FINAL REPORT: CONTRACT NUMBER # 04-308

Air Pollution and Environmental Justice: Integrating Indicators of Cumulative Impact and Socio-Economic Vulnerability into Regulatory Decision-Making

Principal Investigators:

Manuel Pastor Jr., Ph.D.
University of Southern California
Program on Environmental and Regional Equity
Departments of Geography and American Studies and Ethnicity

Rachel Morello-Frosch, Ph.D., M.P.H.
University of California, Berkeley
Department of Environmental Science Policy and Management &
School of Public Health

James Sadd, Ph.D.
Occidental College
Department of Geology and Environmental Science

Prepared by:

Manuel Pastor
Rachel Morello-Frosch
James Sadd

Contact: Manuel Pastor
University of Southern California
Department of Geography
3620 S. Vermont Ave, KAP-462
Los Angeles, CA 90089-0255
Phone: (213) 740-5604
Fax: (213) 740-0056
mpastor@usc.edu

Disclaimer

The statements and conclusions in this Report are those of the contractor and not necessarily those of the California Air Resources Board. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

Acknowledgements

Critical to this project were the people helping our team with grant management, including Sally Jorgenson at CARB, Patti Caton and Betsy Barlow from Brown University, Raul Torres at UC Berkeley, Jeri Muniz, George-Ann Cleary, and Kathleen Speer from the University of Southern California, Susan Welch, Suzanne Ziegler, Kim Durham, Sue Kato, and Leah-Ann Williams from the University of California, Santa Cruz. This project was logistically complicated by both the move of two PIs and the occasional financial crises that froze, then unfroze, then refroze, then defrosted CARB spending; our grants administrators kept us going through it all (or, perhaps, despite it all).

Finally, we want to thank the community advocates for environmental justice, too many to enumerate, who have worked with us on this project. They came to meetings where we previewed methodological approaches and study results, they offered excellent feedback to improve our screening method, and they collaborated on a "ground truthing" effort in Oakland. But perhaps most important is that they have kept our collective eye focused on a key prize: making the science real for communities and decision makers and ensuring environmental health for all Californians. It is a goal that our research team and CARB share with the advocates and we hope that this research will provide a firm scientific foundation for future policy, research and regulatory advancements in this arena.

Table of Contents

List of Tables

<all tables are at the end of the respective chapters>

Table 4.1: Summary of cumulative impact and vulnerability indicators used in the EJ Screening Method, with geographic unit of analysis and data source.

Table 6.1: Community-generated lists of Hazards and Sensitive Receptors

List of Figures

<all figures are at the end of the respective chapters>

Figure 2.1: Locations of Active TRI Facilities Relative to Neighborhood Demographics by Census Tract

Figure 2.2: Population by Race/Ethnicity and Proximity to an Active TRI

Figure 2.3: Percentage of Households within 1 Mile of an Active TRI by Income and Race/Ethnicity

Figure 2.4: 1999 NATA Estimated Excess Cancer Risk (All Sources) by 2000 Census Tracts

Figure 3.1a: Difference in birth weight in grams associated with full pregnancy gaseous pollutant exposures for births within 10 km monitor distance, single and two-pollutant linear models (95% confidence interval)

Figure 3.1b: Difference in birth weight in grams associated with full pregnancy particulate matter exposures for births within 10 km monitor distance, single and two-pollutant linear models (95% confidence interval)

Figure 3.2a: Odds ratio for pre-term birth (29-34 vs. 39-44 weeks) associated with full pregnancy gaseous pollutant exposures for births within 10 km monitor distance, single and two-pollutant linear models (95% confidence interval)

Figure 3.2b: Odds ratio for pre-term birth (29-34 vs. 39-44 weeks) associated with full pregnancy particulate matter exposures for births within 10 km monitor distance, single and two-pollutant linear models (95% confidence interval)

Figure 4.1: Summary of process used to create CI Polygons

Figure 4.2: Map of a portion of the study area showing CI Polygons in white, and areas not scored (including open space, vacant land, industrial land use, etc.) in gray

Figure 4.3: Map of the Health Risk and Exposure category score for the greater Los Angeles urban area

Figure 4.4: Map of the Social Vulnerability category score for the greater Los Angeles urban area

Figure 4.5: Hazard proximity to CI Polygons defined using single buffers (the number of point hazards and land use hazards located within each buffer is summed for each CI Polygon)

Figure 4.6: Map of the greater Los Angeles area showing the distribution of hazard count by CI Polygon (1000 foot buffers, categories determined via Jenks natural breaks)

Figure 4.7: Land use polygons and GIS point hazard features for a portion of the city of Maywood

Figure 4.8: Aerial imagery of the map shown in Figure 4.7 (note the actual polygonal areas of the three air quality hazards shown as points in Figure 4.7).

Figure 4.9: Land use polygons for a portion of the city of Maywood, showing GIS polygons identifying schools, as well as buffered point locations of a school, two childcare facilities, and a health clinic

Figure 4.10: Aerial imagery of the map shown in Figure 4.9

Figure 4.11: Hazard proximity to CI Polygons defined using multiple buffers at 1000 foot intervals

Figure 4.12: Map of the greater Los Angeles area showing the distribution of distance weighted hazard and sensitive land use count by CI Polygon

Figure 4.13: Map of the greater Los Angeles area showing the distribution of distance weighted hazard and sensitive land use count at the tract level (population weighed with categories derived via Jenks natural breaks)

Figure 4.15: Map of the Total Cumulative Impacts score for the South Los Angeles area

Figure 4.16: Map of the Total Cumulative Impacts score for the Inland Empire

Figure 4.14: Map of the Total Cumulative Impacts score for the greater Los Angeles urban area

Figure 5.1: CEC "Demographic Screen" map of census block groups overlain by the location of the proposed Nueva Azalea site and one and six mile buffer

Figure 5.2: EJSM map of cumulative impact score overlain by the location of the proposed Nueva Azalea site, and one and six mile buffer radii

Figure 5.3: Rank of average population-weighted CI scores for census block groups located within a one mile circular buffer surrounding all southern California power plants

Figure 5.4: Rank of average population-weighted CI scores for census block groups located within a six mile circular buffer surrounding all southern California power plants.

Figure 5.5: Rank of percent minority residents for census block groups located within a one mile circular buffer surrounding all southern California power plants.

Figure 5.6: Rank of percent minority residents for census block groups located within a six mile circular buffer surrounding all southern California power plants.

Figure 5.7: Rank of percent residents below the federal poverty level for census block groups located within a one mile circular buffer surrounding all southern California power plants.

Figure 5.8: Rank of percent residents below the federal poverty level for census block groups located within a six mile circular buffer surrounding all southern California power plants.

Figure 6.1: Location of the Hegenberger Corridor study area in east Oakland, bounded by Hegenberger Expressway, International Blvd., San Leandro St. and 92nd Avenue (two views)

Figure 6.2: Results of Hegenberger Corridor microstudy mapping of air quality hazards and sensitive receptors by community residents. (Compare to Figure 6.3)

Figure 6.3: Map of Hegenberger Corridor study area showing locations of air quality hazards and sensitive land uses using data from State regulatory database facilities

Figure 6.4: Comparison of spatial data from Figures 6.2 and 6.3 to emphasize additional air quality hazards and sensitive receptors recorded by Hegenberger community members

Figure 6.5: Location of businesses where auto paint and body work was observed during Hegenberger community field observations, as compared to automobile businesses recorded in the AB2588 "Hot Spots" database provided by CARB

Figure 6.6: Locations of observations of illegal idling by diesel trucks during Hegenberger community field observations.

Figure 6.7: Buffers surrounding air quality hazards from State government agency databases capture twelve sensitive receptors mapped by community residents during the Hegenberger microstudy

Figure 6.8: Reported locations of AB2588 "hot spots" air toxics emissions sources (red squares) compared with actual locations verified by ground truthing (green circles)

Figure 6.9: PM2.5 measurements from 5 sites, CBE East Oakland Phase II, October 7-November 8, 2008

Figure 6.10: PM2.5 weekday measurements from 5 sites, CBE East Oakland Phase II, October 7-November 8, 2008

Figure A.1: View of ArcMap toolbar with tool button and pop-up form with buffer distance specified

Figure A.2: View of ArcMap data view with buffer and selected features

Abstract

This project sought to develop diverse methodological approaches to address environmental justice (EJ) concerns of relevance to air pollution regulation in California. The project consisted of several interlocking research projects, including: 1) a landscape analysis of environmental hazard and air pollution burden disparities in the Bay Area; 2) development of an Environmental Justice Screening Method (EJSM) to identify areas of environmental justice concern with regard to the cumulative impacts of hazard proximity, air pollution, exposure and estimated health risk, and social vulnerability; 3) the implementation of the EJSM to evaluate a hypothetical siting of a power plant facility; 4) a statewide analysis of the association between ambient pollution exposures and adverse perinatal outcomes; 5) the implementation of a community-based participatory ground truthing research project to evaluate the coverage of emissions inventory databases of localized emission sources and sensitive receptors and to build community confidence in the research and regulatory process. Overall, our study results indicate that environmental disparities do exist, even when controls are introduced for spatial dependence. The analysis also suggests that linguistic isolation matters, a variable that may be of special interest for assessing community capacity for civic engagement in the regulatory process. We also found that ambient air pollution is associated with lower birth weight and preterm birth. In our analysis we also assessed for effect modification – that is, we examined whether the relationship between air pollution exposures and poor perinatal outcomes were amplified for certain racial/ethnic groups or by the level of neighborhood poverty. We found that effect modification by area and individual-level measures of race and socio-economic status (SES) were not consistent in terms of changes in effect estimates and statistical significance. We also developed a robust and flexible screening method that offers a scientifically valid and transparent way to examine and rank neighborhoods within regions based on EJ concerns. This method also performs well when used to evaluate potential EJ concerns related to siting decisions which may be useful for future regulatory assessments of the likelihood of community reaction to disproportionate burdens. The method is a screening tool to guide decision-making, not for risk assessment; as the community-based participatory component of this project demonstrates, secondary databases and emissions inventories do not capture the full scope of potentially hazardous emission sources, sensitive land uses, or air quality problems on a localized scale. Instead the EJSM can provide an important first step to guide decision-making regarding further research, community outreach, and regulatory strategies to better address environmental justice concerns related to air pollution impacts across diverse communities in California.

Executive Summary

Background:

Researchers, policy makers and community advocates have been interested in developing a screening method that can identify areas of special concern for the cumulative impacts of environmental and non-environmental stressors in California. Because of previous studies that have shown a pattern of racial and income disparities in air pollution exposures and health risks in the state (Sadd, Pastor et al. 1999; Morello-Frosch, Pastor et al. 2002; Pastor, Sadd et al. 2004; Su, Morello-Frosch et al. 2009), such a method could be useful in implementing environmental justice mandates and could also advance the emerging field of research about the intersection of potential cumulative hazard exposure and social vulnerability.

Methods:

To work toward developing such a screening method, this effort consisted of a set of interlocking and interdependent research efforts, including:

1) A baseline analysis of environmental justice conditions in one part of the state not previously the subject of systematic spatial analysis and multivariate statistical modeling, with an eye toward identifying key determinants associated with patterns of racial and ethnic disparities in lifetime cancer risk and respiratory hazard associated with outdoor air toxics exposure. Innovative spatial techniques were used to associate tract-level demographics with facility locations and spatial autocorrelation controls were used and compared with simple ordinary least square (OLS) strategies. Community members were involved in reviewing the early research as part of the PIs' commitment to incorporating community feedback on the research as it evolved.

2) An analysis of the association between ambient criteria air pollution and adverse perinatal outcomes, utilizing multi-level analysis to assess both individual-level SES variables and neighborhood- level SES variables simultaneously and examining whether they may potentially modify observed associations between air pollution exposures and adverse birth outcomes. We specifically investigate effects for full gestational exposures as well as trimester-specific effects using data collected from air pollution monitors for particulates ($PM_{2.5}$, PM_{10} and coarse PM) as well as CO, NO2, SO_2 and ozone.

3) The development of an Environmental Justice Screening Method (EJSM) utilizing measures of hazard proximity and sensitive land uses, cumulative impact from potential air pollution exposures and estimated cancer and respiratory risks, and indicators of social vulnerability, all at the census tract level, to calculate relative scores for high priority areas. Innovations in the process included the creation of a land use layer for Southern California that focuses on land uses identified as sensitive according to ARB guidelines (CARB 2005), the intersection of this layer with block- level demographics data to drive attention to the finest geographic unit possible, and the creation of a simple distance-decay and population-weighting procedure that generates a tract-level score on hazard location that can then be combined with simple quintile rankings on air pollution

hazards and social vulnerability measures. One specific driver in the process was a requirement that we develop a method that was transparent and accessible to a diverse audience, including regulatory decision-makers and communities. As a result, the EJSM was periodically reviewed throughout its development and evolution by scientific colleagues from an external peer review committee as well as community advocates throughout the state.

4) The use of the EJSM to evaluate the environmental justice dimensions of the hypothetical siting of a power plant and the comparison of those results with the approach to environmental justice screening usually applied by the California Energy Commission (CEC) in a siting process. We specifically chose the Nueva Azalea plant, a project that was proposed for South Gate, California but was never sited due to community concerns that eventually led the company to withdraw its application for the site. To conduct this analysis, we scored the locations of all existing power plants using our EJSM in order to determine a range of feasible sites.

5) The development and implementation, in conjunction with community-based partners, of a "ground truthing" microstudy in the Hegenberger Corridor of Oakland. The project consisted of conducting a local-scale validation study to identify and locate emissions sources that may not be systematically incorporated in state regulatory agency data inventories and measurements of community-level air pollution burdens; the PIs worked with a community organization to train local residents in the ground truthing protocols, and they eventually worked together to conduct some pilot particulate air monitoring as well. Aside from augmenting and validating information from ARB emissions inventories regarding the existence and location of area and point emission sources and sensitive land uses, this project also provided a way to strengthen community engagement in the research and regulatory process related to local air quality and environmental health concerns.

Results:

Each of the interlocking parts of the project produced their own separate but also interacting results that informed each other. For example, the Bay Area EJ analysis and the birth outcomes study helped inform the screening method; the screening method was essential to the siting simulation, and the "ground truthing" exercise (along with community input in devising the Bay Area analysis and reviewing early versions of the screening method) helped to build community credibility for the broad project goal of advancing EJ Methods in ways that better inform regulatory decision-making. The results were as follows:

1) The Bay Area environmental justice analysis indicated that environmental justice disparities were indeed a concern in terms of both community proximity to emission facilities with active air releases as recorded in the Toxic Release Inventory as well as estimated cancer risks and respiratory hazards associated with ambient air toxics. Disparities for African Americans and Latinos persisted in a multivariate regression analysis, controls for spatial autocorrelation tended to slightly attenuate effect estimates but not eliminate their statistical significance, and the degree of linguistic isolation

(measured as the share of households in which no member 14 years old and above speaks English only or English "very well") was a significant explanatory factor in most model specifications. These findings demonstrate that social disparities are persistent even after controlling for potential confounding, and this study also helped elucidate some new important variables, particularly linguistic isolation, as factors to consider in an EJSM incorporating social vulnerability.

2) With regard to birth outcomes, our study builds upon existing work by analyzing the effect of air pollution on birth weight and preterm birth in California. We used California and federal monitoring data for PM2.5, PM10, CO, NO2, SO2, and ozone to assess the relationship between ambient air pollution exposures and birth weight among infants born between 37-44 weeks gestation during the years 1996-2006. We also examined air pollution effects on the risk of preterm birth. We estimated ambient exposures to coarse PM, where coarse particle exposure was defined as the difference in ambient exposures for respirable and fine particles (PM10 - PM2.5.). We assessed the consistency of our effect estimates by testing our models using different nearest monitor radii distance limits for the births we examined from air monitors. Consistent with prior literature, we have shown a modest inverse relationship between ambient criteria air pollutant exposure ($PM_{2.5}$, PM_{10}, coarse PM, CO, NO_2 and O_3) and birth weight among full-term infants as well as risk of preterm birth. Overall, these associations between increasing pollutant exposures and decrements in birth weight and risk of preterm delivery persisted during different trimesters of exposure, although the strongest effects were seen for exposures during the entire gestational period. The results were generally robust in co-pollutant models and across different radii distance limits from air monitors. The results, along with the disparity findings in the Bay Area analysis, suggest reasons why a screening method might be useful for policy makers concerned about environmental health and justice.

3) We developed an EJSM that incorporates multiple indicators of cumulative impacts, including: ambient criteria air pollution exposure, cancer and respiratory hazards associated with modeled air toxics estimates, social vulnerability, and a proximity score (based on a simple distance decay function and population weighting to the tract level). The last of these, the proximity score, provides an indicator of land use (hazard and sensitive land use) which have been identified as important policy and regulatory issues for diverse stakeholders. Because proximity metrics alone do not provide information on potential pollutant exposures or possible human health risks, we also integrated a second dimension of cumulative impact based on air-related measures of potential health risk, including ambient criteria air pollution exposure and cancer and respiratory hazards associated with modeled air toxics estimates. Finally, we also incorporated a third dimension that is based on a range of social vulnerability indicators. These are brought together to create a cumulative score that could identify communities of environmental justice concern. The EJSM is transparent (partly because it relies on intuitive scoring strategies like quintile rankings), and it is adaptable by the agency and sophisticated outside users (as it consists of a few documented Geographic Information Systems (GIS) routines, coupled with a programming exercise); the latter means that agency and other users can weight or add different metrics of cumulative impact and add or update data as

necessary. The method is currently limited to Southern California because of the land use data requirements for our proximity measures but could be adapted to other locations with lower‑quality but still usable land use data. Trainings in the EJSM were conducted with the ARB staff during Spring of 2009, and we have transmitted the relevant data layers and programming syntax. Community advocates were engaged in previewing and reviewing the EJSM to enhance its transparency and ensure its credibility as a tool to guide regulatory decision‑making on EJ issues.

4) The EJSM was used to evaluate the hypothetical siting of the Nueva Azalea plant. We demonstrated that the usual screening procedures of the CEC would likely have rated this site as no worse than many others in terms of EJ issues; the EJSM suggests that the site was one of the "worst" possible areas if one concern was to ameliorate the current pattern of environmental disparity in plant siting. We note that only one plant ranks higher on the environmental justice screening score and that plant has been the subject of major protests concerning its expansion. We conclude that the screening method may have particular strength in predicting likely community reactions, partly because it contextualizes any particular site within a broader landscape of hazard proximity, air pollution burden and social vulnerability.

5) The Hegenberger Corridor microstudy encouraged positive community engagement and enhanced understanding of the research and regulatory process related to local air quality concerns. The study also leveraged community knowledge of local environmental hazards and sensitive receptors to highlight opportunities that would better reduce emissions and protect community environmental health. Most importantly, ground truthing efforts enabled the community to verify the extent to which official emission inventories adequately capture the location and number of emission sources and sensitive receptors, and it can be a way for ARB to periodically evaluate the extent to which data sources are adequately capturing the potential cumulative impacts of multiple emission sources in EJ communities. A small particulate pilot air monitoring study also suggests that there may be local emissions sources of regulatory concern.

Conclusions and Recommendations:

This project entailed multiple components that received ongoing external review and feedback from scientific colleagues and community advocates, particularly for the EJSM, because of its critical application to regulatory decision-making, as state policy mandates that attention be paid to environmental justice issues in the implementation of air quality regulation and, more recently, climate change mitigation strategies as required by AB 32, California's Global Warming Solutions Act .

Going forward, the results of this project highlight several future research and policy opportunities. First, metrics of social vulnerability must be central to the development of scientifically valid methods for assessing geographic disparities in the potential cumulative impacts of pollution burdens across diverse communities in the state. In addition to metrics of SES (such as race/ethnicity, poverty, employment status, etc.), new metrics of community capacity for civic engagement, such as linguistic isolation, are significant and suggest the need

for targeted outreach to better engage language-minority populations in the regulatory process. Second, we think our project's confirmation of the adverse impacts of air pollution on birth outcomes raises the need to consider perinatal health outcomes in future regulatory decision-making regarding criteria air pollutants. Third, we think that the successful integration of community advocates into several aspects of this work – including the Bay Area EJ analysis, the EJSM development, and the Hegenberger microstudy – highlights how combining scientific peer review with ongoing community feedback can ensure that methods development is both scientifically valid and transparent, and that project results are disseminated in ways that productively connect to current environmental justice policy and regulatory concerns.

As for the EJSM, future work should involve developing the land use data (or land use proxy measures) necessary to implement it on a statewide level. Additional metrics and data layers can be integrated into the EJSM to better tailor the tool to specific policy issues. For example, the incorporation of traffic density metrics would be important to build into the screen (given both the social disparity in pollution burdens from mobile sources shown in the Bay Area analysis and the adverse impacts of pollution, much of which comes from mobile emissions, on birth outcomes).[1] Additional data layers should be considered by consulting with scientific colleagues working on environmental and public health tracking initiatives within California and nationally. Finally, alterations to the EJSM, particularly with regard to reweighting data layers or imposing more complicated scoring methods (such as standard deviation breaks or z-scores,) should be weighed against the complexity of implementation and the need to balance scientific validity with transparency in cumulative impacts assessment efforts that are aimed at informing regulatory decision-making.

---

[1] Including traffic metrics was suggested in the original research proposal but not included in the funding after scoping discussions prioritized other aspects of the project; we are in the process of adding a traffic exposure metric to the EJSM as part of a research contract funded by US EPA Region IX but such incorporation was not part of the deliverables for this effort.

Chapter 1. Introduction to the Report

Introduction

Community stakeholders, academic researchers, and others have argued that social inequality, discrimination, and unequal economic opportunity have contributed substantially to a disproportionate burden of environmental hazards on communities of color and the poor. Underlying these concerns, which are frequently articulated using the umbrella term "environmental justice," is the belief that pollution may play an important, yet poorly understood role in the complex and persistent pattern of health disparities among the poor and people of color in the United States. With its emphasis on public health, social inequality, and environmental degradation, environmental justice provides a framework for public policy and scientific debates about the impact of these social and economic factors on the environmental health in these communities.

Whether environmental disparities in the United States do, in fact, exist, has been hotly debated in a series of studies and reviews (Anderton, Anderson et al. 1994; Anderton, Anderson et al. 1994; Bowen 2001; Lester, Allen et al. 2001; Ringquist 2005; Mohai and Saha 2006). The bulk of contemporary evidence supports the view that disparities in proximity to environmental disamenities exists along the dimensions of income and race, and even analysts who remain critical of the national evidence conclude that California is an area where minority communities do appear to bear a disproportionate share of the total burden of pollution exposure and attendant health risks. (see (Bowen 2001) for a review, and (Pulido, Sidawi et al. 1996; Boer, Pastor et al. 1997; Morello-Frosch, Pastor et al. 2001; Pastor, Bullard et al. 2006; Su, Morello-Frosch et al. 2009; Brody, Morello-Frosch et al. Forthcoming) for various presentations of the evidence). As a result of both research and community concerns, the general issue of environmental justice has gained salience with policy makers and regulators in the state.

This pattern of disparate exposure to environmental disamenities by race or income is of particular concern because it is probable that disparities in environmental exposures play an important role in health disparities. The socio-economic stratification of American society is itself a factor (Adler and Rehkopf 2008), as this affects access to care, and it is mirrored by key health outcome measures; the poor are generally less healthy than the rich (Haan, Kaplan et al. 1987; Hahn, Eaker et al. 1996; Ecob and Davey Smith 1999), laborers are more likely to die of heart disease than are members of managerial or professional classes (Navarro 1990; Hemingway, Nicholson et al. 1997; Kawachi and Marmot 1998), and people of color suffer disproportionately from chronic diseases such as cancer, heart disease, and diabetes (USDHHS 1990; Krieger, Rowley et al. 1993). The elimination of health disparities has been identified as a top priority among scientists and public health practitioners, and while community vulnerability and SES factors play important roles, environmental factors, including air pollution, are also important (DHHS 2010). In addition, the direct effect of hazardous social and physical environments can combine with various forms of psycho-social stress to further amplify health disparities along racial and SES lines by enhancing community susceptibility to the effects of toxic substances (Pope, Verrier et al. 1999; Brunner 2000; Gee and Payne-Sturges 2004; Morello-Frosch and Shenassa 2006).

How do we incorporate these scientific insights regarding environmental justice and environmental impacts into regulatory decision making? One central task for the field is to operationalize measures of cumulative impact (including metrics of exposure, hazard proximity, and potential health risk) , which can be integrated with key SES measures that may enhance community vulnerability to the adverse health effects of environmental hazards. Developing concepts of cumulative risk and community vulnerability may pose formidable methodological challenges, but they are crucial for improved regulatory decision-making, enforcement, and intervention (Cal-EPA Advisory Committee on Environmental Justice 2003; National Environmental Justice Advisory Council 2004). Such integrated approaches to assessing cumulative impacts of environmental and non-environmental stressors could lead, as in this project, to the development of screening methods that educate communities about potential drivers of environmental health disparities and that help policy makers determine where they might pay more attention in terms of community outreach, protection, and remediation.

While there were many different components to this research project, the development of an EJSM that integrates measures of environmental and non-environmental stressors was the primary focus of our work. The development of an EJSM posed three related methodological challenges:

1. The need to develop a screening method that is both scientifically valid and readily implemented. This means drawing from the best and most sophisticated environmental justice research but also ensuring that the method remains within the analytic and programming range of those conducting the pragmatic work of siting/permitting/compliance, community outreach, and other policy and regulatory decisions.

2. The need to consider multiple audiences for the screening method. Any method needs to be trusted and understood by community members, thus contributing to an improvement in relationships between agency decision-makers and community leaders. This, in turn, implies the need to involve community members in the creation of the screening method as well as to ensure that the metrics and data are relevant to community environmental justice concerns while also having a foundation in the environmental health, epidemiological, or environmental justice literature. Data layers, scoring methods, and metrics must also be easy to explain and easy to access.

3. Third, because there can be no assumption that any analyst's particular choices with regard to either measures of interest or methods for scoring or weighting factors is definitive, a screening method must be open in its architecture and able to be altered by sophisticated users to integrate and update new data sources as needed.

As discussed below, our research team addressed these challenges in different ways: we developed a combined GIS and programming system and cross-trained CARB staff in its use, and we showed early iterations of the methods to agency staff, external scientific peer reviewers, and community audiences.   To facilitate the adoption of the EJSM, we created a flexible and

open system to enable users to experiment with different weighting schemes and different data layers.

Key Aims of the Project

As noted, this project sought to develop diverse methodological approaches to addressing environmental justice (EJ) concerns of relevance to air pollution regulation in California. One central aim was to develop an EJSM to identify communities of concern with regard to cumulative impacts from environmental and non-environmental stressors, including air pollution exposure and health risk, proximity to sensitive land uses and hazardous emission sources, and indicators of social vulnerability. As an interdisciplinary research team, we brought expertise in environmental health science, epidemiology, economics, social science, spatial analysis and statistical modeling, all with the aim of developing an approach that would enhance regulatory capacity to integrate EJ concerns into regulatory decision-making. We also tried to develop a screening method and research deliverables that innovatively integrated existing ARB data with other relevant data sources in a way that was transparent and scientifically valid to meet the analytical needs of regulators, while also being accessible to community audiences; the latter, as noted above, required that we implement a research and review process that was participatory and clear (see (Cal-EPA Advisory Committee on Environmental Justice 2003).

Thus, this project consisted of a set of interlocking and interdependent components, including:

1. A rigorous baseline analysis of EJ in one area of the state, with the incorporation of spatial controls and the assessment of new independent metrics of social vulnerability;

2. A statewide analysis of the relationship between ambient criteria pollution and adverse birth outcomes, with attention to obtaining a large sample size (approximately 3 million births), assessing the consistency of effect estimates by setting different limits on radii distance from air monitors, testing co-pollutant models, and examining trimester-specific effects along with exposures throughout the full gestational period;

3. The development of an EJSM to identify EJ areas with regard to measures of environmental and non-environmental stressors to assess disparities in cumulative impacts and social vulnerability;

4. The application of the EJSM in a power plant siting scenario to assess its performance against current practice;

5. The implementation of a community-based research project, partly to ground truth the accuracy of statewide secondary databases on emission facilities and sensitive land uses, and partly to build community engagement in the overall project.

Key Aims in Detail

We explain each of these aims and the basic findings in more detail below. Again, the research efforts involved:

1) Establishing a baseline analysis of EJ conditions in at least one part of the state not previously the subject of systematic spatial analysis and multivariate statistical modeling to identify the key determinants that are likely to be associated with existing patterns of racial and ethnic disparities in lifetime cancer and other health risks associated with outdoor air toxics exposure. This analysis was done with application to the San Francisco Bay Area (the area targeted in the original proposal), and the research process had three novel characteristics:

    a. We incorporated community insights into the process by developing the model to be tested in a series of working meetings with local EJ advocates;

    b. We incorporated the best new techniques with regard to both buffering and statistical analysis, particularly with regard to imposing controls for the effect of spatial autocorrelation on the results; and

    c. We explored the use of new variables to get at community capacity for civic engagement, including the degree of linguistic isolation.

This portion of the project had three purposes: to assess with contemporary data whether EJ is a continuing issue in the San Francisco Bay Area, and not one just confined to Southern California (where the bulk of our own research as well as that of the field had been done); to help us explore through statistical analysis new determinants of environmental disparity that should be incorporated into our SES measures for the EJSM (which we did in the form of linguistic isolation); and to engage community advocates early on in providing feedback on the research that would inform the development of the EJSM that in turn, would inform regulatory decision-making on regional and local air quality concerns.

2) Assessing the relationship between ambient criteria pollutant exposures and adverse birth outcomes, notably low birth weight and risk of pre-term birth by tract-level geocoded birth data, with federal and California air quality monitoring information. Previous work in California has examined this issue in Southern California. More recently, statewide analyses have examined birth outcome effects from air pollution exposures using county-level average estimates of pollution exposure with a subset of pollutants over a period of one year or a slightly longer time frame (Ritz and Yu 1999; Ritz, Yu et al. 2000; Woodruff, Parker et al. 2003; Parker, Woodruff et al. 2005; Huynh, Woodruff et al. 2006; Ritz, Wilhelm et al. 2007; Ritz and Wilhelm 2008). We sought to build upon this work by incorporating birth data over ten -year period and to assess the robustness of effect estimates using different radii limits for monitor distances to geocoded births, running co-pollutant models, comparing trimester-specific and full gestational effects, and integrating individual and area-level SES information in statistical models. Moreover, we explored potential effect modification by individual and area-level SES measures.

Consistent with prior literature, our study results shows a modest, yet consistent, relationship between most ambient criteria air pollutant exposures ($PM_{2.5}$, PM_{10}, coarse PM, CO, NO_2 and O_3) and lower average birth weight among full-term infants as well as higher risk of preterm birth. Overall, these associations between increasing pollutant exposures and decrements in birth weight and risk of preterm delivery persisted during different trimesters of exposure, although the strongest effects were seen for exposures during the entire gestational period. Although smaller particles have been the focus of regulatory and scientific attention for its impacts on health (Pope and Dockery 2006), results from this study confirm recent work indicating that exposure to coarse particles may adversely affect birth weight and preterm birth risk (Parker and Woodruff 2008). Prior studies have found evidence for differential effects of air pollution among different socioeconomic groups, such as maternal race (Bell, Ebisu et al. 2007) or neighborhood SES (Ponce, Hoggatt et al. 2005), but we did not find consistent patterns of interaction in our analysis when we examined effect modification by maternal race or neighborhood level poverty rate. Although this study indicates that maternal exposure to air pollution may result in modestly lower infant birth weight and higher risk of preterm delivery, the ubiquity of air pollution exposures and the responsiveness of pollutant levels to planning and regulation efforts suggests the potential implications may be important for infant health and development, and that birth outcomes should be considered in future regulatory decision-making related to criteria air pollution.

3) Using the analytical results to develop and integrate measures of cumulative impact and community vulnerability at the neighborhood level that are transparent, quantifiable, and emerge both from guidance from CARB and engagement with communities. The idea was that these measures could then be systematically integrated into an EJSM that could be used to inform and proactively guide regulatory decision-making, enforcement activities, and community outreach. This turned out to be the most challenging – and the most useful and sought-after – component of the overall project. Working with community leaders, CARB staff, outside experts, and others in an iterative process, we:

   a. Developed a method of creating a land use layer that focuses on land uses deemed sensitive according to ARB guidelines (CARB 2005), intersecting this layer with block level demographics to drive attention to the finest level of geography possible;

   b. Combined the land use data layer with CARB emissions inventory information in a proximity analysis that entailed constructing a multi-buffer approach to distance-weight the number of facilities near any "neighborhood" and then score by neighborhood-based population weights up to the census tract; and

   c. Coupled this hazard proximity score with two other dimensions – air pollution exposure estimates and cancer and non-cancer risk estimates, and social vulnerability measures - to create a cumulative score that could identify EJ communities facing disproportionate cumulative impacts from environmental and non-environmental stressors.

The EJSM that we developed has several virtues: it is relatively simple to implement, it is generally understandable by the non-technical public (partly because, as indicated in the report, we made a strategic choice to use quintile rankings rather than more complicated scoring methods, such as z-scores, strategies that would have required some normalization via logarithmic transformations, as well as simple distance-weighting procedures to characterize hazard proximity), and it is adaptable as necessary by the regulatory staff and by sophisticated outside users (as it consists of well documented GIS routines, coupled with a programming exercise in which weights can be altered and new datasets incorporated). Moreover, those who have reviewed maps of the detailed screening results believe that the pattern portrayed is quite sensible and compares more than favorably to other alternative screening approaches.[2] As part of this exercise, we conducted trainings with the CARB staff and transmitted all relevant data and programming syntax to them.

4) Demonstrated the utility of the screening method by implementing it in the case of a hypothetical siting of an electrical power plant. Working with staff of the CEC, we chose the Nueva Azalea plant, a project that was proposed for the City of South Gate, California but was never sited due to community concerns that eventually led the company to withdraw its application for the site. To conduct this analysis, we:

a. Conducted background research on the facility and the community concerns that were raised about it, demonstrating that environmental justice was a perceived issue;

b. Implemented the traditional CEC method of demographic screening for environmental justice; as we demonstrate below, that method tends to "overscreen" (rule too many areas unsuitable for siting) on the basis of race and "underscreen" (rule too few areas unsuitable for siting) on the basis of poverty, and would have provided an insufficient assessment of potential community resistance to the proposed siting; and;

c. Utilized an alternative approach of scoring potential power plant sites using our EJSM that compared proposed site scores with the scores for all other existing facilities in Southern California, a sample presumably indicative of all other possible sites in terms of proximate land use and other factors.

The exercise shows that the proposed Nueva Azalea site would have been one of among the "worst" possible areas if one concern was to ameliorate the current pattern of environmental disparity in plant siting; indeed, only one plant in Southern California ranks higher on the EJ score and that plant has been the subject of major protest concerning its expansion. We conclude that the EJSM may have particular strength in predicting potential environmental justice concerns related to planned power plant

---

[2] For example, the Environmental Justice Strategic Enforcement Assessment Tool (EJSEAT), a method being developed by the EPA Office of Enforcement and Compliance Assurance to identify areas with potentially disproportionately high and adverse environmental and public health burdens. EJSEAT is currently a draft tool in development, intended for internal EPA use only.

projects, partly because it contextualizes any particular site within a broader landscape of disparate impact and social vulnerability.

5) We conducted a ground truthing microstudy in the Hegenberger Corridor of Oakland to build community trust in this overall research process and to assess how comprehensively state emissions inventories capture local air pollutant emission sources and sensitive receptors of concern within an EJ community. This project consisted of conducting a local-scale, community-based validation study to identify and locate emissions sources that may not be systematically incorporated in state government agency inventories and estimates of community-level pollution burdens. This was conducted in conjunction with a community-based organization, and we incorporated community members as integral partners in data collection, study design and interpretation. This microstudy included:

   a. Workshops where community participants were informed about prior research related to EJ in California, as well as the scientific background necessary to understand cumulative impacts. These workshops also included discussions that led to participants indentifying their local concerns related to cumulative impacts, and making a list of specific facilities of concern that they would be identifying and mapping in the field. Participants were also trained to use GPS receivers and to enter data into the ArcMap GIS application using handheld computers.

   b. Field observation and the location of both air quality hazards and sensitive receptors, followed by spatial analysis and mapping using standard GIS software to examine the concerns of community participants and to evaluate the degree to which CARB air quality and land use guidelines were being met in the Hegenberger corridor, as well as to verify/test the accuracy of facilities reported in state government databases;

   c. A small pilot study involving fine particulate monitoring near emissions sources of concern identified by community members.

Caveats and Modifications

The interrelated projects described above were our original aims and activities in the proposal submitted to CARB. As will be seen in the report, we succeeded in meeting these specified aims. This has, however, been a long research process and project and, along the way, a few modifications were made to our original goals that are key to understanding and gauging the evolution and success of this research program. These include:

1) An acceleration in the development of the EJSM, partly because such a method could be useful in the implementation of AB32 and its mandate to protect disadvantaged communities in any implementation of market mechanisms to regulate greenhouse gas emissions; this fact intensified engagement with ARB staff in the process;

2) An expansion of the ground truthing microstudy, mostly because it was anticipated that new, low- cost particulate monitoring technology would be developed by other researchers at UC Berkeley in a separate CARB-funded contract, and that these devices could be field tested within the Hegenberger micro study. However, the particle monitors were not ultimately developed in time for this proposed field work. Nevertheless, community expectations had been raised and we felt that we had to meet them in the form of some pilot particulate monitoring which this team took on midway through this contract as an add-on to our original project aims.

3) An expansion in the degree of community involvement in guiding the research, something reflected in the incorporation of community groups into early reviews of the Bay Area EJ analysis as well as continuous presentations of various iterations of the screening method to community, professional, and scientific audiences.

Two project pieces that were proposed in our original research proposal were subsequently dropped in the amended contract that was approved by CARB research staff; these two changes were:

1) Traffic analysis: This research team as well as CARB research staff had originally thought that it was important to develop traffic metrics and incorporate them into the EJSM. However, CARB staff ultimately prioritized the birth outcome analysis and decided to eliminate the traffic metrics as part of the project deliverable for the EJSM in the amendment period due to cost considerations. We did receive additional resources from the U.S. EPA and are developing traffic metrics for eventual incorporation into our research and the EJSM. However, those are not part of this report or deliverables for this contract.

2) In developing the prototype for the EJSM, we agreed with CARB staff to focus on Southern California, primarily because of the high quality of the land use data available in that region. Nevertheless, we did apply a version of the EJSM using lower-quality land use data to the Bay Area, primarily as a way to check that the scoring method looked sensible in that region as well. Although the EJSM results for this additional analysis are encouraging, the Bay Area application is not as advanced or reliable at the neighborhood level as is the Southern California application due to the poor quality of the land use spatial data available to use. We also worked to apply the EJSM to San Diego County, work that was not completed because of a decision to focus on attempting to complete the Association of Bay Area Governments (ABAG) region as a cross-check on the method, which was seen by research staff as a higher priority. As noted, that was a very preliminary analysis, mostly for purposes of seeing whether the method made sense in other locales; we chose to focus here on the promised and fully developed single regional case. Application of the method to the rest of the state will require better land use data or land use proxies; we are working on this in other studies and suggest this as an avenue for future research and development.

Road Map to the Report

The report below offers the results from the overall research effort. Because the project was multi-faceted, we present the material in chapter format.

We begin in Chapter 2 with our analysis of the Bay Area EJ "riskscape," examining distributions of point emission sources as well as cancer and respiratory hazards associated with modeled ambient air toxics exposures.

In Chapter 3, we present our analysis of the relationship between ambient air pollution exposures and poor birth outcomes—specifically low birth weight and risk of preterm birth.

Chapter 4 details development of the indicator variables that characterize cumulative impact and community vulnerability as well asthe integration of these metrics into an EJSM, a place-based technique designed to evaluate a large region for cumulative impacts, as measured by environmental exposure and health risk, and socio-demographic inequalities associated with ambient air pollution and its diverse sources.

In Chapter 5, we apply the EJSM to evaluate the EJ implications of a hypothetical case study siting of a proposed electrical power plant, comparing the current CEC practice for identifying areas of EJ concern by a simple demographic screen, with use of the more comprehensive and spatially-specific EJSM; this comparison is extended to include all large power plants in the surrounding region as a way to contextualize the results of the comparison for the case study facility.

In Chapter 6, we describe a community- based participatory research microstudy in East Oakland, where community members were trained to identify and record information on air quality hazards and sensitive receptors using ground truthing field observations, and compare their observations to emissions inventory data used by state government agencies for permitting and regulatory decision-making.. This project also included pilot monitoring for fine particulate air pollutants to explore community concerns regarding air quality in this neighborhood as it relates to emissions sources.

Chapters 7 and 8 provide our conclusions and recommendations, respectively.

Chapter 2. An Environmental Justice Landscape: Disparities and Their Correlates in the San Francisco Bay Area

Introduction to the problem and previous work

The San Francisco Bay Area has a vivid history of activism over issues of "environmental justice." The largely minority community of Bayview-Hunters Point in San Francisco organized for years to force the closing of a (now-shuttered) power plant that was acknowledged by its parent company as one of the "dirtiest" in the state. In West Oakland, activists and residents have expressed concerns about diesel emissions from port-related truck traffic (Palaniappan, Wu et al. 2003; Palaniappan, Prakash et al. 2006). Residents in the inner-ring suburb of Richmond have worried about asthma and other health issues that they associate with nearby oil refineries and other industrial facilities (Larson 2005).

Despite this swell of community activity, to our knowledge, there has been no comprehensive analysis on existing patterns of environmental disparities across the Bay Area as a whole. The gap is striking partly because of the significant body of research on Southern California, to which the authors of this report have contributed (Morello-Frosch, Pastor et al. 2001; Pastor, Sadd et al. 2001), as well as overall studies of California in which all regions are considered together (Pastor, Sadd et al. 2004; Pastor, Sadd et al. 2005).

Whether or not disparities exist in the San Francisco Bay Area is, of course, an open question. In general, the empirical research on EJ can be broadly categorized into two groups (with this study falling in the latter): (1) those very few studies that focus on inferring causation for spatial associations by using the historical context of planning choices and demographic changes within specific localities (Been 1994; Krieg 1995; Yandle and Burton 1996; Pastor, Sadd et al. 2001; Baden and Coursey 2002; Saha and Mohai 2005), and (2) those that seek to investigate the cross-sectional nature of race, income and environmental negatives using increasingly sophisticated statistical methods, sometimes nationwide and sometimes for various regions (Anderton, Anderson et al. 1994; Anderton, Anderson et al. 1994; Been 1995; Boerner and Lambert 1995; Bowen, Salling et al. 1995; Boer, Pastor et al. 1997; Stretesky and Hogan 1998; Morello-Frosch, Pastor et al. 2001; Pastor, Sadd et al. 2005; Morello-Frosch and Jesdale 2006; Mohai and Saha 2007; Downey, DuBois et al. 2008).

Some experts have characterized the evidence of disparities as mixed (see, for example, Bowen, 2001) and it is true that results do vary by region and pollutant. However, this general conclusion remains extant in the field partly because of the initial and influential impact of the Anderton, et al. studies cited above, both of which challenged the geographic focus on zip codes in earlier studies as well as the lack of multivariate analysis; utilizing tracts and multiple regression techniques, these authors found that race was not a significant factor once one controlled for other factors in a nationwide analysis of hazardous treatment sites. Mohai and Saha (2006, 2007), however, show that the failure of Anderton and colleagues to realistically consider geography in their analysis – for example, considering the population of a tract as affected by a facility only if it is inside the tract itself– are flawed because such facilities are

often located on the borders of tracts (perhaps because they are on transit corridors); utilizing data from a previous national study, they demonstrate that this tract-only approach would miss evidence of racial and income disparity while a geographically aware, distance-based approach confirms the racial disparity hypothesis in both simple comparisons and multivariate analysis.[3]

Similarly, in a recent meta-analysis, Ringquist (2005) looked at 49 empirical studies and used newly developed regression techniques to assess common inequity patterns in the various research efforts.[4] The analysis indicates that evidence of racial disparity in environmental hazard burdens exist regardless of "the type of risk examined, the level of aggregation employed, or the type of control variables used in the analysis" (Ringquist 2005): 233); Ringquist suggests that the findings on income are more mixed, although this may reflect a non-linear relationship in which pollution burdens are low for areas with no economic activity, low for areas with great wealth, and peak at income levels somewhere in the middle range (Been and Gupta 1997; Boer, Pastor et al. 1997; Sadd, Pastor et al. 1999; Morello-Frosch, Pastor et al. 2001).

Of course, EJ patterns are also sensitive to the geographic extent of the analysis; evidence of disparity (or lack thereof) is also sometimes masked through national or state pooling of data (which is why the California-wide studies in which differences might be driven by large Southern California gaps could be misleading). The basic insight here is that inequality has to be considered in the context of the industrial clusters, economic development, and traffic patterns that exist in any particular metro area or region. A recent national study on toxic air emissions from large industrial facilities takes this into account by utilizing a fixed effects regression for regions (Ash and Fetter 2004). The authors find that Latinos, for example, are concentrated in metropolitan areas with lower pollution burdens but *within* these metro areas they tend to live on "the wrong side of the environmental tracks" (Ash and Fetter 2004). Downey, et al. (2008) likewise find that that a careful region-by-region analysis suggests that Blacks and Latinos are generally more exposed to hazards than other groups but that this is not always the case; controlling for metropolitan variation and looking at regional particularities is critical.

Even if geographic disparities can be demonstrated, Bowen (2001) and other critics raise legitimate criticisms regarding risk assessment and spatial autocorrelation. On the first dimension, facilities emit different contaminants which vary in physical properties and dispersion characteristics; atmospheric chemistry, facility- type impacts, the potential for human exposures (Foreman 1998), and the health implications of different pollutants widely differ in terms of ubiquity and level of exposures as well as for cancer and non-cancer toxicity. Finally, partly because of the methodological difficulty, very few studies control for spatial autocorrelation (see (Bowen, Salling et al. 1995; Pastor, Bullard et al. 2006)), a systematic process that is common in virtually all regressions data with an underlying spatially contiguous geography that can lead to biased standard errors on regression coefficients and hence an

---

[3] Fisher, Kelly, and Romm (2006) suggest an entirely new approach of using a spatial statistic known as Ripley's K to identify "clusters" of industrial toxic facilities within a region that allows the size or "scale" of the spatial units to vary in order to find the scale at which the most significant spatial clustering is evident. Once identified, the demographic and socio-economic characteristics of the neighborhoods within cluster areas are compared to those of the region in order to validate them as environmental justice sites (or areas of environmental injustice).

[4] Earlier reviews of the literature include Mohai and Bryant (1992), Szasz and Meuser (1997), and Bowen (2001)

overstatement of effect estimates and statistical significance (Anselin and Griffith 1988; Odland 1988, pp. 15-16).

What about research specific to the Bay Area? As noted, there have been some useful studies using community-based participatory research (CBPR) strategies, including Palaniappan, Wu, & Kohleriter (2003), Palaniappan, Prakash, & Bailey (2006), Larson (2005), and May et al. (2004). But these studies, as useful as they may be, are mostly indicative and do not meet the usual scientific standards. A 1995 study performed by the San Francisco Department of Public health did find elevated rates of breast cancer and cervical cancer amongst African American women in Bayview-Hunters Point (Glaser, Davis et al. 1998), which raised concerns over the possible role of industrial sites and led to attempts to halt further industrial development in the area (Kay 1995; Rubenstein 1995). However, elevated rates were not found in a follow up study published in 1998, where cancer rates within Bayview-Hunters Point were found to be equivalent or lower than expected rates based on the total population of San Francisco (Glaser, Davis et al. 1998). The State of California's Environmental Health Investigations Branch has worked on multiple studies analyzing the effects of particulate matter and other air pollutants on Bay Area residents (Lipsett, Hurley et al. 1997; Ostro, Broadwin et al. 2006) and the California Department of Public Health found an association between certain hazardous air pollutants and autism in the Bay Area (Windham, Zhang et al. 2006).

The academic research has been more limited. Fisher et al. (2006) contributed to improved spatial methodology by using GIS to identify the appropriate scale at which to identify the most significant clusters of environmental hazards in one particular county in the East Bay. Szasz and Meuser (2000) developed a detailed analysis of the historical location of industrial sites and minority populations in Santa Clara County but did not utilize multivariate techniques. O'Rourke and Macy (2003) have studied environmental policing in the form of "bucket brigades," citizen taskforces that test local air quality in Contra Costa County and elsewhere. Pellow and Park's work (2002) qualitatively highlights EJ issues in local labor markets by outlining immigrant working conditions in the Silicon Valley electronics industry.

Thus, despite the array of EJ and health outcome studies in the Bay Area, there remains a gap: there is no scientifically reliable quantitative assessment of the *overall* degree of environmental disparity in the region. This project sought to fill that gap, along the way operationalizing a modified version of the GIS approach recommended by Mohai and Saha (2006, 2007), incorporating a full array of modeled risk measures as well as geographic proximity, and controlling for spatial autocorrelation to assure that the relationships found were not spurious. We also sought to introduce new variables as discussed below.

Materials/Methods

<u>Data</u>

We rely here on two air toxics databases and combined them with neighborhood (census tract-level) demographic and other characteristics available from the 2000 Census, as well as some information on land use.

To measure disparities in hazard proximity, we used locational information from the U.S. EPA's Toxic Release Inventory (TRI) for 2003, a collection of self-reported toxic air emissions data from large industrial facilities. In assessing differences in estimated health risk, we relied on the 1999 National Air Toxics Assessment (NATA), a dataset developed by U.S. EPA that estimates annual average ambient air toxics exposures from both mobile and stationary emission sources that can be utilized to estimate potential cancer risk and respiratory hazard at the neighborhood level. All data and analysis was applied to the 2000 census tract level of geography and restricted to the nine-county Bay Area, which includes Alameda, Contra Costa, Marin, Napa, San Francisco, San Mateo, Santa Clara, Solano, and Sonoma counties. Although concentration and health risk estimates derived from NATA can be used to estimate community impacts of air toxics at the census tract level, these estimates cannot be used to accurately assess exposures and health risks for specific individuals, or to identify exposures and risks at smaller geographic units such as census blocks. Despite these limitations, NATA has become a standard database used in environmental justice studies focusing on the tract level and we follow the tradition.

The TRI was mandated under the Emergency Planning and Community Right-to-Know (EPCRA) provisions of the Superfund Amendments and Reauthorization Act (SARA) of 1986, requiring certain industrial and commercial facilities, as well as federal facilities, to report to the U.S. EPA on annual releases and transfers of nearly 650 toxic compounds. There are inherent limitations to the TRI data: emissions are self-reported estimates and not actual measures of releases; small area sources of highly toxic emissions, such as chrome platers, auto body paint shops, and dry cleaners are not required to report; and the TRI does not include releases from mobile sources which are known to significantly contribute to pollution levels and health risks.

Despite its limitations, most of the literature on EJ has taken this database as a starting point in asking questions about the proximity of certain communities to potential hazards. In our consideration of facilities listed in the TRI, we examined only those that report active air releases. Active facilities were located using address-matching (geocoding) of the street address reported to the EPA against the address ranges in high quality spatial data sets recording roads and street statewide. To check for location accuracy and possible errors, each facility address was located using current versions of two different street databases, one from TeleAtlas (2006), and one from Geographic Data Technology (2004).[5]

To measure potential risk, we rely on a second environmental dataset, the U.S. EPA's 1999 NATA, which models concentration estimates based upon an underlying inventory of air toxics emissions including both stationary and mobile sources (U.S. EPA 2009). NATA thus includes modeled ambient air toxics concentration estimates representing the contribution from large industrial facilities as well as smaller- area and mobile emission sources. The greatest contribution to estimated cancer risk from ambient air toxics – over 70% in the Bay Area – are related to mobile emissions.

The NATA inventory is derived from five primary sources, including state and local air quality regulatory agencies, EPA's own air toxics regulatory program and its TRI database, mobile source emissions estimates developed by EPA's Office of Transportation and Air Quality, and other emissions estimates generated from activity data (such as off-road sources). Using the

---

[5] TeleAtlas North America acquired Geographic Data Technology, Inc, in mid-2004.

emissions data as inputs, an air dispersion "fate and transport" model that accounts for movement and atmospheric chemistry of pollutants (due to the effect of winds, temperature, and atmospheric stability) is used to estimate the concentration of each air pollutant for each census tract in the continental United States. The NATA data generated by this process included tract-level concentration estimates for diesel particulates and 177 of the 187 air toxics listed under the 1990 Clean Air Act Amendments. The U.S. EPA also reports figures on cancer risk and respiratory hazard, but these risk estimates do not include diesel and some other air toxics. In our analysis, we combined cancer potency values and respiratory hazard values from U.S. EPA as well as from the California EPA to estimate cumulative lifetime cancer and respiratory risks associated with ambient air toxics exposure. This process enabled us to include contributions from as many pollutants as possible, including the significant effect of diesel.

To calculate cancer risk, we combined air toxics concentration estimates with inhalation unit risk estimates for each carcinogenic compound to estimate overall cancer risks. Estimated cancer risks for each pollutant in each census tract were derived with the formula:

$$R_{ij} = C_{ij} \times IUR_j$$

where R_{ij} is the estimate of individual lifetime cancer risk from pollutant j in census tract i, C_{ij} is the concentration in micrograms of pollutant per cubic meter of air (µg/m3) of the air toxic j in census tract i, and IUR_j is the inhalation unit risk estimate for pollutant j. In accordance with California's AB2588 "Hot Spots" Risk Assessment Guidelines (OEHHA 2003) and EPA's cancer risk guidelines (U.S. EPA 1986; U.S. EPA 1990), cancer risks of each pollutant were assumed to be additive and were summed together in each tract to derive a total individual lifetime cancer risk.

Respiratory hazard was derived by dividing each pollutant concentration estimate by its corresponding Reference Concentration (RfC) to derive a hazard ratio. An RfC for chronic respiratory effects is defined as the amount of toxicant below which long-term exposure to the general population of humans, including sensitive subgroups, is not anticipated to result in any adverse effects. The actual respiratory hazard ratios for each pollutant in each census tract were calculated using the following formula:

$$HR_{ij} = C_{ij}/RfC_j$$

where HR_{ij} is the hazard ratio for pollutant j in tract i, C_{ij} is the concentration in ug/m3 of pollutant j in census tract i, and RfC_j is the reference concentration for pollutant j in ug/m3. An indicator of total respiratory hazard was calculated by summing together the hazard ratios for each pollutant in order to derive a total respiratory hazard index:

$$HI_i = \Sigma_j HR_{ij}$$

where HI_i is the sum of the hazard ratios for all pollutants (j) in census tract i.
For all of the analysis to follow, we used estimates of cancer risk and respiratory hazard that were based on the Hazardous Air Pollutant Exposure Model (HAPEM). HAPEM integrates ambient concentration estimates with information on indoor/outdoor microenvironment

concentrations, penetration of outdoor pollutants into indoor environments, local populations, and individual-level activity patterns to generate an expected range of inhalation exposure concentrations for each census tract before applying the inhalation unit risk estimates and reference concentrations for each pollutant to obtain the final estimates for cancer and non-cancer health risks.

As for the independent variables, it is important to note that the field generally breaks covariates into three broad categories: land use indicators, income indicators, and measures suggesting social vulnerability and/or power. Land use is key because hazards may simply be located where complementary land uses, such as industrial facilities or transportation arteries, are clustered; therefore, any correlation of environmental "bads" with race is an unfortunate byproduct of economic geography. Income is proxy for property values and compensation: more hazardous land uses tend to be where income levels and property values are low, and co-location of the poor and toxics simply reflects the workings of the market system. To the extent that hazard location and poor air quality depends on a community's ability – or inability – to resist placement of undesirable land uses in their neighborhood, markers of social vulnerability, including race, home ownership and other measures are key.

We list the various covariates used in this analysis in Table 2.1, along with definitions and sources. For land use, we utilized a measure from the 2001 U.S. Geologic Survey (USGS) Land Cover Characterization Program that makes use of aerial photo and satellite imagery interpretation to create a land use classification grid at a spatial resolution of 30 meters. This raster dataset was converted to vector polygons, and transferred as an attribute to census tracts using an area-weighting procedure. Unfortunately, the classification method for this data cannot separate industrial, commercial, and transportation land, and combines them into a single class. While such an aggregation is appropriate for explaining cancer risk and respiratory hazard from the broad set of emissions in a database like the NATA, it is less clear that such a measure of land use is appropriate for an explanation of the spatial pattern of the TRI facility location – for which industrial zoning is the driving land use. In the TRI analysis, we therefore use a proxy based on the percentage of area employees in manufacturing from Summary File 3 (SF3) of the 2000 census.

In addition to a direct measure of land use, we use a more indirect measure: population density. We expect, however, a different sign depending on our dependent variables. For estimating TRI proximity, because most of the land area near a TRI is likely to be industrial (despite some residential presence), that should be reflected in the population density measure making it a good (negative) predictor of having a TRI nearby. Our measures of health risk, however, are based on a continuous surface of toxic concentrations from both stationary *and* mobile sources, and denser populations tend to be found in the urban areas of a region where the majority of industrial facilities and mobile emissions are located. Population density therefore should be a good (positive) predictor of the health risk measures.

Population density is also useful because of racial demographic differences between urban and rural areas in the region: if population density is included in the model, then, any observed relationship between minority populations and health risk are not likely to reflect the urban/rural demographic divide. Moreover, our population density measure has greater accuracy than the

direct land use measure because it comes from actual door-to-door counts taken by the short form of the 2000 Census (Summary File 1) while the land use measure is deciphered from satellite imagery. As was done in previous studies (Pastor, Sadd et al. 2001; Pastor, Sadd et al. 2005), we enter the natural log of the measure, given our expectation that the relationship between population density and TRI proximity/health risk is non-linear.

We also tested a variety of additional measures, all from SF3 of the 2000 Census except for the racial demographic variables which come from Summary File 1 (SF1). We include in our analysis both a direct measure of income – per capita income – as well as the percentage population below the federal poverty level. Income is an important measure given the commonly held view that low-income people are more likely to sacrifice air quality for cheaper housing. Moreover, in areas of cheaper residential housing stock and lower incomes, there may be an incentive for businesses (which may or may not bring toxic emissions) to move in due to lower cost real estate and labor, perhaps due to larger pools of available workers (see(Been 1993; Anderton, Anderson et al. 1994; Been 1995).

After testing several income specifications, per capita income dominated in explaining our measures of hazard proximity and health risk. However, the different dependent variables called for alternate specifications of income. The difference in specification lies in the fact that, because our health risk measures are based on ambient toxic concentrations from both stationary and mobile sources over a larger area, they are more correlated with overall economic activity; they tend to be higher in urban areas even when where there are no TRI facilities in the immediate vicinity. For hazard proximity, we would expect any increase in income to make one less likely to live in close proximity to a TRI, but an increase in per capita income from $4,000 to $5,000 is likely to have a greater effect than an increase from $40,000 to $41,000 so we enter the natural log of per capita income into TRI logistical models. For our multivariate models of cancer risk and respiratory hazard, we would expect an increase in income from the very lowest levels to the middle income bands to coincide with increased health risk (due to greater economic activity) but further increases from middle to higher levels of income to be associated with a decline in health risk due to greater political influence in resisting incompatible land uses and increased mitigation costs for polluters, suggesting the U-shaped relationship between income and health risk that has been demonstrated in previous work (see (Boer, Pastor et al. 1997; Morello-Frosch, Pastor et al. 2001; Pastor, Sadd et al. 2005), which is modeled by including the quadratic term.[6]

We also consider home ownership and race. Home ownership approximates the average level of wealth – as opposed to income – in a neighborhood (Krieger and Fee 1994) and, while far from perfect, is the best indicator of wealth that is available in the Census at the tract level. Home ownership may also indicate a greater stake in one's community and, as we have suggested elsewhere, reveal something about political engagement (Morello-Frosch, Pastor et al. 2001; Pastor, Sadd et al. 2005).

Race is at the heart of EJ concerns and so we include measures of the percentage population that is non-Hispanic black (African American), Hispanic (Latino) and non-Hispanic Asian or Pacific Islander (Asian/Pacific Islander). The significance of these measures can be interpreted as

---

[6] Specifically, in these latter models we entered per capita income, normalized as the ratio of per capita income to the mean across all tracts in the Bay Area multiplied by 100, and its square.

picking up the residual correlation between race and either TRI proximity or health risk that remains after controlling for the other important factors in the model.

Finally, we also consider a novel variable as a metric of community power: linguistic isolation. This measure, available in the Census, is the share of all households in which no member 14 years and over speaks only English or speaks a non-English language and speaks English "very well"—in other words, all members of the household 14 years and over have at least some difficulty with English. This measure has been used in in previous work in which we were seeking a control for the presence of more recent immigrants (Pastor, Sadd et al. 2005): the notion was that those with less experience in the country may be less engaged in community activities regarding the siting process for polluting facilities. In general, English language ability is an indicator of ability to navigate the US political system to achieve a positive outcome for the local environment (and it may also indicate the challenges communities can face when receiving information from state agencies about health risks).

Analysis

The basic unit for neighborhood analysis for both the TRI and NATA analysis was the census tract, a standard in both demographic and environmental analysis. For the NATA analysis, the excess cancer risk and hazard ratios were already determined at the tract level. As for the TRI locational research, we concur with Mohai and Saha (2006) that simply designating the tract in which the facility is sited can lead to location misspecification. Therefore, we used a distance threshold procedure by drawing various radii around each facility in GIS; if half of a tract's population fell within each resulting circle, as measured by the populations of the census blocks that fell within or intersected the circle, we considered that tract to be affected. This procedure is similar to the "best practice" approach for characterizing proximity laid out by Mohai and Saha (2006), but expands upon it to make use of the sub-tract level population distribution. The radii utilized were one mile and two and a half miles; we also looked at the area within one half mile of a TRI but the results were quite similar to the demographics for the one mile radius, and so we just present results from the three breaks (within one mile, between one and two and a half miles, and beyond).

To assess the factors associated with point hazard proximity, we utilize a multivariate logistical model that estimates a census tract's probability of being located near a facility. Because such a locational approach has been appropriately criticized for not including mobile sources and not representing exposure or health risk, we assess environmental inequities with respect to two measures of estimated health risk, using, as noted, an iteration of NATA (which includes mobile source emissions) and subjecting that data to multivariate regression analysis.

One crucial part of the analysis involves our investigation of spatial dependence and spatial autocorrelation. Thus, we expand our health risk models by using formal spatial autoregressive (SAR) modeling techniques. This approach is limited to the health risk models partly because the methodology for including spatial effects into regressions in which the dependent variable is dichotomous – as in our logit regressions on TRI proximity – is not well developed and is better suited to continuous variables like health risk. Spatial autocorrelation occurs when a variable is correlated with itself through space (i.e., spatial clustering) and is a characteristic that is observed for most data measured at spatially contiguous units of observation (Odland, 1988). Ordinary

least squares regressions, the general standard utilized in EJ work, can, in the presence of spatially autocorrelated data, exhibit a spatially autocorrelated error term that will bias standard errors and coefficient estimates (Anselin and Griffith 1988; Odland 1988). Such spatial autocorrelation seems particularly likely for racial variables, suggesting that such controls are particularly salient for EJ analysis.

To formally control for spatial dependence, a set of neighbor relationships – or spatial weights – must be constructed. These weights take the form of an NxN matrix $W_{i,j}$ that, for each observation, i ,includes the weights associated with each observation j if they are deemed to be within the realm of spatial interaction – a parameter that is set ahead of time based on some sensible cut – and zero otherwise. In the most basic formulations of such weights, known as first order contiguity weights (including rook and queen weights), the neighbor relationship depends only on whether two observations share a common boundary; under second order contiguity, neighbors of neighbors are included as well and so on. Once all such neighbors are determined, the weights are row-standardized such that the sum of the weights for each row is equal to one, giving each neighbor an equal share of the overall potential spatial influence.

First order contiguity weights are appropriate when the size and shape of the observational units are relatively homogenous (e.g. a rectangular grid). Another approach is preferable for census tracts which, in the Bay Area and elsewhere, come in a wide variety of shapes and sizes. One such approach is inverse distance weighting and it is the strategy we take.

We specifically set the maximum distance for a neighbor effect to occur to be the minimum at which each census tract in the region would have at least one neighbor (10.7 miles) and applied a power function of one, with the distance between tracts measured from tract centroid-to-tract centroid.[7] In a dense urban area, for example, where the distance between tract centroids are typically less than a half mile, this weighting scheme would give more than 22 times as much influence to the "neighbors" within a half mile as to those on the fringes the range of potential influence. At the other end of the spectrum, in rural areas where the census tracts can be large in area, the scheme would give a similar amount of spatial influence to all neighboring tracts within the maximum distance threshold. Such a spatial weights structure is broadly consistent with a visual interpretation of the spatial distribution of our measures of health risk (as well as most of the independent variables considered). In all cases, the weights are row-standardized to one so that the number of weighted observations is not altered through the transformations that occur in the spatial autoregressive models.

Although 10.7 miles may seem like a long distance to allow for spatial effects, we would stress that the decay function at least partially mitigates such concerns; tracts that have large numbers of neighbors tend to be small in area and urban, with a large share of the spatial effect being

---

[7] Centroids for each census tract were determined in ArcMap 9.3 after carefully clipping out water bodies (since it is the center of where people live that is of interest) using "Feature to Point" tool with the "Inside" option activated, which forces the point to remain within the boundaries of the census tract polygon. The census tract and water polygons themselves are from the 2000 Census TIGER/Line data, with all data projected into California Teale Albers with the North American Datum 1983 to obtain accurate distance measurements between centroids. To compute the inverse distance weights, a distance based contiguity matrix was first created in GeoDa (Anselin, Luc, Ibnu Syabri and Youngihn Kho, 2006) and then read into *R* statistical software (R Development Core Team, 2008) where the inverse distance spatial weights matrix was generated using the *spdep* package.

driven by the immediate neighbors. Moreover, setting the threshold distance for spatial effect at the shortest distance that allows all observations to have a neighbor is a common practice to avoid "islands" (e.g. census tracts with no neighbors) (see, for example, (Anselin 2005)). An alternative approach is to set a lower threshold distance and simply allow such "islands" to carry no weight in the spatial weights matrix, but this: (1) assumes no spatial interaction between those observations and their neighbors, which seems to contradict the visual correlation of the values of many of the variables under analysis across the census tracts in such areas, and (2) is somewhat controversial in a literature that suggests that zero weighting any of the observations can create econometric problems. In any case, we also tested all of the models using inverse distance weights specified under 5 mile and 2.5 mile distance thresholds, in which "islands" that resulted from such specifications were assigned a spatial weight of zero; we found no major differences in the key results of the analysis, so we present only the results of what we believe to be the methodologically superior approach of allowing every tract to have a neighbor in the determination of spatial weights.

In the analysis itself, one is attempting to address any spatial autocorrelation in the error term of the regression that is the source of any bias in the coefficients and standard errors of the regression. If the source of the error autocorrelation is an autoregressive process in the dependent variable, then one uses a spatial lag model that models the process by including the spatial lag of the dependent variable (the product of the n x n spatial weights matrix and the n x 1 dependent variable) on the right hand side of the equation. If the source of the autocorrelation is unknown or unobservable, including the omission of an important explanatory variable that is itself spatially autocorrelated or a measurement error in the data that varies systematically with space, then one uses a spatial error model. Tests can be utilized to determine which error model is appropriate, and that is the approach taken below.

Results

The first set of results involves examining the disparities in the location of TRI facilities. The second result looks at issues of health risk in a multivariate analysis. The third set of results examines the impacts of spatial autocorrelation on the regression coefficients and significance levels. We deal with each of these in turn.

Geographic Proximity to TRI Facilities

A traditional approach to examining disparities involves mapping the locations of TRI facilities and looking at the demographics of proximate neighborhoods. Figure 2.1 shows the basic geographic pattern, plotting the locations of facilities with active air releases as recorded in the TRI relative to 2000 census tracts in the Bay Area ranked into terciles by percentage people of color. For visual ease, we focus the map on the more populated sections of the region which include the larger cities of San Francisco, Oakland, and San Jose. (All figures and tables are at the end of the respective chapters).

Figure 2.2 provides a demographic breakdown of the populations by proximity to a TRI release, and it indicates a potential disparity in the racial composition of the proximate population. Table

2.3 considers a range of other variables. As can be seen, there is an income gradient, with increased proximity associated with lower income levels. An even more stark relationship seems to exist between TRI proximity and poverty, with the poverty rate twice as high for tracts situated within a mile of a TRI as they are for tracts greater than two and a half miles away (12 percent and 6 percent, respectively).[8] Likewise, home ownership, a standard measure of wealth, is lower in the more proximate neighborhoods. Unsurprisingly, nearer to TRI's, a greater percentage of land tends to be devoted to commercial, industrial, and transportation uses and the percentage of the local labor force engaged in manufacturing – an indirect indicator of industrial land use frequently employed in the research literature – tends to be higher as well. Population density is lower in neighborhoods that are closer to TRI facilities, something that is partly a function of the fact that in neighborhoods hosting the type of industrial facilities that report to the TRI, some land is devoted to non-residential uses. Finally, figures for immigrants who arrived in the 1980s and 1990s indicate that they are far more as likely to live within one mile of a TRI than to live more than 2.5 miles away (26 percent and 15 percent, respectively). It can also be seen that a household is twice as likely to be linguistically isolated if it is located in a tract situated within a mile of a TRI than if it is more than two and half miles away.

It can also be seen in Table 2.3 that a household is twice as likely to be linguistically isolated if it is located in a tract situated within a mile of a TRI than if it is more than two and half miles away. As we utilize this variable in our regression analysis and pay particular attention to the way it interacts with the race/ethnicity variable, it is useful to repeat the full definition here, particularly as it might be less familiar than the other Census variables to readers. A linguistically isolated household is one in which no member 14 years and over (1) speaks only English or (2) speaks a non-English language and speaks English "very well." In essence, it tags as linguistically isolated those households where all members of the household 14 years and over have at least some difficulty with English.

Figure 2.3 illustrates an exercise in which, using the one mile break, we plotted the share of living near a TRI for each of eight household income categories by race. As can be seen, the likelihood of being near a TRI facility declines as income rises (and so does the disparity between groups). However, the gap between racial and ethnic groups is persistent at each and every level of income, suggesting the need for multivariate analysis.

Table 2.3 shows the results of that multivariate analysis. We utilize a basic binomial logit model, with a response variable that took a value of one for census tracts within a mile of any TRI facility, and zero for tracts more than one mile away from such a facility. The one mile threshold was selected both because it is a reasonable distance to designate "proximate" and because it is standard in previous research, making results generated here more comparable with other EJ analyses using the TRI data in California (for example, (Pastor 1999)). On the left-hand side of the regression, we include homeownership, income,[9] percentage of manufacturing employees,

---

[8] We are using the standard federal poverty rate in these comparisons because this is standard practice in the literature. However, many argue that other measures, such as persons below the 150% and in our development of the screening method, we use 200 % of the poverty level.

[9] Though we examined the performance of different characterizations of income, including testing different specifications (e.g. median household income) and including poverty along with income or alone, we settled on entering per capita income alone given the high collinearity with poverty and the dominance of per capita income

population density, and racial composition of the neighborhood. Because the land use measure combines commercial and transportation land uses with industrial uses and does not distinguish their relative percentages, it works well for pollution burden measures that include mobile and stationary sources, but it is problematic for proxying the more industrial uses associated with TRI's. Thus, here we are forced instead to use a standard but indirect measure, the percent of the local labor force that is employed in manufacturing; this is similar to the strategy used in Anderton et al. (1994) and a number of other studies.

As can be seen in the first column of Table 2.3, home ownership, income, and population density (a proxy for residential land use) are negatively correlated with proximity to a TRI facility, while our proxy for industrial land use is positively associated with proximity to such a facility. Even controlling for all these factors, African Americans and Latinos are significantly more likely to be near a TRI; Asians are less likely, although the result is not statistically significant. We should note that given the pattern of racial differentials in income, there is some degree of collinearity between race and income; the fact that both surface as significant suggests that they likely have independent effects.

Because of the relationship observed in Table 2.2 between TRI proximity, immigration status, and linguistic isolation, we were curious whether more recent migrants were more or less likely to be proximate to TRI releases once we controlled for other factors. The most direct measure to use in addressing this question would be the percentage of recent arrivals, but this measure is highly correlated with the percent Latino and Asian. Thus, after some testing, we decided to separate out one aspect of the argument for immigration status being an important predictor of TRI location: the Census' measure of household level "linguistic isolation," which is defined as households in which no member older than 14 speaks English only or "very well."

In the second set of columns in Table 2.3, we see that linguistic isolation does matter – that is, there is an effect of limited English language capacity even when controlling for all other variables that is significant at the .10 level.[10] Moreover, the coefficient and statistical significance of the percentage Latino falls somewhat when controlling for language, suggesting that part of what looks like racial disparity may reflect the need to convey information and allow groups to mobilize in their native language.

Ambient Air Toxics and Estimated Health Effects

While the previous results demonstrate that facility-based emissions are unevenly distributed, the NATA data on underlying ambient air toxics and diesel particulate exposures from mobile and stationary sources allows us to offer a more complete picture of both cancer risks and respiratory hazard.

---

in terms of significance. The key results are the same with poverty included or under different specifications of income.

[10] Linguistic isolation is likely more significant that one would gather from the table; it carries a p-value of .51 making very near significance at the .05 level.

Figure 2.4 shows a map of excess cancer risks from ambient air toxics in the Bay Area, with the tracts broken by standard deviations from the (logged) mean; this can be visually compared to our earlier map of demographics (Figure 2.1), and such comparison would suggest a correlation.

Table 2.4 is parallel to the approach taken with the TRI data: we offer first simple bivariate comparisons of demographic, income, and other characteristics, for the communities that fall into several categories. Because we cannot simply use proximity, we utilize the standard deviations from the (logged mean) to set breaks, with "least risk" meaning a risk category ranging from zero to one standard deviation below the Bay Area mean, "most risk" being observations ranging upward from one standard deviation above the mean, and "middle range" being everything in between. We use the logged values because the cancer risk and respiratory hazard measures are not normally distributed; by contrast, the natural logs of these measures follow what is close to a normal distribution.

In examining Table 2.4, it can be seen that areas of higher risk are also higher in their proportions of minority and immigrant residents. They have a higher percentage of land devoted to industrial, commercial, and transportation land uses, and they have much lower levels of home ownership. Poverty rises slightly as we go from the "least risk" to the middle range areas and then rises sharply in the "most risk" areas; median per capita income follows a similar trajectory and is much lower in the "most risk" areas. The difference in the pattern between home ownership and income – the "least risk" areas are not necessarily more affluent than the middle band but they are likely more stable and engaged if home ownership is a proxy – suggests that it may be useful to include both in a regression analysis.

For this multivariate analysis, the right-hand-side specification differs slightly from that utilized when examining the TRI pattern, but with a few modifications. First, we enter income with two effects: an initial positive effect at the very lowest levels of income in which we expect that as income rises, cancer risk and respiratory hazard from ambient air toxics will also rise, mostly because it is associated with more economic activity; and a subsequent negative effect in which higher incomes eventually provide a defense, either economically or politically against higher levels of pollution, which is captured by entering income as a quadratic.[11] Second, because ambient air toxics are related to all of the various industrial, commercial, and transportation uses, we can utilize that land use variable in this exercise. Third, under the assumption that transit uses and commercial activity rise with population, population density is expected to be positively correlated with ambient air toxics.

The results are shown in Table 2.5 where we find the anticipated signs with high levels of statistical significance: home ownership is negatively associated with both of our measures of health risk; population density and the percentage industrial, commercial and residential land use is positively associated with health risk; and the coefficient on our income measures are consistent with the U-shaped relationship that was anticipated, showing an initial positive relationship with health risk that is tempered by the negative relationship indicated by the coefficient on relative per capita income squared. The models exhibit reasonable fit, explaining between 56 and 57 percent of the variation in log cancer risk and respiratory hazard across

---

[11] While such a U-shaped pattern is not evident in the simple comparisons drawn in Table 4, it was our hypothesis that it would emerge when controlling for other important measures and that turns out to be the case.

neighborhoods. The race variables are highly significant; when we add linguistic isolation (the second set of columns beneath each measure of health risk), we find it to be important for the cancer risk variable, but insignificant for the respiratory hazard measure.

Controlling for Spatial Autocorrelation

The final table in this section, Table 2.6, reports on the results of regressions in which we try to control for spatial autocorrelation. We utilize the inverse distance (spatial) weights specification discussed earlier. Maximum Likelihood estimation was used to estimate both spatial lag and spatial error models. Diagnostic tests indicated that significant residual spatial autocorrelation remained in the results using the spatial lag approach.[12] Thus, we report the results of the spatial error model in Table 2.6.

Judging by the log likelihood values, the model fit in Table 2.6 improves over the basic OLS approach shown in Table 2.5. The lambda variable represents the spatial lag – or autoregressive component – of the error term in each model. All models suggest positive and highly significant autocorrelation in the residual, meaning that in the OLS specification high values (under-predictions) tended to be clustered in space with other high values and vice versa. Including the term in the regression mitigates the spatial pattern of such under- (over-) prediction by increasing (decreasing) the predicted value stemming from the spatial error model.

As might be expected, the inclusion of spatial effects reduces the significance levels of most of the explanatory variables in the regressions, but the impact is greater in the respiratory hazard models than in the cancer risk models, indicating that there was likely a more significant autoregressive process in the error terms of those models – a notion that is consistent with the slightly higher estimate and lower standard error of the lambda coefficient. However, the basic implications of the results are largely the same with a few interesting nuances.

First, for both the cancer risk and respiratory hazard regressions, the coefficients on African Americans and Asians/Pacific Islanders were reduced substantially more than the Latino coefficient, suggesting that the greater degree of spatial clustering that can be observed for these two groups in the Bay Area may be responsible for greater upward bias in the coefficient estimates and significance levels under the basic OLS specifications. In the initial cancer risk model (first column), for example, the Latino coefficient was reduced by only 20 percent while the African American and Asian/Pacific Islander coefficients were reduced by about 70 percent and 82 percent, respectively.

There is an interesting interaction between the Latino variable, language ability, and space in the cancer risk models. In the two models (spatial and aspatial) that do not take linguistic isolation into account, controlling for spatial autocorrelation causes a reduction in the size of the Latino coefficient. However, in the two models that control for linguistic isolation (the second columns

---

[12] All spatial lag models, spatial error models, and diagnostic tests for residual spatial autocorrelation were carried out using the *spdep* package for *R* statistical software (R Development Core Team, 2008).

of Tables 5 and 6), adjusting for spatial effects has no impact on the Latino coefficient. Rather it is the measure of linguistic isolation that loses explanatory power in the face of spatial controls.

Discussion

Our basic findings are straightforward. First, there is a pattern of disparity by facility proximity, there are differentials in health risks faced by some minority groups, and a reasonable set of covariates does not change these basic results of racial disparity. Second, linguistic isolation reduces but does not eliminate, the racial disparity for Latinos as it seems to do for Asian Pacific Islanders (once we also control for spatial effects); an interesting side-finding is that linguistic isolation matters for cancer risk but not respiratory hazard. Finally, as expected, spatial controls diminish the size of the coefficients for, and the significance of, our independent variables but they do not eliminate statistical significance at traditional levels for the race variables. We deal with each of these main insights in turn.

Even after controlling for income and other factors, race (Black and Latino) seems to matter in our locational measure (TRI proximity) as well as our measures of cancer risk and our measure of respiratory hazard (where Asian Pacific Islander is statistically significant as well). Focusing on the latter measures because they are likely more correlated with health outcomes, it is interesting that African Americans show the strongest association with increased health risk, followed by Asian Pacific Islanders and then Latinos. According to the coefficients, a ten percent increase in the share of African Americans in a tract, holding the other factors constant, is associated with a 13 percent increase in cancer risk and an 11 percent increase in respiratory hazard. The corresponding figures for Asian Pacific Islanders are 6 percent and 7 percent, and 4 percent and 6 percent for Latinos, respectively.

While the inclusion of linguistic isolation diminishes the magnitude and significance of the coefficients on the percentage Latino and Asian Pacific Islander in an OLS model of cancer risk, due to some degree of collinearity with those measures, it does not impact the coefficients one way or the other in the OLS model of respiratory hazards. The indication is that for Latino and Asians Pacific Islander neighborhoods, which are already subject to higher than average cancer risk from air toxics, higher levels of linguistic isolation help distinguish areas that have *even higher* levels of cancer risk.

The same result does not hold for respiratory hazard. Here, Latino and Asians Pacific Islander neighborhoods do experience higher levels of respiratory hazard (again, holding other measures constant), but this by and large is true regardless of the share of linguistically isolated households. We have no completely convincing explanation for the failure of linguistic isolation to obtain significance in this respiratory hazard case, although it is possible that the respiratory hazard data itself has less spatial variability.

However, for Asian Pacific Islanders in the Bay Area, the highly significant positive relationship between Asian Pacific Islanders and both measures of health risk under the basic OLS specifications seems to have been largely driven by spatial dependence in the data. With regard to cancer risk, the spatial error model results indicate that the relationship between Asian Pacific Islanders and cancer risk (evidenced by the positive and significant coefficient on the Asian

Pacific Islander variable in the first column of Table 2.6) may actually be a function of linguistic isolation (see the second column of that table). As for respiratory hazard, the results indicate that removing the impact of spatial error dependence also removes the statistically significant relationship between Asian Pacific Islanders and higher estimated respiratory hazard. This is *not* to say that Asian Pacific Islanders do not live in neighborhoods with higher average levels of respiratory hazard – indeed, all of the bivariate examinations done in this study indicate a positive and significant relationship between the percentage Asian Pacific Islander and respiratory hazard – however, when the degree of disparity is considered in the context of the spatial configuration of the data and controlling for the other measures that are important in explaining respiratory hazard are included in the model, the possibility that the unequal pattern for Asian Pacific Islanders is simply due to chance cannot be ruled out.

Overall, the results indicate – as has been hypothesized by ourselves and others – that spatial autocorrelation is important to consider in developing models that attempt to explain patterns of environmental inequity, and particularly with respect to demographic groups that exhibit a clustered distribution. However, the bottom line is that even after controlling for spatial autocorrelation, when considering the level of estimated cancer risk and respiratory hazard from air toxics concentrations estimated in the NATA data, there is a general pattern of environmental inequity in the Bay Area: densely populated communities of color characterized by relatively low income - and in the case of cancer risk, linguistic isolation - bear a disproportionate share of the hazard and/or risk burden for the region.

We are intrigued by the findings with regard to linguistic isolation. The results may suggest challenges with regard to representation in the siting process but they also suggest remedies: to the extent that the percent Latino and percent linguistically isolated trade-off in significance, what seems like racial disparity may be addressed through communicating information in a more appropriate language.

Before concluding, a large caveat about causality is in order. Those familiar with the literature recognize that the sort of cross-sectional analyses offered here do not explain why the contemporary development of environmental disparities exist (e.g. whether they are due to residential housing choice or discrimination in either housing or facility placement). The results do not even explain the causal pattern; we do not know that race "explains" disparity but rather that the pattern revealed by simple bivariate comparisons is not "explained away" by other important factors that are commonly associated with the spatial distribution of facilities and air pollution. This is essentially a multivariate description of the landscape – although adding controls for spatial autocorrelation is an innovation in that the geographic problems associated with such description are dealt with, to some extent, by accounting for the influence of neighbors.

Conclusion

This portion of the project offers evidence that environmental disparities exist in the Bay Area. In the analysis, we have offer several methodological improvements on past approaches, including better criteria for defining proximity of census tracts to toxic facilities, considerations

of estimated health effects rather than potential exposure, and the controls for spatial autocorrelation in regression testing.

The results of this exercise suggest that in the Bay Area, race and language fluency are significant factors associated with the location of toxic facilities, even when applying econometric methods to control for other important covariates including home ownership, income, population density, and land use. Although the degree of estimated inequity is diminished when we control for spatial autocorrelation in regression tests on estimated cancer risk and respiratory hazard, the percentage African American and Latino remain highly significant in explaining both these measures of health risk, and linguistic isolation remains significant in explaining cancer risk; the results for Asian Pacific Islanders suggest a much weaker or non-existent pattern of disparity.

From a policy perspective, the fact that the race variables for African Americans and Latinos do not lose significance even when we control for other factors and spatial clustering supports the concerns of advocates that have sought to inject the issue of EJ in regulatory decision-making through public testimony and protest. We would, however, also stress the importance of addressing issues of language capacity in both outreach and education; that the influence of race variables used in much of the EJ literature diminish when we control for linguistic isolation suggests both a pattern and a remedy and the importance of communication and outreach.

The rest of the report offers the other building blocks in our own approach to determining a screening method that can tackle issues of environmental health and disparity in productive ways. Before moving on, however, we would stress one aspect of the statistical exercise reported above that is not called out in the report format but was critical to the rest of the project: the role of community participation and feedback in the research process.

The early statistical results were presented and discussed with representatives from over thirty EJ and environmental health organizations in the Bay Area. These organizations were also a base for a community meeting held in November 2006 during which we laid out the entire series of projects covered under this contract, including the goal of developing a screening method. We believe that this interactive process both improved the quality of the research, particularly given community insights about key variables and health concerns, and it also enhanced receptivity and trust for the other projects, particularly the "ground truthing" microstudy in Oakland and the screening method itself.

Table 2.1: Variables Considered in Bay Area Analysis

Variable	Definition (data source)
% owner occupied housing units	% of occupied housing units in a census tract owned by people who are living in them. (2000 Census SF3)
Per capita income	Total income in a census tract divided by total population. (2000 Census SF3)
% persons in poverty	% of population with family income below the federal poverty level. (2000 Census SF3)
Population density	Persons per square mile derived by dividing total population by the area of a census tract. (2000 Census SF1)
% industrial/commercial/transportation land use	% land in a census tract devoted to industrial, commercial or transportation activities. (2001 U.S. Geological Survey (USGS) Land Cover Characterization Program)
% manufacturing employment	% of employed civilian population 16 years and over in a census tract employed in the manufacturing industry. (2000 Census SF3)
% African American	% African Americans (non-Hispanic) residing in a census tract. (2000 Census SF1)
% Latino	% Latinos residing in a census tract. (2000 Census SF1)
% Asian/Pacific Islander	% Asians and Pacific Islanders (non-Hispanic) residing in a census tract. (2000 Census SF1)
% recent immigrants (1980s and later)	% of population that arrived after 1980. (2000 Census SF3)
% linguistically isolated households	% households in a census tract in which no person 14 years and over speaks only English, or speaks English "very well" if they speak a language other than English. In other words, all members of the household 14 years and over have at least some difficulty with English. (2000 Census SF3)

Figure 2.1: Locations of Active TRI Facilities Relative to Neighborhood Demographics by Census Tract
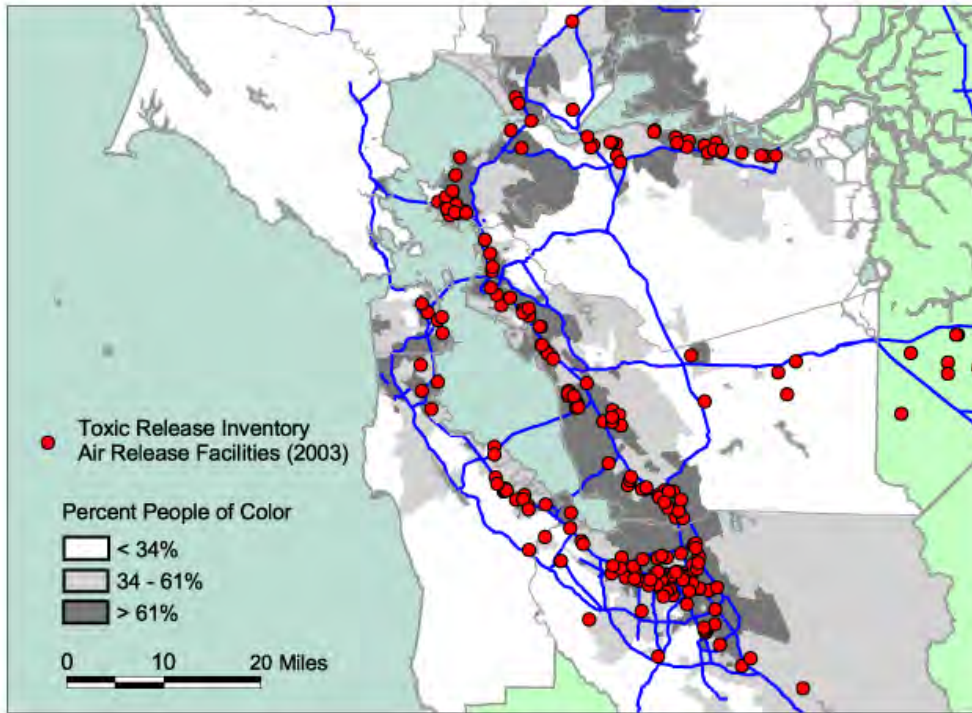
Figure 2.2: Population by Race/Ethnicity and Proximity to an Active TRI
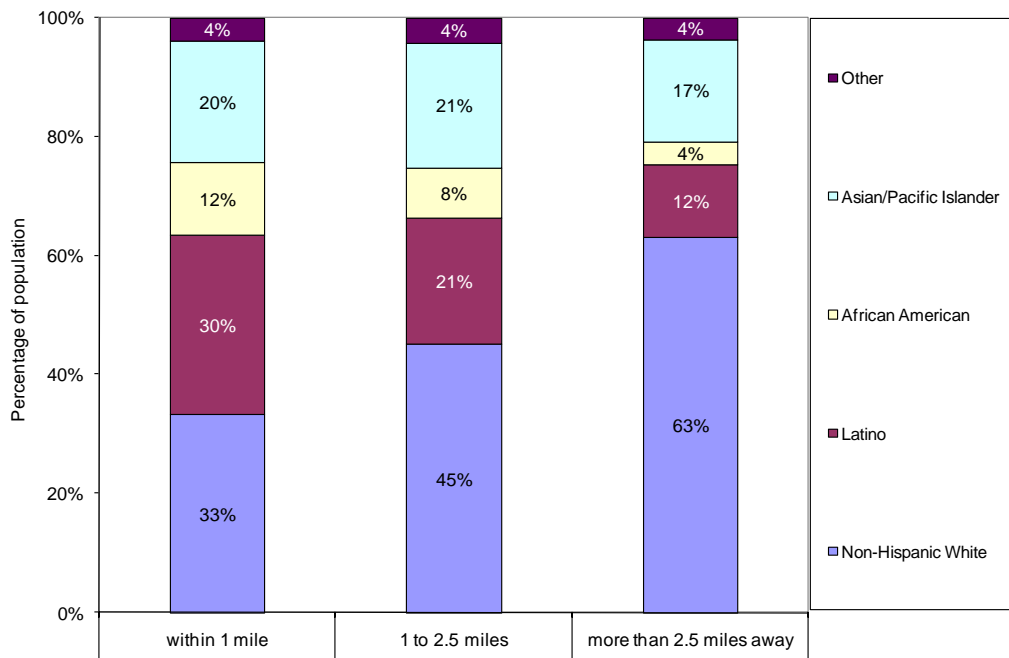
Table 2.2: Demographic and Land Use Characteristics of Tracts (2000) in Relation to Proximity to an Active TRI Facility (2003) in the 9-County Bay Area

Variables	TRI Proximity		
	Less than 1 mile	*Between 1 mile and 2.5 miles*	*More than 2.5 miles away*
% persons in poverty	12%	9%	6%
Median per capita income	$19,702	$25,140	$34,187
% home owner	52%	57%	61%
% industrial, commercial and transportation land use	17%	9%	5%
Population density (persons per square mile)	9,202	10,107	9,748
% manufacturing employment	19%	16%	12%
% recent immigrants (1980s and later)	26%	21%	15%
% linguistically isolated households	12%	9%	6%

Figure 2.3: Percentage Households within 1 Mile of an Active TRI by Income and Race/Ethnicity
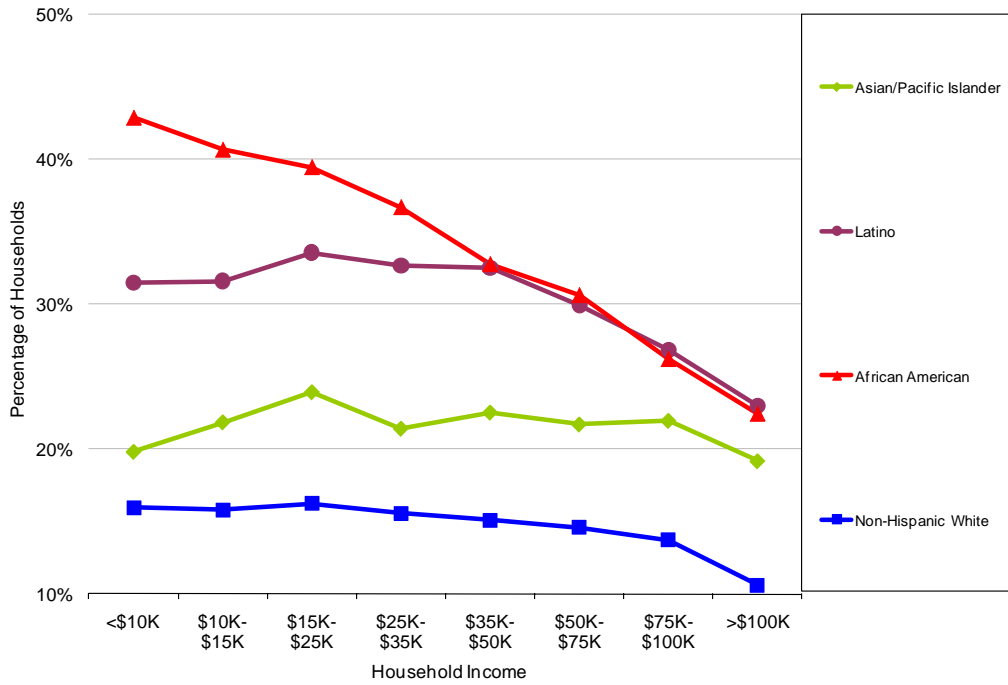
Table 2.3: Multivariate Correlates of Neighborhood Proximity to an Active TRI (Proximate=Within 1 Mile)

Model variables	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.
% owner occupied housing units	-0.8705**	0.3498	-0.5867	0.3823
ln(per capita income)	-0.9722***	0.2933	-0.9035***	0.2965
ln(population density)	-0.1378**	0.0603	-0.1432**	0.0603
% manufacturing employment	6.4811***	0.9542	6.5629***	0.9644
% African American	3.2600***	0.6147	3.3441***	0.6196
% Latino	2.1743***	0.6721	1.5374**	0.7471
% Asian/Pacific Islander	-0.1041	0.6079	-1.0383	0.7841
% linguistically isolated households			2.7777*	1.4229
Log likelihood	-595.58		-593.66	
Percentage predicted correctly	0.8118		0.8190	
N	1403		1403	

***P < 0.01; **P < 0.05; *P < 0.10.

Figure 2.4: 1999 NATA Estimated Excess Cancer Risk (All Sources) by 2000 Census Tracts
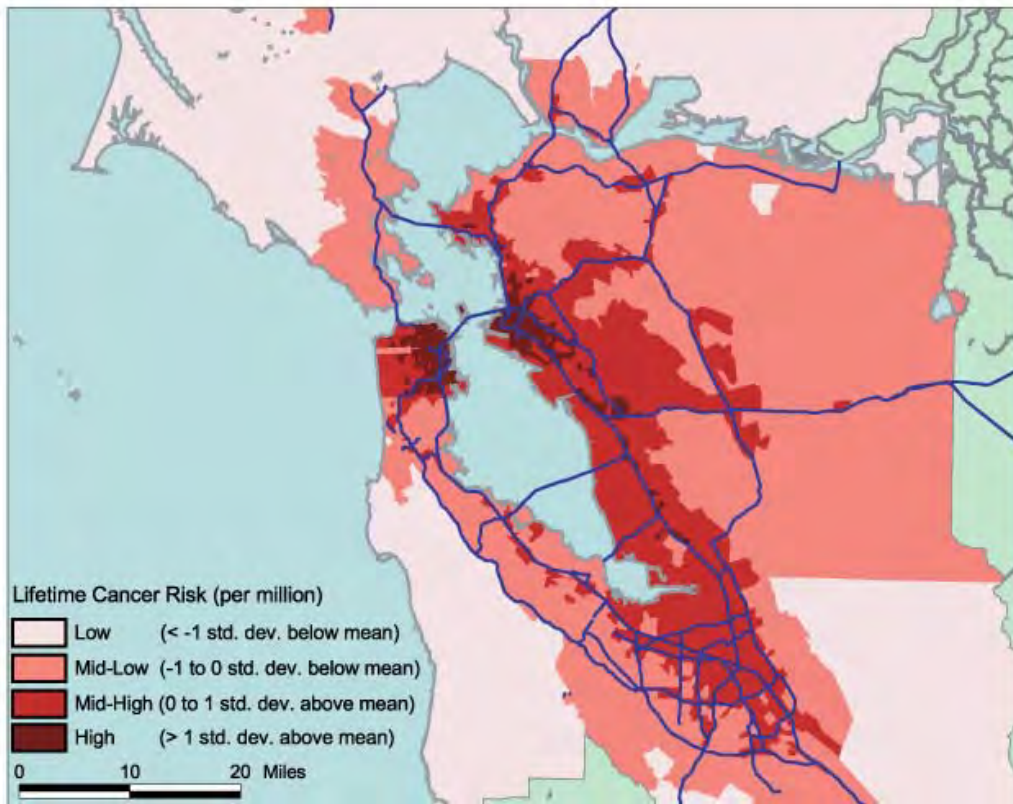
Table 2.4: Demographic and Land Use Characteristics of Census Tracts by Estimated Cancer and Non-Cancer Risk Category

	Cancer Risk			Respiratory Hazard		
Variables	*Least risk*	*Middle range*	*Most risk*	*Lowest hazard ratio*	*Middle range*	*Highest hazard ratio*
% Anglo	68%	48%	39%	66%	49%	33%
% African American	4%	7%	16%	5%	6%	16%
% Latino	17%	20%	17%	18%	19%	24%
% Asian Pacific Islander	7%	21%	24%	7%	22%	23%
% Other	4%	4%	4%	4%	4%	4%
% home owner	70%	61%	28%	71%	59%	34%
Median per capita income	$28,231	$28,187	$22,973	$27,137	$29,329	$20,487
% persons in poverty	7%	8%	15%	7%	8%	15%
Population density (persons per square mile)	2,929	8,175	24,194	2,603	9,346	19,425
% industrial, commercial and transportation land use	3%	8%	17%	4%	8%	20%
% recent immigrants (1980s and later)	10%	21%	24%	10%	21%	26%
% linguistically isolated households	4%	8%	13%	4%	8%	14%

Table 2.5: Multivariate Correlates of Estimated Cancer and Non-Cancer Risk from Air Toxics, Linear Model

Model variables	Cancer Risk				Respiratory Hazard			
	Coeff. Est.	*Std. Err.*	*Coeff. Est.*	*Std. Err.*	*Coeff. Est.*	*Std. Err.*	*Coeff. Est.*	*Std. Err.*
Intercept	3.230***	0.110	3.205***	0.110	-0.115	0.095	-0.115	0.095
% owner occupied housing units	-0.457***	0.045	-0.408***	0.047	-0.230***	0.039	-0.232***	0.041
relative per capita income (tract/region)	0.588***	0.080	0.619***	0.080	0.661***	0.068	0.660***	0.069
relative per capita income squared	-0.001***	0.000	-0.001***	0.000	-0.001***	0.000	-0.001***	0.000
ln(population density)	0.152***	0.008	0.150***	0.008	0.133***	0.006	0.133***	0.006
% industrial/commercial/transportation land use	0.854***	0.079	0.823***	0.079	0.791***	0.068	0.792***	0.068
% African American	1.257***	0.086	1.277***	0.086	1.119***	0.073	1.118***	0.074
% Latino	0.373***	0.086	0.232**	0.096	0.610***	0.074	0.614***	0.083
% Asian/Pacific Islander	0.646***	0.065	0.461***	0.086	0.731***	0.056	0.737***	0.074
% linguistically isolated households			0.643***	0.198			-0.021	0.171
Adj. r-squared	0.5692		0.5721		0.5635		0.5632	
Log likelihood	-465.9250		-460.6600		-251.9300		-251.9250	
N	1403		1403		1403		1403	

***P < 0.01; **P < 0.05; *P < 0.10.

Table 2.6: Multivariate Correlates of Estimated Cancer and Non-Cancer Risk from Air Toxics, Spatial Error Model

Model variables	Cancer Risk				Respiratory Hazard			
	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.
intercept	3.284***	0.269	3.287***	0.268	-0.166	0.358	-0.172	0.360
% owner occupied housing units	-0.126***	0.034	-0.112***	0.035	-0.060**	0.026	-0.055**	0.027
relative per capita income (tract/region)	0.171***	0.061	0.184***	0.062	0.067	0.047	0.072	0.048
relative per capita income squared	-0.000*	0.000	-0.000*	0.000	-0.000	0.000	-0.000	0.000
ln(population density)	0.087***	0.006	0.087***	0.006	0.068***	0.004	0.068***	0.004
% industrial/commercial/transportation land use	0.696***	0.053	0.686***	0.054	0.561***	0.041	0.557***	0.041
% African American	0.382***	0.072	0.392***	0.072	0.147***	0.055	0.150***	0.055
% Latino	0.297***	0.071	0.235***	0.079	0.239***	0.055	0.218***	0.061
% Asian/Pacific Islander	0.115*	0.060	0.034	0.074	0.018	0.046	-0.009	0.057
% linguistically isolated households			0.254*	0.139			0.085	0.107
lambda	0.978***	0.008	0.978***	0.008	0.987***	0.005	0.988***	0.005
Log likelihood	133.088		134.759		496.588		496.903	
N	1403		1403		1403		1403	

*$***P < 0.01; **P < 0.05; *P < 0.10.*

Chapter 3: The Association between Ambient Criteria Air Pollutant Exposure and Adverse Birth Outcomes in California

Introduction

Nearly 6.2% of all singleton births in the U.S. are low- birth weight infants (National Center for Health Statistics 2003). Low birth weight (LBW), defined as a birth weight of less than 2500 grams, is an important predictor of infant mortality and future child health status (McCormick 1985; Sappenfield, Buehler et al. 1987; Mathews and MacDorman 2007), including risk of cardiovascular disease (Barker, Winter et al. 1989; Barker 1995) and impaired cognitive development (Sørensen, Sabroe et al. 1997; Shenkin 2004). Preterm birth, which occurs in just over one in ten American births, is defined as birth before 37 weeks of gestational age (Windham and Fenster 2008). Preterm birth rates have been increasing over the last two decades, from 9.4 percent to nearly 12 percent. Moreover, nearly 440,000 infants are born prematurely on an annual basis, with preterm birth being the second leading cause of infant death and a major contributor to early childhood morbidity (Martin, Hamilton et al. 2003; Moore 2003).

The fetal origins hypothesis posits that *in utero* delays in growth and development as well as preterm birth can increase the risk of many chronic diseases throughout the life course (Osmond and Barker 2000). A strong body of literature has shed much light on the individual-level risk factors (e.g., health behaviors, inter-pregnancy interval, SES, race/ethnicity, and access to adequate health care) (Cramer 1995; Rawlings, Rawlings et al. 1995; Collins, Herman et al. 1997; Hessol, Fuentes-Afflick et al. 1998; Shi, Macinko et al. 2004) as well as place-based factors (e.g. social inequality, neighborhood quality and support networks) (O'Campo, Xue et al. 1997; Buka, Brennan et al. 2002; Morenoff 2003; Huynh, Parker et al. 2005) that are associated with LBW and preterm birth (Parker, Schoendorf et al. 1994; Berkowitz, Blackmore-Prince et al. 1998; da Silva, Simoes et al. 2003).

In the past decade, an increasing number of studies within the United States and elsewhere have identified a relationship between air pollution, preterm birth, and LBW. These studies primarily focus on the commonly monitored criteria air pollutants, including ozone (O_3), particulate matter ($PM_{2.5}$, PM_{10}), carbon monoxide (CO), nitrogen oxides (NO_2 or NO_X), and sulfur dioxide (SO_2). Results from these studies are inconsistent in terms of singling out a particular pollutant that is consistently associated with these adverse perinatal outcomes or elucidating potential windows of susceptibility of the fetus by trimester of exposure. Several reviews have examined the evidence linking air pollution and LBW (Glinianaia, Rankin et al. 2004; Maisonet, Correa et al. 2004; Lacasana, Esplugues et al. 2005; Sram, Binkova et al. 2005; Ritz and Wilhelm 2008; Woodruff, Parker et al. 2009), although inconsistencies in study design have precluded a systemic meta-analysis of the literature. Despite difficulties in synthesizing the literature, reviews have generally concluded that the body of evidence suggests small effects of air pollution exposure on birth weight and preterm birth. Additional investigation is needed to better understand which pollutants and which trimester of exposure appear to cause adverse effects in the fetus.

Air pollution is hypothesized to affect the fetus directly through transplacental exposure or indirectly by adversely impacting maternal health during pregnancy (Glinianaia, Rankin et al. 2004). With the exception of CO which is known to cross the placental barrier and bind efficiently with fetal hemoglobin, the mechanism of toxicity of air pollution on the fetus is poorly understood (Ritz and Yu 1999). Although toxicity mechanisms remain unclear, several have been proposed, particularly for PM effects, including oxidative stress, pulmonary and placental inflammation, blood coagulation, endothelial dysfunction, and changes in diastolic and systolic blood pressure (Kannan, Misra et al. 2006).

California has been the focus of many air quality and birth outcome studies, in part because of its persistent ambient air quality problems. Studies in Southern California have found positive associations between last trimester exposure to CO and particulate matter less than 10 μm in aerodynamic diameter (PM_{10}) and full-term LBW (Ritz and Yu 1999; Wilhelm and Ritz 2005). Two additional California studies found LBW associations for $PM_{2.5}$ but not CO when examining births throughout the entire state (Parker, Woodruff et al. 2005) and for O_3 and CO for births during 1975–1987 in several Southern California cities (Salam, Millstein et al. 2005). For pre-term births, California studies found similar results. Ritz et al (Ritz, Yu et al. 2000) observed a dose-response relationship for PM_{10} exposure during the last 6 weeks of pregnancy and preterm delivery, and reported an increased risk for preterm delivery when controlling for other factors, including other highly correlated pollutants like nitrogen oxides and carbon monoxide. Another California analysis found similar results, with increased preterm birth risks during both the first trimester and final two weeks of pregnancy, as well as total exposure during the entire pregnancy, associated with preterm birth risk (Huynh, Woodruff et al. 2006).

Studies in other areas have found links between air pollution, LBW, and preterm birth. A study in Massachusetts and Connecticut found that an inter-quartile increase in gestational exposure to NO_2, CO, PM_{10} and $PM_{2.5}$ lowered birth weight, and that effect estimates for $PM_{2.5}$ were higher for African American versus White mothers (Bell, Ebisu et al. 2007). A national study linked preterm births to average county-level PM exposures for 2001-2003 and found that results varied markedly by region, with strong associations in the Northwest versus null associations in the Southwest. After controlling for region, the small positive association between PM exposure and LBW in multivariate models became null (Parker and Woodruff 2008). Internationally, results have also been mixed. Studies in Brazil, Australia and Germany found positive associations between exposure to PM and LBW, (Gouveia, Bremner et al. 2004; Mannes, Jalaludin et al. 2005; Slama, Morgenstern et al. 2007) while studies in Canada and Taiwan found null or weak associations (Lin, Li et al. 2004; Dugandzic, Dodds et al. 2006). Other studies found small associations with exposures to other pollutants such as CO, NO_2, and SO_2, and LBW (Wang, Ding et al. 1997; Ha, Hong et al. 2001; Gouveia, Bremner et al. 2004; Mannes, Jalaludin et al. 2005; Dugandzic, Dodds et al. 2006). For preterm birth, studies considering exposure to all criteria air pollutants in the US and Canada, found associations with nitrogen dioxide exposures during the previous six weeks, $PM_{2.5}$ concentrations from the previous week in a retrospective time-series cohort analysis (Darrow, Klein et al. 2009), and average PM2.5 exposure throughout the gestational period (Bobak 2000; Brauer, Lencar et al. 2008).

Different results across studies may be due to differences in how studies control for confounders, regional and national variations in underlying health conditions among populations, differences

in pollution measurement techniques, spatial and temporal differences in exposure assessment, composition of the pollutants examined (e.g. PM composition and size), study sample size, and statistical modeling techniques (Glinianaia, Rankin et al. 2004; Maisonet, Correa et al. 2004; Sram, Binkova et al. 2005; Woodruff, Parker et al. 2009). Although the effects of air pollution on birth weight and preterm birth appear to be small, current findings have important implications for infant health due to the ubiquity of exposures to many of the air pollutants within and outside the United States. Moreover, evidence suggests that certain socio-demographic groups may be more vulnerable to the adverse effects of air pollution on infant health (Ponce, Hoggatt et al. 2005; Morello-Frosch and Shenassa 2006; Bell, Ebisu et al. 2007), although this issue has not been extensively examined.

This study builds upon existing work by analyzing the effect of air pollution on birth weight and preterm birth in California. We used California and federal monitoring data for $PM_{2.5}$, PM_{10}, CO, NO_2, SO_2, and ozone, to assess the relationship between ambient air pollution exposures and birth weight among infants born between 37-44 weeks gestation during the years 1996-2006. We also examined air pollution effects on the risk of preterm birth. Effect estimates- associated ambient exposures to coarse PM were also assessed where coarse particle exposure was defined as the difference in ambient exposures for respirable and fine particles (PM_{10} - $PM_{2.5}$).

Materials and Methods

We calculated pollutant exposures during pregnancy using monitoring data from all monitors within a specified radius of the census tract or ZIP code of the mother's residence. For each birth, we calculated averages for the time periods corresponding to the nine months of pregnancy as well as for each trimester; trimester-specific exposures were examined to identify potentially critical times during pregnancy when pollutants may affect birth weight. We assessed effects for birth weight, measured continuously and categorically. We also analyzed the potential confounding and interaction effects of individual-level and contextual-level measures of SES based on previous work (Bell, Ebisu et al. 2007; Parker and Woodruff 2008).

Natality Data

Data for this analysis came from several sources that were merged using spatial and temporal variables. We acquired tract and ZIP code geocoded birth data from the California Department of Health Services Natality files for 1996-2006 (California Automated Vital Statistics System 2006, unpublished data). California reports locations of maternal residence at both census tract and ZIP code levels. Those births reported with a valid 2000 census tract were assigned that tract code. Remaining births reported with a valid 1990 census tract were assigned that tract code. If neither a valid 2000 tract code nor a valid 1990 tract code was reported on the birth record, then a ZIP code matching a valid census 2000 ZIP Code Tabulation Area (ZCTA) was used as the relevant geocode for the birth. Births that could not be assigned a valid tract or ZCTA location by these methods were excluded from the analysis.

For the LBW analysis, we restricted our analysis to singleton live births, with a gestational age between 39-44 weeks to examine birth weight effects. For preterm birth effects, we restricted our sample to singleton births with a gestational age indicating survival into the third trimester (29-44 weeks gestation). All births in the analyses had information on infant birth weight, sex, date of birth, maternal educational attainment, parity, and a maternal age of 9 to 49 years old. Infants with a reported birth weight that is implausible for gestational age at delivery were excluded from all analyses using the method of Alexander et al. (Alexander, Himes et al. 1996). For example, among full-term births, those with a birth weight of 1,000 grams or less were excluded, as were those with a birth weight greater than 6,000 grams.

LBW was defined for infants delivered full-term as a birth weight of less than 2,500 grams, compared to a birth weight of 2,500 grams or more. Preterm birth was defined as delivery between 29-34 weeks of gestational age with full-term births (gestational age 39-44 weeks) as the reference category. Births occurring between 35 and 38 weeks were excluded in the preterm birth analyses so as to create a clear distinction between preterm and term births.

Because maternal demographics are independently associated with birth weight and preterm birth (Sappenfield, Buehler et al. 1987; Rawlings, Rawlings et al. 1995; O'Campo, Xue et al. 1997; Morenoff 2003) and air pollution (Woodruff, Parker et al. 2003), we added the following measures of maternal characteristics to our multivariate models: maternal age (9-14, 15-19, 20-34, 35-49 years old), educational attainment (</=6th grade, 7- 11th grade, high school diploma or GED, 1-3 years of college, or >= 4 years of college), maternal race/ethnicity (non-Hispanic White, Hispanic, non-Hispanic Black, non-Hispanic Indian/Alaska Native, non-Hispanic Asian Pacific Islander, and non-Hispanic Other Race), and maternal birthplace (Mexico, other or unspecified foreign country, and United States). We also controlled for temporal variables, including calendar year and season of delivery (Jan-March, April-June, July-Sept, Oct-Dec), marital status, parity, and presence of any vs. none of the following pregnancy risk factors: anemia, diabetes, chronic or pregnancy-associated hypertension, and/or herpes). We also used the Kotelchuck index of prenatal care adequacy for the birth weight analysis (no prenatal care, inadequate, less than adequate, adequate, or unknown) (Kotelchuck 1994). For the preterm birth analysis we simply controlled for the initiation of prenatal care for using a trimester of prenatal care initiation (first trimester, second trimester, third trimester or no prenatal care, or unknown).

We also included four measures of neighborhood SES, measured cross-sectionally at the time of the 2000 Census (Krieger, Chen et al. 2003). These measures included: neighborhood poverty rate calculated as the proportion of residents living in households with an income under the federal poverty level (30% and higher, 20% to 29%, 10% to 19%, 5% to 9%, under 5%); neighborhood unemployment rate calculated as the proportion of residents aged 16 years and older in the labor force who were currently looking for work (15% and higher, 10 -14%, 7.5-10%, 5-7.5%, under 5%); home ownership- calculated as the proportion of households owned by their residents (under 20%, 20% to 39%, 40% to 59%, 60% to 79%, 80% and higher); and neighborhood educational attainment rate, a measure of human capital that was calculated as the proportion of residents aged 25 and older with at least a high school education (20% and higher, 15% to 19%, 10% to 14%, 5% to 9%, under 5%). Values for 2000 census tracts and 2000 ZCTAs were calculated from the SF3 file of the 2000 census. Values for 1990 tracts were calculated

using the Census Tract Relationship File to apportion 2000 population characteristics to 1990 tract geographic boundaries (US Department of Commerce. Bureau of the Census 2001).

Exposure Measurements

Information on the ambient concentrations of criteria air pollutants came from two sources, the Environmental Protection Agency's Air Quality System (AQS) (US EPA 2007) and the California Aerometric Information Reporting System (CalAIRS) (CARB 2007). Concentration measurements for gaseous criteria pollutants (CO, NO_2, ozone and SO_2) were usually measured in ppm and particulate air pollutants (PM_{10}, $PM_{2.5}$, and coarse PM) were usually measured in $\mu g/m^3$. PM_{10} refers to particles of 10 micrometers or less and $PM_{2.5}$ represents particles less than 2.5 micrometers in aerodynamic diameter. Coarse PM refers the difference of PM_{10} and $PM_{2.5}$. Concentrations for these pollutants reported in other units (such as ppb) were transformed into the above units. The latitude and longitude of the monitor locations as reported in CalAIRS or AQS were validated by comparing the reported coordinates to address geocoding in Google Earth [Version 4.2.0205.5730, 2007].

Daily values of gaseous pollutants (CO, NO_2, O_3 and SO_2) were calculated by averaging hourly measures, if there were at least 18 hourly measures in a day. Although gaseous pollutants were usually monitored daily, PM was less frequently measured, usually every three to six days. Particulate matter measures were usually reported as daily summaries. When they were not, daily averages of hourly measures were calculated, provided that there were at least 18 hourly measures in a day. If there was at least one valid daily measure of any gaseous or particulate pollutant in a week, a weekly summary for that pollutant was calculated by averaging the daily summaries in that week. Weekly air pollution concentration summaries were assigned to each tract and ZCTA by measuring the distance between the corresponding latitude and longitudes of each active monitoring site and each census block centroid, while accounting for the curvature of the earth. Block-level weekly pollution measurements and distances for each pollutant were then averaged up to the tract and ZCTA levels using the population living within each block as a weighting factor.

Gestational age was reported in the natality file based on the mother's last menstrual period (LMP). We used this information to calculate air pollution exposure for each pregnancy and pollutant for the entire pregnancy and each trimester. For each birth, full pregnancy and trimester-specific exposure measures were calculated by assigning each week of pregnancy the weekly average concentration measure for each pollutant specific to its geocode type (2000 or 1990 census tract, or 2000 ZCTA). Monthly summaries were then calculated by averaging the weekly summaries within each four-week period after conception. If there were fewer than three weekly summaries in a given month, it was not assigned a monthly summary concentration. First trimester summaries were calculated by averaging the first four monthly concentration averages, if none were missing. Second trimester summaries were calculated by averaging the 5th to 7th monthly averages, if none were missing. Third trimester summaries were calculated in like manner, depending on the number of weeks before delivery. Full pregnancy summaries were calculated by averaging all measurements during pregnancy. We assigned a distance to each pregnancy with valid pollutant measures using the maximum distance to an active monitor during any single week of pregnancy. We estimated exposures from monitors within a set of

distance radii- 3km, 5km, and 10km- to assess whether effect estimates were more specific within shorter monitor distance from the mother's geocoded residence.

<u>Analysis</u>

We used linear multivariable models (SAS 9.2) to estimate the impact of air pollutants on birth weight as a continuous measure, and logistic regression models to estimate air pollution effects on birth weight as dichotomous outcome (<2500 grams versus ≥2500 grams). For PM, we estimated the birth weight effect in grams for each 10 μg/m^3 increase in exposure, for CO, the measure was grams of birth weight per ppm, for O_3 and NO_2, the measure was grams of birth weight per part per hundred million, and for SO_2, the measure estimated was grams of birth weight per ppb.

For the preterm birth analysis, we used logistic models to estimate the impact of air pollutants on preterm delivery (based on the four gestational age cut points described above). For PM, we estimated the odds ratio of preterm delivery in response to a 10 μg/m^3 increase in exposure, for CO, the odds ratio of preterm delivery per ppm, and for O_3, NO_2, and SO_2, the odds ratio of preterm delivery per part per hundred million.

In addition to infant sex and gestation age, the maternal factors described above (maternal age, marital status, educational attainment, race/ethnicity, parity, maternal birthplace, prenatal care access, and presence of pregnancy risk factors) along with calendar year, season of delivery and area-level measures (neighborhood educational attainment, poverty rate, unemployment rate, and home ownership) were included in the multivariable models to obtain adjusted estimates. We ran logistic and linear models to examine trimester-specific effects on birth weight and risk of preterm birth as well as effects from full-term pregnancy exposures. We also examined pollution effects on birth weight and preterm birth within strata of maternal race/ethnicity and neighborhood-level poverty rate to assess for potential effect modification. Finally we ran models with two pollutants included simultaneously to assess potential confounding effects of co-pollutants.

Results

Pollutant exposures were estimated for 3,545,177 singleton births for the birth weight analysis and 2,936,711 births for the preterm delivery analysis, although not all births had available monitoring data for all pollutants. 2.3% of births included in the study were under 2,500 grams and 3.1% were preterm with delivery at 29-34 weeks of gestational age. Table 3.1 provides descriptive statistics on the characteristics of the eligible singleton births between 1996 and 2006 in California, whether within 10km of an active monitor throughout pregnancy or not, as well as those included in the study linked within 10km to at least one criteria pollutant for the full pregnancy. Mothers in the study population were predominantly Hispanic or White, over half were born in the United States, and 59% of mothers included in the study had low educational attainment (completed high school or less). Full pregnancy pollutant exposure means and interquartile ranges are shown in Table 3.2. Covariation between gestational exposure estimates ranged from -55% between O_3 and CO to 87% between coarse PM and PM_{10}. Covariation with

an absolute level above 70% consisted of: $PM_{2.5}$ exposures had 0.72 and 0.74 correlation with NO_2 and PM_{10}, respectively; coarse PM had 0.87 correlation with PM_{10}, and CO had 0.79 correlation with NO_2 (data not shown). Pollutant levels varied slightly by year and season of birth (data not shown).

Low Birth Weight Results

In multivariate models, lower birth weight was associated with shorter gestational age; female infant sex and Black, Asian, and Hispanic mothers; younger maternal age; lower maternal educational attainment; lower parity; less access to prenatal care; being unmarried; and living in neighborhoods with lower educational attainment, lower home ownership rates, and higher rates of poverty and unemployment (Table 3.3). Pollution models were adjusted for all of these maternal, infant, and neighborhood risk factors as well as the type of assigned geocode (i.e. 2000 tract, 1990 tract, or 2000 ZCTA),calendar year, and season of birth.

Table 3.4 shows multivariate modeling results for differences in birth weight associated with air pollution exposures for different radii distance from an air monitor. NO_2, O_3, PM_{10}, $PM_{2.5}$ and coarse PM were consistently linked to LBW within all three different distance limits and CO was linked to LBW within 5 and 10 kilometer distance limits in the linear models. NO_2 was associated with increased odds of LBW across the three distance limits, and CO and $PM_{2.5}$ were associated with LBW risks at the higher distance limits in the logistic models. SO_2 was linked to higher birth weights within 5 and 10km distance limits in the linear model, but only within 10km in the logistic model. The associations between birth weight and the trimester-level exposures to criteria pollutants were similar to that between full pregnancy pollutant exposures and birth weight, although trimester effects were reversed or attenuated for some pollutants, such as CO, NO_2, PM_{10}, and coarse PM during the second trimester (Table 3.5) . Overall, the birth weight differences were slightly stronger for the full pregnancy exposures.

Figure 3.1 displays linear model results (within 10 km monitor distance) for each criteria air pollutant alone and also after co-pollutant adjustment for those pollutants with a level of covariation under 70%. Results for all pollutants considered in the multivariate analysis were robust to co-pollutant adjustment remaining statistically significant in all cases, except for CO where effect estimates became insignificant with the addition of PM_{10} and $PM_{2.5}$. Results were also robust across the different distance limits (data not shown).

Based on previous studies we assessed models for interactions by race and neighborhood level poverty rate, but did not find any evidence of effect modification by either of these area or individual level demographic measures (data not shown).

Preterm Birth Results

Multivariate modeling showed that risk of preterm birth was associated with male infant sex; Black, Asian and Hispanic mothers; lower maternal educational attainment; lower parity; late

initiation of prenatal care; being unmarried and living in neighborhoods with lower educational attainment; lower home ownership rates; and higher rates of poverty and unemployment (Table 3.6).

Multivariate results for the association between pollution exposures and risk of preterm delivery are shown with air pollution exposures for different radii distance from an air monitor (Table 3.7). Pollution models were adjusted for all of the above factors as well as infant sex, geocode assignment, calendar year, and season of birth. NO_2, O_3, PM_{10}, $PM_{2.5}$ and coarse PM were consistently associated with an increased risk of preterm delivery within all three different distance limits and CO was linked to preterm birth risk within 5 and 10 kilometer distance limits. SO_2 was associated with preterm birth only at the 3 kilometer radius limit. The association between trimester-specific pollutant exposure and risk of preterm birth was similar to the effects seen for full pregnancy exposures, except that some trimester-specific effects were attenuated for $PM_{2.5}$ and coarse PM. For these pollutants, stronger effects were observed for the full pregnancy exposures.

Figure 3.2 displays logistic modeling results (within a 10 kilometer radii distance) for each criteria pollutant alone, and also after co-pollutant adjustments. Results for the gaseous pollutants were mixed. For CO and ozone, the addition of co-pollutants attenuated effect estimates, in some cases making the effect disappear. For NO_2, the effects remained significant except when $PM_{2.5}$ was added to the model. Conversely, effect estimates for the particulate pollutants considered in the multivariate analysis were robust to co-pollutant adjustment remaining statistically significant in all cases. Results were also robust across the different distance limits (data not shown).

Similar to the LBW analysis, we examined potential interactions of pollution effects on risk of preterm delivery by race and neighborhood- level poverty rate, but we did not find any evidence of effect modification by either of these area or individual- level demographic measures (data not shown).

Discussion

Consistent with prior literature, we have shown a modest relationship between ambient criteria air pollutant exposure ($PM_{2.5}$, PM_{10}, coarse PM, CO, NO_2 and O_3) and birth weight among full-term infants as well as preterm birth. Overall, these associations between increasing pollutant exposures and decrements in birth weight and risk of preterm delivery persisted during different trimesters of exposure, although the strongest effects were seen for exposures during the entire gestational period. Our study results are consistent with previous studies in California which found adverse birth weight effects (Basu, Woodruff et al. 2004; Parker, Woodruff et al. 2005; Salam, Millstein et al. 2005; Wilhelm and Ritz 2005; Parker and Woodruff 2008) and increased preterm birth risk for $PM_{2.5}$ (Ritz, Yu et al. 2000; Ritz, Wilhelm et al. 2007; Brauer, Lencar et al. 2008; Darrow, Klein et al. 2009),CO (Ritz and Yu 1999; Salam, Millstein et al. 2005; Wilhelm and Ritz 2005), and ozone (Salam, Millstein et al. 2005), although the timing of these effects varied in terms of trimester-specific or full gestational exposure. Although smaller particles have been the focus of regulatory and scientific attention for its impacts on health (Pope and Dockery

2006), results from this study confirm recent work indicating that exposure to coarse particles may adversely affect birth weight (Parker and Woodruff 2008). Results for NO_2 and PM_{10} also confirm previous study results in other areas, such as New England (Ha, Hong et al. 2001; Bell, Ebisu et al. 2007). Although prior studies have found evidence for differential effects of air pollution among different socioeconomic groups, such as maternal race (Bell, Ebisu et al. 2007) or neighborhood SES (Ponce, Hoggatt et al. 2005), we did not find such evidence of interaction in our analysis when we examined effect modification by maternal race or neighborhood- level poverty rate. Our future work will re-examine potential effect modification of air pollution birth- outcome relationships by individual and area-level SES factors in a larger population that includes births from several states with a broader range of pollutant burdens and neighborhood conditions.

Although we were able to control for individual and area- level factors, maternal smoking is not consistently reported on California birth records and its exclusion in our study may have changed our results. However, recent studies suggest that smoking does not significantly confound the association between ambient air pollution exposure and adverse perinatal outcomes such as infant mortality and preterm birth (Darrow, Woodruff et al. 2006; Ritz, Wilhelm et al. 2007). Another analysis examining the effect of maternal smoking on the association between particulate matter and birth weight using birth records from Arizona and Florida found minimal changes in the effect estimates for particulate matter exposure and infant birth weight after controlling for maternal smoking (Basu, Parker et al. 2003).

The negative effects of air pollutant exposures on birth weight remained robust to inclusion of other pollutants (except CO) and were consistent for the particulate pollutants in term of preterm delivery effects (less so for gaseous pollutants), although highly correlated pollutants were not included in these models. For example, PM_{10}, $PM_{2.5}$, and NO_2 were found to be highly correlated as well as CO and NO_2 and tend to come from common sources. Thus, this analysis cannot assess whether those pollutants linked to LBW could in fact be proxies for other pollutants with similar emission sources. Future work could deploy methods that better distinguish key common source pollutants that exert adverse effects on LBW. However, this single pollutant approach would not take into account the cumulative impact of exposures to multiple air pollutants, which may be important if in fact chemical mixtures lead to higher health risks than individual chemical constituents. A major source of both gaseous and particulate air pollutants is combustion, and one important area of future inquiry is to take a source-based approach to assessing health effects rather than isolating the impacts of individual pollutants. More can be done to analyze and develop source-specific measures, such as traffic density (Wilhelm and Ritz 2003; Ponce, Hoggatt et al. 2005), which could elucidate opportunities for exposure reduction to multiple pollutants (Woodruff, Parker et al. 2009).

We assessed the consistency of our results by using different distance limits for the births we examined (3, 5 and 10 kilometers). Results for our pollutants remained statistically significant in the linear (birth weight) and logistic (preterm birth) models and results varied more for the logistic models for birth weight effects. Other studies have sought to examine the impact of exposure assessment methods on effect estimates of air pollution impacts on health outcomes. For example, a Los Angeles study demonstrated how within-city gradients of $PM_{2.5}$ exposures produced larger effect estimates for mortality than models comparing the impact of $PM_{2.5}$ across

communities (Jerrett, Burnett et al. 2005). This issue has also been examined in relation to perinatal outcomes in a California study that found that different air pollution metrics (county-wide average, nearest monitor, distance-weighted average of monitors < 5 miles of the mother's residence) affected estimates for air pollution effects on birth weight (Basu, Woodruff et al. 2004), with greater associations between birth weight and $PM_{2.5}$ exposures averaged over counties rather than using monitors closer to a mother's residence. The reasons for this difference remain unclear, however. Nevertheless, these studies suggest that air pollution exposures can vary considerably at smaller scales and that this variation can affect the size of effect estimates.

Although we sought to examine this issue by estimating pollutant effects within different distance limits to monitors, we were limited to the tract and ZIP code-levels which prohibited finer scale assessments of geographical variations in exposure. We averaged weekly exposure estimates to derive trimester-specific and full gestation exposures, so our analysis does not account for differences in the distribution of exposures during the course of a pregnancy. The averaging procedure used to derive exposure measures would not reflect short-term exposures to very high pollutant levels. We used ambient monitoring as a surrogate for personal exposure during the course of pregnancy, which does not account for indoor pollutant levels, occupational exposures, transportation-associated exposures, or other activities not occurring in one's home neighborhood. Such non-differential misclassification of exposure would likely weaken effect estimates toward the null, making the results from this study an underestimate of the true underlying effects. Additionally, birth records only record maternal address at the time of delivery, so we could not account for residential mobility during pregnancy. Studies vary in their estimates of how important the impact of residential mobility may be on effect estimates of air pollution on birth outcomes (Woodruff, Parker et al. 2009). Any misclassification due to this trend is likely to be larger during the earlier stages of pregnancy than during the time period closer to delivery. Also, approximately 1% of the births were excluded from the analysis because they could not be assigned a tract or ZCTA. The effects of these exclusions could have unpredictable impacts on our estimate of the effect of air pollutant exposures on birth weight.

The majority of air pollution and birth outcome studies have focused on criteria air pollutants that are routinely monitored and regulated with national standards, yet there are other pollutants, such as air toxics, that may also be of interest due to their respiratory, reproductive and developmental effects. There is only sparse monitoring data available for air toxics, although modeled annual average estimates are now available for several periods (U.S. EPA 2009). Future studies should expand to include impacts from other categories of pollutants that may exert harm during pregnancy.

Conclusions

In summary, this study indicates that maternal exposure to air pollution may result in modestly lower infant birth weight and higher risk of preterm delivery. Although the effects are smaller than many other exposures, such as smoking, the ubiquity of air pollution exposures, the responsiveness of pollutant levels to planning and regulation efforts, and the fact that the highest pollution levels in California are lower than those regularly experienced in other countries suggests that the potential implications may be important for infant health and development.

Table 3.1: Characteristics of Singleton Births in Study Sample Compared with Overall Population of Singleton Births, California (1996-2006)

	Total Eligible Singleton Births (n=4,776,090)	Study Sample (n=3,545,177)
low birth weight (<2,500 grams)	2.3%	2.3%
maternal age (years)		
15 to 19	9.9%	10.2%
20 to 34	74.2%	74.3%
35 to 49	15.8%	15.4%
educational attainment		
none to 11th grade	30.2%	31.5%
12th grade	27.6%	27.6%
1-3 years college	19.8%	19.4%
4+ years college	22.4%	21.4%
marital status		
married	42.8%	42.0%
not married	22.5%	23.7%
not on form	27.9%	27.4%
missing	6.8%	6.9%
maternal race/ethnicity		
Hispanic	49.6%	51.5%
Black (non-Hispanic)	5.8%	6.3%
American Indian/Alaska Native (non-Hispanic)	0.4%	0.3%
Asian Pacific Islander (non-Hispanic)	11.9%	12.0%
Other Race (non-Hispanic)	0.0%	0.0%
White (non-Hispanic)	32.2%	29.6%
missing	0.1%	0.1%
maternal birthplace		
Mexico	27.6%	28.6%
other or unknown foreign country	18.5%	19.2%
US and her territories	53.8%	52.1%
missing	0.1%	0.1%
parity		
first live birth	39.5%	39.7%
maternal risk factors		
anemia, diabetes, hypertension and/or herpes	4.5%	4.4%
none of the above	86.1%	86.4%
missing	9.4%	9.2%
Kotelchuk index		
no prenatal care	1.6%	1.8%
inadequate	9.0%	9.0%
intermediate	11.8%	11.7%
adequate	44.1%	43.8%
more than adequate	33.5%	33.7%
insufficient information	0.1%	0.1%

Eligible singleton births include singleton births with a gestational age of 37-44 weeks and information for birth weight, sex, date of birth, maternal educational attainment, parity, and a maternal age of 9 to 49 years old.

Study sample includes only those eligible singleton births within 10km of an air monitor active throughout pregnancy.

Table 3.2: Distribution of Pollutant Exposures Averaged Over Length of Pregnancy, as Measured within 10km of Mother's Residential Geocode

pollutant	unit	N	mean	SD	interquartile range
CO	ppm	2,853,245	0.87	0.45	0.56 - 1.09
NO_2	ppm	2,808,662	0.024	0.009	0.017 - 0.031
O_3	ppm	3,303,834	0.024	0.006	0.019 - 0.027
SO_2	ppm	1,167,449	0.0021	0.0011	0.0012 - 0.0028
PM_{10}	$\mu g/m^3$	1,778,579	31.4	11.2	22.6 - 38.7
$PM_{2.5}$	$\mu g/m^3$	1,402,622	16.7	5.5	12.0 - 21.0
PM_{coarse}	$\mu g/m^3$	740,885	15.7	7.5	11.0 - 18.1

Table 3.3: Multivariate Modeling Results for Difference in Birth Weight for Selected Non-Pollution Variables

	change in birth weight (grams) (95% confidence interval)		
mother's race/ethnicity[a]			
Hispanic	-41.8 (	-43.1 ,	-40.6)
nH-Black	-161.5 (	-163.4 ,	-159.6)
nH-AIAN	15.7 (	9.1 ,	22.3)
nH-Asian	-156.8 (	-158.5 ,	-155.1)
nH-Other/Multi	-79.3 (	-97.8 ,	-60.8)
nH-White	(ref)		
mother's birthplace[a]			
Mexico	11.1 (	9.9 ,	12.4)
other foreign	-31.3 (	-32.7 ,	-30.0)
United States	(ref)		
mother's age group			
9 to 14	-57.9 (	-68.4 ,	-47.3)
15 to 19	-43.5 (	-45.1 ,	-42.0)
20 to 34	(ref)		
35 to 49	21.4 (	20.3 ,	22.6)
mother's educational attainment			
less than high school	-16.5 (	-17.7 ,	-15.4)
high school	(ref)		
1-3 years college	17.5 (	16.3 ,	18.7)
4+ years college	20.8 (	19.4 ,	22.1)
parity			
first birth	-108.3 (	-109.2 ,	-107.4)
2nd and higher	(ref)		
mother's marital status[a]			
married	(ref)		
not married	-28.6 (	-29.8 ,	-27.4)
Kotelchuk index of prenatal care[a]			
no prenatal care	-17.8 (	-21.1 ,	-14.5)
inadequate	-50.8 (	-52.3 ,	-49.2)
intermediate	-33.3 (	-34.8 ,	-31.9)
adequate	-7.8 (	-8.7 ,	-6.8)
more than adequate	(ref)		
neighborhood poverty rate[a]			
30% and higher	-18.3 (	-21.3 ,	-15.2)
20% to 30%	-9.3 (	-11.9 ,	-6.8)
10% to 20%	-4.7 (	-6.7 ,	-2.7)
5% to 10%	3.2 (	1.6 ,	4.8)
under 5%	(ref)		
neighborhood owner-occupied housing rate[a]			
20% and lower	5.8 (	3.5 ,	8.1)
20% to 40%	8.7 (	6.8 ,	10.6)
40% to 60%	6.0 (	4.3 ,	7.7)
60% to 80%	3.9 (	2.4 ,	5.4)
80% and higher	(ref)		
neighborhood unemployment rate[a]			
15% and higher	-1.8 (	-4.2 ,	0.6)
10% to 15%	-4.6 (	-6.5 ,	-2.8)
7.5% to 10%	-6.2 (	-7.9 ,	-4.6)
5% to 7.5%	-2.3 (	-3.7 ,	-1.0)
under 5%	(ref)		
neighborhood low education (% aged 25+ with less than a HS education)[a]			
20% and higher	-8.4 (	-10.9 ,	-5.9)
15% to 20%	-3.2 (	-5.2 ,	-1.2)
10% to 15%	-2.6 (	-4.3 ,	-0.9)
5% to 10%	-0.1 (	-1.4 ,	1.2)
under 5%	(ref)		

Table 3.4: Multivariate model results for change in birth weight associated with full pregnancy pollutant exposures measured at 3km, 5km, and 10km monitor distance

	change in birth weight (grams)	odds ratio for low birth weight (<2,500g)
	95% confidence interval	95% confidence interval
CO, per ppm		
at 3km	-2.5 (-5.4 , 0.3)	1.02 (0.98 , 1.07)
at 5km	-5.9 (-7.8 , -3.9)	1.06 (1.03 , 1.09)
at 10km	-5.4 (-6.8 , -4.1)	1.04 (1.02 , 1.06)
NO$_2$, per pphm		
at 3km	-8.3 (-9.6 , -7.0)	1.03 (1.01 , 1.05)
at 5km	-9.7 (-10.6 , -8.8)	1.04 (1.03 , 1.05)
at 10km	-9.0 (-9.6 , -8.4)	1.03 (1.02 , 1.04)
O$_3$, per pphm		
at 3km	-8.9 (-10.6 , -7.1)	1.01 (0.98 , 1.03)
at 5km	-7.0 (-8.2 , -5.8)	0.98 (0.97 , 1.00)
at 10km	-5.7 (-6.6 , -4.9)	0.98 (0.97 , 0.99)
SO$_2$, per ppb		
at 3km	1.7 (-0.3 , 3.8)	1.02 (0.99 , 1.06)
at 5km	2.4 (1.0 , 3.7)	1.01 (0.99 , 1.03)
at 10km	3.1 (2.3 , 3.8)	1.01 (1.00 , 1.02)
PM$_{10}$, per 10μg/m^3		
at 3km	-5.5 (-6.9 , -4.1)	1.00 (0.97 , 1.02)
at 5km	-7.6 (-8.5 , -6.7)	1.00 (0.98 , 1.01)
at 10km	-7.2 (-7.9 , -6.6)	1.00 (0.99 , 1.01)
PM$_{2.5}$, per 10μg/m^3		
at 3km	-9.2 (-12.5 , -5.9)	1.04 (0.99 , 1.09)
at 5km	-11.4 (-13.5 , -9.3)	1.05 (1.02 , 1.08)
at 10km	-12.8 (-14.3 , -11.3)	1.04 (1.02 , 1.07)
PM$_{coarse}$, per 10μg/m^3		
at 3km	-9.4 (-12.8 , -6.0)	1.00 (0.95 , 1.05)
at 5km	-10.1 (-12.2 , -8.0)	0.99 (0.96 , 1.02)
at 10km	-9.3 (-10.7 , -7.9)	0.99 (0.97 , 1.01)

Models adjusted for infant sex, gestational age, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, access to prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home

ownership, educational attainment, unemployment rate)

Table 3.5: Effect of Trimester-Specific Pollutant Exposures on Birth Weight, in grams (95% confidence interval)

	first trimester[a] exposure	second trimester[a] exposure	third trimester[a] exposure
CO, per ppm			
at 3km	-2.2 (-5.0 , 0.7)	5.3 (1.7 , 8.8)	-6.7 (-9.8 , -3.6)
at 5km	-2.4 (-4.4 , -0.4)	3.2 (0.8 , 5.6)	-7.7 (-9.8 , -5.6)
at 10km	-1.9 (-3.3 , -0.6)	2.5 (0.9 , 4.2)	-7.0 (-8.4 , -5.5)
NO_2, per pphm			
at 3km	-2.4 (-4.4 , -0.5)	1.8 (-0.8 , 4.3)	-8.1 (-10.2 , -6.1)
at 5km	-3.1 (-4.4 , -1.8)	0.9 (-0.8 , 2.5)	-7.9 (-9.2 , -6.5)
at 10km	-3.0 (-3.9 , -2.1)	0.6 (-0.6 , 1.7)	-7.0 (-7.9 , -6.0)
O_3, per pphm			
at 3km	-2.9 (-4.4 , -1.5)	-3.1 (-4.6 , -1.6)	-3.0 (-4.4 , -1.5)
at 5km	-2.7 (-3.7 , -1.7)	-2.2 (-3.2 , -1.1)	-2.4 (-3.4 , -1.4)
at 10km	-2.1 (-2.9 , -1.4)	-2.3 (-3.1 , -1.5)	-1.3 (-2.1 , -0.6)
SO_2, per ppb			
at 3km	0.8 (-1.8 , 3.3)	0.4 (-2.7 , 3.5)	0.6 (-1.9 , 3.2)
at 5km	1.8 (0.3 , 3.4)	0.1 (-1.7 , 2.0)	0.4 (-1.1 , 2.0)
at 10km	2.5 (1.6 , 3.4)	-0.1 (-1.1 , 0.9)	0.7 (-0.2 , 1.5)
PM_{10}, per 10µg/m³			
at 3km	-2.6 (-4.3 , -0.9)	-0.3 (-2.2 , 1.6)	-3.1 (-4.8 , -1.3)
at 5km	-2.7 (-3.8 , -1.7)	-1.1 (-2.3 , 0.1)	-4.1 (-5.2 , -3.0)
at 10km	-2.3 (-3.0 , -1.6)	-1.5 (-2.3 , -0.7)	-3.7 (-4.4 , -3.0)
$PM_{2.5}$, per 10µg/m³			
at 3km	-6.9 (-9.6 , -4.2)	-0.5 (-3.6 , 2.6)	-2.4 (-5.2 , 0.4)
at 5km	-6.1 (-7.8 , -4.3)	-2.2 (-4.2 , -0.3)	-3.6 (-5.5 , -1.8)
at 10km	-6.0 (-7.3 , -4.8)	-2.6 (-4.0 , -1.3)	-4.7 (-6.0 , -3.5)
PM_{coarse}, per 10µg/m³			
at 3km	-3.5 (-7.1 , 0.0)	0.3 (-3.5 , 4.1)	-6.7 (-10.1 , -3.3)
at 5km	-4.2 (-6.3 , -2.0)	-1.2 (-3.6 , 1.1)	-5.0 (-7.1 , -2.9)
at 10km	-3.4 (-4.9 , -2.0)	-1.0 (-2.5 , 0.5)	-5.1 (-6.4 , -3.8)

[a]first trimester: first 16 weeks after last menstrual period, second trimester: weeks 17 to 28, third trimester: week 29 to delivery.

Models adjusted for infant sex, gestational age, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, access to prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate)

Table 3.6: Multivariate Modeling Results for Risk of Preterm Delivery Associated with Selected Non-Pollution Variables

Variable	Category	Odds Ratio for Preterm Delivery (24-34 weeks)	95% Confidence Interval	
Mother's age				
	9 to 14	2.71	2.43	3.02
	15 to 19	1.48	1.46	1.51
	20 to 34	(ref)		
	35 to 49	1.32	1.29	1.34
Race/ethnicity				
	Hispanic	1.49	1.47	1.52
	nh Black	2.65	2.59	2.72
	nh AIAN	1.68	1.52	1.87
	nh API	1.34	1.31	1.37
	nh Other	1.44	1.06	1.95
	nh White	(ref)		
	unknown	*0.96*	*0.78*	*1.20*
Mother's Place of Birth				
	Mexico	0.98	0.97	1.00
	Other/Unknown Foreign	0.93	0.91	0.95
	US and territories	(ref)		
Mother's marital status				
	married	(ref)		
	not married	1.63	1.61	1.66
Mother's educational attainment				
	less than HS	1.14	1.12	1.15
	high school	(ref)		
	1-3 years college	0.86	0.85	0.88
	4+ years college	0.62	0.60	0.63
Maternal parity				
	first birth	0.99	0.97	1.00
	2nd to 15th birth	(ref)		
Trimester prenatal care began				
	first	(ref)		
	second	1.51	1.48	1.54
	third or no prenatal care	2.02	1.96	2.08
Any pregnancy risk factor (anemia, diabetes, hypertension and/or herpes)				
	yes	2.83	2.77	2.89
	no	(ref)		
neighborhood poverty rate				
	30% and higher	1.94	1.88	2.00
	20% to 30%	1.64	1.60	1.69
	10% to 20%	1.40	1.36	1.44
	5% to 10%	1.15	1.12	1.19
	under 5%	(ref)		
neighborhood owner occupancy rate				
	20% and lower	1.46	1.41	1.51
	20% to 40%	1.45	1.41	1.49
	40% to 60%	1.36	1.33	1.40
	60% to 80%	1.20	1.17	1.23
	80% and higher	(ref)		
neighborhood unemployment rate				
	15% and higher	1.88	1.84	1.93
	10% to 15%	1.63	1.60	1.66
	7.5% to 10%	1.41	1.38	1.44
	5% to 7.5%	1.27	1.25	1.30
	under 5%	(ref)		
neighborhood low education rate				
	20% and higher	1.77	1.72	1.81
	15% to 20%	1.66	1.62	1.69
	10% to 15%	1.51	1.48	1.54
	5% to 10%	1.34	1.31	1.37
	under 5%	(ref)		

Table 3.7: Multivariate model results for risk of preterm delivery associated with full pregnancy pollutant exposures measured at 3km, 5km, and 10km monitor distance

| Variable	Category	Odds Ratio for Preterm Delivery (24-34 weeks)	95% Confidence Interval	
CO, per ppm				
	3 km	1.03	0.99	1.07
	5 km	1.08	1.05	1.11
	10 km	1.07	1.05	1.09
NO_2, per pphm				
	3 km	1.03	1.01	1.05
	5 km	1.03	1.02	1.05
	10 km	1.03	1.02	1.04
O_3, per pphm				
	3 km	1.04	1.01	1.06
	5 km	1.04	1.02	1.06
	10 km	1.02	1.01	1.04
SO_2, per pphm				
	3 km	1.41	1.06	1.88
	5 km	1.10	0.92	1.33
	10 km	1.04	0.94	1.16
PM_{10}, per $10\mu g/m^3$				
	3 km	1.05	1.03	1.07
	5 km	1.06	1.05	1.07
	10 km	1.06	1.05	1.07
$PM_{2.5}$, per $10\mu g/m^3$				
	3 km	1.16	1.11	1.21
	5 km	1.16	1.13	1.19
	10 km	1.14	1.11	1.16
PM_{coarse}, per $10\mu g/m^3$				
	3 km	1.12	1.08	1.17
	5 km	1.14	1.11	1.17
	10 km	1.11	1.09	1.13

Models adjusted for infant sex, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, initiation of prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate)

Table 3.8: Effect of Trimester-Specific Pollutant Exposures on Risk of Preterm Delivery (24-34 weeks)

	full pregnancy exposures			first trimester exposures			second trimester exposures		
CO, per ppm	Odds Ratio	95% Confidence Interval		Odds Ratio	95% Confience Interval		Odds Ratio	95% Confience Interval	
03 km	1.03	0.99	1.07	1.05	1.02	1.08	0.98	0.95	1.02
05 km	1.08	1.05	1.11	1.08	1.06	1.10	1.02	1.00	1.04
10 km	1.07	1.05	1.09	1.08	1.06	1.09	1.00	0.99	1.02
NO_2, per pphm									
03 km	1.03	1.01	1.05	1.02	1.00	1.03	1.01	0.99	1.03
05 km	1.03	1.02	1.05	1.02	1.01	1.04	1.01	1.00	1.02
10 km	1.03	1.02	1.04	1.03	1.02	1.04	1.01	1.00	1.01
O_3, per pphm									
03 km	1.04	1.01	1.06	1.02	1.00	1.04	1.02	1.00	1.04
05 km	1.04	1.02	1.06	1.02	1.00	1.03	1.03	1.02	1.04
10 km	1.02	1.01	1.04	1.00	0.99	1.01	1.03	1.02	1.03
SO_2, per pphm									
03 km	1.41	1.06	1.88	1.20	0.94	1.52	1.27	1.01	1.61
05 km	1.10	0.92	1.33	1.07	0.92	1.25	1.06	0.91	1.23
10 km	1.04	0.94	1.16	1.02	0.93	1.11	0.97	0.89	1.05
PM_{10}, per $10\mu g/m^3$									
03 km	1.05	1.03	1.07	1.04	1.03	1.06	1.02	1.01	1.03
05 km	1.06	1.05	1.07	1.04	1.03	1.05	1.03	1.02	1.04
10 km	1.06	1.05	1.07	1.04	1.03	1.04	1.02	1.02	1.03
$PM_{2.5}$, per $10\mu g/m^3$									
03 km	1.16	1.11	1.21	1.07	1.04	1.11	1.07	1.04	1.10
05 km	1.16	1.13	1.19	1.08	1.05	1.10	1.07	1.05	1.09
10 km	1.14	1.11	1.16	1.06	1.05	1.08	1.04	1.03	1.06
PM_{coarse}, per $10\mu g/m^3$									
03 km	1.12	1.08	1.17	1.06	1.03	1.10	1.04	1.02	1.07
05 km	1.14	1.11	1.17	1.08	1.06	1.10	1.07	1.05	1.09
10 km	1.11	1.09	1.13	1.07	1.05	1.08	1.06	1.05	1.08

Models adjusted for infant sex, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, initiation of prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate)
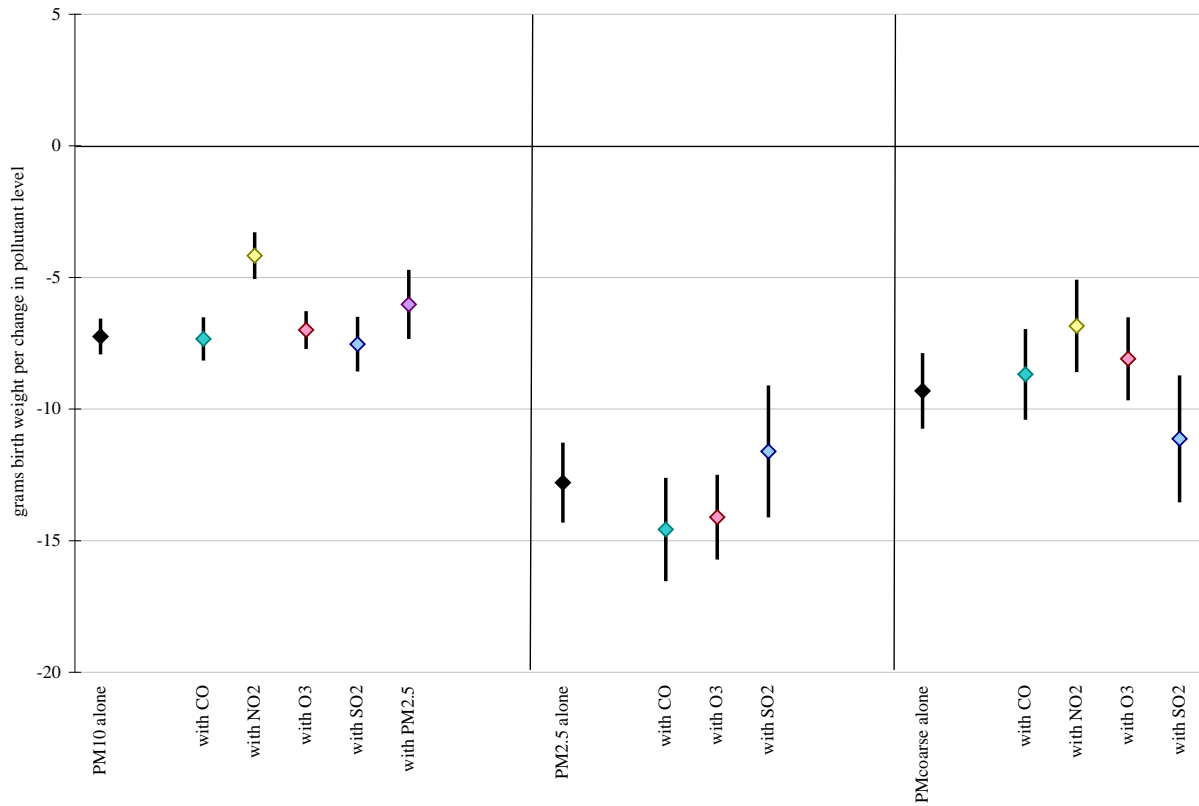
Figure 3.1a: Difference in birth weight in grams associated with full pregnancy gaseous pollutant exposures for births within 10 km monitor distance, single, and two-pollutant linear models (95% confidence interval)
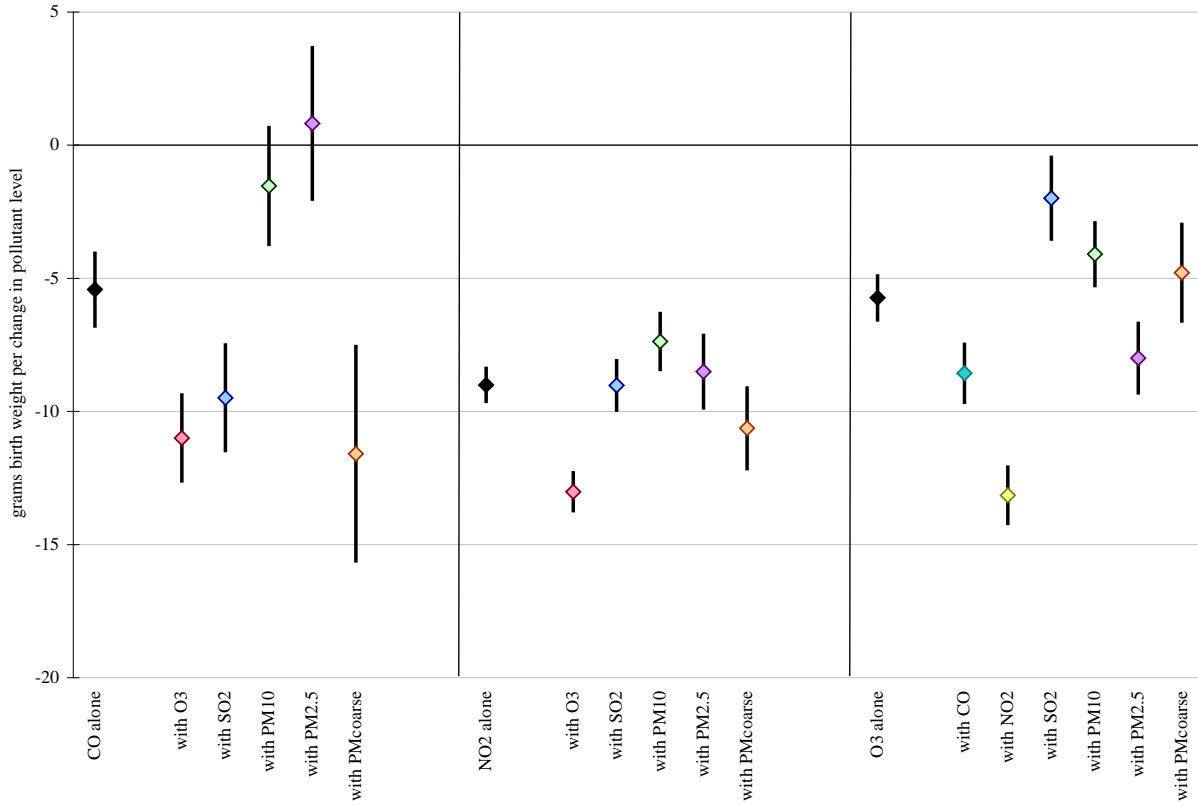
Models adjusted for infant sex, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, initiation of prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate).
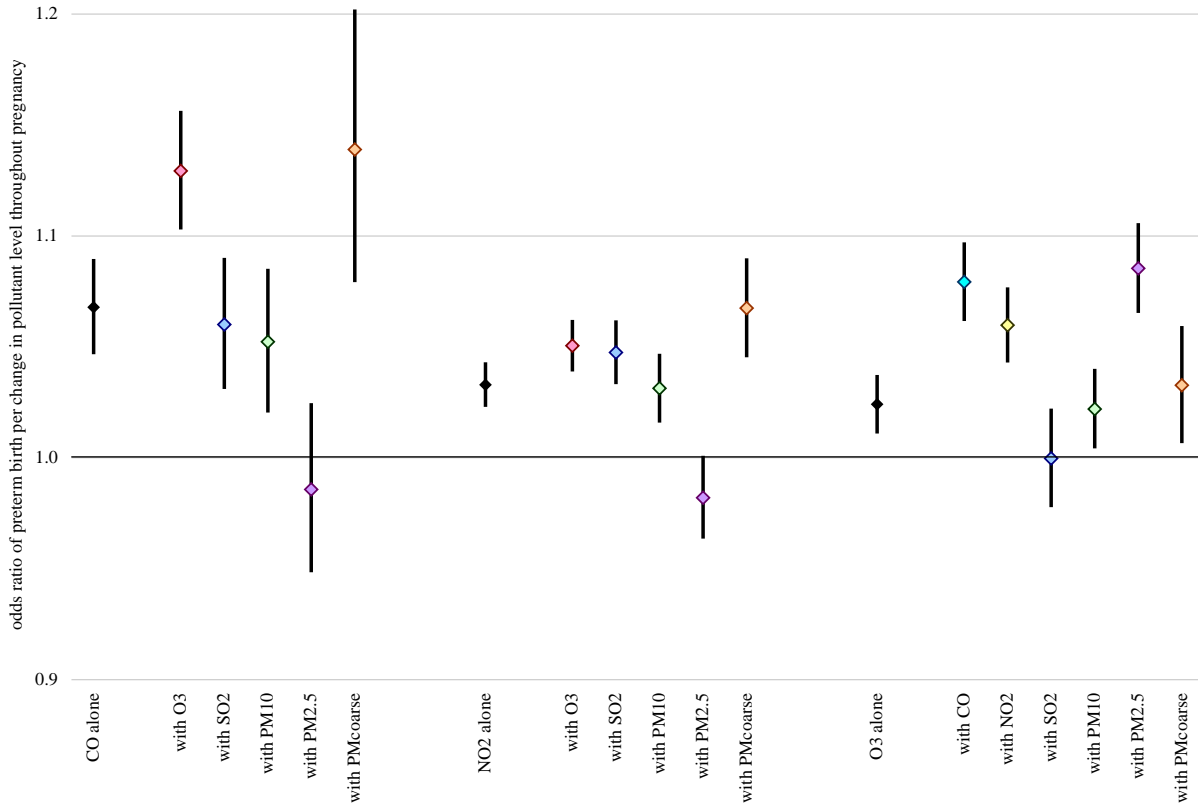
Figure 3.1b: Difference in birth weight in grams associated with full pregnancy particulate matter exposures for births within 10 km monitor distance, single, and two-pollutant linear models (95% confidence interval)

Models adjusted for infant sex, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, initiation of prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate).

Figure 3.2a: Odds ratio for pre-term birth (29-34 vs. 39-44 weeks) associated with full pregnancy gaseous pollutant exposures for births within 10 km monitor distance, single and two-pollutant linear models (95% confidence interval)
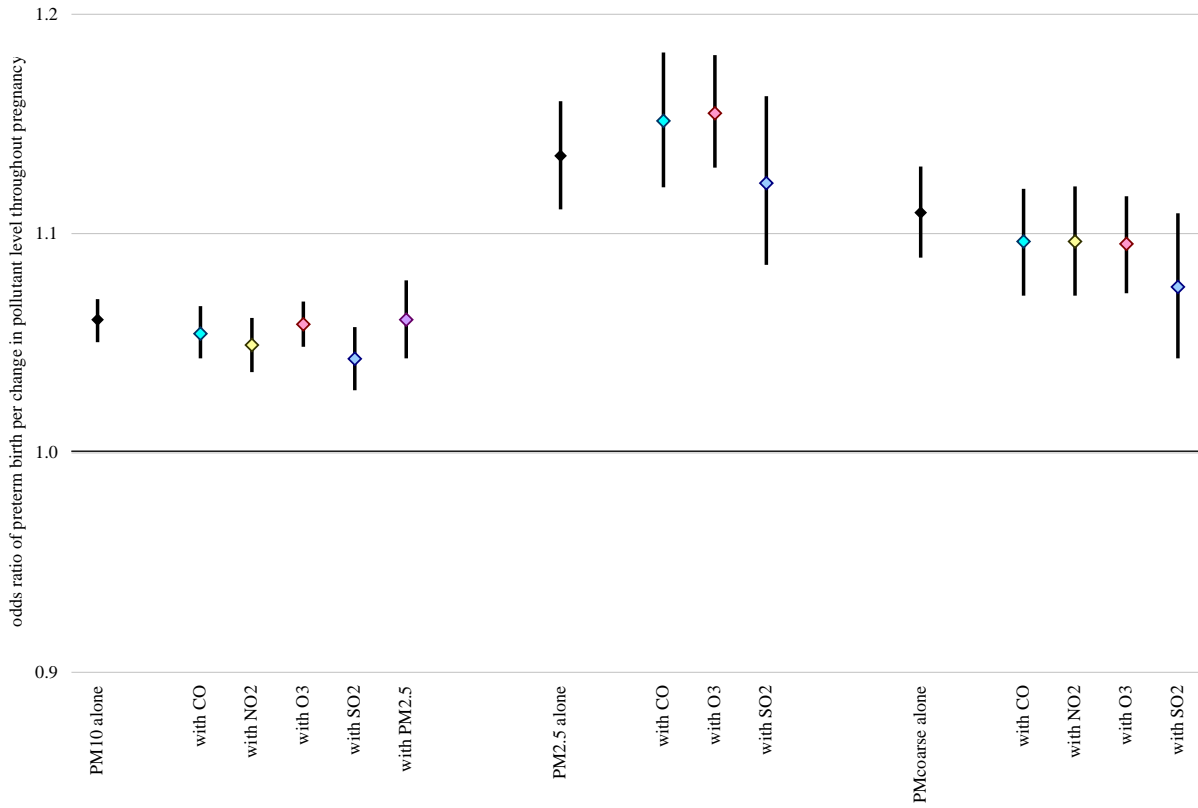

Models adjusted for infant sex, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, initiation of prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate).

Figure 3.2b: Odds ratio for preterm birth (29-34 vs. 39-44 weeks) associated with full pregnancy particulate matter exposures for births within 10 km monitor distance, single and two-pollutant linear models (95% confidence interval)

Models adjusted for infant sex, season and year of birth, parity, maternal factors (race/ethnicity, educational attainment, marital status, initiation of prenatal care, birth place, age) and neighborhood SES measures (poverty rate, home ownership, educational attainment, unemployment rate).

Chapter 4. Implementing an Environmental Justice Screening Method (EJSM)

Introduction to the problem and previous work

As noted earlier, numerous studies in California over the past decade or more have presented a body of strong evidence documenting patterns of disproportionate exposure to both air toxics exposure and associated health risks focused primarily on communities of color and lower-income groups (Burke 1993; Pulido, Sidawi et al. 1996; Sadd, Pastor et al. 1999; Morello-Frosch, Pastor et al. 2001; Morello-Frosch 2002; Morello-Frosch RA 2002). This research is consistent with the long-held concerns of community-based environmental organizations and people living in the most impacted communities. Such organizations frequently note that inequalities in exposure exist for many different pollutants and types of hazards, and that many neighborhoods bear the combined - or cumulative - burden of air emissions from numerous industrial facilities and land uses, as well as emissions from mobile sources on high volume roads and freeways, and emissions associated with smaller facilities that operate either illegally or are not subject to regulatory oversight. Such concentrations may be of particular concern where the exposures affect populations that are, because of age or chronic health conditions, particularly sensitive to air pollution - termed "sensitive receptors" by the California EPA (CARB 2005). These same communities also face challenges associated with low social and economic status, as well as psychosocial stressors, which make it more difficult to cope with exposure and health disparities (Gee and Payne-Sturges 2004; Morello-Frosch and Shenassa 2006).

This problem cannot be fully addressed by current regulatory practice, which uses traditional methods of risk assessment to decide, for example, whether a specific polluting facility can operate under existing law. Risk is typically calculated using single stressors and reported on a chemical-by-chemical, medium-by-medium, and source-by-source basis. Each regulatory authority considers only those projects or facilities within its mandates and jurisdiction, with no integration across jurisdictions. A consequence of framing decisions on magnitude of risk and regulation in this way ignores the fact that, in many communities, residents are exposed to multiple risks with cumulative impacts. The one dimensional facility-by-facility regulatory approach ignores the reality of the multiplicity of factors that affect these communities and, in doing so, fails to adequately protect their health and safety.

Environmental justice advocates have encouraged policymakers and government agencies to look for solutions that take a more holistic and precautionary approach to environmental regulation and public health protection, that addresses the cumulative[13] impacts of multiple

---

[13] The term "cumulative" is widely used in environmental impact assessment to recognize the fact that single pollutants or stressors considered in isolation are inadequate to determine actual health impacts. "Cumulative" is used here within the context of air quality; the term "cumulative impacts" describes the sum of health and nuisance impacts related to a neighborhood's cumulative air pollution emissions from all sources - those that individually comply with existing air quality regulations, and those that do not. The California Environmental Protection Agency and South Coast Air Quality Management District in Greater Los Angeles have begun a process to assess and reduce "cumulative environmental health risk."

hazardous exposures. Cumulative impacts refer to the health effects of multiple exposures to chemical and non-chemical stressors that occur in communities over time and place. These multiple stressors can combine additively or synergistically over the life course to produce ill health (Gee and Payne-Sturges 2004; Morello-Frosch and Shenassa 2006; DeFur, Evans et al. 2007; Clougherty and Kubzansky In Press). Part of developing a more comprehensive approach to regulatory decision-making requires incorporating metrics of social vulnerability. There is mounting evidence that various aspects of social inequality, such as geographic segregation and discrimination, economic disparities, and low levels of civic engagement, contribute to the disparate pattern of environmental hazards exposure and health risk for the communities of color and the poor. For example, air pollution health risk can be exacerbated at the individual or community level by differences in health/disease status, age and genetic variation, and some socioeconomic and community-level factors (Morello-Frosch, Pastor et al. 2002; Morello-Frosch 2002; Gee and Payne-Sturges 2004; Ponce, Hoggatt et al. 2005; Morello-Frosch and Lopez 2006; Morello-Frosch and Shenassa 2006; Clougherty, Levy et al. 2007; Clougherty and Kubzansky In Press). Low SES likewise may limit options for individuals to address and cope with environmental and health problems, for example, the effect of lower income on the likelihood of having health insurance or the impact of language limitations on engagement with public officials (Pastor, et al, 2001). While efforts are beginning to unfold at the federal and state levels to incorporate social vulnerability and non-chemical stressors into risk assessments (National Academy of Sciences 2009; OEHHA 2009), addressing cumulative impacts requires alternatives to traditional risk-based decision making, that encompass more precautionary and place-based approaches to protecting vulnerable communities.

Incorporating these numerous and significant indicators of community and health vulnerability into a comprehensive understanding of environmental health disparities will allow policy makers to identify not only whether a community may be overburdened but also the capacity of that community to cope with the higher exposure and risks, voice its concerns, and seek remedies (Cal-EPA Advisory Committee on Environmental Justice 2003; National Environmental Justice Advisory Council 2004). The term "cumulative impacts" has become a working concept in recent years for regulatory purposes in California. In 2005, building on the recommendations from their Environmental Justice Advisory Committee (CEJAC, 2005), the California EPA has proposed a working definition for "Multi-Media Cumulative Impacts" in its 2005 Environmental Justice Action Plan.[14] The California Office of Environmental Health Hazard Assessment (OEHHA) maintains a Cumulative Impacts and Precautionary Approaches Work Group. The California South Coast Air Quality Management District (AQMD) has a Cumulative Impacts Working Group comprised of representatives of key stakeholder groups from industry and the environmental/community groups who provide input to AQMD staff on the feasibility of rulemaking regarding cumulative impacts of air pollution beyond current AQMD legal requirements.[15] The California Environmental Contaminant Biomonitoring Program[16] plans to

---

[14] "Cumulative Impacts means exposures, public health or environmental effects from the combined emissions and discharges in a geographic area, including environmental pollution from all sources, whether single or multi-media, routinely, accidentally or otherwise released. Impacts will take into account sensitive populations and socio-economic factors, where applicable and to the extent data are available" (See http://www.calepa.ca.gov/envJustice/ActionPlan/)

[15] In 2003, The AQMD Governing Board voted to proceed with a Cumulative Impacts Reduction Strategy. Some preliminary tract-based overlay maps are posted at http://www.aqmd.gov/rules/CIWG/maps.html

[16] Legislation SB 1379 (Perata-Ortiz), September 2006 (See http://oehha.ca.gov/multimedia/biomon/about.html)

measure baseline levels of several environmental chemicals in individuals to serve as a point of comparison for community exposures. In the near term, this data can help predict cumulative exposure and risk, and over a longer term may help measure cumulative health impacts.

For our project we sought to develop a method to screen for disparities in cumulative impacts related to environmental and social stressors. One specific aim was to develop a method that was scientifically valid, yet transparent so that it could be utilized by regulatory agencies as well as community organizations. This required using publicly available data sources and building an analytical structure that allowed flexibility in terms of incorporating new data layers and updating information as necessary. The method also needed to be adaptable to different decision-making scenarios—for example, siting and permitting, standard setting, targeting community outreach and engagement in regulatory issues, etc..

Our effort to develop such a cumulative impacts assessment – what we refer to as an environmental justice screening method (EJSM) - is pioneering in the field. There are, however, other efforts that have attempted to evaluate and map cumulative impacts. The first is more limited both in terms of geography and issue coverage. The Multiple Air Toxics Exposure Study III (MATES III) program of the South Coast AQMD is an intensive assessment of current ambient levels of carcinogenic toxic air pollutants and the attendant estimated health risks, focused on the Los Angeles Air Basin. The program combines air sampling and monitoring, pollutant concentration modeling, and analysis of air pollution emissions inventories to develop metrics of cumulative impact. The current version, MATES III[17], did this evaluation between April 2004 and March 2006, and it was extended because of the anticipated potential effects of near-record rainfall on the pollutant monitoring. The monitoring network was designed to address multiple stakeholder concerns about cumulative impacts by combining data from ten established sites with mobile monitoring stations sampling at neighborhood sites near numerous air toxics emission sources, as well as in areas where community members expressed concern about health risks from air pollution. Additional sampling was done near land uses associated with high levels of air pollution emissions, such as airports, railroads, warehouses, landfills, high-volume roadways, and industrial facilities. These data are interpolated to a relatively coarse 2- kilometer grid covering the air district and offshore areas. Maps generated using MATES III data identify both regional gradients in air quality and health risk, as well as potential "hot spots," expressed cumulatively across multiple pollutant chemicals and sources.

MATES III is an extraordinarily useful and carefully developed dataset, but it is not really a screening method. While it does consider carcinogenic toxic air pollutants, it does not consider other health impacts from particulate exposures nor does it consider indicators of neighborhood social vulnerability. Our screening method includes as indicators two data sources that similarly estimate exposures and health risk from monitored and modeled air pollution: the 2001 CARB cumulative estimated lifetime cancer risk associated with the Community Health Air Pollution Information System (CHAPIS) emissions inventory, and the 1999 U.S. EPA NATA.

---

[17] MATESIII is follows up two earlier generations of the program – MATES in 1986-1987, and MATESII conducted in 1998-1999.

The U.S. EPA is also developing a GIS-based cumulative impacts screening tool, known at the Environmental Justice Strategic Enforcement Assessment Tool (EJSEAT).[18] The goal of EJSEAT is to consistently identify areas with disproportionately high and adverse environmental health burdens, and to do so using nationally consistent data so that results are comparable across states and EPA regions. Similar to our EJSM that we describe below, EJSEAT defines a set of cumulative impacts indicator metrics organized into categories (demographic, environmental, compliance, and health impact), and then uses a simple mathematical methodology to scale these factors and apply to each location a composite EJSEAT score, mapping the results at the census tract level. EPA envisions EJSEAT as a means for both defining and identifying environmental justice communities, with a planned internal use to inform EPA's own enforcement activities.

EJSEAT incorporates key metrics of social vulnerability and therefore represents an advance in methodological approaches to develop tools to inform decision-making related to cumulative impacts and EJ concerns. There are a few variables that our EJSM shares with EJSEAT, including poverty, linguistic isolation, and age, as well as various measures of cancer and non-cancer-related health risk measures and ambient concentrations of ozone and particulates. There are, however, a number of ways in which our EJSM differs from EJSEAT.

First, because EJSEAT relies on datasets that have national coverage, its health vulnerability metrics include two variables (infant mortality and percent low birth weights) that are aggregated to the county level, because the National Center for Health Statistics does not release these data at the tract level due to privacy concerns.[19] Such county-level information may provide sufficient nuance in the eastern United States where counties are smaller, but applying a county-level birth outcomes measure across all tracts in a large county such as Los Angeles does not provide sufficient granularity for a useful screening method. Second, EJSEAT considers a range of enforcement measures, including inspections, violations, and actions, as well as a facility density measure. However, the reliability and meaning of this compliance data is suspect. For example, it is unclear whether more facility inspections reflects a pattern of compliance problems with a particular facility, or in fact reflects effective enforcement activities on the part of EPA. In addition, careful examination of this data has revealed that it shows little geographic variability nationwide. Finally, metrics of facility density must be carefully scrutinized to ensure that population density is also considered. Indeed, high-facility density may not be problematic in an area where population density is very low; we address this methodological issue in our EJSM by developing a measure of hazardous land use and facility proximity that reflects whether facilities are nearer to sensitive land uses such as residences, schools, and health facilities.

A third difference between our EJSM and EJSEAT has to do with the scoring. EJSEAT scores each variable using a "simple mathematical algorithm." This consists of ranking each variable by census tract *within a state* (recall, however, that in the case of the health indicators, all tracts in a county get the same score), taking the average of the scores by each of the four dimensions, and

---

[18] http://www.epa.gov/compliance/resources/policies/ej/ej-seat.html

[19] Although NCHS makes birth data publicly available at the county level, some states, including California, Georgia, Texas and Colorado, among others, do make birth data available by census tract. However, each state has its own IRB requirements for accessing this information and some states, like New York, severely restrict access to this information by outside researchers.

then adding those scores, scaling them, and generating a composite score. Our process is not dissimilar but we do not use a state-level comparison since, as noted in the Bay Area EJ analysis above, most researchers now suggest that comparisons of disparate environmental health impacts should be made at the regional or air basin level (Morello-Frosch, Pastor et al. 2002; Morello-Frosch 2002) . In addition, the spatial resolution of our screening method is much finer (census tracts or smaller), as compared to EJSEAT's use of counties and census tracts.

A fourth and important difference – and one related to the scoring – is the extent of public participation in the development of the EJSM. EJSEAT has had a formal scientific advisory committee that has offered comments after the method was initially developed. We have employed a process that involved significant community and scientific input prior to the development of the EJSM and utilized input and presentations throughout. This was by design as we wanted to develop a method that would have community and scientific credibility and input throughout.[20]

In any case, the EJSM that we have developed is a location-based evaluation of cumulative impacts that includes environmental and social vulnerability measures-- a set of 24 different indicator metrics organized along three dimensions: hazard proximity, air-related exposures and estimated health risk, and social vulnerability. It specifically combines GIS spatial analysis to define and map specific locations using a combination of land use and census polygon data, with analysis of the distribution of a comprehensive set of indicator metrics –proximity to air pollution hazards, health risk measures associated with specific air pollutants and pollutant types, and measures of social and health vulnerability – that together capture both cumulative impacts and vulnerability. The results are then expressed in the form of a cumulative impacts (CI) "score" that applies to a specific location, either census tracts or a unit of geography prepared expressly for use with the EJSM that combines census blocks and land use polygons.

The EJSM differs from other tools for cumulative impacts analysis in a number of important ways. It begins by integrating the two most relevant and geographically detailed layers of information – census polygons and detailed land use data from cities and regional planning agencies – to define the geography of where screening takes place. The EJSM then uses GIS spatial analysis techniques to evaluate proximity of air pollution hazards to residential populations and CARB-specified sensitive land uses. We use standard statistical programming language (Statistical Package for the Social Sciences (SPSS) and SAS) and utilize its ranking procedures to assign integer scores for the various indicator metrics to each geographic unit. We then sum these scores to allow relative comparisons among screened locations. The geographic comparison unit is the region, not the state, for reasons stated above. The scores are then exported into a database for GIS mapping and further spatial analysis.

---

[20] One of the authors is part of the advisory committee for EJSEAT preparing a report to EPA and the National Environmental Justice Advisory Council that analyzes the EJSEAT methodology and makes recommendations for improvements. It is likely that EJSEAT will go through some changes that will make it more similar to the EJSM, particularly by dropping the enforcement data and the county-based health data. There may also be the use of z-scores rather than simple ranking, something which will likely make little difference to the ultimate results and will require the utilization of logs of the variables in order to shift the data to a more normal distribution. Final decisions on these and other changes will be made internally by EPA.

As noted, the U.S. EPA's EJSEAT is the most similar to the EJSM, but there are several important differences. The EJSM uses a greater number of indicators because data quality in California is better than what exists nationwide. The relevance of all indicators selected for use in the EJSM has been documented through analysis of the research literature and as a result of consultation and guidance from stakeholders. The EJSM also makes comprehensive use inclusion of land use and health data unavailable to EJSEAT, and the spatial analysis is done at a significantly greater level of geographic detail. We use distance- and population- weighting to adjust scores to reflect regional variation, and we employ a simpler and more intuitive scoring system.

Materials/Methods

Data

The EJSM Method creates a base map for the geographic analysis by intersecting census blocks with land use polygons. This allows us to isolate those land uses where air quality hazards impact people in passive, non-occupational settings, and to combine this with data that is available at the census tract level about health outcomes, air-related health risk, and social factors. The EJSM reveals regional patterns of cumulative impact and highlights the most impacted and vulnerable communities. The CI scores themselves are not relative measures, representing how a given location compares to other similar areas within a large region in terms of the impact and vulnerability indicators employed.

To test the concept, we have fully developed the EJSM for a single region in which we had relatively good land use data, the six counties covered by the Southern California Association of Governments (SCAG). While the area has multiple air districts, it is mostly under the mandate of the South Coast AQMD and the air districts clearly intersect in terms of effects. Moreover, this region, in aggregate, hosts nearly half of the population of the entire state.[21]

We organize our indicators of cumulative impact and vulnerability in three categories: (a) proximity to air pollution hazards and land uses that are either associated with high levels of air pollution or host sensitive populations (CARB 2005); (b) exposure and health risk measures associated with specific air pollutants and pollutant types; and (c) measures of social and health vulnerability that have been identified from epidemiological literature on social determinants of health as well as EJ literature on the determinants of siting and emissions (see Table 4.1).

Sensitive Land Use and Hazard Proximity Indicators

This category of indicators captures the level of exposure to air toxics emissions and the location of sensitive populations. These indicators and analytical procedures using them were designed to follow the recommendations of the CARB Air Quality and Land Use Handbook (CARB 2005) for definitions of both the specific land uses that require distance from air quality hazards to

---

[21] Los Angeles, Orange, Ventura, San Bernardino, Imperial and Riverside Counties; this area hosts 48.8% of the population of California

protect the health of individuals at the neighborhood level, the land uses and facility types that represent the air quality hazards themselves, and the recommended buffer distances to separate sensitive land uses from land use categories that represent air quality hazards.

Using the definitions in the Air Quality and Land Use Handbook (CARB 2005), land uses requiring protection include residential land and several "sensitive" land uses, including schools, childcare centers, urban playgrounds and parks, and health care facilities.[22] These features were mapped from several sources, including regional land use mapping and geocoded locations from address lists, and they are included in spatial data layers as both point and polygon features. Polygons representing the location and boundaries of schools, childcare centers, urban parks and playgrounds, and health care facilities were extracted from several sources. These include: land use codes 1244, 1252, 1260-1264, 1810-1821, and 1870 from the SCAG 2005 Land Use data layer.[23]

However, not all sensitive land uses are available as polygon features from existing data sources, either because of the spatial resolution of the data or the way in which the data is classified and automated. For example, some commercial and other facilities contain childcare centers or health care facilities that are not mapped separately. To address this shortcoming, point locations for these additional features were added to the EJSM as spatial data layers to develop a complete and accurate collection of sensitive land use features. School location points were automated by geocoding[24] the address list provided by the California Department of Education; public and private schools are included, and we took care to exclude facilities that do not actually host students, such as maintenance facilities and district offices. Childcare point features were automated by geocoding the addresses provided from a search of SIC codes 8350 and 8351 using the D&B (formerly Dunn and Bradstreet) business information service. CARB provided a point data layer of healthcare facilities, presumably from the California Spatial Information Library (Health Care Facilities).[25] To avoid duplication and accept a polygon feature over a point feature as a default, any point feature that intersected with an equivalent polygon feature was dropped – for example, a point location for a school that is located within a SCAG land use school polygon was deleted.

Because representing these features as dimensionless points would result in misclassification of proximity metrics, we assigned to each a minimum area by creating circular buffers surrounding the points. The size of these buffers were selected based upon the area of the smallest equivalent land use in the SCAG Land Use data layer, on the assumption that those automated as polygons by SCAG represented features of an actual size that is compatible with the spatial resolution of the SCAG data layer, and smaller features were simply not mapped. The smallest healthcare facility in the SCAG layer has an area of 5,530 square meters, so point location features were buffered using a radius of 42 meters; childcare facilities (smallest in SCAG = 1007 sq meters)

---

[22] "Senior centers" are also listed as a sensitive land use, but no data for this type of facility was provided to the team by CARB and no funding was set aside to automate these facilities by address geocoding. However, they could be added to the method in the same manner as was accomplished for child care centers.

[23] The classification system for the SCAG land use data is a modified version of the widely used Anderson Land Use Classification, and developed by Aerial Information Systems, Inc.

[24] All geocoding was done within ArcGIS using two street files: Geographic Data Technology 2004, and TeleAtlas 2007; minimum match score of 80 was required to accept location.

[25] http://www.atlas.ca.gov/

were buffered at a radius of 18 meters; and schools (smallest in SCAG = 2260 sq meters) were buffered at 27 meters. This is a very conservative approach to assigning polygon size to a point feature.

On the hazard size, we sought to develop measures of spatial proximity to point- and area-based air pollution sources, such as large industrial facilities; chemical plants and refineries; power plants; aerospace/electronics manufacturing; chemical manufacturing; metal fabrication; chrome plating; and selected types of hazardous waste treatment, storage, and disposal facilities. Point locations include (a) CHAPIS[26] facilities – a subset of California statewide emissions inventories that represent facilities with emissions of primary concern for health impacts, and for which the information is deemed most accurate; (b) chrome-plating facilities identified from the CEIDARS statewide air toxics emissions inventory; and (c) hazardous waste facilities from the California Department of Toxic Substances Control. Following ARB guidelines, this category of indicators considers proximity to land use categories that are associated with high levels of air pollution, including railways and associated facilities, airports, refineries, intermodal distribution facilities, and ports where diesel emissions are concentrated.[27]

CARB provided spatial data layer representing CHAPIS facilities (2001) and chrome plating facilities as point features for specific use in this project. The SCAG study area includes 786 of the 1839 statewide CHAPIS facilities and 137 of the 188 chrome plating facilities. Locations of hazardous waste facilities (2005) were obtained from the California Department of Toxic Substances Control (DTSC); there are also point features, located using geographic coordinates provided by DTSC and verified using address geocoding. We limited the DTSC dataset to facilities that represent the greatest hazard and those associated with sensitive land uses, and we include facilities listed as federal response (including Superfund), state response, voluntary clean-up, military evaluation, or school investigations and clean-up.

As noted, CARB (2005) also identifies several land use categories as significant sources of air pollution and recommends buffer distances for the siting of new sensitive land to protect public health. The following of these land uses were automated as spatial data layers:

- Railways and associated facilities are represented in the spatial data layers as both polygons and line features. Data sources for these features were the SCAG 2005 Land Use data layer land use code 1412 and the National Transportation Atlas (provided by CARB).

- Airports are represented in the spatial data layers as both polygons and line features. Data sources for these features were the SCAG 2005 Land Use data layer land use codes 1273 and 1411, and the National Transportation Atlas (provided by CARB).

---

[26] (http://www.arb.ca.gov/ch/chapis1/chapis1.htm)

[27] CARB also identifies "freeways and major roads" as a land use associated with high levels or air pollution, but does not define a "major road" in a manner that would allow us to build a spatial data layer representing that land use type. However, proximity to roadways of all types is partially captured by one of the health risk variables - the National Air Toxics Assessment cumulative estimated respiratory hazard associated with ambient air toxics for 1999 - which includes the effects of mobile emissions.

- Refineries are represented in the spatial data layers as both polygons. Data sources for these features were the SCAG 2005 Land Use data layer land use codes 1322 and 1325. We also included facilities identified as refineries from the 2006 CARB Emissions Inventory.[28]

- Intermodal distribution facilities are represented in the spatial data layers as both polygons and line features. Data sources for these features were the SCAG 2005 Land Use data layer land use code 1416, and the National Transportation Atlas (provided by CARB).

Ports are represented in the spatial data layers as both polygons and line features. Data sources for these features were the SCAG 2005 Land Use data layer land use code 1411, and the National Transportation Atlas (provided by CARB).

Health Risk and Exposure Indicators

This category of indicators is composed of five metrics of health risk and exposure to identify the range of cumulative impact (see Table 4.1). All are calculated spatially at the census tract level. They include: the 2001 CARB cumulative estimated lifetime cancer risk associated with ambient air toxics exposures from mobile and stationary sources for 2001[29], the NATA cumulative estimated respiratory hazard associated with ambient air toxics for 1999[30], toxicity weighted hazard scores at the tract level for air pollutant emissions from the 2005 Toxic Release Inventory facilities included in U.S. EPA's Risk Screening Environmental Indicators (RSEI) database[31], and census tract-level ambient concentration estimates interpolated from the CARB statewide monitoring network for $PM_{2.5}$ and ozone averaged for 2004-2006.[32]

The RSEI data merits special mention, partly because of the possibility of some confusion over what this data tells us. RSEI is available in two versions. The first is the one most accessible and familiar to the public: it assesses the potential health impact of industrial releases from pounds-based, hazard-based, and risk-related perspectives for individual facilities in the U.S. The second is not generally made available to the public: it is a (1 terabyte) micro-data set that underlies the public version and is fully disaggregated, with a single record for every chemical, every release, and every facility in the United States for 1988 to 2005; while not generally available, some researchers have used it and we are among them.

We term this set the RSEI Geographic Microdata or RSEI-GM. RSEI uses a fate and transport model to chart where each of these facility-chemical-release combinations affects an area within 100 kilometers of any facility in any given year. In the underlying micro-data (RSEI-GM) that

---

[28] Locational accuracy was verified and corrected using the TeleAtlas imagery used as a backcoverage in GoogleEarth.

[29] Risk calculated using the California state emissions inventory and applying a methodology similar to NATA (http://www.arb.ca.gov/toxics/cti/hlthrisk/hlthrisk.htm)

[30] US EPA's National Air Toxics Assessment (NATA) http://www.epa.gov/ttn/atw/nata1999/. Hazard estimates were modified to integrate RELs from the Office of Environmental Health Hazard Assessment.

[31] US EPA's Risk Screening Environmental Indicators model http://www.epa.gov/oppt/rsei/index.html.

[32] This data was provided by CARB for the purposes of this project.

is gridded for the whole United States, there are three measures for each cell-facility-chemical combination: an estimated concentration, a toxicity-weighted concentration, and a RSEI score that takes into account population totals and the age structure. This data can be combined with census data at the block and block group levels for both general EJ analyses and longitudinal studies that can control for changes in regulatory regimes and reporting requirements over time. The measure used here is the toxicity- weighted concentration at the tract level, generated by first area- weighting to the block group level and the population- weighting to the tract level. However, it is different from the RSEI score (which is, in any case, facility-based), as it uses population only to allocate the data to census tract averages and not to calculate relative risk. Even this measure, however, should be seen not as a risk estimates but rather as a relative concentration measure that allows for comparison of relative pollution burdens, exactly appropriate for this task.

Social and Health Vulnerability Indicators

This category of indicators includes metrics identified from epidemiological research literature on social determinants of health and research on EJ, as statistically significant and robust determinants in cases where patterns of disparate impact have been documented (Krieger, Chen et al. 2003; Gee and Payne-Sturges 2004; Morello-Frosch and Shenassa 2006; DeFur, Evans et al. 2007; Adler and Rehkopf 2008). Like the health risk and exposure indicators, all apply spatially at the census tract level. Variables derived from the 2000 U.S. Census include measures of race/ethnicity (% residents of color), poverty (% residents living below twice the national poverty level),[33] wealth (% home ownership using % living in rented households), educational attainment (% population > age 24 with less than high school education), age (% <5 years old and % >60 years old), and linguistic isolation (% residents >age 4 in households where no one over age 15 speaks English well). Non-census metrics include percent voter turnout as a proxy for degree of engagement in local decision-making, which has been shown to be linked to community health status (LaVeist 1992) (% votes cast among all registered voters in 2000 general election)[34] and adverse birth outcomes (% preterm or small for gestational age infants 1996-03[35]) which have also been shown to be sensitive health endpoints that reflect underlying community health status and which recent studies indicate is also affected by air pollution (O'Campo, Xue et al. 1997; Osmond and Barker 2000; Huynh, Parker et al. 2005; Woodruff, Parker et al. 2009).

---

[33] This multiplier is used to account for the generally higher cost of living in California as compared to the rest of the nation.

[34] California Secretary of State data – 2000; data source: UC Berkeley Statewide Database. Voter turnout was calculated as the number of votes cast in the 2000 general election in each census tract divided by the number of voters registered in that tract http://swdb.berkeley.edu/d00/g00.html

[35] California Department of Public Health Natality Files 1996-2003. Adverse birth outcomes from California live birth records, 1996 to 2003, among those with a valid 2000 census tract code, born after at least 24 weeks gestation, and with a plausible birth weight for their gestational age, after the method of Alexander et al. (1996) For tracts with at least 50 live births during 1996 to 2003 we calculated the percentage born prematurely (24 to 36 weeks gestational age), and/or small-for-gestational-age (SGA - defined as below the 10th percentile of the birth weight distribution of all live births for the period (whether having a valid 2000 census tract code or not) of the same gestational age and infant sex. For example, male infants weighing 2,551 grams or less at 37 weeks gestational age were considered SGA, while female infants of the same gestational age were considered SGA if they weighed 2,466 grams or less.

<u>Analysis</u>

The EJSM is a three-step process: (a) GIS spatial assessment to create a mapping base and to conduct spatial analysis to derive hazard proximity metrics, (b) statistical analysis for error-checking and data processing to derive the CI scores, and (c) GIS mapping of the results. Each of these steps is described in detail, below.

The EJSM requires the initial creation of several spatial data layers, beginning with a base map upon which all spatial analysis and results mapping takes place. Consistent with the focus of the CARB Handbook to protect public health from air pollution in non-occupational settings, the base map is constructed from the land use classes specified in the ARB Handbook – residential areas[36] and sensitive land uses - schools, childcare centers, urban parks and playgrounds, and health care facilities as previously described. These land uses are isolated from a vector spatial data layer depicting land use at high spatial resolution and classification detail. In this study, we used the 2005 SCAG land use data layer, which covers a six county area.[37] This layer was merged[38] with the spatial data layers of circular buffers created from sensitive land use point feature data, mentioned earlier. It is necessary to use a GIS union to merge these features.

To create a geographic link between land use and the census tract-based data used in indicators of social/health vulnerability and health risk/exposure, land use polygons were merged with 2000 U.S. Census block polygons using a GIS intersect to create a final base map composed of neighborhood-sized "CI polygons," each with a known land use class and attribute key to census information[39] (Figure 4.1; again, all figures and tables are at the end of each respective chapter).

The resulting base map consists of over 320,000 CI polygons for the study area, with a mean area of 0.017 square kilometers, or about four acres[40]. CI polygons located in urban areas tend to be much smaller than those in sparsely populated outlying areas. The median CI poly size is 0.008 square kilometers, or about two acres; this value is probably a better aggregate descriptive statistic since the distribution of CI poly area is heavily skewed. Because CI Polygons do not cover the entire study area, maps depicting the geographic pattern of CI scores also indicate the unscored areas as a separate class. The geographic extent of CI polygons for a portion of the study area are shown in Figure 4.2; note that unscored areas are shown in subsequent maps in gray.

---

[36] Residential land is identified by selecting land use codes 1100-1152, 1600, 1810-1821, 1870 from the SCAG 2005 Land Use data layer.

[37] (http://www.scag.ca.gov/resources.htm) SCAG uses a modified version of the widely used Anderson land use and land cover classification system (Anderson et al, 1976) – for details see: http://bhweb.civicasoft.com/civica/filebank/blobdload.asp?BlobID=2579

[38] This merge was accomplished using a GIS union to allow buffer circles to be added to areas not covered by land use polygons, and to allow buffer circles to be inscribed into, or intersect, existing land use polygons to avoid double-counting the area of circles that overlap intersect polygons. Multipart to singlepart feature conversion is also necessary to ensure that each CI polygon is a unique record.

[39]The Summary Tape File identifier that contains census tract, block group and block number. In these datasets, this attribute is the field "Arckeybk".

[40] Mean area of census tracts for this same region is 28.8 square kilometers.

The intersection of census blocks and land use allows for a more geographically specific landscape that we think is crucial to the first part of the exercise, the attribution to each resulting polygon of geographically- specific hazard proximity metrics. On the other hand, we note later that much of the rest of the data is only available (or only available easily) at a census tract level. Because focusing just on each polygon could lead to an overinterpretation of the precision of the data at this fine geographic unit (since the information not related to proximity is attached as a tract attribute), we determine a method to surface our hazard proximity scores up to the tract level as explained below. This choice is appropriate given the locational accuracy and spatial resolution of the land use data available. However, we have also experimented with alternative geographic and scoring techniques, including working from a polygon- only map. In any case, once the polygons are created in the base map, the first step is to apply a hazard proximity score. We describe this in the results below.

The second step consists of scoring the various layers. As we note below, the most complicated scoring procedure involves the hazard proximity scores and we report the method in more detail in the results below. The other two scoring procedures are simpler.

Intermediate scores for each of the five health risk and exposure indicators are calculated individually for each CI polygon based on ranking within quintile distributions by tract for all CI polygons in the Southern California study area; as these metrics are at the tract level, each CI polygon receives the value for its host census tract and the ranking, as noted is done at the tract level. For example, a CI polygon that is located in a tract that ranks in the least impacted 20% for each of the five health risk and exposure indicators in the region would receive a health risk and exposure score of 5 (5 scores of 1), whereas one that ranked in the highest quintile for all indicators would have a health risk and exposure score of 25 (5 scores of 5). These summed intermediate scores are then normalized based on quintiles at the tract level to derive the score for this indicator category, which ranges from 1-5. The map of the Health Risk and Exposure category score is shown in Figure 4.3, for a portion of the study area.

Intermediate social and health vulnerability indicator scores are calculated using the same quintile distribution and normalization technique employed for the health risk and exposure indicators, resulting in final social and health vulnerability scores ranging from 1 to 5. To ensure that social and health vulnerability scores are reliable and not distorted by missing data or based upon anomalously small populations, CI polygons in tracts with fewer than 50 people, and those with 5 or more missing indicator values among the 10 in this category were not scored and do not appear in the final CI maps. Those in tracts missing 1 to 4 metrics receive a score calculated as the average quintile ranking using only the non-missing metrics. The map of the Social and Health Vulnerability category score is shown in Figure 4.4 for a portion of the study area.

All scoring is done with SPSS and SAS programming so as to provide a record and to allow for future integration of alternative data layers and/or a different weighting scheme for the various factors considered. We do not attach explicit weights – a process that implicitly assumes that all weights are equal. We do this for several reasons. First, we found it difficult to justify different weights and our external scientific peer review committee agreed that this was the best working assumption due to the paucity of scientific evidence that would point to a specific weighting

scheme. Second, we found it difficult to explain to the public why different weights would be used and, having noticed the challenge that the creators of EJSEAT faced with explaining their method to the public, we opted for transparency in scoring.

After deriving scores ranging from 1 to 5 along the three dimensions of hazard proximity, air pollution exposure/estimated health risk and social vulnerability, we add these up to produce the final total CI score. The scores are then reattached to the underlying polygons (with our preferred method essentially attributing tract-level scores to all these polygons) and mapped. The final results are depicted below.

Results

Scoring hazard proximity and sensitive land uses

Using the CI Polygon layer with the previously described air quality hazard layers, a set of hazard proximity scores is calculated for each CI polygon using the air quality hazards and land uses identified as air pollution sources, described above. To define proximity, we again follow CARB (2005) handbook guidance on specific distance buffers recommended to separate new sensitive land uses from air quality hazards[41]. Because the recommended distance is 1,000 feet for most air quality hazards, the number and type of point location hazard within that buffer distance is determined for every CI polygon, as are all hazards represented as area sources and land use polygon hazards. Similar proximity metrics are also calculated for buffers at 2,000 and 3,000 ft. for use in the distance-weighted hazard scoring procedure, for reasons explained and in a process detailed below.

Intermediate land use and hazard proximity scores are calculated differently than for the above two indicator categories because hazard proximity can be determined at a much higher level of spatial resolution than census tracts, and the data, at least at the polygon level, does not lend itself to simple quantile ranking. Locations hosting sensitive populations are accorded special attention because children, pregnant women, the elderly, and those with existing health problems are especially vulnerable to the non-cancer effects of air pollution, and there is substantial evidence that children are more sensitive to carcinogenic air pollutants (Olshan, Anderson et al. 2000). In addition, those places with proximity to numerous air quality hazards are assumed to have the greatest cumulative impact.

Base map CI polygons are each scored based on the combination of sensitive land use and proximity to potential hazardous sites and/or hazardous land uses. The basic approach to scoring involves attaching to each CI Polygon (or" CIpoly" in the equations, below) two new pieces of information: the number of "hazards" within a certain buffer distance of its boundaries, and a binary dummy variable indicating whether it hosts a sensitive land use (e.g. a school, health facility, etc.) or not. For sensitive land uses, a CI Polygon (CIpoly) receives one point if it is one of the sensitive land use types. For hazards, perimeter band buffers were constructed surrounding each CI Polygon to evaluate hazard proximity. Using the recommended buffer

---

[41] CARB uses research on air pollutant transport and dispersion, and results of community health risk assessment studies to identify suitable buffer distances.

distances in the ARB Handbook, we construct a 1,000 foot buffer around each CI Polygon and determine whether a hazardous land use falls within that buffer and the number of point hazards inside the buffer (Figure 4.5). Figure 4.6 shows the geographic distribution of the hazard count based on this simple distance buffer for a portion of the study area.

To derive an intermediate score for land use and hazard proximity for each CIpoly, we derive the following:

Hazard_CIpoly$_i$ = Number of hazards within a specified buffer of the CIpoly

Sensitive_CIpoly$_i$ = Dummy variable indicating whether a CIpoly is a school, a health care facility, a childcare facility, or a park.

Using a 1,000 foot buffer, the maximum number of hazards any CIpoly has in our Southern California study area is 7 (the vast majority have 0, and the next largest plurality have only 1), so the maximum combined sensitive land use/hazard score a CIpoly could get is 8. The formula for the total score would be:

Hazard_proximity$_i$ = Hazard_CIpoly$_i$ + Sensitive_CIpoly$_i$

The resulting spatial and categorical distribution is shown in Figure 4.6; we utilize Jenks natural breaks to determine the categories to examine.

We see two problems with this simple fixed 1,000 foot buffer approach. The first is that that some features (e.g. the point location hazards) are represented spatially as points but are, in reality, polygons, something that implies that the buffer from a point might not even extend beyond the facility itself (if it were large enough). The second issue is that such a precise scoring method does not recognize the reality of locational accuracy limits[42] of the GIS data and geocoding. Because of this, we experimented with larger buffer distances and then settled on using a distance- weighted approach to account for the issues (discussed further below).

To see the first of these problems, Figures 4.7 and 4.8 compare point location in the data to a visual image of the facilities from satellite photos. As can be seen, hazards represented as point features are actually polygonal areas in the true geography (compare Figures 4.7 and 4.8). A similar situation is shown for a nearby school in Figures 4.9 and 4.10. The degree of locational error can be significant compared to the 1,000 foot buffer used to determine hazard proximity for CI polygons described previously, such that proximity metrics for some CI polygons will be wrong, misclassifying CI polygons with values that are too high, and others too low.

---

[42] The US Census Horizontal_Positional_Accuracy_Report for the 2000 TIGER boundary files states that "The positional accuracy varies with the source materials used, but generally the information is no better than the established National Map Accuracy standards for 1:100,000-scale maps from the U.S. Geological Survey (USGS);thus it is NOT suitable for high-precision measurement applications such as engineering problems, property transfers, or other uses that might require highly accurate measurements of the earth's surface." This suggests uncertainty for census polygon boundaries of about 50 meters, or over 15% of the 1000 ft. buffer distance. Accuracy of some datasets used in this analysis is not known.

The issue of location inaccuracy can also be seen in the figures. In Figures 4.7 and 4.8, we represent the CHAPIS-generated location with a square polygon and the actual location (from street address geocoding and investigation of this particular area) with a circle. Note that of the three facilities pictured in this area, two are inaccurately located, with the degree of inaccuracy more significant for the Saint-Gobain containers facility which is incorrectly located in a residential neighborhood. The sensitive land uses depicted in Figures 4.9 and 4.10 are more locationally accurate, but note that the SouthEast Learning Center, a point feature we located by geocoding, shows up as land "under construction" in the land use spatial data layer.

While we think that improving the geocoding accuracy within the CHAPIS dataset is important for making an EJSM that works well, these types of locational inaccuracies are simply inherent to some of the data used in this method and would require extreme investments in time and effort to resolve for very large areas. Both because of this and the concern about the polygon size of point-designated features, we decided to employ a distance- weighted scoring approach for hazard proximity.

We thus use a distance-weighted approach where the contribution of each individual hazard to the hazard score for a polygon is given a weight that depends on its distance from the CIpoly boundary. For example, for any hazard within 1,000 feet, we assign a weight of one; for a hazard within 2,000 feet but further than 1,000 feet, we assign a weight of 0.5, and for a hazard within 3,000 feet but further than 2,000 feet, we assign a weight of 0.1 (Figure 4.11).

The weighting choices that we use represent a compromise to capture the fact that pollutant concentration is not linear with distance, but that utilizing a formal distance decay function would be computationally costly and might be less easily understood by non-technical users of the EJSM results.

The following formula operationalizes this approach:

$$Hazard\_distance\_score_i = Hazard\_CIpoly\_1000ft_i$$
$$+ 0.5* Hazard\_CIpoly\_1000\_2000ft_i$$
$$+ 0.1* Hazard\_CIpoly\_2000\_3000ft_i$$
$$+ Sensitive\_CIpoly_i$$

Using this method, the summed point totals for each CI Polygon ranges from 0 to 9.8. To calculate the intermediate land use and hazard proximity category scores, we first add the distance- weighted hazard proximity values to the sensitive land use metric to obtain raw scores. The distribution of these values is highly skewed (for example, see histogram in Figure 4.12), and applying the 1-5 score using simple quantiles would result in a highly distorted map that would overemphasize the highest (5) category and create apparent divisions among the low scores that are not reflected in the actual proximity data. Instead, we use the Jenks natural breaks algorithm to classify these values into 5 classes that represent the "best fit."[43] These classes are

---

[43] The Jenks' natural breaks classification scheme determines the best arrangement of continuous data values into classes by iteratively comparing sums of the squared differences between observed values within each class and

then assigned values ranging from 1-5, as is done with the other indicator classes to insure consistency with the other layers.

As compared to Figure 4.6, Figure 4.13 is a better picture of hazard locations and tends to capture a broader area in the higher Jenks breaks. However, we employ a second step in the process to take the result to the tract level.

Results at the tract level

Because our other two dimensions in the EJSM are calculated at the tract level, we have two choices: attach tract level measures to the CI polygons, or surface the CI polygon scores to the tract level. While we tried both, we prefer moving up to the tract level. This process allows us to further deal with issues of locational inaccuracy and it avoids a sense of "false precision" that might be left with a non-expert user looking at variations on the underlying map (variations that would be driven entirely by shifts in one dimension). After all, this is meant to facilitate screening, not assessment.

Indeed, it is important to keep in mind that there is some uncertainty about all the non-census tract-level measures (such as the estimated excess cancer risk and respiratory hazard) and so any method is a guide not a definitive tool. That said, we think that moving up to the tract level at least deals with the fact that the polygon proximity metrics also have some degree of uncertainty and this process smoothes out the pattern.

To surface up to the tract, we can either use area weights or population weights; we prefer the latter as this might better reflect human health concerns. The actual formula involves first assigning the block level population to all the CIpolys within a block according to the CIpoly area. Thus, the variables are:

> $Area_i$ = area of a CIpoly
> $Area\_CIpolys_B$ = the total area for all CIpolys in a block.
> Pop_B = population of a block
> Pop_i = population of a CIpoly

The formula for population in a CIpoly is:

> $Pop_i = Pop_B * (Area_i / Area\_CIpolys_B)$

Letting Pop_T be the total population in the tract, the formula to get to the tract level for the hazard proximity score is generally:

> $Hazard\_proximity\_pop\_wtd_T = \sum Hazard\_proximity_i * (Pop_i / Pop_T)$

---

the mean for that class. The "best" classification identifies breaks in the ordered distribution of values that minimizes the sum of squared differences within each class.

where the summation occurs from i to N, where N is once again the number of CIpolys in the tract. Since we believe that the distance-weighted count to the CIpoly level is most accurate, the formula we prefer is:

$$\text{Hazard\_distance\_score\_pop\_wtd}_T = \sum \text{Hazard\_distance\_score}_i * (\text{Pop}_i / \text{Pop}_T)^{[44]}$$

The results are depicted in Figure 4.13. Note that although we are providing tract-level scores, we have overlaid the map with the non-considered land areas (in grey) so as to draw the viewers' attention to the areas of concern.

Finally, the three intermediate indicator category scores are summed to produce a Total CI Score, which ranges from 3-15 (See figure 4.14). One interesting result of this process is that the resulting distribution is near normal.

Figure 4.15 and 4.16 allow the reader to see the scores at a finer resolution by drilling in to two areas in Southern California.

Discussion

As can be deduced, there are a number of science-policy choices made along the way in deriving a screening approach. Note, for example, that in terms of the land use dimension of this exercise, one could use any variant of buffers (there are three that we look at), distance weighting (and weights for those distance bands), and summarizing (or not) to the tract using either population or simple area weights. These are many permutations and although there are advantages and limitations to the different methods of scoring for land use/hazard proximity data layer, our preferred approach is distance-weighting hazard proximity to the CI Polygons, and then population-weighting the results to get a tract-level score (Figure 4.13).

Using a finer level of geographic resolution – that is, mapping instead back to the CI polygons – could preserve the accuracy of the hazard proximity metrics as they were calculated and this approach may be optimal for land use decisions. However, this approach may also misrepresent the geographic detail of the health risk/exposure and social/health vulnerability metrics, all of which are calculated at the tract level. The fact that we have done this, however, likely has the effect of lowering scores (because of the averaging at the tract level) for CI Polygons that are

---

[44] An alternative way of thinking about this is that we are area-weighting to the block level, then population weighting up to the tract. Note that:

$\text{Hazard\_distance\_score\_pop\_wtd}_T$

$= \sum \text{Hazard\_distance\_score}_i * (\text{Pop}_i / \text{Pop}_T)$

$= \sum \text{Hazard\_distance\_score}_i * [(\text{Pop}_i / \text{Pop}_B) * (\text{Pop}_B / \text{Pop}_T)]$

$= \sum \text{Hazard\_distance\_score}_i * [(\text{Area}_i / \text{Area\_CIpolys}_B) * (\text{Pop}_B / \text{Pop}_T)]$

Where the second step involves multiplying by $(\text{Pop}_B / \text{Pop}_B)$ and the second step involves a substitution. Since we might define a hazard proximity score at the block level as follow:

$\text{Hazard\_distance\_score}_B = \sum \text{Hazard\_distance\_score}_i * (\text{Area}_i / \text{Area\_CIpolys}_B)$

Then it is the case that:

$\text{Hazard\_distance\_score\_pop\_wtd}_T = \sum \text{Hazard\_distance\_score}_B * (\text{Pop}_B / \text{Pop}_T)$

within the high extreme range of the distribution, possibly under-representing cumulative impacts for some neighborhoods.

Another set of methodological choices along the way results from our desire to focus on neighborhood conditions. With our air-relative health risk measures, we generally focus on ambient concentrations rather than estimated health risk that took into account the nature of the population in the area under question. Both the cancer risk estimates we generated from air toxics and diesel PM data from CARB, as well as the respiratory hazard ratios we generated from NATA, reflect the risk estimates from chronic exposures of residences in the area under question. CARB had been interested in having us use mortality data for the ozone and PM fields but this would have diluted any potential observations of exposure disparities as mortality estimates generated by CARB are heavily dependent upon the age of the population in a tract, with older populations likely to have higher mortality estimates even when pollutant concentrations may have been lower. However, we account for the population age characteristics in our social vulnerability dimension and so we thought it appropriate to stick with interpolated concentrations for PM and ozone as well as estimated lifetime risk estimates for air toxics.

We also understand that there are other reasonable approaches that could involve changing the weights on the indicators that we use. Because of this, the method has been designed with maximum transparency in terms of GIS layers and programming to achieve the scores; other users could generate scores using other weights but we believe this to be the best set of procedures and that sense was shared by the Peer Review Committee during our December, 2008 meeting to review the EJSM.

We should also note that there are potential issues with regard to the fact that we take into account hazard proximity rather than simply exposure-related health risk. It is scientifically arguable that one should concentrate on the latter but communities are very concerned about nearby hazards, and any tool that does not include proximity is not likely to meet the community credibility test. Moreover, there are land uses designated by CARB of special concern, and the CARB handbook recognizes proximity relationships of the magnitude we use to be of concern in guiding siting decisions.

Partly because we take the proximity issue so seriously, we have built up the EJSM utilizing the intersection of a land use layer with census block geography. This creates distinct advantages but also one distinct disadvantage: the method relies on reasonably accurate land use data. This data is not uniformly available for all regions of California and we think that future research could be focused on whether lower resolution land use data might be utilized and how that would affect the EJ screening results. That said, we are also convinced that land use data will continue to improve as the years pass and so this is not likely to be a serious liability of the EJSM. We are currently exploring opportunities to expand the EJSM to other regions, such as the Central Valley, where high- resolution land use data is not as readily available but where EJ screening is needed to inform regulatory decision-making.

Finally, any screening method must be followed up with further validation efforts to assess the accuracy of underlying drivers of cumulative impacts and to assess whether additional emissions

sources and land uses are being adequately captured by the underlying data. Moreover, the EJSM can elucidate where there might be areas of regulatory concern for EJ, but determining actual community environmental health risks would require additional local monitoring and modeling efforts.

Conclusion

One of the central goals of this project has been to develop an EJSM that can identify areas affected by cumulative exposure and social vulnerability as a guide to policy makers, regulators, community leaders, and others. As a result, we set out the goal of generating a method that would be based in the scientific literature about environmental health risk and social vulnerability but would also be open to implementation and alteration by diverse users, understandable to the public, and sensible in its results. One key test of this: as part of this project, we conducted trainings with the CARB staff and transmitted all relevant data and programming syntax, stayed in communication through early ARB use, and as a result, staff now have begun to adapt the EJSM for their use in different decision-making contexts including the implementation of the AB32 Scoping Plan.

We believe that the EJSM offered here represents the beginning of a process to develop transparent and scientifically valid tools to inform regulatory decision making that addresses EJ concerns related to disparities in the cumulative impacts of environmental and social stressors. In our view, the EJSM is superior to simple comparisons of health risk (as in the MATES or NATA layers) and it has advantages over the only other similarly comprehensive screening competitor in this arena, EJSEAT. That said, we understand that this is work in its infancy and that future efforts are likely to yield even better products. Because of that, we have tried to not only fully document our own approach and choices but to create the model in such a way – combining GIS layers with SPSS and SAS syntax to program the scores – such that others can build and develop on this work. We believe that a new generation of research and interaction among university, community, and agency researchers will result.
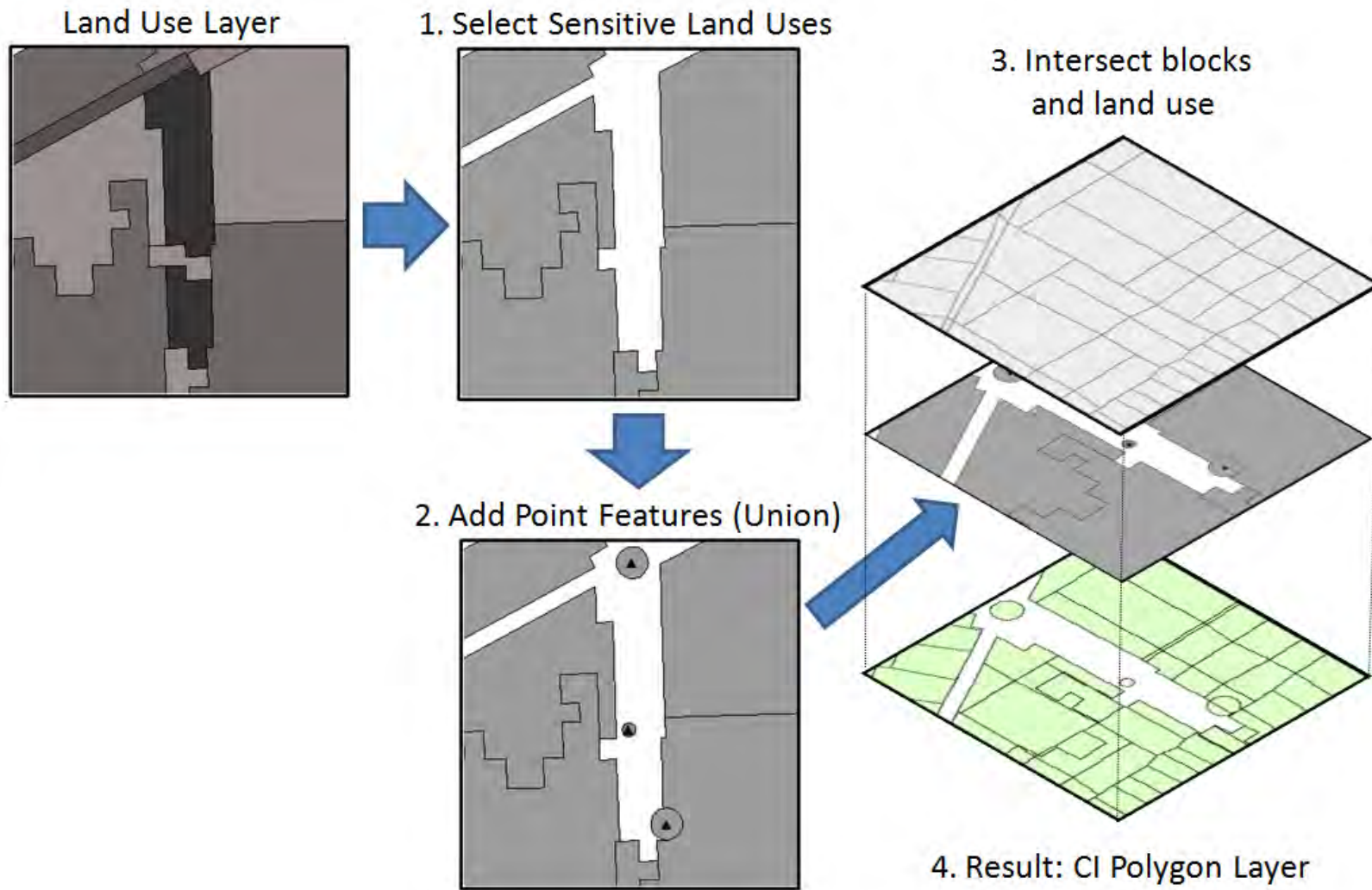
Table 4.1: Summary of cumulative impact and vulnerability indicators used in the EJ Screening Method, with geographic unit of analysis and data source.

INDICATOR	GIS SPATIAL UNIT	SOURCE
Land Use and Hazard Proximity		
Sensitive Land Uses		
Childcare facilities	Land use polygons	SCAG 2005
	Buffered points	Dunn and Bradstreet by SIC code
Healthcare facilities	Land use polygons	SCAG 2005; CaSIL
Schools	Land use polygons	SCAG 2005
	Buffered points	CA Dept of Education
Urban Playgrounds	Land use polygons	SCAG 2005
Air Quality Hazards		
CHAPIS facilities	Point locations	CARB
Chrome-platers	Point locations	CARB
Hazardous Waste	Point Locations	CA Dept. Toxic Substances Control
Hazardous Land Uses		
Railroad facilities	Land use polygons	SCAG 2005
	Line Features	NTAD
Ports	Land use polygons	SCAG 2005
Airports	Land use polygons	SCAG 2005
	Line Features	NTAD
Refineries	Land use polygons	SCAG 2005

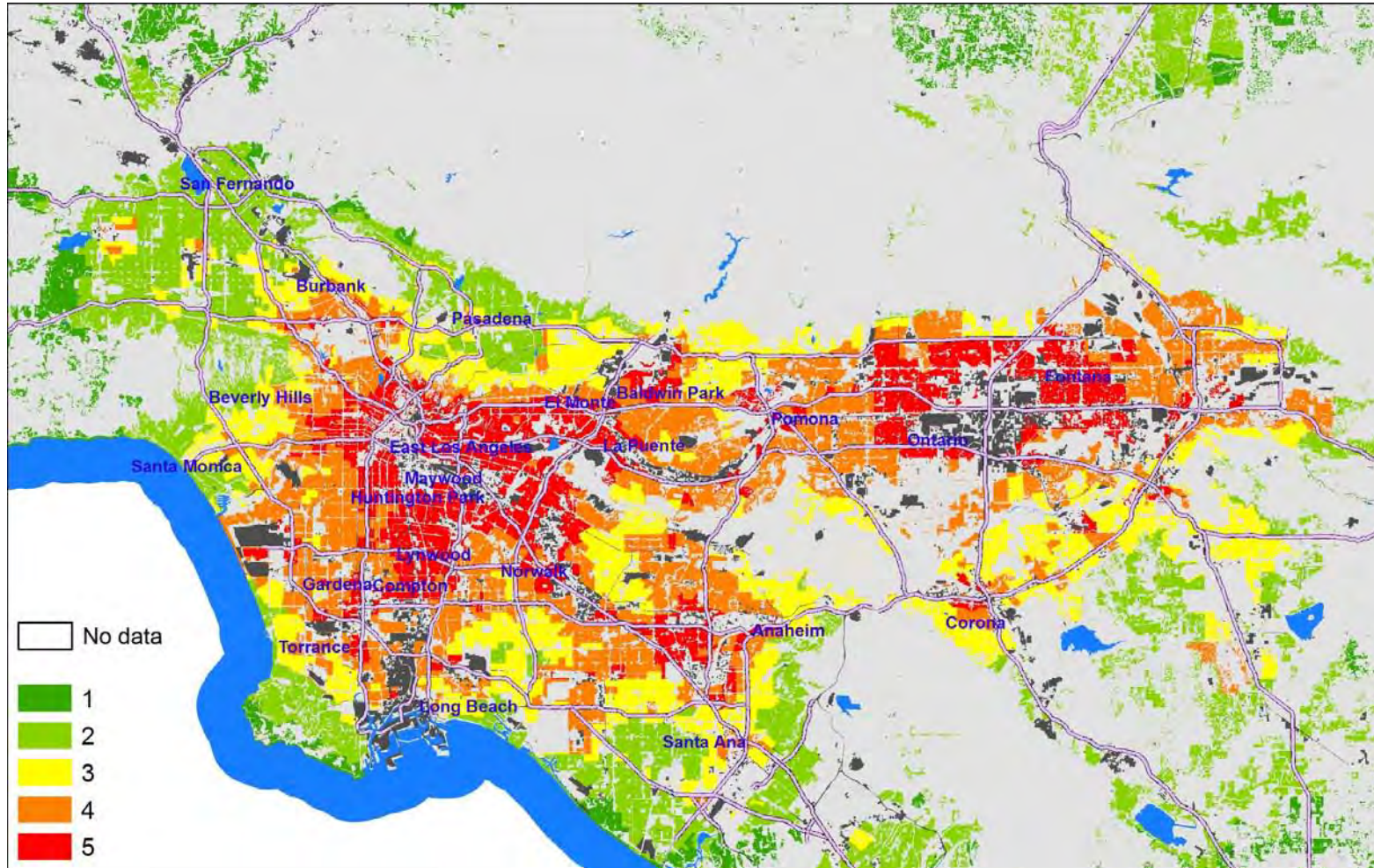|		
Intermodal Distribution	Land use polygons	SCAG 2005
	Line Features	NTAD
Health Risk and Exposure all at Census tract level		
RSEI toxic concentration hazard score		USEPA 2005
NATA respiratory hazard, air toxics, mobile/stationary		USEPA 1999
CARB estimated cancer risk, air toxics, mobile/stationary		CARB 2001
$PM_{2.5}$ estimated concentration from monitoring		CARB 2001-03
Ozone estimated concentration from monitoring		CARB 2001-03
Social and Health Vulnerability all at Census tract level		
% people of color (total pop – non-Hispanic white)		US Census 2000
% below 2X national poverty level		US Census 2000
Home Ownership - % living in rented households		US Census 2000
Housing Value – median house value		US Census 2000
Educational attainment - % >age 24 with <high school		US Census 2000
Age of residents - % < age 5		US Census 2000
Age of residents - % > age 60		US Census 2000
Linguistic isolation -		US Census 2000
Voter turnout - % votes cast in 2000 general election		UC Berkeley Statewide Database
Birth outcomes -% preterm and small for gestational age		CA Dept Public Health 1996-2003

Figure 4.1: Summary of process used to create CI Polygons.

Land Use Layer

1. Select Sensitive Land Uses

2. Add Point Features (Union)

3. Intersect blocks and land use

4. Result: CI Polygon Layer

Figure 4.2: Map of a portion of the study area showing CI Polygons in white, and areas not scored (including open space, vacant land, industrial land use, etc**.) in gray**
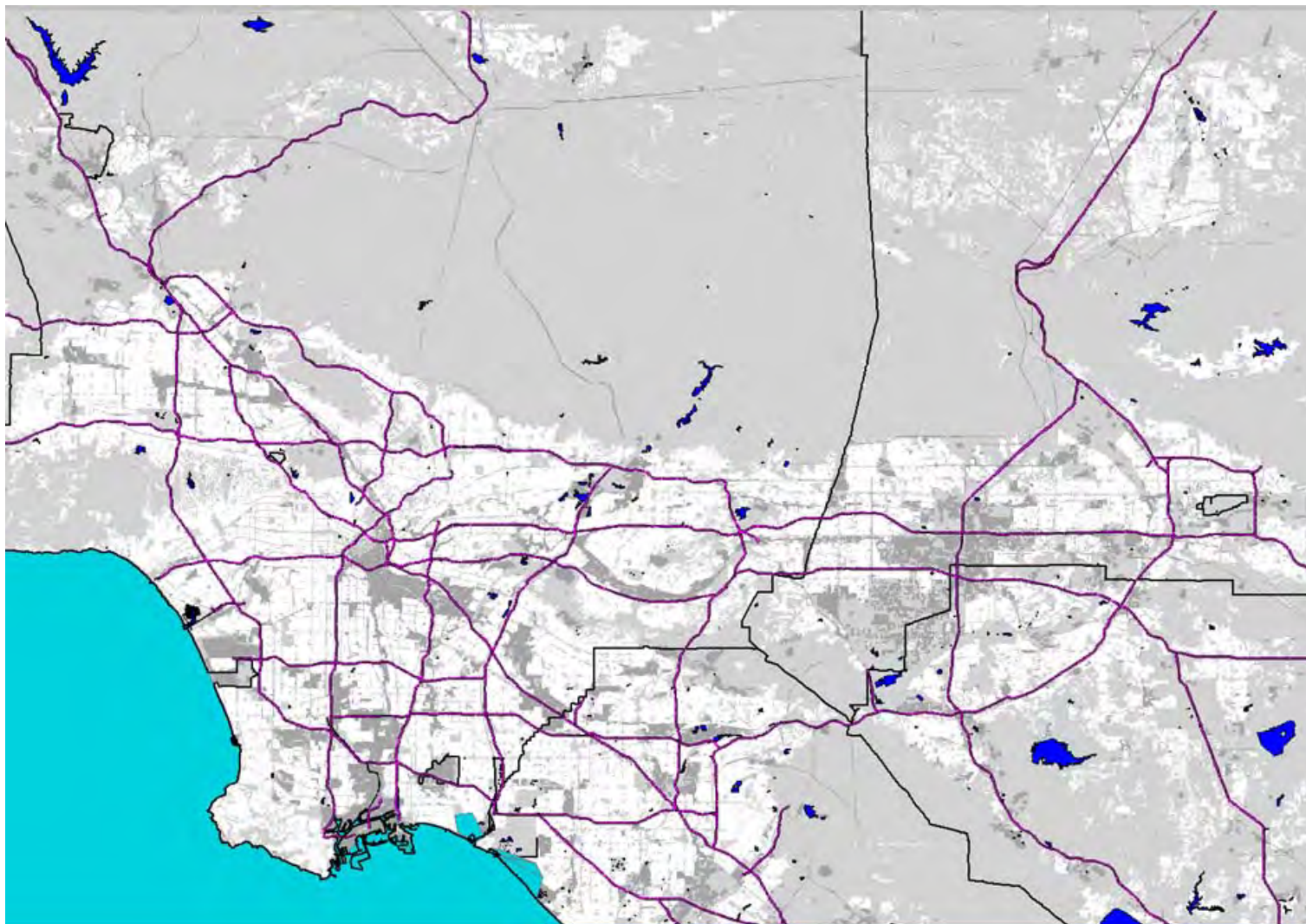
Figure 4.3: Map of the Health Risk and Exposure category score for the greater Los Angeles urban area (land uses not used to create CI Polygons overlay the census tracts, and are shown in gray; mapping is onto CI Polygons using a quintile distribution of the tract level scores)

Figure 4.4: Map of the Social Vulnerability category score for the greater Los Angeles urban area (land uses not used to create CI Polygons overlay the census tracts, and are shown in gray mapping is onto CI Polygons using a quintile distribution of the tract level scores)
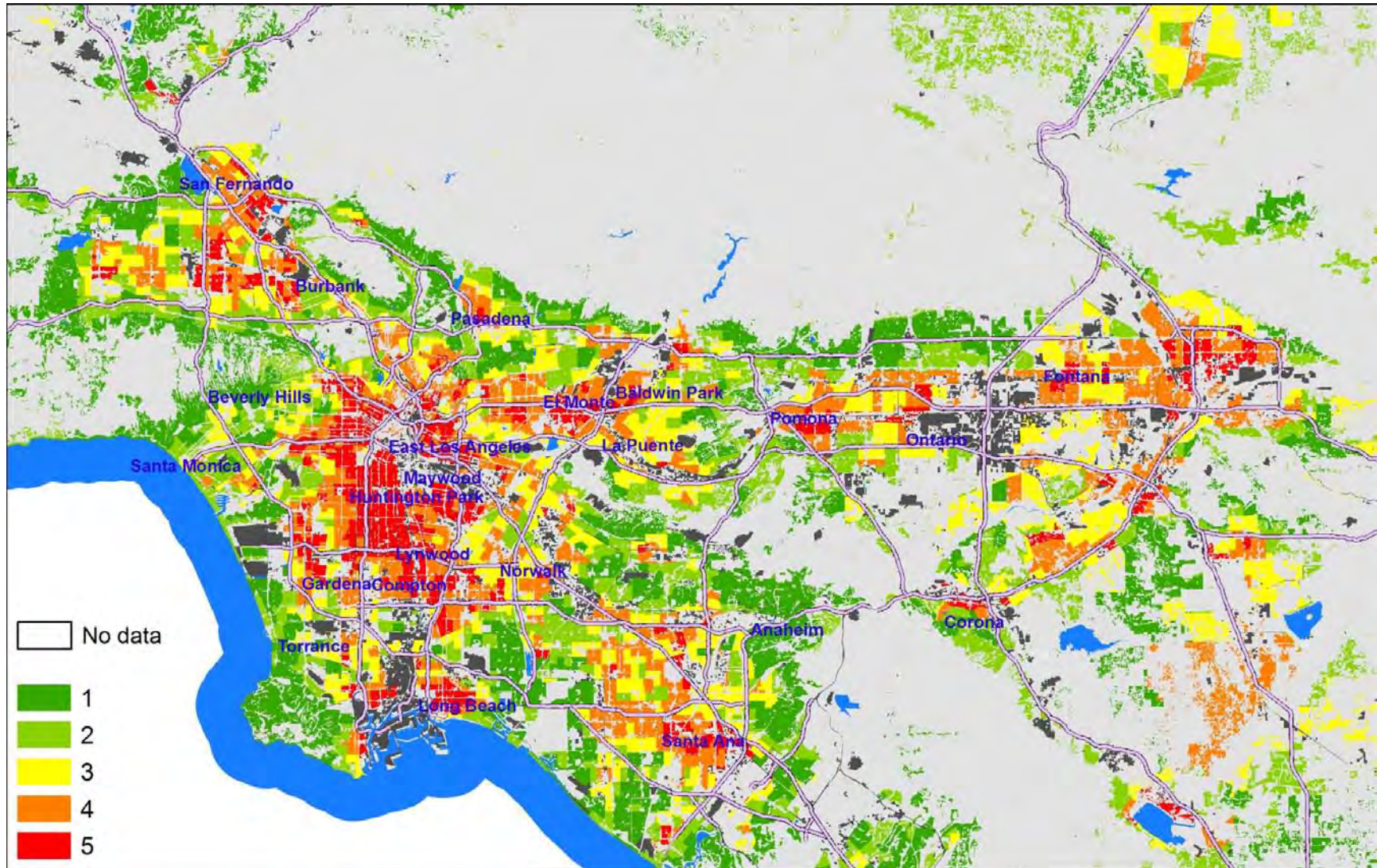
Figure 4.5: Hazard proximity to CI Polygons defined using single buffers (the number of point hazards and land uses hazards located within each buffer is summed for each CI Polygon)

- Buffer CI polygon boundaries

- Hazard proximity based on number of facilities (point-sources) and hazardous land uses inside the buffer

1 PH + 0 LH = 1 proximate hazard

1+0

0+1

PH = Point hazards
LH = Land use hazards
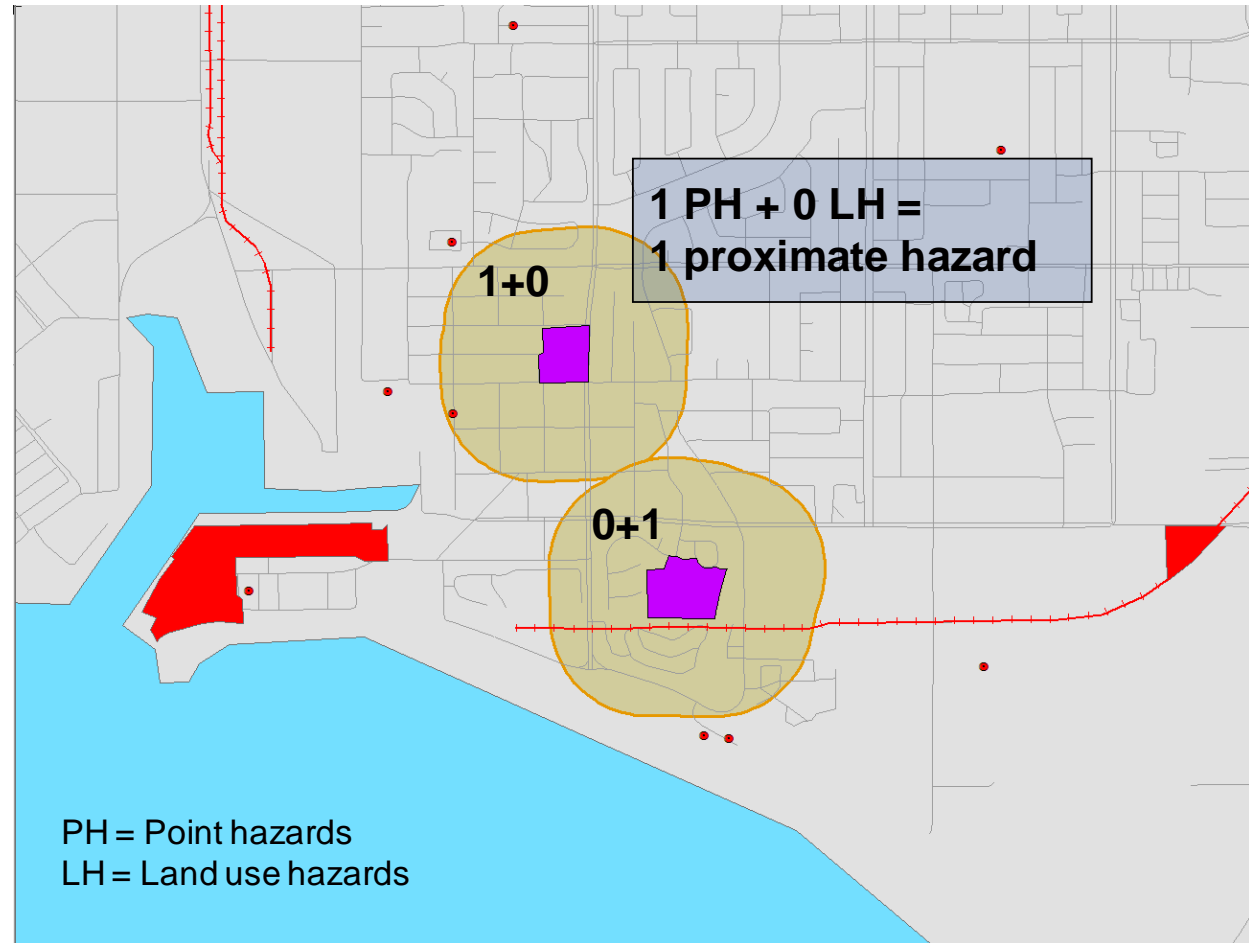
Figure 4.6: Map of the greater Los Angeles area showing the distribution of hazard count by CI Polygon (1,000 ft. buffers, categories determined via Jenks natural breaks)
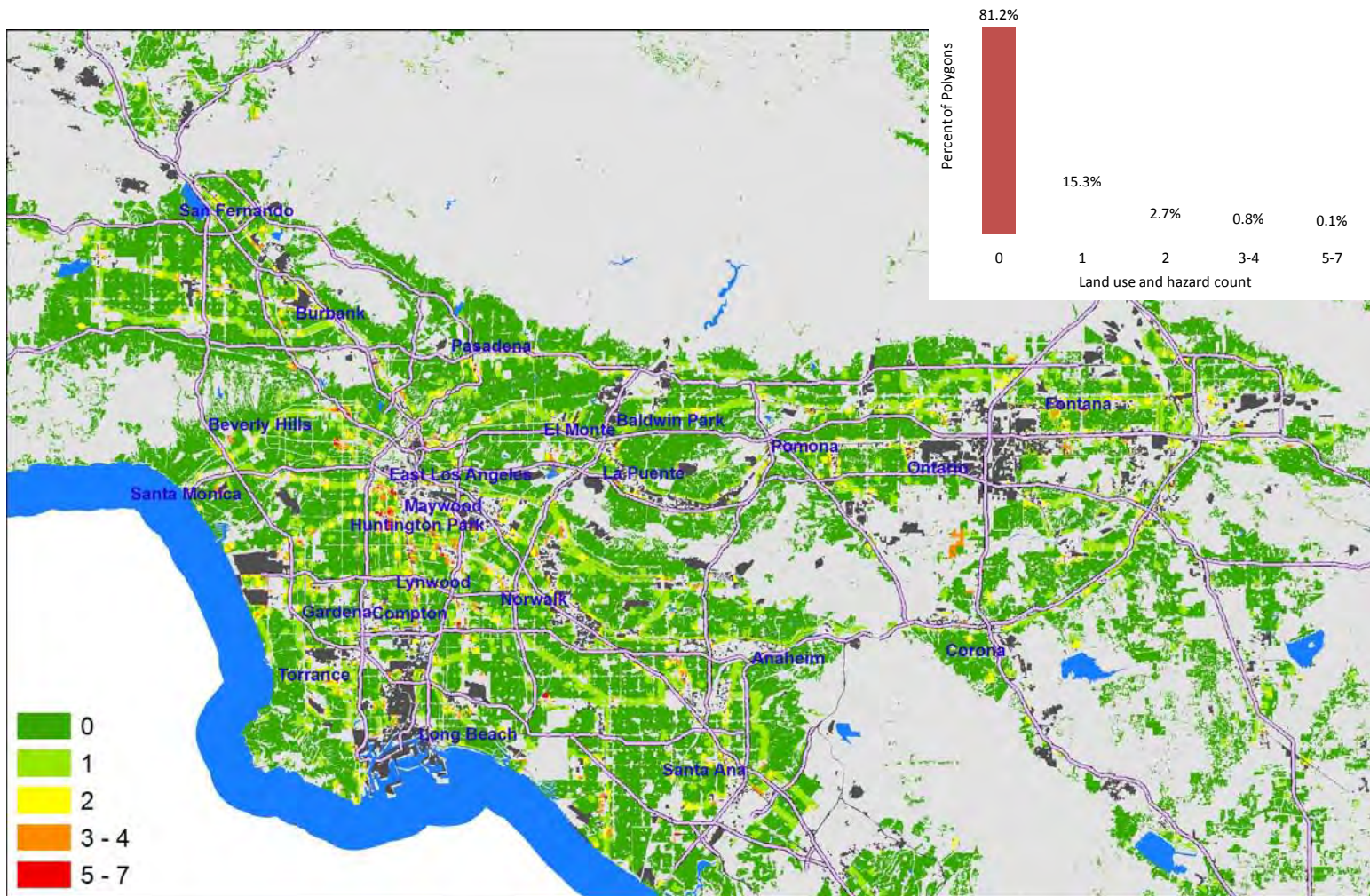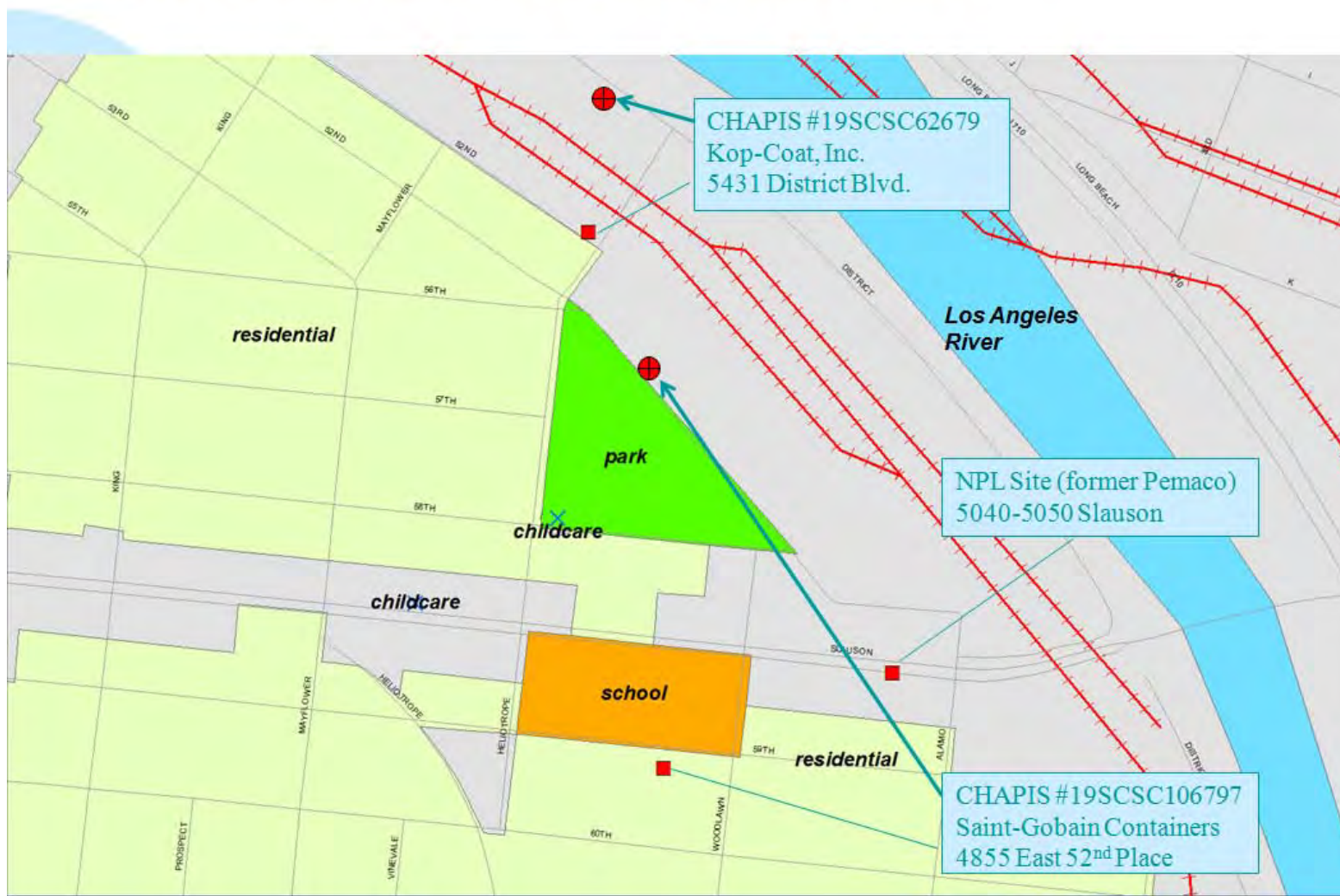
Figure 4.7: Land use polygons and GIS point hazard features for a portion of the city of Maywood

Figure 4.8: Aerial imagery of the map shown in Figure 4.7 (note the actual polygonal areas of the three air quality hazards shown as points in Figure 4.7).

Facilities Mapped as Points Using Available Data But Actual Facility Area is Much Larger

Figure 4.9: Land use polygons for a portion of the city of Maywood, showing GIS polygons identifying schools, as well as buffered point locations of a school, two childcare facilities, and a health clinic
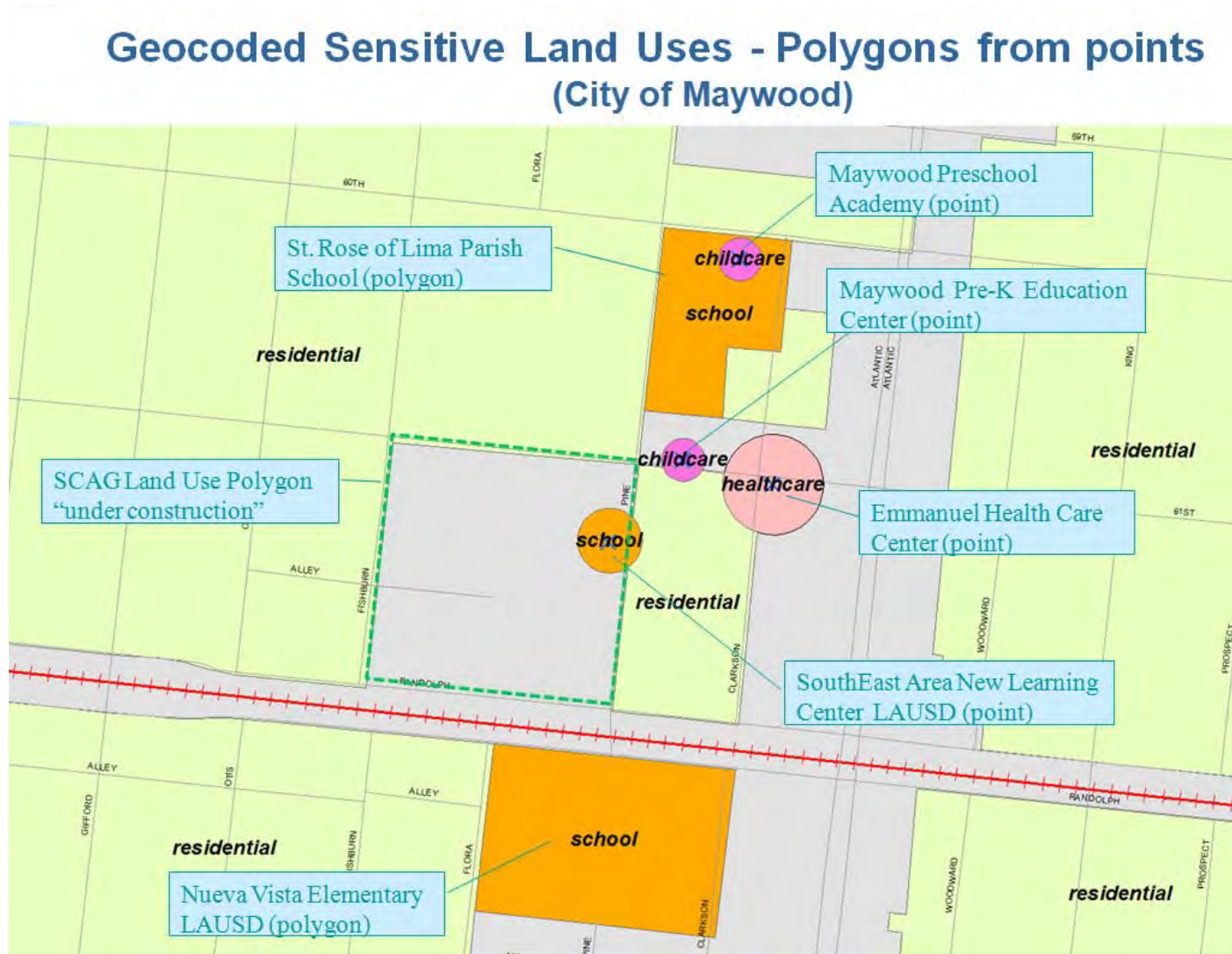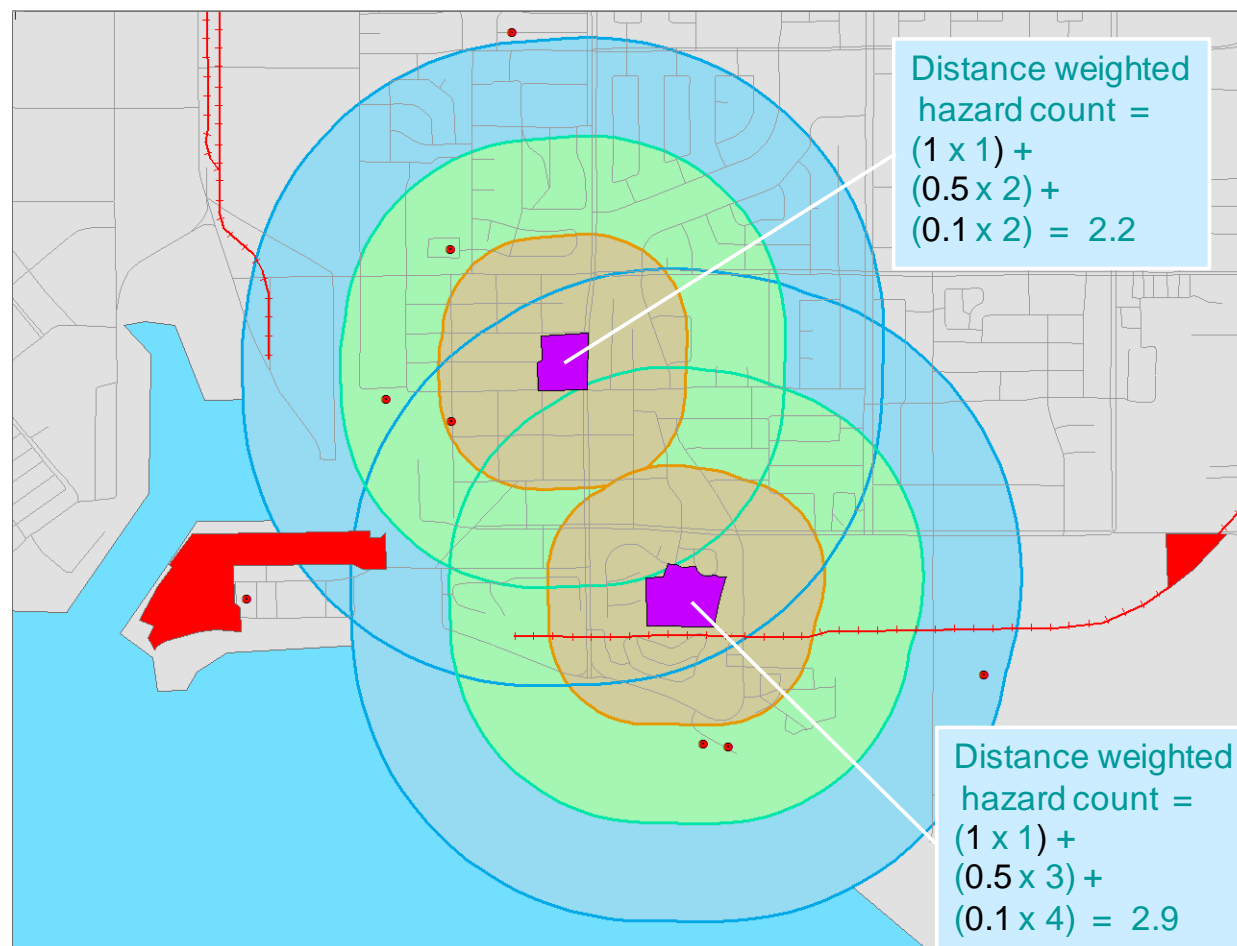
Figure 4.10: Aerial imagery of the map shown in Figure 4.9

## Geocoded Sensitive Receptor Land Uses Polygons from points

Figure 4.11: Hazard proximity to CI Polygons defined using multiple buffers at 1000 foot intervals (the number of point hazards and land uses hazards located within each buffer – exclusive of the interior buffer – is summed for each CI Polygon, appropriately weighted according to distance, and then summed overall for each CI Polygon)
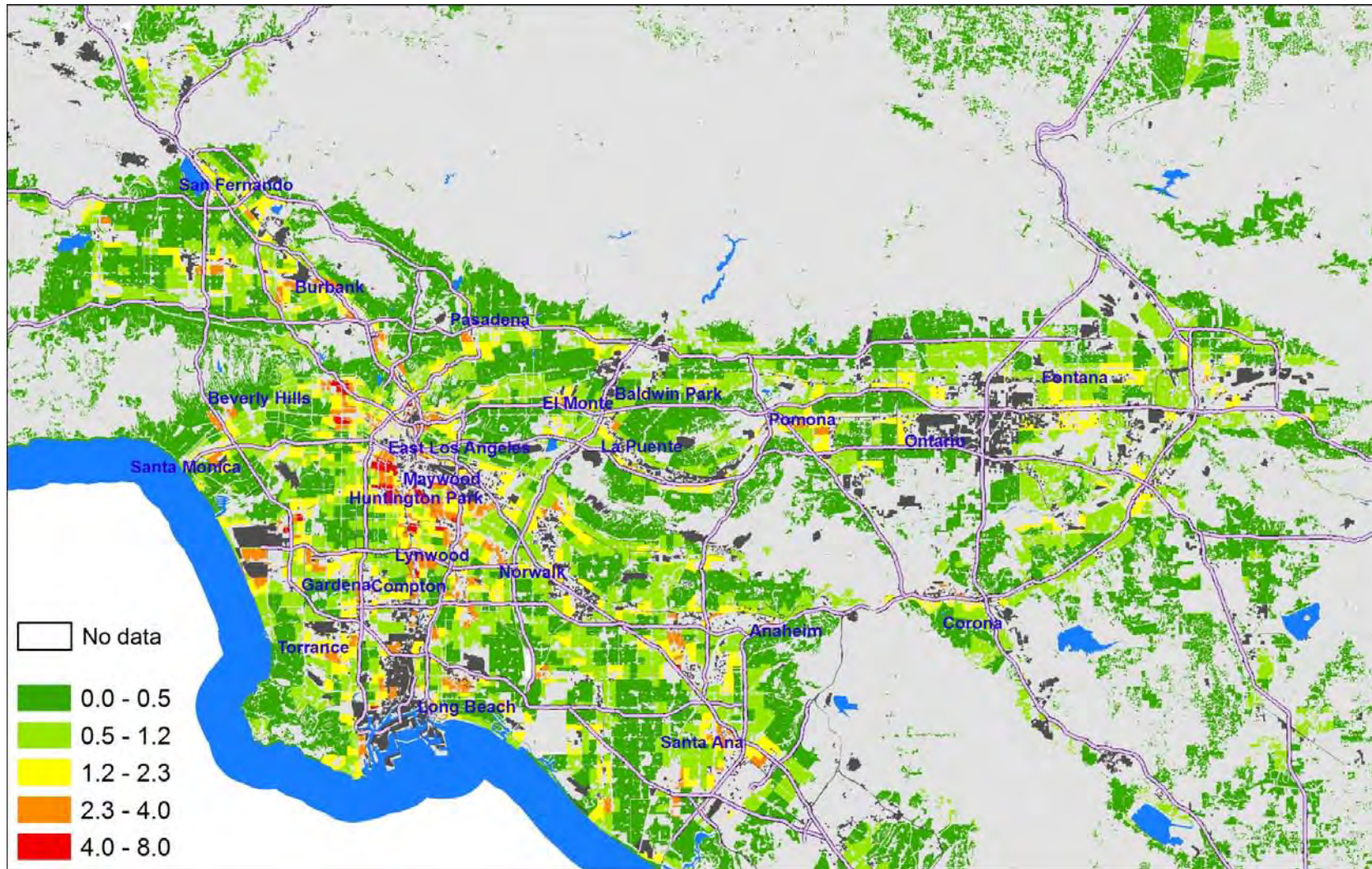

- Buffer CI polygon boundaries <u>at different distances</u>

- Hazard proximity based on number of facilities (point-sources) and hazardous land uses inside the buffer

Distance weighted hazard count =
(1 x 1) +
(0.5 x 2) +
(0.1 x 2) = 2.2

Distance weighted hazard count =
(1 x 1) +
(0.5 x 3) +
(0.1 x 4) = 2.9

Figure 4.12: Map of the greater Los Angeles area showing the distribution of distance weighted hazard and sensitive land use count by CI Polygon
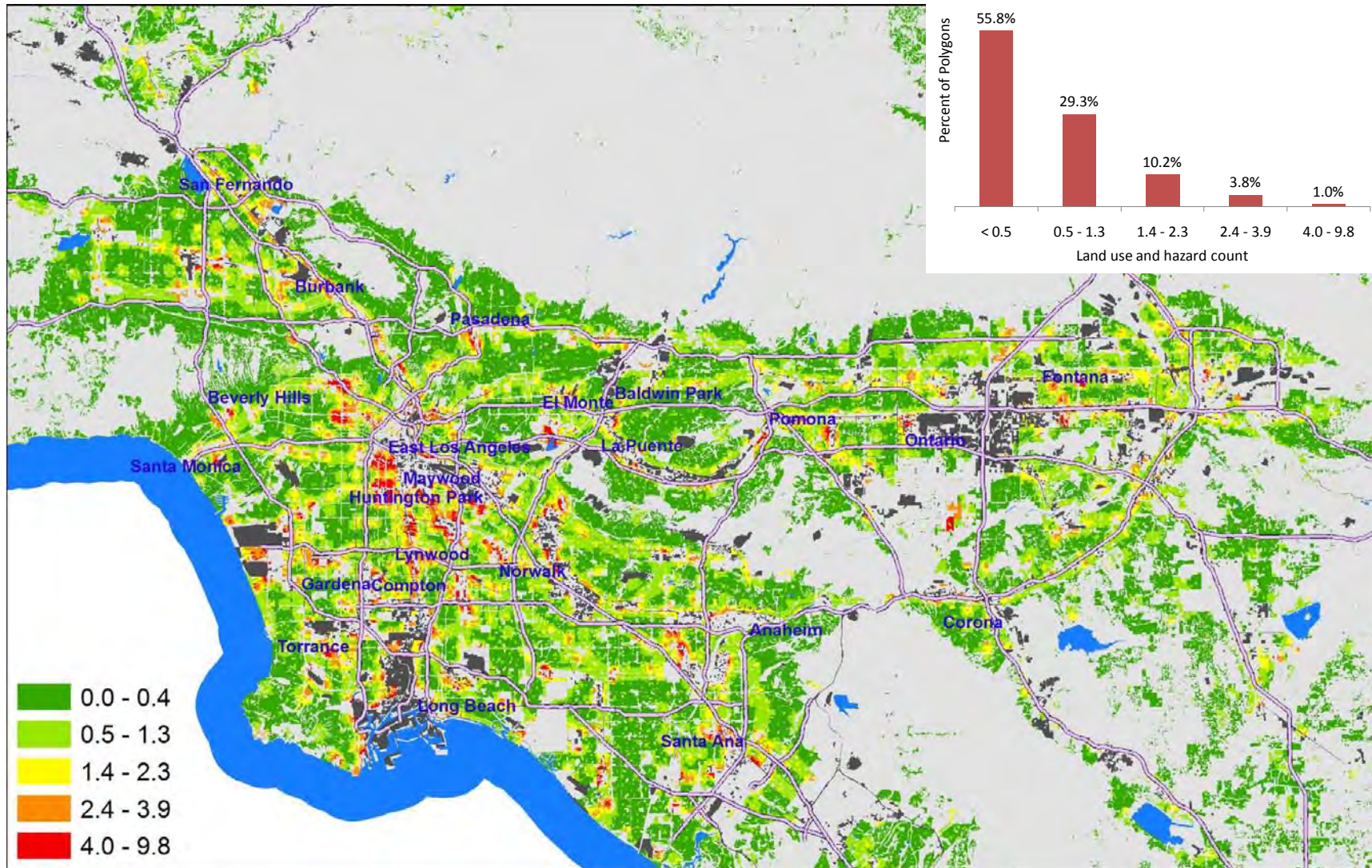
Figure 4.13: Map of the greater Los Angeles area showing the distribution of distance weighted hazard and sensitive land use count at the tract level (population weighed with categories derived via Jenks natural breaks)

Figure 4.14: Map of the Total Cumulative Impacts score for the greater Los Angeles urban area (land uses not used to create CI Polygons overlay the census tracts, and are shown in gray).
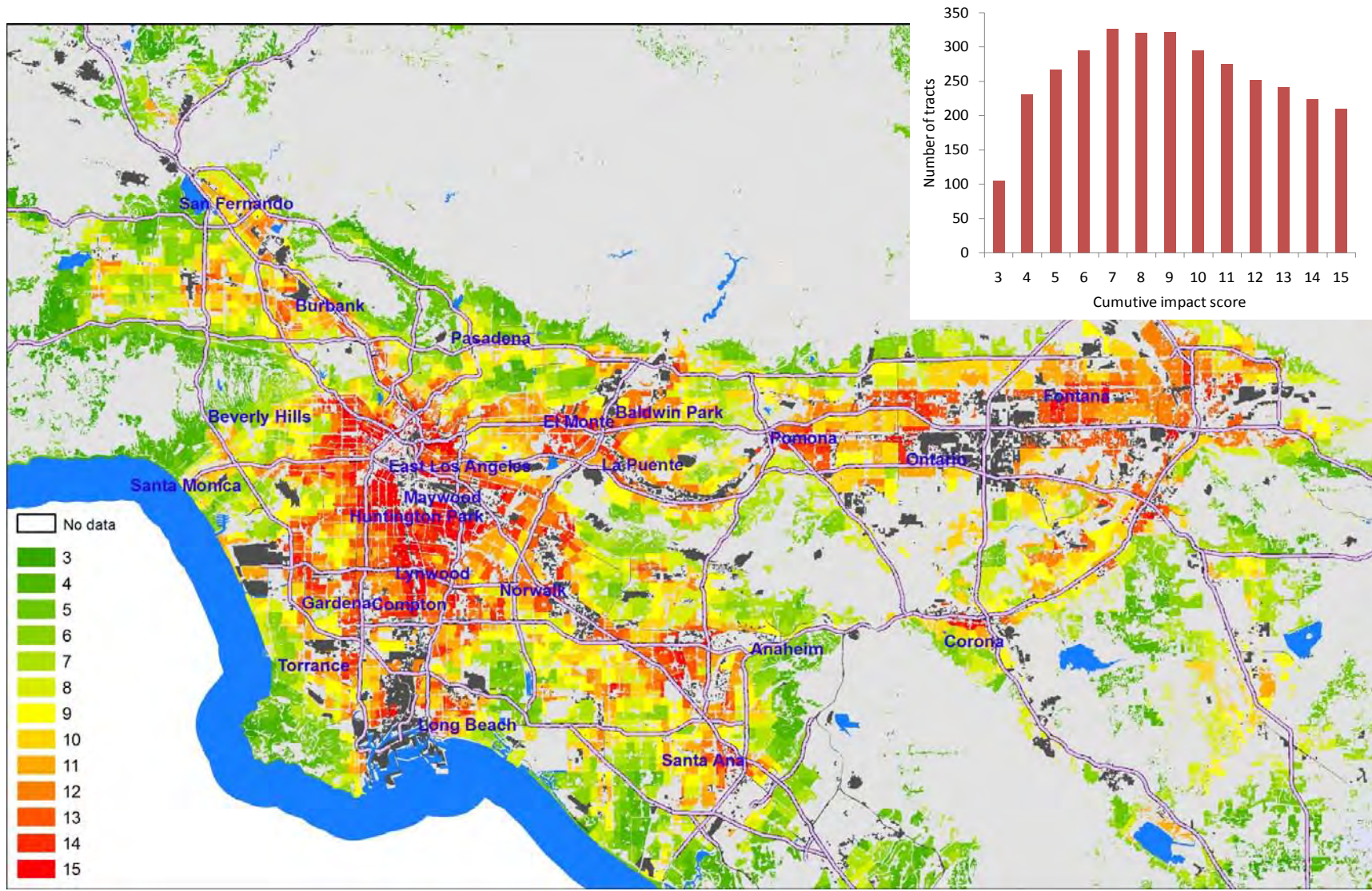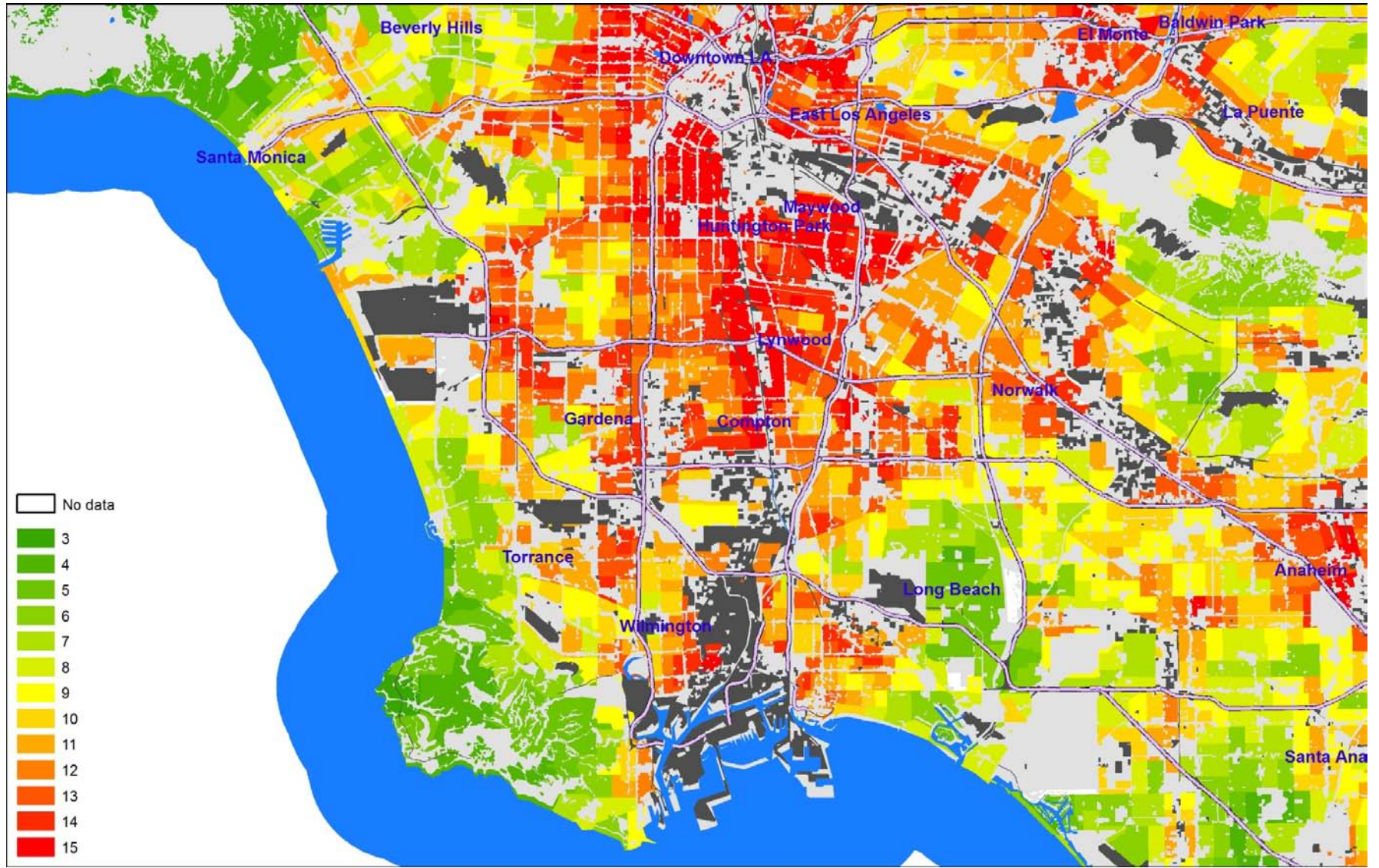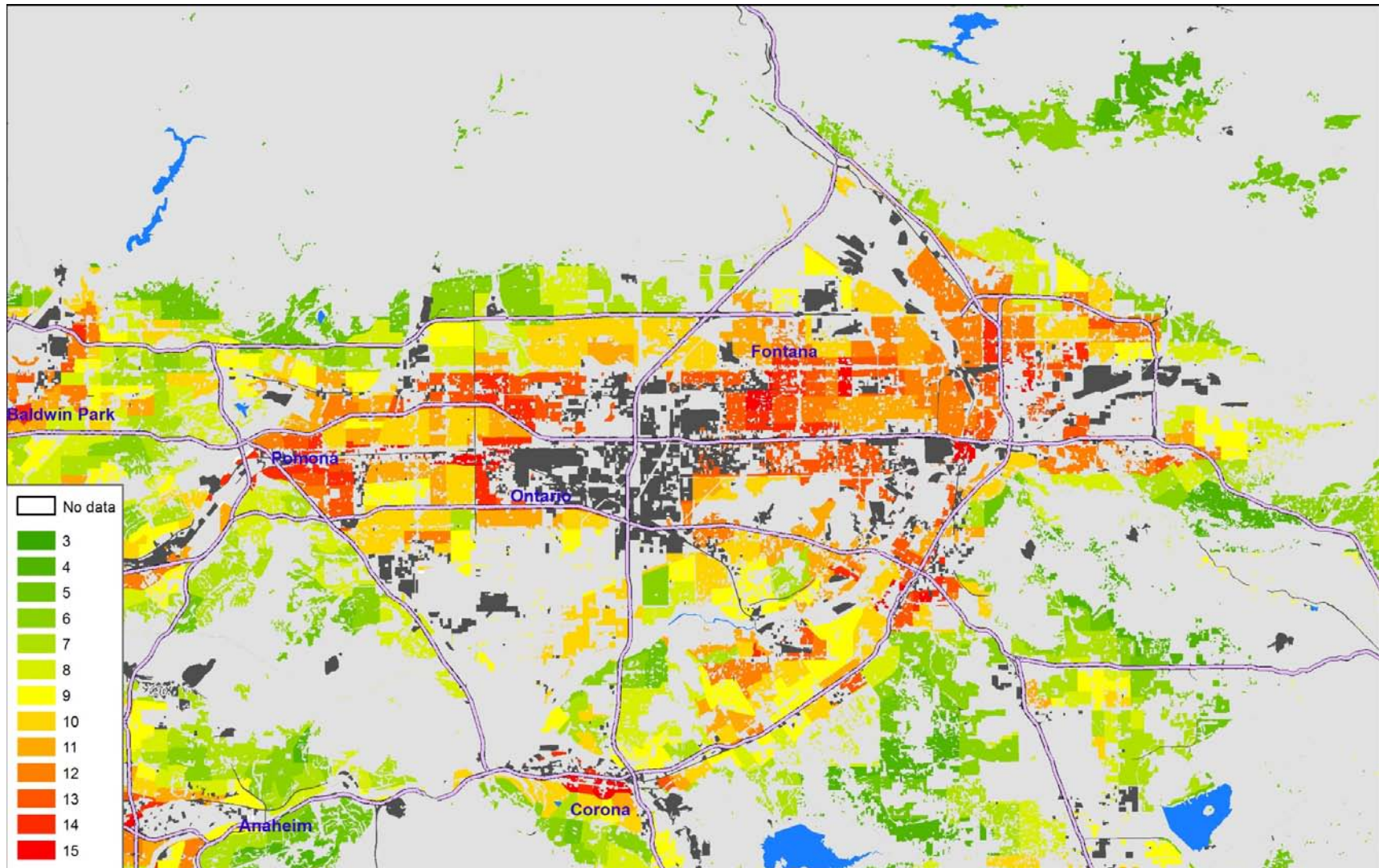
Figure 4.15: Map of the Total Cumulative Impacts score for the South Los Angeles area

Figure 4.16: Map of the Total Cumulative Impacts score for the Inland Empire

Chapter 5. Utilizing the EJSM to Consider Power Plant Siting

Introduction to the problem and previous work

One potential application of the EJSM might be to inform siting decisions for electrical power plants and/or other facilities. Funding for development of the EJSM was provided by the CEC specifically for evaluation of the method's utility in assessing the EJ implications of siting.

The CEC was established by the state legislature in 1975 to streamline the permitting process for energy facilities in California, and it holds the exclusive authority[45] to certify the locations of new power plants of 50 megawatts or more. State law requires that EJ be taken into consideration in the siting process, and the CEC developed its current "Staff Approach to Environmental Justice."[46] That approach follows U.S. EPA and California Environmental Quality Act (CEQA) guidelines for evaluating and addressing EJ concerns. Current CEC practice related to a power plant permit request is for staff to conduct a CEQA analysis around EJ issues for the proposed site. This EJ analysis consists of three components: demographic screening, public outreach, and impact assessment.[47]

- The demographic screen evaluates census demographics within a six-mile buffer around the proposed site[48] following the 1998 U.S. EPA Guidance in defining an EJ community in terms of threshold values for poverty and percent minority residents. For polygon centroids within the buffer, census blocks with populations that exceed 50% minority, or block groups exceeding 50% low income are considered "environmental justice populations."[49] These two values are highly

---

[45] Granted by the Warren-Alquist Act

[46] April 2000, in accordance with Senate Bill 115 and Executive Order 12898. The steps involved in this approach have not always been applied. On December 31, 2001, during the California electricity crisis, the Governor issued emergency orders that allowed for expedited permitting, where projects were exempted from the California Environmental Quality Act, with no environmental justice analysis before permitting, and limited public hearings. Furthermore, Executive Order D-40-01 allowed natural gas-fired plants to operate that did not meet air quality regulations (Latino Issues Forum 2001). Currently, new operating licenses have only been issued through a 12-month review process or the Small Power Plant Exemption process (http://www.energy.ca.gov/sitingcases/index.html).

[47] The information presented here and below is adapted from the CEC's website at: http://www.energy.ca.gov/public_adviser/staff_env_justice_approach.html

[48] The center of proposed site is used for projects less than 20 acres in area, and from the project boundaries for larger facilities.

[49] Council on Environmental Quality and National Environmental Policy Act definitions: minority – sum of self-identified American Indian or Alaskan Native, Asian or Pacific Islander, Black not of Hispanic origin, and Hispanic; low income is defined using the poverty thresholds from the US Census CP-60 calculation, calculated annually at block group level. CEC employs a consultant to monitor and recalculate these race and income measures annually. The guidelines also allow for consideration of areas within the buffer if the percentage of minority or low income population exceeds the percentage of that population in the general population (http://www.energy.ca.gov/envjustice/staff_env_justice_approach.html).

correlated in California. If there is at least one census block within the 6-mile buffer that has a majority minority population, the CEC will actively consider environmental justice in the context of the proposed project.[50]

- The <u>public outreach</u> component, dissemination of information about the proposed project to all local media outlets and public libraries, is conducted in the earliest phases of the project. The CEC's Public Adviser and other CEC staff contact community leaders and organizations to inform them of the project and its details, the process for public participation, and the results of the staff EJ analysis, and they also answer any questions. Local public participation hearings and workshops with language translation are also provided.

- The <u>impact assessment</u> component involves a CEC staff description of the existing setting, analysis of any "unique circumstances" of the affected population, identification of the project's direct, indirect and cumulative impacts, and the assessment and recommendation of appropriate measures of mitigation. This component also determines whether the project creates an unavoidable significant adverse impact on the affected population and, if so, considers whether the impact is disproportionate. It is CEC policy to determine mitigation measures to be applied based on project emissions, and it "mitigates ton for ton" in an effort to ensure that everyone is considered equally.

The question was whether our approach could yield different and/or more sensible results than the current CEC method. To evaluate this, we selected a hypothetical case by mutual agreement: a proposed power plant in South Gate that was never sited. As it turns out, the project was ultimately withdrawn due to community opposition, largely on the grounds of EJ, making it a good case study by which to illustrate whether the EJSM would have highlighted such concerns.

Materials and Methods

<u>Data</u>

In 2000, Sunlaw Energy Corporation proposed the Nueva Azalea power plant project for the city of South Gate, a predominantly minority community about five miles south of the city of Los Angeles. The proposed 550-megawatt, natural gas-fired combined cycle power plant was to be built on a 13.5-acre site at the eastern edge of the city limits. The site is bound by Southern Avenue on the north, East Frontage Road of the 710 Interstate Freeway on the west, and Garfield Avenue and Miller Way on the east; it is largely surrounded by other developed industrial properties on portions of the south, east and west.

---

[50] Dale Edwards, CEC Environmental Justice Coordinator, January 23, 2008

Although power plant projects typically face local resistance when sited in densely populated urban areas, Nueva Azalea seemed timely and promising. Environmentalists were excited because this power plant would be unique in California, using "SCONOx" technology to significantly reduce toxic air emissions. This technology had previously been used in power plants no larger than 33-megawatts, with Nueva Azalea planned to be 17 times larger. Labor unions supported the proposed plant because of the promise of new jobs. At the time it was proposed, a Governor's executive order streamlined the CEC power plant approval process to 21 days in response to statewide energy demands, also raising hopes for quick approval. The "energy crisis" dominated newspaper headlines. Power plant construction quickly became a hot topic and the political climate was geared toward building power plants no matter what the monetary or environmental costs.

However, local opposition came from the mayor and local community organizations who argued that the project placed an unfair burden on South Gate, an area already exposed to many other stationary emissions sources, heavy truck traffic from local industry and nearby freeways, and very poor ambient regional air quality due to its geographic location in one of the most heavily polluted parts of the South Coast air basin. They argued that, regardless of how "clean" the technology, Nueva Azalea was an additional source of air pollution that would add to the cumulative local air quality burden and should not be sited in South Gate. Sunlaw campaigned in support of Nueva Azalea, highlighting the local tax dollars the project would provide and promised to fund neighborhood improvements and provide local scholarships.

Sunlaw believed that many residents were generally supportive of the power plant, and agreed to a non-binding advisory referendum to allow residents to vote on whether they wanted the power plant. Sunlaw agreed to end the project if the referendum voted down. The residents' vote rejected the project by a 2-1 margin, and Sunlaw kept its word and abandoned Nueva Azalea in South Gate, indicating that it would look for other communities interested in hosting the project. In March of 2001, Sunlaw requested a six-month suspension of their application to CEC; after further suspensions in September and November, the company withdrew its application.

In retrospect, considerable resources were expended by Sunlaw and, to a lesser extent, by the CEC on a project proposed for a site that might have been identified early by the CEC as having serious EJ concerns had a useful EJ screening method been available at the time. Using this information, the CEC could have anticipated local resistance to the project and directed the applicant to consider alternate locations early in the process, promoting a more positive outcome for all.

To conduct an analysis of this question, the data required were basic census information on race and poverty as well as a geocoded list of the existing >50MW power plants sites in this region, as we assumed that they represent a good proxy for locations attractive for power plants. The latter data were taken from a 2007 CEC database of all operational power plants in 2007,[51] including all natural gas and coal-fired power plants producing at least 50 online megawatts. Once the set of facilities was determined, the next step was to

---

[51] This database can be found at: http://energyalmanac.ca.gov/electricity/index.html

accurately geocode them. We did this by cross-referencing address location from several sources, including an EPA geospatial shapefile that purportedly contains all the facilities that are listed in the EPA's Facility Registry System (FRS)[52] and an ARB Emissions inventory. A final quality assurance check was carried out by generating a Google Earth (.kmz) file of the facilities and spot checking each one in Google Earth to be sure that it landed on what appeared to be the appropriate location. If it did not, then we first scanned the immediate surrounding area to find an appropriate facility and collect its geographic coordinates. When none was found, we resorted to detailed web searches to locate the facility through a variety of sources including the parent company's website, articles from web media sources, or permit application and review documents.

Analysis

We first evaluated the Nueva Azalea site for EJ concerns using the EJ analysis method currently used by the CEC in power plant siting decisions. We then employed the EJSM to analyze cumulative impacts and community vulnerability in the area surrounding the proposed site. This comparison was further extended to the locations of all power plants >50MW to examine the different conclusions and information offered by the two methods. We did this by first performing a modified form of the standard CEC EJ analysis to the Southern California region, using the preferred 1- and 6- mile buffers. We then compared these results to those derived at the same buffer distances using results from the EJSM. Finally, we applied both methods (our method and a modified form of the CEC method) to the actual locations of the existing >50MW power plants sites in this region, assuming, as noted above, that they represent a good proxy for locations attractive for power plants.

Results

To understand our results, it is important to first consider the CEC method of designating a site as having EJ concerns if it is either fifty percent minority or fifty percent low-income (defined as the percent poor). As it turns out, this is not much of an effective screen. First, we would argue that the underlying information – poverty and minority population – are by themselves insufficient to characterize an EJ community given what we now know about the determinants of environmental injustice (see Chapter 2), and they are of no utility in assessing cumulative impacts. Moreover, at least in Southern California, only one of these two criteria even plays a role in the results of the CEC demographic screen: in the region, while about 58 percent of block groups have a majority minority population, less than 2 percent have more than half the population below the poverty line – and among the 185 block groups that do, all but 6 are majority minority.

Thus, the minority criterion effectively "overscreens" while the poverty criterion "underscreens," and the result in Southern California is that it identifies virtually every

---

[52] The shapefile was downloaded from: http://www.epa.gov/enviro/geo_data.html

single existing power plant of the size that places it under CEC authority (> 50 MW) as having EJ concerns (see discussion below), with income playing virtually no role. Thus, the "demographic screen" is not really very useful for screening at all; in a moderate or densely populated area, it designates practically everywhere as having EJ concerns.

This can be seen by looking at our modification of the CEC method. While it is not currently used in this way, we can use the CEC demographic screen to produce a "screening map" (Figure 5.1; all figures are at the end of the chapter) of the Southern California area. In this map, all the red polygons are areas where the EJ criteria are exceeded.

According to this map, Nueva Azalea would have been literally surrounded by areas of EJ concern; on the other hand, there is little that would distinguish the Nueva Azalea site as being worse or better off than many other potential sites in Southern California. There is nothing represented in that map or criterion, for example, that accounts for the current pattern of disparity in exposure, risk, and vulnerability. The current method would imply a need for public outreach and impact assessment to address any concerns but this would be true of virtually every location in Southern California.

To illustrate how the EJSM could be used as an alternative, Figure 5.2 shows the Nueva Azalea project location against a background of color-classified values of the EJSM-generated CI scores at the 2000 census tract level. As noted in the previous chapter, higher CI scores indicate greater cumulative impacts and community vulnerability.

Note that the area surrounding the proposed Nueva Azalea site has relatively high CI scores. Also shown on the map are all other operating Southern California power plants, nearly all of which are either located in or directly adjacent to areas of high cumulative impact and, therefore, EJ concern.[53] From the map, it is clear that the CI scores in the vicinity of the proposed Nueva Azalea site are much higher than those for the other power plants in the region – the map, in short, is showing more variation than the implied CEC method and therefore gives more nuance to the initial screen.

To give each location and buffer an overall CI score based on the CI scores of the affected tracts, we apply the tract-level scores to all census blocks within each tract and take the population- weighted average CI score across blocks that are within each buffer circle, with blocks designated as being within the buffer circle if their centroid – or geometric center – falls inside of it. This method allows for portions of tracts to be included in the average CI score in cases where the buffer circle intersects a census tract, and it represents the CI score for the average person within the buffer circle rather than the simple average of the tracts that are affected (which seems to be a reasonable approach in that it makes use of the actual population within the buffer).

Figures 5.3 and 5.4 offer the average CI score within the one mile and six- mile buffer distances for all power plants shown in the analysis; while a superior comparison would be to all alternative locations, we did not have that information and thought that the

---

[53] 2007 CEC list of operational power plants in California with generating capacity >/= 50 megawatts

existing power plants would represent a reasonable range of such alternative locations. At both buffer distances, the area has a very high average CI score - 13 out of a possible 15 points - which is much worse than most all other power plants in the region. Indeed, the proposed Nueva Azalea site has the fourth highest CI score at the one- mile distance and second highest at the six- mile distance. From this comparison, one might conclude, on EJ grounds, that the proposed Nueva Azalea site was among the least desirable locations for a new power plant in Southern California (of course, there are many other reasons why a power plant location may or may not be attractive; we are not suggesting that EJ is the only concern but rather are simply offering a different way to measure that concern).

Interestingly, the Nueva Azalea site is outranked in both buffer comparisons by only one power plant- the Malburg plant in Vernon, which is located about five miles northwest of Nueva Azalea. That plant has been the subject of concern by the surrounding community and there was significant opposition when there was a recent proposal to site a new power plant near the facility.[54] Indeed, that project was halted by the community – as was Nueva Azalea – with the City withdrawing its application with the CEC on September 28th, 2009.[55] This pattern suggests that the screen we have developed could be useful in predicting where potential siting plans might adversely affect already overburdened communities and, in doing do, suggest that either special attention in siting or special outreach to affected groups is appropriate.

Discussion

To compare our results to the current CEC requires utilizing a modified version of the CEC approach. We modify it because, as noted above, using the CEC's binary approach – in which EJ concern is triggered by meeting the 50 % minority or 50 % low-income threshold – means that every existing plant or potential location has some concern. We instead wanted to see if we could at least rank concern along these two dimensions and so we derived a method where we drew the buffer and calculate a "score" similar to our CI score. Essentially, this is the percent minority or percent poor within each buffer, derived by counting up the numbers in each of these categories (and the overall population) associated with block group centroids that fall within the respective buffers.

Those results are depicted in Figures 5.5 through 5.8, inclusive. Using the six- mile buffer distance (Figures 5.6 and 5.8), the Nueva Azalea site ranks relatively high both in terms of minority populations and percent living in poverty, but at the one- mile distance (Figures 5.5 and 5.7), it ranks lower in the distribution, making it appear to have fewer EJ concerns, at least when compared to other facilities in the region.[56] In contrast, the EJSM

---

[54] CI scores cannot be shown for High Desert and Long Beach (CA) P&G at the 1 mile distance, because of missing data for indicators of social and health vulnerability in some census tracts falling within this buffer distance.

[55] See *Community Stops Dirty Power Plant in Southeast Los Angeles:* http://www.cbecal.org/campaigns/soeastla.html.
[56] The percentage minority population and poverty rate is not shown for High Desert at the 1 mile distance because there are no people residing in the census tracts falling within the buffer distance.

clearly identifies the area surrounding the proposed Nueva Azalea site as a problematic location in terms of EJ concerns, mostly because it is using several other social factors as well as cumulative exposure indicators, including proximity to other point source hazards and measures of air-related health risk.

The above analysis suggests that the EJSM may offer more nuance and might be usefully incorporated into the CEC siting analysis and permitting process to allow for a comparison of the relative merits of various siting options on an EJ basis. If used in this way – calculating the average of CI scores for a buffer distance around a set of potential site locations – the EJSM could be implemented as a means to inform the CEC process of existing EJ concerns as determined by cumulative impacts and social vulnerability, as well as to suggest the level of potential local objections to power plant siting. The geographic specificity of the EJSM, yielding CI scores at either the CI polygon (census block or smaller) or tract level of geography, also eliminates the need to focus only on an arbitrary buffer distance. The EJSM could further inform the siting process by evaluating the distribution of CI scores at the finer levels of geography (CI polygons or census tracts) to examine the relative contributions of the different indicator classes - hazard proximity, health risk, and social/health vulnerability – to the overall score, in this way obtaining a more comprehensive and nuanced understanding of the relative level of EJ concern that is highly geographically specific.

To help CEC staff become familiar with the EJSM and experiment with its use, we can provide a tool and spatial data layers that allow the user to quickly specify a buffer for any geographic location and select all features located within the buffer for an exploration of CI metrics and an analysis of average population-weighted CI scores. Information on the tool is provided in the Appendices.

Conclusion

The CEC EJ demographic analysis as currently applied is a broad test used to identify areas of "EJ concern" and, in part, define the degree to which the CEC is compelled to carry out measures to address EJ concerns. It does not make comprehensive use of pertinent data available at the state level, nor can it go beyond a site-specific yes/no decision and compare the suitability of alternative locations that might host a new plant in terms of EJ and cumulative impacts. The underlying information – poverty and minority population – are by themselves insufficient to characterize an EJ community given what we now know about the determinants of environmental injustice, and they are of no utility in assessing cumulative impacts. Moreover, at least in Southern California, only one of these two criteria even plays a role in the results of the CEC demographic screen: in the region, while about 58 percent of block groups have a majority minority population, less than 2 percent have more than half the population below the poverty line – and among the 185 block groups that do, all but 6 are majority minority.

In contrast, the EJSM offers a practical and useful alternative to the current CEC staff approach. In addition to consideration of race and poverty, the EJSM incorporates wide

variety of environmental, health, and socio-demographic information identified in the research literature as significant determinants of disparate impact and environmental injustice. It is of considerable utility for screening purposes and decision making around siting, particularly given that it is implemented at a geographic resolution that is small enough to be useful in, for example, incorporating EJ and cumulative impact information into a comparison of the relative merits of alternative site locations (including a "no project" alternative), as is already required under CEQA and CEC regulations.[57] Using the EJSM, neighborhoods within a region can be compared in terms of existing cumulative impacts.

Focusing only on race and income, as per the CEC demographic screen, has the potential of highlighting areas that are not overburdened by air pollution exposure or health risk. For example, the demographics surrounding the Nueva Azalea site should have alerted both Sunlaw and CEC to expect significant local objection to the proposed power plant. Applying the six- mile buffer zone CEC demographic screen to the five power plants with the highest overall percent minority (this includes Nueva Azalea) shows a range in the percentage minority population from 89 to 97 percent, much higher than the remaining power plants in this region which have minority populations below 80 percent. Using this single measure would have suggested potential EJ concerns to the CEC at all five locations. However, the ranking of these five power plants in terms of average CI scores (Figure 5.4) shows that two of the five – Colmac Energy Inc. and Coachella – are at the lower end of the regional distribution of average CI scores. Similar comparisons can be made using poverty rates and with both buffer distance thresholds.

Because the EJSM relies on a wide variety of indicator metrics that are combined into the final CI score, these results are both reliable and robust in terms of the actual determinants of EJ spatial patterns, and less sensitive to small variations in a single measure, such as the percentage minority population or poverty rate. The EJSM also uses a quantile ranking procedures to calculate CI scores, unlike the CEC demographic screen that uses specific threshold which identify EJ concerns in a binary fashion. In our view, there may be some utility to continuing to test the EJSM against hypothetical and actual cases to see what additional information it can provide policy makers, regulators, and community members.
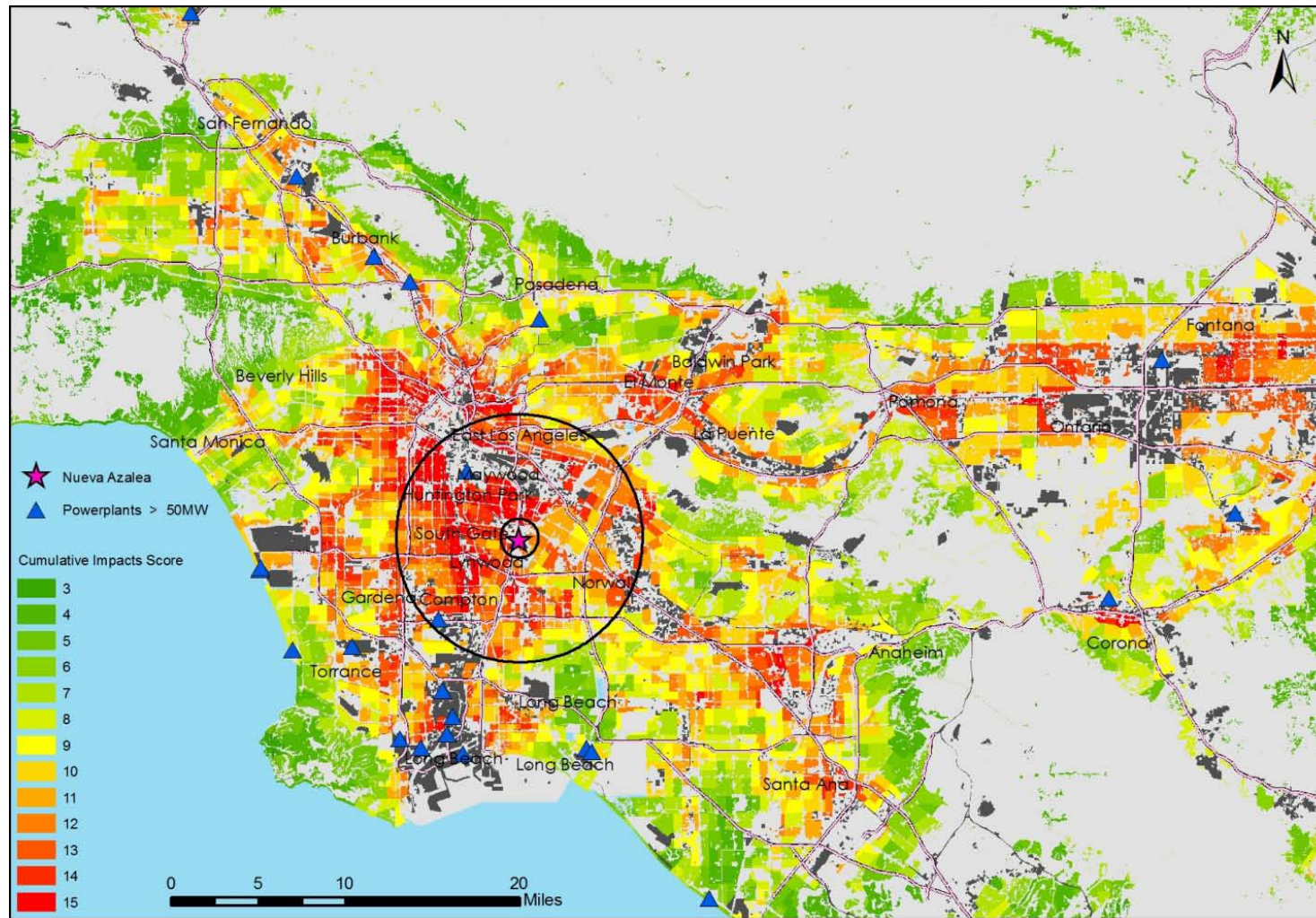
---

[57] California Code or Regulations, title 14, § 15126 and title 20, § 1765.

Figure 5.1: CEC "Demographic Screen" map of census block groups overlain by the location of the proposed Nueva Azalea site and one and six mile buffer radii (areas marked red meet the criteria of hosting "environmental justice populations" – exceeding either 50% minority or 50% low income and buffer radii of one; also shown are locations of all existing power plants (>/=50MW) in the southern California region. Areas in dark gray are industrial while areas in light gray are open space/vacant land uses).
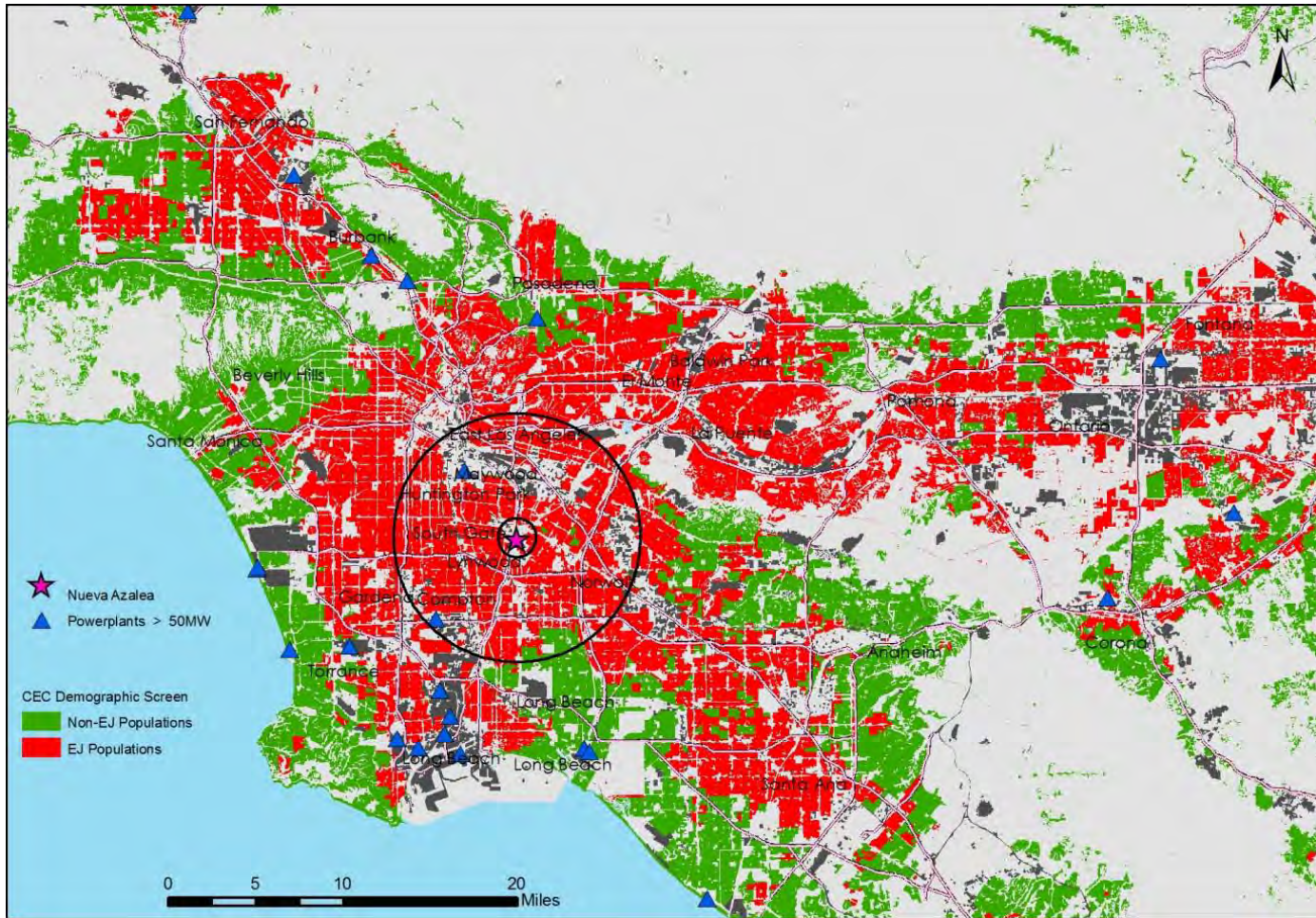
Figure 5.2: EJSM map of cumulative impact score overlain by the location of the proposed Nueva Azalea site, and one and six mile buffer radii (also shown are locations of all existing power plants (>/=50MW) in the southern California region).

Figure 5.3: Rank of average population-weighted CI scores for census block groups located within a one mile circular buffer surrounding all southern California power plants. Nueva Azalea is shown in red.

Population Weighted Cumulative Impact Score Within 1 Mile of Facility

Powerplants in 6-County SCAG Area

Figure 5.4: Rank of average population-weighted CI scores for census block groups located within a six mile circular buffer surrounding all southern California power plants. Nueva Azalea is shown in red.
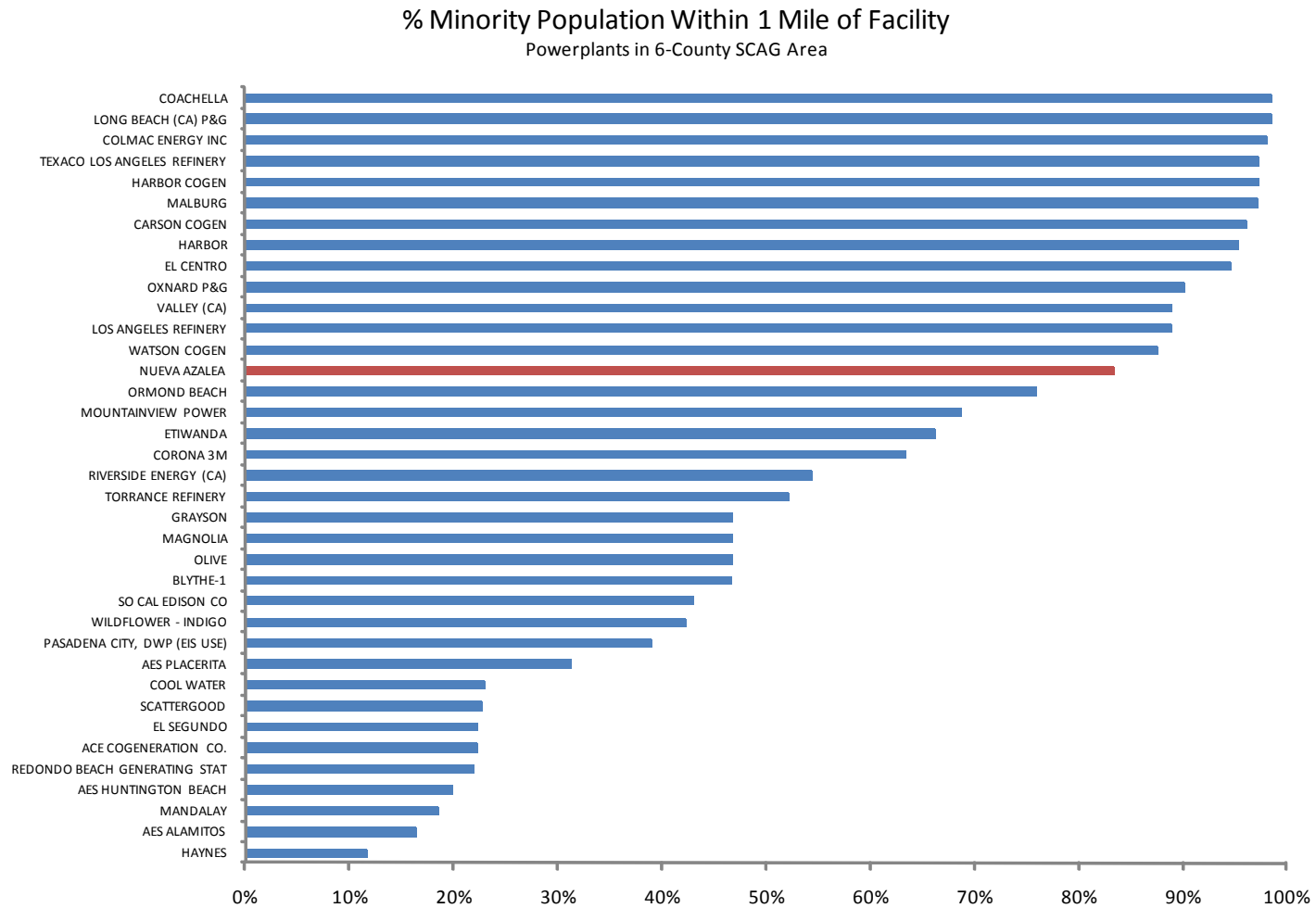

Population Weighted Cumulative Impact Score Within 6 Miles of Facility

Powerplants in 6-County SCAG Area

Figure 5.5: Rank of percent minority residents for census block groups located within a one mile circular buffer surrounding all southern California power plants. Nueva Azalea is shown in red.

% Minority Population Within 1 Mile of Facility

Powerplants in 6-County SCAG Area

Figure 5.6: Rank of percent minority residents for census block groups located within a six mile circular buffer surrounding all southern California power plants. Nueva Azalea is shown in red.

% Minority Population Within 6 Miles of Facility

Powerplants in 6-County SCAG Area

Figure 5.7: Rank of percent residents below the federal poverty level for census block groups located within a one mile circular buffer surrounding all southern California power plants. Nueva Azalea is shown in red.
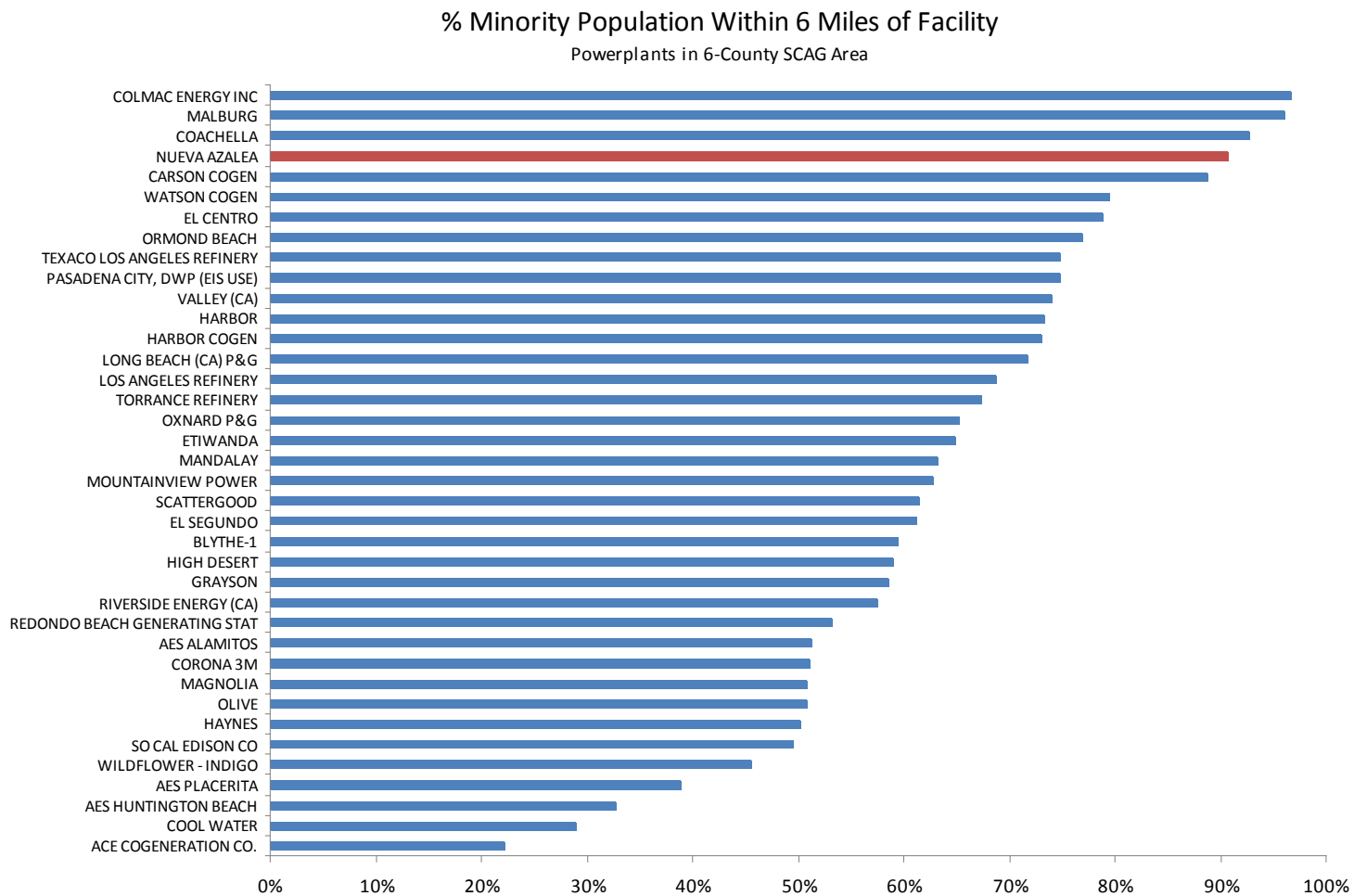
% Population Below Poverty Level Within 1 Mile of Facility
Powerplants in 6-County SCAG Area

Figure 5.8: Rank of percent residents below the federal poverty level for census block groups located within a six mile circular buffer surrounding all southern California power plants. Nueva Azalea is shown in red.
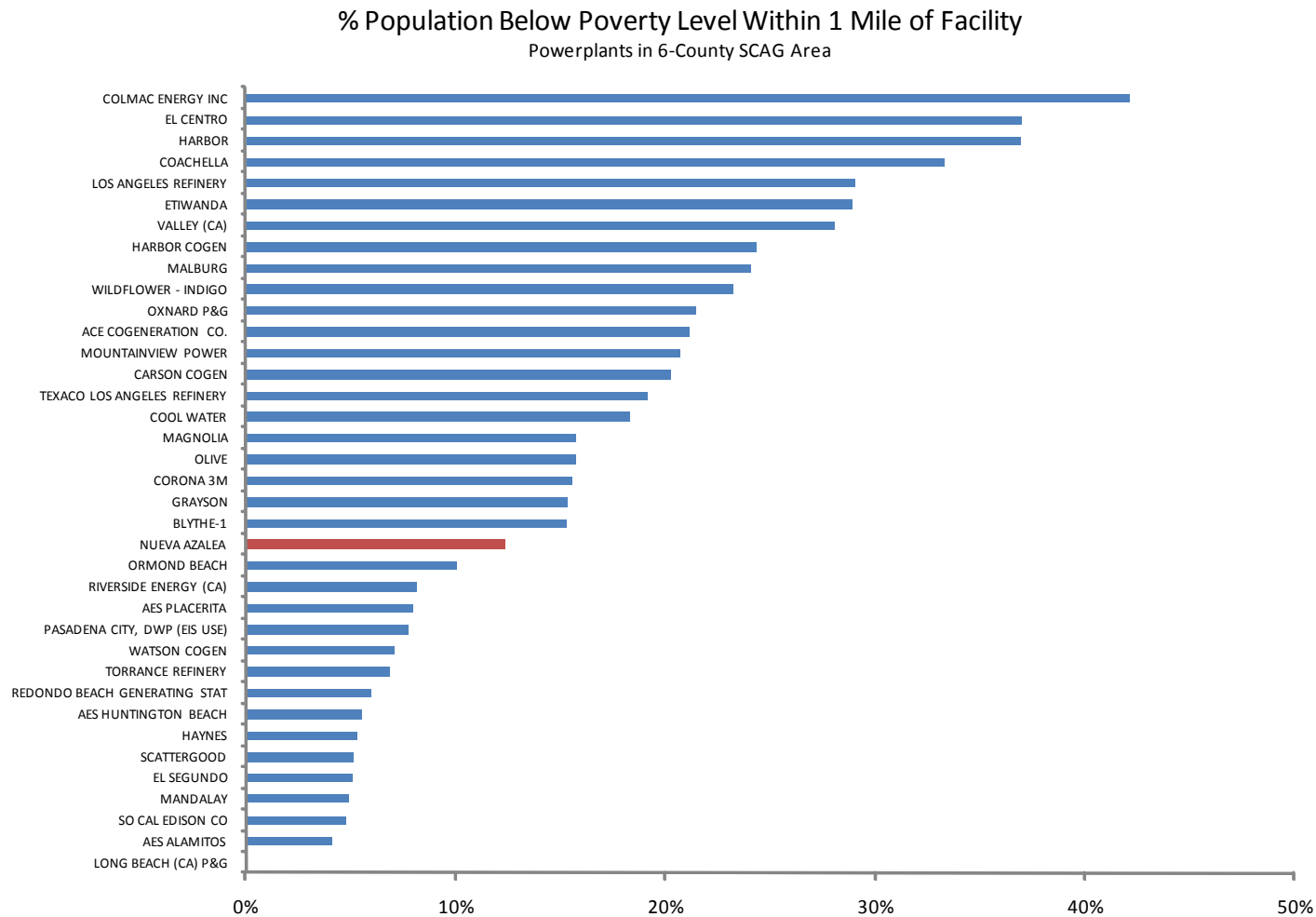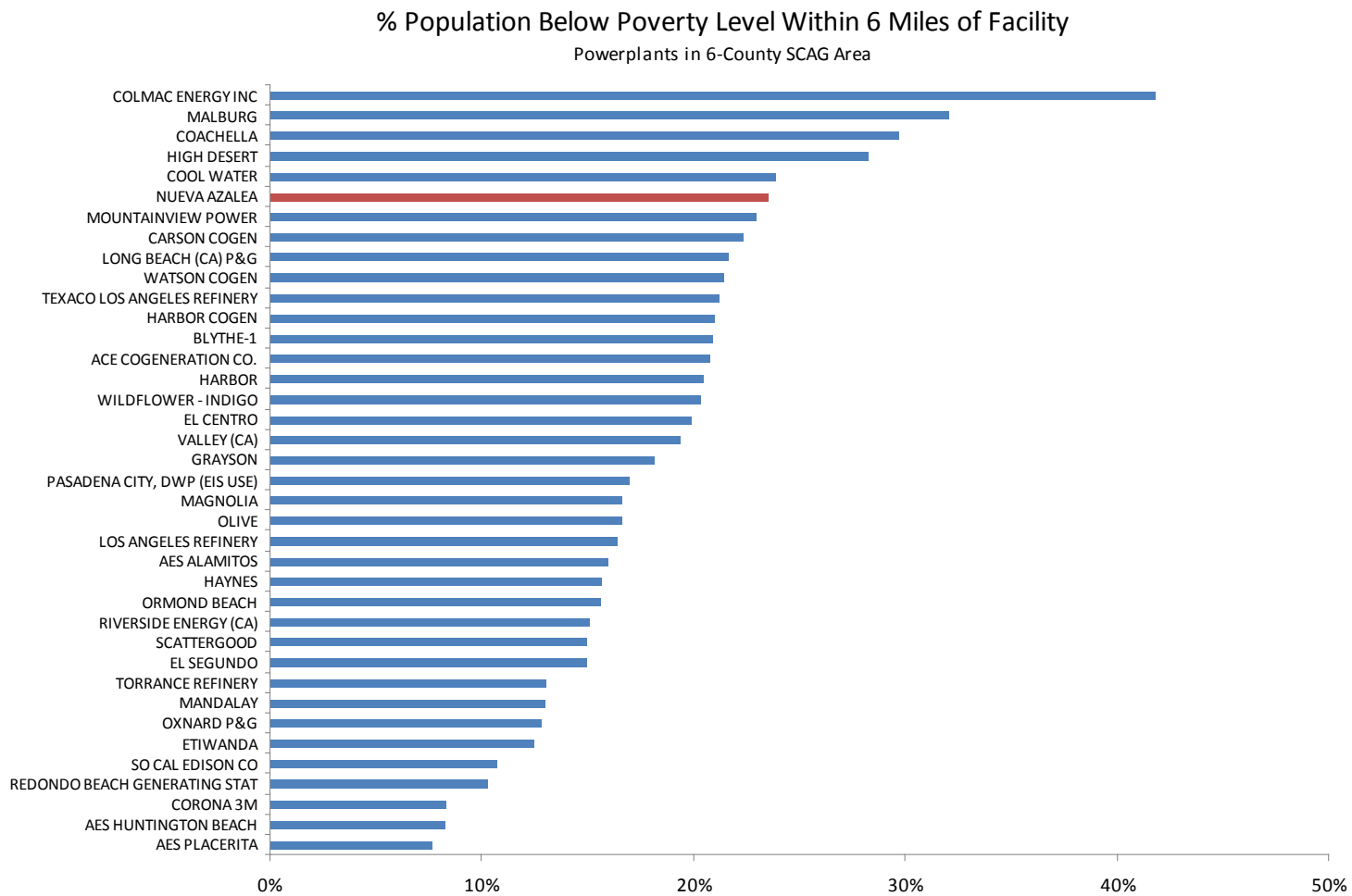

% Population Below Poverty Level Within 6 Miles of Facility

Powerplants in 6-County SCAG Area

Chapter 6. "Ground truthing" in the Hegenberger Corridor of Oakland

Introduction

We designed and implemented a community-based participatory "ground truthing" microstudy in partnership with Communities for a Better Environment (CBE) based in Oakland. A key tenet of community-based participatory research is extensive community involvement in the research process, which can include the design of study protocols, data collection, interpretation of data, and dissemination of results (Israel, Schulz et al. 1998). Community-based participatory research methods seek to incorporate community knowledge and engagement to improve scientific methods in public health research and to ensure that rigorous scientific inquiry is directly linked to policy outcomes and regulatory action (Minkler 2000).

The primary purpose of the microstudy was to locate and map the community's perception of cumulative impacts, including air quality hazards, both stationary facilities and mobile source emissions that may not be identified through traditional regulatory oversight, and sensitive receptors of importance to those who live in the area. This information was then compared to air quality hazard and land use information available from state government agencies. CBE facilitated the research team's collaboration with East Oakland residents who have been active with regulatory agencies on air quality concerns in their neighborhood. Specifically, CBE staff and East Oakland residents were trained to:

1) Assist the project team in identifying and locating environmental hazards and sensitive land uses of concern to augment our understanding of localized approaches for assessing cumulative impact and

2) 'Ground truth' and augment information from existing ARB emissions inventories regarding the existence and location of area and point emission sources.

We worked with CBE through their Leadership Training Institute for community organizers to train community members on the concept and science of cumulative impacts of environmental and non-environmental stressors, as well as on the specific skills needed to collect locational information on emission sources and susceptible receptors. Much of the focus was on verifying sources contained in CARB's emissions inventory and also on identifying pollution sources not captured on current data systems. We collaborated with CBE to develop the data collection protocol and the training curriculum and held several meetings with community members to discuss maps of emission sources based on inventory data sources and to identify the scale and scope of the field study area.

The study area, defined by CBE and referred to as the Hegenberger Corridor, is located in East Oakland near the Oakland International Airport (Figure 6.1). When this area was proposed, we met with CBE representatives to conduct a walk-through of the neighborhood and prepared an atlas of maps of the area using existing databases showing locations of area and point air pollution emissions sources and land use patterns. We identified a study area that was of appropriate size based on the number of community members available to do field work in the prioritization of areas of community concern.

CBE hired a program/community organizer, Nehanda Imara, who lives in East Oakland, as well as staff researcher/scientist Anna Yun Lee to focus on the Hegenberger Corridor project. They led efforts to recruit community members and integrate the microstudy with CBE's Summer EJ Leadership Training Series[58]. This integration was of significant benefit both to the community and the success of the microstudy as it provided both an educational and training forum and recruitment opportunity for community members who were interested in working directly on EJ issues where they live, and in translating study results into strategies for regulatory and policy engagement related to local air quality and emissions sources. The leadership training ran through Summer and Fall of 2007 and consisted of 10 modules that included content education and advocacy training, as well as some scientific and regulatory background training. A CARB staff member attended and participated in some of these training sessions. The presence of a CARB official was highly significant both in educating community participants about the regulatory process and giving them opportunities to discuss their local concerns.

The main goal of this community-based participatory research project was to compare the landscape of air pollution exposure and land use as derived from inventory databases with actual on-the-ground observations and documentation by community members in the neighborhoods they live in. The project was designed to assess the assertion, often articulated in EJ communities, that the cumulative impacts problem of multiple environmental hazards are not fully accounted for in the databases and emissions inventories that are traditionally used by state regulatory agencies to inform decision-making as well as program and regulatory enforcement activities. Often, there are a greater number of air pollution emission sources in neighborhoods that are not captured and monitored by state government regulators. This situation can occur primarily because only facilities that are under the regulatory jurisdiction of these agencies are required to report emissions and are reflected in state regulatory databases. Other smaller emitters, such as auto shops or truck parking areas, are not captured in these databases but may be concentrated in certain communities. Ground truthing can also help identify locational errors for facilities included within regulatory databases.

EJ communities also contend that CARB's definition of "sensitive receptors" is too restrictive and should include other locations such as, retirement and assisted living facilities, churches, and youth recreation program locations that may not necessarily be located in urban parks. It is generally accepted that those most vulnerable to the negative

---

[58] This training series was funded by The California Endowment to support capacity-building, organizing and advocacy work in the Oakland-Hegenberger Corridor in January 2007.

health effects of air pollution are the very young and elderly, as well as those with underlying disease conditions. CARB defines these sensitive populations for operational purposes in terms of sensitive land uses in the 2005 Air Quality and Land Use Handbook (CARB 2005) and recommends a distance buffer of 1,000 ft. from most air quality hazards for siting of new sensitive land uses.

The databases that we used for comparison with community data collection efforts in the Hegenberger microstudy were provided to our research team by CARB staff for integration into our EJSM. These data sources include: 2001 CHAPIS facility locations; chrome plating facility locations; 2004 AB2588 "Hot Spots" air toxics emission sources; and 2004 hazardous waste treatment, storage, and disposal facilities from the California Department of Toxic Substances Control.

For data on the location of sensitive receptors, CARB initially provided us with an incomplete spatial data layer of school locations, which we subsequently replaced with school locations that we geocoded from address lists provided to us by the California Department of Education. We also used locational information on healthcare facilities from the California Spatial Information Library. We obtained the locations of childcare facilities by geocoding from address lists purchased from Dunn and Bradstreet business information service (SIC codes 8350 and 8351).

Originally, the microstudy included two phases. The first phase encompassed the field work with organizers who walked specific neighborhoods and marked, using hand-held GPS devices, the location of hazards and sensitive receptors. This information was then integrated into our existing maps of air quality hazards and sensitive receptors from regulatory data sources. The first phase of the microstudy was conducted during the summer of 2007. The second phase, which was added on by CARB staff after the initiation of this Hegenberger microstudy project, included a pilot $PM_{2.5}$ monitoring component which was carried out in Fall of 2008. CARB staff requested that we expand the microstudy project to include a collaboration with Dr. Kirk Smith of UC Berkeley's School of Public Health that would entail particle monitoring. Specifically, CARB staff asked that we use the Hegenberger collaboration to field test Dr. Smith's UCB-L PM2.5, portable air quality monitors that use commercial smoke-alarm technology and components and custom software (Edwards, Smith et al. 2006). These devices were originally developed by Dr. Smith's research team for deployment in developing countries where fine PM levels are significantly higher than in the United States. CARB supported efforts by Dr. Smith's team to adapt the UCB-L PM2.5 portable air quality monitor for ambient and indoor monitoring in California.

CBE and our community member collaborators were enthusiastic about the prospect of field testing portable, low- cost particle monitors as a way to assess the feasibility of more widespread use of these monitoring tools in other EJ communities in the Bay Area. Therefore, we agreed with CARB staff that funding would be provided for one of Dr. Smith's doctoral students to train and supervise community members in the deployment and use of the UCB-L PM2.5 monitor, provide technical support and logistics during field data collection, and to extract the raw data from the monitors and analyze this data

as part of the Hegenberger microstudy. However, as the project progressed, Dr. Smith informed us of significant challenges faced by his lab and delays in being able to successfully retrofit the UCB-L PM2.5 monitor for field testing in Hegenberger. Performance testing in the lab and field showed that the UCB-L PM2.5 monitors had unacceptably high levels of detection for Bay Area air quality conditions. As a result of these challenges, we were compelled to scrap plans to deploy and pilot test the UCB-L $PM_{2.5}$ monitor for the Hegenberger project. Instead, we implemented a smaller scale $PM_{2.5}$ monitoring phase of this project using monitoring equipment provided by CARB.

Materials and Methods

Data and Training

GPS Locational Work

CBE's Leadership Training Institute encompassed a basic education module for community advocates on the science of air pollution and health, cumulative impacts, environmental health inequalities, and the combined health effects of environmental and psycho-social stressors on community health. This background training module was also leveraged to recruit community residents to assist with the ground truthing data collection activities of the Hegenberger Microstudy.[59] The research team then conducted three trainings to prepare CBE staff and community leaders for the ground truthing field work activities. At the first training, CBE staff and community leaders reviewed data inventory maps provided by the research team to assist them in generating a list of hazards and sensitive receptors to inventory in their East Oakland neighborhoods. Hazards were identified as stationary businesses known to use or emit hazardous chemicals in their processes as well as places generating diesel truck traffic. Sensitive receptors were identified as places where populations particularly susceptible to the adverse effects of pollutant exposures-specifically children and the elderly- congregate and spend a lot of time, such as schools and senior centers. Table 1 shows the lists of the hazard/sensitive receptor categories generated by community leaders.

Many facilities of concern to Hegenberger residents are included in state regulatory databases, but the community list also includes hazards and sensitive receptors that are not included in these data sources. Interestingly, some types of facilities initially suggested for inclusion on the community's more expansive list were ultimately eliminated, such as home-based childcare, because of uncertainties regarding licensing and their precise locations.

---

[59] Participants in the field work included: Nehanda Imara, Organizer, Shana Lazerow, Staff Attorney, and Anna Lee, Staff Researcher/ Scientist, (CBE East Oakland Team); Wafaa Aborashed, Leslie Bowling, Jacquee Castain, Glenda Deloney, Gloria Moy, Maxine Oliver-Benson, Myrtle Washington (Community Leaders).

At the second and third trainings, the research team instructed participants on field work protocols. This included using GPS-enabled portable computers – Hewlett Packard Iraq HX 4700 handheld computers running the Pocket PC operating system, with Teletype CF Model 1653 GPS cards, and ESRI ArcPad v. 6 - to allow participants to automate field observations of environmental hazards and other information necessary to characterize the study neighborhood into a GIS spatial data format. The ArcPad software allowed users to see their location in real time on the display atop spatial data layers that were pre-loaded to provide geographic context for data collection, which included aerial photo coverage of the study area, streets, locations of known air quality hazards, and sensitive receptor land uses. The software also allowed easy recording of a location as an entry into an edit shapefile and a form to enter site attribute information.

Some community leaders had little to no experience with handheld PCs. Field teams took the devices home in between training workshops to practice and become familiar with their operation. The devices worked well for some community leaders, but presented some challenges for others in terms of ease of operation and reliability. For this reason, we backstopped the GPS data collection plan by equipping community users with notebooks containing template worksheets that allowed them to manually record the information needed to ground truth the location and characteristics of the hazards/sensitive receptors. These worksheets contained step-by-step instructions on data collection, data entry forms, maps of known facilities and land uses derived from state government agency databases, aerial photos, and detailed street maps with address ranges. This allowed community leaders to locate facilities manually without depending solely on the handheld computers if they did not feel comfortable with the technology or if the devices did not function adequately. This also enabled them to duplicate their field observations and ensure replication of data collection for future verification of site locations. In many cases, users recorded locations on air photos and with a street address, allowing us to verify the location later with address geocoding.

After training, the ten community leaders were organized in field teams of two, with each team responsible for conducting street-by-street canvassing of a portion of the study area, identifying and locating both hazards and sensitive receptors of concern. Each team was assigned to canvas a portion of the study area using street-by-street field observation; one block overlaps at the boundaries of field team areas were included to ensure that the census was complete. Teams were tasked to:

(a) Verify the location and correct information on all air quality
    hazards recorded in state government agency databases;

(b) Verify the location and correct information on all sensitive
    receptor land uses as defined by CARB (schools, childcare
    centers, playgrounds and urban parks, and healthcare
    facilities); and

(c) Locate and map any additional hazards and sensitive receptors
    that are not included in the regulatory agency databases, using

> the types of facilities identified in the training workshop as guidance.

The name, type of business, address, intersection, category (hazard or sensitive receptor), and other notes about the land use were recorded in a field notes template sheet. All this data was collected from information that could be attained from signs and observations of what could be seen occurring on the property. Teams also recorded observations about types of hazards to residents, specifically idling trucks, trucks passing through residential areas, and large containers on sites that may be filled with chemicals. Overlap of data occurred at boundaries of the designated areas; repetitions in data were omitted. Researchers also identified hazard locations recorded in state regulatory agency databases – CHAPIS facilities and AB2588 "hot spots" emissions sources - to verify locational accuracy.

PM$_{2.5}$ Monitoring

Despite the set-back of not being able to deploy the UCB particle monitors, CBE and Hegenberger community members were still eager to conduct an air monitoring project. Our contract manager, Dr. Álvaro Alvarado, assisted the research team in addressing raised community expectations by working with CARB's Research Division to lend handheld fine PM monitors for a pilot monitoring project in Hegenberger. Dr. Alvarado assisted the team in training CBE staff and community members to conduct air quality monitoring work during the month of October 2008. We held two additional training sessions for CBE staff and community members on how to use the DustTrak monitors to take ambient PM$_{2.5}$ measurements using a TSI Model 8520 DustTrak Aerosol Monitor. The DustTrak is a nephelometer that senses particle scattering of a laser beam and converts signals into a particle mass reading. The PM concentration circumventing the impactor is determined by measuring the intensity of the 90° scattering of light from a laser diode. The instrument sample flow rate is 1.7 L/min and an averaging time of one second is used. The instruments are calibrated at the factory with Arizona road dust (NIST SRM 8632).

CARB's experience with this instrument during their previous school bus study (Fitz, Winer et al. 2003; Sabin, Kozawa et al. 2005) paralleled that of other investigators (Ramachandran, Adgate et al. 2000; Chung, Chang et al. 2001; Yanosky, Williams et al. 2002). In particular, the greatest utility of the DustTrak is to obtain relative measurements of PM$_{2.5}$ with high time resolution, rather than relying on this instrument for absolute PM$_{2.5}$ mass. Instruments based on other methods, such as BAM monitors or tapered element oscillation, (TEOM monitors) do not have adequate time resolution for real-time measurements.

The air monitoring component of the Hegenberger microstudy was an unfunded add-on project that was not originally part of our contract agreement. However, because the expectations of CBE staff and Hegenberger community leaders were raised about the prospect of conducting air monitoring using the UCB PM monitors with support from Dr.

Smith's staff, we believed it was critical to work with community members to conduct some monitoring in sites of high concern. However, we did not have the resources to do the full scale monitoring work originally envisioned in collaboration with Dr. Smith's team.

To carry out our pilot monitoring work, four teams covered five sites of potential concern to residents because they were perceived to be near an emission source or a sensitive receptor (e.g. a school).[60] The protocol aimed to have team members conduct sampling over a two- week period on days when there was no rain or excessive moisture in the air, which can impact PM level readings. Members were asked to monitor three times in the morning during regular business hours, three times in the afternoon during regular business hours, and three weekend visits when the facility of concern was likely to be closed or not running at full capacity. In general, team members were able to follow the protocol and use the monitor and GPS devices (to mark sampling waypoints) with little problem, although one team had some difficulty covering all of the required sampling intervals due to conflicts with their work schedules.

Analysis

On the locational portion of this work, the data collected by CBE staff and community leaders was later transferred to the research team's GIS spatial database by merging the spatial data layers created in ArcPad with geocoded addresses from the field. Duplicates were identified and eliminated, and researchers subsequently visited and recorded the location of each site recorded in state regulatory agency databases using GPS to verify location accuracy. Mapping and spatial analysis was then performed, and the results were reported back to participants in a subsequent workshop to allow them to compare their maps to those created using only state regulatory agency data, and to revisit their hypotheses and discuss their results. On the monitoring side, data was transferred to the researchers and charted for analysis.

Results

GPS Locational Work

The first phase (GPS component) of the Hegenberger Corridor microstudy was completed in late Summer 2007. Community members identified and located air quality hazards, areas of community concern, and sensitive receptors and spatial data sets were created to make comparison maps. Community members identified and located a

---

[60] The five monitored sites in the Hegenberger area included: : 1) Rakha Towing and Auto Repair, 945 - 77th Ave between Hawley St and Spencer St; 2) Tassafaronga Recreation Center, 975 – 85th Ave between E and G Streets; 3) American Brass and Iron Foundry (AB&I), 7825 San Leandro St, 81st Ave and San Leandro; 4) ACORN Woodland Elementary School (near Allen Temple Baptist Church), 1025 81st Ave at Rudsdale Ave.; 5) Grass Valley Elementary, 4720 Dunkirk Ave, Oakland CA 94605 between Golf Links Rd and Glen Artney St (as a location farther away from major emission sources).

significantly greater number of air quality hazards and sensitive receptors (Figure 6.2), compared to those identified from state government agency databases (Figure 6.3).

Comparison of the geographic distribution hazards and sensitive receptors suggests that state government agency databases do not capture the actual number of both hazards and sensitive receptors in the Hegenberger study area. Important differences found in terms of sensitive receptors include microstudy locations of 13 churches (a community facility where the elderly and can spend much of their day according to Hegenberger residents), 8 licensed childcare centers as compared to 5 reported by State of California and Dunn and Bradstreet, 3 schools (vs. 5 reported by the California Department of Education), 1 senior housing facility, and 5 youth-support facilities; participants also confirmed that 2 of the 5 schools reported in the California Department of Education for this area are not active. Among hazards, the community identified 5 locations where illegal diesel vehicle idling was taking place and 39 automobile-related businesses (vs. 9 reported in the AB2588 database – see Figures 6.4 - 6.6).

This comparison also suggests that using only data from state government agency databases may not adequately highlight incompatible land uses where sensitive receptors may in fact be in close proximity to air emission sources. Figure 6.7 shows only these air quality hazards surrounded by 1,000 foot buffers, and the locations of all sensitive receptors identified during microstudy mapping. Note that 12 sensitive receptors are located within the receptor buffer where CARB recommends no such new facilities should be sited. This pattern is particularly apparent along International Boulevard, and includes four childcare centers, four churches (some that offer childcare, as well), one healthcare clinic, and two youth facilities. If the additional community-defined air quality hazards identified during microstudy mapping were included in this analysis, the problem of such receptors located within the buffer distance recommended by ARB as an exclusion zone relative to air pollution sources would likely appear to be more acute and widespread within the Hegenberger study area.

Verification of locational accuracy of CHAPIS and AB2588 facilities showed that the 2 CHAPIS facilities within the Hegenberger study area boundaries are accurately located, but all 16 of the AB2588 facilities reported inside the boundaries are located inaccurately (Figure 6.8). The offset distance ranges among these sites from about 150 ft over 900 ft, except for Waste Management, Inc. which is a reported location inside the study area, but is actually located in San Leandro, over 3 miles to the southeast. These distances are very significant in comparison to the CARB recommended buffer distances recommended for separation of air toxics hazards from sensitive receptors.

Pilot PM$_{2.5}$ Monitoring Results

Figures 6.9-6.10 present results of dust-track monitoring at several sites for all monitoring times as well as the subset of results for weekday monitoring, and they compare them to California's ambient air quality standard and the county average. For most of the sites assessed, median PM levels exceeded the California ambient PM$_{2.5}$

standard and the Alameda county average. Monitoring at Grass Valley was sparse, due to logistical challenges faced by that particular team of community members, making these results less stable than results for other sites where monitoring was more consistent during the weekday. Based on community members' observations, it is possible that the PM levels found in the other four sites are greatly impacted by the large volume of truck traffic at these sites, due to daily deliveries, idling, and major ongoing construction activities (near the Tassafaronga Recreation Center). Further monitoring work should be conducted to verify the robustness of these pilot results.   Additional, longer-term monitoring activities could also entail concurrent indoor and outdoor monitoring to assess whether/how outdoor emission sources may be affecting indoor air quality at the sensitive receptor facilities.

Discussion

The Hegenberger microstudy results suggest that although state databases are critical for understanding the potential impacts of both area and point emission sources on local communities, they may not fully capture the entire range of emissions sources and their potential localized impacts on air quality and community health, and errors in the reported location of some hazards can significantly misrepresent exposure relationships with sensitive receptors. Ground truthing efforts to verify the extent to which official emission inventories adequately capture the location and number of emission sources and sensitive receptors in EJ communities can be a way for ARB to periodically evaluate the extent to which data sources are adequately capturing the potential cumulative impacts of multiple emission sources in EJ communities. Information gathered from these community-engaged ground truthing projects can assist CARB staff in better understanding what facilities may be of significant community and public health concern and whether or not they may be falling below the "regulatory radar screen." For EJ communities, the engagement of researchers and ARB staff with community members to solicit their knowledge of local environmental hazards and sensitive receptor concerns can enhance positive community engagement and trust in the regulatory process and potentially highlight opportunities to better reduce emissions and protect community environmental health. Similar efforts by CARB staff to work with other EJ communities in Wilmington and Barrio Logan (Cal-EPA 2003; Cal-EPA 2004) have been successful in highlighting sources of concern and enhancing community understanding of the regulatory process. This Hegenberger microstudy could be leveraged in ways that enable ARB staff deepen their work with a motivated, trained, and active East Oakland community to address local air quality concerns.

The Hegenberger microstudy also highlighted a number of challenges that can complicate such community- based research projects, as well as lessons for future efforts. The field computer and software technology did not work as seamlessly and reliably as we had planned. The HP/Compaq iPaq handheld PCs, ArcPad GIS software and GPS cards were not adequately user friendly for some community leader field work volunteers, and sometimes they did not work correctly; this decreased enthusiasm for using this GPS technology during the field data collection portion of the study. Some abandoned the

devices, instead relying on recording field information manually on the maps and special forms that we provided for this purpose. The age demographic of community members (most were older) may have made adoption and use of the computers and GPS more difficult. However, the field study also provided some significant benefits. This activity improved community awareness about hazards and sensitive receptors, and the field data collection and map analysis makes the concept of cumulative impact more "real" to community members. This project added momentum toward CBE's organizing goals in the Hegenberger area, as well as CARB's presence (Álvaro Alvarado attended all of the workshops). CARB's participation at community workshops and their willingness to freely discuss community concerns gave community members confidence that their work would be useful and used.
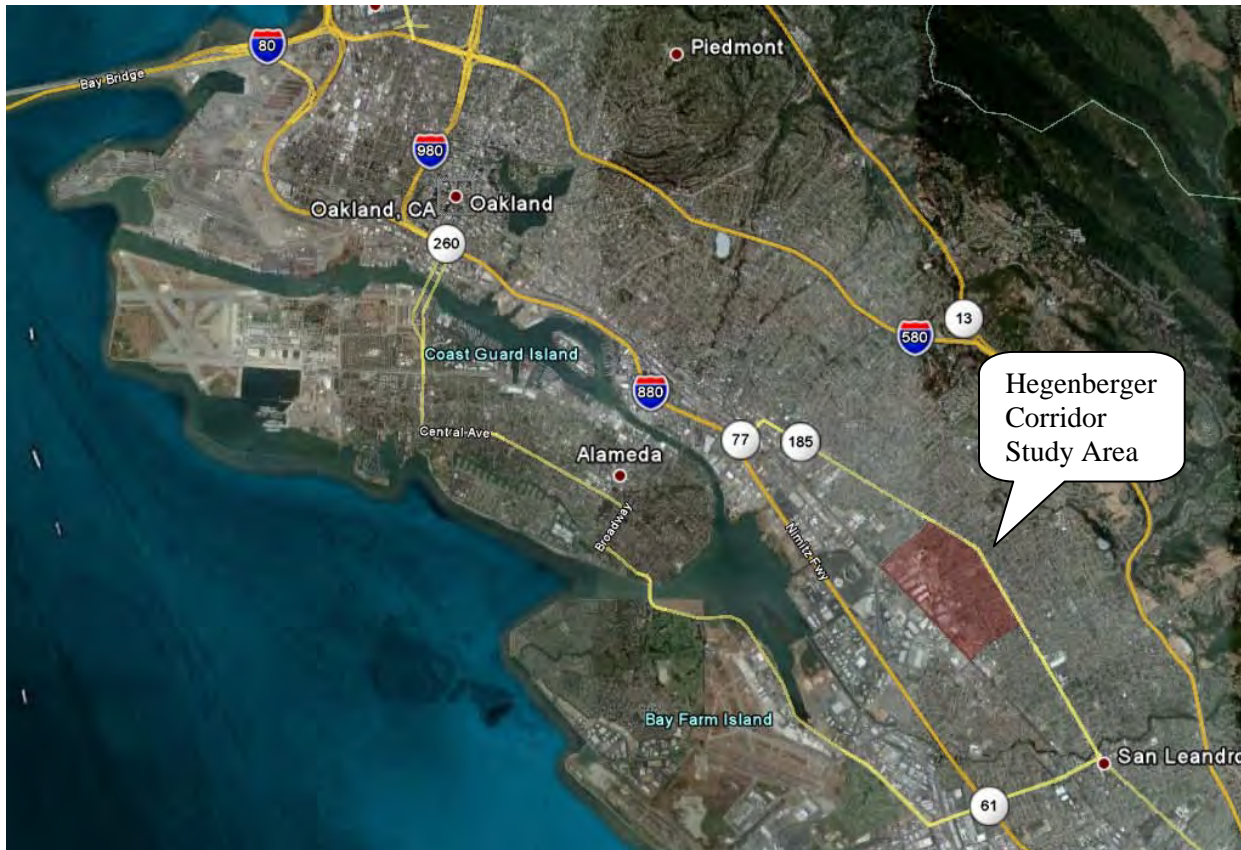
Conclusion

The Hegenberger Corridor microstudy succeeded in enabling community members to record their own observations and define areas of community concern related to air quality and EJ. This included identification and location of environmental hazards and sensitive land uses using previously agreed upon definitions consistent with ARB definitions. This exercise allowed us to augment our understanding of localized approaches for assessing cumulative impact and collaborate with community members to ground truth information from existing ARB emissions inventories regarding pollution sources of concern. Working with CBE and Dr. Alvarado, we educated community members on the concept and science of cumulative impact, and we trained them in skills needed to collect information on local hazards and sensitive receptors. We created spatial data sets and of this information developed a series of maps data sets both to illustrate the community observations and compare these observations with equivalent data from state government regulatory agencies. CBE documented and summarized the results of the microstudy in a downloadable document;[61] CBE staff were the primary authors, and the research team contributed the maps, scientific review, and assisted with the report design. PI Rachel Morello-Frosch and Dr. Álvaro Alvarado also participated in the public release of this document, with the presence of the latter giving community members the sense that their air quality concerns and field work efforts were being respected and taken seriously. We believe that this form of collaboration between researchers, EJ organizers, community leaders, and ARB staff can be a model for future efforts to productively engage EJ communities in regulatory activities.

---

[61] Available at http://college.usc.edu/geography/ESPE/documents/cumulative_impacts_final.pdf
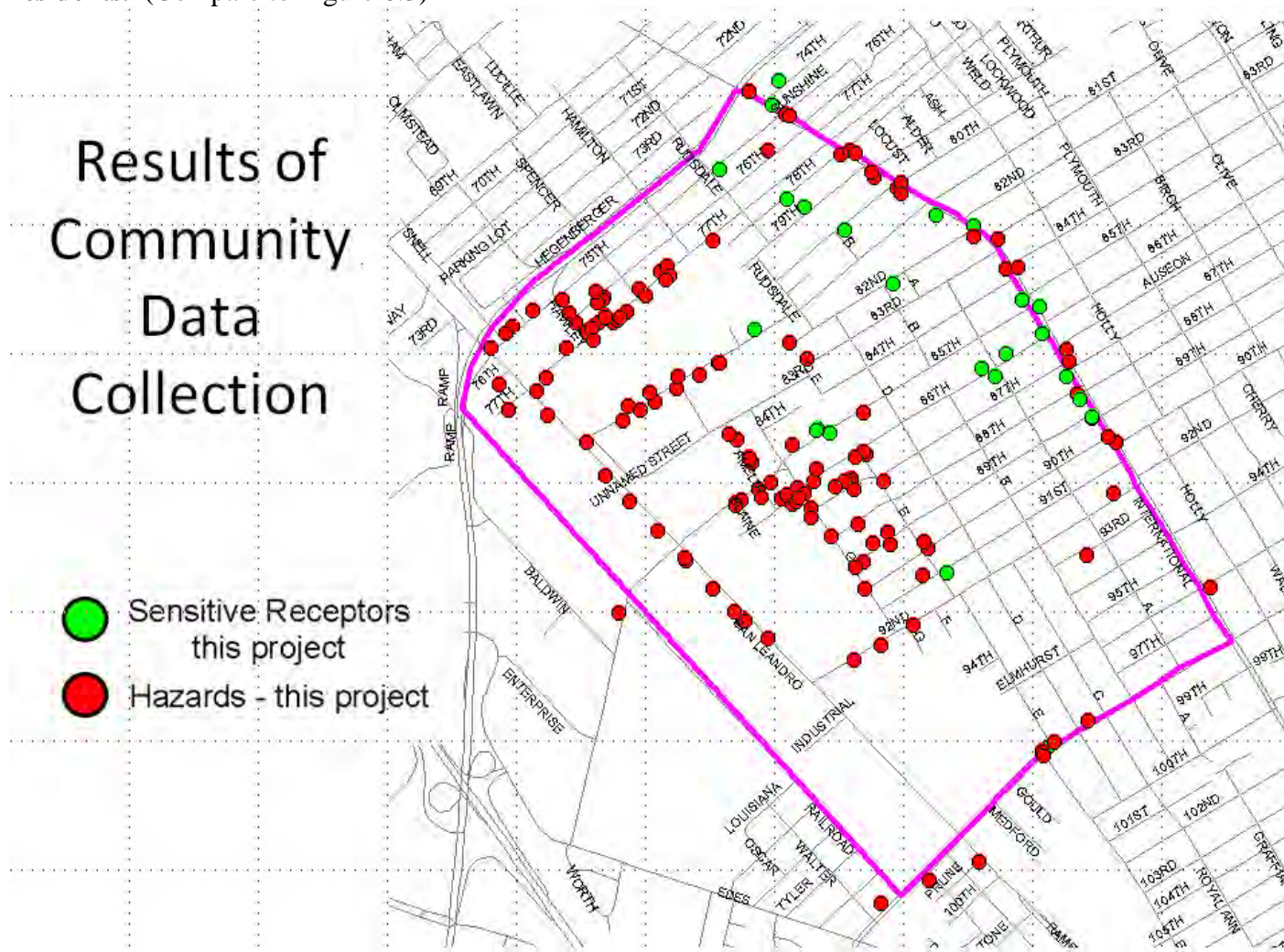
Table 6.1: Community-generated lists of Hazards and Sensitive Receptors.	
Hazards: Polluting Sources	Sensitive Receptors: Populations Affected
- Large Parking Lots - Recycling Facilities - Grocery Stores (or other magnets for Diesel Trucks) - All Sites with Idling Trucks - Gas Stations - Auto Body Shops - Brownfields - Automobile -Repair and -Related Facilities - Dry cleaning facilities	- Children's Centers/Daycare - Senior Centers - Retirement Homes - Playgrounds and Outdoor Recreation Areas - Churches - Healthcare Clinics and Other Facilities - Schools, Head Starts, and Academies

Figure 6.1: Location of the Hegenberger Corridor study area in east Oakland, bounded by Hegenberger Expressway, International Blvd., San Leandro St. and 92nd Avenue (two views)

Hegenberger
Corridor
Study Area

Figure 6.2: Results of Hegenberger Corridor microstudy mapping of air quality hazards and sensitive receptors by community residents. (Compare to Figure 6.3)

Figure 6.3: Map of Hegenberger Corridor study area showing locations of air quality hazards and sensitive land uses using data from State regulatory database facilities.

Figure 6.4: Comparison of spatial data from Figures 6.2 and 6.3 to emphasize additional air quality hazards and sensitive receptors recorded by Hegenberger community members.

Figure 6.5: Location of businesses where auto paint and body work was observed during Hegenberger community field observations, as compared to automobile businesses recorded in the AB2588 "Hot Spots" database provided by CARB.

Figure 6.6: Locations of observations of illegal idling by diesel trucks during Hegenberger community field observations.

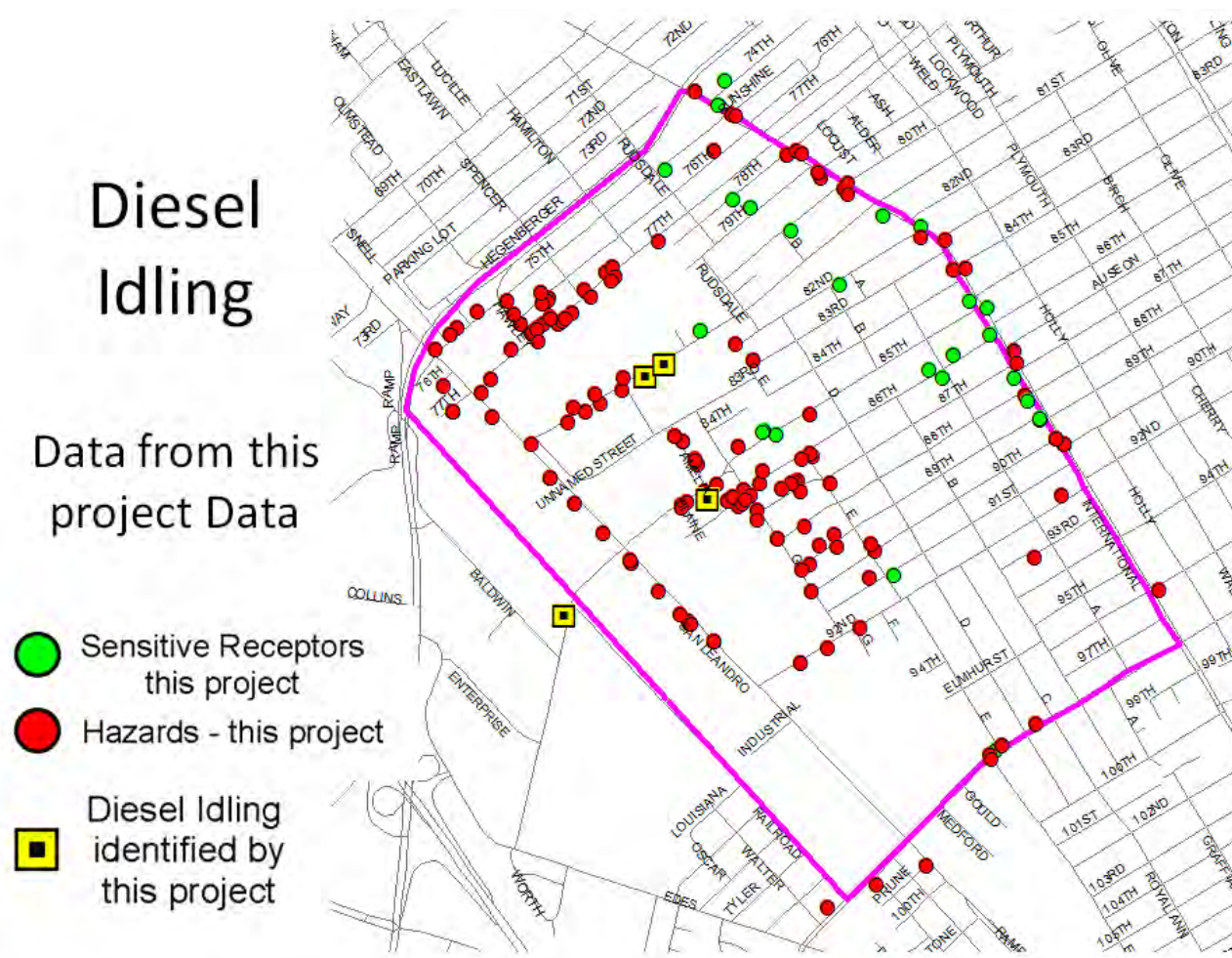**Figure 6.7: Buffers surrounding air quality hazards from state government agency databases capture twelve sensitive receptors mapped by community residents during the Hegenberger microstudy.**

United Outreach Church	1200 75th Ave
Kingdom Hall Jehovah Witness	1057 98th Ave
Allen Temple Baptist Church	8501 International Blvd
Riley Chapel CME Church	1302 80th Ave
Kiddieland Childcare Center	1268 78th Ave
East Oakland Health Care Center	7450 International Blvd
1234 Tots Family Day Care	1234 82nd Ave
Show Me Love Daycare	1279 79th Ave
Barbara Jackson Daycare	1449 74th Ave
Allen Temple Arms	8135 International Blvd
East Oakland Youth Development Center	8200 International Blvd
Ossian Carr Boys and Girls Club	8530 International Blvd

Figure 6.8: Reported locations of AB2588 "hot spots," air toxics emissions sources (red squares), compared with actual locations verified by ground truthing (green circles).

Waste Management Inc.
2615 Davis Street
San Leandro, CA 94577

Figure 6.9: PM$_{2.5}$ measurements from 5 sites, CBE East Oakland Phase II, October 7-November 8, 2008

Includes weekend and weekday monitoring results. Dark colors represent median levels.

Figure 6.10: PM$_{2.5}$ *weekday* measurements from 5 sites, CBE East Oakland Phase II, October 7-November 8, 2008

Chapter 7. Summary and Conclusions

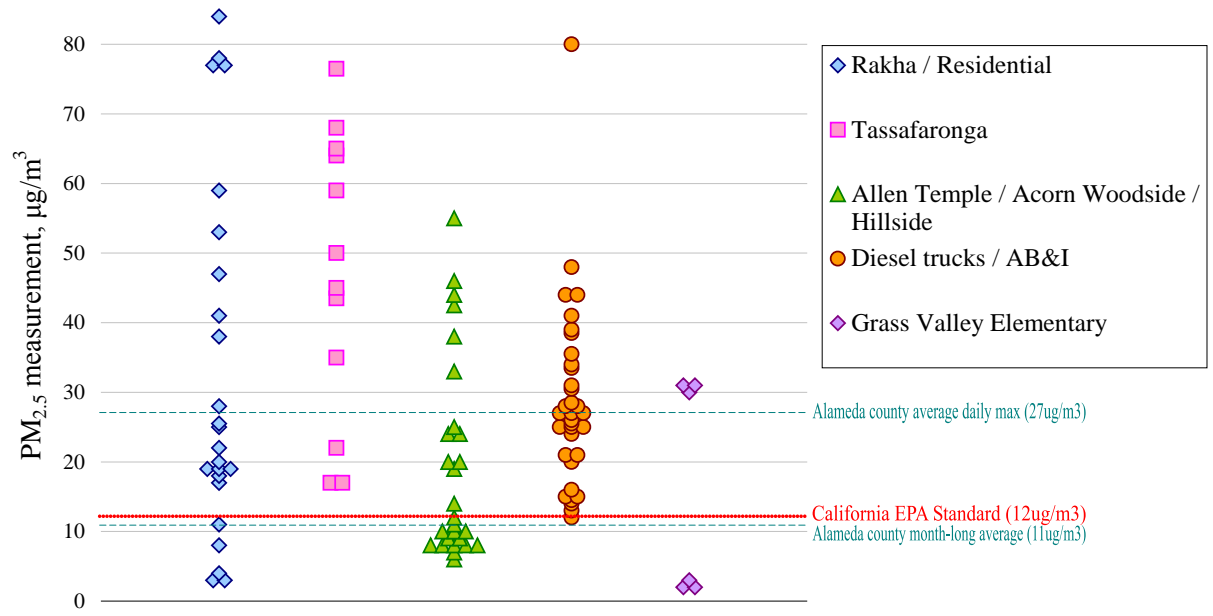This project sought to develop innovative methodological approaches to analyzing environmental justice (EJ) concerns of relevance to air pollution regulation in California. This research initiative included five inter-related components:

1) An EJ analysis in the San Francisco Bay Area that included statistical modeling with controls for spatial autocorrelation as well as the development and assessment of new independent metrics of area-level social vulnerability.

2) A statewide epidemiological analysis to assess the relationship between ambient criteria pollution and adverse birth outcomes (pre-term birth and low birth weight), with attention to assessing the robustness of effect estimates using different exposure assessment techniques and co-pollutant models as well as examining potential effect modification by area- and individual-level socio-economic status variables;

3) Using results from components 1 and 2 to inform the development of an Environmental Justice Screening Method (EJSM) to assess geographic disparities in cumulative impacts of environmental and social stressors within regions and air basins;

4) The application of the EJSM to a power plant siting scenario and to compare these results to current EJ assessment practices used at the California Energy Commission (CEC);

5) The implementation of a community-based participatory ground truthing project in the Hegenberger Corridor of East Oakland to examine how well agency data inventories capture the existence and locations of local emissions sources and sensitive receptors identified by community residents.

Results from the Bay Area EJ analysis were consistent with previous EJ studies in the state indicating that communities of color and low income residents are more likely to host emission facilities and to have higher pollution burdens from mobile and stationary sources, even after controlling for other demographic indicators in multivariate modeling. Although much work has examined EJ issues related to air quality in Southern California, this was the first comprehensive EJ assessment of the San Francisco Bay Area. Results also highlighted the utility of metrics of community civic engagement capacity (i.e. linguistic isolation) when assessing EJ concerns in diverse regions in the state. Results from this study encouraged us to incorporate indicators of neighborhood civic engagement capacity as part of the social and health vulnerability category in the EJSM. The birth outcomes study developed new exposure assessment approaches for assessing the consistency of effects of air pollution on birth weight and preterm birth. Results suggest that birth outcomes are a useful health outcome variable to include in cumulative impact and EJ screening, since this is a sensitive endpoint that may reflect underlying neighborhood health.   Moreover, although effect estimates in the study were modest, the ubiquity of air pollution exposures and the potential responsiveness of pollutant levels to

planning and regulation efforts suggest that the potential implications may be important for infant health and development.

EJSM development was informed by the Bay Area study and the birth outcomes analysis, and it met the goal of assessing disparities in cumulative impacts by integrating metrics of environmental and social stressors in a transparent, yet scientifically valid way that could inform regulatory decision-making and enhance community engagement in local air quality issues of concern. The method integrated three categories of cumulative impact—hazard proximity, exposure and health risk, and social and health vulnerability – and it emphasized screening in residential and land use areas. The benefit of this method is that it is flexible enough to be able to integrate new data sources as necessary and it can be applied to diverse policy and regulatory contexts that require assessments of EJ and cumulative impacts. Moreover, the method has undergone extensive scientific and community review and is already being adapted by CARB staff to inform decision-making in the implementation of the AB 32 Scoping Plan. The method was also tested in a power plant siting case scenario developed by the CEC. Results from the EJSM revealed that the proposed power plant site would have exacerbated environmental disparities in plant siting and highlighted possible ways in which the CEC could strengthen its analytical approach to conducting EJ assessments of future proposed power plant sites.

Finally, the Hegenberger Corridor ground truthing study built upon previous efforts by CARB to engage communities in local efforts to better understand local air quality concerns. This community-based participatory research project compared air pollution emission and land use information as derived from CARB inventory databases with actual on-the-ground observations and documentation by community members. Results suggest that although state databases are critical for understanding the potential impacts of both area and point emission sources on local communities, they may not fully capture the entire range of emissions sources and their potential localized impacts on air quality and community health, and errors in the reported location of some hazards can significantly misrepresent exposure relationships with sensitive receptors. Ground truthing efforts to verify the extent to which official emission inventories adequately capture the location and number of emission sources and sensitive receptors in EJ communities can be a way for ARB to periodically evaluate the extent to which data sources are adequately capturing the potential cumulative impacts of multiple emission sources in EJ communities. Information gathered from such these community-engaged ground truthing projects can assist CARB staff in better understanding what facilities may be of significant community and public health concern and whether or not they may be falling below the "regulatory radar screen." For EJ communities, the engagement of researchers and ARB staff with community members to solicit their knowledge of local environmental hazards and sensitive receptor concerns can enhance positive community engagement and trust in the regulatory process and potentially highlight opportunities to reduce emissions and protect community environmental health. Moreover, the participatory methods and training strategies that we developed for the Hegenberger study can be replicated elsewhere and used specifically to validate results from future efforts to conduct EJ screening for cumulative impacts in other regions.

Chapter 8. Recommendations

Recommendations from this project are for both future research and for policy.

- Further work should address methodological issues of spatial autocorrelation in EJ studies. While the most important result from our EJ analysis of the Bay Area is that the pattern of racial disparity persists in multivariate modeling, it is also the case that effect estimates are attenuated after controlling for spatial error. Future research should refine methods for addressing these spatial correlation issues in EJ studies and should be careful to not overstate the case by failing to control for spatial error.

- The results on linguistic isolation suggest that this metric of civic engagement capacity may be important for regulatory strategies related to EJ. Further research should explore the connections between neighborhood ethnic make-up and linguistic capacity and how these two variables may shape observed patterns of environmental inequity.

- The results on birth outcomes are robust and suggest the need for more systemic regulatory consideration of air pollution impacts on this sensitive health outcome. Future research should examine potential birth outcome impacts from other categories of pollutants, such as air toxics. Moreover, future birth outcome studies should take into account the cumulative impact of exposures to multiple air pollutants, which may be important if in fact chemical mixtures lead to higher health risks than individual chemical constituents. A major source of both gaseous and particulate air pollutants is combustion, and one important area of future inquiry is to take a source-based approach to assessing health effects rather than isolating the impacts of individual pollutants. More can be done to analyze and develop source-specific measures, such as traffic density exposure metrics or metrics that categorize pollutants based on toxicological outcomes (e.g. reproductive or developmental toxicants).

- Future work on screening methods should explore alternative weighting schemes for different metrics and explore the sensitivity of scores to such weights.

- Future work on EJSM development should involve developing the land use data (or land use proxies where high-quality land data is unavailable) necessary to implement the EJSM on a statewide level. This would also facilitate both inter- as well as intra- regional assessments of geographic inequities in the cumulative impacts of environmental and social stressors.

- We also think that traffic metrics, a portion of the screening data that was cut from the original proposal due to costs, would be important to build into the screen (given both the disparity in mobile emissions illustrated in the Bay Area

analysis and the fact that mobile emissions are an important source of criteria air pollutants that are associated with poor birth outcomes).

- Future research work may want to consider adopting elements of the participatory model utilized here. Community participation clearly made a difference for the microstudy, including direct interaction with CARB staff in all community workshops, as well as for informing the Bay Area analysis. But the greatest benefit of community engagement has come in the generally positive reaction to the screening method. When research is likely to inform policy, particularly in controversial areas, early community engagement can enhance understanding and trust in the research and regulatory process.

- In this regard, we note that there are few places in the formal report structure that we have followed to discuss the degree of community engagement and what difference it made to the results. One specific recommendation to ARB might be to consider creating such a section as an optional part of the report format for research projects as appropriate to the research project in question. We understand that not all projects will either need or will incorporate community-based participatory research (CBPR), but having a formal location to report on that aspect of the work might encourage it.

- The often-heard community concern about concentrations of air quality hazards operating under the regulatory radar is consistent with what was observed in the Hegenberger microstudy. The microstudy demonstrated that communities have important local knowledge about cumulative impacts that can contribute to improving regulatory decision-making and enforcement activities. CARB should expand its recent successes in Hegenberger and other communities it has previously engaged and consider ways to systematically access local knowledge when conducting cumulative impact assessments.

- Both the microstudy and the EJSM showed that some of the spatial databases CARB uses are either out of date or contain locational error, presumably reflecting geocoding errors, for some records. Such problems are, of course, not atypical for the types of data that would be useful in this type of analysis and assuming that such errors are randomly distributed, the impacts on broad statistical testing at a regional level is minimal. On the other hand, improved locational accuracy would reduce misclassification and improve scoring for the method we describe. We recommend improving procedures for verifying and error‑checking locational accuracy in data that are used in screening and assessment, including ISO 19115 metadata documentation with the datasets, and making them publically available so that researchers and communities can both use them and help CARB identify potential corrections.

- The EJSM offers many clear advantages over the demographic screen technique used by the CEC to identify areas of EJ concern, by providing greater geographic resolution and detail, a larger and more comprehensive set of indicators, and

superior methods of scoring for comparison among multiple sites. The EJSM is also equally easy to use given the datasets and tools provided under this contract, and it is superior to the U.S. EPA screening tool, EJSEAT. CEC would further benefit from adopting the screening method in evaluating power plant siting in urban areas and modifying it using alternative indicators to evaluate projects in rural areas, such as large alternative energy installations sensitive to other sorts of environmental harm, including wildlife disruption and endangered vegetation.

- The linguistic isolation results suggest the need for continued and improved outreach to language minority populations.

- Any alterations to the EJSM, particularly with regard to reweighting data layers or imposing more complicated ranking systems (such as standard deviation breaks, z-scores, or threshold levels) should be weighed against the complexity of implementation and transparency to an interested public.

- Additional data layers for the EJSM should be considered by consulting with scientific colleagues working on environmental and public health tracking initiatives within California and nationally. This form of cross-agency collaboration on advancing EJ screening development could enhance efforts to address current regulatory challenges in the realms of climate change adaptation and mitigation.

Glossary of Terms and Abbreviations

Term	Explanation
AQMD	Air Quality Management District
AQS	Air Quality System
ARB	Air Resources Board
CalAIRS	California Aerometric Information Reporting System
CARB	California Air Resources Board
CaSIL	California Spatial Information Library (www.atlas.ca.gov)
CBE	Communities for a Better Environment
CBPR	Community-based participatory research
CEC	California Energy Commission
CEIDARS	California Emission Inventory Development and Reporting System
CEJAC	California Environmental Justice Advisory Committee
CEQA	California Environmental Quality Act
CHAPIS	Community Health Air Pollution Information System
CI	Cumulative Impact
CIPoly	Cumulative Impact Polygon
CO	Carbon Monoxide
CO_2	Carbon Dioxide
DTSC	Department of Toxic Substances Control
EJ	Environmental Justice
EJSEAT	Environmental Justice Strategic Enforcement Assessment Tool
EJSM	Environmental Justice Screening Method
EPCRA	Emergency Planning and Community Right-to-Know
GIS	Geographic Information Systems
HAPEM	Hazardous Air Pollutant Exposure Model
IUR	Inhalation Unit Risk
LBW	Low birth weight
LMP	Last menstrual period
MATES III	Multiple Air Toxics Exposure Study
NATA	National Air Toxics Assessment
NO_2	Nitrogen Dioxide
NTAD	National Transportation Data Atlas from the Bureau of Transportation statistics (www.bts.gov)
O_3	Ozone
OEHHA	Office of Environmental Health Hazard Assessment
OLS	Ordinary least square
PM	Particulate matter
PM_{10}	Particulate matter with particles of 10 micrometers or less
$PM_{2.5}$	Particulate matter with particles of 2.5 micrometers or less
RSEI	Risk Screening Environmental Indicators
RSEI-GM	Risk Screening Environmental Indicators Geographic Microdata

SAR	Spatial autoregressive
SARA	Superfund Amendments and Reauthorization Act
SAS	Statistical Analysis Software
SCAG	Southern California Association of Governments
SES	Socio-economic status
SF1	Summary File 1
SF3	Summary File 3
SGA	Small for gestational age
SIC	Standard Industrial Classification
SO_2	Sulfur Dioxide
SPSS	Statistical Package for the Social Sciences
TRI	Toxic Release Inventory
U.S. EPA	United States Environmental Protection Agency
USGS	United States Geological Survey
ZCTA	ZIP Code Tabulation Area

Appendices

Appendix 1. A simple GIS tool

This is a UIToolControl written for ESRI ArcMap that is installed as a button on the ArcMap toolbar and can be used to obtain the CI score from the EJSM for any facility site. To use the tool, the user displays the spatial data (either census blocks or CI Polygons with appropriate EJSM attributes, selects the tool from the toolbar button, uses the mouse select by clicking any location in the data view to select a location around which to construct the buffer (click x,y is center of buffer), specifies buffer size from a dialog box (buffer distance will be set to map units of the queried layer), and presents the results in the spatial data layer attribute table.

All necessary files are posted in a folder called CEC_Tool which is available at www.onlinefilefolder.com . Users need to login with username *ej_screen*; the password has been shared with CARB staff and can be made available as needed.

Installation of the tool is as follows:

1. Open a new ArcMap document.
2. Go to Tools-->Macros-->Visual Basic Editor.
3. Right click on Normal.mxt in the left pane and select Import option.
4. Select ThisDocument.cls file from the files provided, and click open to import the file.
5. Copy and paste the code from ThisDocument.cls to the 'ThisDocument' under 'ArcMap Objects' folder in the left pane.
6. Right click the Normal.mxt in the left pane and click Import file option.
7. Select 'frmBufferTool.frm' and click open to import the file.
8. Now switch to ArcMap window. Go to Tools--> Customize option.
9. Click 'Commands' tab and select 'UIControls'.
10. Click 'New UIControls'. And select 'UIToolControl'. Click 'Create' button.
11. Drag and drop UIToolcontrol to the toolbar. (Figure 5.9)

To use the tool:
1. Click the tool from the toolbar button to display the pop-up form (Figure 5.9)
2. Enter the buffer distance (using map units for distance) and click 'Execute' button.
3. Move the cursor to any location on the data view, press Shift key and left click to define the buffer. The tool will select all features in the selectable layers (set in AM>Selection>Set Selectable Layers
4. Open data table to view results….(check with Justin on method of calculation of average population-weighted CI scores, etc. (Figure 5.10)
5. Double-click to end tool

When using the tool, after you press Shift-click to define the buffered point, the buffer that was created from the previous event is deleted, but this also deletes any graphic elements in the data view, including any text. This line can be commented out in the code if the user prefers.

Figure A.1: View of ArcMap toolbar with tool button and pop-up form with buffer distance specified. In this example, the buffer distance is specified as 6 9659 meters (or 6 miles) as the tool uses the distance units of the Data Frame, which is meters in the datasets provided as part of this project (spatial data layers are in the Teale Albers California projection, NAD83)
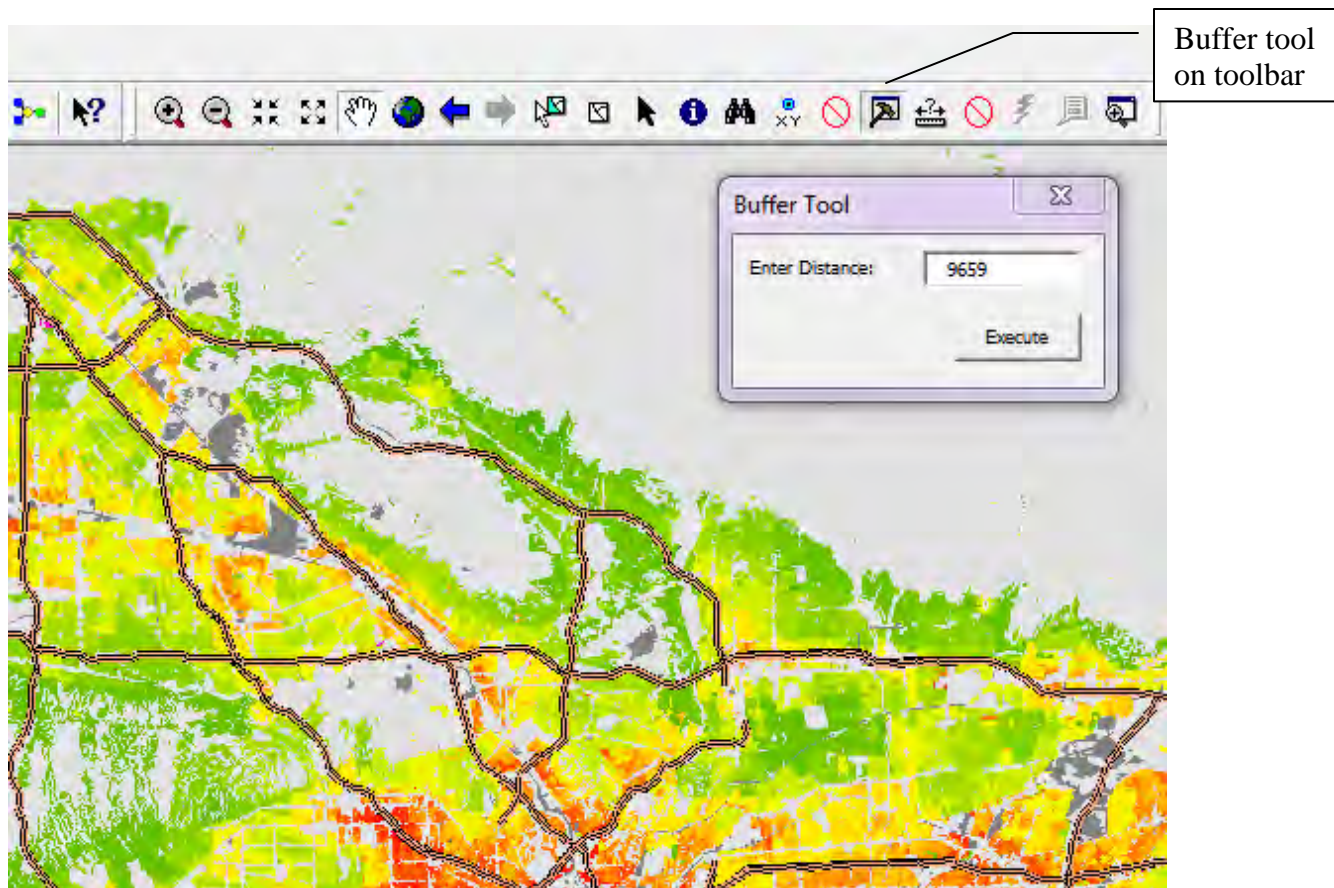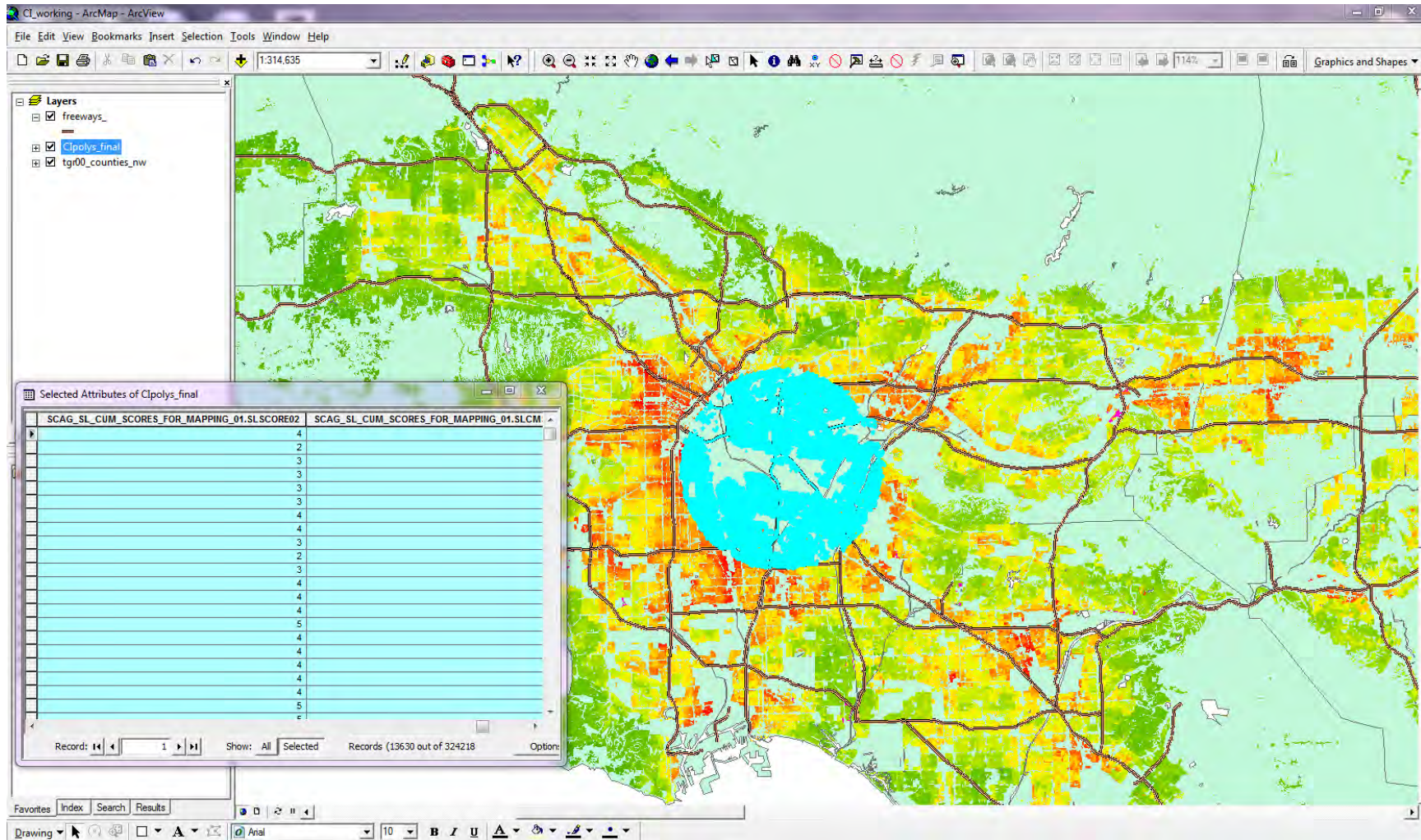
Figure A.2: ArcMap data view showing the resulting buffer and selected features (13,630 of 324,218 CI polygons) relative to a point selected near the proposed Nueva Azalea site. Any point may be selected for buffering using this tool.

References

Adler, N. E. and D. H. Rehkopf (2008). "U.S. disparities in health: descriptions, causes, and mechanisms." Annu Rev Public Health **29**: 235-252.

Alexander, G., J. Himes, et al. (1996). "A United States national reference for fetal growth." Obstetrics & Gynecology **87**(2): 163-168.

Anderton, D. L., A. B. Anderson, et al. (1994). "Environmental Equity: The Demographics of Dumping." Demography **31**(2): 229-248.

Anderton, D. L., A. B. Anderson, et al. (1994). "Hazardous Waste Facilities: "Environmental Equity" Issues in Metropolitan Areas." Eval Rev **18**(2): 123-140.

Anselin, L. (2005). Exploring spatial data with GeoDa™: A workbook. Urbana, Spatial Analysis Laboratory and Center for Spatially Integrated Social Science, Department of Agriculture and Consumer Economics, University of Illinois, Urbana-Champaign.

Anselin, L. and D. A. Griffith (1988). "Do spacial effects really matter in regression analysis?" Papers in Regional Science **65**(1): 11-34.

Ash, M. and T. R. Fetter (2004). "Who lives on the wrong side of the environmental tracks? Evidence from the EPA's Risk-Screening Environmental Indicators Model." Social Science Quarterly **85**(2).

Baden, B. M. and D. L. Coursey (2002). "The locality of waste sites within the city of Chicago: a demographic, social, and economic analysis." Resource and Energy Economics **24**(1-2): 53-93.

Barker, D. (1995). "Fetal origins of coronary heart disease." British Medical Journal **311**(6998): 171-174.

Barker, D., P. Winter, et al. (1989). "Weight in infancy and death from ischaemic heart disease." Lancet **2**: 577-580.

Basu, R., J. Parker, et al. (2003). Does Maternal Smoking Confound the Relationship Between PM2.5 and Birthweight? [abstract]. . American Public Health Association. San Francisco, CA.

Basu, R., T. J. Woodruff, et al. (2004). "Comparing exposure metrics in the relationship between PM2.5 and birth weight in California." J Expo Anal Environ Epidemiol **14**(5): 391-396.

Been, V. (1993). "What's fairness got to do with it? Environmental justice and the sting of locally undesirable land uses." Cornell Law Review **78**: 1001-1085.

Been, V. (1994). "Locally Undesirable Land Uses in Minority Neighborhoods: Disproportionate Siting or Market Dynamics?" The Yale Law Journal **103**(6): 1383-1422.

Been, V. (1995). "Analyzing evidence of environmental justice." Journal of Land Use and Environmental Law **11**: 1-37.

Been, V. and F. Gupta (1997). "Coming to the Nuisance or Going to the Barrios - A Longitudinal Analysis of Environmental Justice Claims." Ecology Law Quarterly **24**(1).

Bell, M. L., K. Ebisu, et al. (2007). "Ambient air pollution and low birth weight in Connecticut and Massachusetts." Environmental Health Perspectives **115**(7): 1118-1124.

Berkowitz, G., C. Blackmore-Prince, et al. (1998). "Risk factors for preterm birth subtypes." Epidemiology **9**: 279–285.

Bobak, M. (2000). "Outdoor air pollution, low birth weight, and prematurity." Environmental Health Perspectives **108**(2): 173-176.

Boer, T. J., M. Pastor, et al. (1997). "Is there environmental racism? The demographics of hazardous waste in Los Angeles County." Social Science Quarterly **78**(4): 793-810.

Boerner, C. and T. Lambert (1995). Environmental justice in the city of St. Louis: The economics of siting industrial and waste facilities. St. Louis, MO, Center for the Study of American Business.

Bowen, W. M. (2001). Environmental Justice Through Research-Based Decision-Making. New York, NY, Garland Publishing.

Bowen, W. M., M. J. Salling, et al. (1995). "Toward Environmental Justice: Spatial Equity in Ohio and Cleveland." Annals of the Association of American Geographers **85**(4): 641-663.

Brauer, M., C. Lencar, et al. (2008). "A Cohort Study of Traffic-Related Air Pollution Impacts on Birth Outcomes." Environmental Health Perspectives **116**(5): 680-686.

Brody, J., R. Morello-Frosch, et al. (Forthcoming). "Linking Exposure Assessment Science with Policy Objectives for Environmental Justice and Breast Cancer Advocacy: The Northern California Household Exposure study." American Journal of Public Health.

Brunner, E. J. (2000). Toward a new social biology. Social Epidemiology. Lisa.F.Berkman and Ichiro.Kawachi. New York, Oxford University Press**:** 306-331.

Buka, S., R. T. Brennan, et al. (2002). "Neighborhood support and the birthweight of urban infants." American Journal of Epidemiology **157**: 1-8.

Burke, L. (1993). "Race and Environmental Equity: A Geographic Analysis in Los Angeles." Geo-Info Systems **October**: 44-50.

Cal-EPA (2003). Community Air Quality Monitoring: Special Studies- Wilmington (Available at: http://www.arb.ca.gov/ch/reports/

wilmington_sb25_report.pdf) Accessed March 3, 2006. . Sacramento, CA, California Environmental Protection Agency, Air Resources Board.

Cal-EPA (2004). Barrio Logan Report: A Compilation of Air Quality Studies in Barrio Logan (Available at: http://www.arb.ca.gov/ch/

programs/bl_11_04.pdf). Accessed March 3, 2006. Sacramento, CA, California Environmental Protection Agency, Air Resources Board.

Cal-EPA Advisory Committee on Environmental Justice (2003). Recommendations of the California Environmental Protection Agency (Cal/EPA)Advisory Committee on Environmental Justice to the Cal/EPA Interagency Working Group on Environmental Justice, Final Report., California Environmental Protection Agency.

CARB. (2005). "Air Quality and Land Use Handbook: A Community Health Perspective. Available at: http://www.arb.ca.gov/ch/handbook.pdf. Accessed June 5, 2009.".

CARB. (2007). "California Air Quality Data Available on DVD-ROM. http://www.arb.ca.gov/aqd/aqdcd/aqdcddld.htm. Accessed 9 July."

Chung, A., D. Chang, et al. (2001). "Comparison of Real-Time Instruments Used to Monitor Airborne Particulate Matter." Journal of the Air and Waste Management Association **51**: 109–120.

Clougherty, J. and L. Kubzansky (In Press). "A Framework for Examining Social Stress and Susceptibility in Air Pollution and Respiratory Health." Environmental Health Perspectives **117**(9): 1351–1358.

Clougherty, J. E., J. I. Levy, et al. (2007). "Synergistic effects of traffic-related air pollution and exposure to violence on urban asthma etiology." Environmental Health Perspectives **115**(8): 1140-1146.

Collins, J. W., A. A. Herman, et al. (1997). "Very-Low-Birthweight Infants and Income Incongruity among African American and White Parents in Chicago." American Journal of Public Health **87**(3): 414-417.

Cramer, J. C. (1995). "Racial and Ethnic Differences in Birthweight: The Role of Income and Financial Assistance." Demography **32**(2): 231-247.

da Silva, A., V. Simoes, et al. (2003). "Young maternal age and preterm birth." Paediatric and Perinatal Epidemiology **17**: 332–339.

Darrow, L., M. Klein, et al. (2009). "Ambient air pollution and preterm birth: a time-series analysis." Epidemiology **20**(5): 689-693.

Darrow, L., T. Woodruff, et al. (2006). "Maternal smoking as a confounder in studies of air pollution and infant mortality (research letter)." Epidemiology **17**: 592–593.

DeFur, P. L., G. W. Evans, et al. (2007). "Vulnerability as a function of individual and group resources in cumulative risk assessment." Environmental Health Perspectives **115**(5): 817-824.

DHHS (2010). Changing Outcomes - Achieving Health Equity: The National Plan for Action. Available at: http://minorityhealth.hhs.gov/npa/templates/browse.aspx?lvl=1&lvlID=31. Washington, DC, Office of Minority Health, Department of Health and Human Services**:** 210 Pages.

Downey, L., S. DuBois, et al. (2008). "Environmental Inequality in Metropolitan America." Organization & Environment **21**(3): 270-294.

Dugandzic, R., L. Dodds, et al. (2006). "The association between low level exposures to ambient air pollution and term low birth weight: a retrospective cohort study." Environmental Health **17**(5): 3.

Ecob, R. and G. Davey Smith (1999). "Income and Health: What is the Nature of the Relationship?" Social Science and Medicine **48**: 693-705.

Edwards, R., K. Smith, et al. (2006). "An Inexpensive Dual-Chamber Particle Monitor: Laboratory Characterization." Journal of the Air and Waste Management Association **56**: 789-799.

Fitz, D., A. Winer, et al. (2003). Characterizing the range of children's pollutant exposure during school bus commutes. Report 00-322 Sacramento, CA, California Air Resources Board.

Foreman, C. (1998). The promise and peril of environmental justice. Washington, DC, Brookings Institution.

Gee, G. C. and D. C. Payne-Sturges (2004). "Environmental health disparities: a framework integrating psychosocial and environmental concepts." Environmental Health Perspectives **112**(17): 1645-1653.

Glaser, E., M. M. Davis, et al. (1998). Cancer Incidence Among Residents of the Bayview-Hunters Point Neighborhood, San Francisco, California 1993-1995. San Francisco, CA, San Francisco Department of Public Health.

Glinianaia, S., J. Rankin, et al. (2004). "Particulate air pollution and fetal health. A systematic review of the epidemiologic evidence." Epidemiology **15**: 36-54.

Gouveia, N., S. Bremner, et al. (2004). "Association between ambient air pollution and birth weight in São Paulo, Brazil." Journal of Epidemiology and Community Health **58**(1): 11-17.

Ha, E., Y. Hong, et al. (2001). "Is air pollution a risk factor for low birth weight in Seoul?" Epidemiology **12**(6): 643-648.

Haan, M., G. Kaplan, et al. (1987). "Poverty and Health: Prospective Evidence From the Alameda County Study." American Journal of Epidemiology **125**: 989-998.

Hahn, R., E. Eaker, et al. (1996). "Poverty and Death in the United States." International Journal of Health Services **26**(4): 673-690.

Hemingway, H., A. Nicholson, et al. (1997). "The Impact of Socioeconomic Status on Health Functioning as Assessed by the SF-36 Questionnaire: The Whitehall II Study." American Journal of Public Health **87**(9): 1484-1490.

Hessol, N. A., E. Fuentes-Afflick, et al. (1998). "Risk of Low Birth Weight Infants Among Black and White Parents." Obstetrics and Gynecology **92**(5): 814-822.

Huynh, M., J. Parker, et al. (2005). "Contextual effect of income inequality on birth outcomes." International Journal of Epidemiology **34**(4): 888-895.

Huynh, M., T. J. Woodruff, et al. (2006). "Relationships between air pollution and preterm birth." Paediatric and Perinatal Epidemiology **20**: 454-461.

Israel, B. A., Schulz, Amy J. , et al. (1998). "Review of Community-Based Research: Assessing Partnership Approaches to Improve Public Health." Annual Review of Public Health **19**: 173-202.

Jerrett, M., R. T. Burnett, et al. (2005). "Spatial analysis of air pollution and mortality in Los Angeles." Epidemiology **16**(6): 727-736.

Kannan, S., D. Misra, et al. (2006). "Exposures to airborne particulate matter and adverse perinatal outcomes a biologically plausible mechanistic framework for exploring potential effect modification by nutrition." Environmental Health Perspectives **114**(11): 1636-1642.

Kawachi, I. and M. Marmot (1998). "Commentary: What Can We Learn from Studies of Occupational Class and Cardiovascular Disease?" American Journal of Epidemiology **148**(2): 160-163.

Kay, J. (1995). Hallinan Plans Roadblock to Halt Bayview Industrial Construction. San Francisco Chronicle. San Francisco, CA**:** Examiner A1.

Krieg, E. J. (1995). "A Socio-Historical Interpretation of Toxic Waste Sites: The Case of Greater Boston." American Journal of Economics and Sociology **54**(1): 1-14.

Krieger, N., J. Chen, et al. (2003). "Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The Public Health Disparities Geocoding Project (US)." Journal of Epidemiology and Community Health **57**(3): 186-199.

Krieger, N. and E. Fee (1994). "Social class: The missing link in US health data." International Journal of Health Services **24**(1): 25-44.

Krieger, N., D. Rowley, et al. (1993). "Racism, Sexism, and Social Class: Implications for Studies of Health, Disease, and Well-Being." American Journal of Preventive Medicine **9**(6): 82-122.

Lacasana, M., A. Esplugues, et al. (2005). "Exposure to ambient air pollution and prenatal and early childhood health effects." European Journal of Epidemiology **20**: 183-199.

Larson, D. (2005) "Breathing Fire, in their own words." Global Community Monitor.

LaVeist, T. A. (1992). "The Political Empowerment and Health Status of African-Americans: Mapping a New Territory." The American Journal of Sociology **97**(4): 1080-1095.

Lester, J. P., D. W. Allen, et al. (2001). Environmental Justice in the United States: Myths and Realities. Boulder, CO, Westview Press.

Lin, C., C. Li, et al. (2004). "Increased risks of term low-birth-weight infants in a petrochemical industrial city with high air pollution levels." Archives of Environmental Health **59**(12): 663-668.

Lipsett, M., S. Hurley, et al. (1997). "Air Pollution and Emergency Room Visits for Asthma in Santa Clara County, California." Environmental Health Perspectives **105**(2): 216-222.

Maisonet, M., A. Correa, et al. (2004). "A review of the literature on the effects of ambient air pollution on fetal growth." Environmental Research **95**: 106-115.

Mannes, T., B. Jalaludin, et al. (2005). "Impact of ambient air pollution on birth weight in Sydney, Australia." Occupational and Environmental Medicine **62**(8): 524-530.

Martin, J., B. Hamilton, et al. (2003). "Births: final data for 2002." National Vital Statistics Report **52**: 1–113.

Mathews, T. and M. MacDorman (2007). "Infant mortality statistics from the 2004 period linked birth/infant death data set." National Vital Statistics Reports **55**(14): -32.

McCormick, M. (1985). "The contribution of low birth weight to infant mortality and childhood morbidity." New England Journal of Medicine **312**(2): 82-90.

Minkler, M. (2000). "Participatory Action Research and Healthy Communities." Public Health Reports **115**: 191-197.

Mohai, P. and R. Saha (2006). "Reassessing Racial and Socioeconomic Disparities in Environmental Justice Research." Demography **43**(2): 383-399.

Mohai, P. and R. Saha (2007). "Racial Inequality in the Distribution of Hazardous Waste: A National-Level Reassessment." Social Problems **54**(3): 343-370.

Moore, M. (2003). "Preterm labor and birth: what have we learned in the past two decades? ." Journal of Obstetrics Gynecology Neonatal Nursing **32**(5): 638–649.

Morello-Frosch, R. (2002). "The Political Economy of Environmental Discrimination." Environment and Planning C **20**: 477-496.

Morello-Frosch, R. and B. Jesdale (2006). "Separate and Unequal: Residential Segregation and Air Quality in the Metropolitan U.S." Environmental Health Perspectives **113**(3): 386-393.

Morello-Frosch, R. and R. Lopez (2006). "The riskscape and the color line: examining the role of segregation in environmental health disparities." Environ Res **102**(2): 181-196.

Morello-Frosch, R., M. Pastor, Jr., et al. (2002). "Environmental justice and regional inequality in southern California: implications for future research." <u>Environmental Health Perspectives</u> **110 Suppl 2**: 149-154.

Morello-Frosch, R., M. Pastor, et al. (2001). "Environmental Justice and Southern California's 'Riskscape': The Distribution of Air Toxics Exposures and Health Risks Among Diverse Communities." <u>Urban Affairs Review</u> **36**(4): 551-578.

Morello-Frosch, R. and E. Shenassa (2006). "The Environmental 'Riskscape' and Social Inequality: Implications for Explaining Maternal and Child Health Disparities." <u>Environmental Health Perspectives</u> **114**(8): 1150-1153.

Morello-Frosch, R. A. (2002). "Discrimination and the political economy of environmental inequality." <u>Environment and Planning C: Government and Policy</u> **20**: 477-496.

Morello-Frosch RA, P. M., Porras C, Sadd J (2002). "Environmental Justice and Regional Inequality in Southern California: Implications for Future Research." <u>Environmental Health Perspectives</u> **110**((Supplement 2).): 149-154.

Morenoff, J. (2003). "Neighborhood mechanisms and the spatial dynamics of birth weight." <u>American Journal of Sociology</u> **108**(5): 976-1017.

National Academy of Sciences (2009). Implementing Cumulative Risk Assessment. <u>Science and Decisions: Advancing Risk Assessment</u>, The National Academies Press**:** 213-239.

National Center for Health Statistics. (2003). "Vital statistics of the United States: 2003, volume I, Natality. Available at: <u>http://www.cdc.gov/nchs/data/statab/natfinal2003.annvol1_26.pdf.</u>"

National Environmental Justice Advisory Council (2004). Ensuring Risk Reduction in Communities with Multiple Stressors: Environmental Justice and Cumulative Risks/Impacts - Draft Report. New Orleans, Louisiana, National Environmental Justice Advisory Council, Cumulative Risks/Impacts Work Group**:** 80 Pages.

Navarro, V. (1990). "Race or Class Versus Race and Class: Mortality Differentials in the United States." <u>Lancet</u> **226**: 1238-1240.

O'Campo, P., X. Xue, et al. (1997). "Neighborhood Risk Factors for Low Birthweight in Baltimore: A Multilevel Analysis." <u>American Journal of Epidemiology</u> **87**(7): 1113-1118.

Odland, J. (1988). <u>Spatial Autocorrelation</u>. Newbury Park, CA, Sage Publications.

OEHHA (2003). The Air Toxics Hot Spots Program Guidance Manual for Preparation of Health Risk Assessments. Oakland, CA, Office of Environmental Health Hazard Assessment

California Environmental Protection Agency**:** 302 pages.

OEHHA. (2009). "Cumulative Impacts and Precautionary Approaches Project. Available at: <u>http://oehha.ca.gov/Listservs/ListSubscribe.asp?LID=79.</u>".

Olshan, A. F., L. Anderson, et al. (2000). "Workshop to identify critical windows of exposure for children's health: cancer work group summary." <u>Environmental Health Perspectives</u> **108**: 595-597.

Osmond, C. and D. Barker (2000). "Fetal, infant, and childhood growth are predictors of coronary heart disease, diabetes, and hypertension in adult men and women." <u>Environmental Health Perspectives</u> **108**(Suppl 3): 545-553.

Ostro, B., R. Broadwin, et al. (2006). "Fine Particulate Air Pollution and Mortality in Nine California Counties: Results from CALFINE." Environmental Health Perspectives **114**(1): 9-33.

Palaniappan, M., S. Prakash, et al. (2006). Paying With Our Health: The Real Cost of Freight Transport in California. Oakland, CA, Pacific Institute.

Palaniappan, M., D. Wu, et al. (2003). Clearing the Air: Reducing Diesel Pollution in West Oakland. Oakland, CA, Pacific Institute.

Parker, J., K. Schoendorf, et al. (1994). "Associations between measures of socioeconomic status and low birth weight, small for gestational age, and premature delivery in the United

States." Annals of Epidemiology **4**: 271–278.

Parker, J. and T. Woodruff (2008). "Influences of study design and location on the relationship between particulate matter air pollution and birthweight." Paediatric and Perinatal Epidemiology **22**: 214–227.

Parker, J., T. Woodruff, et al. (2005). "Air pollution and birth weight among term infants in California." Pediatrics **115**(1): 121-128.

Parker, J. D. and T. J. Woodruff (2008). "Influences of study design and location on the relationship between particulate matter air pollution and birthweight." Paediatr Perinat Epidemiol **22**(3): 214-227.

Pastor (1999). Building social capital to Protect Natural Capital: The Quest for Environmental Justice. Santa Fe, NM, Ford Foundation**:** 35.

Pastor, M., R. D. Bullard, et al. (2006). In the Wake of the Storm: Environment, Disaster, and Race After Katrina. New York, NY, Russell Sage.

Pastor, M., J. Sadd, et al. (2001). "Which came first? Toxic facilities, minority move-in, and environmental justice." Journal of Urban Affairs **23**(1): 1-21.

Pastor, M., J. Sadd, et al. (2004). "Waiting to Inhale: The Demographics of Toxic Air Release Facilities in 21st-Century California*." Social Science Quarterly **85**(2): 420-440.

Pastor, M., J. Sadd, et al. (2005). "The Air is Always Cleaner on the Other Side: Race, Space, and Air Toxics Exposures in California." Journal of Urban Affairs **27**(2): 127-148.

Ponce, N. A., K. J. Hoggatt, et al. (2005). "Preterm birth: the interaction of traffic-related air pollution with economic hardship in Los Angeles neighborhoods." Am J Epidemiol **162**(2): 140-148.

Pope, C. A. I., R. L. Verrier, et al. (1999). "Heart rate variability associated with particulate air pollution." American Heart Journal **138**: 890-899.

Pope, C. I. and D. Dockery (2006). "Health effects of fine particulate air pollution: lines that connect." Journal of the Air and Waste Management Association **56**: 709–742.

Pulido, L., S. Sidawi, et al. (1996). "An Archeology of Environmental Racism in Los Angeles." Urban Geography **17**(5): 419-439.

Ramachandran, G., J. Adgate, et al. (2000). "Comparison of short-term variations (15-min averages) in outdoor and indoor PM2.5 concentrations." Journal of the Air and Waste Management Association **50**: 1157-1166.

Rawlings, J., V. Rawlings, et al. (1995). "Prevalence of low birth weight and preterm delivery in relation to the interval between pregnancies among white and black women." New England Journal of Medicine **332**: 69–74.

Ringquist, E. J. (2005). "Assessing evidence of environmental inequities: A meta-analysis." Journal of Policy Analysis and Management **24**(2): 223-247.

Ritz, B. and M. Wilhelm (2008). "Ambient air pollution and adverse birth outcomes: methodologic issues in an emerging field." Basic Clin Pharmacol Toxicol **102**(2): 182-190.

Ritz, B., M. Wilhelm, et al. (2007). "Ambient air pollution and preterm birth in the environment and pregnancy outcomes study at the University of California, Los Angeles." Am J Epidemiol **166**(9): 1045-1052.

Ritz, B. and F. Yu (1999). "The effect of ambient carbon monoxide on low birth weight among children born in southern California between 1989 and 1993." Environmental Health Perspectives **107**(1): 17-25.

Ritz, B., F. Yu, et al. (2000). "Effect of air pollution on preterm birth among children born in Southern California between 1989 and 1993 " Epidemiology. **11**(5): 502-511. .

Rubenstein, S. (1995). High Rates in Bayview-Hunters Point. San Francisco Chronicle. San Francisco, CA**:** A1.

Sabin, L., K. Kozawa, et al. (2005). "Analysis of real-time variables affecting children's exposure to diesel-related pollutants during school bus commutes in Los Angeles." Atmospheric Environment **39**(29): 5243–5254.

Sadd, J., M. Pastor, et al. (1999). "'Every Breath You Take...': The Demographics of Toxic Air Releases in Southern California." Economic Development Quarterly **13**(2): 107-123.

Sadd, J. L., M. Pastor, Jr., et al. (1999). ""Every Breath You Take... ": The Demographics of Toxic Air Releases in Southern California." Economic Development Quarterly **13**(2): 107-123.

Saha, R. and P. Mohai (2005). "Historical Context and Hazardous Waste Facility Siting: Understanding Temporal Patterns in Michigan." Social Problems **52**(4): 618-648.

Salam, M., J. Millstein, et al. (2005). "Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: results from the Children's Health Study. ." Environmental Health Perspectives **113**: 1638–1644.

Sappenfield, W. M., J. W. Buehler, et al. (1987). "Differences in neonatal and postneonatal mortality by race, birth weight, and gestational age." Public Health Reports **102**(2): 182-192.

Shenkin, S. S., JM Deary, IJ (2004). "Birth weight and cognitive ability in childhood: a systemic review." Psychological Bulletin **130**(6): 989-1013.

Shi, L., J. Macinko, et al. (2004). "Primary care, infant mortality, and low birth weight in the states of the USA." Journal Epidemiology and Community Health **58**(5): 374-380.

Slama, R., V. Morgenstern, et al. (2007). "Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model." Environmental Health Perspectives **115**(9): 1283-1292.

Sørensen, H., S. Sabroe, et al. (1997). "Birth weight and cognitive function in young adult life: a historical cohort study." British Medical Journal **315**(7105): 401-403.

Sram, R., B. Binkova, et al. (2005). "Ambient air pollution and pregnancy outcomes: a review of the literature." Environmental Health Perspectives **113**: 375–382.

Stretesky, P. and M. J. Hogan (1998). "Environmental justice: An analysis of Superfund sites in Florida." Social Problems **45**(2): 268-287.

Su, J. G., R. Morello-Frosch, et al. (2009). "An index for assessing demographic inequalities in cumulative environmental hazards with application to Los Angeles, California." Environ Sci Technol **43**(20): 7626-7634.

U.S. EPA (1986). "Guidelines for carcinogenic risk assessment." Federal Register **51**(185): 33992-34003.

U.S. EPA (1990). Cancer risks from outdoor exposures to air toxics. Washington, DC, U.S. Environmental Protection Agency.

U.S. EPA. (2009). "National Air Toxics Assessment (NATA). Available at: http://www.epa.gov/ttn/atw/natamain/index.html Accessed August 24."

US Department of Commerce. Bureau of the Census. (2001). "CENSUS OF POPULATION AND HOUSING, 2000 [UNITED STATES]: CENSUS TRACT RELATIONSHIP FILES (CTRF) [Computer file]."

US EPA. (2007). "Air Quality System. http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm. Accessed 14 June 2007.".

USDHHS (1990). Age-adjusted Death Rates for Selected Causes of Death, According to Sex and Race: United States, Selected Years, 1950-87. Washington, DC, United States Department of Health and Human Services.

Wang, X., H. Ding, et al. (1997). "Association between air pollution and low birth weight: a community-based study." Environmental Health Perspectives **105**(5): 514-520.

Wilhelm, M. and B. Ritz (2003). "Residential proximity to traffic and adverse birth outcomes in Los Angeles county, California, 1994-1996." Environmental Health Perspectives **111**(2): 207-216.

Wilhelm, M. and B. Ritz (2005). "Local variations in CO and particulate air pollution and adverse birth outcomes in Los Angeles County, California, USA." Environmental Health Perspectives **113**: 1212–1221.

Windham, G. and L. Fenster (2008). "Environmental contaminants and pregnancy outcomes." Fertility and Sterility **89 (2 Suppl)**: 111-116.

Windham, G. C., L. Zhang, et al. (2006). "Autism Spectrum Disorders in Relation to Distribution of Hazardous Air Pollutants in the San Francisco Bay Area." Environmental Health Perspectives **114**(9): 1438-1444.

Woodruff, T., J. Parker, et al. (2009). "Methodological issues in studies of air pollution and reproductive health." Environ Res **109**(3): 311-320.

Woodruff, T., J. Parker, et al. (2003). "Disparities in exposure to air pollution during pregnancy." Environmental Health Perspectives **11**: 942-946.

Yandle, T. and D. Burton (1996). "Reexamining environmental justice: A statistical analysis of historical hazardous waste landfill siting patterns in metropolitan Texas." Social Science Quarterly **77**(33): 477-492.

Yanosky, J., P. Williams, et al. (2002). "A Comparison of two Direct-Reading Aerosol Monitors with the Federal Reference Method for PM2.5 in Indoor Air." <u>Atmospheric Environment</u> **36**: 107-113.