February 1, 2013

sierra research

1801 J Street
Sacramento, CA 95811
Tel: (916) 444-6666
Fax: (916) 444-8373

Ann Arbor, MI
Tel: (734) 761-6666
Fax: (734) 761-6755

CalEnviroScreen
c/o John Faust
Chief, Community Assessment & Research Sectioon
Office of Environmental Health Hazard Assessment
1515 Clay Street, Suite 1600
Oakland, CA 94612

Dear Mr. Faust:

CalEnviroScreen has been developed as a tool to help Cal/EPA identify "California's most burdened and vulnerable communities."[1] The tool is intended to help inform decisions at Cal/EPA by identifying places most in need of assistance.

In developing this tool, Cal/EPA has undertaken a monumental public involvement process. Moreover, the revisions that have been made demonstrate that Cal/EPA has listened to comments and suggestions from various stakeholders. Nevertheless, the tool in its current form still does not meet Cal/EPA's goal of creating a scientifically justified methodology for identifying vulnerable communities. We believe that some of the changes that Cal/EPA made, apparently in response to comments from stakeholders, lack a valid scientific basis. The following comments describe some of the remaining shortcomings and suggest revisions to address them.

The end result of the CalEnviroScreen tool should not be a numerical score.

The current form of the CalEnviroScreen tool, which provides a numerical score for each zip code, is problematic for several reasons and should be replaced with a tool that places communities into categories (the "matrix" approach). The matrix approach, which was recommended by several participants in the academic review panel (with this recommendation restated by two of the panelists at the January 11 workshop), would utilize many of the same indicators as the current model. However, the indicators would be used to categorize each community as "High" or "Low" (or "Medium") Exposure, and "High or "Low" (or "Medium") Vulnerability. Each community would be placed in one cell of a 2x2 or 3x3 matrix. Additional agency attention with respect to resource allocation would be directed toward High Exposure, High Vulnerability communities.

[1] Cal/EPA, *CalEnviroScreen Public Review Draft* (Jan 3, 2013), p. 1.

The numerical score produced by the currently proposed tool invites misinterpretation and misuse. This potential is demonstrated by the following statement in the *Summary of Major Proposed CalEnviroScreen Revisions* that accompanied the current draft: "A decimal place is retained for each score in the calculation of the CalEnviroScreen score to help distinguish ties." There is no need, under any of the uses proposed by Cal/EPA for this tool, to "distinguish ties." The accuracy and precision of the data used to calculate the indicators do not support such distinctions. The numerical score contributes to the erroneous perception of precision, and that the specific rank ordering of communities by CalEnviroScreen score has any meaning.

Adopting the matrix approach simply recognizes the limits of precision and specificity achievable with the screening tool, given data limitations. With this change and a few other adjustments (discussed below), the current CalEnviroScreen tool could be reconfigured to avoid most of the technical objections to the current tool, yet achieve Cal/EPA's goal of identifying high-priority communities.

If a multiplicative approach is retained in lieu of a matrix, the 10x magnitude of modulating pollution burden by population vulnerability is not scientifically justified.

No justification is provided for using a scale of 1-10 for population characteristics as a modulating factor (multiplier) for environmental burden. There is no scientific justification offered for concluding that potential health effects can be as much as 10 times higher in one community over another, given similar exposures, simply due to population vulnerability. For example, two indicators that might be construed to provide information about community health (asthma emergency room visits and low birthweight) show a relative range between 5th and 95th percentile communities of 3.5 and 1.5, respectively—a factor of 2.3 difference in health outcomes attributable to all causes (not just environmental burden). It is simply not credible to assume that population characteristics can magnify environmental effects by as much as a factor of 10.

If the population characteristics score is intended to provide a measure of the relative vulnerability of a community, given similar exposure to pollution, then the potential factor of 10 is much too high.

One of the core tenets of the Environmental Justice movement is that disadvantaged communities receive a disproportionate burden of environmental hazards. The presence of a disproportionate burden is reflected in the screening tool through the various burden indicators. The population characteristics adjustment is therefore necessary only to the extent that those characteristics affect the community's health response to pollution burden. We believe the population characteristics adjustment should be comprised of indicators that reflect the community's inability to mitigate the health effects of pollution (e.g., due to limited access to health care), and increased susceptibility due to pre-existing circumstances (e.g., age). The magnitude of this adjustment should not exceed the observed relative range of health outcomes between communities. The tool should be focused on human health.

The CalEnviroScreen tool, as proposed, is technically and scientifically flawed.

Because of the flaws described in detail below, the current tool produces results that are at best suspect, and at worst absurd. Several of these comments were made in our September 24, 2012 letter to OEHHA, which addressed the previous draft of the tool; those comments that are still relevant to the January 2013 draft of the tool are briefly restated here.

A number of statements in the report do not reflect the conclusions of the cited studies.

There are several instances in which cited references used to support indicators are not consistent with the alleged citation. For example, the Diesel Particulate Matter (DPM) indicator is justified, in part, by reference to the following scientific studies, which do not support the attributed conclusions:

- Statement: "Studies have shown that diesel PM exposure during pregnancy can result in low birth weight, birth abnormalities, and infant mortality"
  Citation: Parker et al., 2006
  Study conclusion: The Parker study evaluated $PM_{2.5}$, not diesel PM. $PM_{2.5}$ is a separate indicator that is already included in the screening tool.

- Statement: "Sensitive populations such as children, the elderly, and those with existing respiratory and cardiovascular disease are particularly susceptible to the harmful effects of exposure to airborne PM, including diesel PM."
  Citation: Sacks et al., 2011
  Study conclusion: The Sacks study evaluated PM, not diesel PM. While the overall conclusion of the study is that PM may contribute to health effects, this study does not support special treatment of diesel PM. Again, $PM_{2.5}$ is already a separate indicator in the screening tool.

- Statement: "Exposure to air pollutants from vehicle emissions has been linked to adverse birth outcomes, such as low birth weight, premature birth and certain birth defects."
  Citation: Ritz et al., 2007
  Study conclusion: The article abstract makes mention of preterm birth; it is possible that the other outcomes (such as low birth weight) were also weakly correlated; however, if so, the correlation was not strong enough for the authors to mention it in the abstract.

The Diesel Particulate Matter indicator is flawed and redundant, and should not be used.

The addition of a DPM indicator is not technically justified. Furthermore, the data used for this indicator are inappropriate for several reasons. The DPM indicator is not technically justified because it is duplicative of, and potentially correlated with, two other indicators already present in the model: traffic density and, to a lesser extent, $PM_{2.5}$

concentrations. Because DPM is not (in fact, cannot be) measured in the ambient air, the proposed indicator is based on surrogates, and its inclusion weakens the tool.

The data selected for use in this indicator are also flawed. The academic panel recommended that indicators be direct applications of primary data. When a model's indicator is an amalgamation of multiple indirect data sources, it creates a lack of transparency and therefore has potential to confound results. The data used to calculate the DPM indicator are the *modeled* annual average concentrations. This would be one of the few indicators based on secondary values derived from primary data, instead of on primary data.

In addition, the NATA data present unreliable results, particularly as used in the tool. For example, DPM emissions from mobile (both on-road and non-road) sources are not inventoried on a census tract (or zip code) basis; rather, they are inventoried at the county level, then disaggregated to smaller geographic sectors using allocation formulas based on surrogates (e.g., population, construction costs [for construction equipment], employees in manufacturing [for industrial equipment]). This defeats the purpose of the tool in terms of identifying differences in community exposures at the sub-county level. EPA has stated that NATA data should not be used "to characterize or compare risks at local levels such as between neighborhoods."[2] The mere existence of values in the NATA database at the census tract level is not evidence of the scientific validity of those values for use at the localized level of the CalEnviroScreen tool.

Furthermore, the NATA database is unreliable and incomplete. The latter criticism is best illustrated by the fact that OEHHA found it necessary to augment the NATA data with modeling of ambient DPM concentrations by the California Air Resources Board (CARB) of ports and railyards. If NATA does not properly account for these obvious and significant sources, then other significant omissions could exist for which OEHHA does not adjust. Furthermore, the database cannot be corrected by simply adding one or more facilities or source categories, because such *ad hoc* adjustments distort the results. The entire goal of this effort is to use a scientifically defensible tool to establish relative priorities. By explicitly adding information about one source category but not others, the results of the analysis are biased away from all other source categories, some of which may be as important as the ones receiving special attention.

Finally, the NATA diesel data are out-of-date. The latest data in the most recent database are from 2005—since then, CARB diesel regulations significantly reducing emissions from stationary and mobile diesel engines have become effective, EPA Tier 3 emission standards have become effective, and transition to Tier 4 has begun. In addition, CARB, the railroads, and the ports have implemented goods movement initiatives that have greatly reduced the emissions that were the basis for the modeling isopleths used to supplement the outdated NATA database, and further reductions are underway. OEHHA has apparently updated the CARB DPM isopleths for ports and railyards to reflect these emission reductions, but the NATA data were not subjected to similar updates.

---

[2] EPA, *An Overview of Methods for EPA's National-Scale Air Toxics Assessment* (January 31, 2011), p. 5.

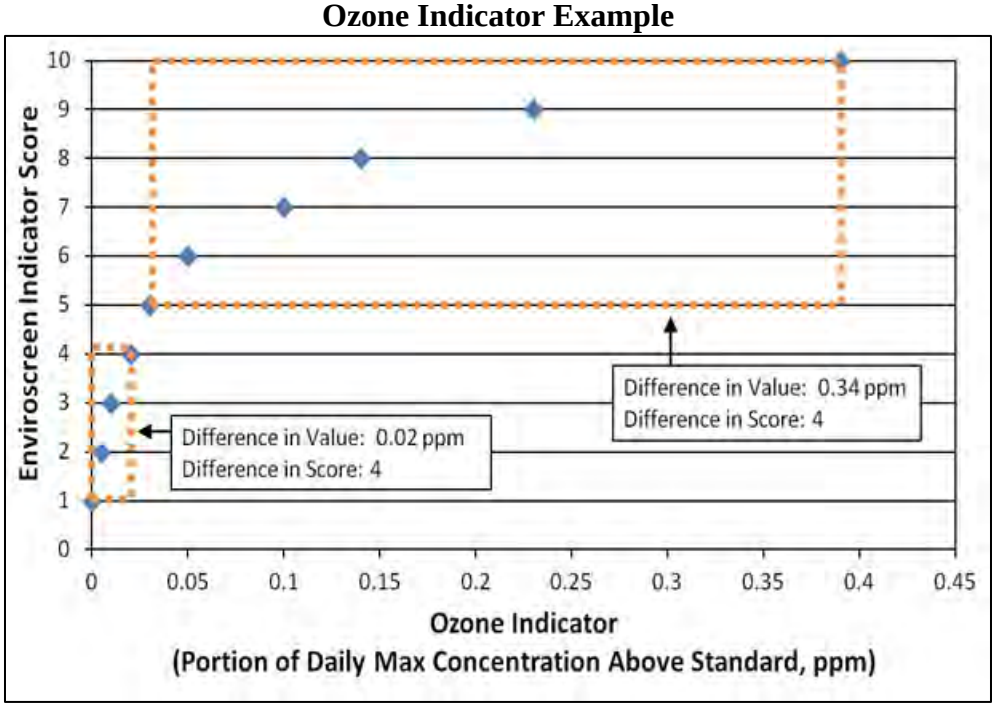
For all of these reasons (age and inaccuracy of the emission estimates, the additional uncertainty introduced by the use of derived rather than measured values, and EPA's own recommendation for use of NATA), the DPM indicator based on NATA should not be included in CalEnviroScreen.

The use of percentile scores for indicators, rather than normalized actual values, contributes to the unexpected predictions of the tool.

The relationship between percentile and absolute value is extremely non-linear for some indicators (e.g., ozone, DPM, pesticide use, toxic releases). For these indicators, a small difference in absolute value can have a big effect on the percentile score at the low end, and a big difference in absolute value can have a relatively smaller effect on the percentile score at the high end.

Using ozone as an example, Figure 1 below shows that the difference in the **value** of the ozone indicator between the 45th percentile (score of 4.5) and the 5th percentile (a score of 0.5) is 0.02 ppm, or less than 30% of the state air quality standard. The difference between indicator **scores** is 4. Figure 1 also shows that the difference in **value** of the ozone indicator between the 95th percentile (score of 9.5) and 55th percentile (score of 5.5) is 0.34 ppm, or nearly five times the state air quality standard. The difference between indicator **scores** is also 4.

Figure 1
Ozone Indicator Example



Indicator values are the low end of the range for each decile as shown in the indicator map on page 18 of the *CalEnviroScreen Public Review Draft (Jan 13, 2013).*

For the first pair of communities in this example (5th percentile and 45th percentile), there is no significant difference in air quality as reflected by ozone concentrations. In contrast, there is a significant difference in ozone air quality between the second pair of communities (55th percentile and 95th percentile). CalEnviroScreen, however, treats these two differences as equally important: their effect on the CalEnviroScreen score is the same.

The use of indicator values that are normalized from the data would retain information about the relative severity of the indicator that is lost using the percentile approach; this is what we recommended explicitly for the ozone and $PM_{2.5}$ indicators, and we believe this recommendation is appropriate for other indicators as well. As was expressed by one of the academic review panelists, the communities that we are most interested in are those at the extreme tail of indicator values. The current scoring approach dilutes the effect of extraordinarily high impacts.

The $PM_{2.5}$ indicator should be based on emissions above the health-based ambient air quality standards, not total PM concentrations.

As recommended by several commenters, the ozone indicator was modified in the January draft to consider only ambient concentrations above established "safe levels," such as ambient air quality standards. The rationale supporting this approach applies to $PM_{2.5}$ as well.

The average of quarterly means of daily average $PM_{2.5}$ concentrations is not a good metric, because it suggests that there is a meaningful distinction in exposures at various concentrations below the state ambient air quality standard. A better indicator of exposure would be something similar to degree days: the sum of 24- hour average $PM_{2.5}$ concentrations in excess of the federal standard. The result would be expressed as $\mu g/m^3$-days, and could be averaged over three years to smooth out weather variations. Again, a maximum score of 10 would be assigned to a fixed (not relative) concentration (e.g., 2 times the federal standard); for example, a score of 10 would be assigned to 2 x 35 $\mu g/m^3$ x 365 days = 25,550 $\mu g/m^3$-days.

Compliant sites should have a score of zero.

Leaking underground storage tanks and other cleanup sites that have been fully remediated should have a weighted score of zero. Similarly, solid waste and hazardous waste facilities that are in compliance with environmental requirements should have a weighted score of zero or close to zero.

The data used for the asthma indicator do not match the rationale provided for the indicator.

Asthma emergency room (ER) visits are included in CalEnviroScreen as markers for the cumulative impacts of environmental and social stressors. However, the justification provided for this use of the indicator does not support its use for this purpose. This indicator is included on the premise that the number of asthma ER visits is a reflection of community susceptibility to pollution, whereas it actually embodies a number of factors,

including access to health care, the actual incidence of asthma in the population, and exposure to allergens and irritants, both indoors and outdoors.

Most of the arguments provided in the report for inclusion of an asthma indicator appear to be related to asthma incidence as an indicator, not emergency room visits; others seem to be holdovers from the proposed use of emergency room visits as an indicator of burden rather than vulnerability. Neither set of arguments supports the hypothesis that emergency room visits is an indicator of vulnerability independent of economic status. If an asthma-related indicator is retained, it should be based on the prevalence of a pre-existing health condition that may be exacerbated by environmental burdens, and not on a metric that combines this with other indicators independently reflected in the model.

Finally, there is a potential problem with this indicator to the extent that patients' residences are in different communities from their workplaces. An asthma attack triggered by workplace exposure has nothing to do with environmental conditions in the residential community or in the community surrounding the workplace.

The Low Birthweight indicator appears to be essentially duplicative of the Poverty indicator, and should be eliminated.

At a minimum, a sensitivity analysis should be performed to determine the extent of overlap. If, as appears to be the case from the CalEnviroScreen maps, the indicator is duplicative, one or the other should be eliminated.

The proposed indicator for Race/ethnicity, as a measure of vulnerability, is not supported by the cited references nor by the rationale provided.

The extent to which race/ethnicity reflects increased vulnerability to environmental effects independent of economic factors is unclear. If OEHHA believes, based on published studies, that this is in fact an independent indicator of vulnerability, the proposed use of this indicator should be clarified. The rationale provided in the draft report indicates that there is a correlation between race and adverse response to specific pollutant exposure, but does not make a scientifically supported case for causation. All of the examples provided indicate increased sensitivity among African-Americans, compared with white populations, except for one citation to a study involving Hispanic children. The rationale goes on to explain that the reasons for the correlation are not well understood, but could be due to cumulative stress or some other cause that has not been captured or accounted for in the studies' designs.

Inclusion of this factor does not meet OEHHA's stated criterion of scientific justification. The justification in the report could, arguably, be used to justify using African-American population as an indicator. It does not support using non-white population as an indicator.

There is no scientific justification of inclusion of "linguistic isolation" as a factor.

The rationale provided for inclusion of this indicator is scientifically weak. If this indicator is retained for other reasons, it should modulate another indicator (such as poverty level), rather than being an independent indicator of population vulnerability.

Information about linguistic isolation may, however, be very useful in designing programs for communities that are identified using legitimate criteria.

The Age indicator (children and elderly) is not consistent with treatment of age used elsewhere by OEHHA (e.g., cancer risk assessment age adjustments).

The definition of children as age 5 or less is inconsistent with the approach OEHHA has used in other screening tools. For example, in other tools OEHHA has defined children as the 14-year span between age 2 and age 16, and assigned an age sensitivity factor of 3 to the child category. The definition of elderly as age 65 or more is inconsistent with the approach OEHHA has used in other screening tools. For example, in other tools OEHHA does not apply an increased sensitivity before age 70.

A number of issues previously raised remain unaddressed.

The issues below were noted in our comments on the previous draft, submitted on September 24, 2012. Please see that comment letter for more details.

- Pesticide use is not a good surrogate for pesticide exposures. Although the revisions made to this indicator (focusing on volatile, hazardous pesticides in agricultural use) improve the indicator, it is still a poor indicator.

- Toxic releases (as reported in the TRI) are not a good surrogate for exposure to toxic air contaminants. Toxic releases to air and water reflect emissions, but not exposures.

- Traffic Density is used as a surrogate for transportation-related emissions which, in turn, are a surrogate for exposure to transportation-related air pollutants. With respect to criteria pollutants, this indicator is somewhat duplicative of the ozone and $PM_{2.5}$ indicators. This indicator does not reflect exposures, which would be affected by both emission rates and dispersion. Exposures depend heavily on whether receptors are predominantly upwind or downwind of roadways. Furthermore, by excluding roads that are locally maintained, exposures to pollutants from many busy traffic intersections may be excluded.

- The method used to characterize Impaired Water Bodies provides a very poor measure of exposure or severity of conditions. For example, a stream that has concentrations of 10 pollutants slightly in excess of applicable water quality standards, and that passes through an unpopulated corner of a zip code area, would be scored 10 times higher than a single, large lake with one pollutant at levels 10 times the applicable water quality standard. This indicator, if retained, should be modified to reflect both the degree of contamination in excess of water quality standards, and the potential for exposure (using either surface area or proximity to potable water as a surrogate).

- The claim is made that educational attainment is an important independent predictor of health. However, no citation is provided to support this statement. It correlates with the poverty indicator, and does not appear to be independent.

Conclusion

The current draft CalEnviroScreen tool does not meet Cal/EPA's objectives. The numerical score presents an illusion of scientific validity that the tool does not merit. Cal/EPA's stated purposes for this tool would be better suited by a non-numeric matrix approach. Such a format would also reduce the likelihood of the tool being misapplied for purposes for which it was not intended.

The work that Cal/EPA has put into this tool has demonstrated the difficulty of creating a scientifically valid methodology for rank-ordering disadvantaged communities. Fortunately, there is no need for such a methodology. With relatively minor changes, the current CalEnviroScreen tool could be reconfigured to avoid most of the technical objections to the current tool, yet achieve Cal/EPA's goal of identifying high-priority communities.

Sincerely,

Gary Rubenstein
Senior Partner