HRS|CDR Health and Retirement Study
Contextual Data Resource

Documentation for

Ozone FAQSD Files

Version 3.0, released January 2025

Jennifer Ailshire, Sung Eun Cho, Jong Woo Nam, and Eun Young Choi

A project of the USC/UCLA Center on Biodemography and Population Health
sponsored by the National Institute on Aging (R21 AG045625, P30 AG017625).

Requested Acknowledgment

If you used the Contextual Data Resource data in a written analysis, please include the following acknowledgement:

> This analysis uses data or information from the Contextual Data Resource (CDR): United States Environmental Protection Agency Ozone FAQSD Files by Census Tract, 2002-2021, Version 3.0 as of January 2025, developed by Jennifer Ailshire, Sung Eun Cho, Jong Woo Nam, and Eun Young Choi at the USC/UCLA Center on Biodemography and Population Health. The development of the CDR was funded by the National Institute on Aging (R21 AG045625, P30 AG017625). For more information, please refer to https://hrs.isr.umich.edu/data-products/restricted-data/available-products.

Suggested Citation

Jennifer Ailshire, Sung Eun Cho, Jong Woo Nam, Eun Young Choi. 2025. Contextual Data Resource (CDR): United States Environmental Protection Agency Ozone FAQSD Files by Census Tract, 2002-2021, Version 3.0. Los Angeles, CA: USC/UCLA Center on Biodemography and Population Health.

Data Sources

United States Environmental Protection Agency (EPA). RSIG and FAQSD Files, https://www.epa.gov/hesc/rsig-related-downloadable-data-files

Version Note

Version 3.0 updates the data with the most recent 2021 measures. Variables names are consistent with earlier versions.

Linking CDR Data with HRS Respondent Data

The Contextual Data Resource is designed to be linked with the Health and Retirement Study Cross-Wave Geographic Information (Detail) file, which is available as restricted data. See https://hrs.isr.umich.edu/data-products/restricted-data for more information about accessing the HRS restricted data. All geographic identifiers are in string format, and match the geographic identifiers in the HRS Cross-Wave Geographic Information (Detail) file.

- To link census tract level data, merge using the 11-digit LINKCEN2010 geographic identifier, which is the 2-digit state + 3-digit county + 6-digit census tract FIPS code. The LINKCEN2010 geographic identifier has been harmonized to the 2010 census tract boundaries.

Note on HRS Cross-Wave Geographic Information File

Respondent locations are carried forward to the next wave and they appear for waves in which the interview may not have occurred. xIWTYPE in the Tracker file indicates whether someone was actually interviewed in that wave, in which case, the location would have been confirmed.

If users want to limit analyses to waves in which the R was interviewed, keep waves of observations where xIWTYPE = 1. In summary, a location appearing on a given line does not indicate that an interview took place in that wave; it is simply the location that was on record at the time.

Dataset Overview

Ozone (O_3) is a reactive gas which can cause irritation and adverse health effects when inhaled. The US Environmental Protection Agency Fused Air Quality Surface Using Downscaling (FAQSD) Files provide estimates for local average concentrations of O_3 (ppb) across the United States using monitoring station data and Community Multiscale Air Quality (CMAQ) output. Data are available at census tract level and include annual, quarterly, and monthly averages for 2002 through 2021.

Data Summary

Dataset Name: Fused Air Quality Surface Using Downscaling (FAQSD) Files
Data Source: United States Environmental Protection Agency (EPA)
Data Source URL: https://www.epa.gov/hesc/rsig-related-downloadable-data-files
Data Collection Method: Primary data collected by the EPA.
Years Collected: All data contains measures between 2002 and 2021.
Geographic Level: Census tract; HRS-CDR uses the daily predictions available at the 2010 census tract boundaries to be comparable across the years.
Geographic Coverage: FAQSD Files cover the contiguous U.S.

Technical Information about Fused Air Quality Surface Using Downscaling

The downscaling fusion model uses both air quality monitoring data from the National Air Monitoring Stations/State and Local Air Monitoring Stations (NAMS/SLAMS) and numerical output from the Models-3/Community Multiscale Air Quality (CMAQ).

The FAQSD model is a Bayesian space-time downscaler model which integrates census-tract level 24-h average monitoring data from the National Air Monitoring Stations and State and Local Air Monitoring Stations (NAMS/SLAMS) with 12 km gridded output from the Models-3/Community Multiscale Air Quality (CMAQ) model. The CMAQ model uses emissions data from the EPA's National Emissions Inventory and includes model emissions, daily continuous emissions monitoring data for significant point sources, and meteorological data.

There are areas with few or no air monitoring sites. In these areas, it may be difficult to adjust gridded CMAQ output to provide accurate predictions of air quality. For such areas, the bias-adjustments have been made based on the available monitoring data which may be located in areas with different emissions and geography.

For further information on the fused air quality surface using downscaling files, please see following published journal papers:

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2010). A bivariate space-time downscaler under space and time misalignment. *The annals of applied statistics*, *4*(4), 1942.

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2010). A spatio-temporal downscaler for output from numerical models. *Journal of agricultural, biological, and environmental statistics*, *15*(2), 176-197.

Berrocal, V. J., Gelfand, A. E., & Holland, D. M. (2012). Space-Time Data fusion Under Error in Computer Model Output: An Application to Modeling Air Quality. *Biometrics*, *68*(3), 837-848.

Bodnaruk, E. W., Kroll, C. N., Yang, Y., Hirabayashi, S., Nowak, D. J., & Endreny, T. A. (2017). Where to plant urban trees? A spatially explicit methodology to explore ecosystem service tradeoffs. *Landscape and Urban Planning*, *157*, 457-467.

US EPA. (2012a). Technical Information about Fused Air Quality Surface Using Downscaling Tool: Metadata Description. Available online: https://www.epa.gov/sites/production/files/2016-07/documents/data_fusion_meta_file_july_2016.pdf

Variable List

Name	Definition	Geography	Years
	Geographic Identifiers		
LINKCEN2010	2010 Census Tract FIPS code (11 digits: 1-2 state, 3-5 county, 6-11 tract)		
	Ozone Annual Average Estimates		
p305tr	Mean O_3 Annual	Tract	2002-2021
	Ozone Quarterly Average Estimates		
p306tr	Mean O_3 (Q1: Jan-Mar)	Tract	2002-2021
p307tr	Mean O_3 (Q2: Apr-Jun)	Tract	2002-2021
p308tr	Mean O_3 (Q3: Jul-Sep)	Tract	2002-2021
p309tr	Mean O_3 (Q4: Oct-Dec)	Tract	2002-2021
	Ozone Monthly Average Estimates		
p310tr	Mean O_3 (Jan)	Tract	2002-2021
p311tr	Mean O_3 (Feb)	Tract	2002-2021
p312tr	Mean O_3 (Mar)	Tract	2002-2021
p313tr	Mean O_3 (Apr)	Tract	2002-2021
p314tr	Mean O_3 (May)	Tract	2002-2021
p315tr	Mean O_3 (Jun)	Tract	2002-2021
p316tr	Mean O_3 (Jul)	Tract	2002-2021
p317tr	Mean O_3 (Aug)	Tract	2002-2021
p318tr	Mean O_3 (Sep)	Tract	2002-2021
p319tr	Mean O_3 (Oct)	Tract	2002-2021
p320tr	Mean O_3 (Nov)	Tract	2002-2021
p321tr	Mean O_3 (Dec)	Tract	2002-2021