September 30, 2016

Submitted via email to SILguidance@epa.gov

RE: Draft EPA Guidance on Significant Impact levels for Ozone and Fine Particles in the Prevention of Significant Deterioration Permitting Program (Revised August 18, 2016)

The American Petroleum Institute (API) submits the attached comments on the Environmental Protection Agency’s (EPA) August 18, 2016 draft Guidance on Significant Impact levels for Ozone and Fine Particles in the Prevention of Significant Deterioration Permitting Program and associated supporting documents.

API represents over 650 oil and natural gas companies, leaders of a technology-driven industry that supplies most of America’s energy, supports more than 9.8 million jobs and 8 percent of the U.S. economy, and, since 2000, has invested nearly $2 trillion in U.S. capital projects to advance all forms of energy, including alternatives. Efficient implementation of the Prevention of Significant Deterioration (PSD) permitting program is important to our members’ ability to construct or modify facilities to meet our nation’s energy needs.

We support EPA’s establishment of Significant Impact Levels (SILs), which provide a balance between protection of air quality and economic development as called for by the PSD program and legislative history. SILs reduce the burden on both permitting agencies and applicants by allowing new or modified sources with a de minimis impact on air quality to avoid costly modeling analyses where it “would only yield information of trivial or no value with respect to the required evaluation of the proposed source or modification.”

In its draft guidance, EPA utilizes a new approach based on the variability of ambient air monitor data to set ozone and PM$_{2.5}$ SILs. After review of this new method, we believe it is more appropriate for EPA to set the ozone and PM$_{2.5}$ SILs using its established precedent of 4%.

---

1 Memorandum from Anna Marie Wood, Acting Director, Air Quality Policy Division, to Regional Air Division Directors, p.11 (June 28, 2010) (guidance on implementing PSD for the 1-hour NO$_2$ NAAQS).
of the NAAQS or PSD increment; the current approach is technically reasonable and protective of air quality.

EPA’s draft guidance for ozone and PM$_{2.5}$ SILs is intended for use in conjunction with other EPA rules and guidance documents, including a revised Modeling Guideline (40 CFR 51 Appendix W) and new Model Emission Rates for Precursors (MERP) guidance. Since these documents are not currently available, we reserve the right to amend our comments once EPA releases its draft MERP guidance and final Appendix W rule.

Thank you for your consideration of these comments. If you have any questions, please contact me at kaliszcz@api.org or at (202) 682-8318.

Sincerely,

Cathe Kalisz

Attachment
American Petroleum Institute (API)

Comments on

“Guidance on Significant Impact Levels for Ozone and Fine Particles in the Prevention of Significant Deterioration Permitting Program” as proposed on August 1, 2016 and revised August 18, 2016.

September 30, 2016
Comments on EPA’s Proposed Ozone and PM$_{2.5}$ SIL Guidance

Table of Contents

1.0 EXECUTIVE SUMMARY AND OVERARCHING COMMENTS .......................................................... 1

1.1 We support EPA’s proposed approach to continue to implement Significant Impact Levels for the New Source Review permitting program. .................................................................................. 1

1.2 We do not support EPA’s approach to tie the magnitude of the SILs to the level of uncertainty in the monitored “design concentrations” .............................................................................................................. 2

1.3 EPA’s bootstrapping approach to estimate the concentration levels for the uncertainty of the monitored design concentrations uses an inappropriate confidence level to establish the SIL level. .... 2

1.4 It is unclear whether EPA’s bootstrapping approach considers certain sources of ozone and PM$_{2.5}$ monitoring uncertainty, such as the instrument accuracy limitations and the expected and measured changes of ozone concentrations within the range of EPA’s probe height guidance. ............. 2

1.5 EPA’s proposed SIL levels are too low and overly conservative. We recommend that EPA establish SILs based on 4% of the NAAQS and PSD increments ............................................................................. 3

2.0 EPA’S PROPOSED USE OF SILS ......................................................................................... 4

2.1 SILs have been used in the New Source Review program for several decades and in many cases, the SIL level has been at or close to 4% of the applicable NAAQS or PSD increment ................. 4

2.2 SILs streamline the permitting approach in a way that appropriately addresses small sources.... 5

3.0 EPA APPROACH TO SET THE MAGNITUDE OF SILS ...................................................... 6

3.1 Up to this point in time, the SILs have generally been set as a small percentage of the applicable NAAQS or PSD increment ......................................................................................................................... 6

3.2 The form of the SIL (highest concentration) is conservative compared to the form of the ozone and the 24-hour PM$_{2.5}$ NAAQS and PSD increment ................................................................. 6

3.3 EPA’s newly proposed technical approach that ties the magnitude of the SILs to uncertainties in the monitored design concentration has limitations that render it less supportable than the current approach using 4% of the NAAQS or PSD increment ................................................................................. 7

4.0 EPA’S USE OF BOOTSTRAPPING TO SET THE MAGNITUDE OF THE SILS .................. 7

4.1 EPA’s bootstrapping analysis is a work in progress that needs updating before it can be considered as a final product for EPA policy decisions ............................................................................. 7

4.2 The confidence interval (50%) that EPA uses to establish the SILs is contrary to normal statistical practice (95%) ......................................................................................................................... 7

4.3 EPA’s bootstrapping analysis using years of monitoring data addresses only one aspect of the design concentration uncertainty ........................................................................................................... 8

5.0 EPA SHOULD USE 4% OF THE NAAQS OR PSD INCREMENT TO ESTABLISH SILs .......... 10

5.1 EPA’s bootstrapping approach uses a confidence interval that is too low .................................. 10
5.2 Other sources of error or bias independently result in high DV uncertainties: instrument measurement errors and concentration gradients as a function of height. .................................................. 10

5.3 Based upon our recommendation that EPA continue to use 4% of the NAAQS for the SIL, we provide the resulting SIL values for ozone and PM$_{2.5}$ that EPA should adopt. .................................................. 10

APPENDIX A: REVIEW OF EPA’S TECHNICAL BASIS FOR THE PM$_{2.5}$ AND OZONE SILS
API Comments on EPA’s August 18, 2016 Proposed “Guidance on Significant Impact Levels for Ozone and Fine Particles in the Prevention of Significant Deterioration Permitting Program”

The American Petroleum Institute (API) provides the following comments on EPA’s proposed guidance for ozone and \( \text{PM}_{2.5} \) Significant Impact Levels (SILs), \(^1\) as released for public review on August 1, 2016, and revised on August 18, 2016. The guidance consists of a “Legal Support Memorandum”\(^2\) supporting the SIL concept, a guidance document\(^3\) that lists the SILs and how they are to be used, and a technical basis document\(^4\) in support of the values of the SILs. Our comments address all of these documents.

OVERVIEW

1.0 EXECUTIVE SUMMARY AND OVERARCHING COMMENTS

1.1 We support EPA’s proposed approach to continue to implement Significant Impact Levels for the New Source Review permitting program.

As stated in EPA’s Legal Support Memorandum for the use of SILs, the goals of the Prevention of Significant Deterioration (PSD) program can be addressed by demonstrating that increased emissions from proposed sources do not “cause or contribute to” a violation of the NAAQS or PSD increment because of their minimal impact on existing air quality. The D. C. Circuit has noted that Congress did not specify in the Clean Air Act’s PSD provisions “how the owner or operator of a proposed source or modification must demonstrate” the project will not cause or contribute to a violation. See Sierra Club v. EPA, 705 F.3d 458, 465 (D.C. Cir. 2013). EPA has filled this statutory gap by specifying how such a demonstration should be made.

As noted in the Legal Support Memorandum, the PSD program is intended to prevent “significant” deterioration in air quality. The legislative history supports a balancing of air quality and economic growth; the PSD program is to “permit the economic development necessary to achieve a steady improvement in our standard of living.” S. Rep. No. 95-127, at 29, reprinted in 3 Environmental Policy Division, Congressional Research Service, A Legislative History of the Clean Air Act Amendments of 1977, at 1403 (1979). In establishing SILs, EPA has acted reasonably to provide the balance between protection of air quality and economic development called for by the statutory

---

\(^1\) Available at https://www.epa.gov/nsr/forms/significant-impact-levels-ozone-and-fine-particles-prevention-significant-deterioration.


purposes and legislative history. In adopting SILs, the Agency has minimized the potential for unnecessary permitting burdens by eliminating a cumulative impact analysis where it “would only yield information of trivial or no value with respect to the required evaluation of the proposed source or modification.” Memorandum from Anna Marie Wood, Acting Director, Air Quality Policy Division, to Regional Air Division Directors, p.11 (June 28, 2010) (guidance on implementing PSD for the 1-hour NO₂ NAAQS).

1.2 We do not support EPA’s approach to tie the magnitude of the SILs to the level of uncertainty in the monitored “design concentrations”.

In the past, the SILs for various pollutants have been associated with a small fraction (e.g., 4%) of the applicable NAAQS or PSD increment. This is a small percentage that has been justified as associated with a “cumulative” factor, such that if the full SIL were taken by the proposed project, it would take a large number (25) of such impacts to sum to the full NAAQS or PSD increment. EPA considers its new technical method as “improved” because it is stated as being based upon the uncertainty in the reported design concentration. However, we note in our detailed comments that the underlying assumptions and selection of an inappropriate confidence interval for this new technical approach result in an underestimate of the uncertainty in the reported design concentration. The procedure has overlooked several issues that would tend to increase the uncertainty. Therefore, we recommend that EPA retain the traditional use of 4% of the NAAQS or PSD increment for specifying the SIL value.

1.3 EPA’s bootstrapping approach to estimate the concentration levels for the uncertainty of the monitored design concentrations uses an inappropriate confidence level to establish the SIL level.

EPA’s bootstrapping approach arbitrarily uses the 50% percentile confidence interval, which is inconsistent with the more universally-used 95% confidence interval. If this non-standard confidence interval is used, the design value uncertainty is significantly underestimated, resulting in a smaller and more conservative SIL. Establishing SILs using the traditional 4% of the NAAQS or PSD increment will provide for slightly larger SILs which will not be overly conservative and will still be protective of the NAAQS.

1.4 It is unclear whether EPA’s bootstrapping approach considers certain sources of ozone and PM₂.₅ monitoring uncertainty, such as the instrument accuracy limitations and the expected and measured changes of ozone concentrations within the range of EPA’s probe height guidance.

EPA has not clearly documented how they consider other sources of uncertainty in the reported design concentrations. These include instrument biases that are inherently present and acknowledged by EPA for both ozone and PM₂.₅, as well as substantial vertical gradients of ozone that occur within the large 2-15 meter interval allowed by EPA for ozone monitor probe placement.

---

5 Design concentrations are associated with the rank and averaging time corresponding to the ozone and PM₂.₅ National Ambient Air Quality Standards (NAAQS) and the PM₂.₅ Prevention of Significant Deterioration (PSD) increments. For example, the design concentration for the ozone NAAQS is the average of three years’ 4th highest daily maximum 8-hour average concentration.
1.5 EPA’s proposed SIL levels are too low and overly conservative. We recommend that EPA establish SILs based on 4% of the NAAQS and PSD increments.

Based upon additional sources of monitor value uncertainty not considered by EPA as well as EPA’s choice of a too-small confidence interval (50% rather than 95%), we conclude that EPA’s proposed SIL levels are too low and overly conservative.

After considering EPA’s uncertainty analysis and other sources of uncertainty that EPA did not consider, we recommend the use of 4% of the NAAQS and PSD increments for the Class II SILs. EPA’s approach to extend the Class II SILs to Class I areas by multiplying the Class II SILs by the ratio of the Class I to Class II increments is reasonable.
SPECIFIC COMMENTS ON EPA’S PROPOSED SIL GUIDANCE

2.0 EPA’S PROPOSED USE OF SILS

2.1 SILs have been used in the New Source Review program for several decades and in many cases, the SIL level has been at or close to 4% of the applicable NAAQS or PSD increment.

The use of SILs is not a new concept, but rather a continuation of an EPA policy that has been in place since 1978. Initially, with the 1978 final rule implementing the 1977 Clean Air Act Amendments, the SILs for SO₂, NO₂, particulate matter, and CO were set to levels at or below the PSD Class I increments. The rule stated that, “these levels shown below are therefore interpreted by the Administrator as representing the minimum amount of ambient impact that is significant.” In the update to the 1977 Clean Air Act Amendments as a result of the Alabama Power vs. Costle court decision, EPA published in the August 7, 1980 Federal Register (45 FR 52707) a discussion about SILs that established a 4%-of-the-NAAQS “rule of thumb” for an insignificant impact because “several sources” in a given area could consume that much of the available concentrations within the NAAQS and not threaten attainment.

Other EPA guidance memoranda over the years refined and extended the use of the SIL. In 1988, the use of the SIL with respect to modeled attainment or nonattainment situations was further explained in an EPA memorandum. In 1991, extension of the SILs to PSD Class I area increments was discussed. Further extension of the SILs to increments was provided in the draft 1990 New Source Review Workshop Manual and in the 1996 proposed New Source Review reforms, which continued the practice of using 4% of the NAAQS or PSD increment for SILs. The October 20, 2010 promulgation of the PM₂.₅ SILs continued this practice for Class II SILs, with a slightly lower percentage, 3.4%, for the 24-hour Class II SIL.

In their PSD implementation guidance for both the 1-hour NO₂ and SO₂ NAAQS, EPA established interim SILs based on 4% of the NAAQS. EPA indicates that they set the interim SILs at 4% “...because

---

7 45 FR 52707-52708, August 7, 1980.
8 EPA memo from Gerald Emison to Thomas Maslany, 1988; available at https://www3.epa.gov/scram001/guidance/mch/saq1.txt.
12 75 FR 64864, October 20, 2010.
13 EPA June 29, 2010 1-hr NO2 NAAQS PSD Implementation Guidance
14 EPA August 23, 2010 1-Hr SO2 NAAQS PSD Implementation Guidance
we believe it is reasonable to base the interim 1-hour SIL directly on consideration of impacts relative to the corresponding 1-hour NAAQS.” (August 23, 2010 SO₂ guidance, p.6)

The use of a 4% level for screening a source’s impact from further review is also present in the Federal Land Managers’ Deposition Analysis Threshold (DAT).

Similar to the SILs, this factor allows for the simultaneous effects of 25 such impacts without exceeding the applicable limit. Accordingly, the use of this cumulative factor has been in use in multiple areas to establish a threshold value that is not expected to threaten the level of concern.

2.2 SILs streamline the permitting approach in a way that appropriately addresses small sources.

It is appropriate for sources with modest emissions that do not materially change air quality to be able to avoid onerous permitting complications. SILs are useful PSD permitting tools that allow relatively small projects to appropriately demonstrate they will not cause or contribute to an exceedance of a NAAQS through a preliminary impact determination instead of through the more rigorous cumulative impacts analysis necessary for a large project. If the new and/or modified sources associated with a project have an impact less than the SIL (and there is room under the NAAQS for the SIL), then the project has a de minimis impact on existing air quality and the source impact analysis is complete. Without the SIL, a cumulative impacts analysis and PSD increment analysis would be required for every project, no matter the size.

SILs have been especially helpful in streamlining the permitting process for pollution control projects. For example, major SO₂ or NOx control projects involving flue gas desulfurization or selective catalytic reduction installations invariably have small PM₂.5 emission increases due to delivery, handling, and storage of materials associated with the new control equipment. The application of SILs avoids permitting delays for these beneficial projects with small PM₂.5 emission increases, and in many instances has resulted in early emission reductions that would otherwise be substantially delayed.

Cumulative analyses demonstrations are cost-intensive and time-consuming - both for a permitting authority and an applicant. Even setting up the associated modeling protocol can require multiple iterations and evaluations with the reviewing authority. The effort needed to conduct a cumulative modeling analysis, especially for Class I areas, is substantial because of the lack of complete emission inventories, with the burden on the applicant to verify source information for other facilities that may not be cooperative in sharing their data. In addition, the conservatism (over-prediction tendency) for modeling sources at their permitted levels distorts the representation of their impacts, and the use of a conservatively high regional background concentration further adds to the over-prediction tendency.

These conservative cumulative analyses are simply not necessary for smaller projects that will not have a significant impact on air quality. Without tools to screen out low-impact projects, state/local permit authorities are overwhelmed with work and are unable to focus on and expeditiously address those projects that have the potential for larger environmental impacts and/or benefits.

---

15 Available at https://www.nature.nps.gov/air/Pubs/pdf/flag/nsDATGuidance.pdf

16 Appendix W stipulates in Section 3.2.2 that approved models cannot be biased toward underestimating concentrations. In addition, Appendix W section 8 specifies that peak (allowable) emission rates are to be used for model input, which is generally an overstatement of the actual emission rate. Together, these considerations result in expected overestimates of modeled concentrations by EPA-approved models.
In the case of the application of a SIL for a NAAQS compliance analysis, a means for assuring that impacts within the SILs will not threaten the current NAAQS attainment status is to assess from available monitoring data the amount of concentration “buffer” that exists below the NAAQS. If the buffer exceeds the SIL, then the entire SIL is available to a proposed emission source. This policy is currently in place, as described in EPA’s PM$_{2.5}$ modeling guidance.\textsuperscript{17}

### 3.0 EPA APPROACH TO SET THE MAGNITUDE OF SILS

#### 3.1 Up to this point in time, the SILs have generally been set as a small percentage of the applicable NAAQS or PSD increment.

As noted in the discussion in Section 2.1, EPA has historically set aside a small percentage (usually 4%) of the available NAAQS or PSD increment for sources with \textit{de minimis} impact to ambient air quality. This percentage is based upon a cumulative factor of 25, in that the NAAQS could accommodate impacts of 25 equivalent sources with impacts below the SIL. This approach, along with the current practice to confirm that there is available buffer under the NAAQS, helps to assure protection of air quality with the SIL program in place. Continuation of this program will continue to allow new or modified sources with low emissions to utilize a streamlined permitting process.

#### 3.2 The form of the SIL (highest concentration) is conservative compared to the form of the ozone and the 24-hour PM$_{2.5}$ NAAQS and PSD increment.

In the past, the form of the NAAQS for short-term averages was often the second-highest concentration. Therefore, using the highest value for the SIL was reasonably consistent with the form of the NAAQS. However, the current ozone NAAQS design value is determined by the 4th highest daily maximum 8-hour average concentration (averaged over three years), and the PM$_{2.5}$ 24-hour NAAQS design value is set by the 98th percentile (8th highest) 24-hour average concentration (averaged over three years). Because of the inconsistent ranking (highest for the SIL, 4th or 8th highest for the ozone and PM$_{2.5}$ short-term NAAQS), this mismatch adds a substantial degree of conservatism if the previous practice of comparing the highest modeled concentration to the SIL were to be retained. From a cursory review of selected concentrations from EPA’s AirData database, it is apparent that the drop in concentration between the highest daily maximum 8-hour average value and the 4th highest daily maximum 8-hour average ozone value can result in an effective reduction of the SIL by up to 10-20\%. The effect is likely higher for PM$_{2.5}$ due to the larger difference between the form of the 24-hour NAAQS (8th highest 24-hour average) vs. the highest value for the SIL. The use of the peak value for the SIL, which could be an outlier value, is more consistent with the use of a 95\% rather than 50\% percentile confidence interval for the design value (DV) uncertainty.

EPA’s current guidance\textsuperscript{18} on the 24-hour and annual PM$_{2.5}$ SILs allows averaging of multiple years of modeled peak impacts for comparison to the SIL. The proposed SIL guidance does not clearly state that multiple-year averages can be used. EPA should clarify in its final guidance that multi-year averages of modeled impacts are the appropriate metric to be compared with the SIL.

\textsuperscript{17} Available at https://www3.epa.gov/ttn/scram/guidance/guide/Guidance_for_PM25_Permit_Modeling.pdf.

\textsuperscript{18} Available at https://www3.epa.gov/ttn/scram/guidance/guide/Guidance_for_PM25_Permit_Modeling.pdf.
3.3 EPA’s newly proposed technical approach that ties the magnitude of the SILs to uncertainties in the monitored design concentration has limitations that render it less supportable than the current approach using 4% of the NAAQS or PSD increment.

Up to the present time, the magnitude of the SIL has been tied to the NAAQS, on the order of 4%. As noted in Section 3.2, the effective percentage is actually less than 4% because the SIL is applied to the highest modeled concentration rather than the statistic associated with the short-term NAAQS. However, EPA has now expressed reservations about this decades-old practice and has decided to move to a new approach to estimate the magnitude of the SIL – tying it to the uncertainty in the value of the design concentration. This uncertainty, as discussed below, can be manifest in a number of forms: instrument accuracy / bias, sensitivity to variations in probe height due to vertical gradients of pollutant concentrations, and to limitations in the statistical variability due to a limited sample size. EPA focuses upon this last issue and does not address the other sources of uncertainty or bias. Due to the assertion by EPA that their computed uncertainty in the design value is in the “noise level” for determining NAAQS compliance, they postulate that a predicted concentration within this noise level cannot alter the NAAQS attainment status.

The new EPA technical approach has aspects that can greatly affect the final result, such as the choice of the confidence interval and the use of a specific database population that could change over time. Other aspects, such as the non-uniform spacing of monitors such that various areas of the United States are not given the same weighting in the analysis, have not been adequately explained by EPA. Furthermore, the whole approach does not address uncertainties and biases in instrument-reported data. The next section discusses these limitations in further detail.

4.0 EPA’S USE OF BOOTSTRAPPING TO SET THE MAGNITUDE OF THE SILS

4.1 EPA’s bootstrapping analysis is a work in progress that needs updating before it can be considered as a final product for EPA policy decisions.

EPA provides a lengthy technical discussion for its development of the ozone and PM$_{2.5}$ SILs in a 50-page main document with additional supporting tables and figures comprising 195 pages of appendices. Appendix A to our comments summarizes our review of EPA’s technical work and provides some recommendations for improvement. Two of the most important limitations in the bootstrapping analysis are summarized in Sections 4.2 and 4.3.

4.2 The confidence interval (50%) that EPA uses to establish the SILs is contrary to normal statistical practice (95%).

For the bootstrapping analysis, a key element in the determination of the uncertainty is the confidence interval selected. As discussed in Appendix A, the standard practice is to use a 95% confidence

---

19 On page 8 of the SIL guidance document, EPA notes “limitations in the rationale” for this long-standing approach.

20 EPA uses this term on page 9 of their Guidance document.


interval rather than a 50% confidence interval. EPA did not provide adequate justification as to why they have departed from the standard statistical practice for this specific application. It is acknowledged that the use of a 50% confidence interval would be more conservative and would limit the value of the SIL to a greater degree than the use of the standard 95% confidence interval. However, since the form of the SIL reflects the very highest value, rather than the more statistically stable form of the NAAQS or PSD increment, a larger confidence interval such as the standard 95th percentile is more appropriate. While the use of a 50% confidence interval implies that there is a 50% chance that the uncertainty is underrepresented, the use of a 95% interval reduces that chance to only 5%.

EPA’s technical analysis indicates that the choice of the confidence interval was a policy decision rather than a technical one. As discussed below, accounting for other sources of uncertainty, either as part of the bootstrapping analysis, or simply considering that reported instrument values have uncertainties that can affect peak values, would result in larger SIL values that those proposed by EPA.

4.3 EPA’s bootstrapping analysis using years of monitoring data addresses only one aspect of the design concentration uncertainty.

The EPA Technical Basis report states that the calculated variability in the observed design value (DV) at each site is due to variations in emissions and meteorology. However, EPA does not mention nor appear to include the contribution of instrument uncertainty (e.g., mean bias and/or imprecision) to the variability of the observed DV. EPA needs to clarify if or how the bootstrapping handles this EPA-allowed instrument bias and uncertainty. The variability determined from bootstrapping, for example, will not account for a systematic bias in the instrument operation.

Limitations in instrument accuracy and EPA-allowed instrument biases introduce a level of uncertainty in ozone and PM$_{2.5}$ observations that EPA must consider for the SILs, both with the bootstrapping analysis and apart from the bootstrapping analysis.

EPA’s technical analysis documentation of the bootstrapping results for design concentration uncertainty due to emissions and meteorological variability does not appear to consider biases in the reported instrument measurements that are allowed by EPA$^{21}$. Apart from the bootstrapping analysis, the instrument-caused errors in peak concentrations that can affect the DV provide an independent source of uncertainty in the reported design concentration.

EPA’s Quality Assurance Handbook for Air Pollution Measurement Systems$^{24}$, Appendix D, provides upper limits for the instrument bias at high concentration levels for any given single calibration exercise. They are +/- 7% for ozone and +/- 10% for PM$_{2.5}$ concentrations above 3 µg/m$^3$. For actual field monitoring, such biases can persist for an extended time in some cases until a manual site audit or calibration is conducted on the instrument. These biases exceed the 4% of the NAAQS levels for the SILs that EPA has adopted since 1980.


For PM$_{2.5}$, an RTI study$^{25}$ that reviewed data from co-located samplers estimated an average measurement uncertainty over a 3-month period in 2004 exceeding 5%. This level of uncertainty is about the same magnitude as the national average bias of 4.3% reported by EPA$^{26}$ for the 2011-2013 three-year period. These levels also exceed the 4% of the NAAQS levels traditionally used by EPA for the SILs.

**Known vertical variations in monitored ozone and PM$_{2.5}$ concentrations over the range of allowed instrument probe heights introduce another level of uncertainty in the monitored concentrations.**

EPA allows ozone monitoring probe heights to vary between 2 and 15 meters, even though the “breathing zone” of nearly all ground level receptors (humans) is below 2 meters. However, since PM$_{2.5}$ and ozone concentrations are reduced near the surface by deposition and absorption onto obstacles and vegetated surfaces, the concentration profile is typically not flat (constant) over this height range.

It is well known that ozone is deposited to the ground by interactions (e.g., chemical reactions, impaction) with vegetation and soil.$^{27,28,29,30}$ The change in concentration with probe height adds to the measurement uncertainty of ozone because of the differences of the probe heights among monitors. For example, API previously presented$^{31}$ on the need to correct ozone measurements to a reference 2-m height. The presentation noted that the average probe height is 5.4 m (although it is 10 m for CASTNET sites).

Field studies have quantified the vertical ozone concentration gradient within the range of probe heights allowed by EPA. For example, Wisbith and Meiners$^{32}$ report the results of a summer field study near

---


Cincinnati, Ohio roadways with ozone concentrations increasing by 10% on average between 2 and 10 meters.

5.0 EPA SHOULD USE 4% OF THE NAAQS OR PSD INCREMENT TO ESTABLISH SILs

5.1 EPA’s bootstrapping approach uses a confidence interval that is too low.

If EPA were to use a standard 95% confidence interval in their bootstrapping analysis, the uncertainty in the reported ozone or PM$_{2.5}$ DV would be increased, as can be computed by applying the results of EPA’s own analysis for the 95% confidence interval:

- the DV uncertainty for ozone (8-hour average) would be 4.4%,
- the DV uncertainty for the PM$_{2.5}$ annual average would be 4.9%, and
- the DV uncertainty for the PM$_{2.5}$ 24-hour average would be 11.3%.

These more appropriate uncertainty levels would increase the SIL estimate to be higher than 4% of the NAAQS. In any case, due to the selection of a non-standard confidence interval and the uncertain influence of other biases in the measurements, the EPA bootstrapping approach needs further development. It is not superior to the practice of using a cumulative factor approach (4% of the NAAQS).

5.2 Other sources of error or bias independently result in high DV uncertainties: instrument measurement errors and concentration gradients as a function of height.

Even without variations in the DV due to emissions and meteorology changes, there are other reported DV uncertainties due to EPA-allowed instrument biases and changes in concentrations with probe height. These uncertainties are generally comparable to or larger than the traditional 4%-of-the-NAAQS rule of thumb for the SILs. Especially for ozone, it is evident from the studies cited above that probe heights above 2 meters could lead to systematic biases of reported concentrations because the ozone concentrations above the breathing zone (up to 2 meters) are generally biased on the high side.

5.3 Based upon our recommendation that EPA continue to use 4% of the NAAQS for the SIL, we provide the resulting SIL values for ozone and PM$_{2.5}$ that EPA should adopt.

The tables provided below list the EPA-recommended SILs based upon its new approach, along with our computation of the SILs with the use of 4% of the NAAQS for ozone and PM$_{2.5}$, but accounting for the current PM$_{2.5}$ SILs that are still in place.

### NAAQS SILs Using EPA’s New Approach vs. the Traditional 4% Approach for Ozone and PM$_{2.5}$

<table>
<thead>
<tr>
<th>Criteria Pollutant (NAAQS level)</th>
<th>NAAQS SIL Concentration (EPA proposal)</th>
<th>NAAQS SIL Concentration (4% Approach)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ozone: 70 ppb (8-hour)</td>
<td>1.0 ppb</td>
<td>2.8 ppb</td>
</tr>
<tr>
<td>PM$_{2.5}$: 35 µg/m$^3$ (24-hour)</td>
<td>1.3 µg/m$^3$; rule is 1.2 µg/m$^3$</td>
<td>1.4 µg/m$^3$; rule is 1.2 µg/m$^3$</td>
</tr>
<tr>
<td>PM$_{2.5}$: 12 µg/m$^3$ (annual)</td>
<td>0.2 µg/m$^3$</td>
<td>0.48 µg/m$^3$; rule is 0.3 µg/m$^3$</td>
</tr>
</tbody>
</table>

* accounts for the PM$_{2.5}$ SIL values still in effect in Section 51.165(b)(2).
### PSD Increment SILs Using EPA's New Approach vs. the Traditional 4% Approach for PM$_{2.5}$

<table>
<thead>
<tr>
<th>Criteria Pollutant (averaging period)</th>
<th>EPA Proposal</th>
<th>4% Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM$_{2.5}$ (24-hour)</td>
<td>0.27 and 1.2 µg/m$^3$ for Class I and Classes II/III</td>
<td>0.27 and 1.2 µg/m$^3$ for Class I and Classes II/III</td>
</tr>
<tr>
<td>PM$_{2.5}$ (annual)</td>
<td>0.05 and 0.2 µg/m$^3$ for Class I and Classes II/III</td>
<td>0.075 and 0.3 µg/m$^3$ for Class I and Classes II/III</td>
</tr>
</tbody>
</table>

Notes:
1) The ratio of the PM$_{2.5}$ 24-hour Class I and Class II increments is 2/9, or 0.222. The resulting 24-hour PM$_{2.5}$ Class I SIL is $1.2 \times 0.222 = 0.27$ µg/m$^3$.
2) The ratio of the PM$_{2.5}$ annual Class I and Class II increments is 1/4, or 0.25. The resulting annual PM$_{2.5}$ Class I SIL is $0.3 \times 0.25 = 0.075$ µg/m$^3$.
3) The Class I increment SILs are based upon the PM$_{2.5}$ Class II SIL values still in effect in Section 51.165(b)(2): 1.2 and 0.3 µg/m$^3$ for the 24-hour and annual averages, respectively. These SIL values are lower than the 4% of the PM$_{2.5}$ NAAQS values, so they are the most constraining values to be used in calculating the Class I increment SILs.
Overview of Comments

The EPA document (‘Technical Basis’) includes a mixture of detailed quantitative statistical calculations (which are generally quite reasonable) and policy decisions. In the sections on statistical analysis, EPA describes investigation of variability (uncertainty) in the observed (monitored) Design Values (DV) for ozone and PM\(_{2.5}\) at each monitoring site in the United States air monitoring (AirData) database. Some sites have been operating for over 15 years and other newer sites have only one or two years of record. There are over 1770 sampling sites in the US for both ozone and PM\(_{2.5}\). To estimate the variability in the observed DV at each site, a standard and well-accepted statistical resampling method (the bootstrap) is used. The method produces an estimate of the full cumulative distribution function (CDF) of the estimated variability of the DV, from which the EPA picks a few standard points (e.g., the 25th, 50th, 75th, 95th, etc. percentiles) to report in tables and figures.

The EPA Technical Basis report states that the calculated variability in the observed DV at each site is due to variations in emissions and meteorology. However, EPA does not appear to consider the contribution of instrument uncertainty (e.g., mean bias and/or imprecision) to the variability of the observed DV. EPA’s allowed uncertainties for the ozone and PM\(_{2.5}\) samplers (discussed in the main text of our comments in Section 5.1) are in the range of 7-10 percent. Because the EPA calculates the DV variability/uncertainty site-by-site rather than over the entire set of samplers at once, if the sampler has a mean bias that occurs for an entire year, the influence of that bias is missed by EPA in their reported uncertainty values provided in the quantitative tables and figures in their report. Thus, even if we were to agree with the EPA’s reasoning that the SIL is proportional to the observed variability in DV, it is essential that the contribution of instrument uncertainty also be considered, and the variances should be added if they are independent.

As done in the past, a single SIL is chosen for each NAAQS (e.g., PM\(_{2.5}\) and ozone). This SIL is applied across the US. The EPA report contains figures and tables that attempt to show that the calculated DV variability does not vary by ‘too much’ across the over 1770 sampler sites for each pollutant. They also calculate the difference in DV variability for sites located close to each other. Variability in DV occurs from one site to another, but the assumption seems to be made that agreement within a factor of about two is satisfactory in the current analysis.

A fundamental issue in the EPA analysis is that they use the 50% confidence interval (determined from the DV variability CDF) rather than the almost universally-used 95% confidence interval. The result of this choice is that the DV variability assumed in this report is a factor of 3 or 4 lower than that used in most other analyses of variability. The EPA report chooses to assume, without providing rationale (other than saying that they must consider policy issues in addition to statistical results), that the SIL is proportional to the bootstrapped estimated ‘50%’ uncertainty range (i.e., determined by the 25th and 75th% points on the CFD). This can be calculated as one half of the bootstrapped 75% minus the bootstrapped 25% DV. The much more common ‘95% confidence interval’ is encompassed by the 2.5th and 97.5th% points on the CDF.
Specific Comments on EPA's Technical Basis Document

(Note: ‘p’ is used as an abbreviation for ‘page’, and ‘par’ for ‘paragraph’ in the text below.)

- page 5, paragraph 2 – The term ‘significant’ is nearly always linked with an assumed confidence level – usually 95%. EPA should acknowledge and state that.

- p 6, par 2 – The variability could be calculated directly, without bootstrapping. For example, suppose there are 10 years of data. Then 10 values of the annual PM$_{2.5}$ concentration can be calculated and the mean and variance and range determined. The purpose of the bootstrap resampling is to improve on that direct calculation of variance from a limited sample size.

- p 6, line 5 from bottom – The bootstrap resampling cannot account for correlations, unless extreme conditions are imposed on the method. EPA should acknowledge or explain this issue.

- p 7, line 13 of first full par – ‘the 50% CI was selected.’ No detailed justification is offered for the selection of 50%, though that is a nonstandard CI selection. Most studies of statistical confidence (e.g., the Hanna reference that EPA relied upon; see next comment) use 95% confidence intervals. For a Gaussian distribution, the range of the 95% CI is about 2.6 times the range of the 50% CI.

- p 7, footnote 26 – The author’s name should be ‘Hanna’, not ‘Hannah’. This reference uses 95% confidence limits.

- p 8, 3rd bullet – It should be emphasized that this is the 98th percentile of the available 24-hour data in a year.

- pp 9 and 10 – Figures 1 and 2 give information on all US samplers, and it is stated that there are over 1770 samplers, so there must be many indistinguishable overlapping dots on the figures. Thus, the reader has only a partial picture from what can be ‘seen’ on these figures. Providing zoomed-in data inserts for regional areas would be helpful in reviewing the distribution of monitoring data sites.

- Section 2.2.1 on General Overview of Statistical Methods (pp 12 – 15) – This is a reasonable overview.

- p 16, first line of last par – ‘monitoring sites are assigned a spatial scale’ – There is much ambiguity in this assignment, and it is unclear how this affects the statistical analyses. It is evident that the monitors are not evenly spaced and that some areas of the country are not as well represented (e.g., portions of the West) as other areas (in the East). EPA should better explain how this non-uniform distribution of monitors is handled in their analysis, especially if monitors in the East tend to bias the results because that part of the country has an overwhelming fraction of all of the sampling sites.

- p 17, line 4 – Sensor uncertainty (error or bias) can affect the calculated variability. Explain how that effect is accounted for.

- p 18, last par of section – What if these changes, such as a commercial development, are not known?

- p 19, footnote 29 – It is possible that the procedures for ‘accounting for multiple monitors, handling of exceptional events’ etc., can influence the bootstrapped distribution functions, so EPA should address this issue in a revised technical document.

- p 20 – The procedure described here is described in other references as Latin Hypercube or blocking. Please provide some literature references. The goal, as EPA states, is to be sure to sample adequately from the separate groupings of data. However, there needs to be adequate rationale provided for the blocking procedure.
• p 20, bottom 7 lines and p 21 top 15 lines – EPA did 20,000 resamples, and found the analysis ‘did not appear to change after 3000-4000 resamples’. But EPA should define what that means. For example, is the ‘change’ criterion defined by 1 %, 0.1%, 0.01%?

• p 21, bottom bullet and footnote 30 – The choice of the 50% CI is not explained here. Although it’s addressed later, the decision to choose a 50% CI should be discussed when it’s first applied. Most analyses of confidence limits use a 95% CI. It becomes apparent later in the document that the choice is a policy decision, but that is also the case for the currently-used 4%-of-the-NAAQS “rule of thumb.”

• p 23 (ozone results) and Figure 4 – Since there are about 1770 samplers, we would expect to see 1770 points on the top and middle parts of Figure 4. However, most of the points overlap and all that we can distinguish is the few outliers. Perhaps the figures could be enlarged to better distinguish the overlapping portions of what is plotted. The bottom part of the figure is more useful, since quantitative results are more easily distinguished.

• p 23 – From this analysis, it appears that only the results for the individual sites are presented. In that case, the listed variability may account for random day-to-day sampler (instrument) errors, but does not account for mean bias errors, which would apply to longer periods or to the entire year. For example, sampler 1 might be biased 1 ppb high and sampler 2 might be biased 1 ppb low. It is known that samplers often exhibit mean bias, of magnitude similar to the SIL, and this is neglected in the current analysis.

• p 23, line 7 from bottom – ‘there is not a trend in the relative variability’. How was this conclusion drawn? Figure 4 needs a legend and a better explanation of what is plotted. Conclusions should be supported quantitatively.

• Fig 4 bottom –The median is about 1.5 for the ‘50% upper’ (chosen later for the SIL) and is about 5 for the ‘95% upper’ (suggested in the Hanna reference relied upon by EPA and used by most other studies). This difference of a factor of 3.3 is close to the factor of 2.6 difference mentioned in the general p.7, line 13 comment above about the expected value for a Gaussian distribution.

• p 25-27 (ozone results, including Figs 5 and 6) – We have similar comments on the PM\textsubscript{2.5} analysis and figures as we noted for the ozone analysis and figures. The plots such as in Fig 5 are difficult to discern because of the 1770+ points and lines on a single diagram.

• p 27, Figure 6 – This figure is useful (as was the bottom panel of Fig 4 for ozone) because the 1770 sites are collected into box plots. We are interested in the difference between the median of the box plots for the ‘50% upper’ and the ‘95% upper’ (1.5 vs 5% for annual and 3 vs 12% for 24-hour averages). Note that the ratio of the 95% to the 50% is about 3 to 4 for the annual and the 24-hour PM\textsubscript{2.5} averages. As stated above, the ratio is about 3.3 for ozone (estimated from the figures). Nevertheless, it is clear that, if the EPA had selected the 95% CI, the variability (and their SIL selection) would be about 3 to 4 times larger.

• Section 3.2.1 on Analysis of PM\textsubscript{2.5} Spatial Variability (pp 28-33, includes Table 1 and Figures 7-9) – We agree that it is useful to know the spatial variability of the calculated statistics and whether nearby sites’ variabilities are correlated. Several quantitative results are listed and plotted. However, some of the subsequent discussion includes subjective statements (hypotheses) and it is unclear if confidence limits were calculated to test the hypotheses.

• p 28 top par – line 3 – The ‘scale’ is arbitrarily assigned. Please provide more explanation for the choice of the scale.
• p 28 - This page should mention that terrain differences can also have a large effect on monitor vs. monitor comparisons. In addition, differences in the angle or direction that the monitors are located with respect to a major source can have a large effect. There could be, for example, two groups of 'close' monitors – in urban areas where there are many distributed sources, and in non-urban areas where a large point source is located. We would expect differences in the metrics for the spatial analysis for these two groups. In addition, probe height differences among monitors can add to the uncertainties.

• p 28, line 2 of 3rd paragraph – 'strong correlation' and line 9 'similar'? What is the confidence in the statement (or the null hypothesis)?

• p 28 last par - We note that the characteristics of the annual wind rose could have a great influence on spatial pollution patterns.

• p 30 – Table 1 – Can EPA further comment on which of these close monitors are 1) in urban areas, and 2) in non-urban areas with single large sources? These two groups should have different spatial variability characteristics.

• p 31, Fig 7 (6 panels) – It is difficult to distinguish between patterns with separation distances indicated by different shades of blue. On the top row (monitors 1 and 2 concentrations), it looks like plus and minus about 2 µg/m² for annual and 5 µg/m² for 24-hour averages. The bottom two rows of points (delta and percent) look like shotgun blast patterns (i.e., large variability with no relation).

• pp 32 and 33 (Figs 8 and 9) – We have the same comment as above for Figs 2 and 3. But in this case there are several key locations with overlapping and obscured dots (e.g., in the Washington to Boston corridor). Can a zoomed-in figure be added for those areas? Also, the captions and legends are unclear. A specific example is the statement in the caption ‘bottom panel shows the difference divided by the mean’. However, we believe that what is shown is 100 times what is stated. That is, EPA is plotting percent rather than fraction.

• Section 3.2.2 – Analysis of the influence of PM₂.₅ Monitor Sampling Frequency (pp 34-36):

  p 34, par 2 – first sentence – Define what is meant by statistics. Is it the mean, median, variance, points on CDF? If EPA is referring to the mean or median, then it is correct that their variability is larger for less frequent sampling. But if EPA is referring to the variance or details such as 95% confidence range, then the values are smaller for less frequent sampling. This is because, for smaller sample sizes, there is less probability that extreme values will be sampled. Please revise this paragraph.

  p 34, last par – The logic in this paragraph is difficult to follow and should be clarified.

  p 35, Fig 10 – 2nd panel – The vertical axis seems like it should be number of samples rather than frequency.

• Section 4 – Application of Air Quality Variability to Determine SILs for the PSD program (pp 37-49). The logic converting the quantitative statistical analysis of the DV variability for the ozone and PM₂.₅ sampler data in the previous sections to suggestions for SILs is unclear. See the discussion in the above ‘Overview of Comments’ section.

• p 37, first sentence of 4.1 – Is it the ‘programmatic and policy considerations’ that led to the choice of the 50% CI rather than the much-more-widely-used 95% CI? The 95% CI is used by EPA for setting the NAAQS based on health impact research. As noted above, theory and the current report’s bootstrapped concentrations show that use of the 95% CI would lead to about a 3 to 4 times larger DV variability (and hence SIL, using the current EPA report’s logic) than use of the 50% CI.
- p 37, line 4 of 4.1 – A ‘single SIL value for each NAAQS’ is selected. However, the basis for the choice of what to consider as acceptable variation from monitor to monitor or year to year is not clear. More justification is needed, including discussion of the effect of sampler errors.

- Section 4.1.1 ‘Selection of the 50% CI for the SIL’ – The justification for choosing the 50% CI is weak. Although EPA indicates on page 38 that ‘... there is no scientific reason to select any one CI over another’, nearly all studies choose the 95% CI, including the EPA’s own documents supporting setting of specific NAAQS. For example, for the SO\textsubscript{2} NAAQS review recently conducted by EPA, the analyses of SO\textsubscript{2} effects on health all used the 95% CI.

- p 43 Fig 12 – all three panels – Is this for all samplers? If so, why aren’t there 1770 points? Note that if the 95% CI had been chosen, the magnitudes of all the ‘absolute and relative uncertainties’ on this figure would be three to four times larger.

- p 44, line 2 from bottom – Long-tailed distributions are common for environmental variables (e.g., wind speed, rain rate, pollutant concentrations), and therefore the mean is expected to be larger than the median.

- p 45, line 5 – ‘based solely on a visual inspection’. As stated earlier, statistical confidence limits should be used in support of hypotheses.

- p 48, Table 4 – Here again we see that the 95% CI is about 3 times larger than the 50% CI.