# EPA's Response to Peer Review of EPA's Power Plant Environmental Justice Screening Methodology

March, 2023

## 1. Introduction

This document responds to comments developed through an external peer review of EPA's Power Plant Environmental Justice Screening Methodology (PPSM). While there are several potential applications of the PPSM and the associated components, the primary objective of this peer review was for an independent group of subject matter experts to review and evaluate the scientific foundations of EPA's PPSM screening-level tool that can be used to rank fossil steam electric generating units (EGUs) in the contiguous U.S. based on their potential to affect more areas of greater potential environmental justice (EJ) concerns.

This peer review focused on the suitability of the PPSM for the purpose of applying various modeling techniques to inform the identification of census block groups potentially exposed to air pollution by power plants, the identification of block groups with potential EJ concern, and the combination of those components to quantify scores and enable ranking of facilities by the potential to affect more areas of greater potential EJ concern.

Given the scope and intended purpose of the PPSM, the charge to the peer reviewers was to:

1) Evaluate the suitability of the PPSM for the intended use, and the reasonableness of the underlying assumptions
2) Identify specific strengths, weaknesses, limitations, and any errors in the formulation, assumptions, and conclusions
3) Propose specific options for correcting errors and fixing or mitigating weaknesses and limitations in the methodology formulation, assumptions, outputs, or conclusions derived

The peer reviewers agreed that PPSM is a useful tool and were supportive of the methods used in the screening tool to understand the potential EJ implications of power plant emissions. They concurred that by ranking EGUs based on their potential to affect more areas of greater potential EJ concern, PPSM fills a void by focusing on specific communities that experience disproportional exposure to air pollution.

The reviewers also identified several areas of improvement. The major comments are summarized below and include EPA responses. The Peer Review Summary Report contains a more detailed discussion of the comments made by each reviewer.

# 2. Addressing the Reviewers' Primary Comments

## 2.1. Trajectory vs. concentration modeling

The reviewers suggested a number of potential refinements to the long-range modeling component of the PPSM. In particular, they questioned whether the "concentration" simulation mode within HYSPLIT would be a more suitable option than the "trajectory" simulation mode that was initially used by EPA. In response, EPA explored this option, and consulted with experts including Dr. Mark Cohen at NOAA.

The two modes track the trajectories of massless parcels of "air" particles throughout the simulation in different ways: the trajectory mode follows a single particle over a period of time, while the concentration mode follows thousands of particles over the same period. So within a single time period the results are therefore quite different – the trajectory mode results in a single path (which could be interpreted as a centerline of a plume), whereas the concentration mode captures many slightly different trajectory paths of thousands of particles and results in a plume.

While the results of these two approaches are different over a single time period, we would expect the results to converge over a large number of simulations. With multiple iterations, each approach will result in a large number of individual particle trajectories (i.e., thousands from the trajectory mode and millions from the concentration mode) that intersect with a comparable set of block groups. However, because the number of particles in the concentration mode is many orders of magnitude larger than the trajectory mode, the concentration mode should be substantially more spatially-resolved toward the edges of the plume.

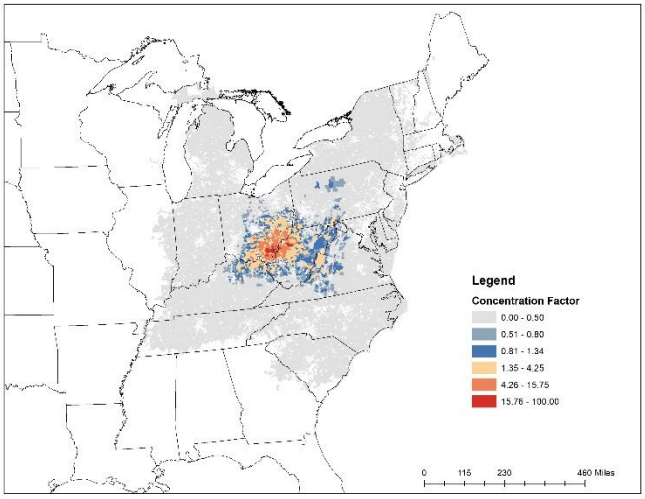
We were also able to qualitatively and quantitatively compare the dispersion patterns between the two methods. For this illustrative analysis, we simulated emissions from a single stack of a facility located in a central location of the Ohio River Valley.[1] We simulated emissions from this stack for three months in 2018 (June 1 to August 31) using the same meteorological data sets used for the trajectory simulations in the analysis shared with the peer reviewers. We created simulations of a 1-hr release which we followed over a 24-hour time-period. We conducted 4 simulated concentration releases per modeling day using 2,500 particles per release with a concentration grid resolution of 0.05 degrees. Figure 1 shows the resulting relative, or "normalized", concentrations (normalized using the concentration at the most-impacted block group) for the concentration mode compared with two versions of the trajectory mode (one focused just on the year 2018 and the other focused on all trajectory simulations from 2018 through 2020).[2,3]

[1] ORIS 8102; 38.9347, -82.1158 latitude and longitude; 253 m stack height

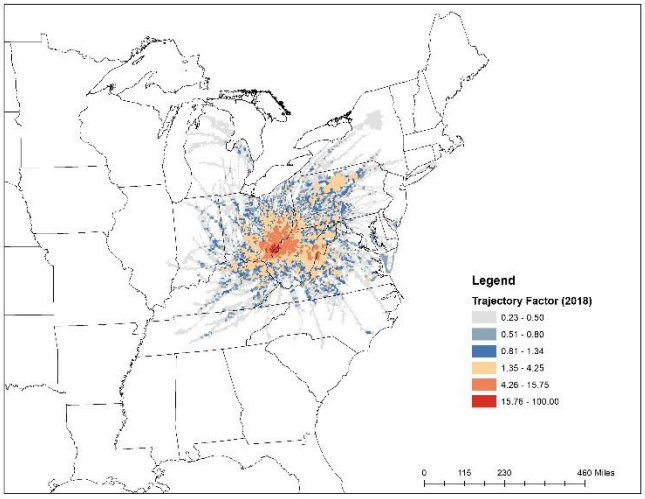
[2] The normalized concentrations were multiplied by 100, so that the block group with the maximum concentration has a value of 100 (rather than a value of 1).

[3] In the concentration simulations, in associating 0.05 degree grid cell concentrations with block groups, we identified the centroid of the grid cell and associated that point with a single block group. Consequently, for grid cells that contain several block groups (I.e., urban areas), only a single block group was identified. This was only done to simplify this comparison (and to ensure that we got representative geospatial representation from both the trajectory and concentration simulations). In future applications, we intend to identify all block groups within the grid cells.

(a)




(b)




(c)

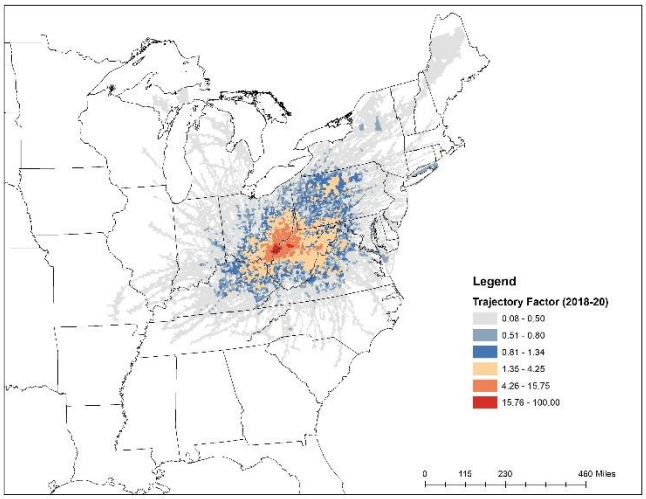



Figure 1. Normalized concentrations for (a) concentration simulations for 2018, (b) trajectory simulations for 2018, and (c) trajectory simulations for 2018-2020.

Qualitatively, the results suggest that the concentration approach identifies similar high- or frequent-impact areas but has substantially more spatially-resolved impacts at farther distances where concentrations have dropped substantially.

In this analysis, we matched the centroid of each simulated grid cell to the census block group containing that point. We then calculated a normalized[4] concentration level for each block group, relative to all the other block groups that have concentration values, with values ranging from 0 to 100. We then binned the normalized concentration levels equally into 20 groups.[5] Next, we assessed the trajectory data set to create concentration values that we could directly compare against the concentration data set. To do this we use the 24-hour trajectory simulations for 2018 to identify the frequency with which trajectories intersect each block group, which is then normalized using a scale of 0-100. This approximates concentrations using the trajectory method allowing us to directly compare the two methods. We compared this normalized concentration information for all block groups found in both concentration and trajectory modeling for the illustrative facility. We compared the normalized concentrations between the two approaches using least-squares linear regression where information for block groups was available in both approaches (Figure 2). We found that the results are highly correlated (slope of 0.93 and a regression coefficient of 0.85).

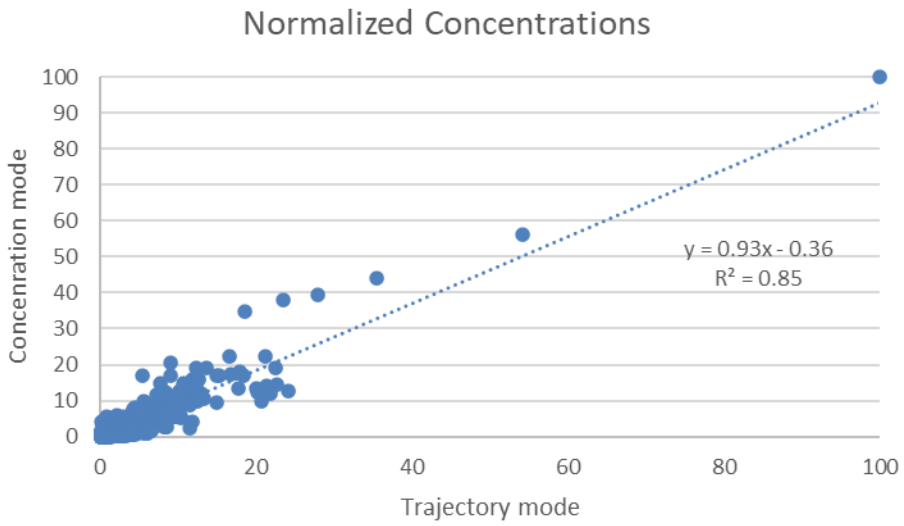



Figure 2. Least-squares linear regression comparison of the normalized concentration at individual block groups using the concentration and trajectory modes.

Next, for the block groups that were identified using both methods, we examined how many block groups were associated with each normalized concentration. We binned the normalized concentrations in 5-unit increments and then counted the number of block groups that are in each bin (Figure 3). A

[4] The concentrations were normalized by dividing the concentration in each block group by the maximum value found across all block groups.

[5] For example, all block groups with a normalized concentration value of 5-10 were placed into a single bin and identified with a normalized concentration of 5.

similar number of block groups with particular concentrations was identified for both the concentration and trajectory modes.

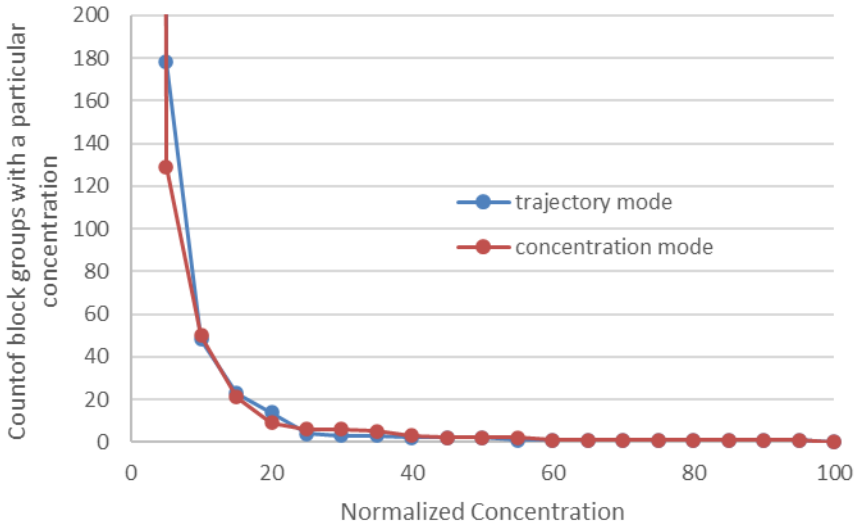



Figure 3. A count of the number of block groups with particular normalized concentrations for the concentration and trajectory modes. Similar patterns are found for both modes.

Based on this analysis, EPA concludes that both the trajectory and concentration method will likely produce comparable and reasonable results when examined over longer time periods. The strong correlation between the normalized intersection frequency and concentration values suggests that we would expect similar results in the HYSPLIT-based PPSM scoring approaches, which effectively weight block groups based on the normalized frequency of intersections discussed above. Therefore, we would not expect the long-range downwind scores developed for this review using the trajectory mode to be significantly different than long-range downwind scores developed using concentration mode.

We also note that running the model in concentration mode offers additional value in that it enables a more-quantitative comparison of potential exposures at various distances downwind. Additionally, the concentration mode appears to provide a more comprehensive accounting of all the downwind areas, particularly the areas farther from each source where potential exposures are relatively small. While this would not have a significant impact on the overall scores reviewed in this version, we expect it would provide a more complete picture of the areas potentially exposed to air pollution from each source. EPA therefore intends to complete a series of HYSPLIT concentration simulations that could be incorporated into the next version of the PPSM.

Additionally, EPA notes that running the model in concentration mode would also respond to reviewer comments regarding the potential use of an exponential decay function to consider the relative exposures in the long-range analysis on a distance or travel-time basis. While EPA's application of trajectory modeling in a frequency of impact approach approximates the expected decay (see for example Figure 3), the concentration modeling would improve this representation. As we note above, we would expect similar results across the different approaches (see comparison of normalized concentrations in Figure 3). We also note that the concentration modeling would not take any chemistry or deposition into consideration, as these aspects are pollutant-specific. We believe that either HYSPLIT-

based approach is a reasonable approximation of potential exposures, and that attempting to incorporate any additional pollutant-specific precision would limit the broad applicability of this screening-level methodology. However, we believe that future users of the PPSM could potentially layer in pollutant-specific information to analyze more specific questions not considered in this initial draft.

## 2.2. Adequacy of 24-hour simulations

In order to determine whether 24-hour trajectories are sufficient to capture long-range pollution patterns, the reviewers suggested that EPA conduct a sensitivity study to examine the effects of simulating different time-periods in order to observe how concentration might decrease as a function of distance from the facility. EPA conducted this study by comparing the peer-reviewed 24-hour trajectory simulations for a particular facility to the 24-hour concentration mode simulations[6] for that facility.

In this analysis, as described in section 2.1 of this document, we matched the centroid of the simulated grid cell to the census block group containing that point. We then, calculated a normalized concentration level for each block group, relative to all the other block groups that have concentration values, with values ranging from 0 to 100. We then binned the normalized concentration levels equally into 20 groups.[7] We also used the 24-hour trajectory simulations for 2018 to identify the average and median length of time each trajectory takes to arrive at each downwind block group. We combined this information for all block groups which intersect with both concentration and trajectory modeling for the illustrative facility.

Figure 4 depicts the relationship between normalized concentration and trajectory hour. The results are consistent with what we see in Figure 1 – generally, the concentrations decreased as a function of distance and time from the facility. The highest concentrations were found in the immediate vicinity of the facility, and the travel times associated with these concentrations is very short (within the first few hours following release). Focusing on some of the smaller concentrations (i.e., the normalized concentration bin of 5-10), the average and median travel times were substantial (around 12 hours, with a standard deviation of about 7 hours). However, for this example facility, for this particular time-period, we did observe some higher normalized concentrations associated with longer travel times (i.e., for the normalized concentration bin of 35-40, the average and median travel times were substantial, around 8 hours of travel time following release). This likely correlates to some of the red areas (shown in Figure 1a) that are around 100 miles downwind of this particular facility. While this is just a single example, potential high downwind exposures that are associated with larger distances (e.g., ~100 mi) and longer travel times (e.g., ~8 hours) illustrate why HYSPLIT long-range simulations may be important to include in this screening methodology.

Overall, we conclude that the current 24-hour simulations likely encompass nearly all of the block group areas that receive the largest impacts, and many (but likely not all) block groups that receive smaller impacts. For the illustrative facility examined in this analysis, the concentrations drop substantially by about 12 hours on average, but that there is substantial temporal variation. For this facility, increased concentrations were observed around 8 hours following release, demonstrating the importance of these

[6] As discussed above, HYSPLIT concentration mode simulations for 2018 from ORIS 8102 (38.9347, -82.1158 latitude and longitude, 253m stack height).

[7] For example, all block groups with a normalized concentration value of 5-10 were placed into a single bin and identified with a normalized concentration of 5.

longer-range simulations. While the normalized concentrations are generally lower at 24 hours, we believe it is important for this EJ screening analysis to be able to consider both higher and lower potentials for exposure. We conclude that a 24-hour simulation represents a reasonable time-period for tthis screening analysis because it identifies the areas that experience the highest impacts and ensures that broad areas that experience lesser impacts are also included. However, outside of this screening analysis, it is important to recognize that pollution from a given release will likely impact downwind air quality beyond the 24-hour simulation period and at longer distances.

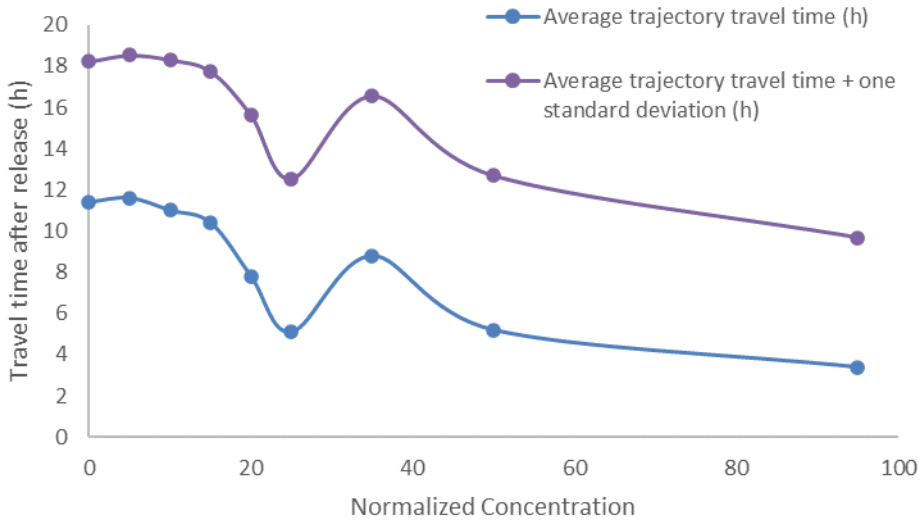



Figure 4. "Plume" travel time following release as a function of the normalized concentration downwind. Generally, the higher the downwind concentration, the shorter the travel time.

## 2.3. Identification of Distances for Proximity Analysis

The reviewers recommended that EPA expand on the discussion of the methods, assumptions, and percentiles chosen to represent different concentrations based on AERMOD dispersion modeling. For example, they recommended explaining the rationale behind defining the "maximum concentration radius" as the average distance to the ten highest concentrations at a facility. Additionally, one reviewer raised questions about why block group receptors are only modeled out to 10 km.

To better understand the assumptions and percentiles chosen in the current proximity analysis it is important to highlight that the goal of this screening-level analysis is to identify census block groups potentially exposed to air pollution by power plants. The "maximum concentration radius" is defined as the average distance to the ten highest concentrations modeled for each source. This definition was selected based on an analysis of the 99$^{th}$ percentile of the concentration distribution at each EGU. The distribution above the 99$^{th}$ percentile at all coal and gas fired EGUs considered in the PPSM contained at least ten concentration values; however, some distributions contained up to 213 values due to a larger sample size (i.e., number of receptors modeled). The number of receptors modeled for each facility was dependent on the spacing of the "gridded" receptors and the density of census-block receptors. To normalize the number of distance values that were averaged in the "maximum concentration radius", the minimum 99$^{th}$ percentile sample size of ten was selected. EPA believes averaging the distances to

the ten highest concentrations accounts for not only the absolute maximum concentration but also additional areas of elevated concentration that may be present in the modeling domain. However, in future updates to the PPSM, EPA may reconsider the definition of the "maximum concentration radius."

Although the area of maximum concentration is of importance in proximity screening, it is not inclusive of all communities that may be impacted by a source. To allow users to screen for additional communities beyond the area encompassed using the maximum concentration radius, the 25$^{th}$, 50$^{th}$, and 75$^{th}$ percentile of the concentration distribution were calculated to provide additional screening distances for each facility. The analysis shared with the peer reviewers used the 50$^{th}$ percentile distance at each facility to represent the "intermediate concentration distance" for purposes of proximity analysis to screen for census block groups that may be exposed to concentration impacts from each EGU. For this analysis, EPA selected the 50$^{th}$ percentile understanding that the radius would capture block groups in areas where concentrations are more than half the maximum concentration. The intermediate proximity analysis can be extended to any percentile within the distribution to gain a better understanding of the relative potential for the facilities included in the analysis to expose areas based on higher or lower concentration levels at different distances, and EPA intends to provide users with the information necessary to do so.

To address the reviewer's question about the receptor strategy, the AERMOD modeling used for PPSM was initially performed for EPA's 2017 Air Toxics Screening Assessment. The modeling used a receptor strategy that included equally spaced "gridded" receptors (1 km in highly populated areas, 4 km otherwise), populated census-block centroid receptors, and monitoring site receptors. The dispersion modeling for point sources explicitly modeled gridded receptors within 50 km and populated census-block centroid receptors within 10 km of any emission point at the facility. Census-block receptors were only modeled within 10 km of the facility to remain consistent with the methodology used in the 2014 National Air Toxics Assessment and to reduce model runtime. A reasonable project timeline was necessary to model all point sources that emitted a hazardous air pollutant in the contiguous United States. For more information on the receptor strategy used in this modeling please refer to the 2017 AirToxScreen Technical Support Document.

The discontinuity mentioned by the reviewers in Figure 2 of the documentation is a result of an increased number of census-block centroid receptors within 10 km of the EGU. Due to the varying size and number of census blocks in any given area, the example distribution provided had an increased number of smaller census blocks just within 10 km of the facility. This discontinuity is not visibly present in all concentration distributions due to its dependency on the number of populated census-blocks within 10 km of the EGU.

## 2.4. Use of RSEI-based Indicators

The reviewers proposed that EPA integrate the Risk-Screening Environmental Indicators (RSEI) model in PPSM, suggesting the replacement of the two EJScreen indicators, cancer risk and respiratory hazard index, which are based on National Air National Air Toxic Assessment (NATA), with information from RSEI for the calculation of the cumulative impact metric.

For the purpose of estimating a cumulative impact metric for use in this PPSM screening analysis, at this time, EPA currently believes that the NATA-based indicators are preferable for two reasons. First, the NATA-based indicators are based on a better source and receptor characterization, using detailed

emissions points for each source and census blocks as receptors. Additionally, the NATA-based indicators consider all known sources of air toxics, including mobile sources, biogenic sources, and wildfires. While there are advantages and disadvantages of RSEI- and NATA-based indicators, at this time we believe that the NATA-based indicators are best suited to inform the cumulative impact metric for the purpose of this screening methodology.

## 2.5. Identifying block groups located partially within proximity analysis buffer

For areas that are only partially within the radius in a proximity analysis, the reviewers suggested that EPA (1) determine the share of the area of each census block polygons that is partially within the radius, and (2) use census block-level population data to weight these partial block groups.

In the current methodology, we are calculating the area share of each block group polygon that is within the radius and using that share to scale the total population of the block group. In future versions of the PPSM, we plan to apply this methodology to block-level population data to better weight the areas that are partially included within a radius.

## 2.6. Application of EJScreen Indicators in PPSM to Identify Block Groups with Potential EJ Concerns (Cumulative Impact Perspective)

The reviewers raised questions regarding the use of several EJScreen indicators in the calculation of the cumulative impact screening metric, including:

- relative weighting of various environmental burden indicators
- potential overlap and double-counting of some environmental burden indicators
- potential overlap and double-counting of some socioeconomic indicators
- the use of unemployment rate to represent susceptibility to air pollution
- the use of age-based indicators
- the use of a single income-based indicator rather than wealth-based indicators, and
- the total count of environmental burden indicators relative to the total count of socioeconomic indicators.

It is important to recognize that quantification of both cumulative impacts and vulnerability is an area of active research, and this initial PPSM methodology should be viewed as an initial screening-level effort to quantitatively evaluate these concepts nationally at a census block group resolution. Additionally, we believe the current methodology utilizes the best national environmental burden and demographic data that is currently available at this resolution. We intend to consider these comments carefully as we incorporate future advances in both the available indicators as well as our understanding of how those indicators can be combined to capture cumulative impacts and vulnerability.

While we believe that the current cumulative impact metric used in this version of the PPSM is a reasonable starting point for screening-level analysis, we also believe that future updates to the methodology could likely be improved by incorporating some, but not all, of these suggestions. In particular, while we considered it carefully, we respectfully disagree with the commenters' suggestion to include age-based indicators. We observe that the "percent over 64" indicator negatively correlates with the other demographic and environmental burden indicators, suggesting that this indicator identifies areas which are relatively well off from both a socioeconomic and environmental perspective.

We also note that it is possible for the user of this work to reassemble the constituent components of the PPSM in various ways to answer different questions. For example, a user could use the long-range downwind or proximity analysis information to explore alternative perspectives of potential EJ concern in the block groups potentially exposed to air pollution by each power plant in this analysis. We intend to explain this further in the updated documentation.

One of the reviewers also questioned the use of percentiles in calculating the cumulative impact and vulnerability metrics, noting that it would not account for extreme outliers. Another reviewer responded that the use of percentiles is reasonable, since the PPSM is not an effort to quantify specific impacts. The EPA agrees with the second reviewer here, and notes that the PPSM is a qualitative effort to identify relative potentials. Since the PPSM is not attempting to quantify specific impacts, we believe it is appropriate to continue using percentile-based indicators. Additionally, per the reviewer's recommendation, we will explain this more clearly in the updated documentation.

### 2.7. Comments regarding the two scoring approaches

The reviewers expressed different opinions on the potential policy value of the two different scoring approaches. The discussion highlighted the fact that each of these two approaches captures different perspectives, each of which can be relevant and important for considering the relative potential for power plants to affect areas with possible EJ concerns. In particular, the different approaches help the user consider both densely populated and rural areas. These comments help confirm EPA's belief that both scoring approaches are important to include in the PPSM.

### 2.8. Documentation, transparency, and clarity

The reviewers offered several comments regarding the documentation describing PPSM, including:

- the importance of clarifying the difference between impact and exposure,
- increasing the transparency of assumptions used, and
- providing more discussion on potential applications of the various PPSM components.

The EPA agrees with these comments and plans to clarify and/or expand on all of these topics in the documentation of the next version of the PPSM.

## 3. Conclusion

EPA is deeply grateful to the reviewers for their time and generosity in evaluating the scientific foundations of EPA's PPSM screening-level tool, as well as their suggestions for strengthening the methodology and documentation. This input will be invaluable to ongoing development of PPSM and an expanded focus on applications. In that context, we are currently working on an updated version of this methodology and documentation that will likely include the updates discussed by EPA above. While we believe that any future updates to this methodology will improve the underlying components of the PPSM, we conclude that the current version is a reasonable and useful methodology for providing a screening-level look at the relative potential for power plants to affect areas with possible EJ concerns, utilizing the best environmental burden and demographic data that is currently available nationally at this resolution. In updating this methodology, we intend to consider the reviewer's comments carefully

as we incorporate future advances in both the available indicators as well as our understanding of how those indicators can be combined to capture cumulative impacts and vulnerability.