Estimating Median Household Income and DAC/SDAC Status for Community Water Systems: A case study of Los Angeles County

Mike Reibel, Department of Geography, Cal Poly Pomona and CSU Water Resources and Policy Initiatives

At the Water Resources and Underserved Communities Symposium, CSU Northridge, May 9, 2019 - mreibel@cpp.edu

Disadvantaged Communities

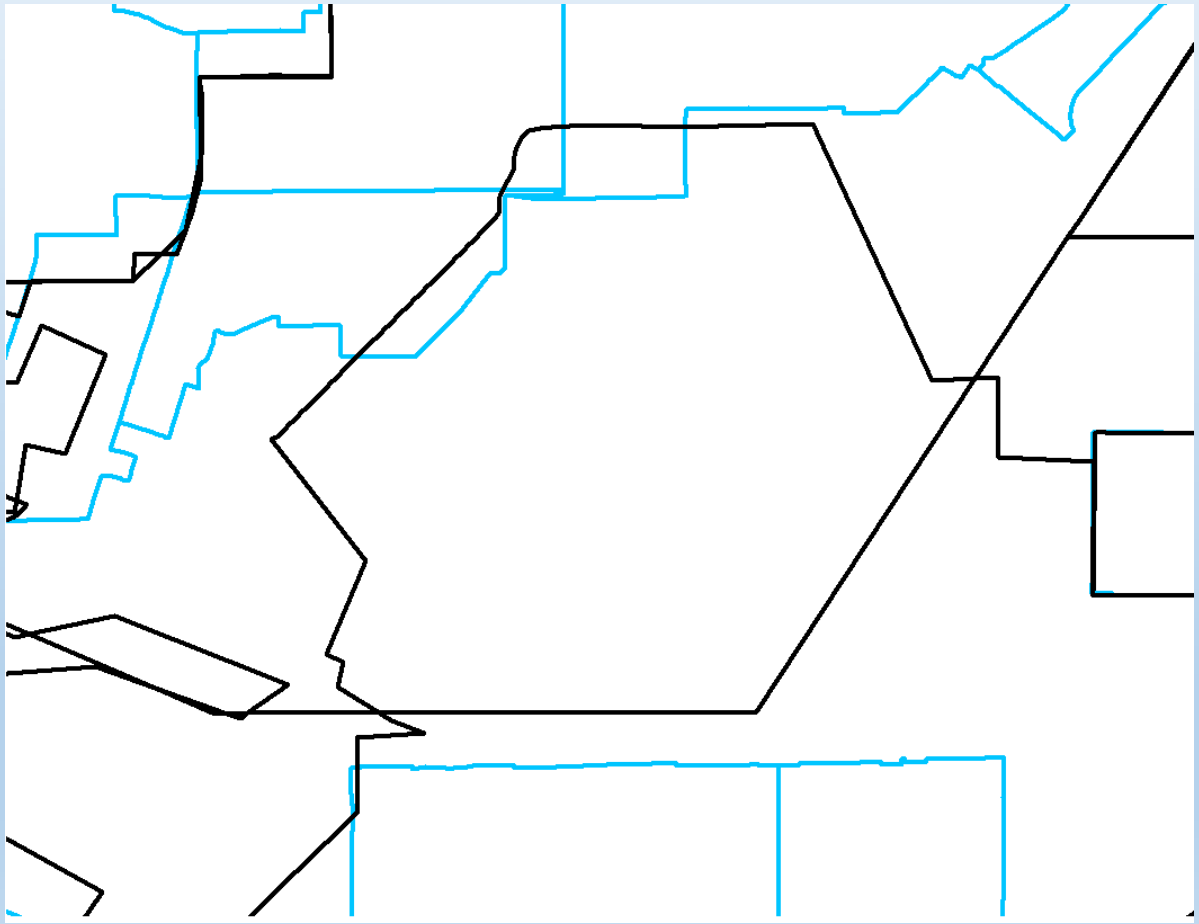
- Section 79505.5a of the Water Code defines disadvantaged communities (DACs) as “a community with an annual median household income that is less than 80 percent of the statewide annual median household income”
- The Health and Safety Code section 116760.20(n) defines a “severely disadvantaged community” (SDAC) as one with a median household income of less than 60 percent of the statewide average
- These definitions are invoked in Water Bond Propositions 1 and 84, which stipulate funding preference for such communities

The Overall Problem

- The median income thresholds should be evaluated for the entire water agency service area
- But income data for water agency service areas are not generally available, and census areas do not normally coincide neatly with water agency service areas
- **DAC determination therefore will always depend on estimates, and estimates are never perfect – all include errors**
- **Estimation is not merely guesswork, however: Some estimation procedures are more accurate than others**
- Accuracy is essential particularly in urban areas where incomes and MHI can vary block by block

The Main Obstacle to Accurate Income Estimates : Boundary Mismatch

Water Boundaries in Black, Census Tract Boundaries in Blue



The Customary Work-Around

- Local water agencies sometimes analyze their service area income levels indirectly and unscientifically, using the method of map overlay:

- "Eyeball" the prevalence of low income households:

  - Use GIS to superimpose the boundary of a service area on a map of census areas with median incomes below specified DAC or SDAC thresholds

  - Conclude from visual impression (not estimates) whether majority of service area corresponds to DAC or SDAC census areas

  - Visual impressions are not even estimates; can be very misleading

Somewhat Better Solutions: Areal Interpolation

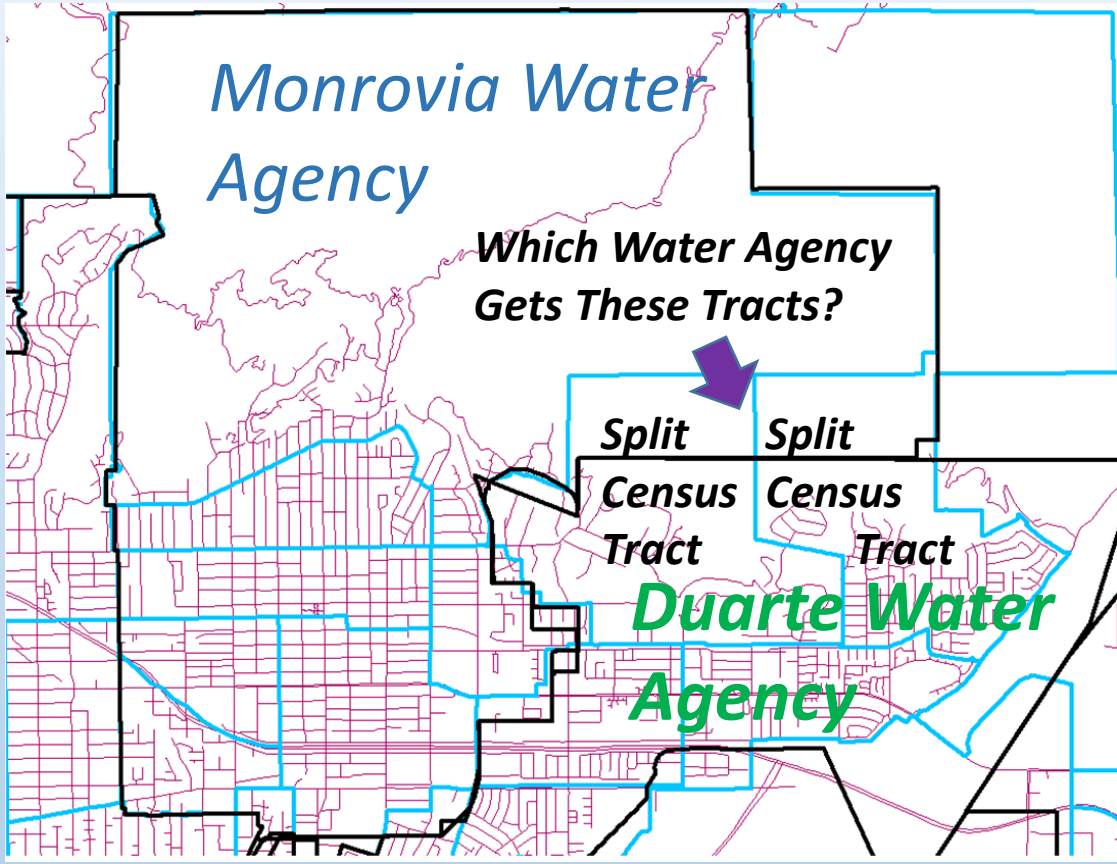
The re-aggregation of data from one set of boundaries (the census zones) to another set of boundaries (the water agencies)

Procedure: Chop up the census areas based on where the water agencies split them and reassign the households (with their incomes) to the water agencies, based on algorithms

Simple Areal Interpolation

Simplest areal interpolation of local medians:

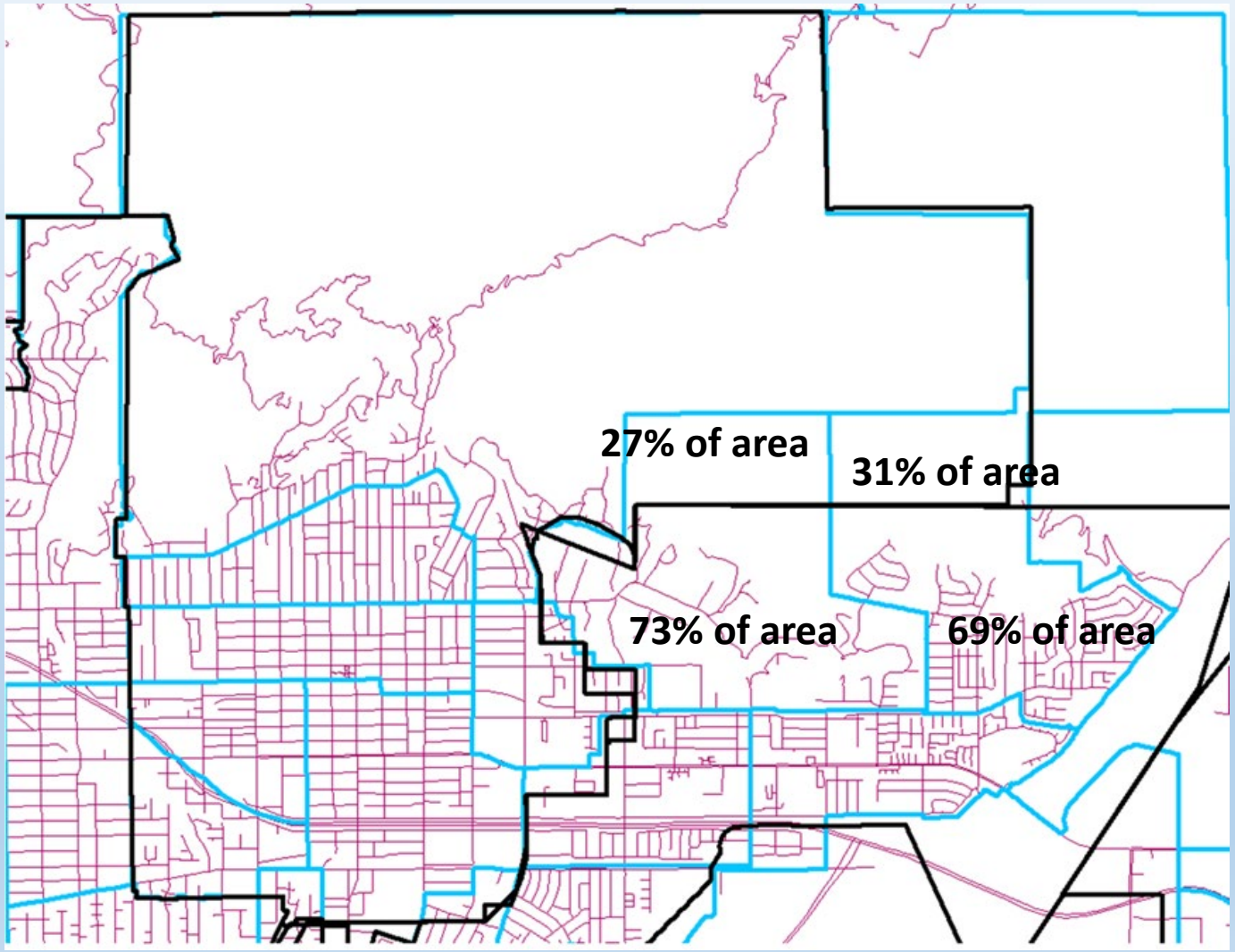
- Average the medians (MHI) for the tracts or block groups [source zones] that correspond to the service area
- Three error sources remain using this approach:
  - What to do with source zones split by two (or more) service areas? Include whole? For which service area? Both?
  - Ignores internal population density variations within source zones (which are hinted at in variations in street grid density in map at right)
  - Mean of medians is error-prone



Water Boundaries in Black, Census Tract Boundaries in Blue

Improved Areal Interpolation

- For partial (split) source zones, use the fraction inside water agency service area. Weight those zones' MHI by the size (area) of the fraction relative to the whole source zone.
- Still ignores internal population density variations within source zones



The State Water Resources Control Board Method

- The SWRCB is the agency that awards technical assistance and infrastructure funding from the water bonds and is responsible for determining if a water system area qualifies as DAC.
- Their method is to interpolate the means of source block group median household incomes twice:

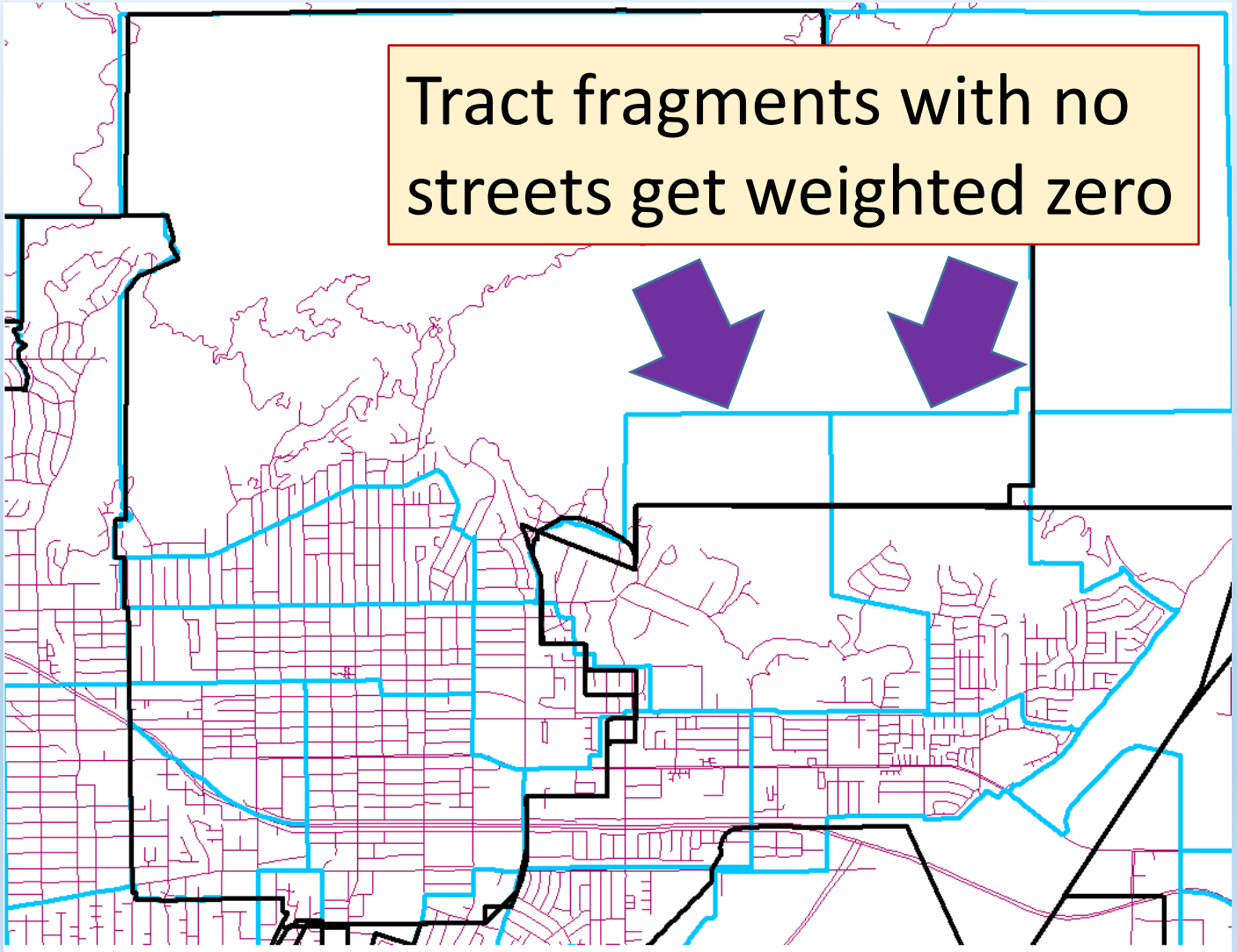
  1) Weighted by proportion of the source zone area inside the target zone (i.e. water system) service area, as in the last slide,

  2) Weighted by the source zone population as a proportion of the total target zone population (these will sum to greater than 100% in cases where there are split tracts)

- In cases where the two methods give conflicting answers re: DAC determination the responsible staffers consult with supervisors and make a decision.

Account for Population Density: The Street Weighted Interpolation Method

- Weight source zone fragments by fraction of the street grid, not by area
- Population only assigned to developed areas
- Reduces error



Another Problem with Most Approaches: *Averaging Averages*

- The problem with taking the mean over a set of medians is it creates more error, by generalizing from a generalization
- The median incomes of a set of block groups are already greatly simplified versions of the actual household incomes represented
- Taking the mean of local medians can result in estimates very different from the *overall median income across the whole service area* – which is what we actually want

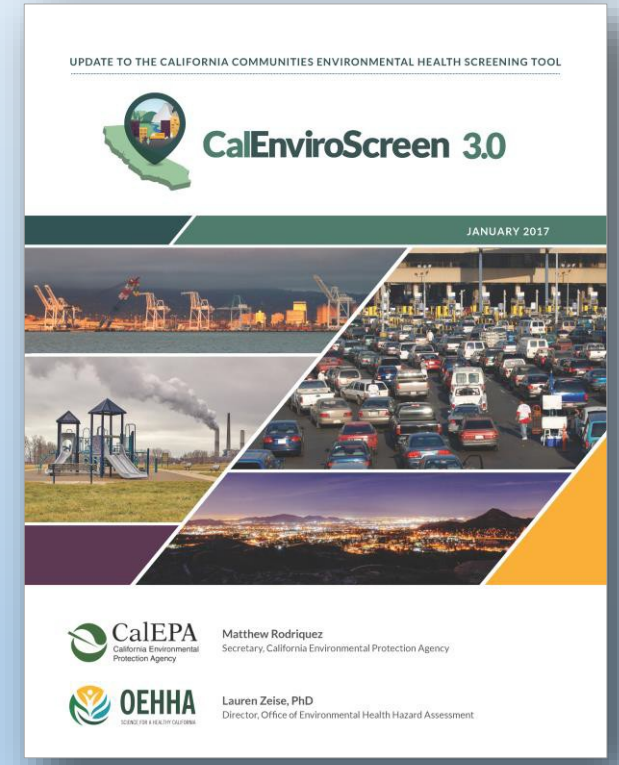
Area or Household	*Income (mean/median)*
Block Group 1	
Household a	**$20,000**
Household b	**$30,000**
Household c	**$36,000**
Household d	**$40,000**
BG 1 Median	**($33,000)**
Block Group 2	
Household e	**$41,000**
Household f	**$200,000**
Household g	**$225,000**
BG 2 Median	**($200,000)**
Mean of Medians 1 & 2	**($116,500)**
Overall Median	**($40,000)**

Next Problem: Census Income Interval Boundaries Don't Match DAC/SDAC Thresholds

- Based on the most recent census data available for small areas (from the 2016 five-year American Community Survey), the DAC threshold of 80% of state median income is $51,026 and the SDAC threshold is 60% of state median income = $38,270
- Those thresholds don't correspond to the census defined income interval boundaries of $50k-$59,999 and $35k-$39,999 for which estimated counts are reported at the level of tracts and block groups
- ***Solution:*** Fractionally assign counts in those intervals to above and below the thresholds based on linear distance from interval boundaries:
- Since $38,270 is 65% of the distance between $35k and $39,999, 65% of the households counted in that interval get put below the SDAC threshold.

CalEnviroScreen 3.0

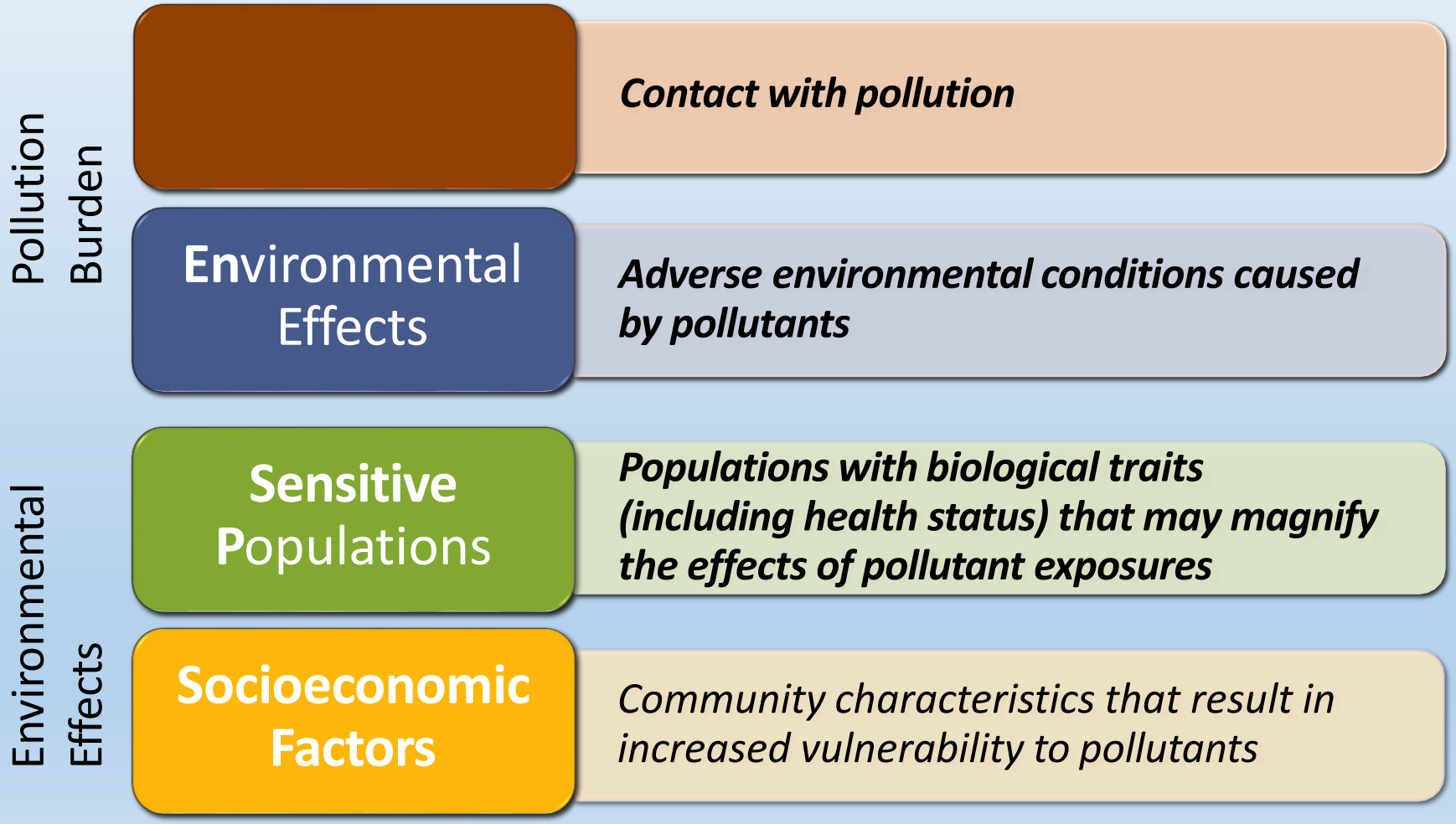
Released January 2017



- Analysis of relative burdens in California communities from pollution and population vulnerability
- 20 indicators combined into a single score
- Census tract scale

Available at: http://oehha.ca.gov/calenviroscreen/report/calenviroscreen-30

CalEnviroScreen Model



CalEnviroScreen 3.0 Indicators

Pollution Burden

Exposures

- Ozone
- PM2.5
- Diesel Particulate Matter
- Drinking Water Contaminants
- Toxic Releases from Facilities
- Traffic
- Pesticide Use

Environmental Effects

- Solid Waste Sites and Facilities
- Cleanup Sites
- Groundwater Threats
- Impaired Water Bodies
- Hazardous Waste Generators and Facilities

Population Characteristics

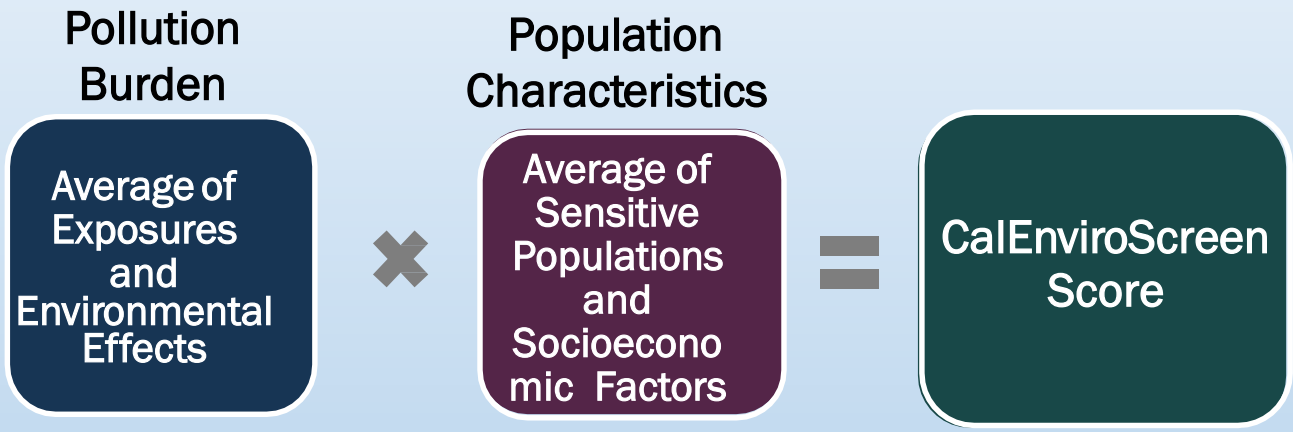
Sensitive Populations

- Asthma
- Cardiovascular Disease
- Low Birth Weight Infants

Socioeconomic Factors

- Educational Attainment
- Housing Burden
- Linguistic Isolation
- Poverty
- Unemployment

Calculating CalEnviroScreen Scores



- CalEnviroScreen score is calculated by combining all indicator Scores, expressed as percentiles; allows for comparison of different areas
- Higher scores mean greater pollution burdens and/or population vulnerability (not clear what the breakdown is)
- The highest 75-100th percentile (top 25%) represent "disadvantaged communities" under SB 535.

CES 3.0 is a complex environmental justice screening tool, not a straight measure of local median incomes

CalEPA, and hence CalEnviroScreen 3.0, define 'disadvantaged communities' very differently from the median income thresholds used by the State Water Resources Control Board and water laws:

According to SB 535 (de León, 2012) and AB 1550 (Gomez, 2016): "CalEPA shall identify 'disadvantaged communities' for investment opportunities based on geographic, socioeconomic, public health and environmental hazard criteria."

Disadvantaged communities for purposes of water subsidies are defined by local median incomes only.

Conclusions

We decided to use the street-weighted interpolation method to derive estimates of community water system median household income. No method of estimation is error-free, but this one is more accurate than other methods in use for three reasons:

1. The street weighting accounts for much of the internal population density variation within census zones, thereby reducing error

2. It interpolates whole income distributions, expressed as counts within intervals, to estimate the water agency service area median, rather than interpolating the medians of smaller areas. This avoids errors from averaging averages

3. It (thus) provides estimates of ***what percentage*** of households are below DAC thresholds in addition to answering the yes/no question whether the median is below the thresholds