



Questions & Answers Part 3

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Gaige Kerr (gaigekerr@gwu.edu) or Carl Malings (carl.a.malings@nasa.gov).

Question 1: This tool (EJScreen) is only available for the U.S.; will it be available for Europe, for example in the future? [Related Question: How does someone outside the United States make use of EJScreen, since most of its data are for the United States?]

Answer 1: Since the EPA is a United States government entity, the Agency largely focuses on environmental issues within the United States. We are not aware of any plans to expand EJScreen for areas outside of the United States. There are some other resources (e.g., [EJAtlas](#)) that might provide information on environmental hazards outside of the United States. If you are interested in understanding environmental and climate injustice outside the United States, you might be able to build your own EJScreen-like tool. Leveraging environmental and demographic data averaged to administrative units (such as [GADM](#)); calculating EJ indices following the formulae discussed in Part 3 of our training; and then visualizing in R, ArcGIS, or an analogous software would allow you to produce a similar type of map as EJScreen.

Question 2: If the EJ Screen does not include NO₂, does the screen add it or does the ArcGIS software include a model for it? We are assessing cancer risk in census blocks.

Answer 2: This was the subject of our EJScreen exercise today; you are free to continue using [the NO₂ dataset we presented](#) in EJScreen after the training today. If you would like access to the data at its native resolution of 1 km², see [this NASA resource](#) detailing the dataset and contact Gaige Kerr (gaigekerr@gwu.edu) with any questions.

Question 3: NASA info regarding air quality (concentration of NO₂ or PM_{2.5}) is available only for the USA? I am interested in Europe.

Answer 3: The specific dataset we are presenting in EJScreen today covers the USA only, since EJScreen only covers the US (see Question 1); However, in general, many satellite



remote sensing datasets and derived (Level 4) air quality datasets are available with worldwide coverage. Please refer back to Part 2 of this training for more details on those.

Question 4: Where can we get resources to locate this type of satellite data map service?

Answer 4: To find other data sources for EJScreen, see the answers to Questions 1 and 2 above.

Question 5: Regarding the NO₂ data and the population in Baltimore, I was wondering if the population data is updated (how often) or if it corresponds to census data?

Answer 5: Sociodemographic Data in EJScreen come from a variety of sources, [described here](#). Population data are derived from the US Census Bureau's American Community Survey. In the current (as of September 2023) version of EJScreen, estimates of the population gathered over the 5-year period spanning 2017-2021 are used. More information on these population estimates can be found [here](#).

Question 6: Do you recommend using EJScreen to look at rural communities or are there more data gaps?

Answer 6: EJScreen is available nationwide, so one can definitely examine rural areas using the screening tool. Note, though, that census block groups—the unit of analysis for EJScreen—in rural areas might be very large in area compared with urban areas, so this large area might smooth over some variations in concentration. For example, high pollution concentrations in one part of a census block group due to a local source might average out to lower concentrations throughout the entire census block group. People may also not be distributed evenly across these larger census block groups, but may live only in one part of it.

Question 7: Regarding the communities with a higher percentage of color in Baltimore, who are subjected to higher levels of air pollution: are they underrepresented or purposely uninvolved politically with the environmental consequences from local industry, or are they mainly living in these higher polluted areas due to financial reasons, since communities near heavy industry will likely be cheaper to live in?

Answer 7: Both of these ideas likely are at play. Communities of color have been often purposely excluded from the decision-making process or may lack the clout needed to ward off the siting of heavy industry. Additionally, if you explore income-related indicators in EJScreen's side-by-side maps, you'll see that the communities we focused on have lower



income (and coupling this information with data on housing prices from, for example, Zillow or Redfin shows that these communities have housing stock that is more inexpensive than elsewhere in the city).

Question 8: What are good resources to learn how to process level 2 data from TROPOMI/VIIRS?

Answer 8: Our previous ARSET trainings on [TROPOMI](#) and [VIIRS](#) are good places to start.

Question 9: Hi, one question, is there any 'fork' of EJScreen to another country (outside EU)? Is it possible to get some core open component of EJScreen (e.g. from GitHub) with given guidelines to import maps and needed datasets for the given country?

Answer 9: The trainers are not aware of a fork of EJScreen for regions outside the United States. However, if you look at Question 1, you can see ways to potentially build an EJScreen-like tool for your area of interest.

Question 10: I'm looking at the NO₂ data for Tucson, Arizona, where I live, and I notice that the Tucson Electric Power Sundt Generation Station doesn't appear as a NO₂ source, though I know it should. The baseload natural gas generators emit NO₂ and the ICE peaking generators definitely emit a spike of NO₂ when they come online. Could the issue be the temporal nature of the satellite data is missing these periodic emissions events?

Answer 10: Thank you for bringing this to our attention; it is entirely possible that certain sources are not represented in this NO₂ dataset, for a variety of reasons. As you note, the timing of satellite passes could be an issue, as well as the time period (this dataset represents 2019 averages). Stack height could also be a factor; although power plants are known emission sources, the height of the emissions could limit their impact on surface-level concentrations in the immediate vicinity, instead contributing to more diluted contributions over a wider area.

Question 11: Is there a repository for the air quality datasets that we can access to use with EJScreen?

Answer 11: If you are referring to the different datasets that are included in EJScreen, you can access them in CSV or geodatabase formats [here](#).



Question 12: What missions provide data with a high resolution of up to 1 square kilometer or more for air quality analysis (NO₂ extraction for example) in Europe?

Answer 12: Please refer back to Part 2 of the training for more information about different satellite missions and their spatial resolutions. There are satellite AOD data products covering Europe from the MODIS and VIIRS instruments. Unfortunately there are currently no missions monitoring NO₂ with such a resolution; this is due to necessary tradeoffs between the spectral resolution needed to retrieve NO₂ information and the spatial resolution. There are, however, high resolution derived (Level 4) NO₂ datasets for surface NO₂ which combine satellite and other data sources. For example, there are the [global annual-average datasets](#) created by the GWU research group.

Question 13: When I open the file with collab, I get this message (and a stack trace) "The file has been corrupted or is not a valid notebook file."

Answer 13: You may want to try re-downloading the file and putting the new one in your drive. Make sure that you extract the ".ipynb" type file from the ".zip" file that gets downloaded.

Question 14: If I'm already familiar with libraries like pandas, geopandas, pyplot, numpy, etc, can I bypass the part where I'd need to use Colab and use my computer instead? Or is the code computationally intensive?

Answer 14: Yes, you can run this code on your own computer; you would need to download the necessary input information from the shared ARSET Google Drive, and update the file paths appropriately for your own system. The code is not exceptionally computationally intensive. We typically use Google Colab for demonstration purposes, to avoid the need to customize codes for different operating systems, etc.

Question 15: If it is confirmed and scientifically verified that a certain underserved community is found to be a victim of environmental justice, what is the next step to resolve the issue?

Answer 15: This is a key question and a bit complex. Community involvement and support are a great starting point. There are also resources to report suspected environmental violations: <https://echo.epa.gov/report-environmental-violations>. Governmental resources, such as those available through the Inflation Reduction Act, might also allow community



members to obtain monitors and instrumentation to further measure and quantify the extent of environmental injustices.

Question 16: Do you have suggestions for how to look specifically at population weighted average for Native Americans? Or is this a limitation of the ACS data?

Answer 16: ACS *does* include [information](#) about the Native American/American Indian population. Since this population subgroup does not comprise a large portion of the population in Maryland, we have not included it in today's training for the sake of time. However, if you were to download the native ACS data, you would find information on Native Americans/American Indians and can examine environmental inequalities for this population subgroup.

Question 17: You need to pay close attention to spelling in folder naming as 'Exercise' is spelled 'Excercise'.

Answer 17: Thank you for pointing that out! Unfortunately that was a mistake; the code assumes that the folder name has the incorrect spelling "Excercise".

Question 18: I have arset_sd4ej_p3_python_excercise as a subfolder to that or having the Excercise_Materials as a subfolder to the arset_sd4ej_p3_python_excercise folder and both are giving me error messages in the first step. It should be /content/drive/MyDrive/SD4EJ_Python_Excercise/SD4EJ_Python_Excercise_Materials right? (It finds SD4EJ_Python_Excercise but nothing below that)

Answer 18: Yes, that's the file path assumed by the Python code. If you are familiar with Python, you can modify the file paths (specified in the first code section of the notebook) to match how your Drive is configured.

Question 19: Are the NO₂ data for the Maryland analyses from ground-based monitoring?

Answer 19: The dataset is generated as a derived or Level 4 data product, which combines ground-based data, satellite data, and land use information. Specifically, NO₂ concentrations from surface level monitors were used to develop a regression based on multiple variables (e.g., population, proximity to point sources, proximity to roadways, etc.) that was applied globally to estimate NO₂ concentrations in regions without monitors.

More details about this dataset were provided in Dr. Susan Anenberg's presentation in Part 1 of this training.



Question 20: Are there plans on conducting satellite data validation measurements by research institutions or other agencies?

Answer 20: Yes, in general, this is commonly done whenever a mission is launched within NASA and other agencies; the data from the new mission is validated and inter-compared with any similar or overlapping data from other missions to ensure that we understand how these different instruments relate to each other, whether they are performing as expected, and how we can create continuity in our long-term datasets using data from multiple successive missions.

Question 21: Is it true that the quality of the air can be tracked only by the concentration of NO₂?

Answer 21: NO₂ is a tracer or proxy for fossil fuel combustion, so looking at NO₂ concentrations or satellite-derived levels can be a powerful, stand-alone way to get a sense for industry, traffic, and other sources (especially in urban areas). However, there are other air pollutants (such as particulate matter and health-relevant air toxics like formaldehyde and benzene) that might have different patterns of spatial variability than NO₂, so NO₂ might not be a “catch all” for all forms of pollution.

Question 22: Is there a way to download datasets from EJScreen for my region and get NO₂ data for other states similar to what was provided for Maryland?

Answer 22: Yes, please see question 11 for how you can download the data used in EJScreen. The data are national-level, so you might need to subset the data to your particular state, region, and/or city.

The NO₂ data (in gridded format, not averaged to underlying census block groups) is also available as a public service [here](#).

Question 23: Are there pitfalls or common assumptions about analyzing environmental justice that people should be aware of? For example, how safe is it to assume a nearby emissions source is the cause of pollution in a neighborhood, and what might lessen uncertainty?

Answer 23: Attributing air pollution to a particular nearby source likely would require model simulations to isolate the influence of this particular source. In terms of other pitfalls,



Satellite Data for Air Quality Environmental Justice and Equity Applications August 23, 30, and September 6, 2023

it is also important to uphold environmental justice frameworks to ensure that research does not further stigmatize marginalized communities. For more information, see this [paper](#).

Question 24: What format does the satellite data have to be in in order to load to EJSCREEN or Python?

Answer 24: Python is pretty flexible with the correct libraries and can accept various formats (geodatabase, Excel files, .csv files, NetCDF files, etc.). For EJScreen, you will need to use the ArcGIS REST API. Creating such a database may require a GIS license and knowledge of the software.