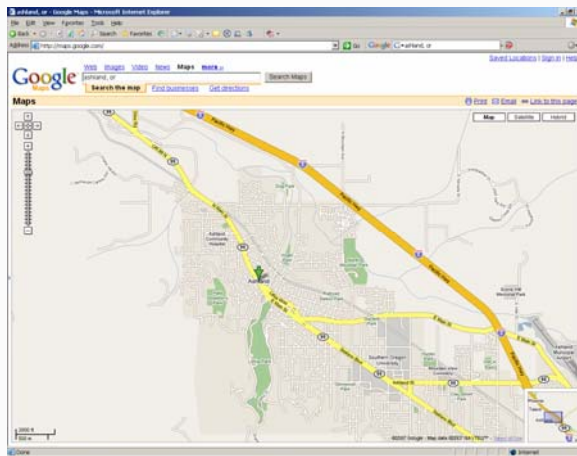
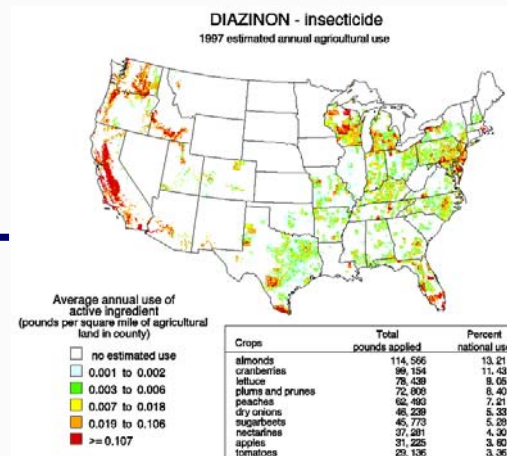


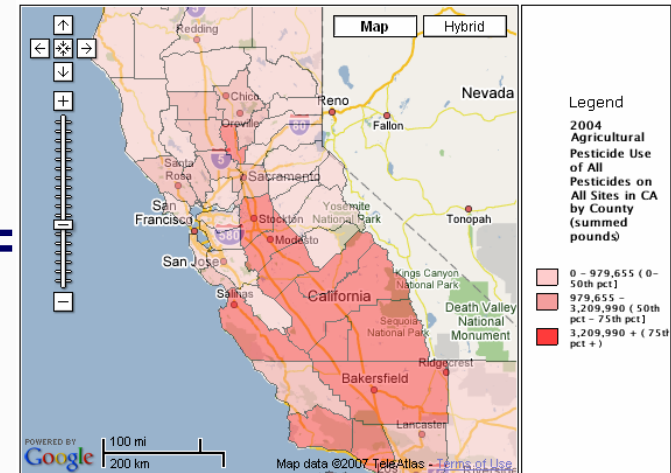
# California Environmental Health Tracking Program Tools: Take Google, Mix In Pesticides, and Voila – a Web-Mapping Service is Born!



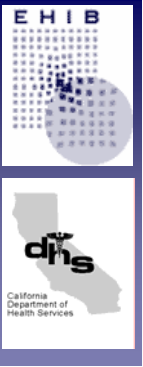
+



=



**Svetlana Smorodinsky**  
WREN 2007 – May 17-18  
Ashland, OR



# Acknowledgements

## CEHTP Team

**Craig Wolff**

Eddie Oh

Eric Roberts

Gala King

Helene Margolis

**Liang Guo**

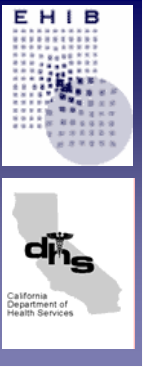
**Makinde Falade**

Michelle Wong

Paul English

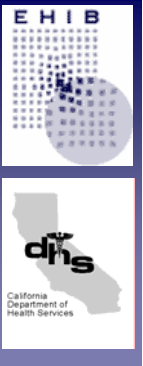
CALIFORNIA  
ENVIRONMENTAL  
HEALTH TRACKING  
PROGRAM

Thank you to CEHTP staff for letting me  
shamelessly steal slides, images, and text



# Topics covered

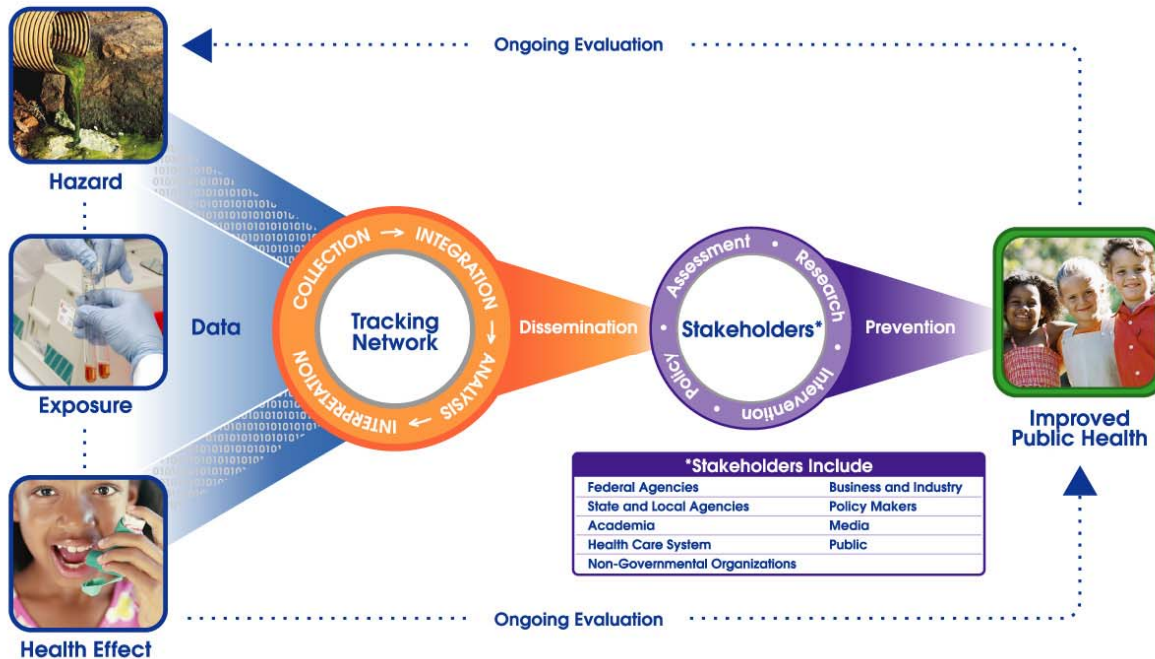
- Environmental Public Health Tracking Network overview
- Infrastructure and tools
- **Demo: geocoding, linkage (traffic and pesticides), pesticide web mapping**
- Health outcome surfaces (preterm birth)



# What's EPHT?

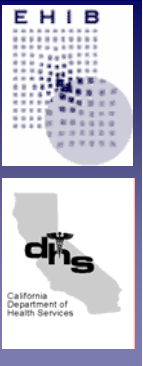
Environmental Public Health Tracking is the ongoing systematic collection, integration, analysis, interpretation, and dissemination of data about environmental hazards, exposure to environmental hazards, and health effects potentially related to exposure to environmental hazards.

## ENVIRONMENTAL PUBLIC HEALTH TRACKING



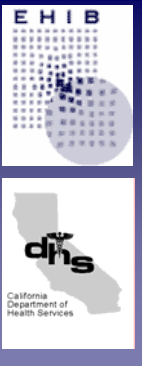
DEPARTMENT OF HEALTH AND HUMAN SERVICES  
CENTERS FOR DISEASE CONTROL AND PREVENTION  
SAFER • HEALTHIER • PEOPLE





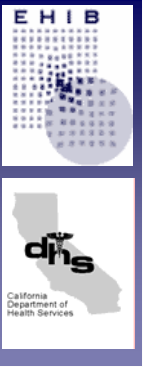
# What will EPHT do?

- EPHT Network will enable direct electronic data reporting and linkage of **health effect, exposure, and hazard data**
- EPHTN will support more effective information sharing and collaboration
  - Provide standards, data exchange mechanisms and tools
- Working definition of the “Tracking Network”:
  - An integrated set of data, standards and tools, developed by CDC, its grantees, and other EPHTN partners. This Network is designed to provide a “one stop” resource to improve participants’ capabilities to identify, understand, track, and respond to environmental public health disease and exposures at the local, state, and national level.



# California Environmental Health Tracking Program

- Contribute to standard-setting with CDC
- Security – managing access to data (role-based access)
- Metadata – managing descriptions of data
- Spatial enhancement – managing the geography of data (e.g. geocoding)
- Integration – combining datasets intelligently (linkage)
- Data exchange – moving data securely
- Visualization and dissemination – provide stakeholders useful tables, maps, charts
- Deploy network gateway – manage access by machines to CEHTP data and services



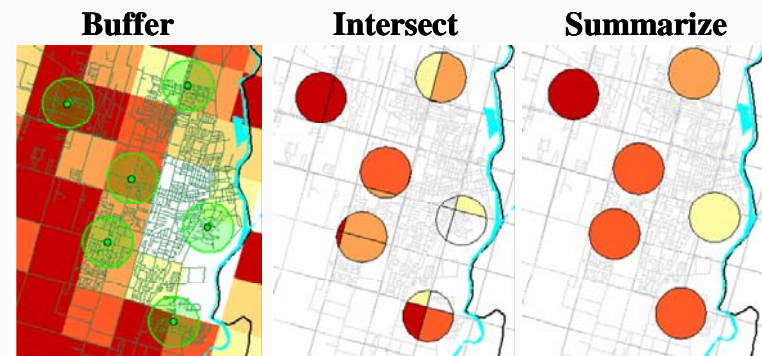
# CEHTP IT/GIS Goal in a Nutshell

- Provide web-based services to public and government agencies to streamline ongoing and systematic environmental health functions
- Focus on simplicity through electronic standards for IT/GIS integration and interoperability

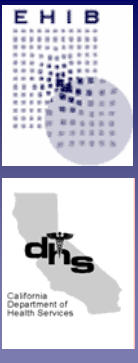


# CEHTP Services and Tools

- Data/System Enhancement
  - Geocoding (yeah!)
  - Geographic feature editing (water systems)
- Record-Level Integration
  - Spatiotemporal linkage (traffic, pesticides)
- Visualization and Dissemination
  - Internet mapping (pesticides, birth surfaces)
  - Google mash-up







# Geocoding Tool

## CEHTP Centralized Geocoding Demonstration

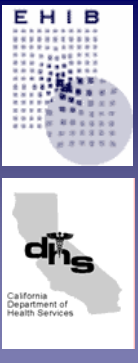
### Background

### Enter Address

### Change Options

### Geocoding Results

- This web page is an example of a client that utilizes the California Environmental Health Tracking Program's (CEHTP) centralized geocoding service.
- Problem
  - Historically, health event data have imperfect geocoding success rates. e.g. Birth Master File = 95%, Medi-Cal = 85%.  
**We want 100%!**
  - High quality address data, or, more preferably, pre-geocoded data, prepares CEHTP to perform ongoing/systematic/automated GIS visualization and spatial linkage with environmental hazard datasets. Geocoding should be done as early as possible in the data collection process.
  - Programs that receive reported data and compile them for official release might not have resources or mandate for geocoding or ensuring high quality address data. Geocoding can be expensive.
  - Programs may desire GIS visualization services for their data product. This is not possible unless the geography of the data is coded and of high quality.
  - Click [here](#) for a workshop presentation on centralized geocoding requirements and implementation.
- [Solution](#)
- [Benefits](#)
- [Advocacy/New Data](#)
- [Client Examples](#)
- [Request Service](#)



# Spatial Linkage Tools

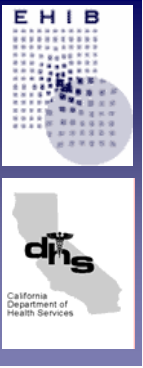
## CEHTP Spatial Linkage Demonstration

### Background

### Enter Buffer Parameters

### Spatial Linkage Results

- This web page is an example of a client that utilizes the California Environmental Health Tracking Program's (CEHTP) spatial linkage web service. This example computes traffic-related metrics on CalTrans Highway Performance Monitoring System (HPMS) 2004 data in California. All geospatial operations are performed on the server-side, while contact between the client (web server) and the GIS server is limited to Simple Object Access Protocol (SOAP) input/output parameters.
- Problem
  - CEHTP partners would like to perform spatial analyses on environmental and health datasets, but often lack the GIS expertise, software/hardware resources, and/or easy/ongoing access to commonly used datasets.
  - Environmental health geospatial analyses have historically been accomplished on a GIS technician's desktop. The vision of the National Environmental Health Tracking Network calls for the creation of standards-based tools that allow clients to perform systematic integration of environmental and health datasets. This means that typical desktop geospatial operations must be ported to an enterprise context.
  - For CEHTP stakeholders who maintain the capacity to perform geospatial analyses in-house, it is commonly necessary to maintain full local copies of spatial-enabled datasets to perform a simple overlay operation. Even when standard electronic data exchange tools are available, desktop geospatial analysis requires the transfer of entire datasets to a GIS technician's local domain. This requirement can become cumbersome for datasets that are updated frequently.
  - Have you ever asked yourself this question when you're at an Internet mapping site: "Great, now how do I analyze this great GIS data relative to my own geospatial data without having to do a huge download and having to learn how to use GIS software?"
- [Solution](#)
- [Benefits](#)
- [Request Service](#)



**To view the geocoding and spatial linkage (traffic) tools,  
visit <http://www.ehib.org>,  
click on Tools**

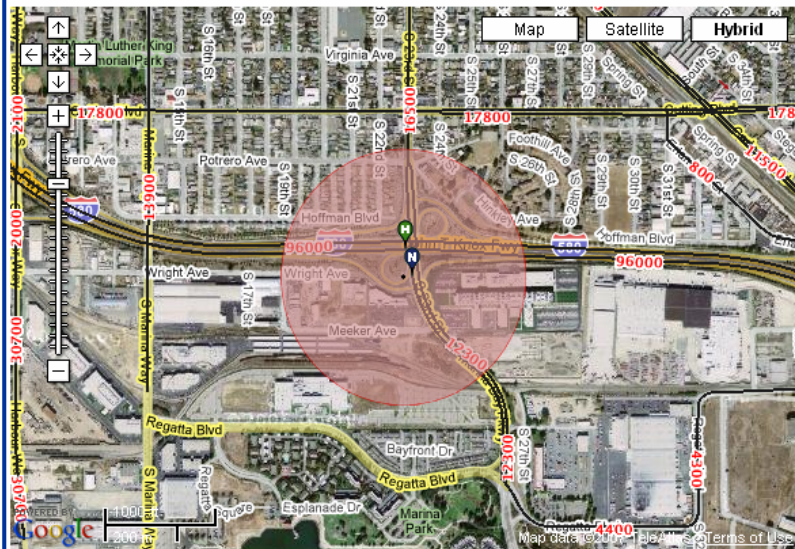
## CEHTP Spatial Linkage Demonstration

Background Enter Buffer Parameters Spatial Linkage Results

Metric	Value
(hdr) Direction to highest segment within buffer (degrees ccw** from east)	88
(ndist) Distance to nearest segment within buffer (meters)	25
(ng) Gauss-adjusted traffic volume of nearest segment within buffer (vehicles/day*)	11,286
(hg) Highest Gauss-adjusted traffic volume segment within buffer (vehicles/day*)	42,210
(sg) Sum of all Gauss-adjusted traffic volumes within buffer (vehicles/day*)	60,352
(ndir) Direction to nearest segment within buffer (degrees ccw** from east)	5
(hdist) Distance to highest segment within buffer (meters)	78
(nu) Unadjusted traffic volume of nearest segment within buffer (vehicles/day*)	12,300
(hu) Unadjusted traffic volume of highest segment within buffer (vehicles/day*)	96,000
(su) Sum of all unadjusted traffic volumes within buffer (vehicles/day*)	124,800

\*average annual daily traffic

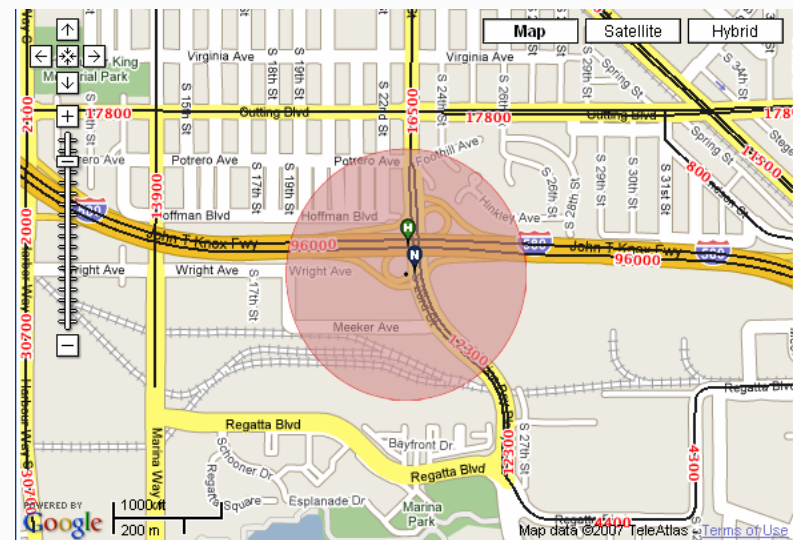
\*\*ccw=counter clockwise



N=nearest segment to buffer center H=segment with highest traffic volume in buffer

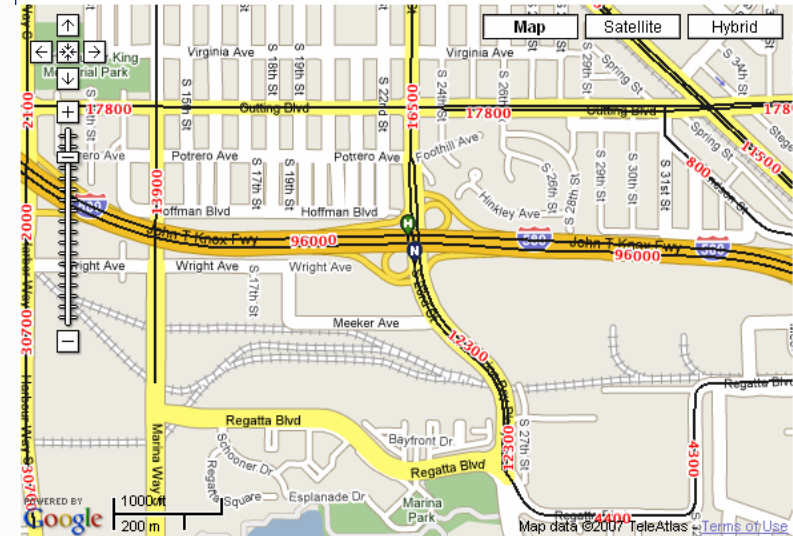
[Hide Buffer](#)

# Traffic Tool



N=nearest segment to buffer center H=segment with highest traffic volume in buffer

[Hide Buffer](#)

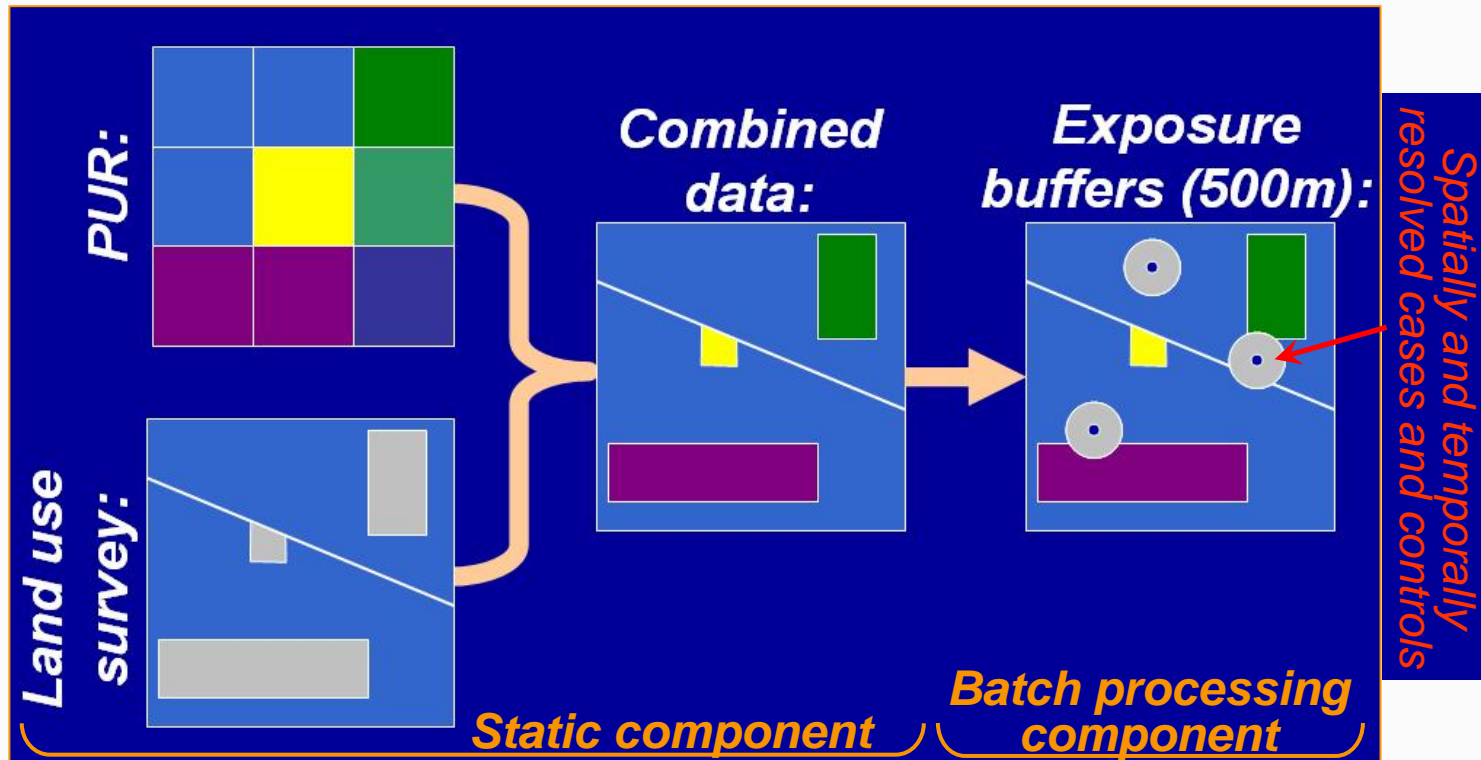


N=nearest segment to buffer center H=segment with highest traffic volume in buffer

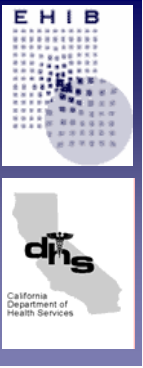
[Show Buffer](#)

# Pesticide Tool

I'm buying a dream house on a golf course in Brentwood, Contra Costa County, California. I heard there are many farms in the area, so do I have to worry about pesticides?



*Spatially and temporally resolved cases and controls*



**Pesticide spatial linkage tool will soon be available at**  
**<http://www.ehib.org>**



# Pesticide Tool

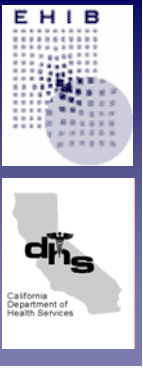
## Agricultural Pesticide Web Mapping Service (WMS)

**Background** **Service Options** **View Map**

This page demonstrates many of the capabilities of the California Environmental Health Tracking Program (CEHTP) Agricultural Pesticide Use Web Map Service (WMS). This implementation is a custom OpenGeospatial Consortium WMS. Originating from the California Department of Pesticide Regulation, County Agricultural Commissioners, Agricultural Growers, and Pesticide Operators, the Pesticide Use Report data has very detailed information across space and time. We attempted to summarize this information flexibly, but without making the interface too complex. Some of the options include:

- Multiple years of reporting
- Various measures of pesticide application intensity
- Various geographic units in which data is displayed
- Maps of individual pesticide chemicals or groups of chemicals
- Maps of individual application sites (crops) or groups of sites
- Statewide maps or maps by individual counties

In the coming months, we will provide metadata, service capabilities, and other documentation for utilizing this service at external websites. In the meantime, please provide us your suggestions or comments by sending an email to [Craig Wolff](#)



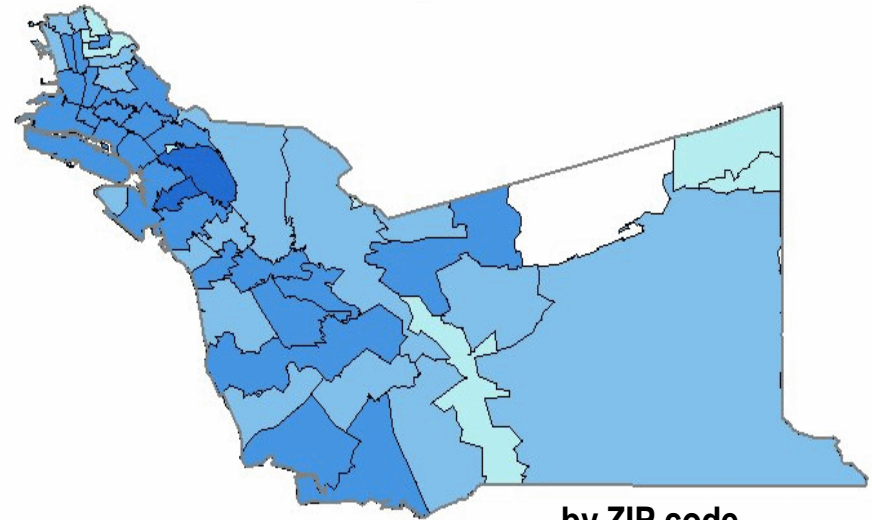
**To view the pesticide mapping tool,  
visit <http://www.ehib.org>,  
click on Tools**



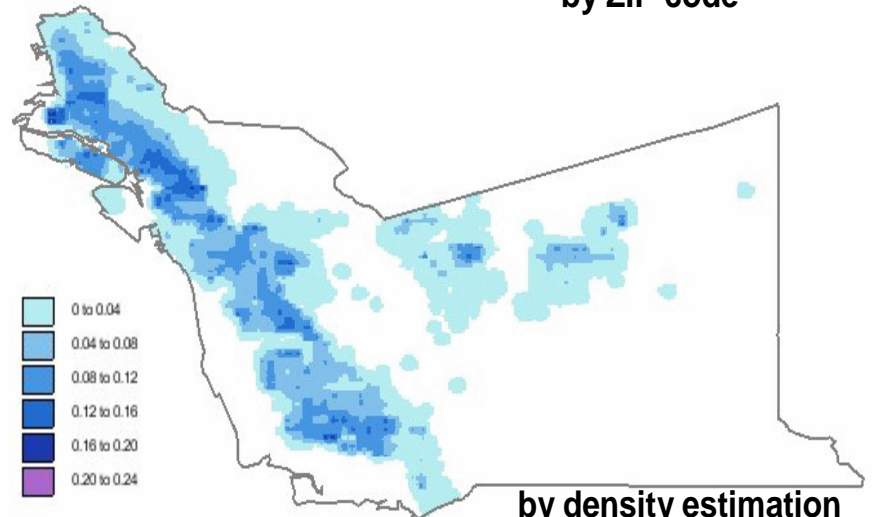
# Analysis, Reporting, Visualization

- Traditional: **Zip code level**
  - Administrative boundaries
  - Capture areas without any population
- Newer, better: **Density estimation**
  - Neighborhood-level rates now visible
  - Statistical significance
  - Visualization not influenced by artificial boundaries
  - Maintain confidentiality

Pre-term birth rate, Alameda County, 2001



by ZIP code

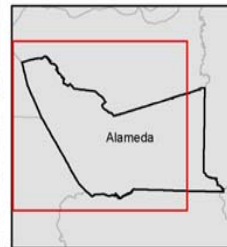
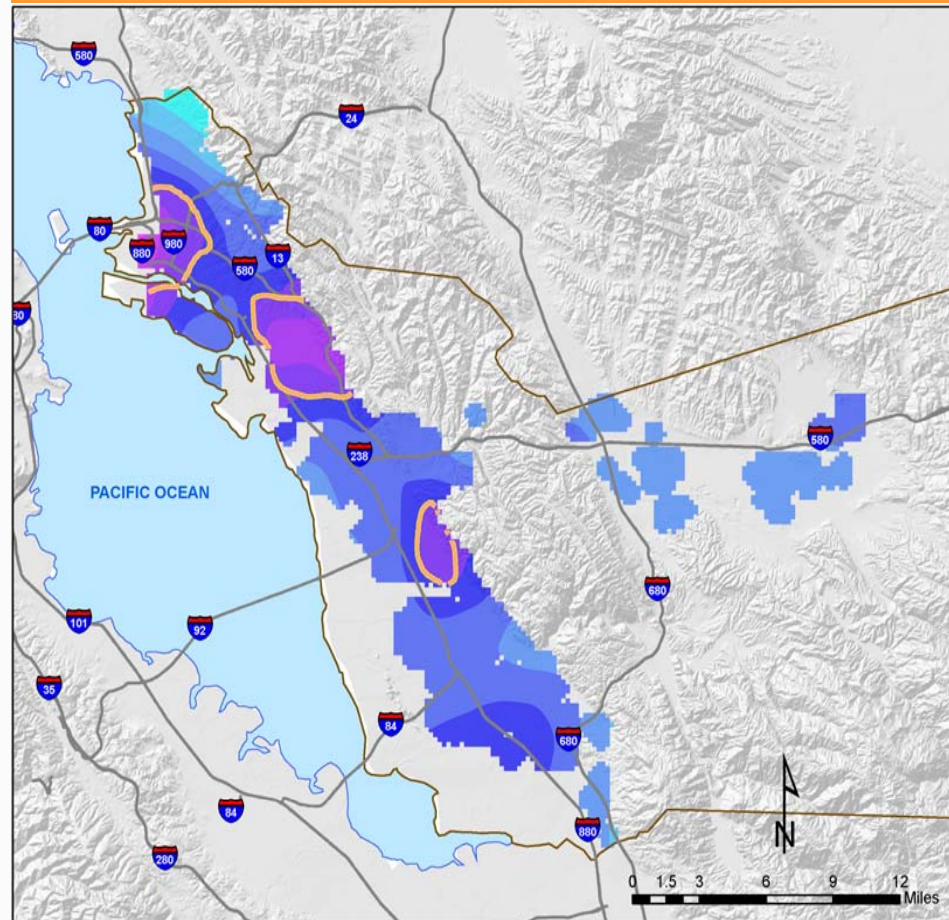


by density estimation

# Analysis, Reporting, Visualization

- New, even better:  
**Smoothed maps**
  - Probability function
  - Regression equation
  - Surface controlled for covariates
  - Neighborhood rates visible
  - Not influenced by artificial boundaries
  - Maintain confidentiality

## Preterm birth risk in Alameda County, 2001 (Loess smooth)

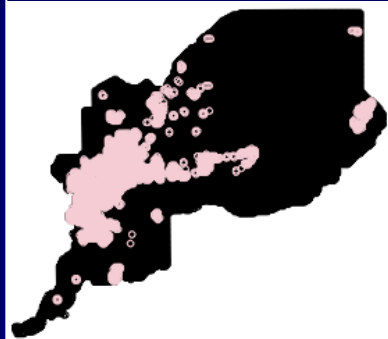


**DRAFT**

# Analysis, Reporting, Visualization

## How?

Geocode records;  
generate  
array of  
evenly-  
spaced points  
where there  
are data



Calculate for  
each point:

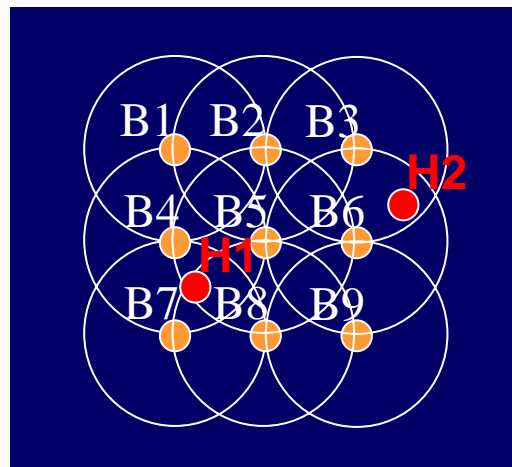
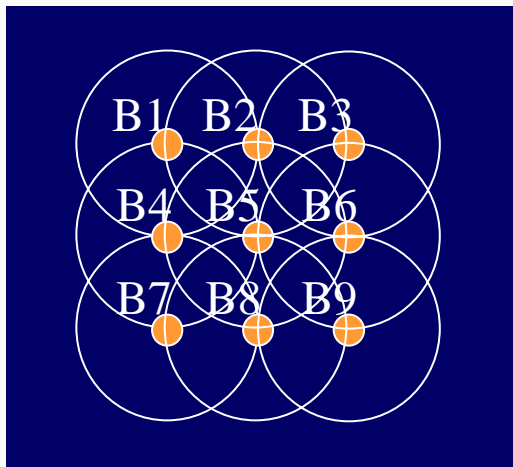
- *Probability  
function*
- *Statistical  
significance*

Interpolate  
between  
points;  
generate  
raster  
surfaces and  
contours

Integrate into  
Web Mapping  
Service

# Analysis, Reporting, Visualization

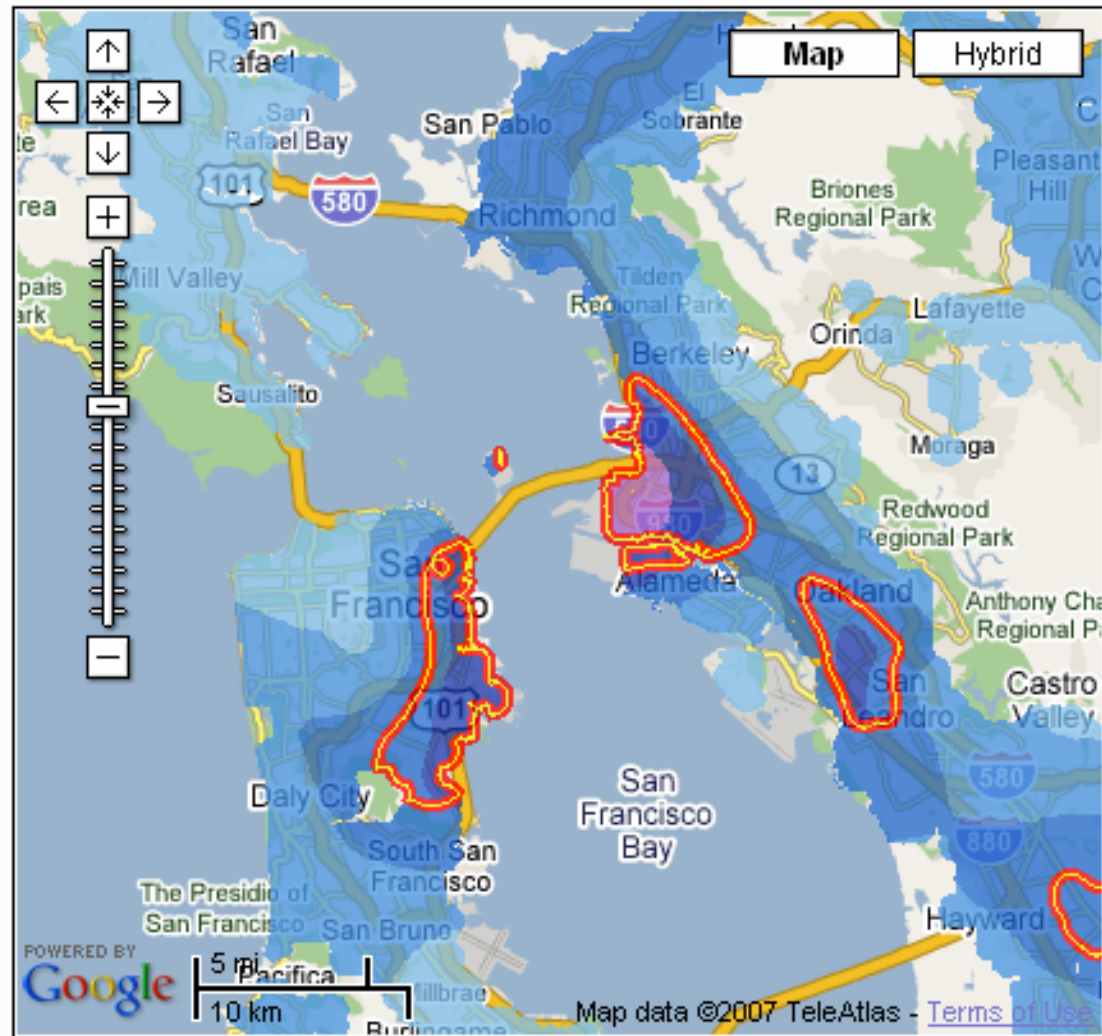
- Buffered grid points
  - Determine starting and ending X/Y of domain in projection that preserves distance and direction (e.g. UTM)
  - Choose spacing for grid points
  - Choose buffer radius
- Intersect buffers with health points
- Create correspondence table describing intersected grid point buffers with health events



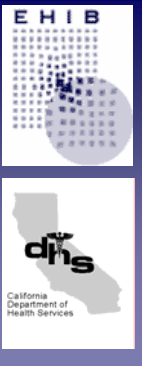
Health ID	Buffer ID
H1	B5
H1	B4
H1	B7
H1	B8
H2	B3
H2	B6

# Analysis, Reporting, Visualization

- Interpolate raster surface
  - Interpolate rate and p-values surfaces against attributed grid points
- Clip the surfaces using mask polygon
  - Get rid of points in water, no population
- Throw into Google
  - Literally (BARGC Google Project)







# To Sum Up



- CEHTP is building an infrastructure that enables stakeholders (government, community, and others) to
  - Enhance the quality of existing data
  - Enhance the ability to analyze data
  - Generate useful and relevant environmental health information, products, and services
- Data enhancement services
  - Automated processes such as geocoding
- Data integration services
  - Linkage and analysis of health indicators and geographically-related exposures to environmental hazards
- Data visualization and dissemination services
  - Creation of dynamic, custom materials in the forms of maps, charts, and reports

*Developing a statewide preterm birth surface as demonstration*