipollutant toxicity-weighted concentration, we summed the 2020 toxicity-weighted concentration of all modeled chemicals in each block group.

Chemical prioritization for filter: Due to data size limitations we could not provide the option to filter the map by all the chemicals included in the multipollutant risk layer (the default view). We prioritized a group of 38 filterable chemicals that (1) contributed the most to risk nationally and/or (2) contributed the most to risk locally. We used 2020 toxicity-weighted concentrations in each census block group as a starting point. For national prioritization, we calculated the national sum of toxicity-weighted concentrations by chemical and ranked them from highest to lowest, selecting the top 15 chemicals each for cancer and non-cancer effects. For local prioritization, we calculated the maximum toxicity-weighted concentration by chemical – the toxicity-weighted concentration in the highest census block group in the country for that chemical – also ranking them from highest to lowest and selecting the top 15 chemicals each for cancer and non-cancer effects. We then merged the top 15 chemicals from the national and local results into final lists. Most chemicals were identified in both the national (total risk) and local (maximum risk) lists, resulting in 20 prioritized chemicals for noncancer effects and 18 for cancer effects. Users should contact dpeters@edf.org with requests for chemicals to add if the map is updated in the future.

Risk from all modeled sources (including petrochemical facilities)

This map overlay visualizes multipollutant toxicity-weighted concentrations from AirToxScreen model data. The AirToxScreen model includes emissions from a broader set of sources than the petrochemical RSEI layer above, including point sources, mobile sources (like cars), fires, and biogenic sources. The model also includes estimates of background and secondary pollutant concentrations.

Data source: [EPA Air Toxics Screening Assessment (AirToxScreen) Model](#)

Years: 2020 (most recent year available at release)

Resolution: 2020 AirToxScreen results are at census block resolution. We averaged the results over census block group geographies.

Toxicity weighting: We applied the same toxicity weighting methodology described above for RSEI to AirToxScreen ambient concentration data. While AirToxScreen does report exposure concentrations in addition to ambient concentrations, we used ambient concentrations to be

consistent with the RSEI toxicity-weighted concentration methodology. AirToxScreen reports concentrations in aggregate for some chemical groups including glycol ethers and PAHPOMs. Following the RSEI [methodology](#), we applied the toxicity weighting for ethylene glycol methyl ether, the most toxic glycol ether based on available toxicity info, to the entire glycol ethers group. For the AirToxScreen PAHPOM category, we applied the toxicity values from the corresponding RSEI chemical group Polycyclic Aromatic Compounds.

Data aggregation: For each chemical modeled by AirToxScreen with available toxicity data, we calculated the toxicity-weighted concentration by multiplying the tox weight by the 2020 modeled concentration in each census block. To calculate the multipollutant toxicity-weighted concentration, we summed the 2020 toxicity-weighted concentration of all modeled chemicals in each block. Then we averaged the toxicity-weighted concentrations by census block group.

Top drivers of risk

Clicking a census block group on the map brings up a popup with the top drivers of risk in that location. The popup shows what percentage of modeled petrochemical risk (RSEI model) is coming from specific facilities and their chemical releases, as well as which sectors contribute most to overall modeled impacts from a broader group of sources like transportation and secondary pollution (AirToxScreen model).

Years: We used RSEI and AirToxScreen data for the year 2020 (same as displayed on map).

Petrochemical facilities apportionment (RSEI model):

We first calculated the total petrochemical toxicity-weighted concentration by summing the toxicity-weighted concentration for all modeled chemicals and facilities in each block group.

We then calculated the percentage of the toxicity-weighted concentration attributable to releases of a specific chemical by a specific petrochemical facility in each census block group by dividing the toxicity-weighted concentration from these individual releases by the total toxicity-weighted concentration from all included petrochemical facilities and chemicals.

We displayed the percentages of the top five facility/chemical combinations contributing to petrochemical risk in each census block group.

Source category apportionment (AirToxScreen model):

The AirToxScreen 2020 model provided concentration estimates for 39 pollution source groups (see Table 2-13 of the [AirToxScreen Technical Support Document](#) for more information). We further grouped these sources into five broad source categories, as shown in Table 1 below. We summed the toxicity-weighted concentrations by source category in each census block and spatially averaged up to census block groups. Then we calculated the proportion of toxicity-weighted concentration attributable to each source category as a percentage of the total for the block group.

Table 1. AirToxScreen source categories

Source category for popup	AirToxScreen source group
Industrial*	pt_stationary_point_conc np_industrial_conc nr_all_other_conc np_oil_gas_conc np_storage_transfer_bulk_terminals_gas_stage1_conc
Road Transportation	or_light_duty_off_network_gas_conc or_light_duty_off_network_diesel_conc or_heavy_duty_off_network_gas_conc or_heavy_duty_off_network_diesel_conc or_light_duty_on_network_gas_conc or_light_duty_on_network_diesel_conc or_heavy_duty_on_network_gas_conc or_heavy_duty_on_network_diesel_conc or_refueling_conc or_heavy_duty_hoteling_conc
Shipping/Rail/Airports	nr_cmv_c1c2_ports_conc nr_cmv_c3_ports_conc nr_cmv_c1c2c3_underway_conc nr_locomotives_conc nr_point_airports_conc nr_point_railyards_conc
Indirect and Biogenic*	fire_conc biogenics_conc secondary_conc background_conc
Other	np_commercial_cooking_conc nr_commercial_equipment_conc np_fuel_combustion_not_rwc_conc np_waste_disposal_conc nr_construction_conc np_agriculture_livestock_waste_conc np_agriculture_livestock_silage_conc nr_agriculture_conc nr_commercial_lawn_garden_conc nr_residential_lawn_garden_conc np_miscellaneous_nonindustrial_conc nr_recreational_inc_pleasure_craft_conc np_residential_wood_combustion_rwc_conc np_solvents_coatings_conc

*Some chemicals form secondarily in the atmosphere from other chemical precursors. The "Indirect and Biogenic" category includes estimates of chemicals like formaldehyde that can be formed as secondary pollutants. Formaldehyde precursors include biogenic and industrial emissions, but the AirToxScreen output does not differentiate between these sources of secondary pollution. Since AirToxScreen doesn't report the amount of secondary formaldehyde that is attributable to industrial releases, the Industrial share is likely an underestimate in places where facilities emit precursors to formaldehyde and other hazardous chemicals.

Monitoring sites

The map displays the location of ambient monitoring sites that are included in EPA's Ambient Monitoring Archive, with a popup that shows measurement values and comparisons to model estimates. Users can filter the map by chemical to view only the monitoring sites that detect a specific chemical.

Data source: [EPA Ambient Monitoring Archive for HAPs](#) - Annual Statistics [file](#)

Years: 2020 (consistent with other data layers on map)

Measured concentration: We used annual mean concentrations that EPA calculated for each monitor and chemical. EPA requires at least three valid quarters to generate an annual average. For more details, see their [methodology](#).

Modeled concentration: To compare measured pollution levels with those estimated by EPA's AirToxScreen model, we included ambient concentration estimates from AirToxScreen for the census block where the monitor was located. Not all measured chemicals are modeled by AirToxScreen, so not all measurements have a corresponding model estimate and bias factor.

Bias factor: To compare the measured concentration with the AirToxScreen modeled concentration, we calculated a bias factor for each chemical where 2020 monitor and model data was available. The bias factor is the ratio of the measured concentration to the modeled concentration (measured / modeled), so a bias factor greater than 1 means that a chemical's measured concentration at that site (and therefore the actual health risk) was higher than EPA's model estimate. [Padilla et al. (2024)](#) presents a more in-depth, multi-year analysis of model bias nationally compared to measurements.

Facility releases

We display the locations of petrochemical facilities that report to the Toxics Release Inventory (TRI) and the pounds of air emissions that they reported releasing in 2020. Users can filter the map by chemical to view only the facilities that release a specific chemical.

Data source: [RSEI Releases Data](#) (based on [Toxics Release Inventory data](#))

Years: 2020 (consistent with other data layers on map)

Facilities included: We selected facilities specific to the petrochemical industry and supply chain if any of their [North American Industry Classification System (NAICS)](#) codes (primary or non-primary) matched the industries specified below.
- Chemical Manufacturing: NAICS 325
- Plastics and Rubber Products Manufacturing: NAICS 326
- Petroleum Sector: NAICS 324
- Other Chemical and Allied Products Merchant Wholesalers: NAICS 424690
- Petroleum Bulk Stations and Terminals: NAICS 424710
- Other Warehousing and Storage: NAICS 493190

Types of releases included: Stack and fugitive air releases.

Data aggregation: For each facility we summed the pounds of 2020 emissions by chemical.

Facility compliance

The facility popup contains information on the recent Clean Air Act compliance history for each facility.

Data source: [EPA Enforcement and Compliance History Online (ECHO) Exporter Version 2.0](#)

Date accessed: December 2, 2024. The file is updated by EPA weekly, and at the time of download it was current to November 30, 2024.

Enforcement metrics: We selected three data points for facility compliance and enforcement history related to the Clean Air Act. Learn more in EPA's [data dictionary](#).
- High priority violator status
  - Status as of November 30, 2024
- Quarters with noncompliance (out of 12)
  - From October 2021 to November 2024
- Number of formal enforcement actions
  - From October 2019 to November 2024

HON facilities

Facilities highlighted in blue are subject to Hazardous Organic NESHAP standards (otherwise known as the HON Rule). Learn more about this 2024 rule on the [EPA website](#) and in the [Federal Register](#).

Data source: [List of Facilities Subject to the HON and Group I and Group II Polymers and Resins NESHAP](#) (EPA)

NAAQS nonattainment areas

While the pollution visualized on the map is limited to hazardous air pollutants, we include a layer on county-level nonattainment for the National Ambient Air Quality Standards (NAAQS) to provide context about other types of air pollutants.

Data source: EPA [Green Book Data Download](#) PHISTORY export

Date accessed: June 7, 2024. At time of download the data was current to May 31, 2024.

Years: 2020 – 2024 (most recent 5 years available)

Method: We identified all counties that were in whole or partial nonattainment for any active NAAQS (includes $PM_{2.5}$, PM_{10}, Ozone, Lead, and SO_2) during any years from 2020 to 2024. For example, a county would be displayed on the map if it was in a nonattainment area for ozone in 2020, or if it was in a nonattainment area for multiple pollutants during multiple years.

Limitations

In addition to the limitations listed below, please see EPA's detailed considerations for using [RSEI](), [AirToxScreen](), and [Toxics Release Inventory]() data.

- Both RSEI and AirToxScreen are modeled estimates based on self-reported industry emissions data and not all facilities meet the thresholds for reporting to TRI.
- There are many limitations and challenges related to modeled and self-reported data. Modeled concentrations based on approximate emission inventories routinely underestimate measured concentrations ([Padilla et al., 2024]())
- Toxicity-weighted concentrations are a relative metric and do not directly represent the rate of a specific health outcome. They are appropriate for relative comparisons of potential impacts, such as between chemicals, sectors or geographic locations. For more details see: [RSEI Toxicity-Data and Calculations]()
- Toxicity-weighted concentrations are based on ambient concentration estimates; they are not adjusted based on exposure factors or population estimates.
- RSEI models within a 50km radius of releases, so potential risks from emissions outside of this range are not accounted for in the model.
- The map only represents RSEI and AirToxScreen data from 2020. Emissions in 2020 were affected by the COVID-19 pandemic, with some TRI facilities [reporting shutdowns or reducing operations](). Emissions and potential impacts in other years may be different.
- Annual average exposure estimates do not capture the possibility of shorter-term higher exposure that may cause more acute health harm. In addition, the toxicity weights only address chronic toxicity from long-term exposure and do not consider acute exposure.
- Not all pollutants have health effects studies to provide the necessary toxicity weights for inclusion in a multi-pollutant risk estimate. Health risk from pollutants without quantified health effect data is not included in this map.
- Census block groups vary in size and may be large in rural areas, averaging over sub-block level spatial gradients and hiding small but concerning hot spots.
- Health risks are grouped generically into cancer and non-cancer without more granular explanation of potential adverse health effects.
- Outdoor air is only one of several environmental media (e.g. soil, water, food, personal products) that expose people to pollution.
- Multipollutant risk is estimated but not cumulative impact (which would include additional stressors not related to chemicals or air pollution).

Data download

The data download contains a csv file with the map data for each census block group. The file has about 240,000 rows (one for each block group with results). The columns of the csv file are described in Table 2 below.

Table 2. Data download file schema

Columns	Column name or prefix	Description
1	GEOID	Unique identifier for 2020 U.S. census block groups
2	state	Two letter U.S. state abbreviation
3	County	Name of U.S. county or parish
4-5	rsei_multipollutant_toxconc_[cancer/noncancer]	RSEI toxicity-weighted concentrations from petrochemical facilities for all modeled chemicals. Separate columns for cancer and non-cancer effects. See Petrochemical industry risk for methodological details.
6-7	rsei_multipollutant_percentile_[cancer/noncancer]	National percentile (between 0 and 1) of RSEI petrochemical multipollutant toxicity-weighted concentrations. Separate columns for cancer and non-cancer effects.
8-47	rsei_top_facilityname_[1:5]_[cancer/noncancer] rsei_top_facilitynumber_[1:5]_[cancer/noncancer] rsei_top_chemical_[1:5]_[cancer/noncancer] rsei_top_percent_[1:5]_[cancer/noncancer]	Top five drivers of petrochemical risk for cancer and non-cancer effects. The facility name, number, chemical are listed as well as the percent of risk attributed to each facility/chemical combination. See Top drivers of risk for methodological details.
48-85	rsei_[chemical]_toxconc_[cancer/noncancer]	Chemical-specific RSEI toxicity-weighted concentrations from petrochemical facilities for a subset of priority chemicals that are included in the map filter. Separate columns for cancer and non-cancer effects. See Petrochemical industry risk for methodological details.
86-87	airtoxscreen_multipollutant_toxconc_[cancer/noncancer]	AirToxScreen toxicity-weighted concentrations for all modeled chemicals. Separate columns for cancer and non-cancer effects. See Risk from all modeled sources for methodological details.
88-89	airtoxscreen_multipollutant_percentile_[cancer/noncancer]	National percentile (between 0 and 1) of AirToxScreen multipollutant toxicity-weighted concentrations. Separate columns for cancer and non-cancer effects.
90-99	airtoxscreen_[sector]_percent_[cancer/noncancer]	Percent of air toxics risk attributed to each AirToxScreen source sector grouping. Separate columns for cancer and non-cancer effects. See Top drivers of risk for methodological details.