EnviroScreen Comments

Scott Fruin [fruin@usc.edu]
Sent: Monday, June 02, 2014 5:00 PM
To: CalEnviroScreen, CalEnviroScreen@OEHHA

Dear Mr. Faust,

The changes and improvements in CalEnviroScreen from version 1.0 to 2.0 were considerable. Changing from zip codes to census tracts was a particularly large improvement.

One important area of remaining possible improvement concerns the use of rankings as multipliers. This should be avoided if possible.

If rankings are used as multipliers, it distorts importance differences in the relative changes in the underlying distributions of the variables of concern. For example, one common aspect of distributions of concentrations of environmental contamination is that the concentrations on the high end have an extended tail and are often lognormally distributed. If a lognormal distribution is converted to percentiles, from low to high, there is a large compression of the high end differences in concentration where health concerns are greatest. For example, there is a much larger difference in concentration going from the 90th percentile to the 99th (a 10% change in rank multiplier) than from the 50th to the 55th (also a 10% change in rank multiplier).

Furthermore, when using percentiles as multipliers, small changes in concentration on the low end have a disproportionate effect on the multiplier (regardless of distribution type) regardless of how small they are. For example, going from 10th percentile to 20th percentile doubles the multiplier even if the 10th and 20th percentile concentrations are low and are significantly less than a factor of two different.

If rankings must be used, using a range further away from 0 will help reduces these distortions, such as 100 to 200 or 1000 to 1100, etc.

I hope these examples illustrate my point. Please feel free to contact me if my suggestion is not clear or you would like any more detail.

Best regards,

Scott Fruin
Assistant Professor
Preventive Medicine
USC Keck School of Medicine