EPIDEMIOLOGIC INVESTIGATION TO IDENTIFY CHRONIC EFFECTS OF AMBIENT AIR POLLUTANTS IN SOUTHERN CALIFORNIA

Prepared for the California Air Resources Board and the California Environmental Protection Agency
Contract No. 94-331

Principal Investigator

John M. Peters, M.D., Sc.D.
Department of Preventive Medicine
University of Southern California
Los Angeles, CA

Participating Researchers

Edward Avol, M.S., Kiros Berhane, Ph.D., W. James Gauderman, Ph.D., Frank Gilliland, M.D., Ph.D., Michael Jerrett, Ph.D., Nino Künzli, M.D., Ph.D., Stephanie London, M.D., Rob McConnell, M.D., Bill Navidi, Ph.D., Edward Rappaport, M.S., Duncan Thomas, Ph.D.
Department of Preventive Medicine
University of Southern California
Los Angeles, CA

Frederick Lurmann, M.S., Paul Roberts, Ph.D., Siana Alcorn, Tami Funk
Sonoma Technology, Inc.
Petaluma, CA

Henry Gong, Jr., M.D., William S. Linn, M.S.
Los Amigos Research and Education Institute
Downey, CA

Glen Cass, Ph.D. (in memorial)
California Institute of Technology
Pasadena, CA

Helene Margolis, Ph.D.
California Air Resources Board
Sacramento, CA

May 14, 2004

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 projects.

Acknowledgements

The contributions of the California Air Resources Board staff, particularly Dane Westerdahl and both John Holmes and Bart Croes as Chiefs of Research, are greatly appreciated. The contributions of Clint Taylor, Ken Bowers, Steve Mara, Cindy Stover, Barbara Weller and Richard Bode, to acquisition of air quality and exposure data are also greatly appreciated.

We thank the staff, students and parents of the schools for their cooperation. We also appreciate the cooperation of the school districts in the 12 communities.

We wish to acknowledge valuable contributions from our External Advisory Committee composed of David Bates, Morton Lippmann, Jonathan Samet, Frank Speizer, Jack Spengler, Arthur Winer, James Whittenberger, and Scott Zeger.

We thank Tracy Bastain especially for her tireless efforts to keep us on track.

We thank the local air pollution control districts for operation of the air monitoring stations, including the South Coast Air Quality Management District, the Mojave Desert Air Quality Management District, the San Luis Obispo County Air Pollution Control District, and the San Diego County Air Pollution Control District.

We also acknowledge the efforts of John Collins (Collins Consulting) and David Bush (Technical and Business Systems, Inc.) in establishing and maintaining the quality assurance and quality control programs for the Children's Health Study (CHS).

We are grateful for the important contributions of the following individuals to the exposure assessment elements of the project: Susanne Hering of Aerosol Dynamics, Inc.; Jun Wu and Arthur Winer of the University of California, Los Angeles; Karen Anderson and Ken Clark of Los Amigos Research and Education Institute; Stephen Colome of Integrated Environmental Services; Jamie Shauer of the University of Wisconsin; and Lynn Salmon of California Institute of Technology.

We gratefully acknowledge support for this research by the California Air Resources Board, the California Environmental Protection Agency, the National Institute of Environment Health Sciences, the National Heart, Lung, and Blood Institute, the U.S. Environmental Protection Agency, and the Hastings Foundation. Additional support was provided by funding contributions from Toyota USA and the South Coast Air Quality Management District. This Report was submitted in fulfillment of ARB Contract 94-331, "Epidemiologic Investigation to Identify Chronic Effects of Ambient Air Pollutants in Southern California," by the University of Southern California under the sponsorship of the California Air Resources Board. Work was completed as of May 14, 2004.

Table of Contents

List of Figures

List of Tables

Abstract

A prospective study of about 6000 children living in 12 Southern California communities of varying ambient air pollution profiles was initiated in 1993. The primary purpose of the study was to determine whether air pollution causes chronic adverse respiratory health effects. Particulate matter (hourly PM_{10}, two-week-integrated $PM_{2.5}$, and several constituents including elemental and organic carbon, metals, and ions), ozone (O_3), nitrogen dioxide (NO_2), and acid vapor (primarily nitric) were measured in each community during the study period. Health outcomes assessed were annual pulmonary function tests (maximal spirometry), annual questionnaires on respiratory conditions and symptoms, and school absence monitoring. Demographics, housing characteristics, time-activity patterns and exposure to tobacco smoke were also assessed annually by written questionnaire. Study results indicated that children's lung function growth was adversely affected by air pollution, new cases of asthma and asthma exacerbations were associated with ambient air pollution levels, and school absences from acute respiratory illnesses followed rises in ozone levels. We conclude that current levels of ambient air pollution in Southern California are associated with clinically important chronic health effects that have substantial health and economic impacts. These findings indicate the need for cleaner air for our children to breathe.

1. Executive Summary

1.1. Background

Air pollution in Southern California continues to pose significant challenges to regulatory agencies and to health professionals. Several million persons living in the region are exposed to pollution levels that have been associated, in laboratory and field investigations, with acute and sub-acute respiratory effects. When the Children's Health Study (CHS) began in the early 1990s, it was known from laboratory observations that acute exposure to air pollutants produced decrements in pulmonary function, increased prevalence of respiratory symptoms, and respiratory tract inflammation. The paramount CHS research question has been whether *chronic* respiratory disease occurs as a result of breathing polluted ambient air.

In short-term exposure studies of humans in controlled exposure chambers, among common air pollutants, ozone shows the strongest evidence of adverse effects. Numerous laboratory exposure-response studies in human volunteers have shown that lung function losses, respiratory irritant symptoms, and increases in bronchial reactivity result from ozone exposure levels commonly observed in the South Coast Air Basin (SoCAB), either from comparatively brief (~1 hour) exercise at "alert" concentrations of 0.2 ppm and higher or from prolonged exercise at concentrations near the California ambient air quality standard of 0.09 ppm (US Environmental Protection Agency 1986; Folinsbee et al. 1988; Lippmann 1989; Lippmann 1991). Recovery to normal function levels typically takes several hours after ozone exposure ceases. Some effects of short-term exposure persist for more than 24 hours. At the time the study began, similar acute effects had not been seen from other pollutants at the levels encountered in Southern California [nitrogen oxides (NO_x), particulate matter less than 10 microns in diameter (PM_{10}) or acid vapors]. Since the study began, many hundreds of papers have been published demonstrating the relationship between pollutant levels and morbidity and mortality. This literature is well summarized by Brunekreef and Holgate (2002).

Studies of humans conducted in Southern California have suggested the possibility of chronic respiratory effects from air pollution (Detels et al. 1987; Abbey et al. 1991; Sherwin 1991; Sherwin and Richters 1991), but because of population attrition in the Detels studies, reliance on questionnaire data in the Abbey study, and possible confounding in the Sherwin study, conclusions are uncertain. When the Children's Health Study began, essentially no human data on children existed on chronic respiratory effects resulting from specific components of air pollution. The large number of persons in Southern California exposed to air pollution, the existing data on acute effects, and the available air monitoring data have provided a unique opportunity to examine chronic health effects resulting from air pollution in humans. The identification of health effects plus the generation of dose-response data provides regulators with highly valuable information for risk management.

Children were selected as the study population for several reasons: they often spend more time outdoors; they exercise more than adults; they do not smoke (at least the young ones); they do not have hazardous occupations; they are more likely to have spent their entire lives in Southern

California; their growing lungs may be more sensitive to the effects of air pollution and they are accessible in large numbers through schools.

1.2. Methods

Community selection was based on air pollution levels and exposure patterns plus demographic data of a group of census tracts in 86 communities. The basic principle governing the selection of communities was to select a group of communities having widely divergent exposure characteristics. A second principle we followed was that the communities being compared should be similar with respect to potential confounding variables. Following these principles, we selected 12 communities in 6 Southern California counties.

Participating study schools were selected based on: (1) location in a pre-selected community of interest based on air pollution levels and patterns; (2) sufficient population of target-aged children; (3) preponderance of children attending school from the immediate neighborhood; (4) demographic similarity with other potential and participating community school sites; (5) absence of localized air pollution sources such as close proximity to factories or freeways; and (6) proximal location to a fixed-site air monitoring station. The design approach specified child entry into the study at the fourth, seventh, and tenth grades and required the enrollment of at least four schools in each community (two elementary schools, a junior high school, and a senior high school).

Three cohorts were established in 1993; one with about 900 tenth grade students, another with about 900 seventh grade students, and still another with about 1800 fourth grade students. These cohorts are referred to in this report as cohorts A, B, and C. In 1996, about 2,000 additional fourth grade students were enrolled in the study. This cohort is referred to as cohort D. In each case, students who continued to reside in the twelve communities were evaluated annually through high school graduation (twelfth grade). Students in cohort D will graduate from high school in 2004. In this report, we report on data collected through high school graduation on cohorts A, B, and C. We also report on four years of follow-up data on cohort D.

The CHS written questionnaire was composed of several sections: demographics, a medical history, a housing survey, exposure to tobacco smoke, exposure to pets and pests, and a time-activity assessment. An extensive set of questions was asked about the history of respiratory diseases. These included asthma, bronchitis and pneumonia and associated symptoms such as cough, phlegm production and wheezing. The initial questionnaire collected information on the past history of these conditions and symptoms including frequency and time of onset. Asthma questions considered physician diagnosis, severity and medication use. Each annual follow-up questionnaire concentrated on adverse respiratory health experiences during the past year and allowed us to ascertain the incidence of new-onset conditions such as physician-diagnosed asthma and bronchitis.

Lung function testing took place in the spring of 1993 and in each subsequent spring to minimize seasonal confounding with intercurrent summer or winter acute air pollution episodes. The subjects were asked to perform at least 3 satisfactory maximal expiratory maneuvers. A maximum of 7 efforts were attempted. Six testing units (spirometers), operated by trained lung function technicians, were dispatched to conduct field-testing in a given community. Each

community was visited at least twice (at least one month apart) with half the participating subjects being tested each visit. The annual follow-up pulmonary function tests were planned to achieve as close to a 12-month interval between testing as possible. Heights and weights of the subjects were measured at the time of each lung testing.

The absence monitoring activity was designed to collect data to determine the frequency and severity of respiratory illnesses in relation to concurrent ambient air pollution levels and to compare respiratory disease patterns between communities and by exposure to various pollutants. Because schools were required by the State Department of Education to keep data on absences in order to receive capitation funding for students for most of the effective study period, there was motivation for schools to collect accurate data. We used documented school absences to trigger an investigation of the reason for the absence. This involved phoning the student's home to interview the parent or guardian. By this approach we were able to classify whether the illness was respiratory. We also asked whether the child had seen a doctor related to the reported absence, and if so, the doctor's diagnosis was noted. We asked about use of medications since this might provide an indication of the severity of the illness.

Monitoring stations were established in each of the twelve communities. This was accomplished by augmenting seven existing regional air monitoring stations and creating five new stations in late 1993 and early 1994. Continuous hourly measurements of ozone (O_3), nitrogen dioxide (NO_2), and PM_{10} were made at each station. Integrated measurements of particulate matter less than 2.5 microns in diameter ($PM_{2.5}$) mass, PM chemical constituents, and acid vapors were made using a multi-legged two-week sampler (TWS) designed for the study (Hering et al. 1994; Lurmann et al. 1994). The PM chemical constituents included $PM_{2.5}$ sulfate, nitrate, and ammonium and PM_{10} elemental carbon (EC) and organic carbon (OC). The main carbon sampling leg did not have a size-selective inlet; however, testing indicated the particle size-cut was approximately 10 μm. A second carbon sampling leg was implemented in 2001 with a 2.2 μm size cut for comparison purposes. Throughout this report, references to $PM_{2.5}$ EC and OC concentrations refer to concentrations derived from the ~PM_{10} EC and OC measurements by application of suitable adjustment factors. The TWS also collected samples for determination of concentrations of nitric acid, hydrochloric acid, formic acid, and acetic acid (collectively described in this report as acid vapor). These measurements were made throughout the study period, 1994-2001. Additional measurements of carbon monoxide (CO), particle number (PN), $PM_{2.2}$ EC, $PM_{2.2}$ OC, and $PM_{2.2}$ elements by x-ray fluorescence (XRF) were implemented in most communities in 2000 and 2001. Also in 2001, the measured nitric oxide (NO) concentrations were retrospectively added to the database. Hourly temperature and relative humidity were measured at some of the CHS air monitoring stations for some of the years to complement the air quality data. These data were supplemented with meteorological data collected at locations near the CHS communities.

Information on usual time-activity profiles and household characteristics were collected annually for all CHS participants. These variables were used directly in health models as potential confounders or effect modifiers, and they were used indirectly in models of microenvironmental concentrations of O_3, PM_{10}, $PM_{2.5}$, and NO_2 in homes, schools, and vehicles to derive estimates of individual exposures to these pollutants. Traffic density data on freeways and major arterial roadways were combined with meteorological data, using line-source dispersion models, to

predict local pollution concentrations at all CHS participants' homes and all schools. These model-based predictions were supplemented by measurements of NO_2 concentrations during two 2-week periods in 287 homes across the 12 CHS communities.

Multi-level random effects models were used for the statistical analysis of the health outcome data in relation to air pollution and other risk factors. This approach provides a unified and valid way to assess associations at three levels of comparison: over years, between individuals, and between communities.

1.3. Results

Our findings demonstrated an association between breathing polluted air in Southern California and significant chronic deficits in lung function among adolescent children. We observed air pollution effects on lung function level at study entry (youngest cohort, age 10yrs), on 4-year lung function growth (age 10-14 years) in two independent cohorts, on 8-year lung function growth (age 10-18 years) in the original fourth grade cohort, and on the maximum rate of lung function growth during adolescence (over the study period). Air pollution exposure over the 8-year (from fourth grade to twelfth grade) study period was also linked to clinically significant deficits [forced expiratory volume in one second (FEV_1) below 80% predicted] in lung function at age 18 years. We found that there were three to five times more children with clinically significant deficits in lung function living in communities with high outdoor air pollution levels compared to communities with low pollution levels. In a subset of children who moved away from their original study community, we observed consistent associations of changes in lung function growth rates with corresponding changes in ambient air pollution exposure between their former and current communities of residence. The pollutants most closely associated with lung function deficits were NO_2, acids (either inorganic, organic, or a combination of the four acids monitored), PM_{10}, and $PM_{2.5}$. Several constituents of $PM_{2.5}$, including EC, nitrate, and ammonium, also showed associations with lung function growth. However, the inter-correlation among PM pollutants, and their high correlations with NO_2 and acid, limited our ability to distinguish the independent effects of any one of these pollutants.

Our findings demonstrated effects of air pollution on both new onset asthma and asthma exacerbations. Prior to the performance of the CHS, the prevailing scientific view was that air pollution made existing asthma worse but that it did not *cause* new cases to develop. Study data showed that new cases of asthma are much more likely to occur in high ozone communities, especially among those children who exercise regularly and at elevated levels. Additionally, our analyses regarding exposure to traffic-related air pollution have found associations between proximity to high traffic density (a marker for pollutant exposure) and increased risks for prevalent asthma among children.

We have demonstrated that air pollution is related to bronchitic symptoms and that asthmatics are more likely to be affected than non-asthmatics. Evaluation of the longitudinal data implicated NO_2 and organic carbon as being responsible for the observed effects.

Our results showed that short-term changes in O_3, but not NO_2 or PM_{10}, were associated with a substantial increase in school absences from both upper and lower respiratory illness. Absences were significantly increased 2 to 3 days after exposure and reached a peak on day 5 after

exposure. Because exposures at the levels observed in this study are common, the increase in school absenteeism from respiratory illnesses associated with relatively modest day-to day changes in O_3 concentration documents an important adverse impact of O_3 on children's health and well-being.

Our data also demonstrate an association between ozone levels and birth weight of children. High ozone levels during the second or third trimester of pregnancy are associated with lower birth weight. Other manuscripts resulting from this study have demonstrated the important health effects associated with maternal smoking, environmental tobacco smoke, genetics, obesity, and dietary factors.

1.4. Conclusions

Our main conclusion is that current levels of air pollution in Southern California are associated with several serious health effects that are costly to children's health and to the state. Lung function was found to be consistently associated with a package of highly correlated pollutants that include particulates, NO_2, and acids, but not ozone. This impact of vehicle-related pollution on children's lung function is likely to have life-long adverse health sequelae. The demonstration of strong evidence linking exposure to new cases of asthma (the most common chronic disease of childhood) to ozone is another striking association. It is also important to note that most of these associations extend to pollution levels below current ambient air standards and may exert significant health effects. Taken as a whole, the results from the Children's Health Study should provide scientific support for aggressive and accelerated efforts to achieve clean air for our children to breathe.

2. Introduction

2.1. History

The Children's Health Study (CHS) began in 1991 and has continued through 2004 with support from the California Air Resources Board. The study has been divided into three phases. The first phase (Phase I) was designed to develop and evaluate the best methodologies to measure exposures, assess health effects and select study populations. To provide expertise on each of these important issues, three blue ribbon panels of outside advisors were convened. Phase I was completed in 1992 (Peters 1992).

Phase II was a cross-sectional study of children living in 12 Southern California communities during the 1992-1993 school year. The Phase II cross-sectional study was conducted to provide early information on the possible chronic effects of air pollution in Southern California children and to determine, if effects were found, which pollutant (or pollutants) was responsible. To do this, about 3600 children from 12 communities with differing air pollution patterns were enrolled in the study from public schools within the respective communities. About 50% of the children were in fourth grade (Cohort C), about 25% in seventh (Cohort B) and about 25% in tenth (Cohort A).

Annual questionnaires were completed which covered health history, residential history, housing characteristics, and history of exposure to other possibly harmful agents, such as tobacco smoke (both active and passive smoking). In addition, the physical and outdoor/indoor activity of each subject was ascertained. The lung function of each subject was assessed and school absences were recorded to determine frequency and severity of respiratory illnesses.

Monitoring of air pollutants was conducted in the communities, the schools and (through a separately funded project) at a sample of the subjects' residences. Ozone, PM_{10}, $PM_{2.5}$, NO_x, and acid (nitric, formic, acetic, and hydrochloric) concentrations were determined at the community level. Ozone, PM_{10}, $PM_{2.5}$, formaldehyde, air exchange rates and airborne acids were measured at a sample of residences. Ozone was measured inside and outside the schools. The Phase II final report was submitted in 1996 (Peters 1997).

The third phase (Phase III) of the Children's Health Study consisted of the follow-up of the three cohorts established in 1993 plus the establishment of a new cohort of about 2000 fourth graders (Cohort D) in 1996 in the same twelve communities.

This report emphasizes the results of the follow-up of the four cohorts but also synthesizes the results over the entire period of study.

2.2. Background

Air pollution in Southern California continues to pose significant challenges to regulatory agencies and to health professionals. Several million persons living in the region are exposed to pollution levels that have been associated, in laboratory and field investigations, with acute and sub-acute respiratory effects. At the time this study began, it was known from laboratory

observations that acute exposure to ozone produced decrements in pulmonary function, increased prevalence of respiratory symptoms, and respiratory tract inflammation. The paramount research question has been whether chronic respiratory disease occurs as a result of breathing polluted ambient air.

In short-term exposure studies of humans in controlled exposure chambers, among common air pollutants, ozone shows the strongest evidence of adverse effects. Numerous laboratory exposure-response studies in human volunteers have shown that lung function losses, respiratory irritant symptoms, and increases in bronchial reactivity result from ozone exposure levels attained in the South Coast Air Basin (SoCAB), either from comparatively brief (~1 hour) exercise at concentrations of 0.18 ppm and higher or from prolonged exercise at concentrations near the California ambient air quality standard of 0.09 ppm (US Environmental Protection Agency 1986; Folinsbee et al. 1988; Lippmann 1989; Lippmann 1991). Recovery to normal function levels typically takes several hours after ozone exposure ceases. Some effects of short-term exposure persist for more than 24 hours. Similar acute effects have not been seen from other pollutants at the levels encountered in Southern California (NO_x, PM_{10} or acid vapors), although acute effects of ambient pollution associated with hospital admissions have been reported (Katsouyanni et al. 1997; Spix et al. 1998; Samet et al. 2000a; Le Tertre et al. 2002; Zanobetti et al. 2002). Changes in PM_{10} levels have been associated with trends in mortality in the United States and other parts of the world (Brunekreef and Holgate 2002).

Studies of humans conducted in Southern California have suggested the possibility of chronic respiratory effects from air pollution (Detels et al. 1987; Abbey et al. 1991; Sherwin 1991; Sherwin and Richters 1991), but because of population attrition in the Detels studies, reliance on questionnaire data in the Abbey study, and possible confounding in the Sherwin study, conclusions are uncertain. At the time the Children's Health Study began, essentially no human data existed on chronic respiratory effects resulting from specific components of air pollution. In fact, chronic effect studies are still rare (Sunyer 2001; Brunekreef and Holgate 2002) and the focus has usually been on mortality in adults, whereas chronic effects of air pollution on morbidity remains a major need of research (Dockery et al. 1993; Pope et al. 1995; Abbey et al. 1999; Hoek et al. 2002; Pope et al. 2002; Pope et al. 2004). The large number of persons in Southern California exposed to air pollution, the existing data on acute effects, and the available air monitoring data provided a unique opportunity in Southern California to examine chronic health effects resulting from air pollution in humans. Careful construction of lifetime exposures, taking into account temporal and spatial patterns of physical activity; sensitive measures of pulmonary responses; and thoughtful consideration of study design, confounding, and bias provided the framework for addressing critical public health questions regarding long-term exposure to ambient air pollution.

The focus of the investigation has been public school-aged children. We decided to study children for several reasons: they often spend more time outdoors; they exercise more than adults; they do not smoke (at least the young ones); they do not have hazardous occupations; they are more likely to have spent their entire lives in Southern California; their growing lungs may be more sensitive to the effects of air pollution and they are accessible in large numbers through schools. Over the course of the study, it became evident that it is also crucially important to identify any highly susceptible subgroups. The mechanistic hypotheses discussed below, for

example, suggest that individuals with specific genotypes may be particularly sensitive to the effects of air pollution. Air quality standards, which aim not only to protect the population as a whole but also the most sensitive individuals, will need to take such findings into account.

2.3. Public Health Significance

The U.S. Environmental Protection Agency is currently engaged in a process for updating the National Ambient Air Quality Standard for particulate pollution and for ozone. Data from the Children's Health Study have figured prominently in the draft revision of the Criteria Document (US Environmental Protection Agency 1997c; US Environmental Protection Agency 2003). Unlike much of cancer risk assessment, which relies heavily on extrapolation from animal toxicology studies, the national standards for particulate pollutants have been based primarily on epidemiologic evidence. Until recently, the majority of such epidemiologic data have related to time series studies of the short-term association between daily fluctuations in mortality, hospital admissions, asthma severity, and other outcomes on daily fluctuations in air pollution levels (and confounders like weather) (Pope et al. 1991; Romieu et al. 1996; Schwartz et al. 1996; Delfino et al. 1998; Samet et al. 2000b; Yu et al. 2000). These studies provide little guidance about the impact of air pollution on *children's* health. The Harvard Six-Cities Study (Dockery et al. 1993) and the American Cancer Society's cohort (Pope et al. 1995; Pope et al. 2002) provide longitudinal data on the chronic health effects of air pollution, but both of these studies were confined to adults. Only limited cross-sectional data have been available on the effects of air pollution on childhood lung function and respiratory illnesses, with all the usual problems of interpretation of such data: it is possible for example, that individuals who have suffered from respiratory effects of air pollution in the past may decide to move to cleaner communities, so that clean communities may over-represent sensitive individuals and polluted communities may under-represent them, leading to a spurious negative association. Other confounding factors (e.g., pollens) and differences in diagnostic patterns between communities could also account for such a finding. Only longitudinal observations within communities and within individuals can overcome such selection bias and between-community confounding. This has been the rationale for conducting the Children's Health Study, and why longitudinal data, such as the type collected in the CHS, is likely to continue to be of major importance for setting air pollution standards.

Perhaps the most important regulatory question concerns the identification of the specific pollutants or source of pollution that are responsible for any observed health effects. Little controversy remains that current levels of air pollution have demonstrable effects on health, but there is great controversy over the contribution of specific pollutants to these observed effects. Regulations aimed at controlling the wrong pollutant could have counter-productive effects.

2.4. Biological Rationale

Given the large number of variables under study – health endpoints, air pollution measures, confounding and modifying factors – we feel strongly that our analysis must be guided by an understanding of the underlying biological mechanisms to avoid degenerating into data dredging. This framework is laid out most clearly in the paper by Gilliland et al. (1999b), which outlines our understanding of mechanisms of lung injury and response at the cellular and molecular levels.

We hypothesize that exposure to air pollution produces oxidative and nitrosative stress, which mediates the adverse effects of air pollution Figure 2.4-1. Air pollution contains a wide variety of oxidants, such as O_3, NO, NO_2, and reactive organic compounds. In addition, reactive oxygen species (ROS) are generated during the metabolism of polyaromatic hydrocarbons (PAHs), by redox cycling of quinones and transition metals, and by the inflammatory process which occurs in response to air pollution exposure. The respiratory tract has enzymatic and non-enzymatic antioxidant defenses that counteract the effects of oxidants. These defenses include antioxidant vitamins and enzymes that metabolize ROS and detoxify oxidation products. Chronic respiratory effects arise from exposure to gaseous air pollutants (O_3, NO_2, assorted acids) and particulates that produce chronically increased oxidative stress, alterations in immune regulation, and repeated pathologic inflammatory responses that overcome lung defenses and disrupt the normal regulatory and repair processes. Because oxidative stress is central to the pathogenesis of adverse respiratory outcomes, individuals who have an imbalance between oxidative stress and their antioxidant defenses may be at greatest risk for adverse respiratory health outcomes.

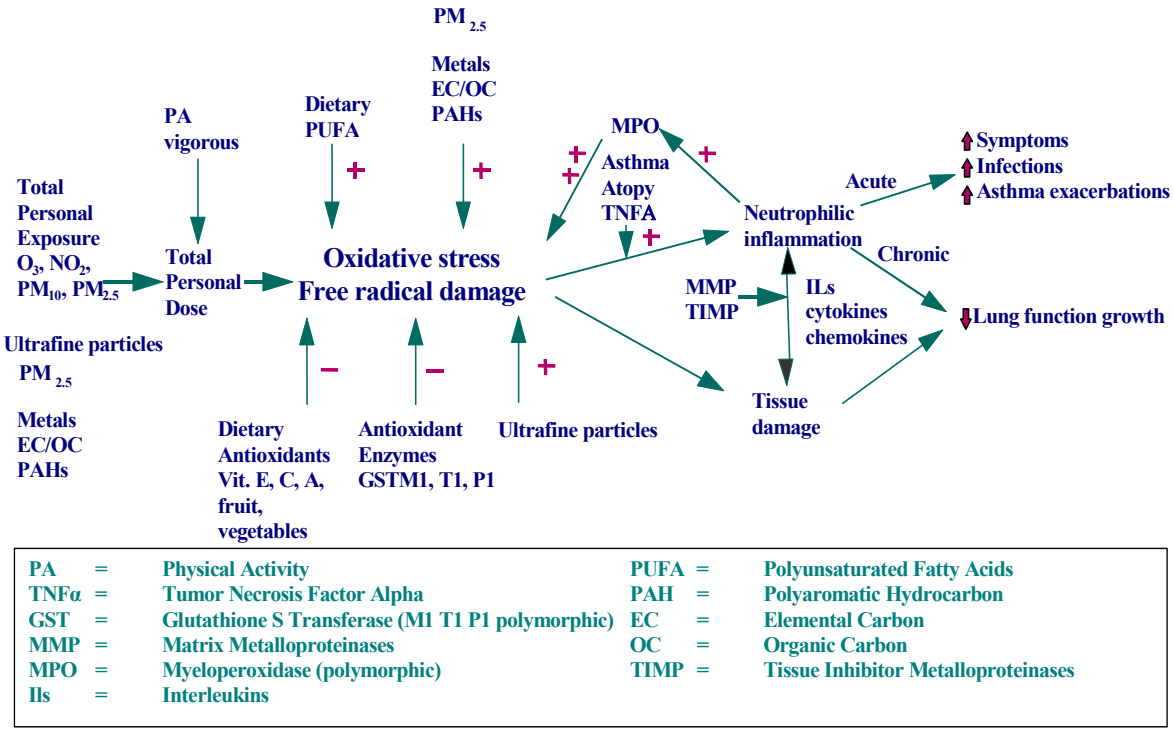


Figure 2.4-1. Biological impact pathway for ambient air pollutants, acute and chronic respiratory effects and susceptibility factors

Multiple factors, including lifestyle (diet and physical activity) and genetic variation determine the intensity and biologic effects of oxidative stress following exposure to air pollution. The first line of antioxidant defense is the respiratory extracellular lining fluid (RELF). For gases and soluble liquid droplet aerosols, clearance occurs by dissolution and reactions with RELF components (e.g., antioxidants) or absorption into epithelial cells. For example, O_3 is highly reactive with unsaturated lipids, antioxidants, proteins, and mucopolysaccharides. Insoluble particles are cleared from the proximal airway by the mucociliary elevator, whereas those in the distal airways and alveoli are dissolved or cleared by alveolar macrophages, but some may be deposited around small airways. It appears that some ultrafine metallic particles are retained for longer periods in the distal lung. Retained particles may contribute to adverse effects by enhancing free radical production in these regions. The level of oxidative stress resulting from the interactions with the RELF is affected by the level of antioxidants and antioxidant enzymes such as glutathione-S-transferases (GSTs), glutathione peroxidases (GPXs), and superoxide dismutases (SODs). Beyond the RELF, cells are equipped with extensive oxidant defenses that are built from the same antioxidants and antioxidant enzymes. In addition, other enzymes detoxify oxidation products to reduce the damage from oxidation that inevitably occurs despite the defenses.

Dietary intake contributes to antioxidant defenses by modulating levels of antioxidant vitamins and antioxidant enzymes. Children with low fruit and vegetable intake, low antioxidant intake, or high polyunsaturated fat intake may be at increased risk for adverse effects. Dietary intake of antioxidants, fruits and vegetables may therefore be important in protecting the airway from oxidant and radical air pollutants, especially among children with at-risk genotypes. We hypothesize that genetic variants that increase oxidative stress or decrease antioxidant defenses are likely to increase the occurrence of adverse outcomes. A better understanding of the genetic polymorphisms in oxidative stress pathways may allow identification of children at greatest risk of adverse respiratory effects from air pollution.

2.5. Hypotheses

As the Children's Health Study progressed, our original research questions evolved into six major hypotheses. They are presented here to guide the reader and appear again with commentary in Section 5 – Discussion and Synthesis.

- Lung growth in children is permanently affected by air pollution in Southern California, leading to permanent deficits in lung function.
- Respiratory illnesses are more frequent and severe in children living in areas of higher pollution.
- Children with more frequent and severe respiratory illness suffer more severe deficits in pulmonary function.
- Subsets of children can be identified who are more susceptible to the effects of air pollution.
- If chronic effects are detected, they can be attributed to a specific pollutant or combination of pollutants.
- If effects are seen, and a pollutant can be identified as the agent, thresholds below which effects are not detectable can be identified.

In the following sections, the methods and results from the past decade of the Children's Health Study are highlighted. Numerous findings from the CHS have been published (see Appendix) in peer-reviewed journals. The following report provides the reader with an overview of the methodologies employed, the results observed, the outcomes determined, and the investigators' interpretation and synthesis of what the Children's Health Study means for public health.

3. Methods

3.1. Introduction

This methods section describes the overall approach to carrying out the study and is divided into five major components: Sections 3.2 and 3.3 consist of a description of community and school selection and student enrollment; Section 3.4 describes the exposure assessment procedures; Section 3.5 highlights the health effects assessment; Section 3.6 outlines the statistical approaches; and Section 3.7 covers quality assurance and quality control.

3.2. Study Design

3.2.1. Community Selection

Community selection was based on air pollution levels and exposure patterns plus demographic data of a group of census tracts in 86 communities. The approach used to estimate past exposures to air pollutants was based on data collected in 1986-1990 by monitoring stations existing prior to our own study. This allowed estimates of O_3, PM_{10}, NO_2, and acid vapor [defined as the sum of both strong acids (nitric and hydrochloric) and organic acids (formic and acetic)] with spatial interpolation as needed, particularly for acid vapor. The demographic data consisted of about thirty variables for each census tract within 5 kilometers of each monitoring site, including racial and age composition, housing characteristics, and educational levels. These data were derived from the 1990 U.S. Census and available in tabular form, and were analyzed by Dr. Glen Cass at the California Institute of Technology (CalTech).

The basic principle governing the selection of communities was that greater dispersion in the design variables tends to allow for more accurate estimates of effects. This principle dictated that the site selection algorithm should be formulated to select a group of communities having widely divergent exposure characteristics. Because we also wanted to be able to distinguish the effects of different pollutants, we also sought to select communities that would minimize the correlations between pollutants.

A second principle we followed was that the communities being compared should be similar with respect to potential confounding variables. Because the real confounders cannot be predicted in advance and because routine data on them may not be available at the level of census tracts, we were restricted to using those variables available from census data. Perfect balance between communities on these variables was unlikely to be achievable and if it were, it would not have guaranteed comparability of the real confounders. Hence, we relied on two corollaries of this principle: first, that heterogeneous communities are preferred because they would be more likely to exhibit overlapping distributions of risk factors, thus improving the prospects for making adjustments for confounding in the analysis; and second, replication of exposure profiles was highly desirable, to improve the chance of including demographically comparable communities and to allow estimation of residual variance within pollution profiles.

Because of the large number of variables, it was necessary to devise an orderly procedure for considering them. (See Navidi et al. 1994, for details). We decided to make an initial selection on the basis of the occurrence of the main pollutants and then use the demographic data to select the final communities from amongst candidates with comparable exposures.

Two basic approaches were considered for the site selection algorithm. We refer to them as the *distance approach* and the *factorial approach*. Both were designed to yield a selection of communities varying greatly in exposure characteristics. Both relied on a common set of measured and interpolated ambient concentrations for 86 air monitoring sites in the study area for the period from 1986 to 1990. Because of differences in the number of locations at which pollutants were measured and the frequency and type of measurements made, we judged that the data were more reliable for ozone than PM_{10}, and more reliable for PM_{10} than NO_2 and more reliable for NO_2 than acid vapor. First, as a central point for comparison, the regional average concentration for each pollutant was calculated. Second, the deviation from the regional average concentration of each pollutant was determined for each site. The deviation was expressed as the difference from the average, measured in standard units. Other methods of calculating deviations were considered. Some of them involved transforming the original data to a log scale, others involved expressing deviations as a percentage above or below the average. We found these alternative calculations to have virtually no impact on the final selection of communities.

In the distance approach, the standardized deviations of the site-specific five year means of each of the three pollutants from their respective overall averages were squared and added to produce a single measure of the "distance" from that site to the central point in three-dimensional pollution space. Communities whose distances from the center were great were those whose pollution levels were generally well above or well below average. In the factorial approach, each site was categorized as high or low for each pollutant (as determined by whether the level of the pollutant is above or below average at that site identified as "hi/low" or "+/-" in subsequent discussions). For each pattern of low and high exposure, an attempt was made to select two communities whose pollution profiles matched that pattern. The data were computerized and analyzed according to both approaches.

Groups of communities selected by the distance method turned out to be less satisfactory than groups of similar size selected by the factorial method because the distance method resulted in an overrepresentation of sites exhibiting high ozone levels, reducing the power to determine the specific contribution of ozone level to observed differences in health outcomes between communities.

Replication of study sites was considered highly desirable for both statistical and other study design reasons mentioned previously. We attempted to select the two sites representing similar pollution profiles with as much geographical distance between them as possible but discovered that this was often not possible because the same pollution profile tended to occur in the same local areas.

After studying the pollution patterns and the available communities, we identified seven distinct air pollution profiles based on the four pollutants of interest. We attempted to choose more than one community for each pollution profile when that was possible.

Two communities with similar exposure profiles could differ systematically in health outcomes, for reasons such as differences in confounding factors (e.g. demographics), differences in the degree of systematic measurement error such as might be caused by machine malfunction, and differences in biases due to events such as viral epidemics, local sources of pollution or local school administrative practices. Replication of communities was necessary to separate the effects of pollution from the effects of these other factors.

Furthermore, without replication it would have been impossible to estimate accurately at the aggregate level the proportion of variation in health effects due to the pollutants under study as opposed to other unmeasured factors. Additional methodologic details are provided in the original cross-sectional publications (Peters et al. 1999a; Peters et al. 1999b).

The final set of communities selected for study is presented in Table 3.2-1 along with their deviations from regional average long-term pollutant concentrations. Demographic characteristics of the communities are described below in relation to the samples of study subjects actually enrolled and are presented in Table 3.3-1.

Table 3.2-1. Deviation* from regional average long-term concentrations in selected communities.

Communities Selected	Peak O_3	PM_{10}	HNO_3	NO_2
San Dimas	+1.8	+1.2	+1.3	+1.1
Upland	+1.5	+2.2	+1.5	+1.5
Mira Loma	+1.6	+2.8	-0.9	+0.7
Riverside	+1.6	+2.8	-0.9	+0.7
Alpine	+0.9	-0.5	-0.8	-0.4
Lancaster	+0.5	+0.2	+1.5	-0.7
Long Beach	-0.8	+0.4	+0.7	+1.5
Lake Elsinore	+1.0	+1.1	+0.4	-0.5
Lake Arrowhead	+1.8	-0.4	+0.5	+0.1
Santa Maria	-1.5	-1.1	-1.1	-1.0
Atascadero	-0.1	-1.1	-1.1	-1.0
Lompoc	-1.3	-1.0	-1.1	-1.4

* Deviation = $\frac{\text{[site mean - regional mean]}}{\text{regional standard deviation}}$

3.3. Study Populations

3.3.1. Cohorts

Three cohorts were established in 1993; one with about 900 tenth grade students, another with about 900 seventh grade students, and still another with about 1800 fourth grade students. These cohorts are referred to in this report as cohorts A, B, and C. In 1996, about 2,000 additional fourth grade students were enrolled in the study. This cohort is referred to as cohort D. In each case, students who continued to reside in the twelve communities and attend participating study schools were evaluated annually through high school graduation. Students in cohort D will graduate from high school in June 2004. For the purposes of this report, we will be reporting on data collected through high school graduation on cohorts A, B, and C. We will also report on four years of follow-up data on cohort D.

3.3.2. School Selection

We contacted potential schools to seek participation in the study. Criteria for selection required (1) location in a pre-selected community of interest based on air pollution levels and patterns; (2) sufficient population of target-aged children; (3) preponderance of children attending school from the immediate neighborhood; (4) demographic similarity with other potential and participating community school sites; (5) reasonably stable communities in terms of residential migration, to improve chances for longitudinal follow-up with participants; (6) absence of localized air pollution sources; and (7) proximal location to a fixed-site air monitoring station.

The design approach specified child entry into the study at the fourth, seventh, and tenth grades (nominally in the following age groups: 9-10 years, 12-13 years, and 15-16 years, respectively) and required the enrollment of at least four schools in each community (two elementary schools, a junior high school, and a senior high school), linked by student promotion, so that the selected elementary school would feed students into the participating junior high or middle school, and so on. Within each community, the study population consisted of about 150 fourth-graders, 75 seventh-graders, and 75 tenth-graders.

We contacted the appropriate school district superintendents to establish cooperative understandings. When appropriate, we met with the district personnel to present an overview of the study, including the rationale, the goals, and the envisioned level of effort required by participating schools and students for project success. Early in the site negotiations process, we reviewed and visited schools being considered for study participation along with district personnel to explain the program to the local administration and obtain school-based administrative support for the study. When insurmountable difficulties were identified, alternative sites were identified by school district personnel.

Following school district approval for participation, we scheduled meetings with the on-site administrators and potentially affected teachers to discuss the project and obtain support for its performance. Scheduling for administration of the annual questionnaire, lung function testing, and exposure monitoring was done to minimize the disruption of school activities and to facilitate efficient planning of field operations.

3.3.3. Student Selection and Recruitment

Complete classrooms of students were recruited. The cooperation of the student and the granting of informed parental (or legal guardian) consent for participation by minors were required. Recruitment consisted of presenting a brief overview of the study to potential classrooms. The presentation emphasized the personal health nature of the study and the opportunity to learn about individual health and the lungs. To minimize bias, the presentation downplayed the aspect of air pollution. The study procedures and protocols were explained to enable children to gain insight into the level of participation that would be required (completion of questionnaires once per year, lung function testing once per year and absence monitoring with call back of parents).

3.3.4. Student Enrollment

The health history questionnaires were first circulated during early 1993, the enrollment point for subject entry into the study. They were completed by a combination of parents and children for the older subjects and by parents for the younger subjects. Enrollment was contingent upon receipt of the written informed consent from the child's parent or legal guardian (whose request for signature appears on the first page of the questionnaire) and upon return of the health questionnaire. A code number uniquely identified each student and his/her personal information remained confidential, divulged and identified only by unreferenced subject number for analytical purposes.

Table 3.3-1 presents data on demographic characteristics of the enrolled subjects in cohorts A, B, and C, as reported by parents, by community (analogous characteristics of cohort D are provided in section 4.2.4). Whites were a plurality in all communities, and a majority in all except Long Beach, Riverside, and Santa Maria. Hispanics (who might be of any race, but usually placed themselves in the white or "other" categories) were a majority in Santa Maria. Sizable (> 10%) minorities of a specific race were present in Riverside (African Americans) and Long Beach (African and Asian Americans). Family income was higher in Upland, and lower in Santa Maria. A large majority of subjects (ranging from 72 to 95% across communities) lived in single-family houses. Slightly more than half of the children were female.

Table 3.3-1 Demographic characteristics of the enrolled students in cohorts.

Community	Eligible Subjects	Subjects with Baseline Questionnaire Information (%)		White	Black	Asian	Other	Hispanic	Male	Income > $50,000
Alpine	396	298	(75)	84.0	0.4	0.8	14.8	12.8	49.8	37.5
Atascadero	371	260	(70)	84.1	0.4	0.4	15.1	11.4	40.5	36.3
Lake Elsinore	397	316	(80)	76.9	2.3	1.7	19.1	23.8	53.1	25.6
Lake Arrowhead	402	347	(86)	83.8	0.6	1.5	14.1	16.2	48.5	36.3
Lancaster	350	266	(76)	70.4	5.8	2.5	21.3	26.8	45.7	29.4
Lompoc	410	305	(74)	72.5	8.7	0.8	18.0	19.3	50.0	32.6
Long Beach	414	325	(79)	37.9	16.1	21.8	24.2	22.3	47.9	31.1
Mira Loma	438	308	(70)	66.8	1.1	1.4	30.7	34.0	46.3	29.4
Riverside	469	369	(79)	43.9	14.0	6.4	35.7	38.4	47.1	21.4
San Dimas	397	303	(76)	61.6	5.8	8.8	23.8	29.7	47.8	34.3
Santa Maria	371	300	(81)	46.3	1.6	2.8	49.3	60.1	48.1	12.9
Upland	428	279	(65)	69.4	2.6	8.7	19.3	16.7	49.6	65.6
Total	4,843	3,676	(76)	66.1%	5.1%	4.9%	23.9%	26.0%	47.9%	32.2%

(Source: (Peters et al. 1999a))

3.4. Exposure Assessment Methodology

The approach for characterization of exposure to air pollution of ambient origin incorporated a prospective ambient air monitoring program in the twelve communities, a retrospective analysis of pre-CHS study exposures, and a modeling analysis of personal exposure that included exposure to traffic-related pollutants. The methodologies for each of these elements of the approach are described in this section.

3.4.1. Ambient Air Quality Measurement Methods

3.4.1.1. Network Objectives and Operations

The objective of the ambient air monitoring program was to obtain seasonal and annual average concentrations for all of the CHS air pollutants at a single representative location in each community for use in the analysis of chronic air pollution health effects. An additional objective was to obtain hourly and daily data for a subset of pollutants that could be used in assessment of potential chronic effects resulting from repeated short-term exposures. Continuous hourly measurements of ozone, NO_2, and PM_{10} were made at the central CHS air monitoring station in each community. Integrated measurements of $PM_{2.5}$ mass, PM chemical constituents, and gas-phase acids were made using a multi-legged two-week sampler (TWS) designed for the study (Lurmann et al. 1994). The PM chemical constituents included $PM_{2.5}$ sulfate, nitrate, and ammonium and PM_{10} elemental carbon (EC) and organic carbon (OC). The filters collected for carbon analysis were initially collected without a specified size-cut, and later using dual sampling legs with a $PM_{2.2}$ size cut and the undifferentiated leg (for comparative purposes). The TWS also collected samples for subsequent laboratory determination of two-week average concentrations of nitric acid, hydrochloric acid, formic acid, and acetic acid (collectively identified as acid vapor). These measurements were made throughout the study period, 1994-2001. Additional measurements of carbon monoxide (CO), particle number (PN), $PM_{2.2}$ EC, $PM_{2.2}$ OC, and $PM_{2.2}$ elements by x-ray fluorescence (XRF) were implemented in most communities in 2000 and 2001. Also in 2001, the measured nitric oxide (NO) concentrations were retrospectively added to the database. Hourly temperature and relative humidity were measured at some of the CHS air monitoring stations for some of the years to complement the air quality data. These data were supplemented with meteorological data collected at locations near the CHS communities.

The CHS air quality monitoring network was established by augmenting existing stations and creating new stations in late 1993 and early 1994. Seven of the air monitoring stations were existing sites where air pollution control agencies monitored ozone, NO_2, or PM_{10}. The existing sites included Atascadero, Santa Maria, North Long Beach, Lancaster, Upland, Lake Elsinore, and Alpine. Five new air monitoring sites were established for the study. These new sites were located at the U.C. Riverside Agricultural Station in Riverside, Jurupa Valley High School in Mira Loma, Rim of the World High School in Lake Arrowhead, Cabrillo High School in Lompoc, and Gladstone Elementary School in San Dimas. All the station locations met EPA siting requirements. Several stations were relocated during the study. The San Dimas measurement station at Gladstone School was shut down in 1996, and the CHS air monitoring equipment was moved to the nearest district-operated regulatory air monitoring station,

Glendora, which is approximately 4 km from the original San Dimas sampling site. Samplers and instruments were operated concurrently for 12 months to characterize typical air quality differences between the San Dimas and Glendora locations. Several important and consistent differences were observed in data collected at the two sites (and these are discussed in Sections 3.4.1.3 and 4.1.1.1.1 through 4.1.1.4.4).

Secondary pollutant levels were similar at the two monitoring locations; however, primary pollutant levels, such as NO, were much lower at Glendora than San Dimas due primarily to differences in proximity to traffic. In 2000, the ARB moved the Santa Maria air monitoring station 0.4 km southwest of the original location. In 2001, the Antelope Valley Air Quality Management District moved the Lancaster station 2.3 km south of its original location. Concurrent sampling data were not collected when the Santa Maria and Lancaster stations were moved. However, both the new and old sampling locations in Santa Maria and Lancaster were similar with regard to roadway proximity, traffic density, spatial topology, and general land use (in terms of potential local sources), so differences in measured ambient air concentrations between the old and new sampling locations were likely to be small.

Responsibility for field operations and data management evolved over the study period. Initially, the field operations were divided among Sonoma Technology, Inc. (STI), ARB, San Luis Obispo County Air Pollution Control District's (SLOAPCD), South Coast Air Quality Management District's (SCAQMD) and San Diego County Air Pollution Control District (SDCAPCD). In 1995, ARB contracted with the local air pollution agencies and RM Environmental Inc. for field operations of the stations through 2001. The air quality data flowed from the field operations group to the exposure data manager, which was STI in 1993-1994 and ARB in 1995-2001. Filters and coated-denuders were handled by the Environmental Health Service at the Los Amigos Research and Education Institute (LAREI). Gravimetric analyses of $PM_{2.5}$ mass and ion chromatography for determination of sulfate, nitrate, ammonium, formate, and acetate were conducted in the LAREI laboratory. The thermal/optical transmittance laboratory analyses of EC and OC were performed at the California Institute of Technology. The XRF elemental analysis of $PM_{2.2}$ was conducted at the Desert Research Institute (DRI). The exposure data manager delivered all the data to STI for quality assurance and data validation prior to delivery of the data to USC. All measurements and data processing were conducted in accordance with documented quality assurance procedures (Bowers and Taylor 2000).

3.4.1.2. Instrumentation

The CHS air monitoring equipment was state-of-the-art at the time the network was established. Ozone was monitored hourly using EPA-approved UV photometric instruments. NO_2 was determined hourly from EPA-approved chemiluminescent instruments measuring NO_x and NO, and NO_2 was calculated as the difference ($NO_2 = NO_x - NO$). Automated daily calibration systems were employed with all of the ozone and NO_x instruments. Hourly PM_{10} mass was measured using the tapered element oscillating microbalance (TEOM) instrument (Patashnick and Rupprecht 1991). The standard operating procedure for the TEOMs included heating the incoming air to 50^{o}C in order to evaporate water from the aerosol prior to collection. The PM_{10} concentrations measured by TEOMs were subsequently adjusted to account for the evaporation of ammonium nitrate and organic compounds. Additional instruments were installed in 2000 and 2001 to obtain data for directly emitted pollutants. Carbon monoxide analyzers were installed at

all of the stations except Alpine, and condensation particle counters (CPC, TSI Model 3022A) were installed at all of the stations to assess ambient ultrafine particle number concentration.

A new aerosol/acid sampler (the two-week sampler, or TWS) was developed to collect integrated samples of $PM_{2.5}$ mass, $PM_{2.5}$ chemical components, and selected gas-phase acids. This development effort was needed because the existing aerosol/acid technologies were designed for 4- to 24-hr sampling intervals, and the use of short-interval sampling was prohibitively expensive in a long-term study. Thus, a major effort was undertaken to develop an acceptable two-week aerosol/acid sampler. Details of the sampler development and evaluation effort are provided in Lurmann et al. (1994) and Hering et al. (1994). The TWS was initially designed with three sampling legs (A, B, and C) and subsequently modified in 2000 to incorporate two additional sampling legs (D and E). The standard operating procedures for the new sampler are described in Taylor et al. (2003).

Leg A consists of a single-jet Teflon impactor that removes particles with diameters larger than 2.5 μm, followed by a carbonate-coated glass honeycomb denuder that collects nitric and hydrochloric acids, followed by a Teflon filter and a carbonate-impregnated quartz filter for fine-particle mass and ions. Ion chromatography is used to measure nitrate and chloride in denuder extraction which determines the nitric acid and hydrochloric acid concentrations. The Teflon filter is analyzed gravimetrically for $PM_{2.5}$ mass and by ion chromatography for $SO_4^{=}$, NO_3^{-}, and NH_4^{+}. The Na_2CO_3 coated quartz backup filter is used to collect volatilized particulate nitrate, and it is also analyzed by ion chromatography. The flow rate is 0.4 L/min. $PM_{2.5}$ mass is determined from the sum of the mass measured on the Teflon front filter and the mass of ammonium nitrate determined by nitrate measurements on the backup filter. Total nitrate ion concentration is determined as the sum of front and backup filter nitrate. The total ammonium ion concentration is the sum of the front filter ammonium and the ammonium presumed to be associated with the measured backup filter nitrate (i.e., 18/62 of the backup filter nitrate).

Leg B utilizes the filter pack methodology of Solomon et al. (1988) for determination of two gas-phase organic acids. It has a Teflon pre-filter to remove particles, followed by two potassium hydroxide-impregnated quartz fiber filters. The material collected on the filters is extracted and analyzed by ion chromatography for formate and acetate to give formic acid and acetic acid concentrations. The acetic acid is an upper-limit value because this measurement method is known to have positive interference from peroxycetylnitrate (PAN) (Grosjean and Parmar 1990). Sample flow rate is 0.4 L/min.

Leg C contains a baked quartz fiber filter from which OC and EC are determined by thermal/optical reflectance laboratory analyses. The flow rate is 1.3 L/min. It is comprised of an inlet line (6 mm diameter by 2.5 cm long) and a Teflon filter holder (Savilex). It does not contain a size-selective inlet, yet Salmon et al. (2001) demonstrated that Leg C of the sampler collects an effective PM_{10} sample. Comparison testing was also performed to establish an effective equivalent collection value for a 2.5 micron diameter size cut (Glen Cass, private communication). The equivalent comparisons were determined to be:

$PM_{2.5}$ EC = (0.98)(Leg C OC), and $PM_{2.5}$ OC = (0.78)(Leg C OC).

For all exposure and health analyses subsequently reported in this CHS final report, the effect size cut for EC and OC was a $PM_{2.5}$ cut.

The laboratory analysis procedure of the collected samples is similar to the NIOSH method (NIOSH 1996) and is described by Birch and Carey (1996).

Leg D is identical to Leg C except it incorporates a size-selective inlet (Mingchih et al. 1999). It consists of an aluminum $PM_{2.2}$ impactor followed by a baked quartz fiber filter from which $PM_{2.2}$ OC and EC are determined by thermal/optical reflectance laboratory analyses. The flow rate is 1.3 L/min. The impactor has a conical cavity with a glass fiber substrate that is coated with two drops of mineral oil to collect large particles (Mallinckrodt N.F. white mineral oil #6358). Taylor (2000) demonstrated that the mineral oil does not interfere with the OC/EC measurements.

Leg E contains a $PM_{2.2}$ impactor followed by a Teflon filter from which elements from aluminum to uranium are determined using x-ray fluorescence (XRF) analyses. The flow rate is 1.3 L/min. The laboratory analysis method is described by Watson et al. (1999).

3.4.1.3. Concentration Adjustments

PM_{10} concentrations measured using TEOMs with heated inlets are biased low due to the volatilization of ammonium nitrate and certain organic compounds (Allen et al. 1997). To account for the loss of semi-volatile species in the heated TEOM inlets, regression relationships were developed relating the 24-hr average PM_{10} concentrations measured by TEOM to similar HiVol sampler-based (the Federal Reference Method-FRM-for PM_{10}) measurements. Collocated measurements of PM_{10} concentrations using the TEOM and HiVol samplers were available at four locations in the study area: Rubidoux, Lancaster, Atascadero, and Long Beach. The results are summarized in Table 3.4-1. The TEOM adjustment factor was highest (1.25) at Rubidoux, located east of Los Angeles between Mira Loma and Riverside. This adjustment factor was used for all of the low elevation inland sites in the South Coast Air Basin (SoCAB), including San Dimas, Glendora, Upland, Mira Loma, and Riverside. The TEOM adjustment factor (1.02) derived from the Atascadero data was used for low PM_{10} concentration sites, including Santa Maria and Lompoc. The TEOM adjustment factor (1.11) derived from the Lancaster data was used for the other low PM_{10} concentration sites, including Alpine and Lake Arrowhead. The TEOM concentrations at Lake Elsinore were adjusted using the Long Beach adjustment factor (1.18).

Table 3.4-1. Relationship between daily HiVol Sampler and TEOM PM_{10} concentrations used to determine PM_{10} adjustment factors.

Sites with Collocated Samplers	Adjustment Factor = Regression Slope (α)[a]	Coefficient of Determination (r^2)	Other CHS Communities For Which the Adjustment Factor Was Applied
Rubidoux	1.25	0.69	Mira Loma San Dimas/Glendora UC Riverside Upland
Lancaster	1.11	0.63	Alpine Lake Arrowhead
Atascadero	1.02	0.89	Lompoc Santa Maria
Long Beach	1.18	0.67	Lake Elsinore

[a] [Adjusted PM_{10}] = α [TEOM PM_{10}]

After the San Dimas air monitoring station was closed (1996), concentration estimates for the San Dimas community were estimated from measurements obtained at the Glendora station (on the outskirts of Glendora and about 3 km from the freeway). A comparison of one to three years of concurrent monitoring at San Dimas and Glendora indicated that all the pollutant concentrations were highly correlated, and significant bias (>10%) existed for NO, NO_2, nighttime ozone, $PM_{2.5}$ EC and $PM_{2.5}$ nitrate. The average EC and nitrate concentrations were 29% and 19% higher, respectively, at San Dimas than Glendora. Average differences in $PM_{2.5}$ mass and ammonium were less than 10%, so no adjustments were made. The largest bias was in NO concentrations which were 100% higher on a 24-hr average basis and ~500% higher for the 1-hr morning maximum on average at San Dimas than at Glendora. Much smaller differences were found for NO_2, which was higher at San Dimas than at Glendora, and for nighttime ozone which was lower at San Dimas than at Glendora. Regressions for each hour of the day were used to estimate San Dimas NO_2 concentrations from measurements at Glendora and Azusa, and to estimate San Dimas ozone concentrations from measurements at Glendora.

3.4.1.4. Two-week Sampler Accuracy and Precision

The precision and accuracy of TWS prototypes were determined by comparison of two-week average results with 14 daily samples collected with the SCAQS samplers (Lurmann et al. 1994). In-use accuracy was determined by comparisons with collocated $PM_{2.5}$ FRM data. Comparison of $PM_{2.5}$ mass data for 19 two-week periods in 2000 and 2001 in Long Beach, where collocated FRM sampler data were available, indicates the TWS $PM_{2.5}$ mass measurements are within ± 2.7 $\mu g/m^3$ and ±15% of the FRM measurements on average, and are biased by -2.3 $\mu g/m^3$ and -12.5% on average. Note that, for some periods, the two-week averages of FRM measurements were determined from 13 days rather than 14 days of data, which may contribute to the bias. Figure 3.4-1 shows the measurements are well-correlated (r^2 = 0.80), but biased low presumably due to loss of volatile compounds over the two-week sampling period. The comparison with

FRM data is comparable to those with the SCAQS samplers collected in 1993. Motallebi et al. (2003) found similar results when they compared 2-week average TWS mass concentrations with the mean of ten daily FRM measurements made within 2-week periods at Long Beach in 1999 and 2000.

The in-use precision was determined from 13 site-years of collocated data. Two extra samplers were operated at Riverside, Mira Loma, Lake Elsinore, and/or Alpine between 1994 and 2001. The 1994 data were excluded from the precision analysis because the designation of primary and secondary sampler was unclear. Table 3.4-2 and Figure 3.4-2 show the comparison between the primary and secondary samples collected at all locations in 1995-2001. Based on the data from all sites that exceeded the detection limits, the mean absolute differences in collocated $PM_{2.5}$ mass, nitrate, sulfate, and ammonium concentrations were ±2.0, ±0.7, ±0.2, and ±0.3 $\mu g/m^3$ and ±11%, ±10%, ±9%, and ±13%, respectively. The mean absolute differences in collocated acetic acid, formic acid, nitric acid, and hydrochloric acid were ±0.30, ±0.14, ±0.36, and ±0.16 ppb and ±9%, ±12%, ±12%, and ±20%, respectively. The precision is best determined by the pooled coefficients of variation. The TWS pooled coefficients of variations were 10%, 11%, 11%, and 15% for $PM_{2.5}$ mass, nitrate, sulfate, and ammonium, respectively, and 8%, 12%, 13%, and 22% for acetic acid, formic acid, nitric acid, and hydrochloric acid, respectively. All the TWS measurements, except those for the hydrochloric acid, met the original precision goal of ±15%. The ambient concentrations of hydrochloric acid were quite low and variable. For this reason, discussions of ambient acid vapor levels in this report typically refer to the sum of nitric, formic, and acetic acids only.

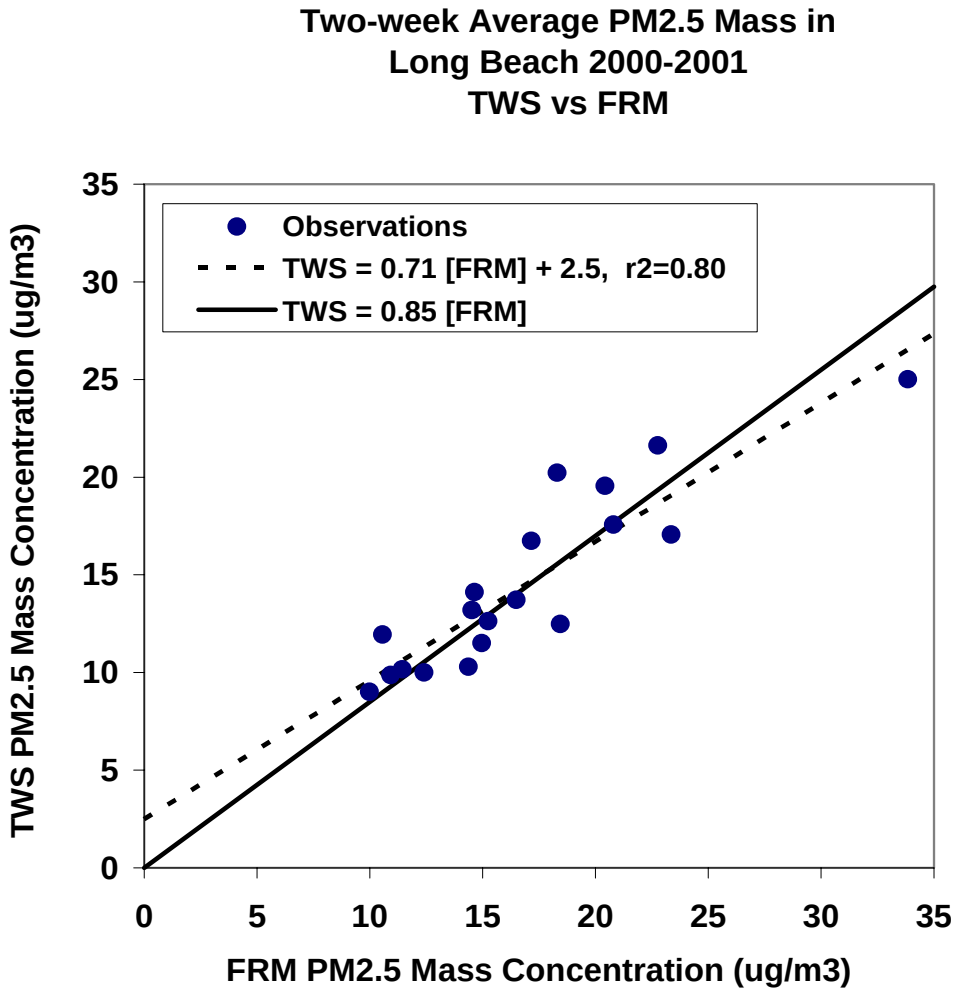


Figure 3.4-1. Comparison of $PM_{2.5}$ mass measurements collected by the two-week sampler and FRM sampler in Long Beach in 2000 and 2001.

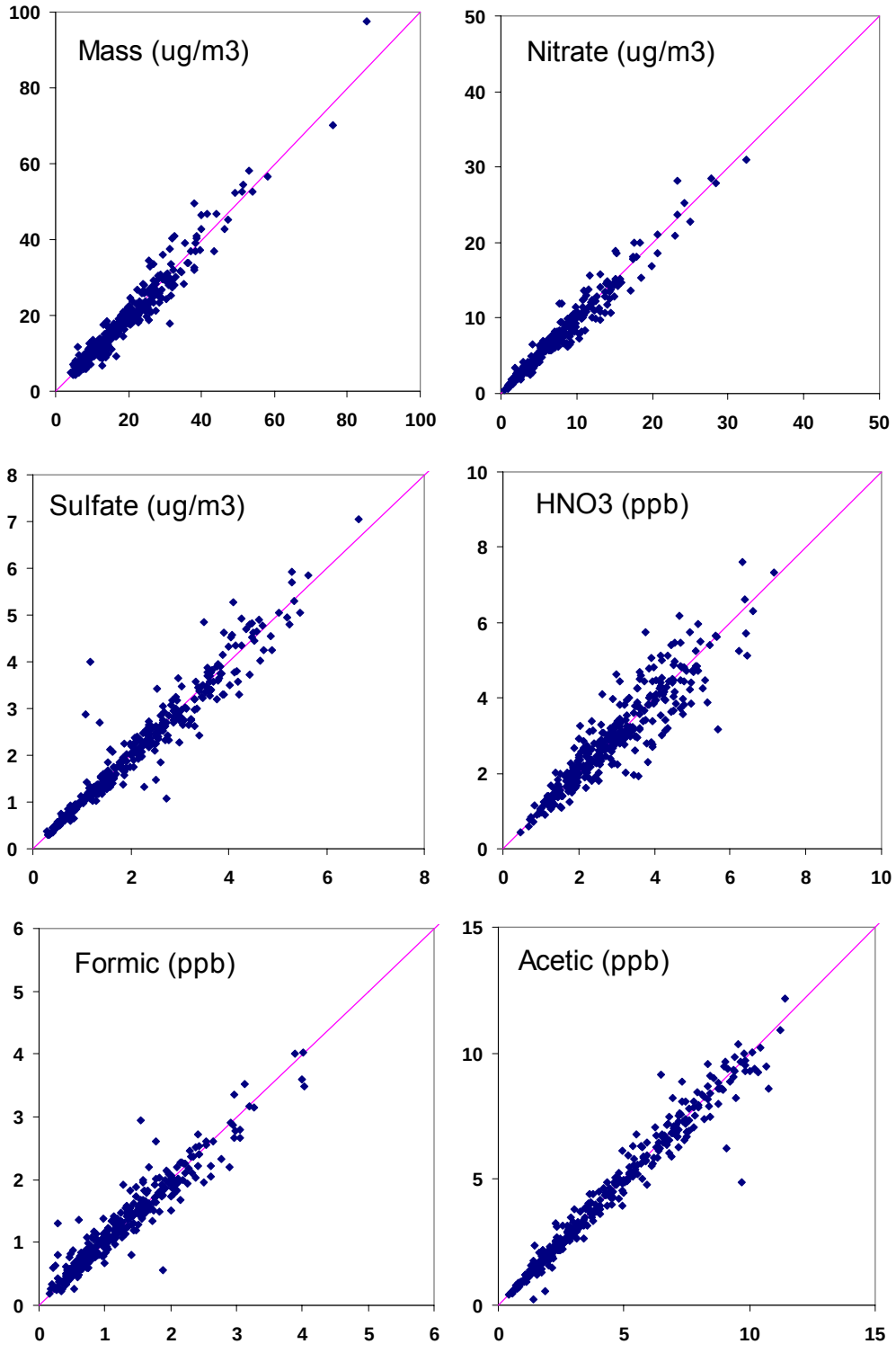


Figure 3.4-2. Comparison of collocated two-week sampler concentrations of $PM_{2.5}$ mass, $PM_{2.5}$ nitrate, $PM_{2.5}$ sulfate, nitric acid, formic acid, and acetic acid from 1995-2001.

Table 3.4-2. Comparison of collocated two-week sampler concentrations of $PM_{2.5}$ mass, $PM_{2.5}$ nitrate, $PM_{2.5}$ sulfate, nitric acid, formic acid, and acetic acid collected 1995-2001.

Species	Mass ($\mu g/m^3$)	Nitrate ($\mu g/m^3$)	Sulfate ($\mu g/m^3$)	Ammonium ($\mu g/m^3$)
Number of Samples	336	352	345	352
Primary Sampler Mean Concentration	19.80	7.22	2.17	2.70
Collocated Sampler Mean Concentration	19.56	7.16	2.16	2.68
Mean Difference	0.24	0.06	0.01	0.02
Mean Normalized Difference (%)	2.81	1.40	1.54	4.30
Mean Absolute Difference	1.97	0.71	0.20	0.28
Mean Normalized Absolute Difference (%)	11.39	10.45	9.06	12.59
Pooled Standard Deviation	1.97	0.80	0.24	0.40
Pooled Coefficient of Variation (%)	10.03	11.19	11.15	14.92

Species	Nitric Acid (ppb)	Hydrochloric Acid (ppb)	Formic Acid (ppb)	Acetic Acid (ppb)
Number of Samples	359	46	355	366
Primary Sampler Mean Concentration	2.91	0.81	1.26	4.40
Collocated Sampler Mean Concentration	2.91	0.77	1.27	4.37
Mean Difference	0.00	0.03	-0.01	0.03
Mean Normalized Difference (%)	1.41	7.71	-0.75	2.86
Mean Absolute Difference	0.36	0.16	0.14	0.30
Mean Normalized Absolute Difference (%)	12.44	20.35	12.15	9.02
Pooled Standard Deviation	0.37	0.17	0.15	0.37
Pooled Coefficient of Variation (%)	12.67	21.87	12.16	8.34

3.4.1.5. Metrics of Exposure to Ambient Air Pollution

A variety of metrics were selected at the beginning of the study for characterization of exposure to ambient air pollution. The parameters listed in Table 3.4-3 were determined on a daily, monthly, and annual basis from the hourly data. Monthly and annual average concentrations were determined for $PM_{2.5}$ and acids from the two-week sampler integrated measurements. A 75% completeness criterion was applied at each level to determine averages. For example, 18 or more hours of valid data were needed to determine a valid 24-hr average concentration; 22 or 23 valid daily averages were needed to determine a valid monthly average; and 9 or more monthly averages were needed to determine a valid annual average.

Because completeness of the annual average pollution data is critically important for this assessment of chronic health effects, two methods were used to determine annual average values. The Method I annual average values for a specific year were computed only from the valid monthly average values for that year. Method II annual averages were based on the valid monthly values for the specific year and, where valid monthly data were missing, monthly estimates obtained from observations for that location and month of the two surrounding years. For example, if the monthly average concentration for Riverside in June 1998 was missing, the Method II annual average was computing using an estimate for June 1998 that was derived by averaging the Riverside data for June 1997 and June 1999. Method II averages were constructed because the seasonal variations in concentrations of some pollutants are quite large and annual averages determined with data or estimates for every month of the years were less likely to be biased than Method I estimates that were missing data for the three lowest concentration months or the three highest concentration months. Method II annual average values were used for health effects assessment useless otherwise noted.

For selected analyses, monthly and annual average coarse PM ($PM_{10-2.5}$) concentrations were estimated by subtracting the TWS $PM_{2.5}$ mass from the PM_{10} mass determined from the adjusted TEOM measurements. The coarse PM estimates have larger relative uncertainties than either the $PM_{2.5}$ or PM_{10} mass concentrations.

Table 3.4-3. Daily, monthly, and annual exposure parameters determined from continuous measurements.

Parameter	Ozone	PM_{10}	NO_2	NO	CO	CPC
24-hr average concentration	X	X	X	X	X	X
Daily maximum 1-hr concentration	X	X	X	X	X	X
Daily maximum 8-hr average	X			X	X	X
10 AM-6 PM PST average concentration	X	X				
6 AM -6 PM PST average concentration			X			
Sum of hourly concentrations greater than Threshold #1	>30 ppb	>30 $\mu g/m^3$	>15 ppb			

Sum of hourly concentrations greater than Threshold #2	>60 ppb	>60 μg/m^3	>30 ppb			
Sum of hourly concentrations greater than Threshold #3	>90 ppb	>90 μg/m^3	>45 ppb			
Sum of hourly concentrations greater than Threshold #4	>120 ppb	>120 μg/m^3	>60 ppb			
Sum of hourly concentrations greater than Threshold #5	>150 ppb	>150 μg/m^3	>75 ppb			
Number of daily 1-hr NAAQS exceedances	X					

3.4.2. Pre-study Exposure Assessment Methods

In order to examine the relationships between health status and cumulative exposure to air pollution, estimates of the population's exposure to key air pollutants from birth to the time of enrollment in the study were calculated. The database was based on EPA's Aerometric Information Retrieval System (AIRS) ambient air quality data collected in the United States from 1975 through 1995. Specifically, the database contained monthly average values of the CHS exposure metrics (see Table 3.4-3) for ozone, NO_2, and PM_{10} for each location that could be satisfactorily geocoded. Both the AIRS station locations and air quality data were subjected to limited quality control checks for acceptance into the database (no duplicates, no unreasonably high or persistent values, etc.).

3.4.2.1. Geocoding Residence Locations

To assign lifetime or historical exposures for individual subjects, CHS residential histories were linked to available exposure monitoring data. The residential history data indicated that CHS subjects lived in 46 states and more than 40 foreign countries prior to enrolling in the study. No attempt was made to assign historical exposures for foreign locations due to a lack of data. Historical residence locations were geocoded with city, state, and ZIP code information. Street addresses were not collected so traditional geocoding of exact residential locations was not feasible. If valid and consistent ZIP codes and city names were available, the locations were assigned the coordinates of the geographic center of the ZIP code region—or ZIP code centriod (ZCC). If only city and state were available and it was a relatively small city, the centroid of the geographic city boundary was assigned. This was referred to as the city centroid (CC) method. If only the city and state were available and the city had a population greater than 500,000, the ZCC with the highest population was assigned. This is referred to as the ZIP Code Weighted Centroid Coordinates (ZWCC). In cases where the city name and ZIP code were inconsistent, the geocoding was based on the city name because it was assumed to be more accurate than the ZIP code. Some historical locations were military bases rather than cities. The ZIP code boundaries containing the largest area of military base land were used to assign geographic locations to study participants. Lastly, the coordinates of the CC were assigned in cases where the ZIP code coordinates could not be found. Crestline, Lake Arrowhead, and Lake Gregory residences fell into this category and most residences in these areas could not be located with sufficient accuracy to include them in the traffic and exposure modeling. No assignments were made for about 2% of the cases because of missing city names and invalid ZIP codes; however, 98% of domestic residences were successfully geocoded.

3.4.2.2. Pre-study Air Quality Database

The spatial coverage of the historical air quality monitoring networks is shown in Figure 3.4-3 through Figure 3.4-6. Ozone and total suspended PM (TSP) have the most extensive spatial coverage, while NO_2 has the least extensive spatial coverage. The TSP monitoring network was dense by 1975 and did not change significantly over the period shown. The spatial coverage for ozone, NO_2, and PM_{10} improved during the period. PM_{10} data are not available prior to 1982, and only a few sites reported PM_{10} data before 1988. PM_{10} concentrations for 1975 through 1987 are estimated from the more widely available TSP data. Collocated PM_{10} and TSP data from 1988 through 1992 were used to create PM_{10}-to-TSP ratios and regression relationships by site. The 24-hr average PM_{10} and TSP are correlated ($0.51 < r^2 < 0.88$ in different states that have at least two years of collocated data). The mean PM_{10}-to-TSP ratios for each station with more than 20 daily samples are shown in Figure 3.4-7 and Table 3.4-4. There are station-to-station and regional variations, with the Midwestern region having the lowest ratios on average (0.48 ±0.08) and the Southeastern region having the highest ratios (0.63 ±0.08), compared to the national mean of 0.54 ±0.10. The variance in the mean PM_{10}-to-TSP ratios within regions is modest (coefficient of variation is less than 18%). PM_{10} concentrations prior to 1988 are estimated using the TSP observations and region-specific PM_{10}-to-TSP ratios (or the national ratio for sites in Hawaii and Alaska).

The infrequency of PM sampling is a concern regarding historical particle exposures. While the monthly ozone and NO_2 exposure metrics are based on hourly measurements and are invalidated unless there are 24 or more days per month with valid measurements, the TSP and PM_{10} data are based on 24-hr samples collected every sixth day. Monthly TSP and PM_{10} means are usually based on only 4 or 5 daily measurements per month and are, therefore, inherently less robust than the metrics derived from continuous measurements.

Another concern is that NO_2 data collected with early chemiluminescence monitors are less accurate than more recent data. The ARB's recommendation to reduce the archived 1975 through 1980 NO_2 concentrations by 12% to account for errors in converter efficiency calculations was implemented in this data set (California Air Resources Board 1979).

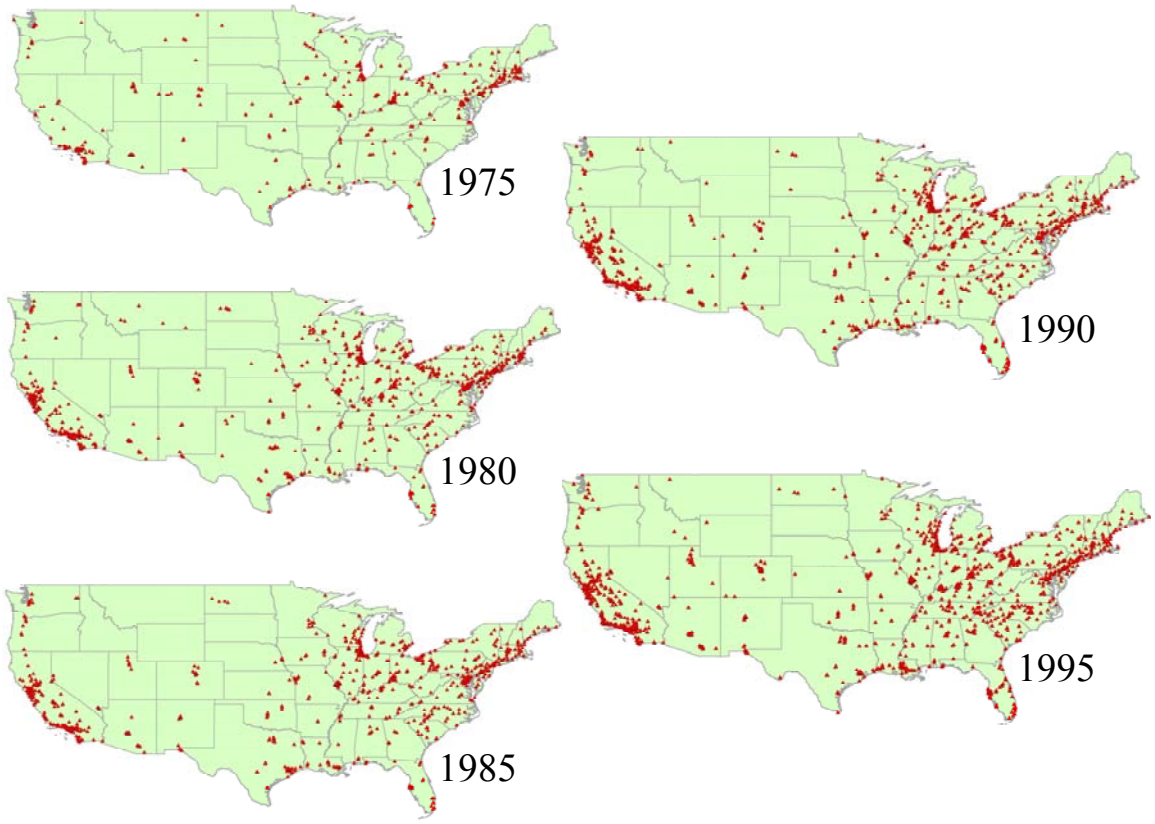


Figure 3.4-3. The ozone monitoring locations in 1975, 1980, 1985, 1990, 1995.

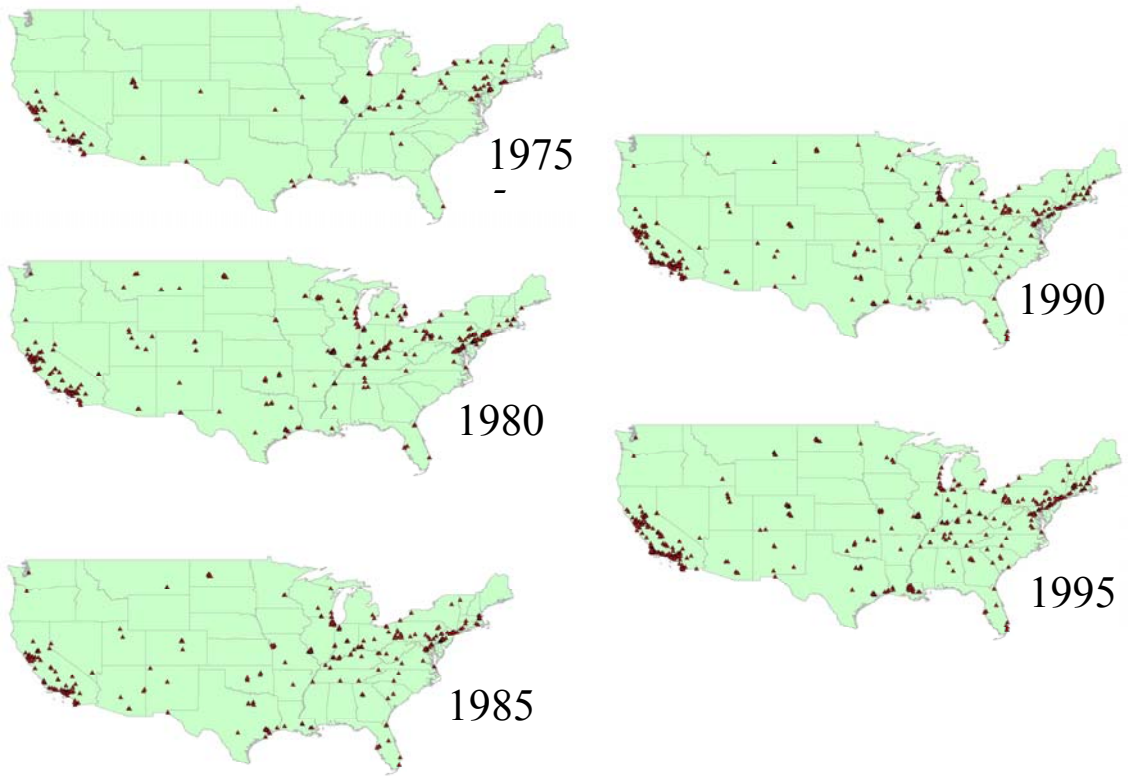


Figure 3.4-4. The NO_2 monitoring locations in 1975, 1980, 1985, 1990, 1995.

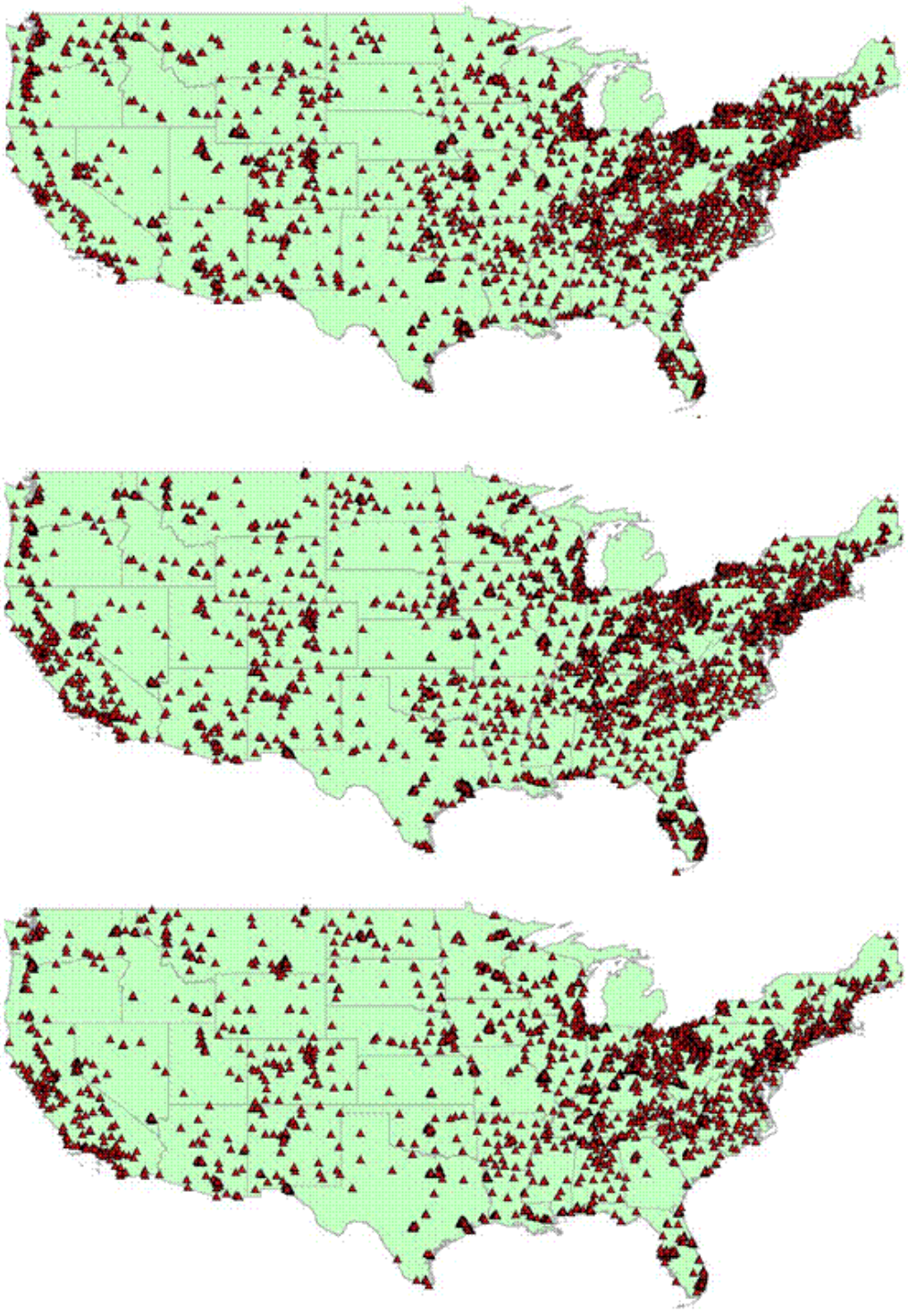


Figure 3.4-5. The TSP monitoring locations in 1975 (top), 1980 (middle), and 1985 (bottom).

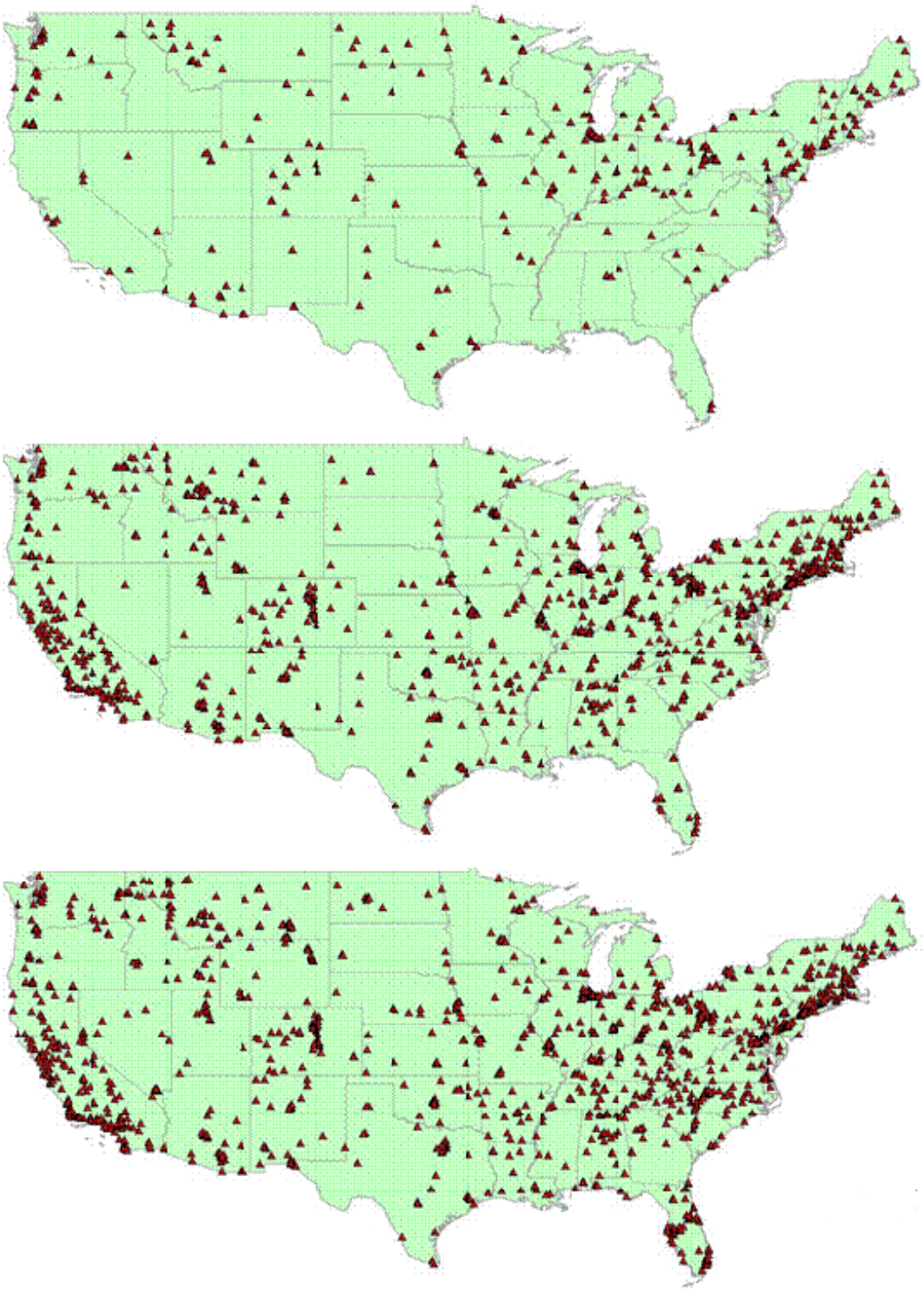


Figure 3.4-6. The PM_{10} monitoring locations in 1985 (top), 1990 (middle), and 1995 (bottom).

Table 3.4-4. Mean PM_{10}-to-TSP ratios in different regions of the United States.

Region	Region Code	Average Ratio	Standard Deviation	Number of Stations
Midwest	3	0.48	0.085	155
Northeast	4	0.54	0.092	220
Northwest	1	0.60	0.075	40
Southern California	2	0.53	0.068	44
Southeast	5	0.63	0.084	99
Entire United States	7	0.54	0.100	570

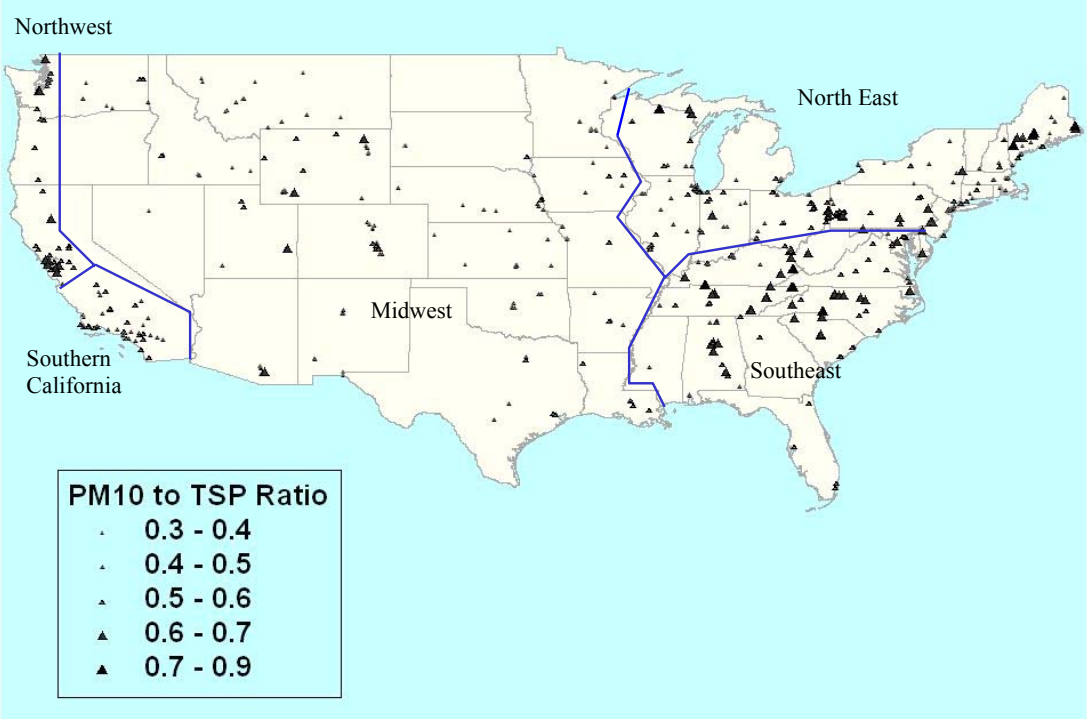


Figure 3.4-7. Average ratios of PM_{10} to TSP concentrations in various regions of the country.

3.4.2.3. Spatial Mapping

The monthly concentrations of pollutions at specific historical locations were estimated using spatial interpolation. An inverse distance weighted (IWD) method was used and the quality of each interpolated value was ranked according to the following definitions:

- Highest quality spatial interpolation (Code 1). Spatial interpolations are based on inverse distance weighting of data from the three closest stations located within 5 km of the point of interest (residence location). This method is used whenever one or more stations located within 5 km of the point of interest have valid data.
- Second highest quality spatial interpolation (Code 2). Spatial interpolations are based on inverse distance weighting of data from the three closest stations, with the closest station located between 5 km and 25 km of the point of interest (residence location). This method is used when there are no stations within 5 km, and there are one or more stations located within 25 km of the point of interest with valid data.
- Lowest quality spatial interpolation (Code 3). Spatial interpolations are based on inverse distance weighting of data from the three closest stations located between 25 km and the maximum interpolation radius of the point of interest (residence location). This method is used when there are no stations within 25 km, and one or more stations are located within the maximum interpolation radius km of the point of interest with valid data.
- Spatial interpolations not feasible (Code 9). No monitoring stations with valid data are located within the maximum interpolation radius of the point of interest (residence location).

Ozone data were spatially interpolated using 100-km maximum interpolation radii; NO_2 and PM_{10} data were both spatially interpolated using 50 km.

The spatial mapping errors were evaluated by comparing monthly observations with estimates spatially interpolated to each specific measurement station location without using the data from the specific measurement station. Table 3.4-5 shows the results for 1-hr daily maximum, 10 a.m.-6 p.m. average, 8-hr daily maximum, and 24-hr average ozone metrics for all years and states combined. The evaluation indicates there is little bias in the ozone estimates (less than 1 ppb and 2%) and the errors average 13% to 21%. The coefficients of determination (r^2) are 0.73-0.87 for the 1-hr maximum, 0.69 to 0.84 for both 8-hr metrics, and 0.58 to 0.68 for the 24-hr average ozone, depending on the quality ranking of the interpolations. The spatial mapping errors were evaluated using 25-km, 50-km, and 100-km maximum interpolation radii and found to be very similar. Spatial maps of the observed and estimated concentrations were prepared and reviewed by subregion. The spatial coverage with interpolations made with a 25-km maximum radius was sparse compared to ones made with 50-km and 100-km radii. Based on the review, the interpolations made using the 100-km maximum radius were adopted.

The spatial mapping errors for NO_2 and PM_{10} were evaluated using 25-km and 50-km maximum interpolation radii. The results for 25 km were too sparse to meet the needs of the study. Table 3.4-6 shows the results for the 24-hr average, 1-hr daily maximum, and 6 a.m.-6 p.m. average NO_2 metrics. The spatially mapped NO_2 estimates have little bias (less than 1 ppb), but the errors average 19% to 37%. The results for 24-hr PM_{10}, shown in Table 3.4-6, indicate, again, there is little bias (less than 1.7 $\mu g/m^3$) between the estimated and observed values, and the errors are 25% to 32% depending on quality code. The coefficients of determination (r^2) indicate that

the spatially mapped NO_2 estimates explain more than 50% of the variance in the observed NO_2 for interpolation quality rankings 1 and 2. The r^2 values are highest for the monthly average daily maximum NO_2 concentrations; 74% of the variance in estimated daily maximum NO_2 can be explained by the variance in observed values of stations within 5 km. The PM_{10} estimates with quality rankings 1, 2, and 3 explain 42%, 41%, and 29% of variance in the observed PM_{10} values. The NO_2 and PM_{10} estimates are less accurate than the ozone estimates because there are fewer NO_2 and PM_{10} measurement locations and because local sources influence their concentrations. The measurement errors for historical exposure assignments for different pollutants need to be considered in the analysis of associations with health effects.

Table 3.4-5. Inverse distance weighted interpolation[2] (IDWI) of monthly average 1-hr, 8-hr, and 24-hr ozone 1975-1995 in 50 states.

Averaging Time	IDWI Quality Ranking[a]	Max Radius (km)	Number of data Points	Mean Estimated[b] Concentration (ppb)	Mean Observed Concentration (ppb)	Mean Bias (ppb)	Mean Normalized Bias (%)	Mean Error (ppb)	Mean Normalized Error (%)	R^2
1-hr Max	1	5	5708	48.62	49.23	-0.61	1.11	7.28	16.21	0.73
1-hr Max	2	25	52756	50.55	50.85	-0.29	1.26	6.08	13.15	0.86
1-hr Max	2	50	52756	50.58	50.85	-0.27	1.29	5.91	12.78	0.87
1-hr Max	2	100	52756	50.59	50.85	-0.26	1.32	5.89	12.74	0.87
1-hr Max	3	50	14355	47.76	47.68	0.08	1.02	5.97	13.35	0.80
1-hr Max	3	100	20780	47.58	47.27	0.31	1.66	5.97	13.56	0.78
10am-6pm	1	5	5401	39.25	39.81	-0.56	1.35	6.51	18.74	0.70
10am-6pm	2	25	51495	39.97	40.43	-0.46	1.06	5.58	15.80	0.83
10am-6pm	2	50	51495	39.98	40.43	-0.45	1.08	5.41	15.30	0.84
10am-6pm	2	100	51495	39.99	40.43	-0.44	1.10	5.39	15.25	0.84
10am-6pm	3	50	14232	38.23	38.56	-0.33	0.49	5.59	16.30	0.77
10am-6pm	3	100	20647	38.17	38.44	-0.27	0.84	5.50	16.20	0.76
8-hr Max	1	5	5712	40.31	40.94	-0.63	1.21	6.69	18.26	0.69
8-hr Max	2	25	52799	41.22	41.63	-0.42	1.19	5.56	15.07	0.83
8-hr Max	2	50	52799	41.22	41.63	-0.41	1.16	5.39	14.60	0.84
8-hr Max	2	100	52799	41.23	41.63	-0.41	1.19	5.37	14.56	0.84
8-hr Max	3	50	14374	40.04	40.39	-0.34	0.53	5.58	15.28	0.78
8-hr Max	3	100	20782	39.98	40.26	-0.28	0.89	5.53	15.31	0.77
24-hr Avg	1	5	5274	27.86	28.43	-0.57	0.72	5.39	20.61	0.58
24-hr Avg	2	25	49421	26.21	26.77	-0.56	0.59	4.76	19.13	0.66
24-hr Avg	2	50	49421	26.17	26.77	-0.60	0.43	4.59	18.44	0.68
24-hr Avg	2	100	49421	26.18	26.77	-0.59	0.46	4.57	18.39	0.68
24-hr Avg	3	50	13822	26.58	27.64	-1.07	-1.04	5.30	20.77	0.64
24-hr Avg	3	100	19965	26.65	27.84	-1.19	-1.21	5.29	20.49	0.63

[a] See text for explanation of spatial interpolation quality ranking

b For purposes of evaluating the spatial mapping errors, estimates are made for each measurement station location without using the data from the specific measurement station in the interpolation.

Table 3.4-6. Inverse distance weighted interpolation (IDWI) of monthly average 24-hr, 1-hr max, and 6 am to 6 pm NO_2 for 1975-1995 in the 50 states.

Averaging Time	IDWI Quality Ranking[a]	Number of Data Points	Mean Estimated[b] Concentration (ppb)	Mean Observed Concentration (ppb)	Mean Bias (ppb)	Mean Normalized Bias (%)	Mean Error (ppb)	Mean Normalized Error (%)	R^2
24-hr NO_2	1	4126	25.6	26.0	-0.38	2.29	6.09	24.9	0.51
	2	19605	26.9	26.6	0.31	5.98	6.14	26.3	0.64
	3	5383	22.2	21.3	0.97	10.49	6.72	35.0	0.21
Daily Max NO_2	1	5385	37.6	38.1	-0.40	1.49	6.80	19.1	0.74
	2	22266	45.0	44.2	0.79	6.48	8.65	23.1	0.73
	3	6176	39.0	36.8	2.16	14.21	10.39	33.6	0.30
6am-6pm NO_2	1	3807	25.1	25.5	-0.38	3.67	6.85	28.9	0.47
	2	18230	26.5	26.3	0.28	6.55	6.66	28.8	0.62
	3	5129	21.1	20.0	1.11	11.23	6.74	36.9	0.22

[a] See text for explanation of spatial interpolation quality ranking.

[b] For purposes of evaluating the spatial mapping errors, estimates are made for each measurement station location without using the data from the specific measurement station in the interpolation.

Table 3.4-7. Inverse distance weighted interpolation (IDWI) of monthly average 24-hr PM_{10} 1975-1995 in the 50 states.

Averaging Time	IDWI Quality Ranking[a]	Number of Data Points	Mean Estimated[b] Concentration ($\mu g/m^3$)	Mean Observed Concentration ($\mu g/m^3$)	Mean Bias ($\mu g/m^3$)	Mean Normalized Bias (%)	Mean Error ($\mu g/m^3$)	Mean Normalized Error (%)	R^2
24-hr PM_{10}	1	231938	34.8	34.2	0.64	7.71	8.22	24.9	0.42
	2	175967	32.5	30.8	1.66	12.55	7.78	27.1	0.41
	3	52650	31.6	31.1	0.51	10.23	9.09	31.8	0.29

[a] See text for explanation of spatial interpolation quality ranking.

[b] For purposes of evaluating the spatial mapping errors, estimates are made for each measurement station location without using the data from the specific measurement station in the interpolation.

3.4.3. Exposure Modeling and Traffic Assessment Methods

3.4.3.1. Exposure Modeling Methods

Modeling was conducted with the Individual Exposure Model (IEM) to estimate the long-term average personal exposures of CHS participants to pollutants of ambient origin. The personal exposure estimates were designed to complement the ambient air quality data collected at central monitoring sites in the communities. The microenvironmental modeling approach incorporated into the IEM was designed to account for differences in the amount of time subjects spent in different microenvironments and differences in the air pollutant concentrations of ambient origin present in those environments, including differences due to variability in subjects' exposure to traffic-related pollutants (Wu et al. 2003). The exposure modeling approach is similar to that used in other microenvironmental models (Duan 1982; Ott et al. 1988; Burke et al. 2001); however, air quality dispersion models are used to characterize neighborhood-scale concentrations of traffic-related pollutants.

The IEM model uses the microenvironmental approach where the integrated exposure is estimated in Equation 3-1 as the sum of exposures in each microenvironment occupied by the individual for the time period of interest (Sexton and Ryan 1988).

$$E_{ij} = \sum_{k=1}^{T} \sum_{m=1}^{M} C_{ijkm} \Delta t_{ijkm} \qquad \textbf{(3-1)}$$

where

E_{ij} =integrated exposure of the i^{th} individual on day j

C_{ijkm} = concentration in the m^{th} microenvironment during hour k of day j when it is occupied by the i^{th} individual

Δt_{ijkm} =amount of time the i^{th} individual spent in the m^{th} microenvironment during hour k of day j

The exposure model uses the hourly and two-week average concentration data available in each community for each year as inputs and employs a Method I averaging approach.

In order to express exposure in the familiar units of concentration, rather than concentration time, the IEM uses the time-weighted exposure (E'), which is calculated using Equation 3-2:

$$E'_{ij} = \frac{1}{\sum_{k=1}^{24} \sum_{m=1}^{M} \Delta t_{ijkm}} \sum_{k=1}^{24} \sum_{m=1}^{M} C_{ijkm} \Delta t_{ijkm} \qquad \textbf{(3-2)}$$

The IEM was applied to estimate personal exposures of CHS participants to NO_2, $PM_{2.5}$ mass, EC, OC, PM_{10} mass, and ozone. The subjects were assumed to stay in their community throughout the year and divide their time between the following five microenvironments: residential outdoor, school outdoor, residential indoor, school indoor, and in-vehicle. A schematic of the components of the model used for the CHS application is shown in Figure

3.4-8. The methods used to determine the time-activity patterns (Δt_{ijkm}) and microenvironmental concentrations (C_{ijkm}) are described below.

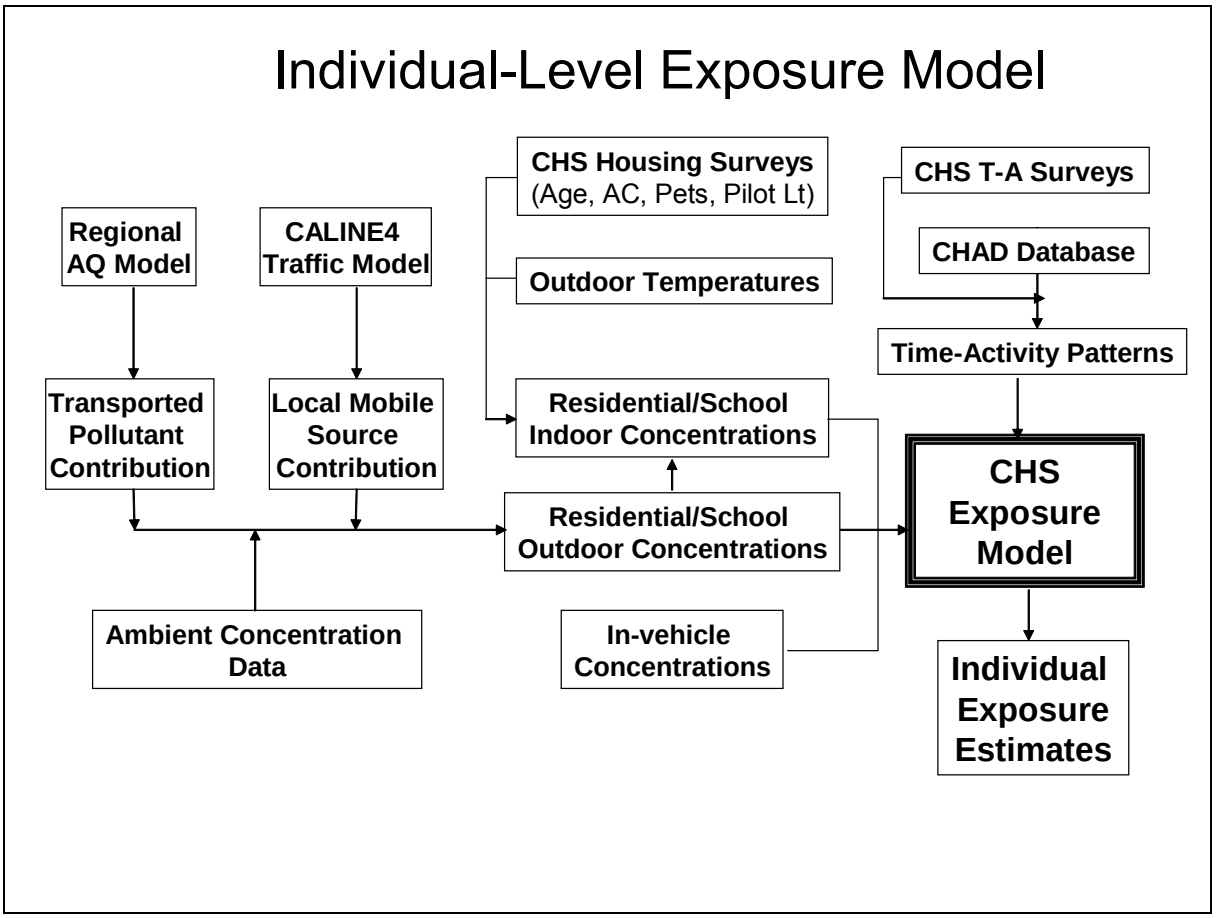


Figure 3.4-8. Schematic of the individual exposure model.

3.4.3.1.1. Time-activity Patterns

Time-activity surveys were administered twice a year to CHS subjects. The survey asked the children how much time (in 5 categories) they spent outdoors in the afternoons (12-6pm), on the weekdays and the weekend days, and during the summer. The survey also asked the children if they spent more than 15 minutes daily traveling between school and home and by what means. The survey data were useful for ranking the individual children based on time-use characteristics; however, they were insufficient for detailed exposure modeling that covered 24 hours of the day.

The Consolidated Human Activity Database (CHAD) developed by the U.S. EPA provides 24-hour time-activity patterns (with 1-minute time resolution) based on telephone recall diaries (US Environmental Protection Agency 1997a). A time-activity submodel was developed to create 24-hr time-activity series for each child in the CHS cohort by using information from both the CHS surveys and the CHAD database. The CHS subjects were assigned to one of 48 time-activity categories distinguished by age (9-12 and 13-18 years old), gender, time spent outdoors (low, high), day type (non-summer weekdays, non-summer weekend days, and summer days), and time spent in vehicles (low, high). The assignments were based on the average characteristic determined from the two CHS surveys in each year. The children were categorized with respect to time in travel based on whether they reported more than or less than 15 minutes vehicle travel time from school to home. Fifteen minutes is an appropriate grouping variable here since the 1995 National Personal Travel Survey also reported an average of 15 minutes for children's travel to schools. Detailed time-activity patterns were extracted from the CHAD database for the same 48 categories and aggregated into 15-minute intervals. The CHAD profiles for each age, gender, and day-type were stratified into high- and low-time-outdoors and time-in-vehicle subgroups. The median time outdoors and time in vehicles are shown in Table 3.4-8. Only CHAD time-activity data collected in the ARB (Wiley et al. 1991) and NHAPS (Nelson et al. 1994) surveys were used to ensure the quality and consistency of the data. Categories with less than 20 profiles were combined to avoid biases caused by sampling from too small a data set. The large number of CHAD codes for locations and activities were aggregated into the five corresponding microenvironments (residential indoor, residential outdoor, school indoor, school outdoor and in-vehicle) considered in this study. All other locations were grouped into another category that was treated the same as the residential indoor microenvironment. The IEM model sampled from the subset of CHAD profiles on each day-type that was appropriate for each child, based on the child's age, gender, and ranking of time outdoors and in vehicles.

Table 3.4-8. The median time outdoors and time in vehicles by age, gender, and day type for time-activity patterns extracted from CHAD. CHS children were assigned to groups with low or high times in these two locations based on the CHS surveys.

Age Group	Gender Group	Day Type	Median Time Outdoors[1] (min)	Median Time In-Vehicles[1] (min)
7 - 12	Male	Nonsummer Week Days	60	30
		Nonsummer Weekend Days	90	45
		Summer	105	30
	Female	Nonsummer Week Days	30	45
		Nonsummer Weekend Days	45	60
		Summer	75	45
13 - 18	Male	Nonsummer Week Days	15	45
		Nonsummer Weekend Days	90	45
		Summer	60	45
	Female	Nonsummer Week Days	<7[2]	60
		Nonsummer Weekend Days	<7[2]	75
		Summer	15	30

1. Times are truncated to nearest 15 minute interval. The median was the cut-point for the low and high times in these locations.
2. Based on 15 minutes interval time-activity data. For these two groups, all profiles with zero time outdoors were grouped into "low" category and all others with greater than zero values were grouped into "high" category.

3.4.3.1.2. Outdoor Microenvironmental Concentrations

The NIH pilot study of ambient NO_2 levels in CHS communities demonstrated that there is substantial within-community variation of ambient concentrations (see Section 4.1). The CHS exposure modeling attempts to estimate the neighborhood-scale spatial variations in outdoor concentrations. The outdoor concentrations at a specific location in a community are assumed to be a result of contributions from local source emissions and pollutants transported from upwind regions. For the pollutants addressed in this study, local traffic-related emissions are hypothesized to be the primary source of within-community spatial variations. Local stationary source and off-road mobile emissions are likely to be less important than on-road mobile sources in most communities. Thus, the ambient concentrations at specific locations are decomposed into the components from (1) local mobile source emissions and (2) transported pollution plus local stationary sources. The ambient pollutant concentrations for each residence and school are calculated as the sum of these two components:

$$C_{outdoor} = C_{local\,traffic} + C_{transported} \quad \textbf{(3-3)}$$

$$C_{outdoor} = C_{CALINE\,4} + \alpha C_{Measured}^{Central\,Site} \quad \textbf{(3-4)}$$

where

$C_{local\ traffic}$ = estimated concentration from all mobile source emissions in the community as determine from an air quality dispersion model (CALINE4)

$C_{transported}$ = pollutant transported into to the community determined as a fraction (α) of the concentration measured at the community central air monitoring site

The local traffic component is estimated using the CALINE4 dispersion model which is specifically designed to estimate concentrations near roadways. The fraction (α) of central-site measured concentrations due to transport and other local sources can be estimated by regional air quality models or from intracommunity measurement data.

The CALINE4 model, developed by the California Department of Transportation (Caltrans) and the U.S. Federal Highway Administration (Benson 1989), is one of several Gaussian line source dispersion models that is designed to estimate local-scale pollutant concentrations from motor vehicle emissions. Often it is used to estimate worst case 1-hr or 8-hr maximum CO concentrations from a congested roadway. It has primarily been evaluated for inert traffic-related pollutants such as CO. It was selected for the CHS analyses because it has a credible scientific formulation and it is recommended by Caltrans for analysis of CO imposed from transportation projects in California (Garza et al. 1997). The objective of the CHS application is to characterize long-term exposures (annual or warm season and cool season) from vehicle emissions on all roads in a community. The principal inputs to the model are meteorological conditions, traffic volumes, roadway geometry, and vehicle emission rates. The approach used to specify these inputs for long-term exposure is somewhat different than is used for the worst case analyses.

A climatological approach is used to estimate long-term average concentrations. The model is applied for a wide range of meteorological cases in each community, and the seasonal or annual concentrations are calculated by weighting the results for individual cases by the frequency of occurrence of the conditions in the community. Typically, the model is run for 6 wind speeds, 16 wind directions, and 3 atmospheric stability conditions (288 cases total). The frequency of occurrence of meteorological conditions is determined from five years of surface observations (1995-1999) measured in or near the communities. The cases are run for an inert pollutant using daily average hourly traffic volumes. The results are post-processed to incorporate not only the meteorological frequency of occurrence but also the pollutant-specific emission factors, diurnal and day-of-week variations of traffic volumes, and chemical conversion (for NO to NO_2).

Annual average daily traffic counts data for Interstate freeways, other principal arterials, minor arterials, major collectors, and minor collectors in 2000 were obtained from Caltrans. The annual traffic counts are based on continuous measurement data for freeways and intermittent measurements (usually every three years) on other arterials and some collectors. Annual traffic volumes for 1994-2000 were backcast from the 2000 data assuming a 2 percent per year growth rate. Diurnal traffic volume variations and day-of-week variations for light-duty and heavy-duty vehicles on freeways were obtained by averaging Caltrans weigh-in-motion (WIM) data for Southern California. Diurnal variations for collectors were determined from more limited traffic measurements (Chinkin et al. 2002).

Additional databases were used to apportion the total traffic volume to heavy-duty diesel vehicles (HDV) and light-duty gasoline vehicles (LDV). For the seven communities located within the SoCAB, including Orange County and portions of Los Angeles, Riverside, and San Bernardino Counties, HDV fractions of total traffic volumes were derived from travel-demand transportation model simulations. The Southern California Association of governments (SCAG) performed separate simulations of LDV and HDV traffic volumes in the SoCAB area for 1997. The simulations incorporate an accurate link-node roadway network for freeways and a surrogate link-node roadway network for nonfreeways. Despite concerns regarding this network, it provides much better spatial resolution for the fractions of LDV and HDV than any other database. HDV fractions of total traffic volume in the five communities located outside the SoCAB were derived from the Caltrans statewide truck-traffic-volume database for freeways and state highways for 1998. It is linked to a third roadway network, the state post-mile roadway system. Both databases rely heavily on data from the WIM sensors at selected freeway locations. HDV fractions (~3%) for collector streets were determined from more limited traffic measurements (Chinkin et al. 2002).

ArcInfo was used to preprocess the Caltrans roadway link and traffic count data. The Caltrans roadway geometries were mostly based on TIGER files and were often inaccurate. Comparison to global positioning system (GPS)-accurate TeleAtlas Roadway Network data showed that the Caltrans' TIGER roadway links occasionally had 250-m discrepancies from actual roadway locations. Figure 3.4-9 shows an example of a freeway location difference in Riverside. The roadway geometry errors were random and affected roadways of all sizes in most communities. Zhu et al. (2002a; 2002b) reported ten-fold differences in measured concentrations of traffic-related pollutants between 30 m and 200 m downwind of Southern California freeways. Errors of 100 m to 250 m in the location of major roadways relative to residence are not acceptable for

neighborhood-scale assessments of traffic effects. Since the Caltrans roadway location data did not have sufficient accuracy for our intended use, the Caltrans annual traffic volumes were transferred to the GPS-accurate TeleAtlas roadway network. The HDV fractions of total traffic volumes were also mapped from their original network to the TeleAtlas network. The TeleAtlas roadway database incorporated both more accurate and more precise location information. For example, each direction of travel on moderate and large roadways is represented as a separate link in the TeleAtlas database.

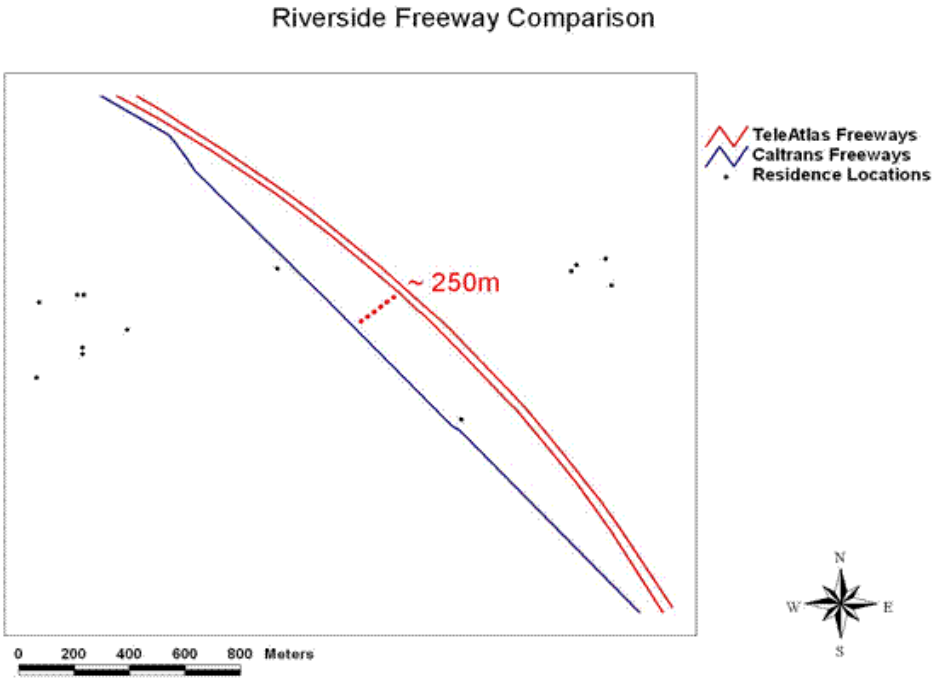


Figure 3.4-9. Comparison of Caltrans and TeleAtlas roadway location data for a freeway in Riverside.

The street address of residences and schools of CHS residences were geocoded on the TeleAtlas roads. Considerable effort was made to standardize the ~8500 participant addresses, which resulted in more than 98% of the residences having the highest quality geocoding match quality. Proper addresses are geocoded between cross-streets based on their street numbers and at a distance of 13 m from the roadway on the correct side of the street. In densely populated areas with uniform lot sizes and house numbering, the valid addresses are mapped at a GPS accuracy of ±20 meters.

Vehicle emission factors were obtained from the ARB's EMFAC2002 vehicle emissions model. County vehicle registration data and community monthly average temperatures were used in the model to estimate the fleet average NO_x, CO, and PM emission rates for light-duty gasoline and high-duty diesel vehicles traveling at speeds ranging from 20 to 70 mph in 1994 through 2000. The emission factors for nonfreeways were based on the average of the emission factors for 20, 30, and 40 mph. Vehicle emission factors on freeways were based on the average of the emission factors for 50, 60, and 70 mph. The EC and OC fractions of exhaust PM emissions were based on composite profiles from Gillies and Gertner (2000). Paved road-dust emission

factors for $PM_{2.5}$ and PM_{10} were based on Southern California in-roadway measurements (Fitz and Bufalino 2002). The EMFAC model also estimates the PM emissions from brake wear and tire debris.

Table 3.4-9. Average transport factors for NO, NO_2, $PM_{2.5}$ Mass, and EC in CHS communities.

Community	NO		NO_2		$PM_{2.5}$		EC	
	Summer	Winter	Summer	Winter	Summer	Winter	Summer	Winter
Atascadero	0.75	0.74	0.87	0.88	0.91	0.90	0.85	0.84
Santa Maria	0.75	0.74	0.87	0.88	0.91	0.90	0.85	0.84
Lompoc	0.75	0.74	0.87	0.88	0.91	0.90	0.85	0.84
Lancaster	0.34	0.21	0.38	0.32	0.89	0.84	0.74	0.60
Lk Arrowhead	0.25	0.45	0.52	0.72	0.96	0.86	0.85	0.71
San Dimas	0.48	0.55	0.77	0.84	0.97	0.94	0.91	0.88
Upland	0.38	0.46	0.68	0.73	0.95	0.89	0.85	0.76
Mira Loma	0.48	0.60	0.63	0.85	0.98	0.96	0.91	0.86
Riverside	0.33	0.45	0.52	0.72	0.96	0.86	0.85	0.71
Long Beach	0.75	0.74	0.87	0.88	0.91	0.90	0.85	0.84
Lk Elsinore	0.18	0.23	0.29	0.50	0.97	0.97	0.87	0.84
Alpine	0.10	0.16	0.29	0.50	0.97	0.97	0.87	0.84

The outdoor pollutant concentrations due to transport from upwind and local non-mobile sources are estimated as a fraction of the ambient concentrations in the community. The fractions can be derived from regional air quality models or from analysis of ambient data and emissions. UCLA researchers made episodic simulations with the MM5+SMOG models (Lu et al. 1997a; Lu et al. 1997b) for Southern California where the on-road mobile source emissions for one community (~15-km x 15-km area) were removed. Comparison with baseline cases suggest the transport fraction, α, ranges from 0.55 to 0.99 depending on pollutant, community, and time of day. The UCLA study did not simulate factors in all of the communities and used outdated emissions estimates, yet it confirmed the importance of transport for neighborhood-scale analyses within urban areas. The transport factors used for IEM applications to the CHS are shown in Table 3.4-9.

3.4.3.1.3. Indoor Microenvironmental Concentrations

Human beings spend the majority of time indoors which is important for exposure assessment. Indoor pollution may be due to indoor sources or infiltration of outdoor ambient air. Exposure to pollution of ambient origin is most relevant for the CHS objectives so modeling is needed to separate the contributions of outdoor air and indoor sources to personal exposures (Wilson et al. 2000). A single-compartment, steady-state mass balance equation is used to calculate indoor concentrations:

$$C_{indoor} = \frac{paC_{out}}{a+k} + \frac{Q_{is}}{(a+k)V} \quad \textbf{(3-5)}$$

where:

p = penetration coefficient;
a = air exchange rate (h^{-1});
k = lumped deposition rate and chemical decay rate (h^{-1});
Q_{is} = the emission rates from indoor sources (μg/h); and
V = the volume of the (m^3)

The first and second terms represent the contributions of outdoor air and indoor sources to indoor concentrations. Indoor sources can be important contributors to personal exposure but were not included (Q_{is}=0) for the CHS application since the epidemiologists planned to separately examine factors associated with indoor sources (e.g., gas stove usage) as categorical variables in their statistical analyses. Therefore, the personal exposures obtained from the present study are exposures due to pollutants of ambient origin only.

The penetration factor represents the losses that occur as pollutants enter a building. For all gaseous pollutants, a unit penetration coefficient was used. The penetration coefficients for PM were represented as triangular distribution with minimum, mode, and maximum values of 0.9, 0.95 and 1.0, respectively (Koontz et al. 1998). NO_2 and ozone are reactive and tend to deposit on indoor surfaces rapidly. A decay rate of 1.0 h^{-1} was used for NO_2 (Yamanaka 1984), which was within the range of 0.2-1.3 h^{-1} found by Nazaroff et al. (1993). A decay rate of 2.8 h^{-1} was used for ozone (Lee et al. 1999). A zero decay rate was assigned to CO because of its inert properties. The deposition rates of PM_{10} and $PM_{2.5}$ were represented as normal distributions with means and standard deviations of 0.65 h^{-1} (±0.28) and 0.39 h^{-1} (±0.16), respectively (Ozkaynak et al. 1994). The deposition rates of EC and OC were the same as those for $PM_{2.5}$ mass. Air exchange rate measurements are not available for most CHS homes. Colome et al. (1994) found in the California Residential Indoor Air Quality Study that air exchange rates were closely associated with home volume, cooking type, and heating type. A mean air exchange rate was assigned to each CHS home according to the house type (single-detached and attached) and cooking facilities (electric stove, gas stove without pilot light, gas stove with pilot light) based on Colome et al. (1994).

3.4.3.1.4. In vehicle Microenvironmental Concentrations

ARB's recent studies of in-vehicle concentrations show there is a wide range of conditions on roadways, including ones with concentrations that are much higher than commonly observed at routine ambient air monitors (Rodes et al. 1998; Fitz et al. 2003; Fruin 2003). Vehicle occupants traveling behind and/or inside high-emitting vehicles are exposed to high levels of traffic-related pollutants. Even though children typically spend less than one hour per day in vehicles, the high concentrations in this microenvironment may contribute significantly to their total exposure. Existing data are insufficient to accurately characterize the distributions of exposures in vehicles. Hence, mean concentrations were used for communities with low and high traffic conditions. The mean concentrations derived by Rodes et al. (1998) and Fruin (2003) for urban conditions in Los Angeles and rural conditions in Sacramento were used for the high traffic and low traffic CHS communities, respectively. The mean concentrations are listed in Table 3.4-10.

Table 3.4-10. In-vehicle concentrations for the low- and high-traffic communities.

Pollutant	High Traffic Community	Low Traffic Community	Reference
CO (ppm)	4.5 ppm	1.7 ppm	(Rodes et al. 1998)
NO_2	3 × ambient	3 × ambient	(Fitz et al. 2003)
PM_{10} mass	60 ($\mu g/m^3$)	21 ($\mu g/m^3$)	(Rodes et al. 1998)
$PM_{2.5}$ mass	49 ($\mu g/m^3$)	11 ($\mu g/m^3$)	(Rodes et al. 1998)
$PM_{2.5}$ EC	7 ($\mu g/m^3$)	4 ($\mu g/m^3$)	(Fruin 2003) (Zhu et al. 2002b)
$PM_{2.5}$ OC	0.5 × $PM_{2.5}$	0.5 × $PM_{2.5}$	(Sioutas 2003)
Ozone	0.1 × ambient	0.1 × ambient	(Chan et al. 1991)

3.4.3.2. Traffic Assessment Methods

Numerous studies have shown associations between traffic and respiratory health end-points (Duhme et al. 1996; Oosterlee et al. 1996; Brunekreef et al. 1997; van Vliet et al. 1997; Ciccone et al. 1998; Kunzli et al. 2000b; Hoek et al. 2002; Lin et al. 2002). There was an interest in examining the relationships of CHS participant health status with traffic indicators separately from the CALINE4 dispersion model estimates of concentrations from mobile source emissions. Dispersion modeling provides refinements but introduces additional uncertainties compared to analysis of traffic alone. Hence, a three-level hierarchical approach was adopted for traffic assessment that considered (1) the distance of residences to nearest roadways of various types, (2) GIS-mapped traffic density assignments at residences, and (3) the CALINE4 dispersions model estimates of traffic-related pollutant concentrations at residences (as described above).

The database of Caltrans annual average daily traffic volumes and HDV volume fractions mapped to the GPS-accurate TeleAtlas roadway network (described above) was used for the analyses. In order to ensure the accuracy of the traffic exposure assignments, CHS baseline and annual survey addresses were geocoded using the same TeleAtlas database and software. Figure 3.4-10 shows a map of the updated residence locations along with traffic volumes in Long Beach. It conveys the typical extent of coverage for roads with traffic data and the within-community variations in proximity to roadways. The first sets of traffic metrics were the distances from residences to the nearest roadways of different types and the associated LDV and HDV traffic volumes on those roads. GIS tools were used to calculate the distance to the nearest (1) interstate freeway, U.S. highway, or limited access highway; (2) other highways; (3) arterial roads; (4) collector roads; and (5) local roads. The database of distances and volumes metrics were compiled and considered for use in analysis of asthma incidence and prevalence. Initially, these data were used to obtain an understanding of which subjects were living within the zone of influence (e.g., 50 m, 100 m, 150 m, 200 m, etc.) of busy roads.

The second approach for characterization of traffic exposures was to calculate traffic densities, which vary more smoothly in space than the distances to nearest roads. They also capture the effects of intersection and multiple roadway influences that are missed using only distance to the nearest roadways. The link-based traffic volumes are used to generate maps of traffic density using the ARCInfo Spatial Analyst software. Traffic density maps are created with a Gaussian decay function that has traffic densities decreasing by ~90% between the roadway and 150 m away (perpendicular) from the roadways, which is consistent with the characteristics observed by Zhu et al. (2002b). Identical mapping procedures are used in all the communities so that the results are comparable across communities. The densities reflect proximity to traffic without consideration of differential exposures caused by meteorology (a limitation of the ARCInfo Spatial Analyst software). The traffic densities are mapped as if the wind speeds and directions were uniformly distributed across all quadrants. The traffic density map for Long Beach is shown in Figure 3.4-11. It clearly shows high densities in narrow bands along the freeways, moderate densities along major arterials, and lower densities in the suburban neighborhoods. A database of densities at all of the CHS residences was compiled for use in the health analysis.

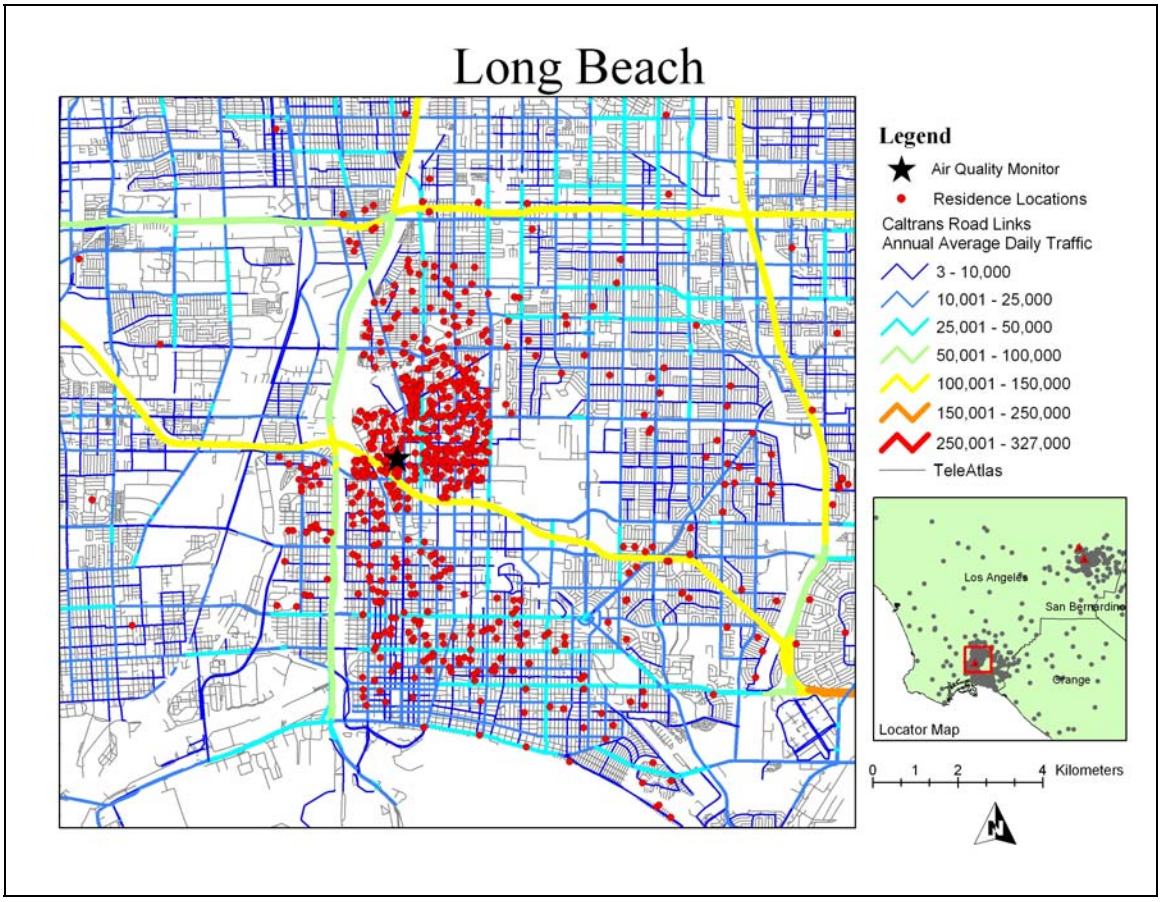


Figure 3.4-10. Annual average traffic volumes in Long Beach, CA and CHS residences.

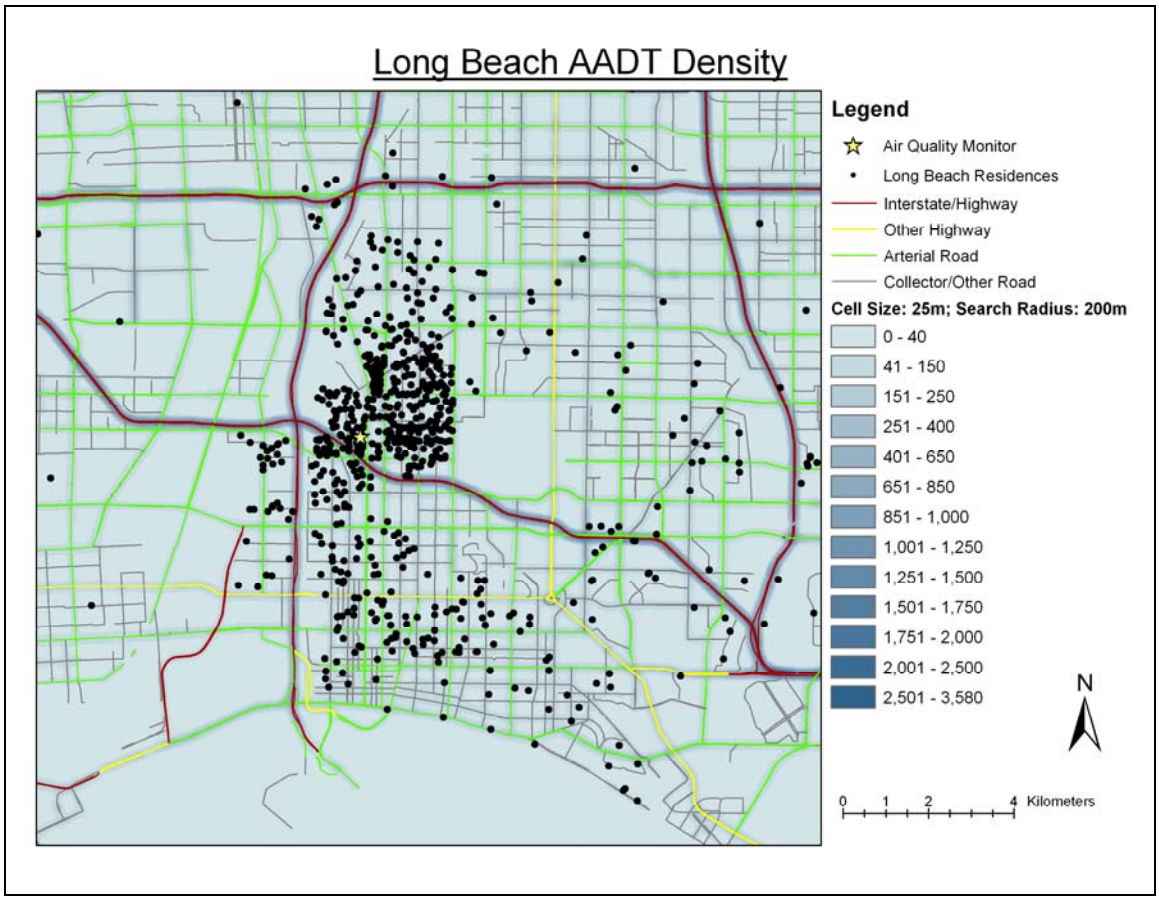


Figure 3.4-11. GIS-mapped traffic density in Long Beach, CA.

3.5. Health Outcomes/Endpoints

3.5.1. The Questionnaire

The questionnaire was composed of several sections: demographics, a medical history, a housing survey, history of exposure to ETS, pests and pets, and a time-activity assessment. Early in Phase II, we pilot-tested the questionnaire among families of USC staff and other project volunteers ineligible for main study participation (e.g., neighbors of staff, community friends). During Phase I, pilot testing of an earlier version was performed among volunteers recruited from local pediatricians' offices seeing large numbers of asthma cases.

Two questionnaires were given simultaneously to participating subjects in 1993, the first school year of study. One survey contained questions about the subject's medical history and housing history, while the other contained questions about physical activity. In later years, an activity questionnaire was administered in the early winter and additional activity questions were asked of subjects in the spring during visits for lung function testing in the schools.

In administering the questionnaire, we contacted all project participants in a given community during a single visit and enlisted the active participation of the selected schools. The project field staff arrived on-site to meet with teachers regarding circulation of the survey material prior to its distribution. Questionnaires were precoded by community classroom.

Students were instructed (by teachers) to take the questionnaires home that evening, to help their parents or legal guardians accurately complete as much of the survey as possible, and to return the completed and sealed survey to school the following day (the parents and guardians of fourth-graders completed the surveys in their entirety, whereas the seventh and 10th graders usually completed portions of the questionnaires themselves). Spanish translations of the questionnaire were available.

Questionnaires were logged in each day at school by project staff personnel or locally-recruited volunteers. Participants not returning questionnaires were encouraged to complete and return them as soon as possible. Iterative logging and evaluation of the return/completion rate was updated frequently.

Questionnaires were returned to study personnel at USC via personal courier or expedited mailing. As soon as the questionnaires were logged in at USC, we decided either to accept the return rate or to institute an additional distribution. Incentives were offered to classrooms achieving high participation rates. These took the form of games, books or recreational equipment.

An activity assessment portion of the questionnaire was administered during lung function testing field operations in the spring. Subjects were given the time/activity pattern assessment surveys by project testing staff following their lung function testing battery, asked to complete the written surveys, and return the documents to field testing staff within the subsequent 24-hours. Seventh and tenth graders generally completed these at the time of testing in school.

Fourth graders were asked to take them home. Due to low 4th grade return rates, a large number of 4th grade surveys were completed with parents via telephone interview.

An extensive set of questions was asked about the history of respiratory diseases. These included asthma, bronchitis and pneumonia and associated symptoms such as cough, phlegm production and wheezing. The initial questionnaire collected information on the past history of these conditions and symptoms including frequency and time of onset. Asthma questions considered physician diagnosis, severity and medication use. Each annual follow-up questionnaire concentrated on adverse respiratory health experiences during the past year and allowed us to ascertain the incidence of new-onset conditions such as physician-diagnosed asthma and bronchitis.

3.5.2. Lung Function Testing

Lung function testing took place in the spring of 1993 and in each subsequent spring to minimize seasonal confounding with intercurrent summer or winter acute air pollution episodes. Spirometers (Morgan Spiroflow Model 132 Rolling Seal Spirometers) and associated computerized interfaces and data logging capabilities were used to document lung function performance. Testing was performed in a predetermined and administration-approved area of the school (nurse's office, library, multi-purpose room). To the extent possible, we performed testing during the morning hours to avoid possible acute effects of potential daily peaks in ambient air pollution.

Project staff tested the subjects individually in a consistent, prearranged manner: (1) A subject initially removed his/her shoes. (2) We determined standing height and weight to provide necessary information for standardizing predictions of lung function performance based on gender, race, height, and weight; then we began testing. (3) A subject was then privately questioned regarding smoking; this approach has proven more effective than written responses in prior epidemiologic investigations (Speizer, personal communication, 1992). (4) In a seated position, pinching their nose shut with their fingers, the subject was asked to perform at least 3 satisfactory maximal expiratory maneuvers. A maximum of 7 efforts were attempted. (5) After testing, a subject was given a time/activity assessment survey to complete at school or at home that evening and he/she was instructed to return the completed survey to school that day or the following day (either to project staff or local volunteers).

Experience has shown that with proper training and instruction, virtually all 10-year-old children can satisfactorily perform pulmonary function tests. Our system employs visualization aids that assist the technician in achieving performance characteristics that meet or exceed the criteria set by the American Thoracic Society (ATS). Our technicians are carefully trained and monitored to maintain excellence.

Up to six testing units (spirometers), operated by trained lung function technicians, were dispatched to conduct field-testing in a given community. To facilitate testing, the six units were assigned as needed to fourth-grade, seventh-grade, and tenth-grade testing. In some instances, a local volunteer at the testing location provided assistance in obtaining students for testing and measuring of students' heights and weights. Pulmonary function was remeasured on a sample of about 10% of the subjects to provide data for test-retest reliability

Students left their classrooms in groups of about five or six, at approximately 10-minute intervals, for testing. Staggering the flow of students ensured that technician time was efficiently used and that student time away from classroom activities was minimized.

Testing was scheduled to conform with classroom schedules (Monday through Friday). To accommodate this schedule, field staff traveled on Sunday or Monday, set up equipment and carried out testing on Monday through Friday mornings. Friday afternoons were reserved for data management, administration, and travel.

Each community was visited at least twice (at least one month apart) with half the participating subjects being tested each visit. This approach minimized the potential confounding of intercurrent acute pollution episodes, outbreaks of illness (e.g. flu), or some other unidentified and uncontrolled variable; it also facilitated completion of testing within one week of on-site presence. The study population was arbitrarily divided (by classroom, or other convenient division) into two roughly equal subgroups, one of which was tested during the initial visit, and the other tested later in the same testing year.

The annual follow-up pulmonary function tests were planned to achieve as close to a 12-month interval between testing as possible. Most of the intervals fell between 11 and 13 months. Extensive procedures were established to maximize accuracy and reproducibility and included frequent calibration of equipment and frequent evaluation of technicians. These procedures have been described and published in the scientific literature to share our successful approaches with others (Linn et al. 1998; Gilliland et al. 1999a; Enright et al. 2000).

3.5.3. Absence Monitoring

The school absence monitoring activity was designed to collect data to determine the frequency and severity of respiratory illnesses in relation to concurrent ambient air pollution levels and to compare respiratory disease patterns between communities. Because schools were required to keep data on the nature and type of absences to receive capitation funding for students up until the late 1990s, there was motivation for schools to collect accurate data.

We used documented school absences to trigger an investigation of the reason for the absence. This involved phoning the student's home to interview the parent or guardian. In deciding what questions to include, literature from studies that used symptom questionnaires in children was reviewed. We gave parents an opportunity to report their opinion of the illness, but also we asked a list of symptoms for each call. A caller could skip the symptom list if the parent said that the illness was due to a musculoskeletal injury, or if the parent said that the absence was for social reasons (coded as a non-illness absence). By this approach we were able to classify whether the illness was respiratory. We also asked whether the child had seen a doctor and if they had, we asked about the doctor’s diagnosis. We asked about use of medications since this might provide an indication of the severity of the illness.

In August of 1993, the 6 field technicians began contacting school district offices to acquire data on rates of absences. In the fall of 1993, a pilot study of absence monitoring was implemented to assess the feasibility of the approach and to work out details.

In their initial contacts with the schools, the field technicians did the following:

1. Ascertained how the districts excused absences: note from home, call from home, or both.
2. Ascertained the method the schools used to record absences: computer data base versus paper records.
3. Ascertained how frequently the various schools were willing to provide data: weekly, bi-weekly, monthly, quarterly.
4. Attempted to ascertain the coding format schools used for absences: excused/not excused, ill/not ill.
5. Arranged how the data would be sent to USC - generally by postage paid mailers.

Because the data were obtained from schools in different formats, the ease of using the data varied by school. In some schools, students in our study were subsetted so that the absence data included only these children. In most schools, data on all children were included and thus the field team was obliged to enter names into our database to ascertain if each child was a study participant. The schools used different codes for the data and different report formats so that it turned out to be more time consuming to input the data than originally anticipated. For each eligible absence, the field technician filled out the top of a calling form with the child's name, name of parent who signed the form, phone number and dates of absence. Calling sheets were divided among the field personnel. Using a script and structured questionnaire the field technicians made calls to the participant's parents or guardians.

We also learned that most parents had difficulty remembering absences that were more than 4 weeks in the past, and that 4-8 p.m. was the most efficient time to make telephone contact. Parents were generally willing to provide absence information although some were concerned about how frequently they might get called. The caller often reached another adult in the household who appeared fully knowledgeable about the absence. Because it can be difficult to reach the person who signed the consent form and repeated calls are bothersome to the other household member who is able to provide the information, we decided that after one try to reach the signer, we would accept another adult over the age of 18 as the respondent.

A project coordinator and two part-time interviewers, dedicated to the absence monitoring activity, were hired and trained in January 1994. The project coordinator was responsible for overseeing the receipt and entry of all absence monitoring data, generating a listing of students that needed to be called, distributing them to the two callers, receiving completed surveys and making quality control calls. Our pilot study had demonstrated that the delay created by having the district collate the data would result in most students having absences more than 4 weeks in the past at the time we received the data. Since we had limited resources to make calls, we established a priority system which focused our emphasis on the youngest group. More detail is provided in the publication by Gilliland et al. (2001b).

3.6. Statistical Approaches

Exposure-response relationships between the various health endpoints and ambient pollution levels have been analyzed using a three-level hierarchical random effects model, essentially a generalization of the pioneering work of Laird and Ware (1982). Brief summaries of this analysis

approach for the relevant endpoints have been provided in substantive publications (see especially Gauderman et al. (2000) and Gilliland et al. (2001b)), and a more comprehensive technical discussion of all the relevant statistical issues is provided in Berhane et al. (in press). A discussion of our approach to accounting for indoor sources of air pollution to the risk of asthma can be found in McConnell et al. (2002a). Here we briefly summarize the basic approach.

In general, our modeling approach has been to include variables with known associations to outcome (e.g., height for pulmonary function testing), and design variables or potential confounders that may vary by community (e.g. field technician, race). Outcome selection has been based on what is most interesting (e.g. FEV_1, FVC, doctor-diagnosed asthma) with additional variables added depending on the focus of the paper. For missing data, all analyses are based on complete-case observations (with both outcomes and covariates), rather than using imputation or other data fill-in procedures. Missing data was a factor in the choice of covariates in some models, in that we tended to favor a model with fewer covariates and more subjects rather than more covariates and fewer subjects. "Hayfever," for example, was a variable with a lot of missing data (i.e. approximately 9%); thus, although it was a potential confounder for some outcomes, it was typically considered in sensitivity analyses but not included in the final models.

In order to exploit the three levels of comparison (between times, between subjects, and between communities) in an integrated manner, we use the basic approach of generalized linear mixed models with a hierarchical structure of random effects (Laird and Ware 1982; Breslow and Clayton 1993). Conceptually, this is easiest to understand as if the analyses were conducted in three stages as follows (Figure 3.6-1): in the first stage, one fits a separate regression (for each subject) of the outcome variable (e.g., annual FEV_1 measurements) on age and other time-related covariates [e.g. height, weight, and (for lung function measurements) spirometer used]; in the second stage, these fitted slope (and possibly intercept) estimates for each subject are then regressed on a set of indicator variables for each community and subject-specific covariates (e.g., race, date of birth, residential activity, baseline asthma status); in the third and final stage, the estimated community effects are then regressed on the ambient air pollution levels and possibly other community-specific covariates (e.g. altitudes). The effect of ambient air pollution is thus assessed at the last of these stages, thereby properly reflecting the effective number of observations for this level of comparison and its appropriate between-communities residual variance. To summarize in its simplest form, suppose we let Y_{cit} denote the outcome variable for community c, individual i, at age t, and the mean ambient pollution level by X_c, then ignoring covariates other than air pollution and focusing on the regression of rates of change, we could write the model in this three-stage fashion as:

$$Y_{cit} = a_{ci} + b_{ci}t + e_{cit} \quad (3.6.1)$$

$$b_{ci} = B_c + e_{ci}' \quad (3.6.2)$$

$$B_c = \beta_0 + \beta X_c + e_c'' \quad (3.6.3)$$

It is also possible to assess the effects of air pollution at the individual level by including the deviations of personal exposure x_{ci} from the community mean (e.g., $x_{ci} - X_c$) in the second equation, or at the temporal level by including the deviations of annual exposures from the long-term averages (e.g., $X_{ct} - X_c$) in the first equation. It is also possible to assess the effects of air pollution both on the overall level of the outcome variable (using a similar regression model for

the intercept estimates a_{ci}) and on the rates of change (using the regression of the slope estimates b_{ci}). It is conceptually straight-forward to add other levels of comparisons, such as between schools within a community or between cohorts. Finally, the approach can be applied to binary and event-time data as well through appropriate specification of the link function (logit or log, respectively). Although we have described the conceptual approach as a multi-stage model, there are compelling advantages to fitting the three (or more) equations simultaneously, with several different error terms. In this way, the model described above becomes:

$$Y_{cit} = a_{ci} + (\beta_0 + \beta X_c + e_c'' + e_{ci}')\, t + e_{cit} \qquad (3.6.4)$$

where a_{ci} might be treated as either a fixed or random effect and e_c'', e_{ci}' , and e_{cit} are the three random effects. These calculations are implemented in SAS procedures MIXED and GLMMIX, as well as more specialized programs like ML_n. In the remainder of this report, we refer to this as a "multi-level mixed model."

To focus the analysis on the effects of air pollution on rates of change, we have treated the person-specific intercepts a_{ci} as fixed effects so as to avoid having to make any distributional assumptions about them. However, in some analyses, we have also been interested in the effects of ambient air pollution on levels of lung function (e.g., at baseline or at the end of the study). A convenient way to accomplish this is to treat both the slopes and intercepts as random effects. This simply entails supplementing equations (3.6.2) and (3.6.3) by additional equations of the same form for the a_{ci} and community-mean intercepts A_c. Details can be found in Berhane et al. (in press) and the variants of this approach that have been applied in specific analyses are described in the Results section below. Other analyses have been applied to cross-sectional data (typically the baseline observations). In this case, Eq. (3.6.1) is not used, and the resulting two-level model is given by Eqs. (3.6.2) and (3.6.3), taking as b_{ci} the baseline observation for child i in community c.

When applied to binary outcome data, the first level of the model is simply replaced by a logistic model of the form

$$\text{logit Pr}(Y_{cit} = 1) = a_{ci} + b_{ci}\, t$$

(plus additional time-specific covariates as needed). This approach is appropriate for repeated binary data, such as individual school absences, as described in Rondeau et al. (submitted). For the analysis of incidence, however, we adopt a multi-level Cox regression approach, as in (Ma et al. 2003). Here, the basic model is of the form

$$\lambda_{ci}(t) = \lambda_0(t) \exp(\alpha Z_{cit} + \beta X_c + A_c)$$

where Z_{cit} denotes person- and time-specific covariates and Ac is a random effect for community. Finally, in some analyses, we have aggregated over subjects or over times, as in the analyses of school absences (Gilliland et al. 2001b; Berhane and Thomas 2002). For a time series analysis of daily absences as function of daily air pollution, the basic model takes the form

$$Y_{ct} \sim \text{Poisson}[R_{ct} \exp(s_c(t) + A_c + \beta X_{ct})]$$

where Y_{ct} denotes the total number of incident absences in community c on day t, R_{ct} the corresponding number "at risk," $s_c(t)$ is a community-specific smooth function of time, A_c a random effect, and β is the log-relative risk parameter of interest. On the other hand, for analysis of the chronic effects of air pollution on individuals' absence rates, we use a Poisson regression model of the form

$$Y_{ci} \sim \text{Poisson}[R_{ci} \exp(A_c + \alpha Z_{ci} + \beta X_{ci})]$$

where Y_{ci} now denotes the total number of absences for child i in community c, and R_{ci} an expected number of absences based on the days he or she was at risk and the community average daily absence rates on these days.

To allow for proper examination of the lag structure of the pollution effects, we also used the polynomial distributed lag approach (Almon 1965). This approach allows for the possibility that the effect of pollution levels on any given day may persist for several subsequent days, i.e., health outcomes on any given day are influenced by pollution levels of the preceding several days. Hence, the model takes the form

$$Y_{ct} \sim \text{Poisson}[R_{ct} \exp(s_c(t) + A_c + \Sigma_k g_{ck} \Sigma_j (X_{c(t-j)} - Xc) j^k]$$

where $k=0, ...,D$ (the degree of the polynomial) and $j=1, ..., L$ (the maximum lag considered). In this model the effects of the several lagged exposures are assumed to follow a polynomial structure. This makes the model parsimonious and overcomes the problem of collinearity that would arise by simple additive inclusions of the several lagged exposure terms in the model.

The basic multilevel linear model described above is appropriate for relatively short periods of time (e.g., the first four years of follow-up), over which growth rates are approximately constant. When applied to longer periods of follow-up (e.g., the entire 8-year period of the study for cohorts C and D), we have found it necessary to allow for nonlinear growth curves. Various approaches to this problem are illustrated in the results below and extensive discussion of the statistical methods can be found in (Berhane in preparation-a; Berhane et al. in press). Because it may not be sufficient to simply add additional quadratic or cubic terms to the first level model described above, we have adopted a flexible modeling approach based on cubic splines (DeBoor 1974). In essence, this entails replacing the first level model (3.6.1) by

$$Y_{cit} = A_c + s(t,\mathbf{B}_c) + s(t,\mathbf{b}_{ci}) + e_{ci}$$

where $s(t,\mathbf{b})$ is a cubic spline function with parameters $\mathbf{b}$. The first of these splines describes the average growth curve for community c and is modeled in relation to air pollution in the third level; the second spline is included simply to allow for nonlinear deviations in growth rates between children within communities owing to their different ages at maturation. From the community average parameter estimates $\mathbf{B}_c$, we can compute various "functionals" $\mathbf{B}_c^*$, representing quantities of interest such as the maximum attained lung function, the age at peak growth rate, and the peak growth rate itself. These quantities are then taken as the dependent variable in Eq. (3.6.3).

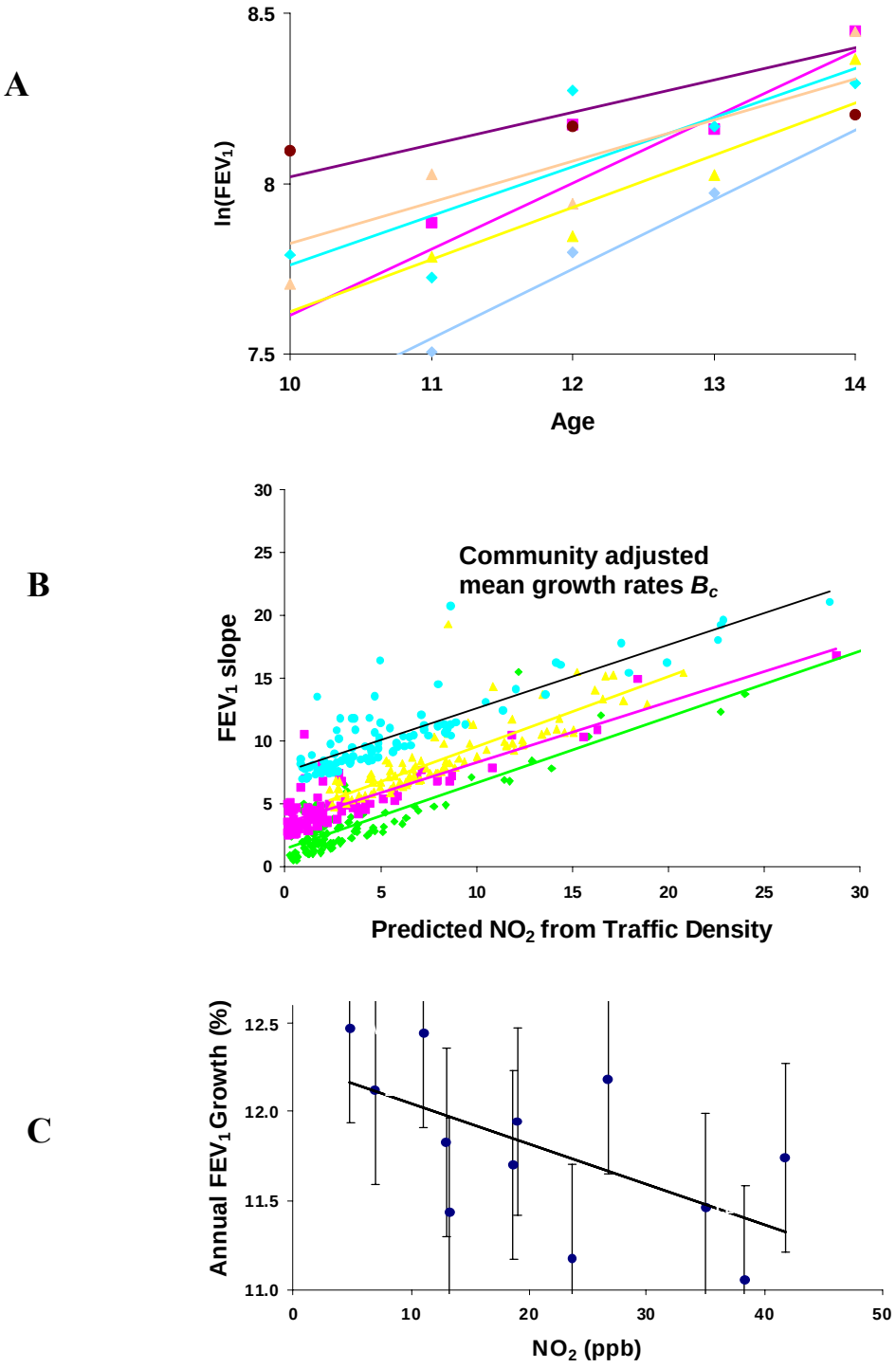


Figure 3.6-1. Hypothetical example of the three-level model.

Panel A: repeated individual FEV_1 measurements for a sample of children from a single community, plotted as a function of age; child-specific slope estimates are taken to level 2. Panel B: individual growth rate estimates for a sample of communities, plotted against one covariate; community-specific mean covariate-adjusted growth rates are then taken to level 3. Panel C: community-mean growth rates plotted as a function of ambient air pollution; this regression forms the basis of the inference of the effect of ambient air pollution on lung function.

3.7. Quality Assurance/Quality Control

The quality assurance (QA) and quality control (QC) procedures developed for and applied to the Children's Health Study (CHS) utilized an overlapping set of QA/QC elements for application to the three component areas of the study (exposure assessment, health assessment, and data management). A complete description and discussion of the procedures have been presented in greater detail in previous study documentation (including the Study Quality Assurance Plan, the Phase I Final Report, the Phase II Final Report, and a separate-cover Appendix A on Exposure Assessment Methodology) and in a published manuscript (Bush and Taylor 1999). In addition, several published manuscripts and internal reports have been prepared on various QA/QC aspects of this project (for example, seeLinn et al. 1996; Linn et al. 1998; Gilliland et al. 1999a; Bowers and Taylor 2000; Enright et al. 2000; Taylor 2003). An overview of the study-critical QA/QC perspective is highlighted in the ensuing discussion, whose purpose is to illustrate the breadth and range of the overall process and programs in place.

3.7.1. QA/QC Overview

As previously mentioned, the quality assurance (QA) and quality control (QC) aspects of the epidemiologic study involved several different components. Written protocols for the study designs, approaches, and procedures were developed. Standard Operating Procedures (SOPs) were prepared and approved for study use. Criteria for personnel qualification (for example, for working with subjects and testing in schools) were established and followed. Written records - most typically in the form of logbooks, chain-of-custody field data collection sheets, or electronic computer transaction files - were developed to document various aspects of the project. A carefully maintained data management system was developed and used, with discrete data validation procedures, editing, and data releases. A standardized and systematic external quality assurance review procedure was instituted for each of the three study components.

Collection of ambient air quality information required the establishment, deployment, and operation of assorted air monitoring instrumentation for both real-time and time-integrated data quantification. Real-time data acquisition involved operation of a number of commercially available analytical monitors, in strict conformance with state and federal operating guidelines. Exposure-related time-integrated data collection typically involved filter or denuder sample collections, which involved the operation and use of chemical and analytical support laboratories, each of which adhered to written standard operating protocols for the media preparation, handling, and analyses being performed. These components are collectively referred to as the Exposure Assessment portion of the study.

The Health Assessment component of the study involved collection of various health outcomes information through physical evaluations, clinical interviews, questionnaire self-administrations, and reviews of medical absence records from participating schools.

Data from exposure and health assessments were collated into a master study database, whose access, validation, editing, storage, and use involved an additional number of discrete QA/QC applications.

For each of the study elements, a written compendium of procedures employing standard operating protocols (SOPs) was developed and utilized to document study methods, provide for

consistent performance in multiple study locations, and create a standardized training resource for new, returning, or re-assigned staff. Specific training and certification of qualified personnel assured that study operations were consistently performed by staff trained and experienced in the appropriate SOPs. Relevant information collected through the exposure, health, or data management portals of study entry was preserved in either hard-copy or electronic format (or both). When electronic data records were utilized, transaction files were maintained to document any validation, editing, and cleaning procedures applied to achieve an evolved and cleared data set for analytical use. Details of electronic data manipulations, from raw data to final versions, were specified in a data management plan maintained by the study data management team.

A considerable level of effort was devoted in the study to the use of multiple layers of quality assurance oversight, including the routine use of external study auditors. As the primary ambient air quality data manager for the study, the ARB staff compiled and maintained the SOPs. Updates and changes in the SOPs were reviewed by the CHS QA officer in a timely manner.

3.7.2. Exposure Assessment

Exposure assessment activities for the study initially involved site selection, instrument development and deployment, site preparation, field operations, and data collection. For the large majority of study time, field operations (data acquisition) and data collection were the primary foci of exposure assessment. (It should be noted that some sampling site relocations did, in fact, occur over the project period of performance. These are discussed in the Exposure Assessment section of the report, Section 3.4.1.1).

Air monitoring data collected in the study underwent three levels of review. The first level of review was performed by the field equipment and laboratory operators on a monthly basis. Calibration data were reviewed and appropriate calibration factors were applied to the data. The data were reviewed for consistency with the SOPs and invalidated if any obvious operational or laboratory problems occurred. All of these "QA Level 0.5" data were transferred from the collecting agencies and laboratories to the ARB in Sacramento.

The second level of review occurred after the data from all locations were merged. ARB staff reviewed the continuous data and integrated data for consistency across the network. ARB reviewed and documented the reasons data were missing or invalid, and often explored ways to fill in data gaps where feasible. ARB staff delivered their "QA Level 1"data to STI annually for the third level of review.

At STI, the data were checked to assure that they were within expected ranges of concentrations, with no anomalous persistent values or significant shifts in the instrument baseline concentrations. Drifting baseline concentrations of ozone, NO_x/NO, and PM_{10} were adjusted when baselines were greater than ±4 ppb for gases and ±5 $\mu g/m^3$ for PM_{10}. Time series plots of all the hourly parameters measured at each site were examined for consistency of relationships between pollutants. Concentrations from each new year were compared against concentrations from previous years on a monthly and annual basis to identify possible anomalies. Average diurnal concentration profiles were compared to data from prior years. Potentially anomalous CHS station data were compared to data collected at nearby stations for consistency when ever possible.

A significant effort was made to estimate missing concentrations as part of the third level of review and processing. Single hours of missing data were estimated as the average of the concentrations measured during the hour preceding and following the missing hour. When nearby stations had valid measurements for periods with missing CHS data, hourly concentrations were estimated using regression relationships between pollutant concentrations at a specific hour of the day at a CHS site and a nearby site. This method was primarily used for ozone and NO_2 rather than PM_{10}. No attempt was made to fill in missing two-week average acid and $PM_{2.5}$ concentrations and hourly CO and CPC concentrations due to the lack of adequate nearby measurements.

Data quality codes were assigned to each hourly and two-week average measurement based on the overall quality assurance assessment. The codes indicated whether the data were valid (code=1), estimated (code = 2 or 3), suspect (codes = 6 or 7), or invalid (code = 9). Daily, monthly and annual average values were computed using all valid, estimated, and suspect data.

3.7.3. Health Assessment

On the health assessment side, a rigorous, analogous plan involving overlapping and systematic QA oversight, audits, and reporting was routinely applied. Field testing personnel were initially trained by project investigators, in conformance with pre-established testing protocols. Technicians were required to demonstrate testing proficiency to the health research team prior to being given clearance for actual subject testing. In the field, a daily calibration routine was performed prior to and following every day of lung function testing on every spirometer used for testing. In addition to pre-and post-calibration, re-calibration during any given day's testing day was performed if and when room temperature in the location of lung function tested changed by more than two degrees Celsius from the previous calibration.

Within the lung function testing software, additional QA/QC evaluations were present which required the testing technician to objectively grade subject test performance for each attempted test effort. All testing results were logged into the data set and subsequently reviewed by a trained external QA officer. A specifically designed software program, originally developed at the University of Arizona for Lung Health Study, was utilized by the external Health QA Officer to evaluate technician testing performance and to provide regular specific written guidance for (a) improvement of personal testing performance, and (b) consistency of testing approach with the entire group of six testing technicians. In addition to the computerized and off-line review of lung function testing results, unannounced on-site visits were also routinely performed by the Health QA Officer (in each of the 12 CHS communities) to observe technician performance in the field.

Back-up copies of all lung function testing results were routinely created for each testing system at the end of each testing day, and these floppy disk copies were returned to the data management team at USC. (Each testing system in the field also maintained an updated lung function results data file in the respective resident hard drive of the computer assigned to each testing spirometer).

Questionnaires completed by study participants were reviewed in the field by research staff and returned to the data management group at USC. At USC, each questionnaire was reviewed for

completeness and intra-questionnaire consistency prior to processing. A sub-set of previously identified critical questions were evaluated on each questionnaire, and if responses were missing or judged to be inconsistent, the written responses were reviewed by the health research team and a decision was made to contact the subject in question for a brief telephone interview to clarify the written response.

Assessment of the overall health field testing protocol and performance was performed annually by Dr. Paul Enright of The University of Arizona, an international expert in lung function testing quality assurance. On-site observation with interactive comments was typically followed by a group discussion reviewing observed performance and recommended modifications, if and as they were observed.

Absence monitoring data from reporting schools were reviewed by project staff and entered into a standardized database. Entries were double-entered and compared for consistency. A subset of subjects' parents was re-contacted following telephone interviews to confirm the information originally provided during previously conducted interviews.

3.7.4. Data Management

In the data management area, raw data was submitted daily or bi-weekly from the health testing, and from the exposure assessment section following data review by the exposure assessment sub-contractor. Data submissions were evaluated for adequate/accurate subject identifiers, data completeness, and out-of-range values. A series of software programs and some manual reviews were performed with incoming data. Any changes or modifications to data were documented in a data transaction file. Copies of historical databases were maintained and archived on a controlled-access server. Updated data releases, containing corrected, imputed, and more complete data sets, were periodically provided to study investigators. When this was done, each dataset was carefully identified by version and release date.

4. Results

4.1. Exposure Assessment

4.1.1. Ambient Air Quality

4.1.1.1. Spatial Patterns

The CHS communities were selected to take advantage of naturally occurring differences in community air pollution profiles for the four ambient air pollutants (ozone, particles, nitrogen dioxide, and acids) of study health interest. Maximum differences in ambient pollution levels and pollutants mix were achieved by selecting communities across a 300-km x 400-km area in southern California. The locations of the communities, shown in Figure 4.1-1, extend from Atascadero in southern San Luis Obispo County to Alpine, in southeastern San Diego County. Lompoc, Santa Maria, and Atascadero are small towns north of Los Angeles with relatively few industrial air pollution sources. Long Beach is a densely populated coastal city with a major shipping port, extensive industrial sources, and a high volume of traffic. San Dimas, Upland, Mira Loma, and Riverside are cities located east of Los Angeles in a downwind region known for adverse air pollution. Lake Arrowhead is a small community in the mountains east of Los Angeles at an elevation of 5,200 ft. Lancaster (elevation 2600 feet) is located in the high desert area north of Los Angeles, and Lake Elsinore is located in the low desert southeast of Los Angeles.

4.1.1.1.1. Spatial Patterns of Gaseous Pollutants

The spatial pattern of ambient ozone concentrations in the 1994-2001 study period is illustrated in Figure 4.1-2 and Table 4.1-1 and Table 4.1-2. Lake Arrowhead stands out as having substantially higher ozone concentrations than the other CHS communities and almost all other communities in the United States. It ranks highest for all ozone exposure metrics, with an eight-year average 1-hr daily maximum concentration of 81 ppb. Exceptionally high ozone levels occur at Lake Arrowhead because it is located at the optimum distance downwind of the SoCAB high emissions area for "same day" photochemical production of ozone and at an elevation where nighttime ozone scavenging by urban NO_x emissions is minimized by meteorological decoupling of the polluted surface layer (lower elevation) from the aloft layer (higher elevation). Conversely, Long Beach, Lompoc, and Santa Maria have consistently low ozone levels compared to other communities, with eight-year average 1-hr daily maximum concentration of 36 to 42 ppb. These communities are located in coastal areas that typically have lower ozone because they are in or upwind of the NO_x and VOC source areas. The 1-hr daily maximum ozone concentrations in Riverside, Mira Loma, Upland, San Dimas, Lake Elsinore, and Alpine (all inland communities) were higher, on average, than the 12-community mean ozone level. These communities are all located downwind of significant urban NO_x and VOC emission areas. Lancaster and Atascadero had levels slightly below the 12-community mean 1-hr maximum ozone, on average. The spatial pattern of 10 AM-to-6 PM, 8-hr average ozone levels was very similar to the pattern of 1-hr maximum ozone levels; however, the pattern of 24-hr average ozone levels was quite different. Because ozone is scavenged at night by fresh NO emissions in

the larger urban areas, inland urban communities such as San Dimas, Upland, Mira Loma, and Riverside have above-average daytime ozone levels, but very low overnight ozone levels. This phenomenon results in below-average 24-hr ozone levels. Lake Elsinore, Lancaster, Alpine, and Lake Arrowhead experienced above-average 24-hr ozone levels. The 1994-2001 patterns of ozone exposure were similar to those expected and predicted at the start of the study, with the most notable difference being lower than expected ozone levels in San Dimas and Upland, especially in the cooler seasons.

Ambient concentrations of NO_2 were highest in Upland, where an eight-year average concentration of 38 ppb was observed during the CHS (see Figure 4.1-2 and Table 4.1-1). San Dimas and Long Beach recorded the next highest NO_2 levels on average (~34 ppb), and Mira Loma and Riverside had above average NO_2 levels (26 ppb). All of these communities are in and/or downwind of the high NO_x emission areas in southern California. In contrast, average NO_2 levels in Lompoc were an order of magnitude lower than the highest community, Upland (4 ppb vs 38ppb, respectively). This ten-fold spatial variation in NO_2 across the network was much larger than the three-fold spatial variation observed for ozone. Lake Arrowhead, Santa Maria, and Alpine had moderate NO_2 levels, on average. The different metrics for NO_2 exposures, such as the 1-hr daily maximum, 8-hr daily maximum, and 6 AM-to-6 PM average concentrations, were highly correlated with the 24-hr average value at each site. Spatial patterns of the other NO_2 metrics were very similar to the 24-hr spatial pattern. The range of concentrations and spatial pattern of NO_2 observed during the CHS were in close agreement with the design expectations.

The distribution of NO concentrations was somewhat different than that observed for NO_2. The highest NO concentrations (~39 ppb) occurred in San Dimas and Long Beach. Mira Loma, Riverside, and Upland had relatively high levels (18 to 28 ppb), on average. The communities with high NO also had relatively high NO_2 and traffic densities. Lompoc had low NO levels (2 ppb) because few nearby NO_x sources existed. Lake Arrowhead and Alpine have low NO levels (1 to 2.4 ppb) because they are located significant distances downwind of the major NO_x emissions areas and the relatively high ozone levels efficiently scavenge NO in these communities. In communities with moderate and high NO, the station location relative to major roadways may have significantly influenced the observations. For example, the average daily 1-hr maximum NO concentration measured at the Glendora station was six times lower than the NO measured at the San Dimas station, even though they are within a few miles of each other and are both "in the same community." NO_2 was only 18% lower at the Glendora station, compared to levels assigned for San Dimas. Therefore, between-community NO comparisons should be interpreted cautiously.

The spatial pattern of formic acid concentrations was almost identical to that of NO_2 with the highest concentration occurring at Upland and the lowest concentration at Lompoc (Figure 4.1-3 and Table 4.1-1). Upland and San Dimas had an eight-year mean concentration of 2.2 ppb formic acid, compared to 0.3 ppb at Lompoc. The 12-community mean concentration was 1.1 ppb; thus there was a seven-fold relative difference across the network. The communities of Mira Loma, Riverside, Lake Elsinore, and Long Beach had above-average levels of formic acid, while Lancaster, Alpine, Lake Arrowhead, Atascadero, and Santa Maria have below-average

levels. Ambient sources of formic acid are not well-characterized, but the similarity of the spatial distribution of formic acid and that of NO_2 suggests it may be emitted by similar sources.

Acetic acid was the most abundant gas-phase acid observed in the CHS (3.6 ppb 12-community mean concentration). The spatial pattern of acetic acid concentration was between that for ozone and NO_2, suggesting there may be primary and secondary (photochemical) sources of this weak organic acid. The highest concentrations were observed in Mira Loma, where the eight-year mean value is 5.9 ppb (Figure 4.1-3 and Table 4.1-1). The nearby communities of Riverside, Upland, and San Dimas had relatively high levels (4.5 to 5.3 ppb) on average. Lompoc had the lowest level (1 ppb), followed by Santa Maria at 1.3 ppb. Atascadero, Lake Arrowhead, Alpine, and Long Beach had below average acetic acid concentrations, while Lancaster and Lake Elsinore had levels slightly above average. There was almost a six-fold relative difference across the network.

Nitric acid was the second most abundant acid observed in the CHS. The pattern of ambient nitric acid concentrations, shown in Figure 4.1-3 and Table 4.1-1, is similar to that of acetic acid. The highest levels (3.9 ppb eight-year average) occurred in Upland and San Dimas. The nitric acid concentrations were above the 12-community mean of 2.2 ppb in Riverside, Mira Loma, Lake Elsinore, and Long Beach and slightly below the 12-community mean in Lake Arrowhead, Alpine, and Lancaster. In contrast, the average nitric acid concentration was only 0.4 ppb in Lompoc and less than 1 ppb in Santa Maria and Atascadero. There was a ten-fold relative difference in nitric acid across the network, as was the case for the nitric acid precursor, NO_2. The spatial pattern of nitric acid is similar to the design expectation except for Riverside and Mira Loma, where higher-than-expected nitric acid levels are observed. Decreases in ammonia emissions from dairy operations in the inland valleys between 1986 and 1994-2001 are the most probable cause of higher nitric acid levels in these communities.

Finally, ambient hydrochloric acid levels were more spatially uniform and lower than other measured acids at all sampling locations. The 12-community mean concentration was 0.4 ppb (±0.15). Atascadero, Lancaster, and Lake Arrowhead had the lowest levels on average (<0.25 ppb), and Long Beach, San Dimas, and Upland had the highest levels (0.53 to 0.74 ppb). The reported levels may be biased high because the concentrations were often below the limit of detection during periods with good air quality. Industrial sources and chemical reactions of the marine aerosol with nitric acid are probably responsible for the high hydrochloric acid levels in Long Beach. The reaction of sea salt with nitric acid is also probably responsible for Lompoc having higher hydrochloric acid levels than the other clean communities located farther inland.

4.1.1.1.2. Spatial Patterns of Particles

The spatial pattern of airborne particles was quite different from that observed for ozone, but similar to that observed for NO_2 and gaseous acids. The observed ambient particle spatial distribution varied somewhat with particle size and chemical composition. The long-term (eight-year) average PM_{10} mass concentrations shown in Figure 4.1-2 and Table 4.1-3 and Table 4.1-4, indicated higher concentrations within the South Coast Air Basin (SoCAB) than in the air basins to the north and south. The highest PM_{10} concentrations occurred at Mira Loma, where the 67 $\mu g/m^3$ eight-year average PM_{10} concentrations exceeded the 50 $\mu g/m^3$ PM_{10} annual average National Ambient Air Quality Standard (NAAQS). The PM_{10} peak at Mira Loma was 60%

higher than observed levels at the nearest upwind and downwind communities, Upland and Riverside, where PM_{10} levels were 41 $\mu g/m^3$ on average. Long Beach, San Dimas, and Lake Elsinore also had above-average PM_{10} levels, while Lancaster and Alpine had levels slightly below the 12-community mean of 32 $\mu g/m^3$. The lowest PM_{10} concentrations occurred at Lompoc (15 $\mu g/m^3$). Relatively low concentrations (20-21 $\mu g/m^3$) also occurred at Santa Maria, Atascadero, and Lake Arrowhead. The spatial variation in long-term PM_{10} concentrations between CHS communities was four-fold across the network. The spatial pattern of PM_{10} was entirely consistent with the CHS study design.

The spatial distribution of $PM_{2.5}$ mass concentrations, shown in Figure 4.1-2, Table 4.1-3 and Table 4.1-4, were generally similar to the PM_{10} pattern. The highest $PM_{2.5}$ mass concentrations were observed at Mira Loma (28 $\mu g/m^3$ average), while the lowest concentrations were observed at Lompoc (5.6 $\mu g/m^3$). Riverside and Upland, with long-term $PM_{2.5}$ concentrations of 22 $\mu g/m^3$, and San Dimas and Long Beach, with long-term $PM_{2.5}$ concentrations of 18 $\mu g/m^3$, exceeded the $PM_{2.5}$ annual NAAQS level of 15 $\mu g/m^3$. Lake Elsinore recorded $PM_{2.5}$ levels (12.9 $\mu g/m^3$) slightly below the 12-community mean of 14 $\mu g/m^3$. Santa Maria, Atascadero, Lake Arrowhead, and Lancaster had relatively low $PM_{2.5}$ levels (7 to 8.5 $\mu g/m^3$). There is a five-fold relative difference in long-term $PM_{2.5}$ mass concentrations across the CHS communities.

The two most abundant chemical constituents identified in collected $PM_{2.5}$ were nitrate and organic carbon (OC). The 12-community mean concentrations were 4.5 $\mu g/m^3$ of nitrate and 4.7 $\mu g/m^3$ of OC. If one considers that nitrate has ammonium or sodium associated with it (which adds 30% of the nitrate mass) and that the OC has hydrogen and oxygen associated with it (which typically adds 40% of the OC mass), the combined contribution to $PM_{2.5}$ mass was over 80% (on average) in the CHS communities. Not surprisingly, the spatial distributions of nitrate and OC were very similar to that of $PM_{2.5}$ mass with the highest concentrations occurring at Mira Loma and the lowest concentrations occurring at Lompoc, as shown in Figure 4.1-3 and Table 4.1-3. The $PM_{2.5}$ nitrate levels varied from 0.76 $\mu g/m^3$ at Lompoc, where nitrate was 13% of $PM_{2.5}$ mass, to 11.5 $\mu g/m^3$ at Mira Loma, where nitrate was 40% of $PM_{2.5}$ mass. Nitrate levels were relatively high (6.7 to 9 $\mu g/m^3$) in Riverside, Upland, and San Dimas where the nitrate fractions of $PM_{2.5}$ mass were in the 35% to 40% range. Nitrate was relatively low (1.3 to 2.2 $\mu g/m^3$) on average in Santa Maria, Atascadero, Lancaster, Lake Arrowhead, and Alpine. As with NO_2 and HNO_3, there is about a ten-fold variation in nitrate between CHS communities.

A striking feature of the $PM_{2.5}$ OC distribution was the extremely high concentrations observed in Mira Loma (12 $\mu g/m^3$). The long-term average Mira Loma OC concentration was twice as high as the next highest set of communities (Upland, Riverside, and San Dimas). OC clearly contributed significantly to the high $PM_{2.5}$ mass in Mira Loma. The source of the excess OC in Mira Loma is not likely to be secondary formation because this would influence all of the high PM inland sites (e.g., Upland, Riverside, San Dimas, and Mira Loma) in a similar manner. Thus, the excess OC is likely to be associated with primary sources immediately near Mira Loma. The lowest concentrations of OC occurred in Lompoc (1.4 $\mu g/m^3$), and relatively low OC concentrations were measured in Lake Arrowhead (2.5 $\mu g/m^3$), Santa Maria (2.6 $\mu g/m^3$), and Alpine (2.7 $\mu g/m^3$).

Ambient $PM_{2.5}$ sulfate concentrations were generally low in the CHS communities, and the spatial pattern was mostly similar to that of $PM_{2.5}$ mass, as shown in Figure 4.1-3 and Table 4.1-3. The eight-year 12-community mean sulfate concentration was only 1.7 $\mu g/m^3$. Atascadero had the lowest sulfate, while Long Beach had the highest sulfate loadings. The CHS sulfate concentrations occurred in three distinct groups. Atascadero, Lancaster, Lake Arrowhead, and Lompoc had similarly low concentrations (0.8 to 1 $\mu g/m^3$); Santa Maria, Alpine, and Lake Elsinore had moderate sulfate (1.3 to 1.8 $\mu g/m^3$); and Riverside, Upland, San Dimas, Mira Loma, and Long Beach had higher sulfate (2.3 to 2.8 $\mu g/m^3$). Santa Maria had slightly higher sulfate levels than other clean communities due to industrial activity in the area.

Lastly, the $PM_{2.5}$ elemental carbon (EC) concentrations showed a ten-fold variation across the CHS communities. As with most other $PM_{2.5}$ constituents, the lowest and highest EC levels were observed at Lompoc (0.12 $\mu g/m^3$) and Mira Loma (1.3 $\mu g/m^3$). However, the EC concentrations in Mira Loma were not exceptionally high compared to other communities. Long Beach had EC levels equal to those in Mira Loma, and Upland, San Dimas and Riverside had slightly lower EC levels (1 to 1.2 $\mu g/m^3$). Lake Elsinore, Lancaster, Alpine, Atascadero, Lake Arrowhead, and Santa Maria had below-average EC levels (0.3 to 0.7 $\mu g/m^3$). The spatial pattern of EC was generally similar to those of $PM_{2.5}$ mass and NO_2.

4.1.1.1.3. Spatial Correlation of Pollutants

The spatial correlation of the eight-year average air pollutant concentrations is illustrated in Figure 4.1-4 and Table 4.1-5. Long-term average 24-hr average ozone concentrations were not correlated with any of the other pollutants. Ozone was noticeably anti-correlated with NO (r=-0.61), NO_2 (r=-0.56), EC (r=-0.51), sulfate (r=-0.50), and hydrochloric acid (r=-0.50). All other pollutants are correlated. NO was strongly correlated with NO_2, EC, and formic acid. NO_2 was spatially correlated with all species except ozone, including formic acid (r=0.95), EC (r=0.94), nitric acid (r=0.85), sulfate (r=0.87), and $PM_{2.5}$ mass (r=0.81). PM_{10} mass was highly correlated with $PM_{2.5}$ mass (r=0.95) and most of the $PM_{2.5}$ chemical constituents (e.g., with OC [r=0.97] and NO_3 [r=0.94]). PM_{10} was also strongly correlated with the gas-phase acids, especially acetic acid (r=0.86). The correlation between $PM_{2.5}$ mass and the measured PM constituents was high—the correlation coefficient ranging from unity for ammonium and 0.99 for nitrate to 0.86 for sulfate. $PM_{2.5}$ mass was also correlated with formic, acetic, and nitric acid (r>0.8), and these three acids were highly correlated with one another.

The collinearity in air pollutant concentrations results from co-emissions by many sources (similar emission sources) and by common atmospheric chemical reactions. For example, NO_x and EC are co-emitted in motor vehicle exhaust and are highly correlated (r=0.94) in ambient air. Likewise, the high correlation of NO, NO_2, nitric acid, and nitrate (r>0.77) is understandable given that nitrate is formed from nitric acid, nitric acid is formed from NO_2, and NO_2 is formed from NO. Thus, only two truly independent pollutant profiles in the CHS were identified: (1) ozone and (2) a complex mixture of NO_2, acids, and particles.

4.1.1.2. Temporal Air Quality Trends

Population growth, industrial expansion, implementation of air pollution control programs, and meteorology may influence year-to-year changes and long-term trends in air quality. Southern

California has one of the most successful air pollution control programs in the world, and pollutant levels were expected to decline over the course of the CHS. The 1994-2001 period included some years with anomalous weather conditions, including adverse conditions that produced high PM in 1995 and favorable conditions associated with "El Nino" that reduced PM in 1997-1998.

The trends in annual average ozone and NO_2 at stations within and outside the SoCAB are shown in Figure 4.1-5. Table 4.1-6 lists the species and communities for which statistically significant (at the 95% confidence interval) temporal trends were observed from 1994-2001. The annual average 10 AM-to-6 PM ozone concentrations showed statistically significant decreases over time in the seven communities with the highest ozone levels (Lake Arrowhead, Riverside, Mira Loma, Upland, San Dimas, Lake Elsinore, and Alpine). The annual ozone concentrations in the five cleaner (lower ozone) communities did not show notable year-to-year variations or trends. The downward trend in daytime ozone is stronger and more widespread than trends in any other pollutants monitored during the CHS. The decline in ambient ozone in communities located within the SoCAB is also consistent with the SCAQMD's air quality management plan (AQMP), which focused on reducing 1-hr maximum ozone concentrations during this period. Additionally, the extreme peak ozone readings in the SoCAB were reduced during the CHS, such that the ambient ozone concentrations achieved on the worst air quality days of the year were lowered, compared to previous years. In contrast, annual average NO_2 concentrations did not show strong declining trends in most communities. In fact, after declining in the early years, NO_2 in most communities increased in 1998-2000. The exceptions were in Long Beach, where NO_2 did show a statistically significant downward trend, and in Upland, where the downward trend was almost statistically significant. No other communities showed NO_2 trends (except Lompoc where NO_2 concentrations are so low that the trend is not meaningful).

The temporal trends in the annual average $PM_{2.5}$ and PM_{10} concentrations are shown in Figure 4.1-6. Modest year-to-year changes were observable, but no overall trend was evident. At Mira Loma, where the highest PM levels occurred, the concentrations at the beginning and end of the period were virtually the same, even though the $PM_{2.5}$ was lower in most of the intermediate years and PM_{10} vacillated in the intermediate years. The trends in PM_{10} were statistically significant only in San Dimas (decreasing) and Santa Maria (increasing), two sampling locations potentially confounded by station relocations during the sampling period. The trends in $PM_{2.5}$ were statistically significant in Atascadero (increasing) and Lake Arrowhead (decreasing), which may not be meaningful given the low levels of $PM_{2.5}$ in those communities.

$PM_{2.5}$ chemical constituents showed temporal patterns similar to those of $PM_{2.5}$ mass. Figure 4.1-7 and Figure 4.1-8 depict NO_3, SO_4, OC, and EC at most sites within the SoCAB, which declined from 1994 through 1998, and then essentially returned to circa 1994 levels by 2001. The steep decline in PM component concentrations from 1995 to 1998 in the SoCAB was similar to the decline in ozone concentrations. Temporal trends were not evident in communities outside of the SoCAB during this period.

Annual average concentrations of nitric acid and formic acid revealed either flat or downward trends during this period, while concentrations of acetic acid increased in all communities, as shown in Figure 4.1-9. Formic acid concentrations decreased by up to 50% between 1995 and

1998 in some communities and remained near the reduced levels through 2001. The formic acid trends were quite similar to those of ozone. The relative decreases of nitric acid were smaller than those of formic acid. If one combined the decreasing formic acid data with the increasing acetic acid data, no trend in total organic acid was evident. The downward trend in nitric acid concentrations was statistically significant in Long Beach, San Dimas, Lake Arrowhead, and Lompoc. Likewise, the declining formic acid trend was statistically significant in Long Beach, San Dimas, Upland, Riverside, and Mira Loma which are the same sites for which ozone showed significant downward trends. The acetic acid increases were statistically significant mostly in the communities with lower primary emissions pollution levels (Atascadero, Santa Maria, Lancaster, Lake Arrowhead, and Alpine).

4.1.1.3. Seasonal and Diurnal Patterns

The seasonality of air pollution in the CHS was clearly evident in the monthly data. The monthly average ambient concentrations of ozone (1-hr maximum), NO_2, NO, $PM_{2.5}$, and PM_{10} mass are shown in Figure 4.1-10 through Figure 4.1-14, respectively, for 1994-2001. These figures also illustrate the year-to-year variability in monthly average concentrations.

Ozone is formed photochemically and generally reaches its highest concentrations in the summer in urban areas. In the clean troposphere, ozone levels often peak in the spring when increasing sunlight combines with precursors that have built up over the winter to form the highest ozone concentrations of the year. Ten of the 12 CHS community have seasonal ozone profiles typical of urban areas, where ozone levels are near background levels in the winter and increase to their maximum levels in the June/August time period. Lompoc and Santa Maria have profiles similar to clean air sites, with small variations over the year and minor peaks in the spring and fall. Note, the ozone levels in December and January were near background levels in all of the communities, and increased steadily until July, declining thereafter (on average).

The seasonal pattern of NO_2 concentrations in most communities was "U-shaped," with minimum monthly average concentrations occurring between May and July, and maximum monthly concentrations occurring in November–January. This pattern was evident in 10 of the 12 communities and consistent with seasonal variation in meteorological conditions. Wind speeds, mixing heights, and solar radiation are lower in winter than in summer and lead to higher NO_2 concentrations in winter. NO_2 in the rural mountain communities of Lake Arrowhead and Alpine showed a different pattern, where concentrations were slightly higher in the summer and/or fall months than in the spring and winter months.

Like NO_2, the monthly average concentrations of NO were lowest in the summer and highest in the winter. Observed concentration differences between summer and winter were much larger for NO than NO_2, especially in the high traffic communities like Long Beach and Riverside where winter NO levels were an order of magnitude higher than summer levels. The seasonal pattern of NO was similar in all of the communities except Alpine and Lake Arrowhead.

Seasonal variations in $PM_{2.5}$ mass were not dramatic at most sites. The mountain communities of Alpine and Lake Arrowhead had low PM concentrations in winter and high concentrations in summer, much like ozone. Several communities with low or moderate $PM_{2.5}$ levels, such as Lompoc, Santa Maria, Lancaster, and Lake Elsinore, had similar levels of $PM_{2.5}$ in all seasons.

Most other communities tended to have higher levels in the September–December period than any other time of the year due to meteorological conditions. The conditions most conducive to $PM_{2.5}$ build up, slow air transport (stagnation), limited vertical mixing, and cool temperatures, occur most frequently in the fall. Woodsmoke is also more common in the fall and winter, expecially in Atascadero. In the highest reporting $PM_{2.5}$ site, Mira Loma, occasional high values were also reported in March through May. The data suggest that year-to-year fluctuations in monthly average $PM_{2.5}$ can be large and obscure the seasonality.

The seasonal patterns of PM_{10} concentrations were more consistent than those of $PM_{2.5}$. Most communities experienced lower PM_{10} levels in the spring and summer than in the fall and winter. The seasonal variations were not as large as those observed for ozone and NO. The higher PM_{10} values in the fall were consistent with dry dusty conditions in southern California during that time of year, which enhances re-suspended dust emissions. In Mira Loma, the community with the highest PM_{10} levels, PM_{10} concentrations tended to be high in every season except the spring.

The diurnal variations in annual average hourly concentrations of ozone, NO_2, NO, and PM_{10} are shown in Figure 4.1-15 through Figure 4.1-18, respectively. The ozone profiles showed maximum values between 1 PM and 3 PM (PST) in the afternoon in all communities due to daytime photochemical production. The shapes of the diurnal profiles in each community were quite consistent year-to-year. In fact, Long Beach and Atascadero had virtually identical profiles every year; the year-to-year variability in the average hourly concentrations was near the precision of the instruments (~3 ppb). Urban communities with significant NO_x emissions, such as Long Beach, San Dimas, Upland, Mira Loma, and Riverside, had low morning and nighttime values due to NO scavenging effects. The rural mountain communities of Lake Arrowhead and Alpine had much higher morning and nighttime ozone levels and experienced broader daytime peaks than the urban sites. Communities with generally low ozone levels, such as Lompoc and Santa Maria, had higher morning and nighttime ozone levels than the more urban communities. These patterns reinforce concerns over elevated ozone exposures to children who exercise outdoors in the afternoon.

The average diurnal patterns of NO_2 concentrations varied more by community and year than the ozone patterns. The NO_2 patterns in 10 of the 12 communities had morning and evening peaks and achieved their minimum concentrations in mid-day (12 to 2 pm PST). The timing and shape of the peaks were quite varied. Long Beach had a single late morning (10 am PST) peak and Lake Arrowhead a single evening (7 pm PST) peak. The shapes of the diurnal profiles in each community were consistent from year to year. NO_2 concentrations varied by a factor of 2, on average, over the diurnal cycle. The typical NO_2 diurnal pattern is most likely the result of the combination of the diurnal variations in meteorology and mobile source emissions. NO_2 concentrations tend to be high in the morning and evenings when traffic emissions are high and wind speeds and ventilation are low, in contrast to mid-day when faster transport and dispersion conditions result in lower NO_2 concentrations.

The highest NO concentrations occurred at 6–7 am (PST) in almost all communities. Early morning and mid-day NO levels were often much lower than the morning peak values. The average diurnal NO profiles for San Dimas and Glendora, shown in Figure 4.1-17 and Figure 4.1-19, illustrate the large differences in NO between a station located near an urban freeway

(San Dimas) compared to a suburban site ~4 km from the freeway (Glendora) in the same CHS community.

The diurnal profiles of PM_{10} were quite varied. Atascadero, Lancaster, Mira Loma, Upland, and Lake Elsinore typically experienced morning and evening PM_{10} peaks. Alpine and Lake Arrowhead had afternoon peaks, while Long Beach, Riverside, San Dimas, Lompoc, and San Maria had broad mid-day PM_{10} peaks of varying shapes. These diurnal differences will surely influence population exposure. Children exercising at mid-day in Lake Elsinore, for example, were likely to be exposed to levels lower than the "Lake Elsinore average PM_{10}", whereas children exercising at mid-day in San Dimas and Long Beach were likely to be exposed to levels higher than the daily averages for these communities. Mira Loma, the community with the highest average PM_{10} concentrations, experienced its high values from 6 pm to 8 pm (PST) in the evening. Lompoc, the community with the lowest PM_{10}, had similar values throughout the daytime period.

4.1.1.4. Special Studies

A number of special studies were performed to further characterize the nature of the air pollution profiles in the CHS communities. Air pollution data were collected for one or more years in most or all of the communities to obtain a "snapshot" of annual concentrations that could be used in cross-sectional analyses of health effects. Furthermore, if one ignores the generally small trends and year-to-year variability, the special study data can also be used in longitudinal analyses of health effects. The special studies reported in this section include (1) measurements of elemental concentrations, including trace metals, in $PM_{2.5}$ samples, (2) measurements of particle-phase organic compounds in PM_{10} samples and PM_{10} source apportionment, (3) measurements of carbon monoxide, and (4) measurements of particle number concentrations.

4.1.1.4.1. Elemental Concentrations

The ambient concentrations of 40 elements potentially detectable in $PM_{2.2}$ were measured throughout 2001 using two-week samplers. The annual average concentrations of the elements measured by XRF are listed in Table 4.1-7 and shown in Figure 4.1-20 and Figure 4.1-21. Only those concentrations that exceeded the analytical limit of detection, listed in Table 4.1-7, should be interpreted quantitatively. The XRF method, as run at DRI for these samples, was not able to reliably quantify the levels of sodium (Na), magnesium (Mg), phosphorus (P), gallium (Ga), arsenic (As), rubidium (Rb), yttrium (Y), palladium (Pd), silver (Ag), cadmium (Cd), indium (In), antimony (Sb), barium (Ba), lanthanum (La), gold (Au), mercury (Hg), thallium (Tl), and uranium (U) in most of the $PM_{2.5}$ samples. For the compounds with 1 to 5 ng/m^3 limits of detection, the results indicated these compounds were either not present or present in very small quantities. It must be noted that the limits of detection for the two lightest elements analyzed, sodium and magnesium, were much higher and these results should not be interpreted to suggest the elements were not present in the samples. Sodium and magnesium probably exist in most fine particles samples, based on other measurements of southern California aerosol composition (Watson et al. 1994).

The most abundant element (quantified by XRF) in fine PM collected was sulfur (S) across all CHS communities. Annual average sulfur levels ranged from 500 to 1300 ng/m^3, which was consistent with the sulfate concentration data. The next most abundant group of elements were

ones commonly present in mineral dust or crustal material: silicon (Si), iron (Fe), aluminum (Al), potassium (K), and calcium (Ca). The sum of these crustal elements was between 200 and 800 ng/m^3 on an annual basis across the communities. The next most abundant element in most communities was chlorine (Cl). Relatively high chloride ion concentrations were observed in coastal communities like Long Beach (112 ng/m^3) and Lompoc (57 ng/m^3), where smaller sea salt particles contributed chloride ions, and also in inland areas like Upland (59 ng/m^3) and Mira Loma (270 ng/m^3). Chlorine was below the 4.5 ng/m^3 limit of detection in Lake Arrowhead (1.7 ng/m^3). Iron was by far the most abundant trace metal, with annual concentrations ranging from 18 ng/m^3 in Lompoc to 187 ng/m^3 in Mira Loma. Zinc (Zn) was also as abundant as some of the crustal elements, such as aluminum, in many of the communities. Annual zinc concentrations ranged from 4 ng/m^3 in Lompoc to 33 ng/m^3 in Upland. Annual concentrations of titanium (Ti), vanadium (V), copper (Cu), and manganese (Mn) ranged from 1 to ~10 ng/m^3. The annual concentrations of other trace metals, such as chromium (Cr), cobalt (Co), and nickel (Ni), were less than 3 ng/m^3. The annual average lead (Pb) concentrations ranged from 1.2 ng/m^3 in Lompoc to 9 ng/m^3 in Glendora and Upland. There were also small but detectable amounts of selenium (Se), bromine (Br), strontium (Sr), zirconium (Zr), molybdenum (Mo), and tin (Sn) in the fine particles. The most abundant of these, tin, had annual concentrations ranging from 0.5 ng/m^3 in Santa Maria to 10 ng/m^3 in Glendora.

The extent to which spatial patterns of exposure of these elements differed from other indicators of fine particle pollution is potentially important for studying health effects. Sulfate, nitrate, ammonium, OC, and EC were all correlated with $PM_{2.5}$ mass (Table 4.1-5). A comparison of the total mass of elements measured by XRF as a group and $PM_{2.5}$ mass is presented in Figure 4.1-22. The coefficient of determination (r^2) was 0.82. The sum of the crustal elements adjusted for the probable mass of oxygen associated with them (mass of crustal oxides = 1.89[Al] + 1.57[Si] + 1.2[K] + 1.4[Ca] + 1.43[Fe]) was compared to $PM_{2.5}$ mass in Figure 4.1-23. The crustal mass concentrations were also related to $PM_{2.5}$ mass ($r^2 = 0.75$). Many of the individual elements were also spatially correlated, although not as strongly as these sums. None of the more abundant elements had patterns that strongly differed from that of $PM_{2.5}$ mass.

4.1.1.4.2. Particle-phase Organic Compounds and Source Apportionment

Atmospheric PM samples collected on Leg C of the two-week sampler in 1995 were analyzed using gas chromatography/mass Spectrometry (GC/MS) techniques by Manchester-Neesvig et al. (2003). The 26 two-week samples from each community were combined into samples for three periods: December 29, 1994–May 3, 1995; May 3–November 1, 1995; and November 1–December 28, 1995. Ninety-four particle-phase organic compounds were quantified in the samples, including n-alkanes, fatty acids, polycyclic aromatic hydrocarbons (PAHs), hopanes, steranes, aromatic di-acids, resin acids, aliphatic di-acids, methoxyphenols, and levoglucosan. The annual average concentrations of these compounds are listed in Table 4.1-8. Analyses of collocated data showed the combined sampling and analysis procedures were as precise as the other parameters measured with the TWS (as described in Section 3.4.1). Annual averages for San Dimas were based on data for the second and third periods only because the samples for the first period were invalidated. Significant seasonal and spatial variations were evident in the data. The average concentrations during the third period, November–December, were two to three times higher than average concentrations during the first and second periods. November–December 1995 was the period with the highest monthly $PM_{2.5}$ mass in the eight-year study

period. For many of the compounds, five- to ten-fold spatial variations in the annual concentrations were observed in the year-long data set.

A source apportionment analysis was conducted by Manchester-Neesvig et al. (2003) using the molecular-marker source apportionment model developed by Schauer et al. (1996) and Schauer and Cass (2000). The measured concentrations for each period were used in the model to estimate the primary source contributions to $PM_{10,}$ OC, and PM_{10} mass. The estimated source contributions to PM_{10} mass are shown in Table 4.1-9 and Figure 4.1-25 for six important types of air pollution sources. The source types that could be distinguished with this particular data set included gasoline-powered motor vehicle exhaust, diesel vehicle exhaust, wood smoke, vegetative detritus, tire wear, and natural gas combustion. Re-suspended road dust (from paved and unpaved roads) and meat cooking smoke were also identified but could not be reliably quantified. Because paved and unpaved road dusts are important sources of PM_{10} and were not included in the model, the model estimates could not apportion 100% of the PM_{10} concentrations. Seasonal variation in the gasoline vehicle contribution was compared to the diesel vehicle contribution. Gasoline and diesel vehicles were estimated to contribute about 15% and 85%, respectively, of the vehicle exhaust PM_{10} in summer and fall. In winter and early spring, the gasoline vehicle percentage increased to about 45%. The seasonal difference was hypothesized to be due to higher cold-start PM emissions in winter than in summer, as seen in dynamometer test of light-duty vehicles in Denver, Colorado (Cadle et al. 1999). Another notable feature was the larger wood smoke contribution in fall and winter, compared to summer, using leucoglucosan as a wood smoke marker.

4.1.1.4.3. Carbon Monoxide

Continuous carbon monoxide (CO) analyzers were installed in 2000 in all CHS communities except Alpine. Data suitable for assessing long-term average concentrations were available in Atascadero, Santa Maria, Lompoc, Lancaster, and Long Beach in 2000 and in 11 CHS communities in 2001. Figure 4.1-26 displays annual average (24-hr) CO concentrations. Lompoc and Lake Arrowhead had lower CO levels (0.1 to 0.2 ppm) compared to all other communities while Long Beach and Glendora had higher CO levels (0.7 to 0.8 ppm) on an annual basis. These are also the communities with the lowest and highest traffic volumes. The other communities had moderate annual CO levels of 0.3 to 0.5 ppm.

Average diurnal profiles of hourly CO concentrations are illustrated in Figure 4.1-27 for representative communities. All of the diurnal profiles had a morning peak that occurred between 7 am and 9 am (PST). Mid-day CO levels tended to be 20% to 70% lower than morning peak levels. An evening CO peak comparable to the morning peak occurred in a few communities (e.g., in Atascadero and Santa Maria). In most communities, the magnitude of the evening CO peak was lower than the morning CO peak. These diurnal profiles were similar to NO profiles and consistent with motor vehicle emissions contributing the largest amount to ambient concentrations. The morning CO peak was most likely due to the concurrence of high traffic volume and adverse meteorological conditions (i.e., low wind speeds and limited ventilation) during the morning rush hours' commute.

CO is not necessarily a unique indicator of air pollution since it is co-emitted with NO_x from vehicles and other combustion sources. Figure 4.1-28 shows the relationship of morning CO and

NO hourly concentrations in four representative communities. The hourly concentrations were reasonably well-correlated in Santa Maria ($r^2 = 0.73$), Atascadero ($r^2 = 0.79$), Lancaster ($r^2 = 0.82$), and Long Beach ($r^2 = 0.88$). The spatial patterns and trends of CO concentrations during the eight-year study in the CHS communities were similar to NO concentrations, especially in the morning. However, the downward trend in CO was steeper than that for NO. This was reflected in the SoCAB transitioning from NAAQS CO nonattainment status to attainment status during the CHS study period.

The high correlation of CO and NO has implications for San Dimas where NO concentrations near the center of the community were six times higher than those measured on the outskirts of the community in Glendora (see Figure 4.1-19). Since CO and NO were well-correlated in Glendora ($r^2 = 0.51$), the spatial differences in NO might imply that San Dimas had the highest CO of any CHS community.

4.1.1.4.4. Ultrafine Particulate Matter

Ultrafine PM includes some of the smallest atmospheric particles, and is defined as particles with diameters less than 0.1 μm (or 100 nm). The atmospheric loading of ultrafine PM is therefore small when expressed in terms of mass, but very large when expressed in terms of particle number concentrations. Atmospheric particle number distributions are dominated by ultrafine particles, but particles with diameters larger than 0.1 μm rarely contribute more than a small percent to total particle number concentrations. Spatial and temporal variations in ambient particle number concentrations might therefore be quite different than in $PM_{2.5}$ or PM_{10} mass.

Increasing concern about the possible health effects of ultrafine PM, coupled with a dearth of available information about the spatial variability of ultrafine particles in the urban atmosphere, led ARB staff to install condensation particle counters (CPCs) in CHS monitoring stations in fall 2000. Sampling difficulties related to the instrument's operating solvents and routine continuous operation of laboratory-grade instrumentation limited the valid data capture rate in the first year of sampler deployment. During the first year (2001) of operations, data were collected for 9 to 12 months in Atascadero, Santa Maria, Lompoc, Lancaster, and Alpine, and for 1 to 5 months in Lake Arrowhead, Glendora, Upland, Mira Loma, Riverside, Long Beach, and Lake Elsinore. The available data were, therefore insufficient to compare long-term average particle number concentrations across the CHS network. However, more limited comparisons were possible and are discussed below.

Figure 4.1-29 shows the annual average particle number concentrations in the five communities with nine or more months of data. The highest annual particle number concentration occurred in Santa Maria, where the average level was 20,700 particles/cm^3. This is substantially higher than in the neighboring communities of Lompoc and Atascadero where annual concentrations were 4,500 and 13,200 particle/cm^3, respectively, and may reflect local source contamination of the sampler. Lancaster and Alpine reported annual particle concentrations of 11,700 and 13,400 particles/cm^3, respectively.

Most of the CPCs operated in October–December 2001. Figure 4.1-30 shows a comparison of the monthly average particle concentrations during this period. The highest monthly average particle number concentrations occurred in Long Beach, Upland, Mira Loma, and Riverside

where levels ranged from 17,000 to 19,000 particle/cm^3 in October, from 21,000 to 28,000 particle/cm^3 in November, and from 28,000 to 37,000 particles/cm^3 in December. The levels at Santa Maria (25,000 particles/cm^3) were almost as high as at Riverside in December, but lower in the other months. The monthly average concentrations were in the 10,000-to-20,000 particles/cm^3 range in most other communities, except in Lompoc and Lake Arrowhead where concentrations ranged from 4,100 to 9,300 particles/cm^3. Note, the levels in Glendora were comparable to those in Lancaster and Alpine, which was lower than expected.

Diurnal variations in particle number can be quite large. The average diurnal profiles of particle number concentrations in October–December are shown in Figure 4.1-31. The average hourly concentrations varied by factors of 3 or 4 over the course of the day in many communities. For example, the average hourly concentration in Upland increased from 13,000 at 3 am to 43,000 at 7 am, declined to 19,000 by noon, and increased again to 35,000 by 6 pm. Diurnal particle number profiles generally displayed two characteristic shapes: one with a broad mid-day or afternoon peak, and another with morning and evening peaks. The broad mid-day or evening peak diurnal profiles occurred in Long Beach, Santa Maria, Glendora, and Lake Arrowhead. All other communities had morning and evening peaks, with the magnitude of the morning peak exceeding that of the evening peak, on average. The shapes of the particle number diurnal profiles shared similarities with those of NO and CO in many communities. Fresh emissions of ultrafine particles from motor vehicles were the mostly probable source of observed morning and evening peaks in particle number concentrations. Long Beach is a high-traffic-density community with coastal meteorology where a strong morning peak concentration was expected; however, the observed diurnal profile of particle number in Long Beach was more similar to the diurnal profile of heavy-duty vehicle traffic than light-duty vehicle traffic on Interstate-710 (Chinkin et al. 2003). Interstate-710 in Long Beach has a much larger percentage of heavy-duty vehicles (diesel truck) than other freeways, and the mid-day peak in traffic volume and particle number may suggest that truck emissions were the dominant source. The diurnal profile and relatively high particle number concentrations in Santa Maria (an otherwise low pollution site) may reflect the influence of mid-day (e.g. lunchtime) meat cooking at nearby fast-food restaurants. The mid-day peak in Lake Arrowhead was probably related to the importance of pollution transport, rather than any substantive local pollution sources, in this downwind mountain community.

The relationship of particle number concentrations to other parameters measured in the CHS varied by parameter and site. The data were too sparse in 2001 to reliably examine spatial correlation across all communities. However, within-community temporal correlations were informative. On a short-term basis, there was little observed relationship between PM_{10} mass and particle number ($0.01 < r^2 < 0.08$), which was consistent with results for other areas. Relationships were observed between morning concentrations of primary species, such as NO and CO, and morning particle number concentrations. Figure 4.1-32 and Figure 4.1-33 show the relationship of hourly particle number and primary species between 5 AM and 9 AM. The particle number concentrations correlated reasonably well with NO and CO in the morning in Atascadero and Lancaster ($0.61 < r^2 < 0.74$), and less so with NO and CO in Long Beach and Santa Maria ($0.25 < r^2 < 0.43$). The relationships between morning concentrations of these primary pollutants were stronger in communities that had large morning peaks in their diurnal profiles, suggesting a common source, (fresh motor vehicle emissions) for these primary

pollutant emissions. The lower correlation in Santa Maria may be due to the generation of ultrafine particles from meat cooking rather than from motor vehicles that co-emit NO_x and CO. Understanding particle number concentrations in Long Beach may be particularly complex because of uncertainties in contributions from port activities and other industrial sources as well as gasoline- and diesel-fueled vehicles.

The intra-community variability and spatial representativeness of the particle number as well as CO and NO primary species data are uncertain. The CHS station locations within the communities were not selected on the basis of their positions relative to roadways. For example, the Riverside monitoring station was located at an agricultural research field much farther away from major roadways than most CHS residences. Likewise, the Glendora station is more than 3 km from a busy freeway, yet most CHS residences are within 1 km of the freeway in San Dimas. Proximity to local sources is likely to have significant influence on the concentrations of NO, CO, and ultrafine PM, so these data must be interpreted cautiously. Nevertheless, as shown in Table 4.1-10, the relative ranking of communities by primary species concentrations was quite consistent. These rankings must be considered preliminary because of the limited data records for CO and particle number.

4.1.2. Pre-study Exposures

The ambient concentrations to which the CHS participants were exposed from birth (1975–1986) to the time they enrolled in the study (1993 or 1995) were derived from the U.S. EPA's database of historical air quality. There were approximately 4500 unique domestic locations where participants were born or lived for one or more months. Figure 4.1-34 shows that the pre-study locations were spread across the entire United States in 46 of the 50 states. A significant number of the participants lived in California before enrollment, as shown in Figure 4.1-35. Participants also lived in foreign countries before the study, but insufficient data were available to reliably estimate the exposure occurring in these countries. The monthly average air pollution metrics for ozone, NO_2, and PM_{10} were estimated for all of the domestic locations between 1975 and 1995. Figure 4.1-36 through Figure 4.1-39 show time series plots of key exposure metrics from 1975–1993 in the CHS communities. They clearly illustrate the diversity of exposure levels in the different communities prior to commencement of the study. The range of exposure levels across the CHS communities was as large or larger (depending on pollutant) than the range between low- and high-pollution areas outside southern California. This occurred because the CHS communities with low (e.g., Lompoc) and high pollution (Lake Arrowhead or Mira Loma) are among the lowest and highest in the country historically. The cumulative exposures for individuals were estimated using the monthly ambient concentrations for each location and time period indicated in their residential histories.

4.1.3. Characterization of Traffic and Exposures to Motor-Vehicle-Related Pollutants

4.1.3.1. Distance to Nearest Roadways

The proximity of residences to roadways is commonly used in epidemiologic studies of traffic and respiratory health (Oosterlee et al. 1996; Brunekreef et al. 1997; English et al. 1999). For the CHS, the distances from residences to the nearest roadways of different types were calculated

and the associated LDV and HDV traffic volumes on those roads were tabulated. GIS tools were used to calculate the distance to the nearest (1) interstate freeway, U.S. highway, or limited access highway; (2) principal arterials or other highways; (3) minor arterials; (4) major collectors; and (5) minor collectors. Multiple classes of roadways were included in order to capture a range of potential exposures. It was particularly important to include arterials because even though they have lower traffic volumes than freeways, residents often live closer to arterials than freeways and may experience similar local-scale traffic impacts. Also, not all of the CHS communities have freeways (Lompoc and Lake Arrowhead do not have freeways).

The baseline and subsequent residential locations (~8500) were included in the database of distances and traffic volumes. The results for all CHS residences are summarized in Figure 4.1-40 and Figure 4.1-41. Figure 4.1-40 shows the distribution of distances between the nearest interstate freeway, U.S. highway, or limited access highway and CHS residences. The calculations indicate that 1.7%, 5.7%, and 6.5% of CHS residences were located within 100 m, from 100 to 300 m, and from 300 to 500 m, respectively, of major freeways. Zhu et al. (Zhu et al. 2002a; Zhu et al. 2002b) reported traffic-related pollutant concentrations are substantially elevated within 100 m of Los Angeles freeways and concentrations decline asymptotically to urban background levels with 300 m of the freeway. Thus, approximately 7% of the CHS residences were close enough to freeways for participants to most likely be exposed to elevated concentrations of traffic-related pollutants. About 18%, 43%, and 26% of residences were located from 500 to 1000 m, 1000 to 3000 m, and more than 3000 m from a major freeway, respectively. Figure 4.1-41 shows the distribution of distances to principal or minor arterials in all communities. Approximately 13% and 26% of CHS residences were located from 0 to 100 m and from 100 to 300 m, respectively, of principal or minor arterials. Arterials typically have 60% to 80% lower traffic volumes than freeways, but many more residences (~39%) were located near arterials.

Proximity to roadways and traffic varied considerably across the CHS communities. Table 4.1-11 and Table 4.1-12 show the percentage of residences in each community located various distances from freeways and arterials. They also show the community-average daily traffic volumes on the nearest roads and the average fractions of HDV, which are predominately diesel-fueled vehicles. Twelve to twenty percent of the CHS residences in Atascadero, Riverside, Long Beach, and Alpine were located with 300 m of a major freeway. In contrast, only 1% of CHS residences in Lancaster were with 300 m of the freeway. The traffic volumes on the freeways in Long Beach and Riverside were much higher than those in Atascadero and Alpine. Traffic volumes on the freeways varied from 40,000 to 224,000 vehicles per day, and the HDV fractions on freeways varied from 5% to 10% between these communities. From 14% to 22% of the CHS residences in Long Beach, San Dimas, Upland, Mira Loma, Riverside, Lancaster, and Alpine were located within 100 m of a minor arterial with 11,000 to 38,000 vehicles per day. More than 60% of the residences in these same communities were located within 300 m of minor arterials. Less than 30% of the CHS residences in Lompoc, Lake Arrowhead, and Atascadero were located within 300 m of a principal or minor arterial. Hence, there was considerable heterogeneity in the proximity to traffic between and within the communities. Long Beach stands out as having high traffic volumes and more CHS participants living close to busy arterials than in other communities. Atascadero stands out as having the most subjects living close to the freeway.

4.1.3.2. Traffic Density

The second approach for characterization of traffic exposures was spatial mapping of traffic densities. The spatially mapped traffic density estimates tended to vary more smoothly in space than the distances to nearest roads. The traffic density estimates also captured the enhancement effects of intersection and multiple roadway influences that are missed using only distance to the nearest roadways. The spatially mapped traffic densities in the CHS communities are shown in Figure 4.1-42 through Figure 4.1-47. They show high densities in narrow bands along the freeways, moderate densities along major arterials, and lower or zero densities in the suburban neighborhoods. The maps also show the proximity of CHS residences to roadways.

Traffic densities at the CHS residence locations were extracted from the mapped fields. The distribution of traffic densities in the communities is shown in Figure 4.1-48 and Figure 4.1-49. Zero or low (<10) traffic densities are estimates for a large percentage of the residences. For example, more than 40% of the residences in Lake Elsinore, Lompoc, Alpine, and Mira Loma had zero traffic densities, which indicates they were located in low-density suburban neighborhoods for which Caltrans does not supply traffic volume data. The majority of residences in Lake Elsinore, Lake Arrowhead, Lompoc, Atascadero, Lancaster, Alpine, Upland, and Mira Loma had low (<10) traffic densities. This result is also consistent with the predominantly suburban land use in these communities. Nevertheless, there are small percentages of residences in many communities that had high traffic densities. Thirty-seven percent of the residences in Long Beach and 11 to 14% of the residences in San Dimas, Upland, Riverside, and Santa Maria had high (>70) traffic densities. These distributions suggest there were potentially large differences in exposures to traffic for children within the same communities as well as in different communities.

The average traffic densities in the different communities are shown in Table 4.1-13 and Figure 4.1-50. There was more than a sevenfold difference in mean traffic densities across the communities. The community-mean traffic densities were low at residences in Lompoc, Lake Arrowhead, Lake Elsinore, and Mira Loma, and were high at residences in Long Beach, Riverside, and San Dimas. The relative ranking of communities with low and moderate traffic densities varied somewhat depending on whether all residences or only residences with nonzero traffic density estimates are considered; however, the three high traffic density communities were clearly high regardless of the choice of metrics.

There are notable differences between-community ranking for traffic density and air pollutant concentrations. Santa Maria stood out as having relatively high traffic densities yet ambient concentrations of air pollutants, including motor-vehicle-related pollutants, were generally low at the Santa Maria central air monitoring site. Atascadero also had somewhat higher ranking for traffic density than most ambient air pollutant concentrations. In contrast, Mira Loma had among the highest air pollutant concentrations and motor vehicle PM source apportionment, yet the traffic density estimates for Mira Loma were quite low. Lake Elsinore also had a lower ranking for traffic density compared to air pollutant concentrations. Both Mira Loma and Lake Elsinore are growing communities where the traffic data may be underestimated. Also, emissions from trucks idling at distribution centers may be important in Mira Loma, but are not represented in the traffic density estimates.

4.1.3.3. Line Source Dispersion Modeling of Traffic-Related Emissions

Motor vehicle emissions have been recognized as contributing to local-scale (e.g., CO) and regional-scale (e.g., ozone, NO_2, $PM_{2.5}$) air pollution problems and health effects for many years. If motor vehicle pollutants influence health as a function of proximity to roadways, it is important to establish which pollutants are responsible for the effects as well as how much different types of roads and vehicles are contributing. The third approach for characterizing traffic-related exposures in the CHS extends the proximity to traffic analysis to include specific pollutant concentrations. Air quality dispersion modeling is used to estimate the contributions of local traffic-related emissions to ambient concentrations of specific pollutants at residences in the CHS communities. The CALINE4 line source dispersion model is adapted to estimate long-term average concentrations of NO_2, NO_x, $PM_{2.5}$ EC, and $PM_{2.5}$ OC. The model calculations are made to estimate the contributions of two classes of vehicles, gasoline-fueled light-duty vehicles and diesel-fueled heavy-duty vehicles, on two types of roadways, freeways and arterials plus collectors (or nonfreeways).

The estimated range of long-term average ambient concentrations from local mobile source emissions at CHS residences is illustrated in Figure 4.1-51. The box-whisker plots show the annual average NO_2, NO_x, $PM_{2.5}$ EC, and $PM_{2.5}$ OC concentrations at CHS residences estimated by the CALINE4 model. These particular results are for simulations made with vehicle emission factors and traffic volumes for 1997, and meteorological data for 1994-1999. The boxes show the 25th, 50th, and 75th percentiles, while the whisker shows the high and low values. The inner-quartile ranges of concentrations for Long Beach, Riverside, and San Dimas were substantially higher than in other CHS communities. There was a lot of overlap in the ranges of estimated exposures in most communities; however, the estimated exposures of NO_2 and $PM_{2.5}$ EC at all residences in Lompoc were lower than the lowest estimate for the high traffic community of Long Beach. The model results suggest there were "hot spot" residences that had much higher contributions from local sources than the average for the community. In most of the communities, the hot spot concentrations were three to five times higher than the median concentrations, yet there were certain communities like Alpine that showed greater disparity. The residences with the highest estimated local traffic impacts in Alpine had concentrations that were more than 10 times the median concentrations at residences in the community. The relative rankings and dispersion of the estimated concentrations in the different communities were similar for the different pollutants. The between-community variations in the relative amounts of HDV and LDV activity did not have an overwhelming effect on the distributions of estimated NO_2, NO_x, $PM_{2.5}$ EC, and $PM_{2.5}$ OC concentrations. Communities with high NO_2 estimates also had high $PM_{2.5}$ EC and $PM_{2.5}$ OC estimates. Note, additional simulations (not shown) were made with emission factors and traffic volumes for 1994 and 2000 which showed slightly higher and lower levels, respectively.

Figure 4.1-51 also includes comparisons of the estimated concentrations from local traffic emissions to the 1995-1998 average ambient concentrations observed at the central air monitoring station in each community. Model-estimated incremental concentrations from local traffic were substantially lower than the central site-measured concentrations at almost all residences in the study. This result is not unexpected for these species since regional transport in all communities and other local sources in some communities are likely to be important contributors. The comparison may also indicate that local vehicle emission rates are

underestimated. This comparison of the model results to measurements is limited and should be viewed as necessary but not sufficient to evaluate model performance. Many more locations with reliable measurements of local source indicators are needed to reliably evaluate this type of neighborhood-scale modeling.

Figure 4.1-52 and Figure 4.1-53 show comparisons of CALINE4 model estimates for the air monitoring station locations to the 4-year average observed ambient concentrations at the stations. There was correlation between the model estimates and the observations for all pollutants. The dispersion model estimates local motor vehicle emissions contributed 24% of the observed NO_2 concentrations, on average, and the coefficient of determination was 0.57. The model's estimates for NO_2 in Long Beach were higher than for other stations and were a larger percentage (40%) of the observed ambient concentration. The high predicted concentration for Long Beach was consistent with high traffic volumes in the community and the high ratio of predicted traffic contribution to total observed NO_2 was consistent with Long Beach's coastal location where regional pollutant transport had less influence. Long Beach also had more local sources of NO_2 (e.g., from ship, trains, and other stationary sources) than most other CHS communities, but the importance of the local sources is not clear from this comparison. The model estimates for NO_x were more strongly correlated with observations ($r^2 = 0.67$) than those for NO_2. This result was consistent with the dispersion model being formulated for chemically non-reactive species and, therefore, the model was more accurate for NO_x than NO_2. Long Beach results appeared to be outliers; the model results were more strongly correlated when the values for Long Beach were excluded ($r^2 = 0.81$ for NO_2 and $r^2 = 0.97$ for NO_x).

The average estimated contribution of local motor vehicle emissions to the observed $PM_{2.5}$ EC concentrations was 26% and the coefficient of determination was 0.57. The predicted $PM_{2.5}$ EC concentrations for Long Beach were much higher than those predicted for any other community monitoring station, yet the average predicted $PM_{2.5}$ EC concentration was only 40% of the observed concentration. The relationships of the model results to the observed concentrations were very similar for $PM_{2.5}$ EC and NO_2. This was somewhat surprising given that EC emissions from vehicles are not well characterized and most emissions experts consider PM (and EC) emissions rates to be far more uncertain than those for NO_x.

The model estimates for $PM_{2.5}$ OC were not well correlated with the observed concentrations ($r^2 = 0.11$). The model estimated that the highest $PM_{2.5}$ OC from local mobile sources occurred in Long Beach. The model estimate for the Mira Loma station location was quite low compared to the very high OC concentration observed in the station. The correlation with the observed value was higher ($r^2 = 0.71$) if the data from Long Beach and Mira Loma were excluded. The comparison of estimated and observed $PM_{2.5}$ OC was confounded by not only the contributions of regional transport and other local sources, but also the contribution of secondary OC to the measured OC.

The average estimated NO_2, NO_x, $PM_{2.5}$ EC, and $PM_{2.5}$ OC concentrations at all residences and at the top 10% of traffic impacted residences are shown in Table 4.1-14 and Table 4.1-15 for each community. Residences in Long Beach, San Dimas, and Riverside had the highest concentrations, on average, for all pollutants, followed by residences in Upland and Mira Loma. Residences in Lompoc and Lake Arrowhead had the lowest estimated concentrations from local

traffic. Mean concentrations at the top 10% of traffic-impacted residences were 50% to 200% higher than the mean concentrations at all residences in most communities. In Alpine, mean concentrations at the top 10% of traffic-impacted residences were 270% to 340% higher than the mean concentrations at all residences. The top 10% of traffic-impacted residences in Riverside were estimated to have higher NO_2 concentrations than those in Long Beach (also shown in Figure 4.1-51).

The estimated relative contributions of light-duty and heavy-duty vehicle activity on freeways and non-freeways (arterials and collectors), shown in Table 4.1-14 and Table 4.1-15, suggest light-duty vehicle activities generally contribute more NO_2, NO_x, $PM_{2.5}$ EC, and $PM_{2.5}$ OC than heavy-duty vehicle activities. One notable exception was in Long Beach where heavy-duty vehicle activities were estimated to contribute the majority of the $PM_{2.5}$ EC at the top 10% of traffic-impacted residences. The results consistently indicated heavy-duty vehicle traffic on freeways contributed more to residences than heavy-duty vehicles traffic on non-freeways. The relative importance of freeway contributions for light-duty vehicles varied considerably by community. Light-duty vehicle emissions from freeway traffic were estimated to have little or no impact at residences in Lancaster, Santa Maria, Lompoc, and Lake Arrowhead, on average. However, in Riverside and Long Beach it was not uncommon for the light-duty vehicle freeway contribution to exceed the non-freeway contribution. Both of these communities have two freeways with high traffic volumes. Detailed examination of the model estimates for individual residences within the top 10% of traffic-impacted residences indicated the dominant contributions could come from heavy-duty freeway, light-duty freeway, or light-duty non-freeway activities. Hence, no single class of vehicles or type of roadway was exclusively responsible for the highest estimated concentrations at residences.

Lastly, Figure 4.1-54 shows the relationships between local mobile source NO_x estimates and the previously described traffic density metric. The correlation was relatively strong in Atascadero ($r^2 = 0.75$) and Riverside ($r^2 = 0.71$), but weak in San Dimas ($r^2 = 0.39$). The principal differences between the approaches are that the CALINE4 model incorporates meteorology (albeit simplistically) and predicts a more gradual decline in concentrations with distance from roadway than the traffic density model. Neither approach has been satisfactorily evaluated against observations, yet both represent plausible characterizations of traffic-related exposures.

4.1.4. Exposure Modeling

The Individual Exposure Model (IEM) was used to estimate the long-term average exposures of CHS participants to pollutants of ambient origin. The model considers exposures occurring indoors and outdoors at residences, indoors and outdoors at schools, in vehicles, and in other microenvironments. Indoor exposures are based on estimates of the contribution of outdoor air that infiltrates to the indoor microenvironment. The contributions of indoor sources to personal exposure are ignored in these simulations because of the CHS's focus on ambient air pollution. Model estimates were made for each subject in each year for NO_2, ozone, $PM_{2.5}$ mass, $PM_{2.5}$ EC, $PM_{2.5}$ OC, and PM_{10} mass. The volume of model output is large. The model results for 1997 are used here to illustrate characteristics of the personal exposure estimates and, most importantly, to compare them to the ambient exposure data that are used in the health analyses. Although 1997 is presented as a singular case example here, the relationship between personal exposure estimates and ambient exposure levels is consistent for other study years. Long-term trends in

pollution spatial patterns have been previously described and reflect generally consistent inter-community relationships (see Section 4.1.1.1). Accordingly, the discussion presented herein is representative of and relevant to the larger overall study exposure database.

Table 4.1-16 lists the mean and standard deviation of the estimated annual average personal exposures by community in 1997. Figure 4.1-55 shows box-whisker plots of the distributions of annual average estimates in each community and the observed Method I annual average ambient concentrations. The inner-quartile range of estimated personal NO_2 exposures are below the mean ambient observations in all of the communities except Riverside. In Riverside, the median personal NO_2 exposure is comparable to the measured value at the central site. Excluding Riverside, the median personal NO_2 exposures in the communities range from 50% to 80% of the ambient concentrations measured at the central site. These results are consistent with personal exposure and indoor/outdoor observations in Los Angeles at homes without indoor combustion (Colome et al. 1989; Drye et al. 1989). Typical differences between the personal exposure estimates and ambient concentrations are illustrated by pairs of sites with similar ambient concentrations. For example, the annual average ambient NO_2 concentration was 34 ppb in both Upland and Long Beach in 1997, yet the mean personal NO_2 exposure concentrations were 21 ppb in Upland and 28 ppb in Long Beach. Likewise, the mean ambient NO_2 in both Mira Loma and Riverside was 22 ppb, yet the mean personal NO_2 exposure concentrations were 13 ppb in Mira Loma and 23 ppb in Riverside. The differences are primarily due to higher local traffic impacts on concentrations outdoors of schools and residences in Long Beach and Riverside.

The personal exposure estimates for ozone are simulated without accounting for local-scale impacts of traffic. The exposure modeling results are driven by the ambient concentration data in each community, yet account for differences in housing characteristics and time-activity among individuals. The median "all hours" personal ozone exposures are consistently estimated to be 35% to 40% of the observed ambient levels in the community. The inner-quartile ranges in personal exposure estimates are small, 3 to 5 ppb on a 24-hr basis. The strong correlation of personal and ambient ozone suggests both exposure metrics would behave similarly in health effects analyses.

The estimated personal exposures for $PM_{2.5}$ and PM_{10} mass of outdoor origin track the ambient concentrations fairly well. The PM_{10} estimates are more strongly correlated with the ambient concentrations than the $PM_{2.5}$ estimates. The median estimated personal $PM_{2.5}$ exposures in the communities range from 70% to 95% of the ambient concentrations on an annual basis for 1997. The median personal PM_{10} estimates in the communities range from 60% to 80% of the ambient concentrations. The inner-quartile ranges of annual $PM_{2.5}$ exposures are below the ambient concentrations in all communities except Long Beach, where the inner-quartile range overlaps the ambient $PM_{2.5}$ concentration. The inner-quartile ranges of $PM_{2.5}$ personal exposure estimates are very similar in Mira Loma, Riverside, and Long Beach, yet the ambient $PM_{2.5}$ concentrations are 20% and 30% lower in Riverside and Long Beach, respectively, than in Mira Loma. Local traffic contributions to personal $PM_{2.5}$ exposures are estimated to be higher in Riverside and Long Beach than in Mira Loma. Another interesting characteristic of the $PM_{2.5}$ personal exposure estimates is that although variations in estimated exposures exist within each community, there is no overlap between the estimates in the seven cleaner communities and the

five more-polluted communities (San Dimas, Upland, Mira Loma, Riverside, and Long Beach). This stratification is less evident in the PM_{10} estimates than in the $PM_{2.5}$ estimates.

The model estimates for the long-term average personal $PM_{2.5}$ EC and $PM_{2.5}$ OC exposures are often higher than the ambient concentrations. The median personal $PM_{2.5}$ EC estimates are higher than the mean ambient concentrations in every community except Lancaster and Lake Elsinore. Even the 25th-percentile personal $PM_{2.5}$ EC estimates are higher than the mean ambient concentrations in every community except Lancaster, Lake Elsinore, and Alpine. Yet, the median estimated personal exposures levels are still within ±25% of the observed ambient levels in all communities. The extremes in individual $PM_{2.5}$ EC exposures in the communities are estimated to be 50% to 100% greater than the ambient concentrations. The median $PM_{2.5}$ OC personal exposure estimates are higher than the ambient concentrations in San Dimas, Riverside, and Long Beach, and lower than the ambient concentrations in the other communities. The extremes in individual personal exposure estimates for $PM_{2.5}$ OC are up to double the ambient concentrations. The highest ambient $PM_{2.5}$ OC occurs in Mira Loma; however, all of the personal $PM_{2.5}$ OC exposure estimates for participants in Mira Loma are lower than the ambient level.

The estimated contributions to mean personal exposures from outdoor local on-road mobile sources and from regional transport and other outdoor local sources in the CHS communities are shown in Figure 4.1-56 and Figure 4.1-57. Regional transport and non-mobile local sources are estimated to be the dominant contributors to mean personal exposure for most pollutants in most communities. The local mobile sources are estimated to have a greater impact on NO_2, $PM_{2.5}$ EC, and $PM_{2.5}$ OC than for $PM_{2.5}$ or PM_{10} mass. The contributions of local mobile sources to mean personal exposure are highest in San Dimas, Riverside, and Long Beach. Upland also has relatively high traffic contributions to personal exposure. These results are consistent with the characterization of traffic in the different communities.

The estimated mean personal exposure to pollutants of outdoor origin occurring while subjects were indoors and outdoors at residences and schools, in vehicles, and in other microenvironments are displayed in Figure 4.1-58 and Figure 4.1-59. The microenvironment where the largest contribution to mean personal exposure occurs is indoors at residences for all pollutants except ozone. This is consistent with time-activity data, which indicates more time is spent indoors at residences than any other microenvironment. Exposures occurring in vehicles are estimated be the second most important microenvironment for $PM_{2.5}$ EC personal exposures. Exposures occurring indoors at schools are the second most important microenvironment for most other pollutants. Exposures occurring in "other" microenvironments are also significant for most of the pollutants, especially ozone. Overall, the microenvironmental distributions of exposures are consistent with the time-activity data for children.

For the purposes of the health outcome analyses, there was concern that no personal exposure data for CHS subjects were available to verify the exposure model and that the potential benefits of having individual exposure estimates would be offset by increased uncertainties and possibly biases in these estimates. The concern for uncharacterized uncertainties outweighed possible benefits, so individual exposure model results were not used for any health outcomes analyses. Rather, these analyses were performed using community monitor data, with application of various categorical variables to account for individual differences.

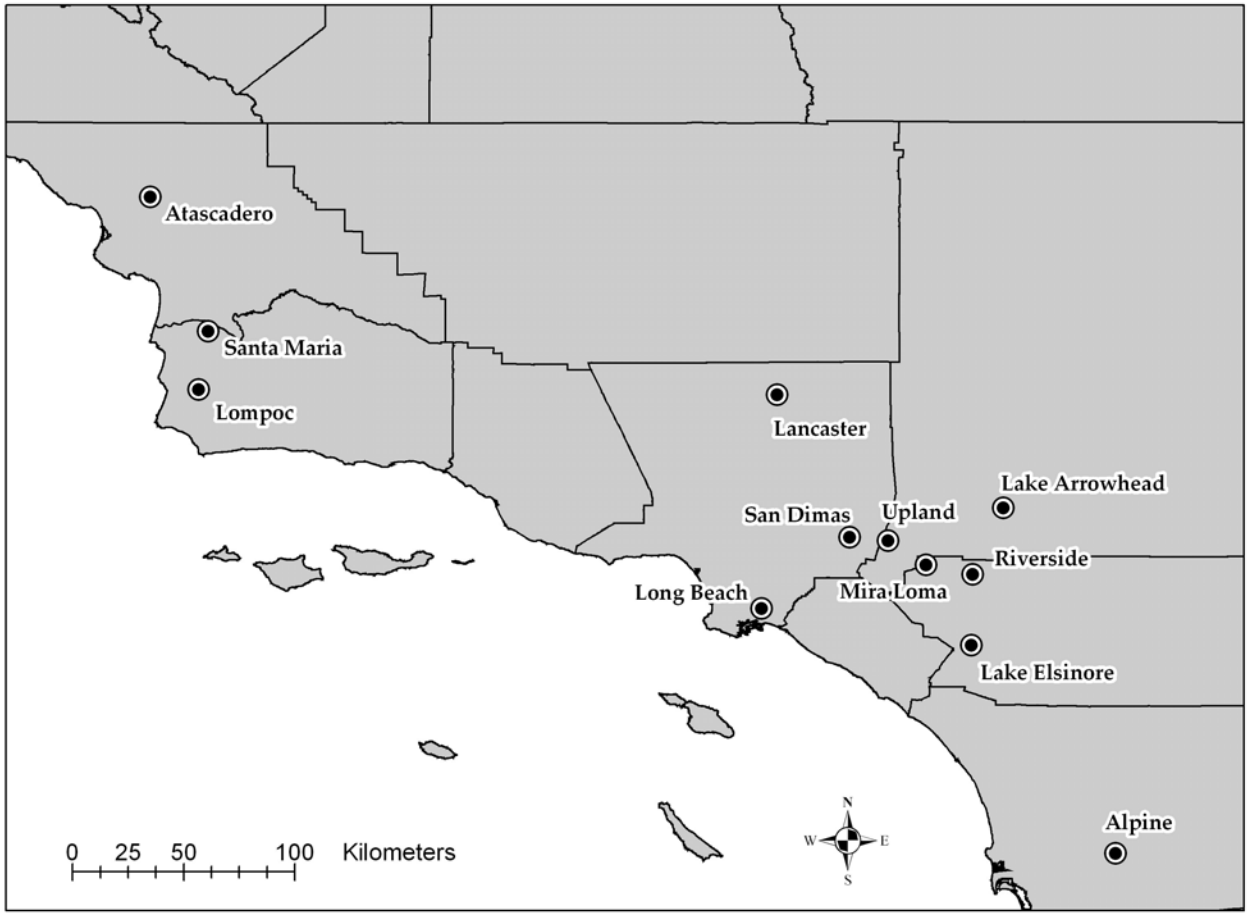


Figure 4.1-1. Locations of Children's Health Study communities.

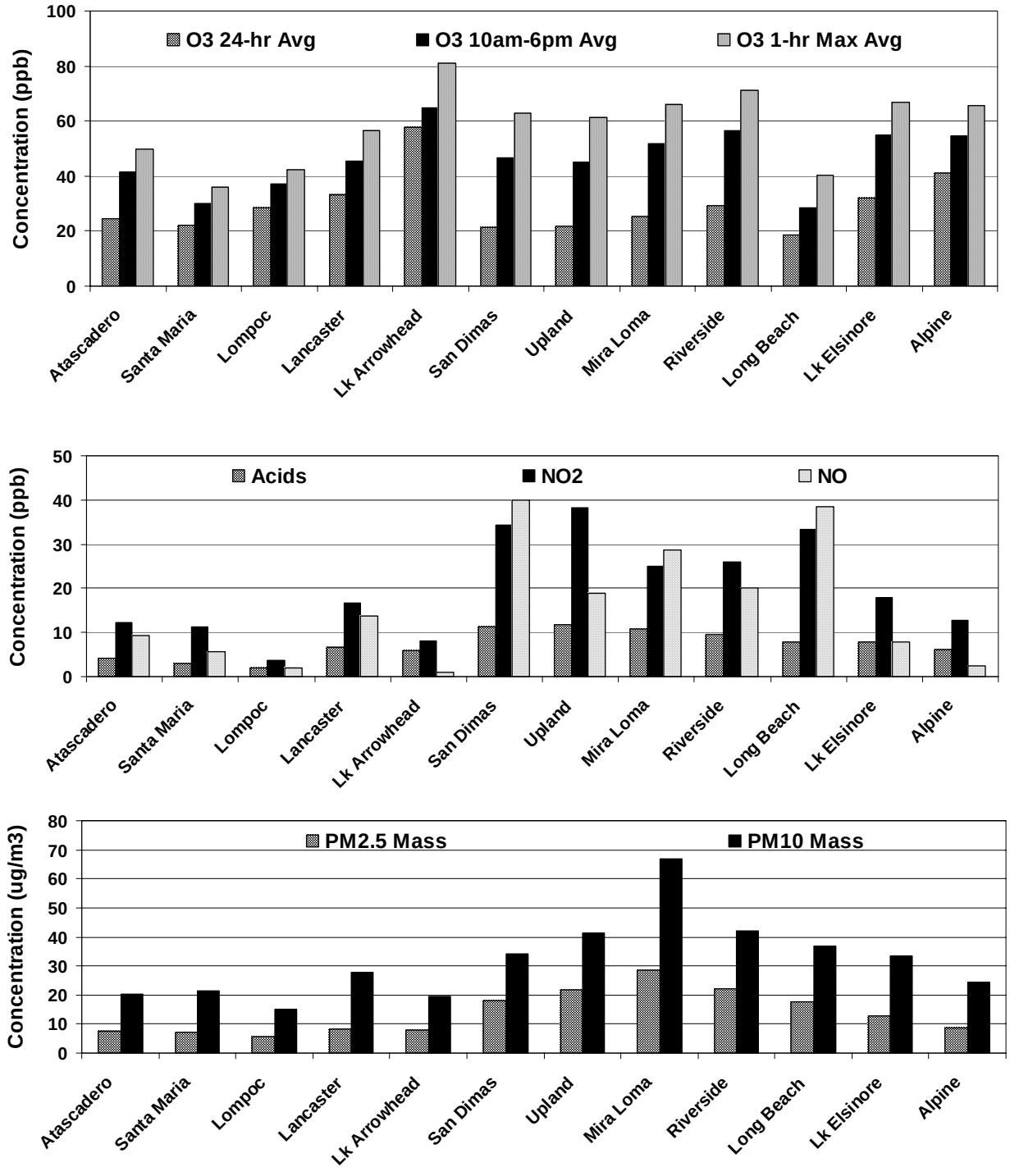


Figure 4.1-2. Eight-year average ambient concentrations of ozone, NO_2, Acids, $PM_{2.5}$ mass, and PM_{10} mass in the CHS communities.

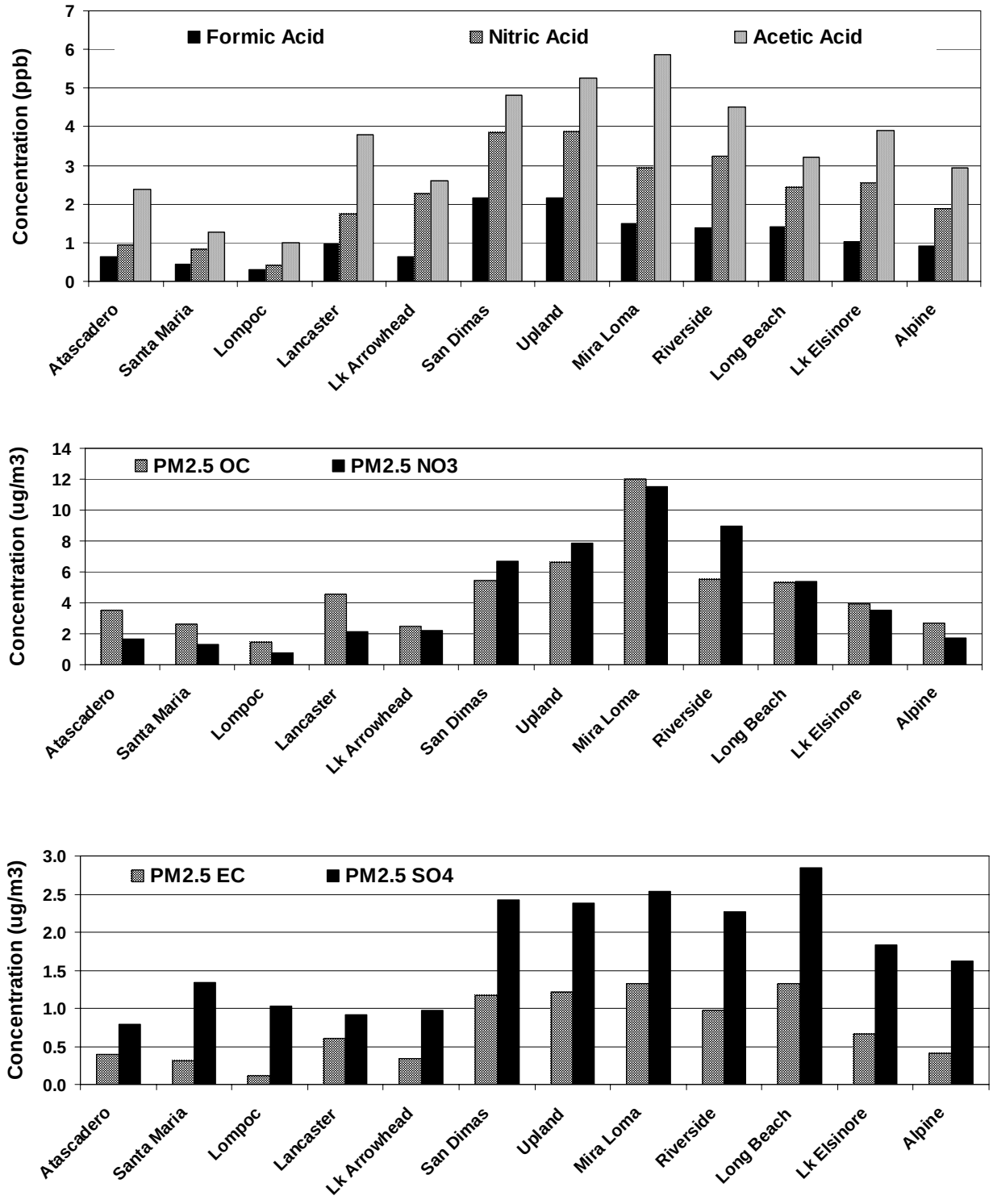


Figure 4.1-3. Eight-year average ambient concentrations of Formic Acid, Acetic Acid, Nitric Acid, $PM_{2.5}$ OC, $PM_{2.5}$ NO_3, $PM_{2.5}$ EC, and $PM_{2.5}$ SO_4 in the CHS communities.

Table 4.1-1. Eight-year average ambient concentrations (ppb) of ozone, NO_2, NO, Formic Acid, Acetic Acid, Hydrochloric Acid, Nitric Acid, and Total Acid (formic + acetic + hydrochloric + nitric) in the CHS communities.

Community	Ozone 24-hr	Ozone 10am-6pm	Ozone 1-hr Max	NO_2	NO[a]	Formic Acid	Acetic Acid	HCl	Nitric Acid	Total Acids
Atascadero	24.6	41.5	49.8	12.4	9.2	0.63	2.37	0.21	0.94	4.1
Santa Maria	22.2	30.2	35.9	11.3	5.5	0.43	1.26	0.36	0.83	2.9
Lompoc	28.5	37.3	42.3	3.7	2.1	0.30	0.98	0.36	0.42	2.1
Lancaster	33.3	45.5	56.3	16.6	13.7	0.97	3.79	0.22	1.74	6.7
Lk Arrowhead	57.7	64.9	80.9	8.0	1.0	0.64	2.60	0.25	2.28	5.8
San Dimas	21.3	46.5	62.7	34.3	39.9	2.15	4.81	0.54	3.86	11.3
Upland	21.9	45.1	61.4	38.3	18.8	2.17	5.26	0.53	3.87	11.8
Mira Loma	25.4	51.7	65.9	25.0	28.6	1.51	5.86	0.43	2.92	10.7
Riverside	29.2	56.7	71.1	25.9	20.2	1.37	4.51	0.44	3.23	9.6
Long Beach	18.8	28.6	40.3	33.3	38.5	1.40	3.22	0.74	2.43	7.8
Lk Elsinore	32.0	54.8	67.0	17.9	7.8	1.02	3.91	0.47	2.55	8.0
Alpine	41.2	54.5	65.7	12.8	2.4	0.91	2.92	0.44	1.89	6.2
Mean	29.7	46.4	58.3	20.0	15.6	1.12	3.46	0.41	2.25	7.2
Standard Deviation	10.8	10.9	13.6	11.2	13.8	0.61	1.52	0.15	1.14	3.24

[a] NO data are less complete than those for other species.

Table 4.1-2. Standardized deviations of the eight-year average ambient concentrations of Ozone, NO_2, NO, Formic Acid, Acetic Acid, Hydrochloric Acid, Nitric Acid, and Total Acid (formic + acetic + hydrochloric + nitric) from the CHS 12-community mean concentrations.

Community	Ozone 24-hr	Ozone 10am-6pm	Ozone 1-hr Max	NO_2	NO[a]	Formic Acid	Acetic Acid	HCl	Nitric Acid	Total Acids
Atascadero	-0.47	-0.46	-0.62	-0.68	-0.47	-0.81	-0.71	-1.36	-1.15	-0.96
Santa Maria	-0.69	-1.49	-1.64	-0.77	-0.73	-1.13	-1.44	-0.34	-1.25	-1.35
Lompoc	-0.11	-0.84	-1.17	-1.45	-0.98	-1.35	-1.63	-0.39	-1.61	-1.60
Lancaster	0.34	-0.08	-0.14	-0.30	-0.14	-0.25	0.22	-1.31	-0.45	-0.16
Lk Arrowhead	2.59	1.70	1.65	-1.06	-1.06	-0.79	-0.57	-1.10	0.03	-0.46
San Dimas	-0.77	0.01	0.32	1.28	1.76	1.67	0.89	0.81	1.41	1.27
Upland	-0.72	-0.13	0.23	1.63	0.23	1.70	1.19	0.76	1.43	1.42
Mira Loma	-0.39	0.49	0.56	0.45	0.94	0.62	1.58	0.12	0.59	1.07
Riverside	-0.05	0.94	0.94	0.53	0.33	0.41	0.69	0.17	0.86	0.71
Long Beach	-1.01	-1.64	-1.32	1.19	1.66	0.45	-0.16	2.12	0.16	0.17
Lk Elsinore	0.22	0.77	0.64	-0.19	-0.57	-0.17	0.30	0.35	0.27	0.22
Alpine	1.07	0.74	0.54	-0.63	-0.96	-0.36	-0.35	0.17	-0.31	-0.34

[a] NO data are less complete than those for other species.

Table 4.1-3. Eight-year average ambient concentrations (μg^3) of PM_{10} Mass, $PM_{2.5}$ Mass, $PM_{2.5}$ SO_4, $PM_{2.5}$ NO_3, $PM_{2.5}$ NH_4, $PM_{2.5}$ OC, and $PM_{2.5}$ EC in the CHS communities.

Community	PM_{10} Mass	$PM_{2.5}$ Mass	$PM_{2.5}$ SO_4	$PM_{2.5}$ NO_3	$PM_{2.5}$ NH_4	$PM_{2.5}$ OC [a]	$PM_{2.5}$ EC [b]
Atascadero	20.2	7.6	0.78	1.67	0.65	3.50	0.39
Santa Maria	21.5	7.0	1.34	1.33	0.61	2.61	0.32
Lompoc	15.2	5.6	1.04	0.76	0.38	1.44	0.12
Lancaster	27.7	8.2	0.92	2.14	0.88	4.55	0.60
Lk Arrowhead	19.7	7.7	0.97	2.20	0.96	2.51	0.33
San Dimas	34.4	18.2	2.42	6.72	2.46	5.42	1.17
Upland	41.3	21.8	2.38	7.87	2.98	6.63	1.21
Mira Loma	66.8	28.5	2.53	11.49	4.03	12.02	1.33
Riverside	42.1	22.2	2.27	8.98	3.29	5.54	0.98
Long Beach	36.9	17.7	2.85	5.41	2.26	5.31	1.33
Lk Elsinore	33.4	12.8	1.83	3.54	1.53	3.95	0.66
Alpine	24.4	8.5	1.62	1.72	0.93	2.67	0.41
Mean	32.0	13.8	1.75	4.49	1.75	4.68	0.74
Standard Deviation	14.1	7.6	0.73	3.54	1.22	2.78	0.44

[a] $PM_{2.5}$ OC estimated from 0.78*PM_{10} OC
[b] $PM_{2.5}$ EC estimated from 0.98*PM_{10} EC

Table 4.1-4. Standardized deviations of the eight-year average ambient concentrations of PM_{10} mass, $PM_{2.5}$ mass, $PM_{2.5}$ SO_4, $PM_{2.5}$ NO_3, $PM_{2.5}$ NH_4, $PM_{2.5}$ OC, and $PM_{2.5}$ EC from the CHS 12-community mean concentrations.

Community	PM_{10} Mass	$PM_{2.5}$ Mass	$PM_{2.5}$ SO_4	$PM_{2.5}$ NO_3	$PM_{2.5}$ NH_4	$PM_{2.5}$ OC [a]	$PM_{2.5}$ EC[b]
Atascadero	-0.83	-0.81	-1.32	-0.79	-0.90	-0.42	-0.78
Santa Maria	-0.74	-0.89	-0.56	-0.89	-0.93	-0.74	-0.95
Lompoc	-1.19	-1.08	-0.97	-1.05	-1.12	-1.17	-1.40
Lancaster	-0.30	-0.75	-1.13	-0.66	-0.71	-0.05	-0.31
Lk Arrowhead	-0.87	-0.80	-1.06	-0.65	-0.65	-0.78	-0.92
San Dimas	0.17	0.57	0.93	0.63	0.59	0.27	0.98
Upland	0.66	1.06	0.87	0.96	1.01	0.70	1.07
Mira Loma	2.47	1.93	1.07	1.98	1.87	2.64	1.35
Riverside	0.72	1.11	0.72	1.27	1.26	0.31	0.55
Long Beach	0.35	0.51	1.51	0.26	0.42	0.23	1.34
Lk Elsinore	0.10	-0.13	0.11	-0.27	-0.18	-0.26	-0.17
Alpine	-0.53	-0.70	-0.17	-0.78	-0.67	-0.72	-0.74

[a] $PM_{2.5}$ OC estimated from 0.78*PM_{10} OC
[b] $PM_{2.5}$ EC estimated from 0.98*PM_{10} EC

Table 4.1-5. Spatial correlation (r) in the eight-year average ambient concentrations of ozone (24-hr), NO_2, PM_{10} mass, $PM_{2.5}$ mass, $PM_{2.5}$ SO_4, $PM_{2.5}$ NO_3, $PM_{2.5}$ NH_4, $PM_{2.5}$ OC, $PM_{2.5}$ EC, formic acid, acetic acid, hydrochloric acid, and nitric acid in the CHS communities.

	Ozone	NO_2	NO[a]	PM_{10}	$PM_{2.5}$	SO_4	NO_3	NH_4	OC	EC	Formic Acid	Acetic Acid	HCl	HNO_3
Ozone	1.													
NO_2	-0.56	1.												
NO[a]	-0.61	0.85	1.											
PM_{10}	-0.35	0.67	0.65	1.										
$PM_{2.5}$	-0.41	0.81	0.73	0.95	1.									
SO_4	-0.50	0.87	0.81	0.77	0.86	1.								
NO_3	-0.36	0.77	0.71	0.94	0.99	0.81	1.							
NH_4	-0.35	0.79	0.72	0.94	1.	0.85	1.	1.						
OC	-0.37	0.64	0.65	0.97	0.90	0.67	0.90	0.89	1.					
EC	-0.51	0.94	0.90	0.85	0.92	0.91	0.88	0.90	0.82	1.				
Formic	-0.41	0.95	0.78	0.68	0.80	0.80	0.78	0.80	0.65	0.89	1.			
Acetic	-0.20	0.78	0.64	0.86	0.87	0.69	0.88	0.88	0.85	0.85	0.87	1.		
HCl	-0.50	0.73	0.69	0.45	0.57	0.90	0.49	0.55	0.33	0.71	0.62	0.37	1.	
HNO_3	-0.12	0.85	0.65	0.67	0.80	0.76	0.80	0.82	0.61	0.82	0.93	0.89	0.55	1.

[a] NO data are less complete than those for other species.

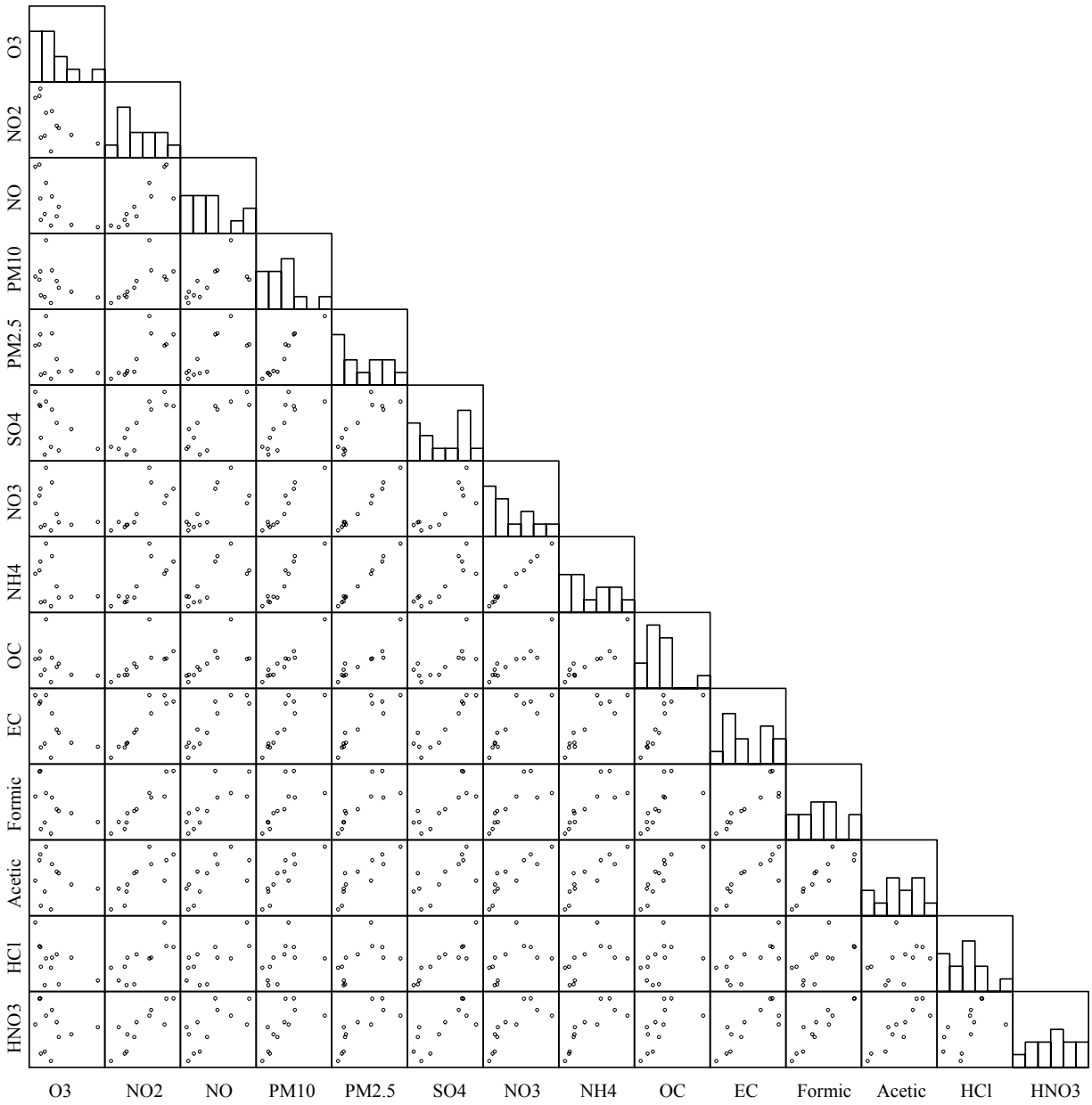
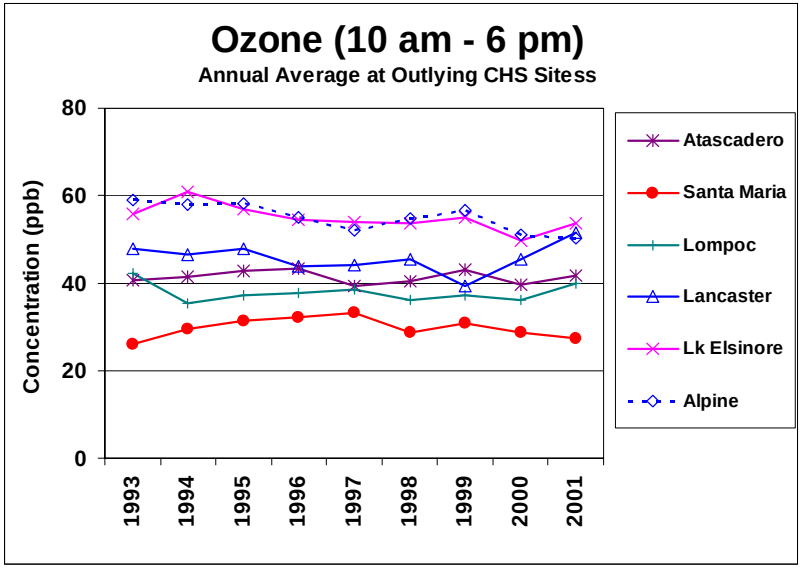


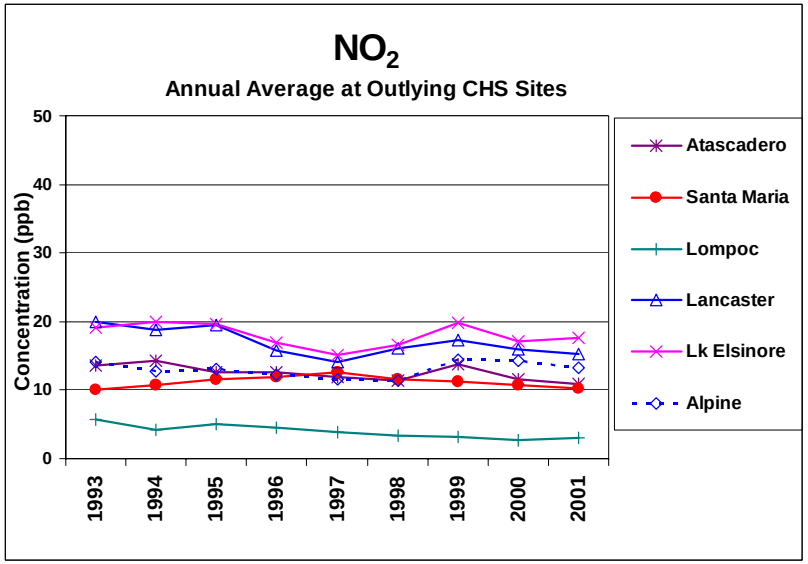
Figure 4.1-4. **Spatial correlation of eight-year average ambient concentrations in the CHS.**

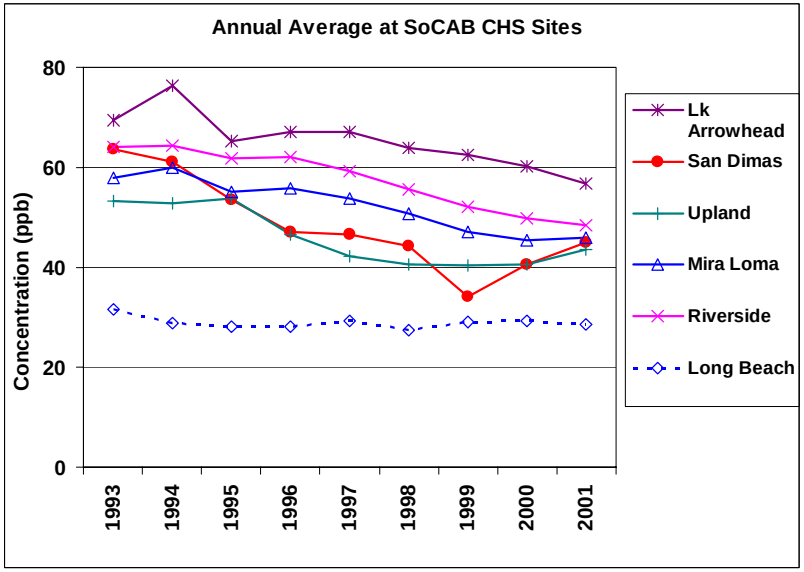
Table 4.1-6. **Pollutants and communities for which the eight-year trend in annual average concentrations show significant decreases (Dec) and increases (Inc).**

Community	Ozone	NO_2	PM_{10}	$PM_{2.5}$	SO_4	NO_3	EC	OC	Formic Acid	Acetic Acid	HNO_3
Atascadero	–	–	–	Inc	–	–	–	–	–	Inc	–
Santa Maria[a]	–	–	Inc[a]	–	–	–	–	–	–	Inc[a]	–
Lompoc	–	Dec[b]	–	–	–	–	–	–	–	Inc[b]	Dec[b]
Lancaster[a]	–	–	–	–	–	–	–	–	–	–	Dec[a]
Lk Arrowhead	Dec	–	–	Dec	Dec	Dec	–	–	–	Inc	Dec
San Dimas[a]	Dec[a]	–	Dec[a]	–	–	–	–	–	Dec[a]	–	Dec[a]
Upland	Dec	–	–	–	–	–	–	–	Dec	–	–
Mira Loma	Dec	–	–	–	–	–	Inc	–	Dec	–	–
Riverside	Dec	–	–	–	–	–	–	–	Dec	Inc	–
Long Beach	–	Dec	–	–	–	–	–	–	Dec	–	Dec
Lake Elsinore	Dec	–	–	–	–	–	–	–	–	–	–
Alpine	Dec	–	–	–	–	–	–	–	–	Inc	–

[a] The air quality monitoring stations relocated during the period.
[b] The levels are so low at Lompoc that the trends are not meaningful.







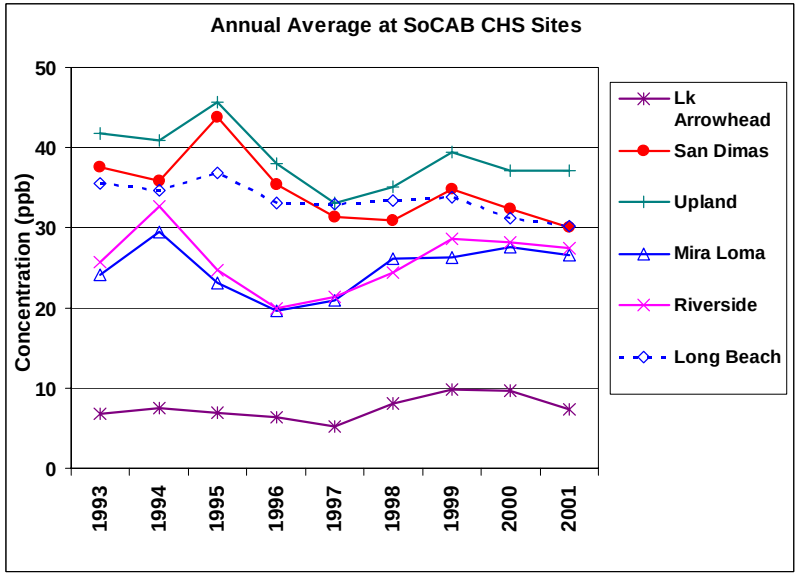


Figure 4.1-5. Temporal trend in annual average 10 AM–6 PM ozone and 24-hr NO_2 concentrations from 1994-2001.

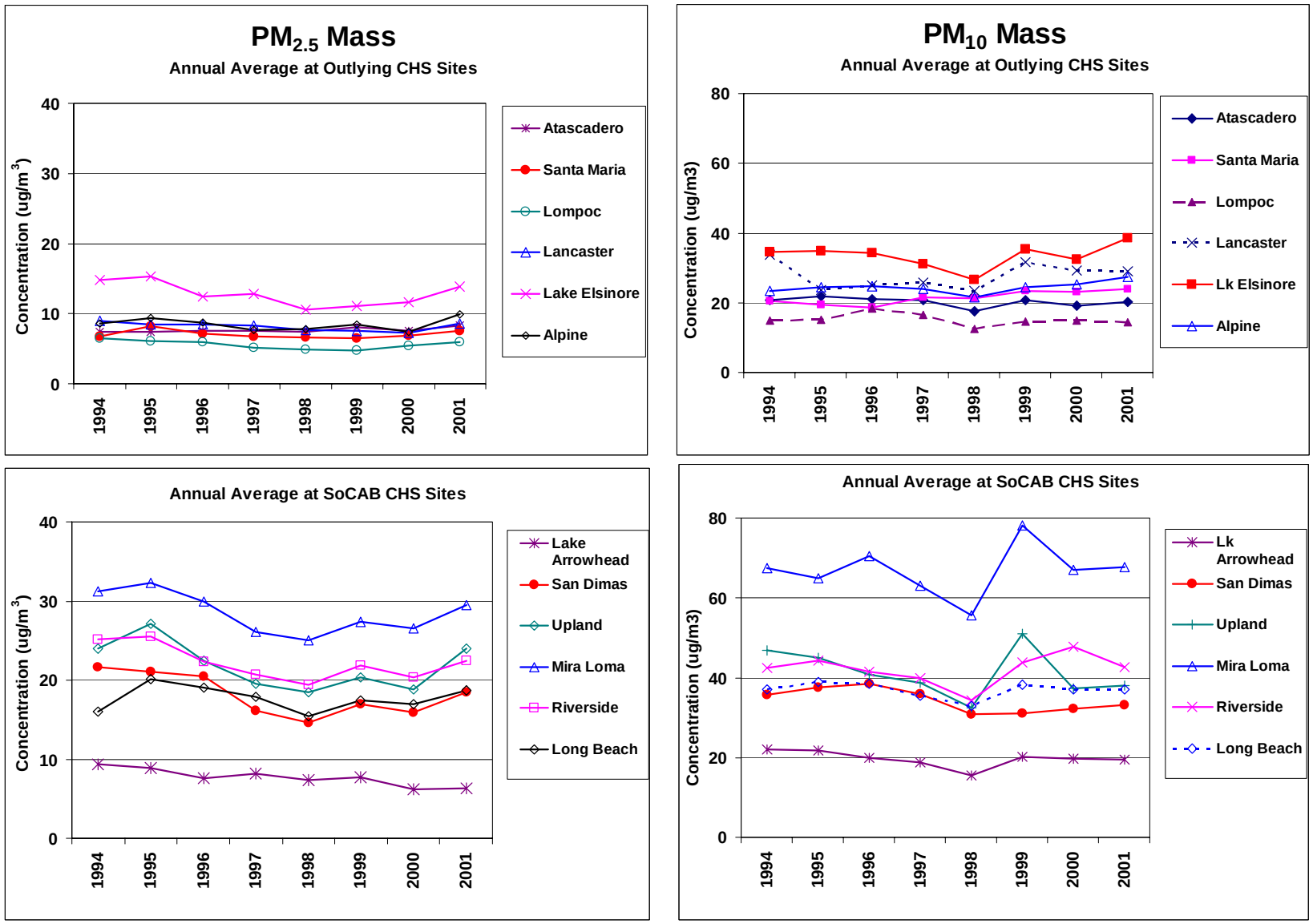


Figure 4.1-6. Temporal trend in annual average $PM_{2.5}$ and PM_{10} mass from 1994-2001.

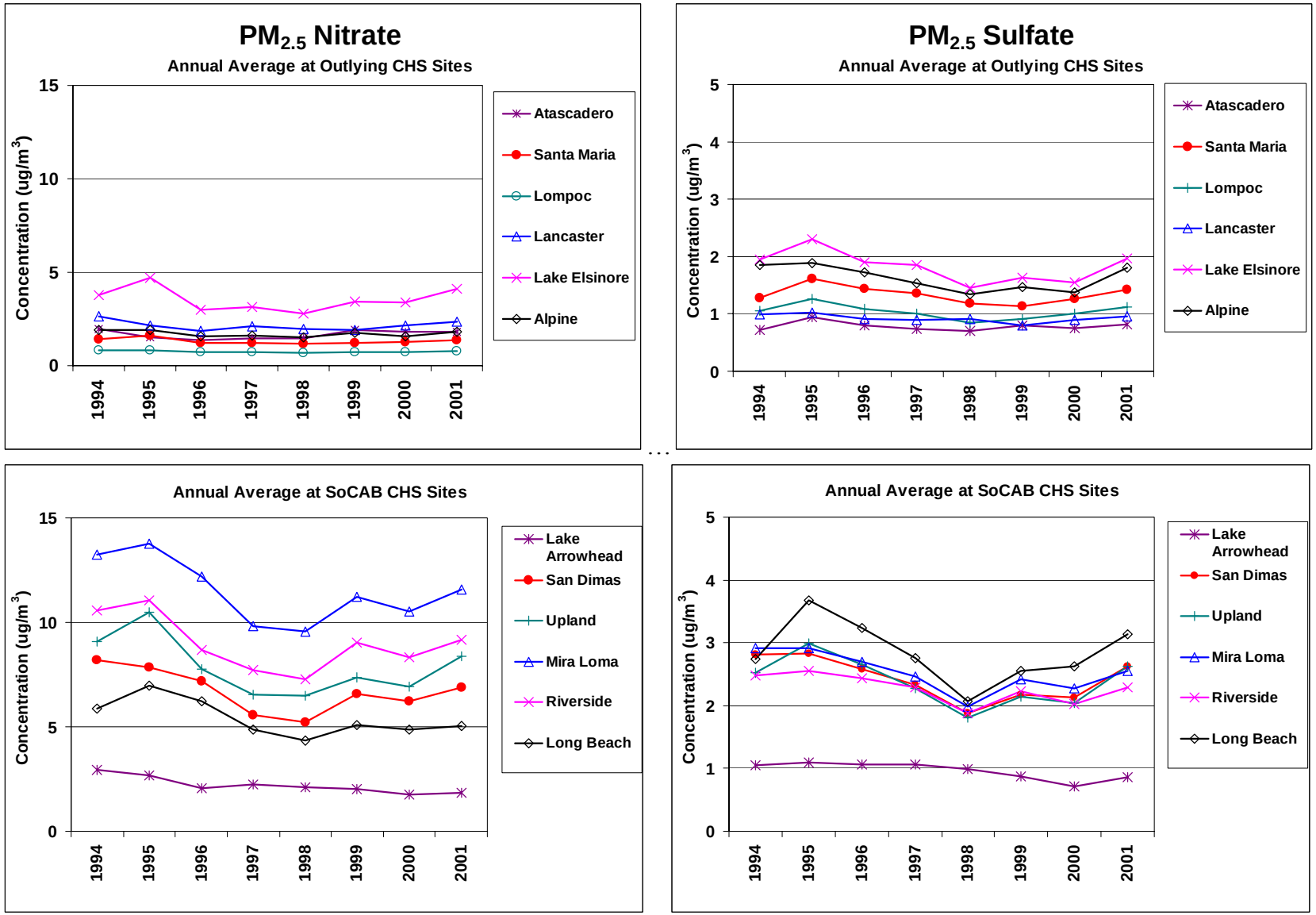


Figure 4.1-7. Temporal trend in annual average $PM_{2.5}$ nitrate and sulfate concentrations from 1994-2001.

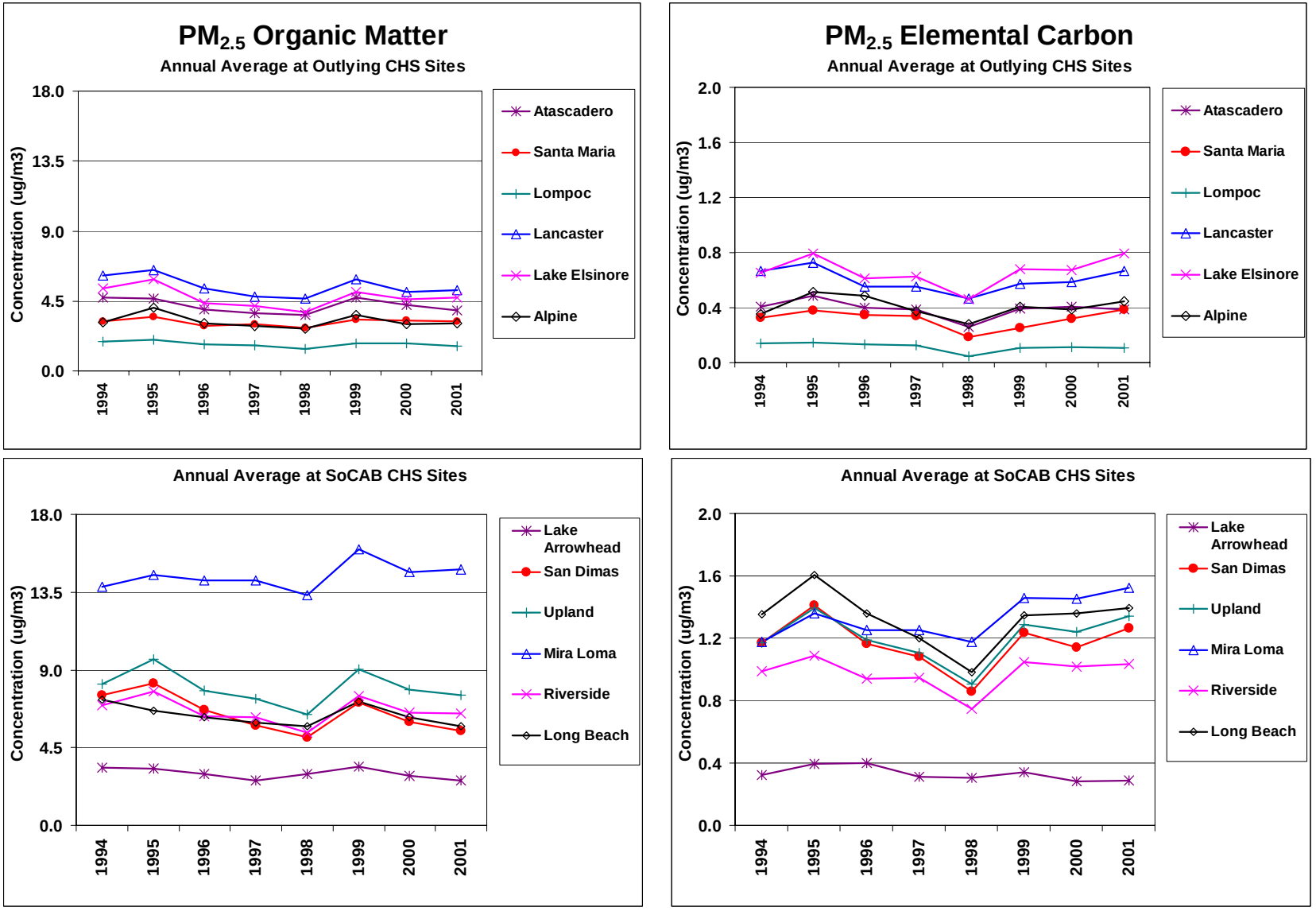


Figure 4.1-8. Temporal trend in annual average $PM_{2.5}$ organic and elemental carbon concentrations from 1994-2001.

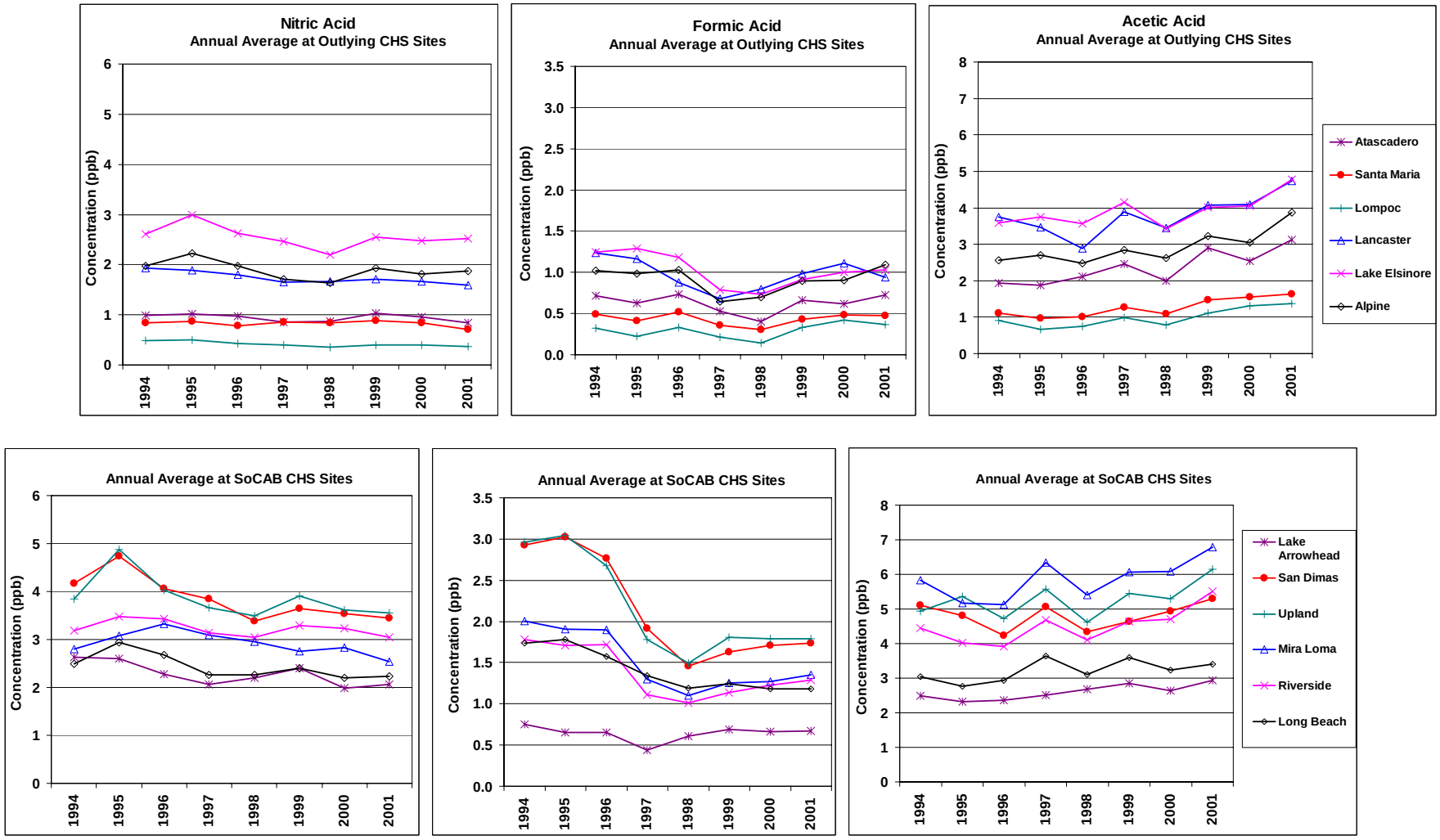


Figure 4.1-9. Temporal trend in annual average nitric, formic, and acetic acid concentrations from 1994-2001.

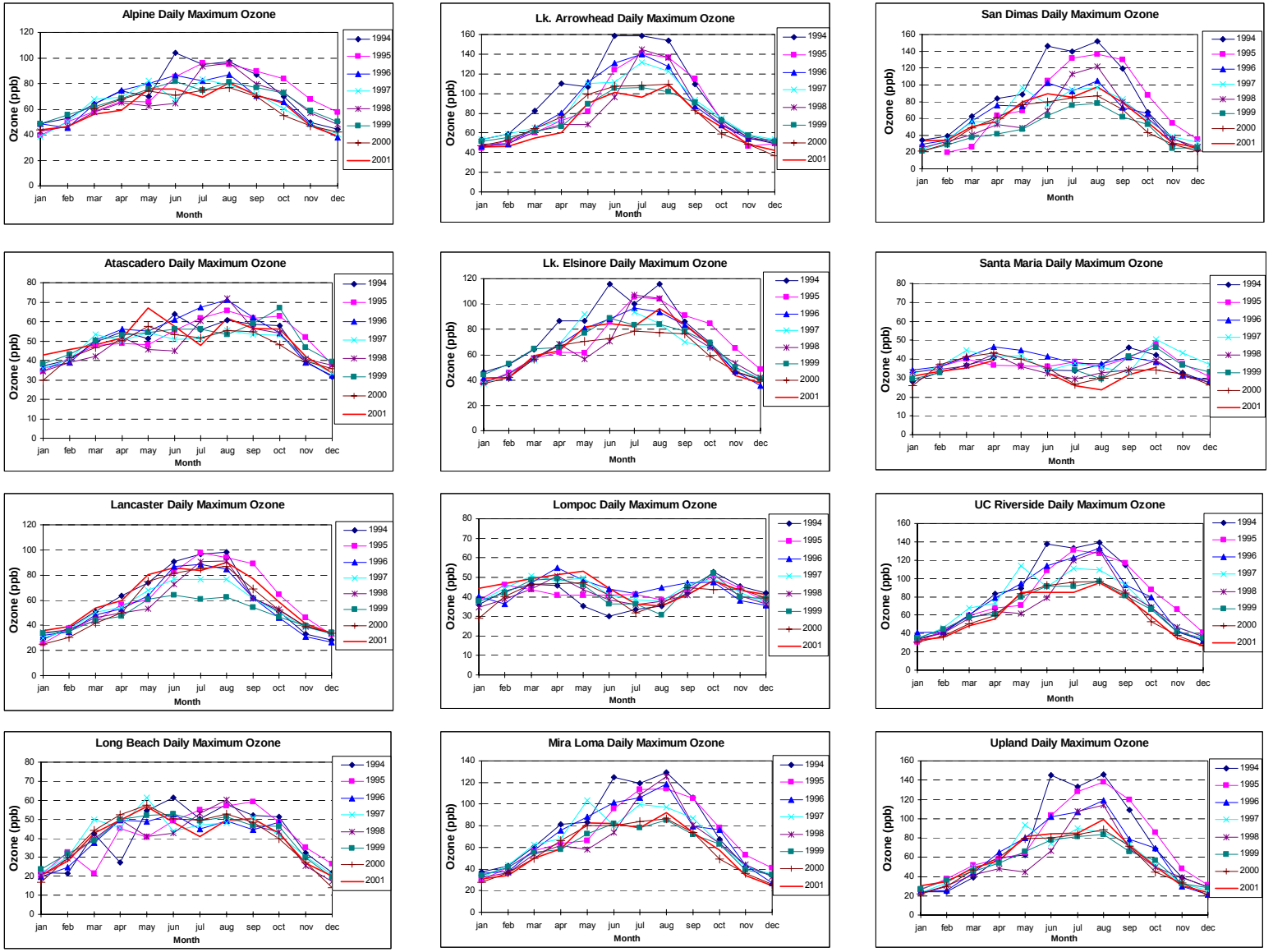


Figure 4.1-10. Monthly average ozone daily 1-hr maximum concentrations for 1994 through 2001.

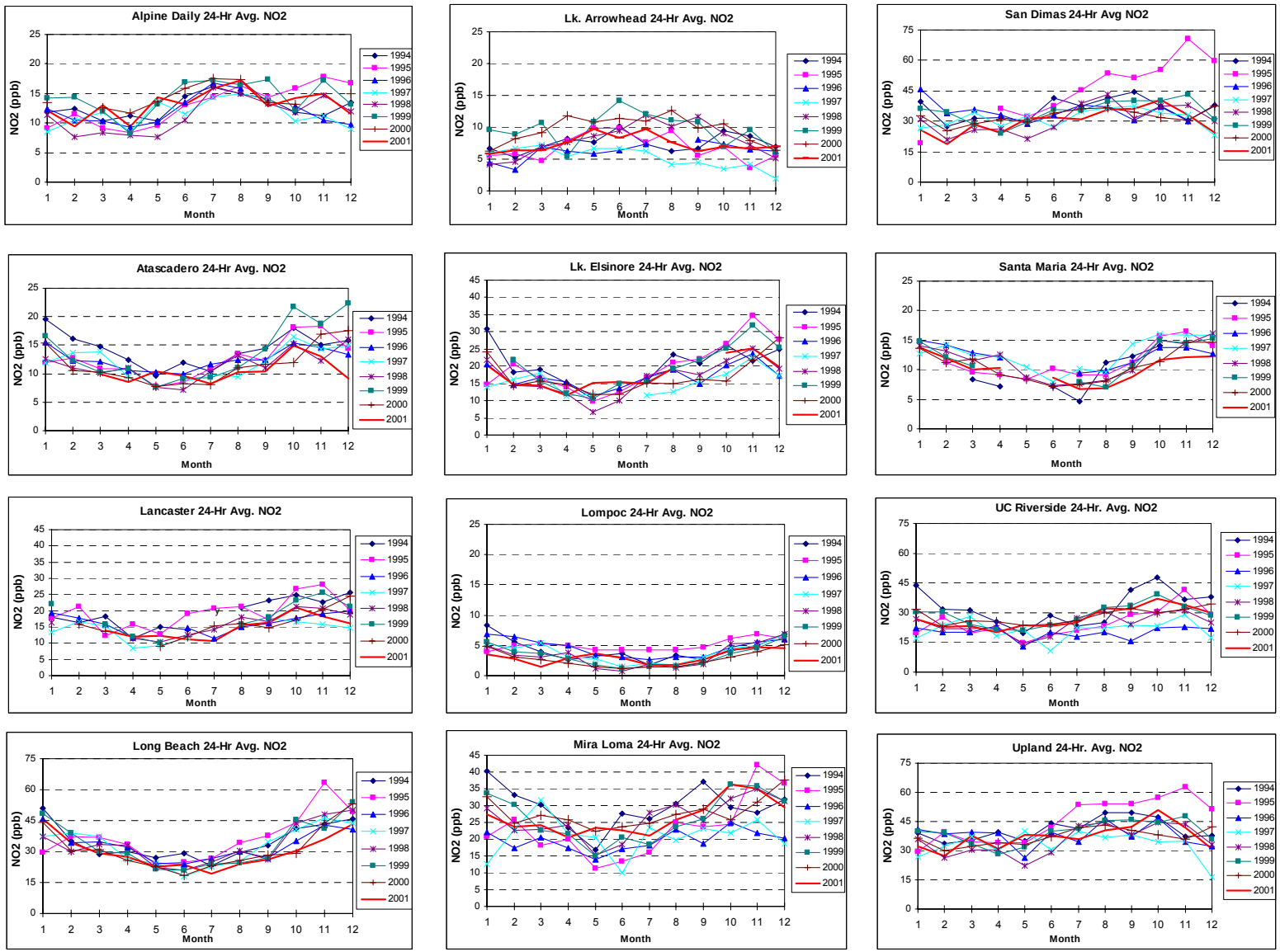


Figure 4.1-11. Monthly 24-hr average NO_2 concentrations for 1994-2001.

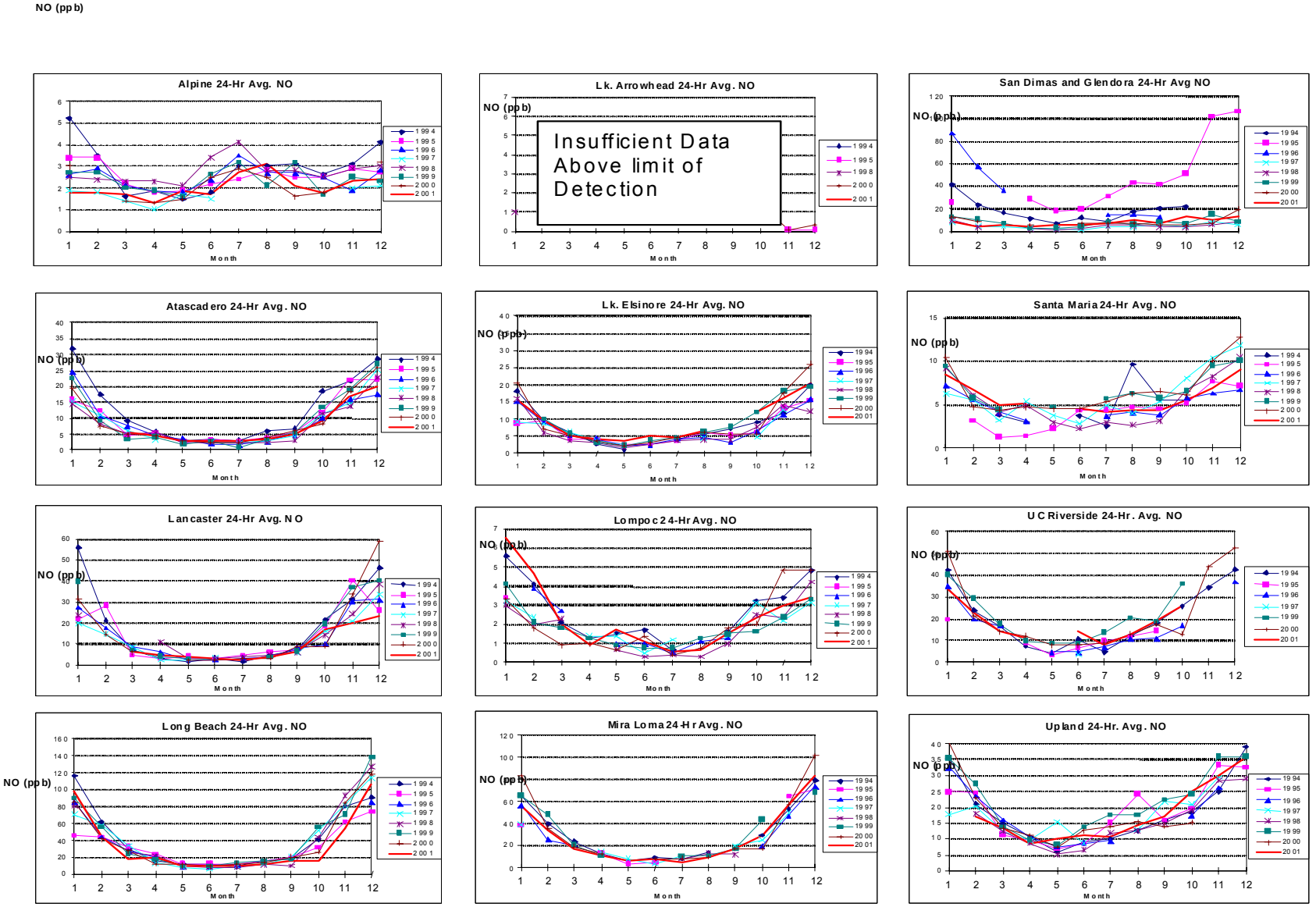


Figure 4.1-12. Monthly 24-hr average NO concentrations for 1994 through 2001.

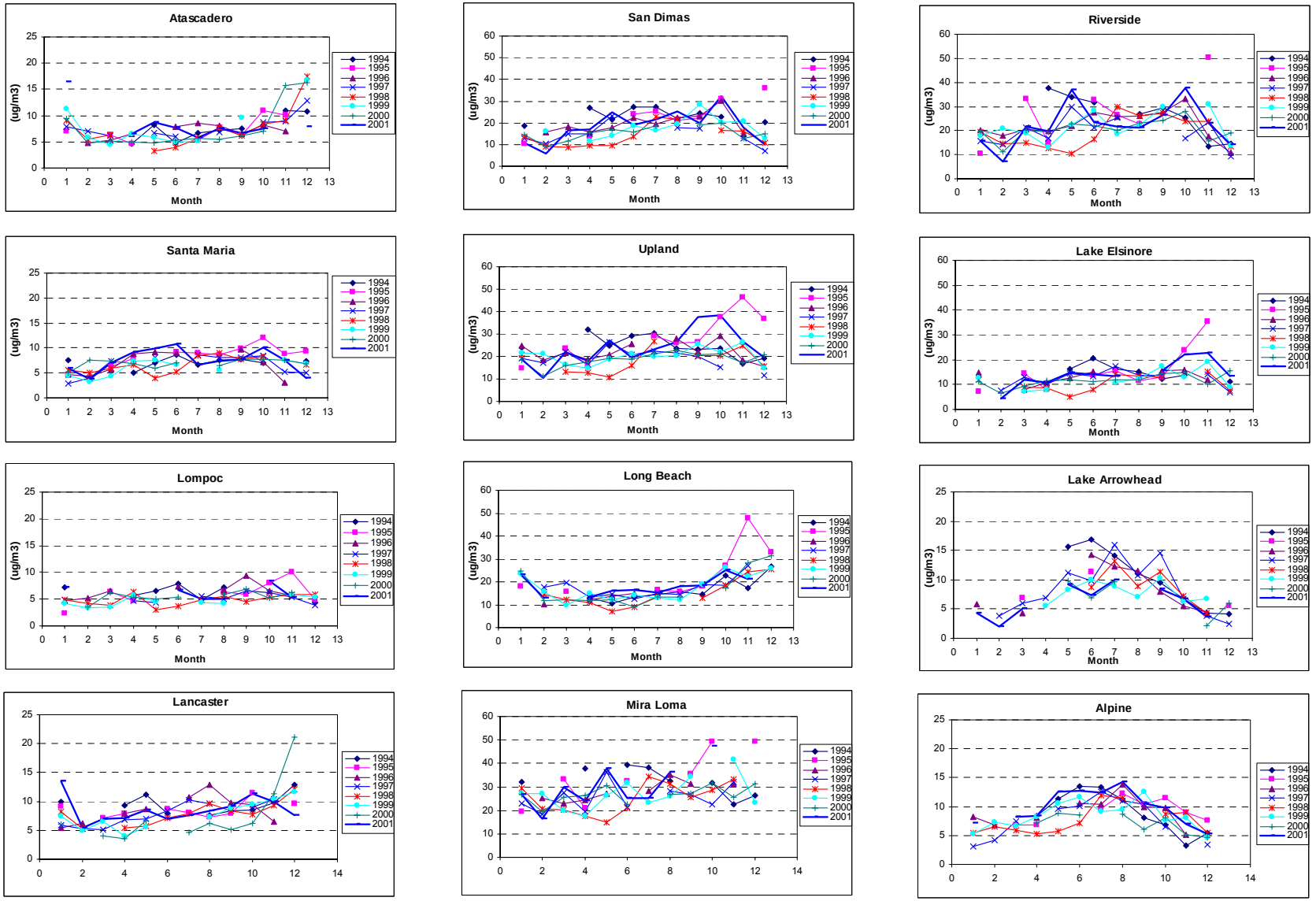


Figure 4.1-13. Monthly 24-hr average $PM_{2.5}$ concentrations for 1994 through 2001.

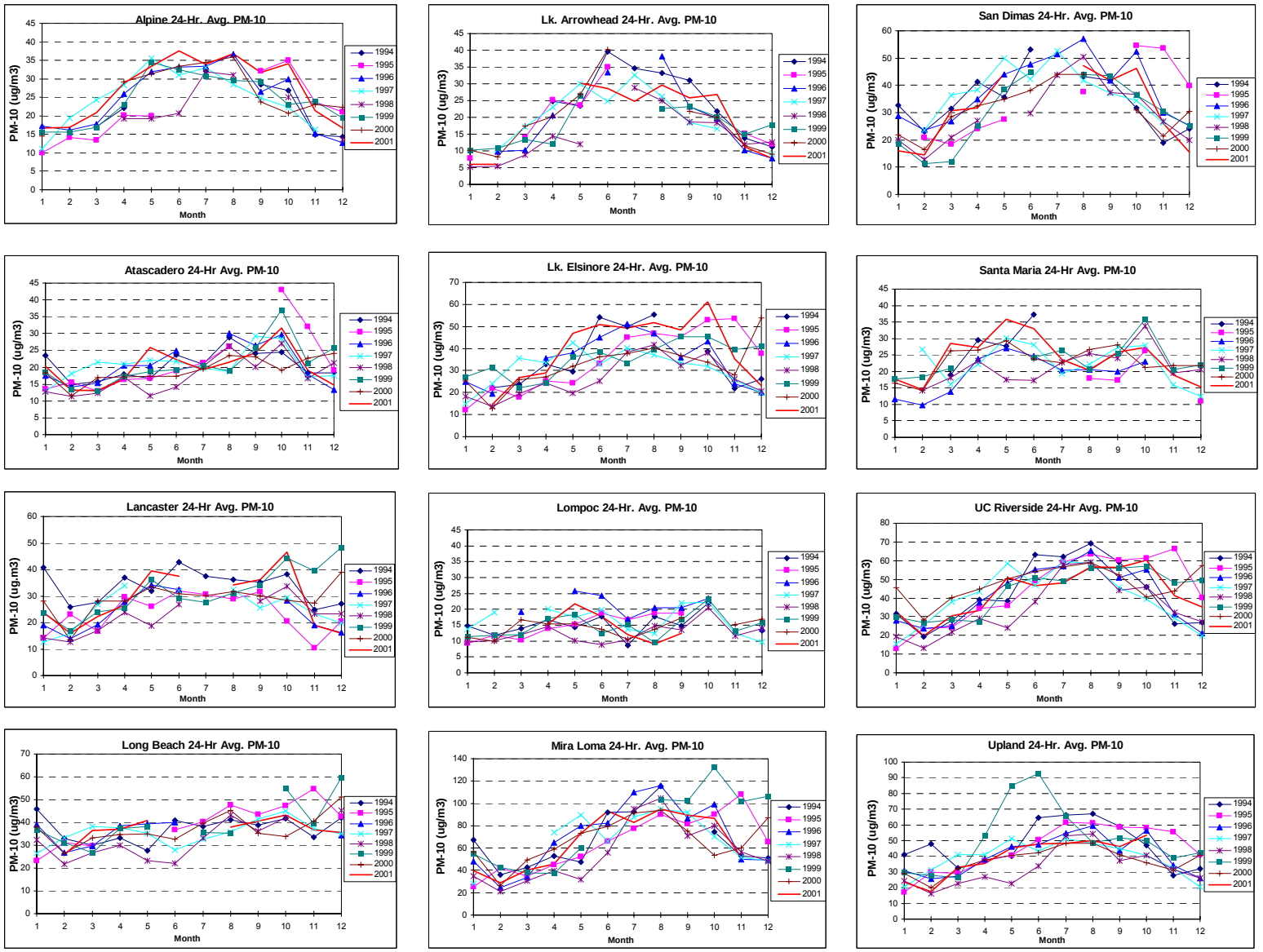


Figure 4.1-14. Monthly 24-hr average PM_{10} concentrations for 1994 through 2001.

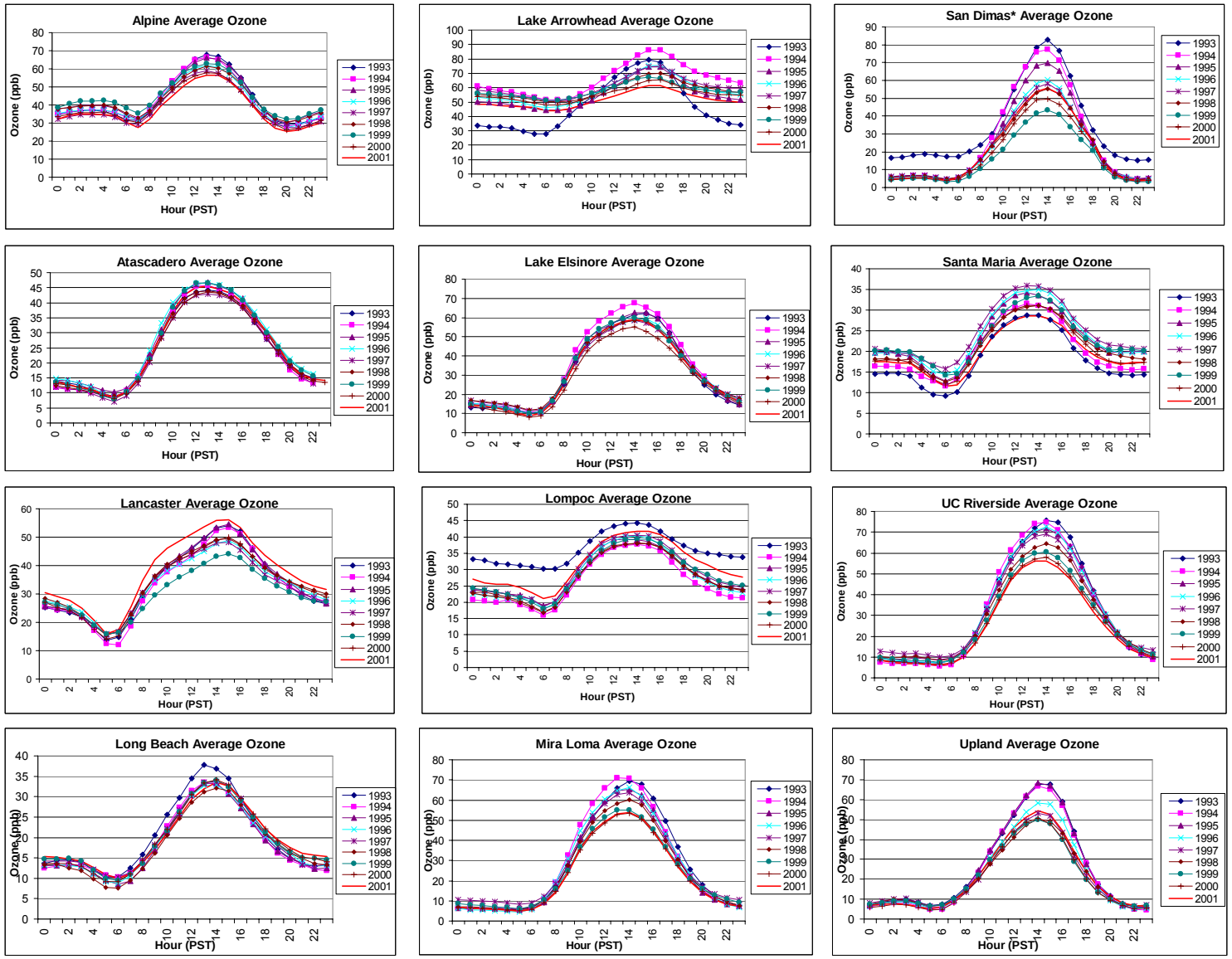


Figure 4.1-15. Average diurnal profile of ozone for 1994 through 2001. The scales vary from site to site.

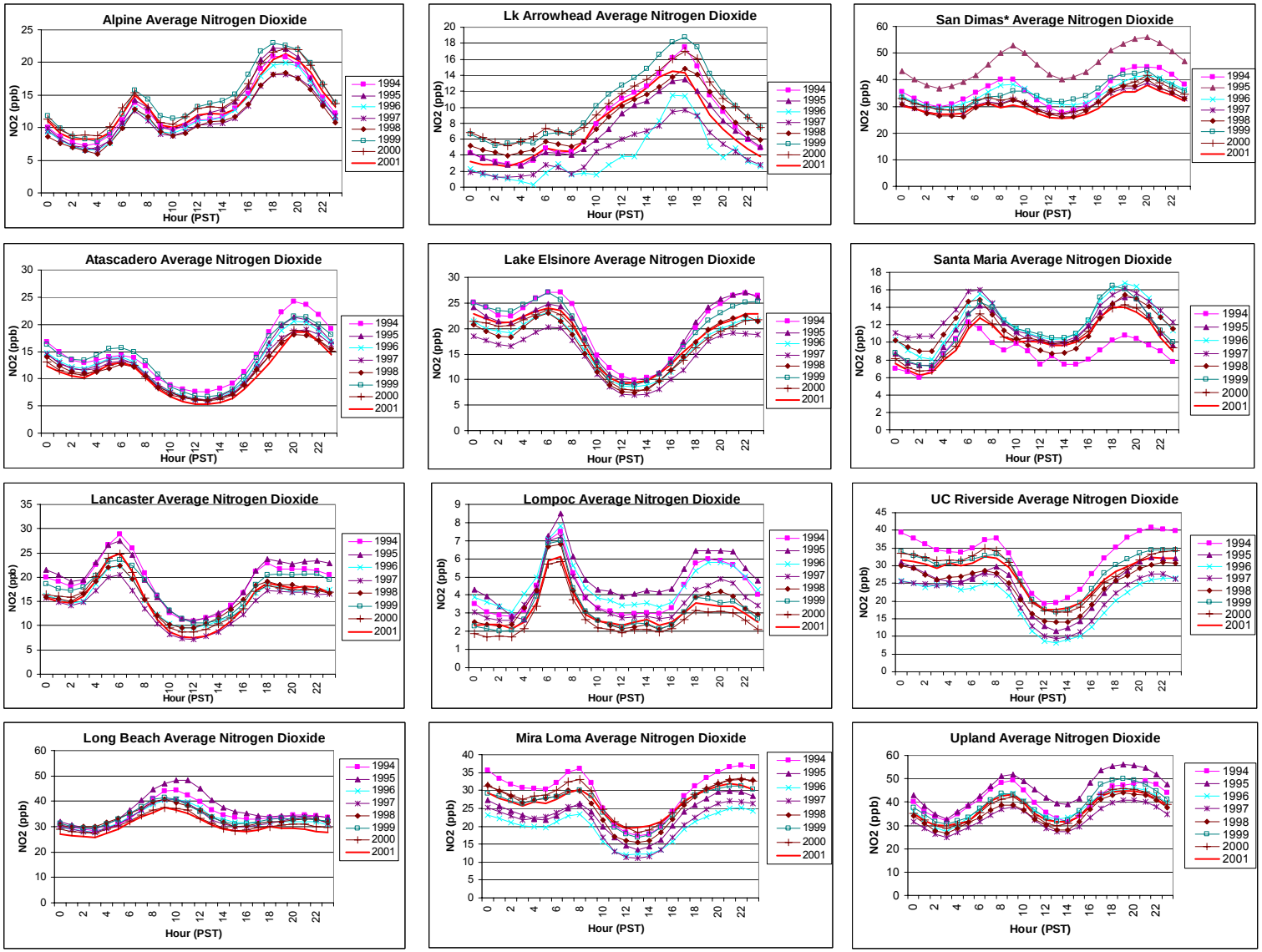


Figure 4.1-16. Average diurnal profile of NO_2 for 1994 through 2001. The scales may vary from site to site.

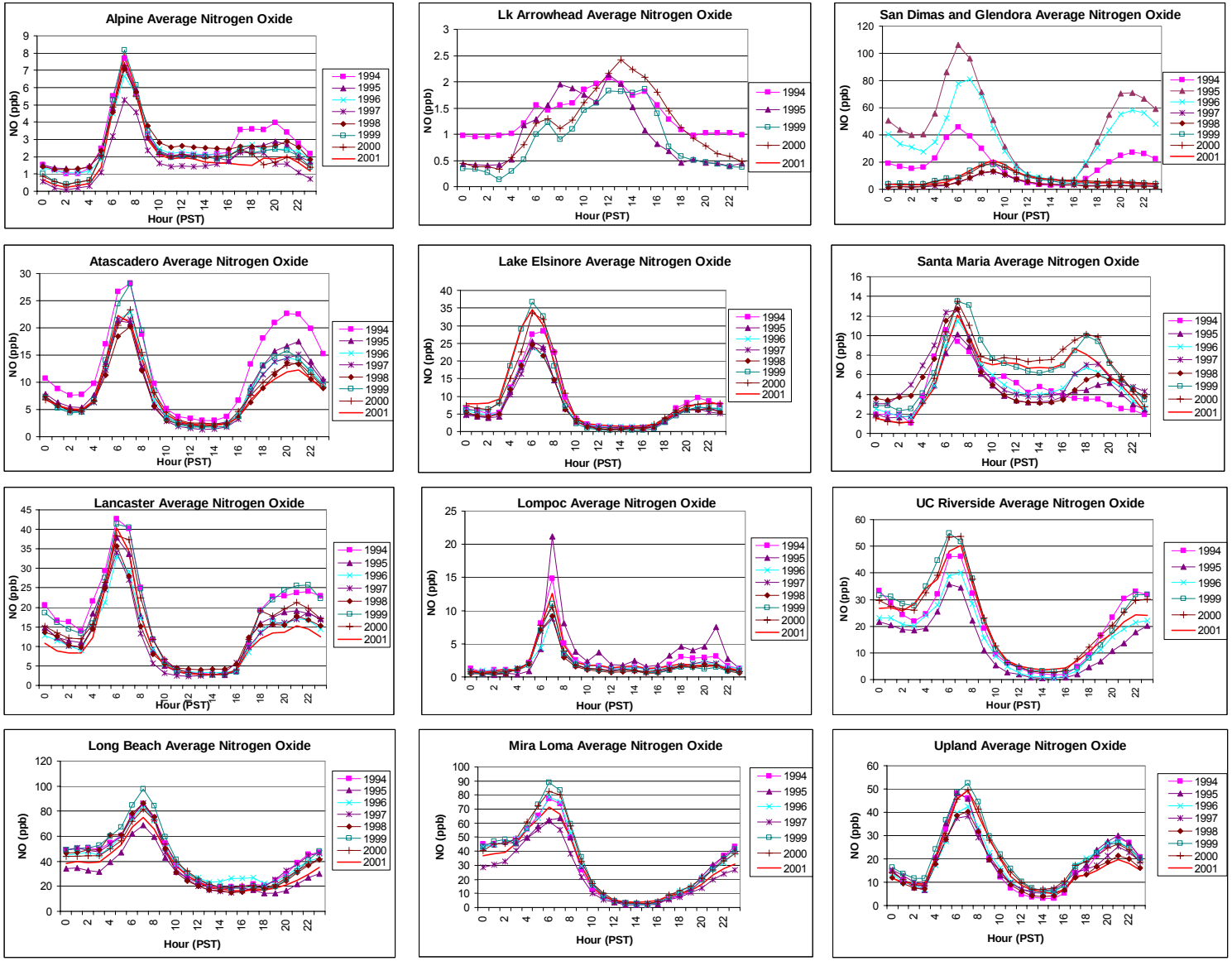


Figure 4.1-17. Average diurnal profile of NO for 1994 through 2001. The scales vary from site to site.

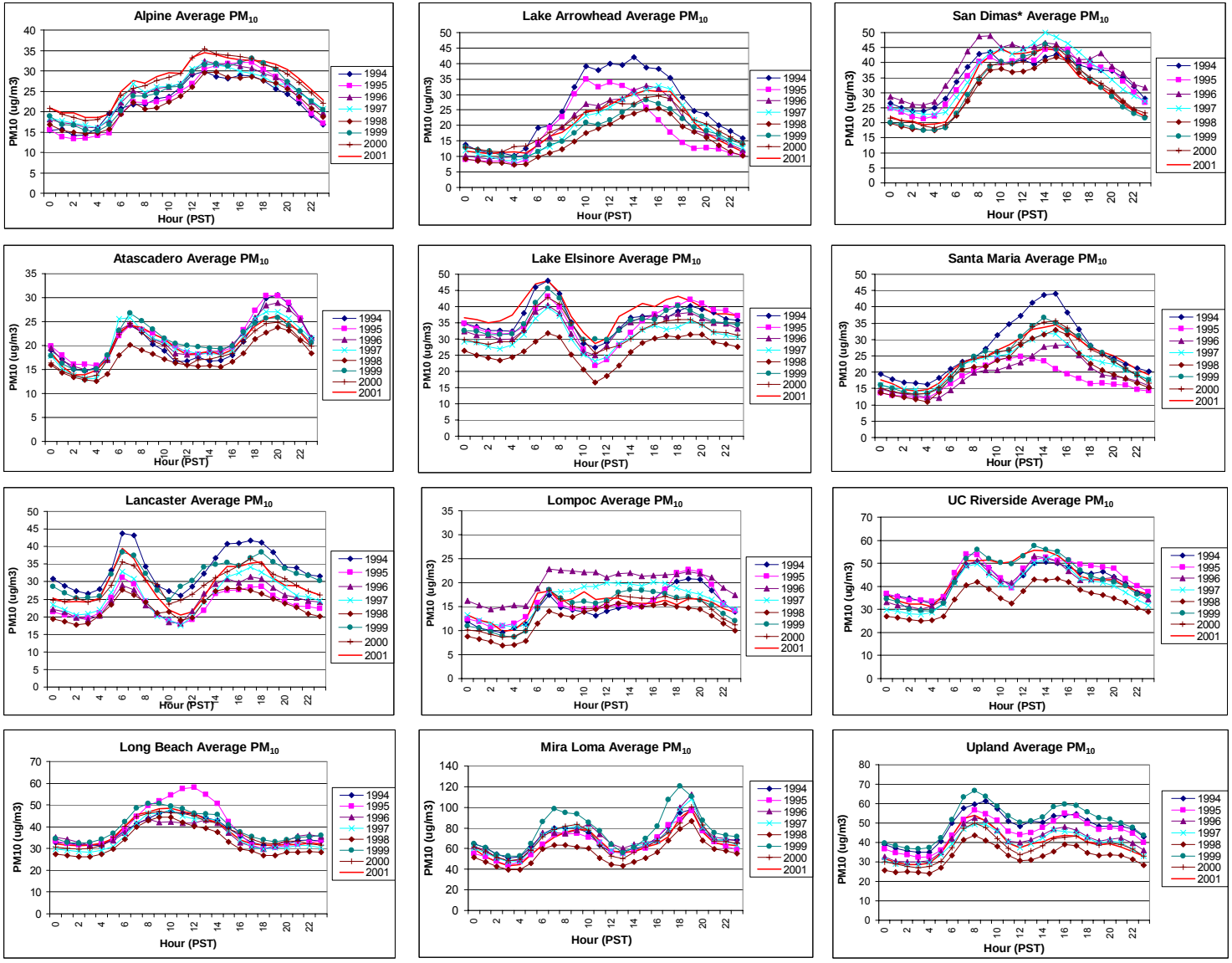


Figure 4.1-18. Average diurnal profile of PM_{10} for 1994 through 2001. The scales may vary from site to site.

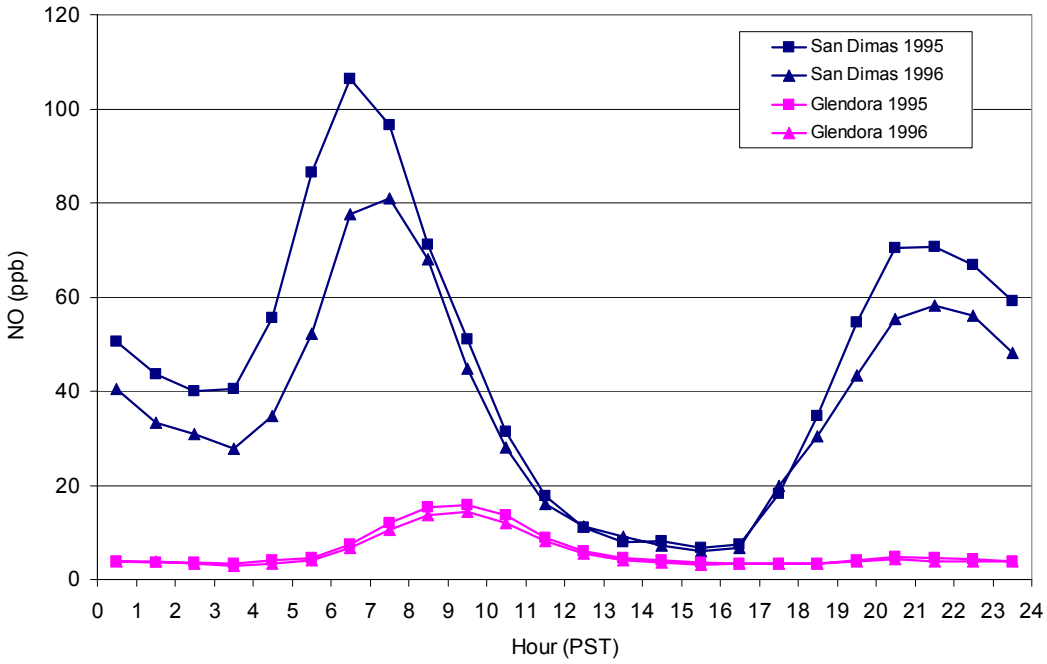


Figure 4.1-19. **Average diurnal profile of NO in 1995 and 1996 measured at the San Dimas and Glendora air monitoring stations in the community of San Dimas.**

Table 4.1-7. Annual average concentrations (ng/m^3) of elements measured by XRF in CHS communities in 2001.

Element	LOD[a]	Community											
		Atascadero	Santa Maria	Lompoc	Lancaster	Long Beach	Glendora	Upland	Mira Loma	Riverside	Lake Elsinore	Lake Arrowhead	Alpine
Na	>>	199.74	292.24	283.92	111.58	230.02	186.69	167.76	152.45	192.68	185.96	150.50	240.72
Mg	>>	36.17	51.61	53.49	39.06	60.66	46.92	32.34	45.18	56.52	39.48	38.13	37.27
Al	4.5	23.71	21.47	14.80	45.98	30.27	36.34	36.46	52.80	36.14	45.63	32.59	31.14
Si	2.8	99.69	81.90	44.12	145.14	79.91	104.59	108.11	185.72	119.56	147.76	91.42	87.35
P	2.5	1.04	1.43	2.32	0.58	1.83	0.70	1.28	1.60	0.71	0.27	1.62	0.39
S	2.2	460.47	752.32	590.27	511.52	1490.1	1338.4	1231.0	1195.6	1098.4	1050.4	536.84	929.83
Cl	4.5	23.55	28.89	57.45	8.14	111.85	11.57	59.10	270.56	43.50	17.58	1.70	10.63
K	2.7	80.62	58.55	44.51	84.81	94.30	171.78	116.87	154.82	88.14	101.75	44.12	50.56
Ca	2.0	35.07	35.43	26.60	53.63	46.90	46.87	63.50	159.17	68.10	46.88	30.06	34.94
Ti	1.3	1.94	1.68	0.69	5.03	7.68	8.74	5.86	8.33	4.88	4.64	2.46	4.13
V	1.1	0.54	1.30	1.03	0.81	12.99	5.31	4.19	5.23	3.91	3.46	1.11	4.14
Cr	0.8	0.13	0.17	0.15	0.31	1.28	0.94	0.87	0.80	0.38	0.52	0.32	0.55
Mn	0.7	1.04	0.90	0.40	2.28	3.34	3.68	4.73	5.99	3.15	2.59	1.84	1.78
Fe	0.7	52.15	40.78	17.70	101.33	115.57	106.60	145.79	187.20	129.02	95.64	50.51	63.28
Co	0.4	0.29	0.28	0.17	0.47	0.79	0.67	0.80	1.01	0.79	0.59	0.30	0.37
Ni	0.4	0.52	0.79	0.73	0.50	4.50	2.30	2.27	2.39	1.89	1.48	0.53	1.78
Cu	0.5	3.27	4.24	3.34	4.67	6.86	6.43	7.83	6.35	4.56	4.19	1.42	2.17
Zn	0.5	6.42	6.03	4.12	8.75	26.65	24.23	33.03	28.55	16.00	11.44	7.10	7.68
Ga	0.8	0.27	0.11	0.20	0.27	0.22	0.37	0.34	0.23	0.27	0.27	0.45	0.31
As	0.7	0.57	0.33	0.29	0.48	0.57	0.32	0.64	1.02	0.34	0.74	0.38	0.35
Se	0.5	0.47	0.59	0.43	0.45	0.90	3.68	2.00	0.86	0.81	1.11	0.29	0.32
Br	0.4	2.47	2.47	2.52	3.27	5.57	6.10	5.90	6.97	5.45	4.77	3.13	3.42
Rb	0.4	0.32	0.36	0.33	0.40	0.34	0.28	0.37	0.44	0.35	0.38	0.28	0.32
Sr	0.5	0.70	0.92	0.60	1.53	1.95	3.40	2.18	2.85	1.69	1.08	0.84	0.97
Y	0.6	0.51	0.56	0.58	0.51	0.42	0.59	0.50	0.43	0.50	0.50	0.46	0.41
Zr	0.8	0.68	0.79	0.83	0.95	0.80	0.80	0.93	1.04	1.06	0.82	0.65	0.96
Mo	1.2	1.36	1.39	1.05	1.42	1.66	1.92	1.73	1.39	1.19	1.24	1.43	1.37
Pd	4.9	0.38	0.35	0.49	0.59	0.24	0.26	0.39	0.29	0.33	0.57	0.37	0.32
Ag	5.4	0.89	1.34	0.89	0.98	0.79	0.75	0.44	0.76	0.67	0.52	0.68	0.42
Cd	5.4	0.91	0.88	0.64	0.76	0.42	1.03	0.63	1.01	0.84	0.35	0.32	0.30

Table 4.1-7. Annual average concentrations (ng/m^3) of elements measured by XRF in CHS communities in 2001.

Element	LOD[a]	Community											
		Atascadero	Santa Maria	Lompoc	Lancaster	Long Beach	Glendora	Upland	Mira Loma	Riverside	Lake Elsinore	Lake Arrowhead	Alpine
In	5.8	1.33	0.71	1.02	1.32	1.41	0.81	0.92	1.71	0.59	0.54	1.05	1.01
Sn	7.6	1.27	0.51	1.40	2.22	4.30	9.89	7.49	3.02	2.66	1.11	1.82	1.83
Sb	8.0	2.03	2.64	2.08	3.25	2.56	4.17	3.49	3.43	2.81	2.47	2.85	3.42
Ba	23.2	13.80	18.17	10.71	14.37	16.50	20.90	21.02	16.19	19.78	12.76	10.79	8.97
La	27.6	17.54	9.84	17.92	13.49	9.23	9.31	11.25	10.53	10.27	11.44	14.32	11.85
Au	1.4	0.29	0.48	0.45	0.50	0.37	0.55	0.36	0.31	0.32	0.33	0.74	0.52
Hg	1.2	0.54	0.50	0.43	0.70	0.37	0.40	0.42	0.32	0.39	0.43	0.32	0.41
Tl	1.1	0.24	0.36	0.41	0.25	0.39	0.37	0.51	0.38	0.29	0.31	0.33	0.57
Pb	1.3	2.00	2.00	1.23	3.04	6.07	9.46	8.72	6.52	5.55	4.17	2.41	3.95
U	1.0	0.69	0.74	0.70	0.64	0.46	0.64	0.70	0.54	0.81	0.43	0.52	0.81

[a] LOD is the analytical limit of detection. The shaded annual average concentrations are below the LOD.

Table 4.1-8. Annual average concentration (ng/m^3) of organic compounds in ambient PM_{10} in 1995 in CHS communities.

Compound	Atascadero	Santa Maria	Lompoc	Lancaster	Lake Arrowhead	San Dimas*	Upland	Mira Loma	RiverSide	Long Beach	Lake Elsinore	Alpine
	Alkanes											
Tetracosane	2.95	2.48	1.29	1.92	2.17	2.89	4.15	2.92	3.12	5.07	2.29	2.04
Pentacosane	1.98	1.33	0.46	2.05	0.81	3.72	3.52	3.90	3.27	5.98	2.18	1.47
Hexacosane	1.65	1.52	0.52	1.89	0.65	4.24	3.21	3.66	4.08	6.05	1.83	1.23
Heptacosane	2.66	2.11	0.88	2.24	1.21	4.05	4.12	5.08	5.26	5.36	2.70	2.29
Octacosane	1.59	1.90	0.78	1.97	0.96	3.25	2.87	4.01	4.59	4.52	1.69	1.59
Nonacosane	5.34	4.00	1.65	4.01	2.32	5.60	6.31	9.40	7.78	5.61	4.77	4.37
Triacontane	1.00	1.14	0.59	1.24	0.41	1.86	2.09	2.58	2.97	2.61	1.12	0.81
Hentriacontane	3.30	2.33	0.98	2.82	1.20	3.92	4.56	11.21	5.72	3.81	3.00	2.16
Dotriacontane	0.61	0.63	0.34	0.83	ND	1.23	1.38	1.73	1.82	1.66	0.70	0.46
Tritriacontane	0.63	0.74	0.30	1.15	ND	1.30	1.64	2.49	1.85	1.80	1.02	0.69
Tetratriacontane	0.56	0.84	0.18	1.66	ND	0.41	0.69	1.70	1.94	2.17	1.01	0.17
Total	22.25	19.02	7.99	21.77	9.74	32.46	34.55	48.70	42.40	44.65	22.30	17.27
	Saturated Cycloalkanes											
Pentadecylcyclohexane	0.042	0.036	0.013	0.055	0.028	0.066	0.064	0.033	0.061	0.053	0.034	0.029
Hexadecylcyclohexane	0.078	0.055	0.021	0.089	0.027	0.061	0.091	0.071	0.078	0.073	0.034	0.033
Heptadecylcyclohexane	0.065	0.065	0.021	0.088	0.032	0.080	0.100	0.094	0.085	0.113	0.044	0.036
Octadecylcyclohexane	0.097	0.073	0.025	0.091	0.032	0.100	0.116	0.109	0.080	0.136	0.052	0.034
Total	0.281	0.228	0.080	0.324	0.119	0.308	0.370	0.306	0.304	0.375	0.164	0.132
	Alkanoic Acids											
Tetradecanoic acid	6.69	4.90	2.72	4.41	3.60	6.61	8.67	7.59	5.59	6.37	5.48	5.65
Pentadecanoic acid	3.87	2.88	1.78	2.75	2.47	4.97	5.42	4.69	3.62	3.66	3.81	3.99
Hexadecanoic acid	33.08	23.34	8.60	21.64	5.47	48.74	88.54	63.37	33.06	47.41	20.56	18.97
Heptadecanoic acid	2.84	2.03	1.08	2.28	1.42	3.89	5.94	5.27	3.07	3.50	2.50	2.44
Octadecanoic acid	15.79	12.06	3.85	13.84	5.07	27.74	52.21	106.50	29.54	25.92	11.15	9.15
Nonadecanoic acid	0.93	0.61	0.25	0.76	0.35	0.92	1.30	1.28	0.95	1.07	0.87	0.90
Eicosanoic acid	2.60	1.42	0.64	1.80	0.69	2.11	2.77	5.05	2.18	2.28	1.64	1.69
Heneicosanoic acid	1.07	0.49	0.25	0.72	0.17	0.55	0.78	1.15	0.64	0.86	0.54	0.66
Docosanoic acid	5.28	2.21	1.14	2.99	0.77	2.01	2.80	5.41	2.08	2.33	1.72	2.32

***** San Dimas site values are the average of the second and third sampling periods.

Table 4.1-8. Annual average concentration (ng/m^3) of organic compounds in ambient PM_{10} in 1995 in CHS communities.

Compound	Atascadero	Santa Maria	Lompoc	Lancaster	Lake Arrowhead	San Dimas*	Upland	Mira Loma	RiverSide	Long Beach	Lake Elsinore	Alpine
	Alkanoic Acids (continued)											
Tricosanoic acid	1.31	0.54	0.24	0.82	0.15	0.50	0.73	1.69	0.55	0.79	0.50	0.83
Tetracosanoic acid	6.60	2.43	1.40	2.70	0.70	2.00	2.61	5.81	2.16	1.94	2.08	2.44
Pentacosanoic acid	0.54	0.25	0.11	0.47	0.10	0.22	0.34	0.92	0.29	0.28	0.36	0.39
Hexacosanoic acid	2.44	0.96	0.40	1.55	0.42	1.07	1.35	4.54	1.13	0.80	1.74	1.22
Heptacosanoic acid	0.15	0.03	ND	0.23	0.04	0.08	0.08	0.69	0.13	ND	0.13	0.19
Octacosanoic acid	0.75	0.40	0.12	1.02	0.23	0.46	0.53	4.22	0.71	0.42	0.98	0.60
Nonacosanoic acid	ND	ND	ND	0.13	ND	0.02	0.03	1.24	0.05	ND	0.10	0.05
Triacontanoic acid	0.40	0.16	0.03	0.72	0.11	0.34	0.32	9.93	0.52	0.19	0.43	0.29
Total	84.34	54.71	22.60	58.83	21.75	102.21	174.43	229.34	86.27	97.82	54.58	51.76
	Alkenoic Acids											
9-Hexadecenoic acid	0.857	1.463	0.915	0.935	1.147	0.333	0.375	0.185	0.813	0.126	0.433	1.157
9,12-Octadecadienoic acid	0.399	0.074	0.020	0.342	0.085	ND	0.029	ND	0.291	0.018	0.011	0.162
9-Octadecenoic acid	0.887	1.620	1.273	0.997	1.091	0.281	0.503	1.000	0.928	0.504	0.315	1.359
Total	2.142	3.157	2.208	2.273	2.323	0.614	0.907	1.185	2.031	0.649	0.758	2.679
	Resin Acids											
Pimaric acid	0.183	0.050	0.009	0.132	0.046	ND	0.179	0.102	0.054	0.132	0.043	0.042
Sandaracopimaric acid	0.076	0.016	ND	0.098	0.017	ND	ND	0.013	0.060	0.105	0.029	0.034
Isopimaric acid	0.138	ND	ND	0.158	0.050	ND	0.293	0.922	0.114	ND	ND	0.149
Dehydroabietic acid	17.139	5.127	3.231	10.357	1.063	5.568	9.053	8.976	4.568	12.072	2.960	1.710
Abietic acid	0.060	ND	ND	0.079	ND	ND	ND	ND	ND	ND	ND	ND
Abieta-6,8,11,13,15-pentaen-18-oic acid	0.400	0.136	0.017	0.311	0.021	0.160	0.505	0.423	0.385	0.448	0.147	0.037
Abieta-8,11,13,15-tetraen-18-oic acid	0.464	0.149	0.096	0.260	0.073	0.239	0.536	0.584	0.366	0.488	0.122	0.082
7-Oxodehydroabietic acid	35.849	13.260	12.209	15.134	5.542	13.298	24.808	22.888	14.675	31.570	15.955	8.420
Total	54.307	18.739	15.562	26.529	6.811	19.265	35.374	33.908	20.221	44.813	19.256	10.473

* San Dimas site values are the average of the second and third sampling periods.

Table 4.1-8. Annual average concentration (ng/m^3) of organic compounds in ambient PM_{10} in 1995 in CHS communities.

Compound	Atascadero	Santa Maria	Lompoc	Lancaster	Lake Arrowhead	San Dimas*	Upland	Mira Loma	RiverSide	Long Beach	Lake Elsinore	Alpine
	Alkanedioic Acids											
Propanedioic acid	2.22	0.76	0.44	1.71	8.02	5.66	2.85	1.19	6.97	4.20	3.45	5.32
Butanedioic acid	17.93	13.50	8.95	12.62	24.44	4.48	7.14	8.99	20.12	30.63	12.31	18.60
Methylbutanedioic acid	5.00	4.71	2.81	4.04	7.09	3.18	2.69	3.30	5.95	6.52	3.19	5.54
Pentanedioic acid	10.13	8.50	5.81	7.97	16.24	7.01	7.55	5.61	10.15	12.08	7.96	10.69
Hexanedioic acid	4.46	3.12	2.36	4.30	6.69	2.88	3.31	2.35	4.65	3.46	3.32	4.33
Heptanedioic acid	2.93	1.68	1.65	2.17	1.89	0.81	1.22	0.79	1.61	0.89	1.14	1.68
Octanedioic acid	6.87	3.96	3.01	4.78	4.45	1.86	2.71	2.71	3.29	1.71	2.46	3.24
Nonanedioic acid	13.13	5.78	3.27	8.07	2.60	3.36	6.70	3.87	4.27	3.32	2.58	3.22
Total	60.46	41.26	27.87	43.95	63.40	23.58	31.32	27.62	50.05	58.60	32.96	47.30
	Aromatic Acids											
1,2-Benzenedicarboxylic acid	4.133	1.974	1.389	2.988	8.172	4.541	4.073	3.193	4.514	3.515	3.010	4.606
1,4-Benzenedicarboxylic acid	0.658	0.700	0.297	0.555	0.691	0.292	0.303	0.358	0.380	0.256	0.338	0.514
1,3-Benzenedicarboxylic acid	0.449	0.385	0.235	0.341	0.616	0.322	0.287	0.255	0.377	0.319	0.233	0.359
4-Methyl-1,2-Benzenedicarboxylic acid	1.416	0.865	0.470	1.485	3.157	1.926	1.857	1.670	2.212	1.402	1.314	1.818
Benzenetricarboxylic acids	0.478	0.197	0.104	0.301	1.004	1.065	0.637	0.118	0.590	0.523	0.667	0.930
Total	7.134	4.120	2.496	5.670	13.640	8.147	7.157	5.595	8.073	6.015	5.562	8.228
	Hopanes											
22,29,30-Trisnorhopane	0.066	0.040	0.037	0.083	0.046	0.056	0.067	0.096	0.068	0.125	0.055	0.054
17α(H),21β(H),29-Norhopane	0.285	0.248	0.151	0.322	0.162	0.340	0.351	0.483	0.329	0.543	0.311	0.253
18α(H),29-Norneohopane	0.102	0.092	0.049	0.110	0.059	0.086	0.122	0.135	0.139	0.157	0.108	0.080
17α(H),21β(H)-Hopane	0.346	0.362	0.213	0.430	0.208	0.537	0.538	0.673	0.435	0.704	0.412	0.301
22S, 17α(H),21β(H),30-Homohopane	0.152	0.153	0.080	0.184	0.172	0.179	0.222	0.286	0.205	0.305	0.179	0.142
22R, 17α(H),21β(H),30-Homohopane	0.094	0.165	0.115	0.165	0.080	0.173	0.172	0.306	0.202	0.294	0.166	0.142
22S, 17α(H),21β(H),30,31-Bishomohopane	0.057	0.086	0.033	0.112	0.061	0.108	0.128	0.193	0.133	0.222	0.097	0.081
22R, 17α(H),21β(H),30,31-Bishomohopane	0.043	0.052	0.021	0.098	0.028	0.068	0.073	0.138	0.104	0.144	0.085	0.073
Total	1.144	1.198	0.700	1.505	0.816	1.547	1.672	2.310	1.616	2.493	1.414	1.127

***** San Dimas site values are the average of the second and third sampling periods.

Table 4.1-8. Annual average concentration (ng/m^3) of organic compounds in ambient PM_{10} in 1995 in CHS communities.

Compound	Atascadero	Santa Maria	Lompoc	Lancaster	Lake Arrowhead	San Dimas*	Upland	Mira Loma	RiverSide	Long Beach	Lake Elsinore	Alpine
	Steranes											
20R+S, 5α(H),14β(H),17β(H)-Cholestane	0.277	0.287	0.127	0.304	0.106	0.345	0.334	0.403	0.288	0.470	0.222	0.198
20R, 5α(H),14α(H),17α(H)-Cholestane	0.205	0.190	0.096	0.234	0.083	0.251	0.250	0.359	0.239	0.390	0.188	0.180
20R+S, 5α(H),14β(H),17β(H)-Ergostane	0.169	0.174	0.061	0.180	0.055	0.202	0.229	0.309	0.226	0.375	0.172	0.131
20R+S, 5α(H),14β(H),17β(H)-Sitostane	0.266	0.276	0.121	0.263	0.078	0.315	0.311	0.437	0.276	0.512	0.201	0.177
Total	0.917	0.926	0.405	0.981	0.323	1.114	1.124	1.508	1.030	1.748	0.782	0.687
	PAHs											
Fluoranthene	0.070	0.092	0.007	0.091	0.042	0.132	0.153	0.107	0.117	0.141	0.050	0.038
Acephenanthrylene	0.001	ND	ND	0.006	ND	0.004	0.006	0.007	0.006	0.004	0.004	ND
Pyrene	0.064	0.067	0.007	0.098	0.030	0.139	0.158	0.108	0.108	0.143	0.047	0.033
Retene	0.624	ND	ND	0.443	ND	ND	0.727	0.638	0.242	0.808		ND
Benzo[ghi]fluoranthene	0.041	0.034	0.003	0.066	ND	0.073	0.099	0.065	0.048	0.081	0.028	0.012
Cyclopenta[cd]pyrene	0.010	ND	ND	0.014	0.004	0.016	0.013	0.020	0.007	0.028	0.006	0.004
Benz[a]anthracene	0.014	0.011	0.000	0.032	0.037	0.037	0.040	0.059	0.030	0.185	0.016	0.008
Chrysene/Triphenylene	0.101	0.075	0.017	0.128	0.057	0.146	0.168	0.160	0.124	0.332	0.054	0.028
Benzo[k]fluoranthene	0.131	0.058	0.012	0.182	0.479	0.143	0.173	0.231	0.124	0.334	0.092	0.027
Benzo[b]fluoranthene	0.083	0.057	0.009	0.178	0.666	0.109	0.151	0.231	0.116	0.411	0.098	0.046
Benzo[j]fluoranthene	0.002	ND	ND	ND	ND	ND	ND	ND	ND	0.002	0.005	ND
Benzo[e]pyrene	0.088	ND	ND	ND	ND	0.180	0.051	0.052	ND	0.075	ND	ND
Benzo[a]pyrene	0.028	0.029	ND	0.049	0.280	0.050	0.059	0.092	0.055	0.213	0.011	0.003
Perylene	0.002	ND	ND	0.004	0.008	ND	0.001	ND	0.002	0.005	0.001	ND
Indeno[cd]pyrene	0.031	0.031	0.003	0.100	0.088	0.065	0.115	0.164	0.087	0.148	0.050	0.031
Benzo[ghi]perylene	0.074	0.069	0.002	0.251	0.340	0.195	0.271	0.286	0.157	0.314	0.110	0.062
Indeno[cd]fluoranthene	0.014	ND	ND	0.031	ND	0.016	0.039	0.028	0.021	0.027	0.017	0.004
Coronene	0.040	0.036	0.000	0.110	ND	0.121	0.184	0.109	0.055	0.105	0.013	0.004
Total	1.417	0.558	0.060	1.785	2.031	1.426	2.410	2.357	1.297	3.358	0.602	0.300

* San Dimas site values are the average of the second and third sampling periods.

Table 4.1-8. Annual average concentration (ng/m^3) of organic compounds in ambient PM_{10} in 1995 in CHS communities.

Compound	Atascadero	Santa Maria	Lompoc	Lancaster	Lake Arrowhead	San Dimas*	Upland	Mira Loma	RiverSide	Long Beach	Lake Elsinore	Alpine
	Oxy-PAHs											
1H-Phenalen-1-one	0.064	0.038	0.004	0.067	0.012	0.051	0.068	0.029	0.038	0.079	0.022	0.003
Anthracen-9,10-dione	0.112	0.202	0.027	0.132	0.011	0.140	0.198	0.121	0.159	0.218	0.077	0.075
1,8-Naphthalic Anhydride	0.221	0.250	0.031	0.207	0.030	0.894	0.932	0.517	0.620	0.816	0.400	0.266
Benz[de]anthracen-7-one	0.107	0.013	0.002	0.064	ND	0.095	0.092	0.092	0.052	0.140	0.027	0.004
Benz[a]anthracene-7,12-dione	0.029	0.009	ND	0.038	ND	0.041	0.050	0.059	0.028	0.098	0.007	0.012
Total	0.535	0.512	0.064	0.508	0.052	1.220	1.340	0.819	0.897	1.350	0.533	0.360
	Other Compounds											
Levoglucosan	300.8	132.3	89.7	144.5	20.6	109.1	123.3	207.0	104.2	105.6	81.9	55.6
Squalene	1.406	3.602	0.850	1.419	2.323	7.692	5.290	0.323	3.644	1.266	1.627	4.783
Cholesterol	0.159	0.223	ND	ND	ND	ND		ND	0.248	ND	ND	ND
Total	302.4	136.1	90.5	145.9	22.9	116.8	128.6	207.4	108.1	106.9	83.5	60.4
	Carbon											
Organic carbon	5137	3822	2074	7136	3215	9981	10605	15533	8675	7390	6028	4458
Elemental carbon	510	426	150	786	455	1700	1507	1478	1179	1719	815	512
Total	5647	4248	2224	7922	3671	11681	12111	17010	9854	9109	6843	4970
	Mass											
PM_{10} mass	21,750	19,290	15,190	23,680	21,750	37,550	45,020	64,090	44,320	38,870	34,730	24,350

* San Dimas site values are the average of the second and third sampling periods.

Table 4.1-9. Average source contributions (µg/m^3) to PM_{10} mass by CHS community in 1995.

Source	Atascadero	Santa Maria	Lompoc	Lancaster	Lake Arrowhead	San Dimas	Upland	Mira Loma	Riverside	Long Beach	Lake Elsinore	Alpine
	Period 1 (12/29/94 to 5/3/95)											
Wood Smoke	1.99	0.51		0.80	0.13	na	0.99	0.21	0.51	0.68	0.39	0.34
Diesel Exhaust	1.00	0.66	0.26	1.54	0.90	na	2.25	2.16	1.68	2.91	1.36	0.75
Gasoline Vehicle	1.19	0.61	0.70	0.73	0.66	na	0.68	0.80	0.43	1.11	0.57	1.05
Vegetative Detritus		0.15	0.08	0.13	0.19	na	0.22	0.39	0.25	0.27	0.13	0.12
Natural Gas Comb.				0.02	0.05	na				0.01		
Tire Wear Debris	0.31	0.82	0.37	0.35		na	0.71	0.86	0.73	1.75	0.63	
Sum of Primary Sources	4.49	2.76	1.41	3.57	1.94	na	4.85	4.42	3.60	6.73	3.09	2.25
	Period 2 (5/3/95 to 11/1/95)											
Wood Smoke	0.49	0.24	0.07	0.11	0.10	0.40	0.40	0.29	0.30	0.12	0.45	0.24
Diesel Exhaust	0.89	0.98	0.28	1.39	1.33	3.71	3.73	3.03	2.55	3.66	1.98	1.44
Gasoline Vehicle	0.28	0.22	0.08	0.38	0.15	0.25	0.19	0.37	0.39	0.28		0.28
Vegetative Detritus	0.41	0.18	0.11	0.25	0.36	0.43	0.50	1.06	0.71	0.35	0.21	0.48
Natural Gas Comb.		0.01		0.01	0.10	0.01	0.01		0.01	0.05	0.01	
Tire Wear Debris	0.21	0.24		0.88		0.41	0.61	0.89	0.93	0.73		
Sum of Primary Sources	2.28	1.87	0.55	3.01	2.04	5.21	5.43	5.64	4.89	5.20	2.65	2.45
	Period 3 (11/1/95 to 12/28/95)											
Wood Smoke	5.92	0.26		2.17	0.22	1.64	2.38	3.72	1.81	1.19	0.39	0.28
Diesel Exhaust	1.46	1.03	0.83	2.32	0.88	4.57	3.94	5.57	3.45	4.72	2.17	1.70
Gasoline Vehicle		0.23	0.11	0.46	1.06	0.33	0.41	1.18		0.51	0.61	0.28
Vegetative Detritus	0.25	0.29	0.11	0.48	0.08	0.28	0.87	1.53	0.50	0.44	0.88	0.33
Natural Gas Comb.		0.01		0.04	0.02	0.02	0.03			0.06	0.01	
Tire Wear Debris	0.43	0.32		0.74		0.61	0.83	1.31	1.45	1.38	1.58	
Sum of Primary Sources	8.06	2.14	1.05	6.21	2.27	7.45	8.47	13.37	7.22	8.31	5.64	2.59

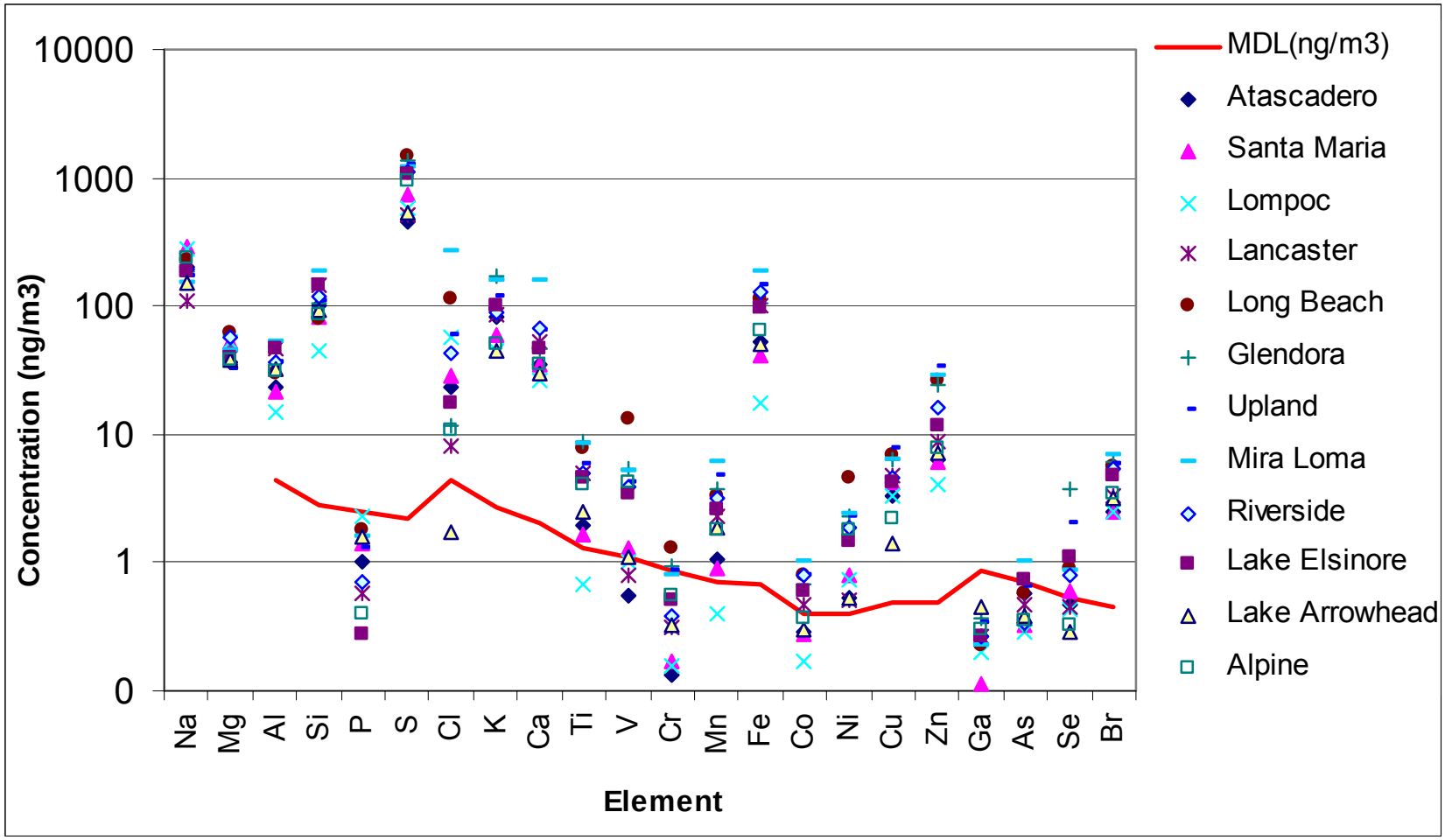


Figure 4.1-20. Annual average concentrations of selected elements in $PM_{2.2}$ in the CHS (Part A for Na – Br).

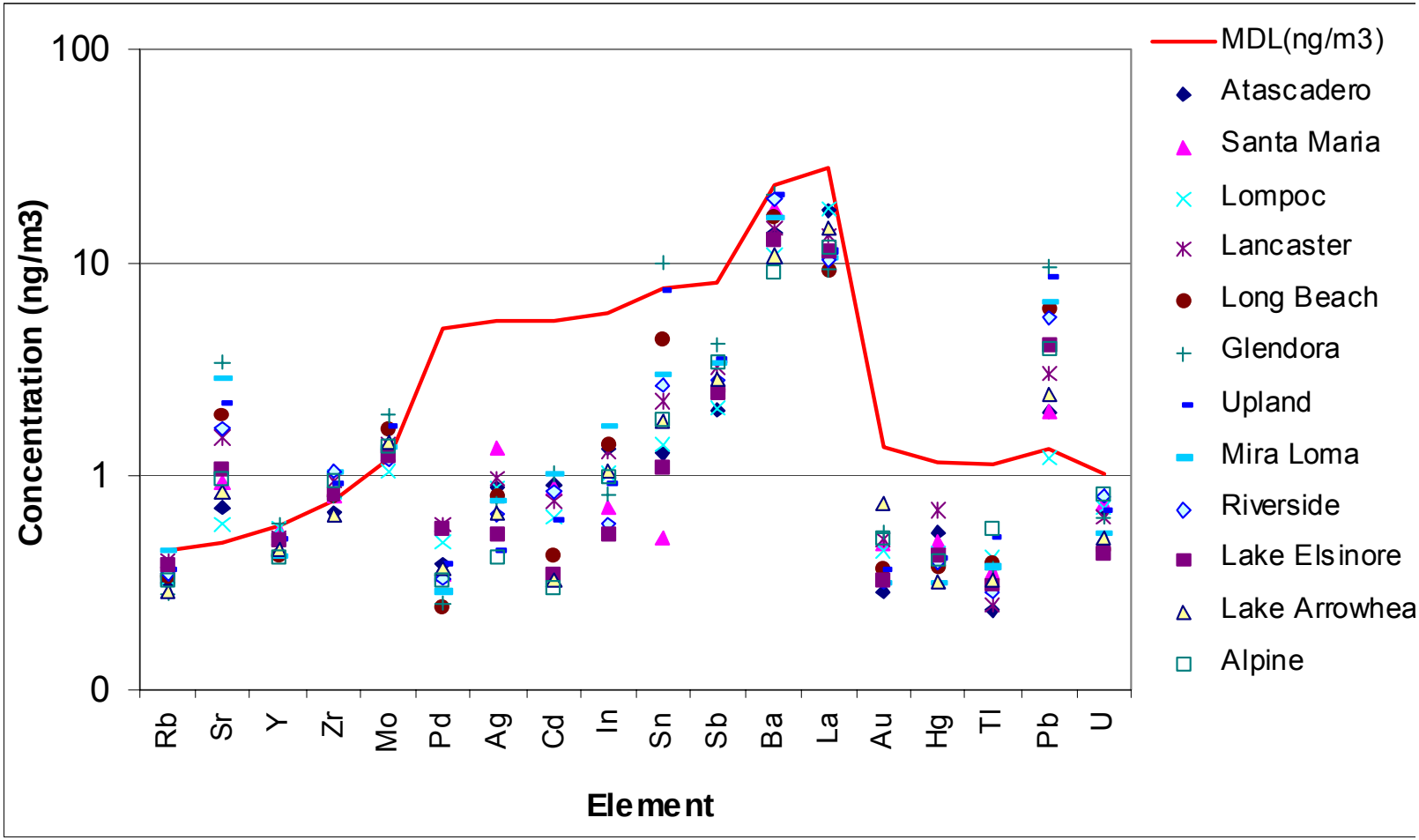


Figure 4.1-21. Annual average concentrations of selected elements in $PM_{2.2}$ in the CHS (Part B for Rb-U).

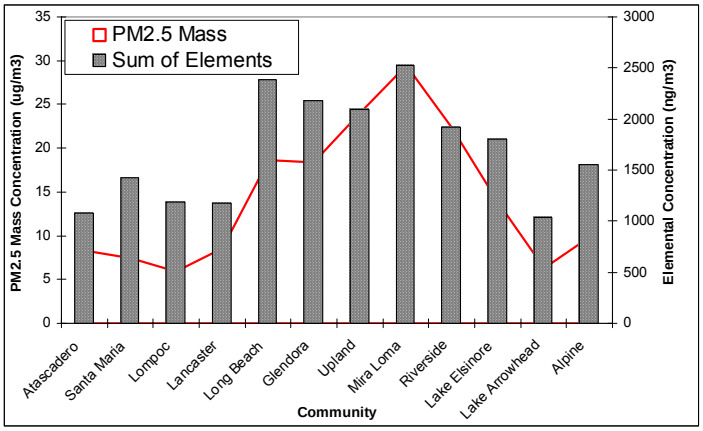


Figure 4.1-22. A comparison of the sum of the elements measured by XRF and $PM_{2.5}$ mass by CHS community in 2001.

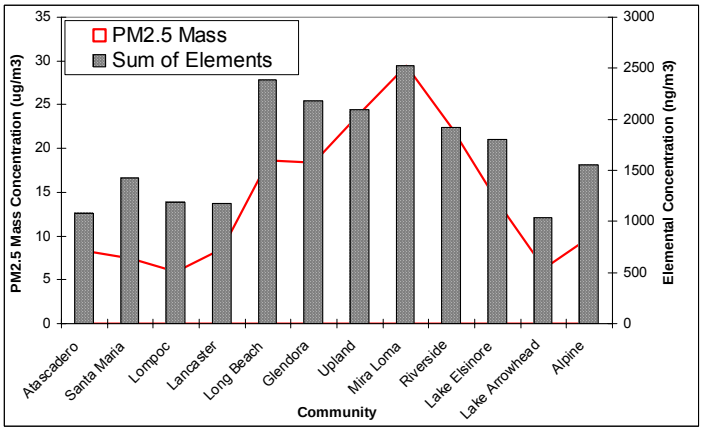


Figure 4.1-23. A comparison of the sum of $PM_{2.2}$ crustal elements and $PM_{2.5}$ mass by CHS community in 2001.

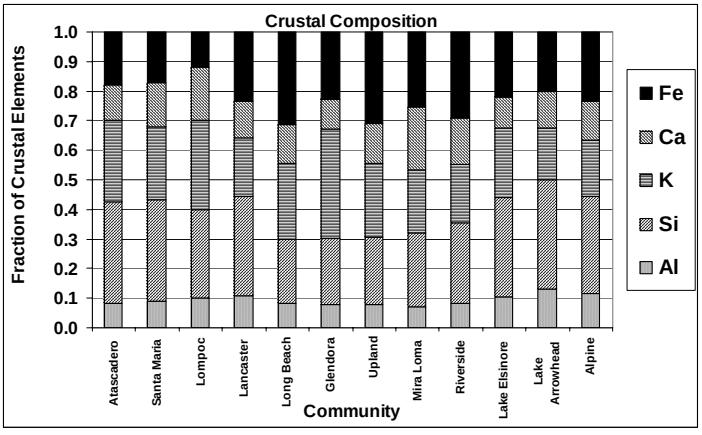


Figure 4.1-24. The 2001 relative composition of the crustal element mass by community.

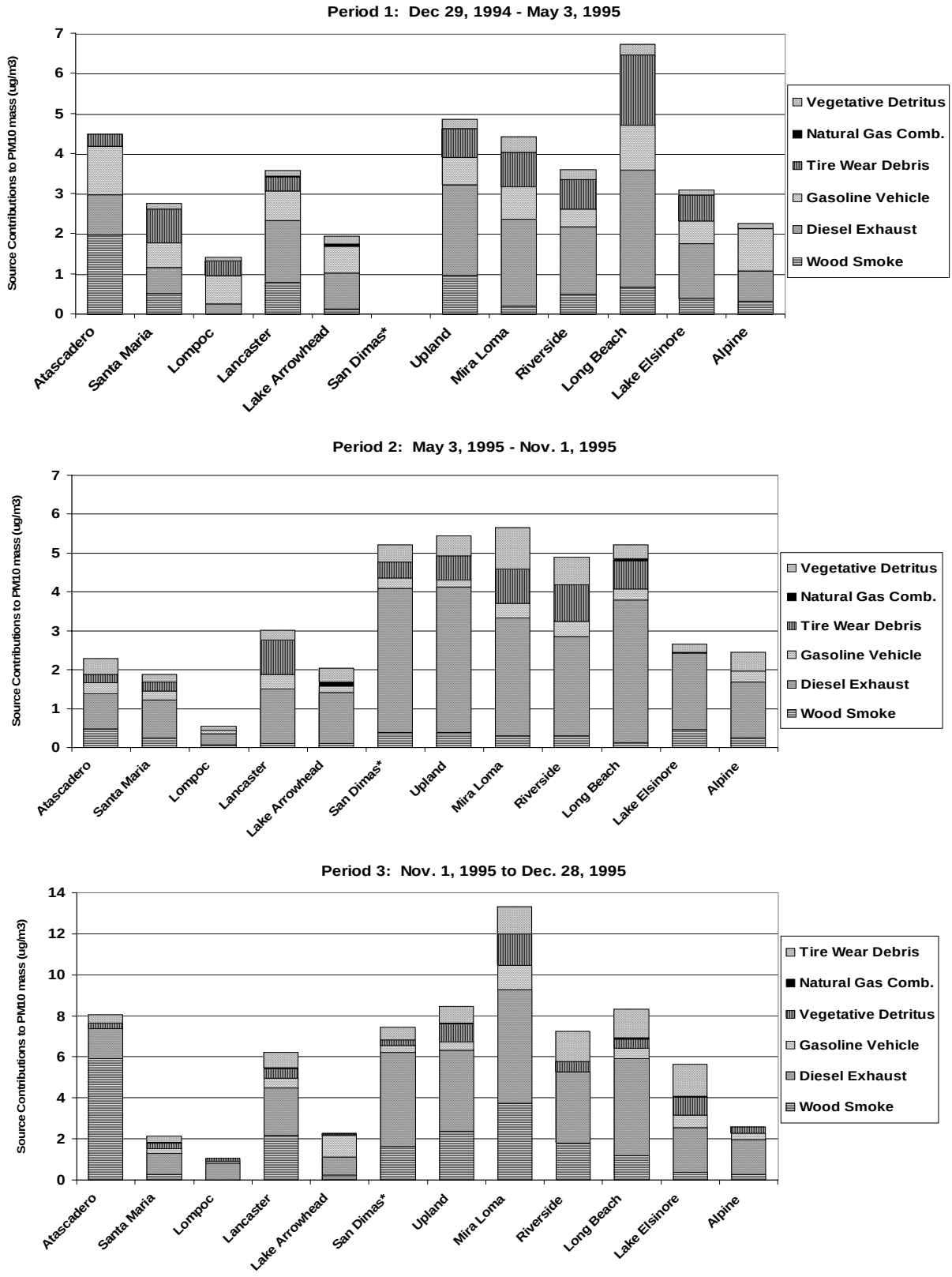


Figure 4.1-25. Average contributions (µg/m^3) to PM_{10} mass by selected source types in three time periods in 1995.

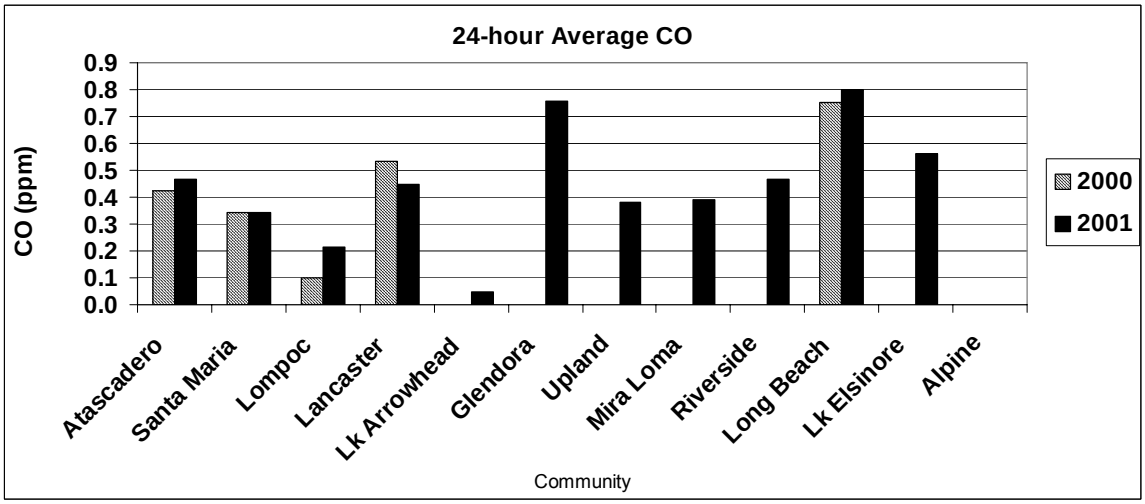


Figure 4.1-26. **Average CO concentrations in CHS communities in 2000 and 2001.**

Note, missing data for Glendora, Upland, Mira Loma, Riverside and Lake Elsinore are a result of data quality issues. No CO measurements were collected in Alpine.

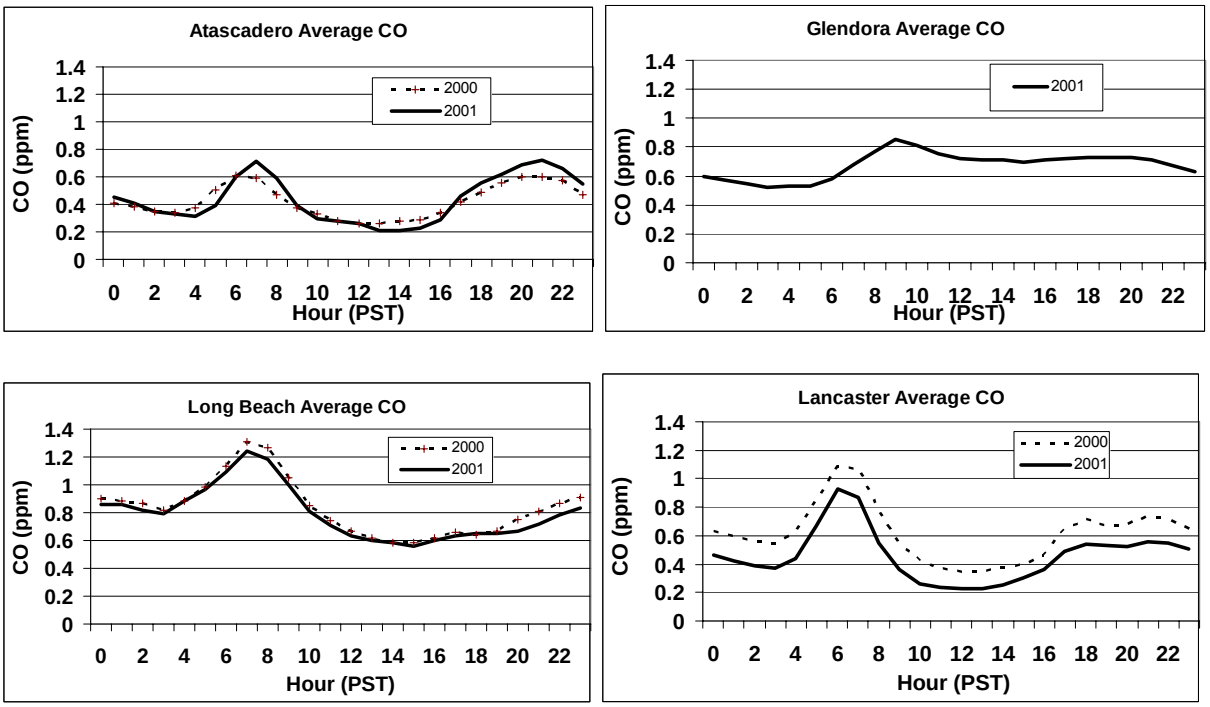


Figure 4.1-27. **Average diurnal profiles of hourly CO concentrations in Atascadero, Glendora, Long Beach, and Lancaster in 2000 and 2001.**

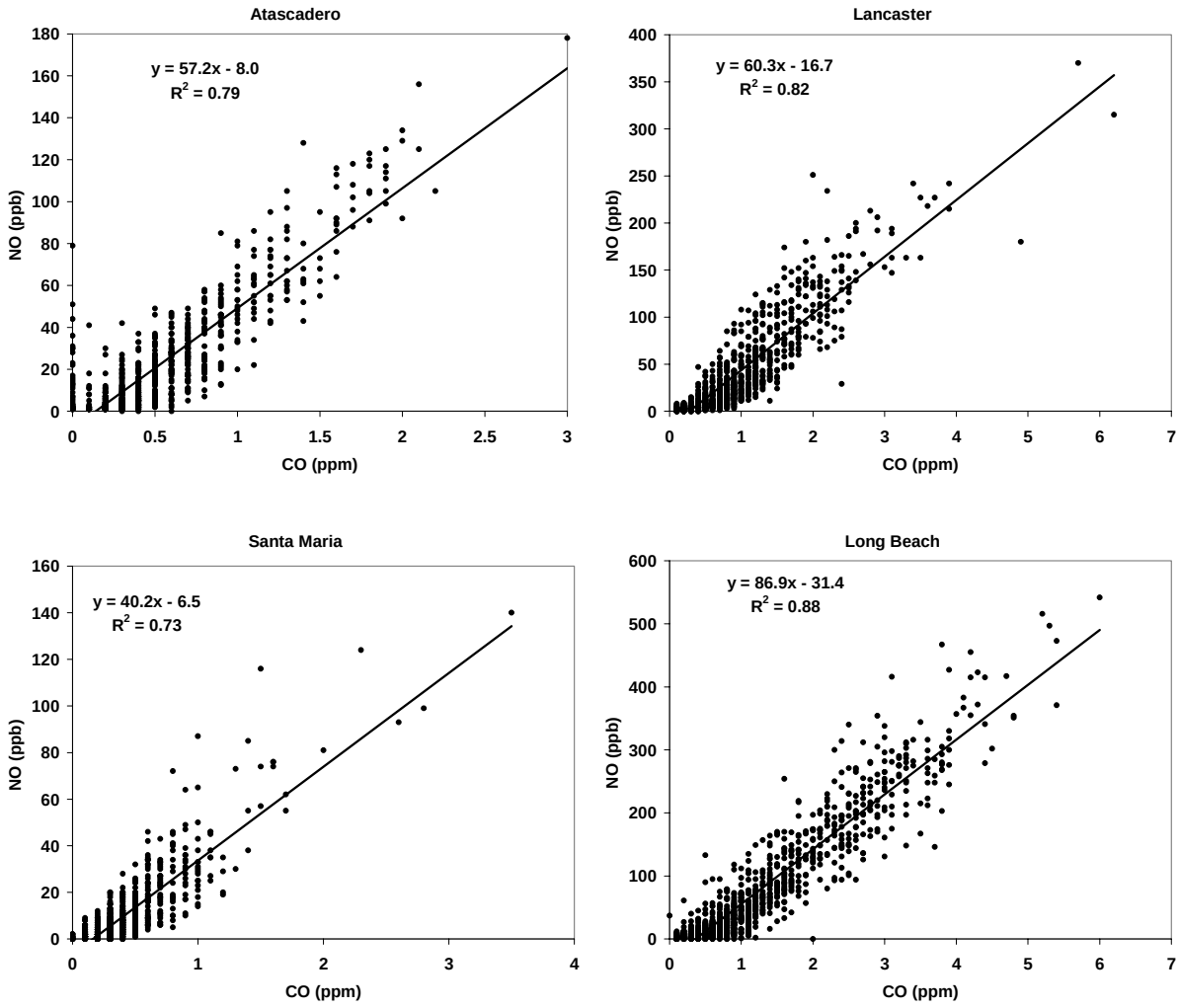


Figure 4.1-28. Comparison of 5–8 AM (PST) CO and nitric oxide hourly average concentrations at Atascadero, Santa Maria, Lancaster, and Long Beach in 2000-2001.

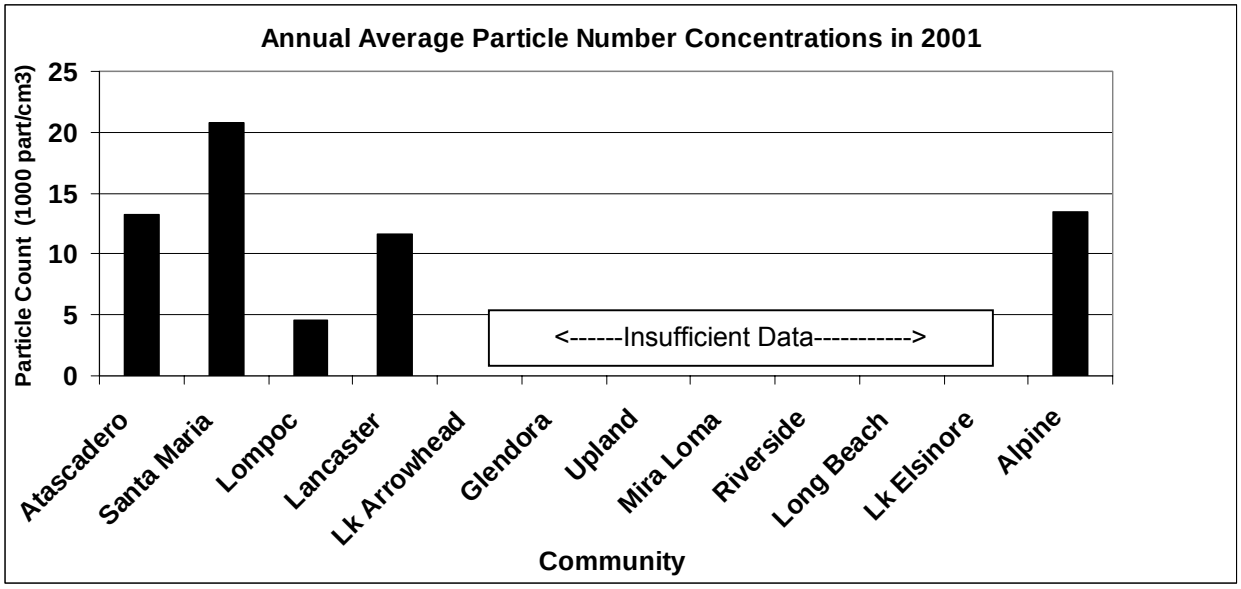


Figure 4.1-29. Annual average particle number concentrations in five CHS communities in 2001.

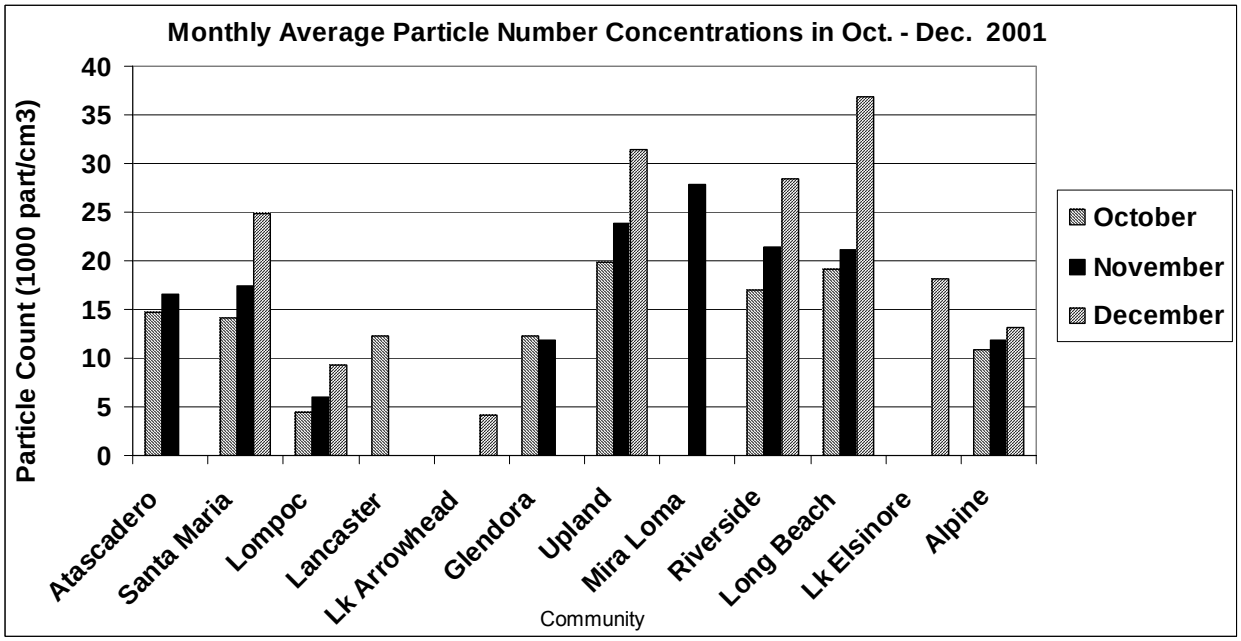


Figure 4.1-30. Monthly average particle number concentrations in 11 CHS communities in October–December 2001.

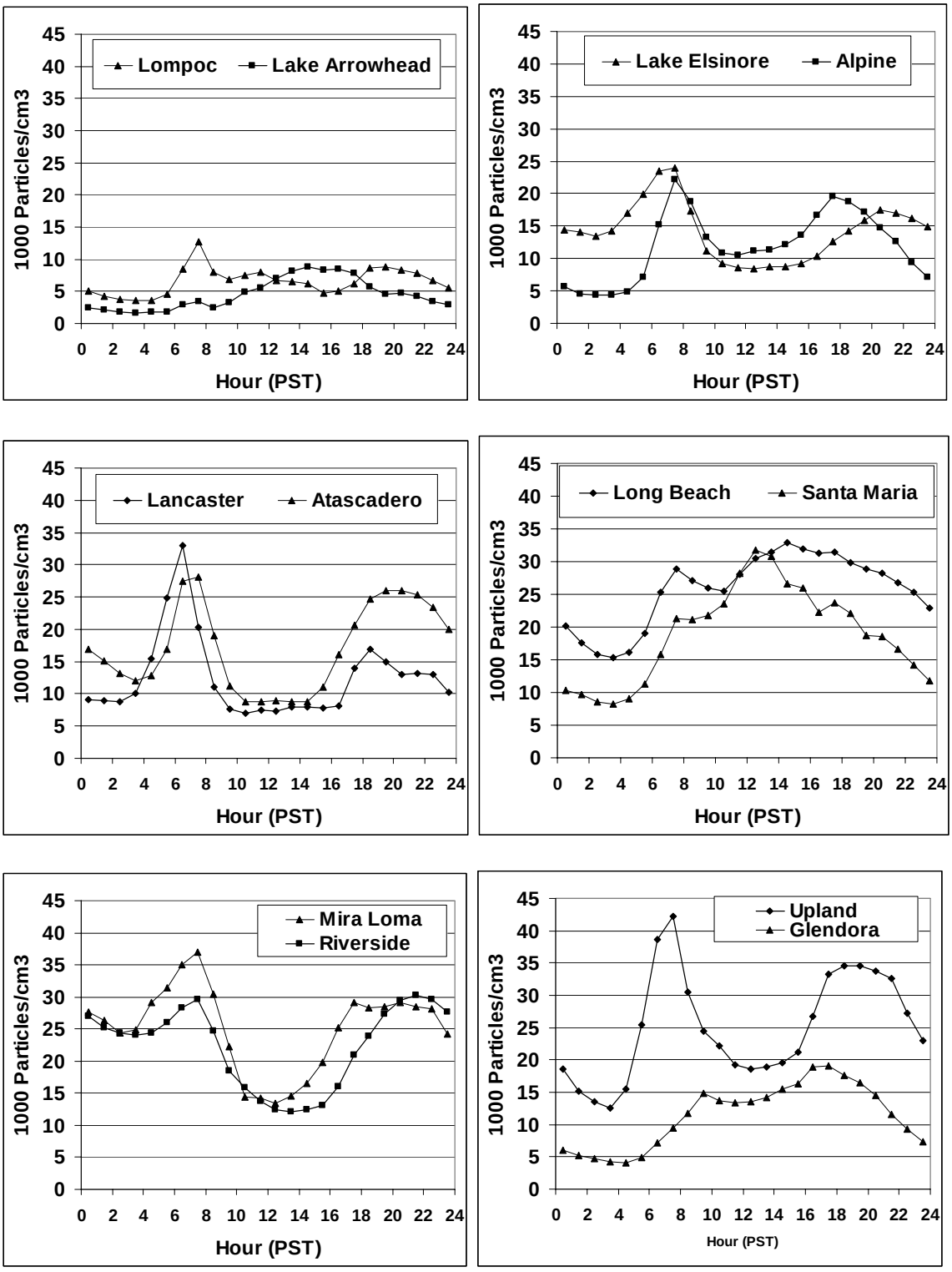


Figure 4.1-31. Average diurnal profiles of hourly particle number concentrations in October–December 2001 in CHS communities.

The profiles for Lancaster, Lake Arrowhead, and Mira Loma are based on sparse data.

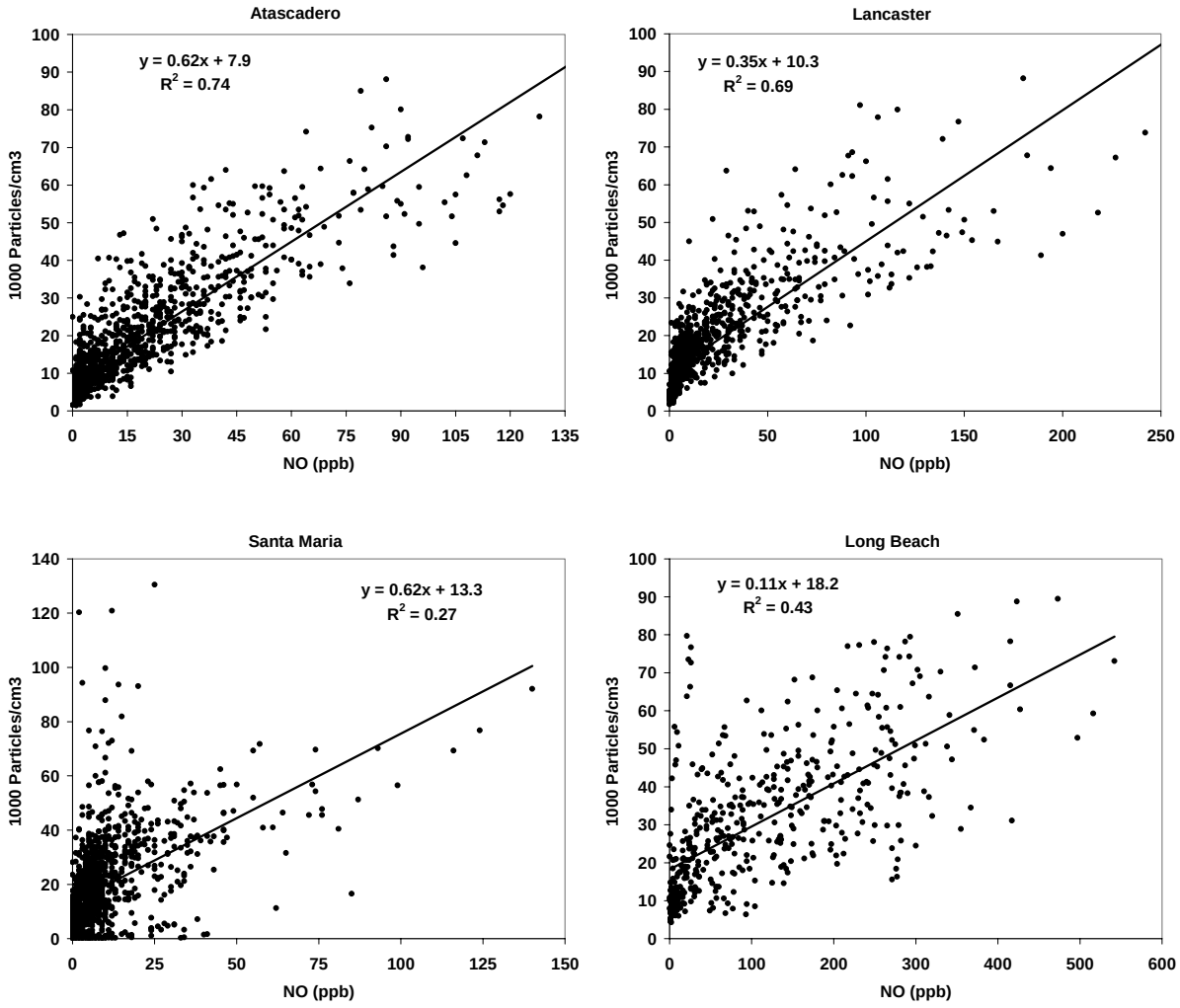


Figure 4.1-32. Comparison of 5–8 AM (PST) particle number and nitric oxide hourly average concentrations at Atascadero, Santa Maria, Lancaster, and Long Beach in 2001.

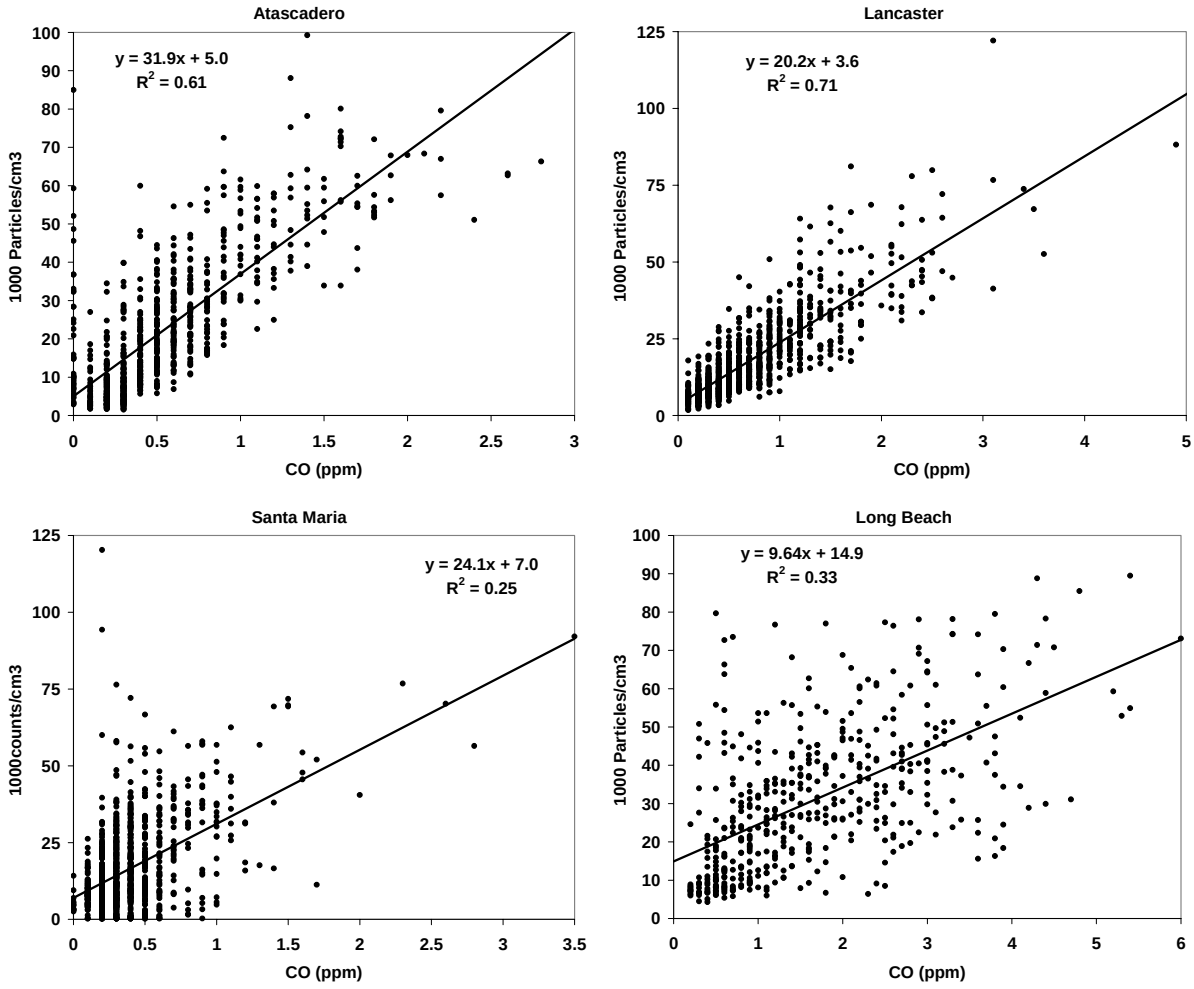


Figure 4.1-33. Comparison of 5–8 AM (PST) particle number and CO hourly average concentrations at Atascadero, Santa Maria, Lancaster, and Long Beach in 2001.

Table 4.1-10. **Relative ranking of primary pollutant concentrations in CHS communities.**

Community	NO	CO	Particle Number[a]
Lompoc	Low	Low	Low
Lake Arrowhead	Low	Low	Low
Alpine	Low	NA	Moderate
Atascadero	Moderate	Moderate	Moderate
Santa Maria	Moderate	Moderate	Moderate
Lancaster	Moderate	Moderate	Moderate
Lake Elsinore	Moderate	Moderate	Moderate
San Dimas/ Glendora	High/ Moderate	NA/ High	NA/ Moderate
Upland	High	Moderate	High
Mira Loma	High	Moderate	NA
Riverside	High	Moderate	High
Long Beach	High	High	High

[a] Rankings for CO and especially particle number are probably less accurate than those for NO because of the short data record.

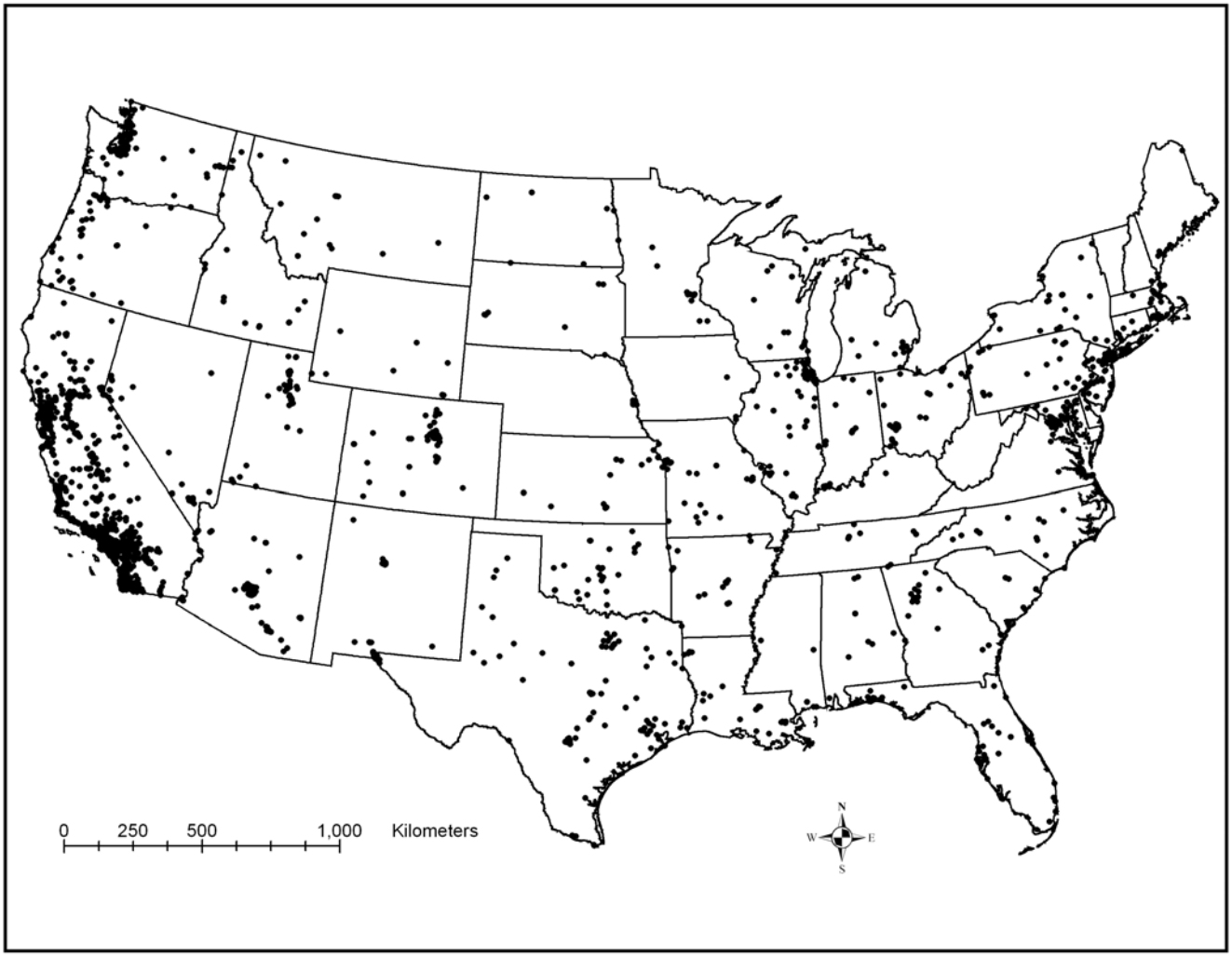


Figure 4.1-34. Locations within the contiguous United States where CHS residents lived for one or more months prior to enrollment in the study.

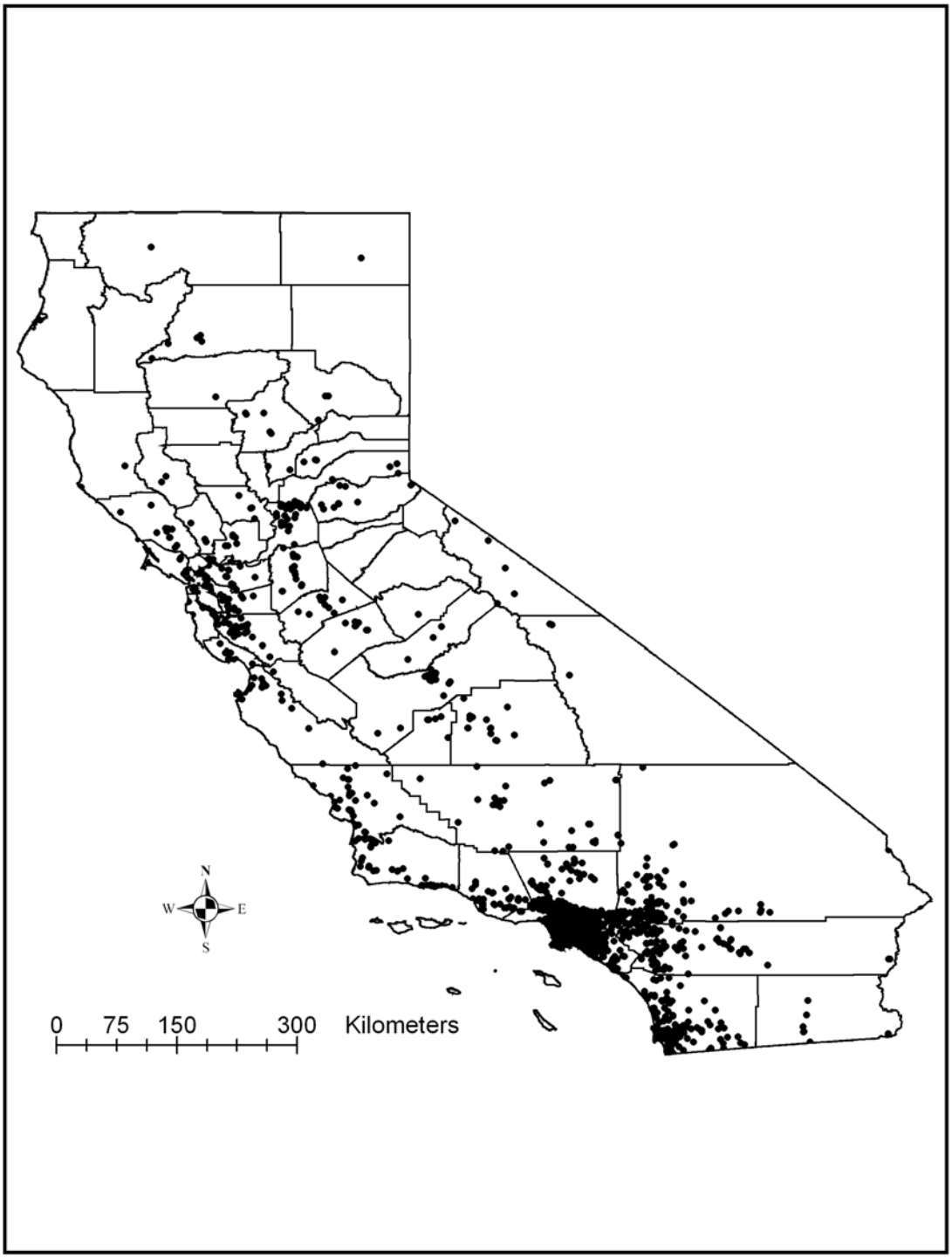


Figure 4.1-35. Locations within California where CHS residents lived for one or more months prior to enrollment in the study.