

3.0 REQUIREMENTS FOR AIR QUALITY MODELING SYSTEMS AND DATA ANALYSIS

Data requirements for CCOS are determined by the need to drive and evaluate the performance of modeling systems, which include the following components.

- 1) Meteorological model, which provides winds fields, vertical profiles of temperature and humidity, and other physical parameters in a gridded structure.
- 2) Emissions inventory and supporting models that provide gridded emissions for area sources such as on- and off- road motor vehicles and vegetation.
- 3) Air quality model, which simulates the chemical and physical processes involved in the formation and accumulation of ozone.

In evaluating modeling system performance, the primary concern is replicating the physical and chemical processes associated with actual ozone episodes. This necessitates the collection of suitable meteorological, emissions, and air quality data that pertain to these episodes. The data requirements of CCOS are also driven by a need for complementary, independent and corroborative data analysis so that modeling results can be compared to current conceptual understanding of the phenomena replicated by the model.

This section describes the data requirements for meteorological and air quality models and for developing temporally and spatially resolved emissions inventories of reactive organic compounds and nitrogen oxides. The input data required for each modeling component (Section 3.1) are described separately from the data required to evaluate model performance and validity of modeling results (Section 3.2). Section 3.3 describes the potential approaches and methods for observation-based data analysis and verification of emission inventories.

3.1 Modeling System Inputs

The output of the meteorological model and gridded emission inventory serve as inputs to the air quality model. Meteorological modeling and emission inventory development are discussed first in Sections 3.1.1 and 3.1.2, and the output of these first two components of the modeling system are discussed in the context of inputs to the air quality model in Section 3.1.3.

3.1.1 Meteorological Modeling

The specification of the meteorological fields that drive the transport and dispersion of atmospheric pollutants is the critical component in mesoscale air quality modeling. The primary objective is to obtain wind fields over the model grid and determine mechanical and convective mixing depths. The simplest way is to use field measurements and interpolate the values over the entire domain. However, field measurements are generally spatially and temporarily sparse, and can be especially inadequate in areas with complex terrain and land-sea interactions. Another way is to use diagnostic and prognostic meteorological models to estimate meteorological fields from existing data and then to adjust these fields through parameterizations of physical processes. And best results are often achieved when the methods are combined such

that interpolated or gridded field measurements are used as input to a prognostic model at regular time intervals in the model simulation in a process called Four-Dimensional Data Assimilation (4DDA). After a general description of this process and a few of the models involved, specific data requirements for setup of and input to the meteorological model are listed.

Transport and the dynamics and thermodynamics of the atmospheric boundary layer (ABL) govern dispersion of atmospheric pollutants. The main difficulties in dispersion estimates arise with topographic complexity and increasing atmospheric stability. Turbulence in the ABL is created by wind shear and destroyed by buoyancy and dissipation. Since these effects are nearly balanced in the stable ABL, turbulence intensities are usually low and intermittent. In some cases of stagnant stable conditions the horizontal diffusion of the plume can be of the same magnitude or larger than the actual transport. Moreover, turbulent velocities are frequently affected by gravity waves and the stable ABL undergoes non-stationary evolution. Additional dispersion due to wave phenomena also needs to be resolved. During stable conditions the ABL flow usually decouples from the synoptic winds and the local circulations dominate its flows. In some cases a low-level jet can develop at the top of the surface stable layer and the fate of pollutants at various elevated layers can be completely different over a very small vertical separation. In contrast, during stable and stagnant conditions the winds close to the surface are weak and sometimes below the detection limit of usual instrumentation.

Radiation and advection can also cause fog and cloud formation and significantly change the rate of chemical reactions for some species and deposition processes. The depth of the stable ABL is of the order of 100 m, and radiation processes, as well as local effects such as urban effects, vegetation, soil properties, and small-scale topographic features, can significantly influence ABL characteristics. Plume meandering is frequently observed during stable conditions in topographically complex terrain, and use of data from limited measurement sites can yield erroneous conclusions. All these effects significantly modify transport and dispersion as well as removal of pollutants. Consequently, an extensive measurement network is necessary in order to capture the spatial and temporal structure of the ABL in a mesoscale domain. The success in any type of dispersion calculation will be limited to appropriate capture and input of atmospheric parameters.

One of the diagnostic models recommended for this study is CALMET. Wind fields in CALMET are calculated in a user specified number of vertical levels by taking into account the influence of terrain on the atmospheric flow and applying an inverse weighting scheme. The initial terrain-adjusted domain mean horizontal components of the wind at each grid point are modified to obtain the final interpolated wind components. CALMET can calculate a spatially variable initial guess field using objective analysis of the measurements. Moreover, CALMET allows use of gridded wind fields created by a prognostic atmospheric model, such as the Penn State University Meteorological Model (MM4, MM5), as “initial guess” fields or as substitutes for observations. CALMET has detailed algorithms for the depth of the convective layer as a function of the potential temperature lapse rate in the layer above the mixing depth, the time step, and the temperature discontinuity at the top of the mixed layer. The daytime mechanical mixing depth is determined from the Coriolis parameter, the friction velocity and the Brunt-Vaisala frequency in the stable layer above the mixed layer. The nighttime depth of mechanical mixing is determined from the friction velocity. CALMET uses an upwind-positioned averaging scheme to smooth out the mixing depths through use of determined weighting factors. Since CALMET

has detailed algorithms for wind fields and mixing depth, and furthermore allows initialization with the output from the prognostic atmospheric model, it is an optimum tool for obtaining the initial meteorological fields necessary for estimation of transport, dispersion, and chemical transformation of atmospheric pollutants.

Based on the complexity of terrain in northern and central California, the MM5 model developed by Penn State University and the National Center for Atmospheric Research (NCAR) represents an appropriate tool for resolving dynamics and thermodynamics if used on the scale of 1-2 km horizontal resolution with nesting capabilities. It should include a nonhydrostatic option and full parameterization of physical processes including turbulent transfer. MM5 uses an advanced four-dimensional data assimilation scheme connected to either measurements or synoptic fields. However, uncertainty exists as to the extent to which MM5 can be used to infer turbulence properties in complex terrain.

Some advantages in using a prognostic model such as MM5 are:

- High resolution in horizontal and vertical directions.
- Topography with resolution of 30 seconds embedded within the model structure.
- Detailed prognostic fields of meteorological parameters (wind, temperature, humidity, turbulence, radiation, clouds).
- Physically-based estimate of mixing depth in both convective and stable cases, with full spatial and temporal variability.
- Detailed structure of the small-scale local flows that cannot be resolved through simplified parameterization.
- Detailed vertical structure of meteorological parameters and stability, which is especially important near sources and receptors and along the transport path.

Some disadvantages in using a prognostic modeling approach are:

- The models are fairly complex and expensive to run. Usually they are limited to certain case studies.
- The models have assumptions in simplification of basic differential equations and numerical techniques and in the parameterizations of physical processes.

Meteorological modeling for the CCOS must focus on the characterization of the origin and fate of atmospheric pollutants in central California. The main difficulties are due to complex terrain and a number of significant sources within the large region and its surroundings. Integration with dispersion models may also be a problem. It is also desirable that atmospheric modeling adequately treat formation and evolution of fog and clouds, which are essential determinants for liquid-phase chemistry, that may be important in coastal regions or during precipitation events that may clean the air after ozone episodes (e.g., monsoonal flow in the San Joaquin Valley). However, these processes are often not well-represented over a large scale

domain like that planned for CCOS. Despite these difficulties, atmospheric models are useful tools in understanding the structure and evolution of boundary layer dynamics and providing meteorological fields as input for dispersion models. The results from wind field modeling will be used as input for dispersion and chemical modeling of relevant pollutants also for the entire intensive study period.

Necessary decisions and data requirements for the setup of the meteorological model include:

- Domain size (coordinates in latitude and longitude or UTM).
- Topographic information for the orography of the domain.
- Spatial resolution, i.e., grid size, and any nesting of grids of different sizes.
- Land use per surface grid square with information on soil moisture, albedo, roughness height, and vegetation.

Typical land use categories include high rise urban, residential, deciduous forestland and evergreen forest categories. Surface resistances and deposition velocities are a function of land use and vegetation data. Measurements of deposition for important species over representative categories of land use and vegetation type are planned for a subsequent follow-on study after summer 2000.

Necessary and/or highly desired field measurements for input to the meteorological model include:

- Routine surface network measurements of wind speed and direction, temperature, humidity, pressure, solar radiation, and precipitation.
- Routine aloft network measurements of wind speed and direction by radiosonde and existing radar wind profilers (RWP), and sodars; of temperature by RWP with Radio Acoustic Sounding Systems (RASS), of humidity and pressure by radiosonde, and of precipitation by WSR-88D (NEXRAD), which can also provide some wind information as a radial component of the Doppler radar beam).
- Supplemental measurements as part of the CCOS and CRPAQS study design. These may include: additional surface and upper sites making routine measurements listed above, turbulence by sonic anemometer (CRPAQS 100-m tower at Angiola), actinic flux for photolysis of key species, and soil moisture for key land use types.

3.1.2 Emission Inventory Development

For a study covering as large of an area as CCOS, the development of the emission inventory will be a major effort. In addition to the size of the study domain, the emission inventory required for modeling is more detailed than the routine summer seasonal emission inventory. The emission inventory (EI) needed to support the CCOS modeling will be a series of day-specific, hourly, gridded emission inventories that cover each day of the ozone episodes

captured during the field study. To help coordinate the EI development effort, the Emission Inventory Coordination Group (EICG) has been established to determine the process for preparing the emission inventories needed to support air quality modeling for CCOS. Participants in the group include many local air districts, several local councils of government, Caltrans, California Energy Commission, and the ARB. Local air districts participating to date include San Joaquin Valley Unified APCD, Bay Area AQMD, Sacramento Metropolitan AQMD, Mendocino County AQMD, Northern Sierra AQMD, Yolo-Solano AQMD, Placer County APCD, San Luis Obispo County APCD, and Monterey Bay Unified APCD. Other local air districts will also be participating.

There are about 30 districts in the CCOS modeling domain. Each local air district in the state updates a portion of the emission inventory for their area. An emission inventory is needed to estimate emissions for each day that will be modeled in the summer of 2000. Table 3.1-1 provides a preliminary outline of the proposed roles and responsibilities of the various agencies and groups involved in the development of the emission inventory. Special effort is being made to assist smaller districts that have limited staffing resources. The following subsections discuss point, area, on-road, and biogenic sources; and emissions forecasts, a two-track approach for the modeling schedule, quality controls, and proposed studies for the EI development effort

3.1.2.1 Point Sources

The local air districts have primary responsibility for controlling stationary sources of air pollution within their jurisdictions. In fulfilling this responsibility, the districts issue permits, inspect and evaluate point sources, and estimate emissions for these sources. Point source emissions occur at facilities that can be identified by name and location. Generally, facilities with emissions greater than 10 tons per year of any pollutant are reported as point sources. Districts may include smaller facilities in their point source inventory if they choose. The emissions from smaller facilities that fall below the point source reporting threshold are estimated as part of the area source emission inventory.

Districts gather information from their point sources and generally update the inventory annually. Point sources are estimated using a “bottom up” approach. Point source activity levels usually relate to a facility-specific process rate. Emission factors are derived from tests that relate emissions to the process causing the emissions. The districts then pass the emissions and other related information to ARB. The statewide point source emission estimates are stored and maintained in the California Emission Inventory Development and Reporting System (CEIDARS).

Modeling inventories require more data than are needed to calculate seasonal or annual average emissions. For point sources, the location in UTM coordinates is used to place the emissions within a grid cell. Temporal data are used to adjust the annual emissions for a particular month, day and hour. Stack data such as height, diameter, exit temperature and velocity are used to estimate the plume rise from the stack to place the emissions in the correct vertical cell of the model.

The local air districts and, if needed, private contractors will gather day-specific point source data. The local districts will gather information from large point sources, particularly

those that have significant seasonal variations. The CCOS EICG will agree upon a standard emission reporting threshold for collecting day-specific, and when available hourly specific, point source emission information for all districts in the CCOS domain. Air districts will have the option to report day-specific data below the standard threshold.

For the remaining point sources for which day-specific data are not collected, the ARB will use facility-specific information contained in CEIDARS. The emissions are adjusted to represent a specific month and a specific day of the week (a Tuesday in August, for example). Daily emissions are distributed over each hour. These adjustments are based on temporal profiles that are stored in CEIDARS. Each district assigns temporal profiles using survey data or knowledge of the source. The local districts and the ARB will work through the EICG to develop comprehensive point source updates for 1999 and 2000 emission inventories.

3.1.2.2 Area Sources

There are currently approximately 360 emission source categories for which area source methods are used to estimate emissions. These categories include both stationary and off-road mobile sources. The source categories are divided into four types of emission sources. *Aggregated point sources* are many small point sources, or facilities, that are not inventoried individually but are estimated as a group and reported as a single source category. Examples include gas stations and dry cleaners. *Area-wide sources* include source categories associated with human activity and emissions take place over a wide geographic area. The emissions data are developed only at the aggregated level. Consumer products and agricultural operations are examples of area-wide sources. *Natural (non-anthropogenic) sources* generally include source categories with naturally occurring emissions such as geogenic sources and wildfires. *Off-road mobile sources* include categories such as farm equipment, off-road recreational vehicles, and aircraft. Collectively, these types of sources are referred to as area sources.

Emissions from area sources are generally estimated using a “top down” approach. This approach generally uses a process rate or amount of a substance being used multiplied by an emission factor. An example of a process rate for dry cleaners would be the amount of perchloroethylene used in a county. Emissions are estimated for each county. ARB and the districts generally update the emissions from area sources once every three years. Updates are based on the latest data available including surveys and the results of research projects. The statewide area source emission estimates are stored and maintained in CEIDARS.

The ARB staff is responsible for developing methods and estimating emissions from over 260 categories. The local air districts are responsible for developing methods and estimating emissions for the remaining categories. The local districts also have the prerogative to use their own methods for area source categories that are normally the responsibility of ARB. ARB’s Mobile Source Division is developing an off-road mobile source emission inventory model, OFFROAD, which will be used to calculate the emissions from off-road mobile sources. The ARB and the local districts will work together through the CCOS EICG to revise and/or update all major area source categories for use in preparing the 1999 and 2000 emission inventories.

To grid area sources, emissions for a particular source category are spatially distributed within a county through the use of surrogates. In this context, a surrogate is an economic or

demographic parameter, which is correlated to emissions for a particular source category and has a known spatial distribution. Some common surrogates used for this purpose are population, housing units, and employment. As with point sources, temporal data are used to adjust the annual emissions for a particular month, day and hour. Annual emissions from all sources are checked for reasonableness and completeness. The gridding surrogates are planned for update as part of the CCOS emission inventory development process.

Day-specific data can also be collected for area source categories. For this study, data will be collected for wildfires, agricultural burns and prescribed burns. ARB will gather information needed to estimate wildfire emissions including the date, time, and location of the burn along with the amount and type of material burned. Either the local districts or a contractor will collect similar information for agricultural and prescribed burns.

3.1.2.3 On-Road Motor Vehicle Sources

The ARB has primary responsibility for developing emission estimates for on-road mobile sources. On-road mobile sources are motor vehicles that travel on public roads. This category consists of gasoline-powered and diesel-powered passenger cars, light-duty trucks (6,000 lbs. gross vehicle weight [GVW] or less), medium-duty trucks (6,001-8,500 lbs. GVW), heavy-duty trucks (over 8,500 lbs. GVW), urban buses, and motorcycles. Emissions from motor vehicles include exhaust, evaporative, crankcase, and tire-wear emissions. The Mobile Source Division of the ARB develops emission factors for California. The Emission Inventory Branch of the Planning and Technical Support Division of the ARB has the responsibility for developing the motor vehicle activity components such as vehicle miles traveled (VMT), starts, and vehicle population, which are used by the models.

The emission factor model to be used for the on-road motor vehicle inventory is EMFAC99. This model, which is expected to be approved by the ARB in October 1999, is an integrated, modular system of programs. These modules perform the functions of the computer programs CALIMFAC, WEIGHT, EMFAC and BURDEN that were used in previous versions of the model. Although the EMFAC99 modules have been integrated into a single program, the functions of EMFAC99 will be discussed in terms of the individual modules to make it easier to understand the model.

The “CALifornia Inspection/Maintenance emission FACtor” (CALIMFAC) module computes base emission rates for each technology group, with and without inspection and maintenance (I/M) benefits. The base emission rates consist of a zero mile rate and a deterioration rate for each model year for each pollutant. The “EMission FACtor” (EMFAC) module computes fleet composite emission factors by vehicle class, speed, ambient temperature, emissions process and technology for a calendar year. The WEIGHT module provides EMFAC with activity weighting fractions for individual model years so that the fleet composite emission factors can be produced. The WEIGHT module also provides the accumulated mileage by model year for any particular calendar year in order to calculate the “deteriorated” emission rate for a model year. The BURDEN module calculates the emission estimates (in tons per day) by multiplying the composite emission factors from EMFAC by activity estimates. Vehicle miles traveled, vehicle type distribution, the number of registered vehicles, and speed distributions, which are the input data for the BURDEN module, are obtained from the California Department

of Motor Vehicles (DMV) and the California Department of Transportation (Caltrans). Transportation data also are provided by the local councils of government (COGs). The BURDEN module selects the proper emission rate according to the speed, ambient temperature, relative humidity, and heavy-duty I/M benefits of a county or sub-region of a county.

Motor vehicle emission rates consist of exhaust emissions of TOG, CO and NO_x from running and continuous starts and TOG emissions from hot soak, diurnal, running and resting evaporative processes. Running exhaust emissions include emissions from the tailpipe and through the crankcase after the vehicle has warmed up and has reached a stabilized mode of operation. Running exhaust emissions also include exhaust particulate matter and particulate matter from tire wear. Exhaust emissions from continuous starts result from operation of the vehicle before the engine has reached the stabilized mode of operation. Continuous starts replace the concept of hot and cold starts used in older versions of the model. The continuous start rate is a function of the engine time off. This is estimated in increments of 1, 5, 10, 20, 40, 60, 90...720 minutes, with the 1 minute rate being very close to a hot stabilized value, and the rate for 720 minutes being essentially a cold start. Hot soak emissions result from gasoline vaporization from elevated engine and exhaust temperatures after the engine is turned off at the end of a trip. Emissions for the running exhaust and running evaporative processes are assumed to be proportional to vehicle miles traveled. The emission rates for hot soaks and continuous starts are expressed in terms of grams per start rather than as grams per vehicle mile traveled. Diurnal evaporative emissions are expressed in terms of grams per vehicle and result from the daily changes in the ambient temperature due to expansion of the air-fuel mixture in a partially filled fuel tank. Running evaporative losses are releases of gasoline vapor from the fuel system during vehicle operation. Resting loss evaporative emissions are due to fuel line hose or fuel tank permeation.

There are two ways of characterizing inventories available from the ARB: daily average and planning. The average daily emissions are expressed as an emission rate in tons per average day, determined by dividing annual emissions by 365. Countywide and basinwide totals are provided by source categories. This inventory is updated and published annually by the ARB. EMFAC99 can produce planning inventories that take into account the effects of diurnal and seasonal variations in temperature and activity patterns. The planning inventories from EMFAC99 will provide emission estimates for various periods specified by the user. Hourly resolution will be possible during an average summer day, winter day, season, month or annual average.

The Direct Travel Impact Model (DTIM) is the analog to EMFAC99's BURDEN module for gridded inventories. DTIM was developed jointly by Caltrans and the ARB. DTIM is designed to use the output of travel demand models as its activity input for vehicle miles traveled, trips (or starts) and speed distributions. These activity estimates from the travel demand model are at the individual link or zone level and therefore provide spatial allocations for motor vehicle emissions. DTIM applies the same fleet composite emission factors to the motor vehicle activity, as does the BURDEN module. The user defines an input file of parameters such as hourly temperatures by grid cell to calculate and distribute emissions into the grid system. Data are provided by the COGs, particularly the numbers of trips by zone, volumes of vehicles by link, and link-specific speed estimates. Factors to allocate trip and VMT estimates from the model to individual hours are also an input to DTIM.

3.1.2.4 Biogenic Sources

Many studies suggest that VOC emissions from plants may play a role in the formation of ozone and fine particulate matter. Unfortunately, biogenic hydrocarbons remain a potentially large and yet not well understood source of emissions. The ARB has developed a GIS-based biogenic hydrocarbon emissions inventory model (BEIGIS) for use in California. BEIGIS has been applied in southern California and can be applied in central California if region-specific databases are available. BEIGIS has four primary data inputs, which are discussed below. Once these inputs are assembled, ARB will generate a speciated, gridded emission inventory for biogenic hydrocarbons for the CCOS domain.

The primary data inputs for BEIGIS are: 1) maps showing the distribution of vegetation species, 2) vegetation biomass distribution maps, 3) leaf mass constants, and 4) plant-specific emission rates. For the first input, all currently available GIS coverages of the distribution of plant species for agricultural, natural and urban areas will be compiled and meshed into a seamless coverage for the CCOS domain. A list of the plant species ranked by acreage will be developed. This ranking will assist in identifying those key species to be targeted for leaf mass constants and emission rate measurements and thus focusing on the region-specific data needed for the CCOS domain. For the second input, biomass can be estimated using the leaf area index (LAI) in forested areas where species distributions are relatively homogeneous. LAI can be thought of as the number of leaf layers directly above a given ground surface area. ARB currently has a 1 km resolution LAI database for California based on satellite imagery. As part of this study, it is hoped that a 30 meter resolution database can be developed using satellite imagery. The third input listed is leaf mass constants. Biomass cannot be accurately estimated using indirect methods such as LAI in areas that contain heterogeneous vegetation species distributions, such as urban areas. In these areas, biomass is most accurately determined directly using ground-based botanical surveys. In this methodology, biomass is estimated using the canopy volume (cubic meters) for each tree surveyed in the field, as well as leaf mass constants (grams per cubic meter) for those trees. However, leaf mass constants have only been determined for less than 5% of the 6000 tree species in California. It is hoped for this study that leaf mass constants can be developed for the most important vegetation species in the CCOS domain. The final input listed, plant-specific emission rates, will be compiled in a database for the most commonly occurring species in the CCOS domain. At a minimum, the database will include emission rates for isoprene, monoterpenes, and oxygenated compounds. For those species for which emission rate measurements are not available, it is hoped during this study to assign emission rates using a taxonomic method.

3.1.2.5 Emission Forecasts

Emission backcasts and forecasts are made by the ARB using base year emissions in conjunction with estimates of growth and emission control effectiveness. Emission growth is based upon available projections of socio-economic trends. Local districts supply growth data, often using information from local planning organizations. When local data are not available ARB uses default values from research projects that use surrogates based on business activity. Emission reductions are based upon adopted regulations and control measures. Backcasts and forecasts are also available for planning inventories.

3.1.2.6 “Fast Track” and Final Modeling Inventory Development Process

A gridded emission inventory is a critical input to photochemical air quality models. Historically, modeling emission inventories have been criticized for using incomplete day-specific data and for lagging behind the development of other data such as meteorological and air quality data that are collected in air quality studies. This subsection describes a two-track approach.

To make a modeling inventory available in a shorter time frame, the EICG is taking a two-tier approach to the development of the modeling emission inventory that will provide modelers with a “fast track” inventory for preliminary modeling. This inventory will be followed by a more detailed modeling inventory that will incorporate all of the day-specific information collected during the field study.

The districts and ARB will prepare a seasonal emissions inventory for 1999 during calendar year 2000. The 1999 annual emissions for point sources will be forecasted to 2000 and then spatially and temporally distributed to create a modeling inventory for a summer weekday. Area source emissions for 1999 will also be forecasted to 2000. The forecasted emissions will be spatially distributed using the spatial surrogates used in the 1994 Ozone State Implementation Plans and temporally distributed to represent a summer weekday. On-road mobile sources will also be gridded based on the 1994 Plans. Modelers will use this “fast track” inventory to begin testing the performance of the photochemical model. When the 2000 annual average inventory is ready, it will be gridded to replace the forecasted 2000 inventory. Day-specific data will replace the gridded annual average inventory. Results from contracts will be incorporated into the annual average inventory and/or the day-specific inventories.

3.1.2.7 Requirements for Quality Control and Quality Assurance of Emission Inventory

The EICG is placing special emphasis on ensuring consistency in emission inventory preparation procedures between all the districts in the modeling domain. The Air Resources Board will take the lead in coordinating and ensuring quality control and quality assurance standards for the emission inventory portion of the study. Specific guidelines will be outlined for assisting state and local agencies for implementing more uniform and systematic approaches to collecting, compiling (formats and accuracy), and reporting emission inventory data. Also, to ensure good quality control practices, the EICG will be developing a comprehensive quality control and quality assurance plan. The plan will include items such as:

- Quality control checks for collecting non-emission data, updating activity data, and using appropriate emission factors for calculating emissions;
- Emission calculation methodology;
- Updating or revising categories to ensure that the latest methodology is used;
- Data submittal;
- Data correction;

- Documentation; and
- Quality review of the entire inventory.

3.1.2.8 Proposed Contractor Studies for Emission Inventory Development

The collection of most of the emission inventory data discussed in the previous sections will be provided by public agencies as in-kind contributions to the CCOS study. However, there are some critical elements of the CCOS emission inventory that the members of the EICG do not have sufficient resources and /or expertise to provide. In general, these needs relate to the spatial and temporal distribution of emissions from motor vehicles and other area sources. The emission inventory elements, which will need to be supported by the CCOS study, are listed below.

Day-Specific Traffic Count Information

Day-specific traffic count information will be gathered during the field study. These traffic counts are used to create factors to adjust the travel demand volumes to hourly vehicle volumes. A scoping study is planned to review the availability of traffic data and develop a sampling plan using statistical methods for the placement of traffic counters for CCOS. Based on the results of the scoping study, a field sampling plan will be developed to collect and analyze hourly traffic count information for the CCOS domain. The day-specific data will be incorporated into the transportation model.

Gridding of On-Road Mobile Source Emissions

The first step needed to create an on-road motor vehicle inventory is to integrate the transportation data and run DTIM (Direct Travel Impact Model) for the entire CCOS domain. After compilation of the necessary transportation data from the local COGs who have transportation planning responsibilities, approximately thirty transportation networks can be meshed into a single uniform network for CCOS. Once a uniform mesh is created, the original VMT and speed data will be reassigned to the uniform mesh. ARB's MVEI (Motor Vehicle Emission Inventory) models and DTIM can then create gridded, hourly emission estimates of NO_x, TOG, CO and PM for on-road mobile sources.

Gridding of Area Source Emissions

This project is to develop base year and future year gridding surrogates for spatially distributing area source and off-road emission source categories. The emissions from these categories are estimated at the county level. For use in photochemical models, area and off-road emissions are spatially distributed within a county through the use of surrogates. In this context, a surrogate is an economic or demographic parameter that is correlated to the emissions for a specific area or off-road source emissions and has a known spatial distribution. Some common surrogates used for this purpose are population, housing units, and zoning designations. The last update of the spatial surrogates for area and off-road sources was done for the 1994 plans for attaining the one-hour ozone standard and was based on 1990 data. These surrogates need to be

updated to reflect the actual conditions in the year 2000 using the most recent data available. Surrogates will also need to be projected for future years for use in attainment demonstrations.

Biogenic Emission Model Expansion

The ARB is working to develop a statewide biogenic hydrocarbon inventory. A GIS-based biogenic inventory has been created for southern California. ARB is extending the biogenic inventory further north. However, data on some of the major plant species are missing for the Central Valley and the Bay Area. Biomass and emission factor data will be needed for these species.

Day-Specific Information on Agricultural and Controlled Burns

This project would be the collection of day-specific data for controlled burns and agricultural burns. Local air districts track some information about controlled and agricultural burns. Because the burn may not occur on the day that the permit is issued, there is some concern about the accuracy of the data for day-specific use. This research project would review available data and collect information needed to estimate emissions daily during the study period from controlled and agricultural burns.

3.1.3 Air Quality Modeling

This section discusses commonly used air quality models and their chemical mechanisms, advection schemes, grid setup, and boundary and initial conditions. Specific data needs for input into the air quality model are then presented. These are grouped by meteorological files (output from the meteorological model), gridded emission inventory files, and air quality files (for boundary and initial conditions).

There are several air quality models that are well established and within the state-of-the-art. The models include CIT, CALGRID, UAM-V, CAMx, CALPUF and EPA's MODELS-3. State-of-the-art air quality models can be applied to simulate photochemical air pollution and aerosol particles. Many of the models can be used to investigate air pollution from urban to regional scales. The models are grid models and some allow two-way nesting so that the model can simulate air pollution on a larger regional scale and on a more highly resolved urban scales simultaneously.

Three different chemical mechanisms are in wide use; carbon bond IV (CBM-IV), SAPRC 1997 and RADM2, all of these mechanisms were validated by extensive testing against environmental chamber data. The chemistry solvers used in the models are relatively fast but more work is needed to make them more general so that it is easier to make changes in the chemical mechanisms. Aerosol formation is not yet understood on a fundamental level. The formation of aerosol particles from the reactions of NO_x, SO₂ and from volatile organic compounds (VOC) are often treated by an empirical thermodynamic approach.

The advection of trace gases has seen some improvement in recent years. There are several possible horizontal advection solvers. The Bott advection solver and the Smolarkiewicz solver are two typical choices. The Bott solver is becoming more popular than the

Smolarkiewicz solver because Bott is faster and more accurate with less numerical diffusion than the Smolarkiewicz solver.

The typical files required by an air quality model are shown in Figure 3.1-1. Gridded data files should be available with resolution of at least 1 by 1 km if possible. An air quality model requires files with gridded land use and surface cover over the modeling domain. Typical land use files divide the land use into at least 11 categories (such as urban, rocky, industrial, etc.). A file of the UV surface albedos over the modeling domain is also required.

Meteorological Files for Input to Air Quality Model

Typical meteorological data files for air quality modeling are listed in Table 3.1-2. As discussed in Section 3.1.1, these gridded fields may be generated through the use of objective analysis of observations or through the use of a meteorological model. The height of the modeling grid must be greater than the mixing height field. The mixing height must be specified relative to the ground surface because the grid follows the local ground elevation. Nighttime conditions for neutral and unstable conditions are directly calculated by the model. The hourly wind components (x, y, z) must satisfy Equation (3-1) so that the flow fields are mass consistent.

$$\nabla \bullet \mathbf{u} = 0 \quad (3-1)$$

The rate of chemical reactions is highly dependent upon the temperature. The surface temperature field is used by the chemical solver to determine the correct chemical rate constants. The absolute humidity field is needed by the chemical solver because many reactions, especially the generation and loss of HO_x radicals, are humidity dependent. The solar radiation scaling field is needed to calculate the atmospheric stability. The ultraviolet radiation scaling field is required to calculate photolysis rates. Both the solar radiation scaling field and ultraviolet radiation scaling field are affected by the presence of clouds and this effect should be incorporated into the data files for these quantities.

Emissions Inventory Input to Air Quality Model

A great amount of care must be taking in preparing the emissions inventory. Without a reasonably accurate emissions inventory, it is not possible to accurately model the photochemical formation of ozone and other pollutants. There are two required emissions files; the area source emissions inventory and the elevated point source emissions inventory. Table 3.1-3 shows typical chemicals that need to be included in the emissions inventory. However, the organic chemical emissions should be given in as much detail as possible, rather than just alkane, alkenes etc, and should address biogenic sources.

Area source emissions may be considered as part of the boundary conditions (McRae, 1992). The area source file requires that the emissions be given in units of ppm meter / min. The height of the surface grid box is used to calculate the emissions in units of ppm / min, which is often called an “emissions frequency” since it has the same units of inverse time.

Point sources are assumed to be volume sources in the model filling an entire grid square. If the emissions are vented through a stack, the stack height, diameter, exhaust temperature, flow rate and source strength are needed to calculate the plume rise height and emissions rate.

Air Quality Input to Air Quality Model

Air quality files of initial conditions and boundary conditions are required. Each grid square and each layer in the modeling domain starts with a concentration for each chemical species. One of the difficulties of determining the initial conditions is that there are not enough measurements available to determine the three dimensional chemical fields for starting the simulations, thus these initial concentrations must be estimated from a sparse network of ambient measurements. It will be necessary to interpolate surface concentrations to estimate vertical distributions. If there are even a few measurements of vertical chemical distributions, e.g., from aircraft, the relative changes in chemical concentration with height could be used for this interpolation. Otherwise, typical values for other urban areas will have to be used (Milford et al., 1989). The CCOS modeling team should examine the different methods by which initial conditions are estimated from surface and airborne measurements, and state the equations, assumptions, inputs, and uncertainties for these methods. The team should then determine which methods are most applicable to the single point and aircraft measurements, and determine which methods are most useful for approximating initial conditions for integrated air quality modeling.

For the boundary conditions it would be desirable to set the boundaries at locations with relatively low emission rates. VOC concentrations are usually assumed to be constant or negligible at the western and northern boundaries and at the top of the study domain. As shown in Section 2.8, boundary conditions, particularly for formaldehyde, significantly affected model outputs in the SARMAP modeling, resulting in over-prediction of ozone levels in the Bay Area. Thus, sufficient field measurements of VOC must be made to realistically represent boundary conditions in the air quality model. Aircraft measurements of the western boundary are deemed essential.

3.2 Model Evaluation

The reliability of model outputs is assessed through operational and diagnostic evaluations and application of alternative diagnostic tools. Operational evaluations consist of comparing concentration estimates from the model to ambient measurements. Diagnostic evaluations determine if the model is estimating ozone concentrations correctly for the right reasons by assessing whether the physical and chemical processes within the model are simulated correctly. This broad task also involves reconciliation of data analysis and observation-based results with modeling results. It includes the evaluation of modeling uncertainties, processes, and assumptions and their effect on observed differences among model results, measurements, and data analysis results.

3.2.1 Evaluation of Meteorological Modeling

Components of meteorological model evaluation include collection of sufficient measurements to support the following generic evaluation tasks:

1. Compare the simulated atmospheric processes using MM5 in both fully predictive and data assimilation modes. Since a number of airborne and remote sensing upper-air measurements will be available, it is desirable to place emphasis on the data assimilation mode for the entire field program period. The fully prognostic mode may be desirable for understanding basic characteristics of specific weather episodes with high pollutant concentrations, but results from the two approaches, with and without 4DDA, should be compared.
2. Determine and compare modeled and observed flow patterns in northern and central California with acceptable horizontal and vertical resolution, using physical parameterization of the main atmospheric processes (radiation, moisture, clouds, and fog).
3. Compare detailed information on the vertical wind and temperature structure of the atmospheric boundary layer with modeled results during the case studies. From aircraft observations, determine elevated layers with specific stability and dynamics, and compare to concentrations, stability and dynamics in modeled layers. The vertical structure is essential for estimates of transport and dispersion of atmospheric pollutants, as well as for determination of the amount of decoupling of local flows from the air aloft. Determine to what extent differences are attributable to lack of vertical resolution in the model.
4. Determine properties of land-sea breezes, urban circulations, local flows (slope and drainage), and diurnal variation of thermal stability and shear, and compare modeled results to any available observation-based estimates of these features.
5. Determine spatial characteristics of mixing depth for both convective and stable conditions and compare to aircraft spirals or other observation-based methods of determining mixing heights (e.g., profiler return amplitude data).
6. Estimate properties of turbulence transfer and associated vertical fluxes in the boundary layer.
7. Conduct sensitivity tests of the input parameters (topographic resolution, model grid, synoptic fields vs. radiosonde network, range and variation of sea/surface temperature, urban effects/roughness, sinks and sources of heat) to estimate uncertainties in model results.
8. Quantify differences between the predicted and observed wind fields and stability parameters for specific case studies.

3.2.2 Evaluation of Emission Inventory Estimates

Section 3.1.2 described the planned development of the CCOS emission inventory data. Uncertainties in precursor emissions are largely viewed as one of the weaker links in the air quality modeling process. Independent evaluations of emission inventory estimates must be an integral component of air quality modeling studies. If ambient measurements and emission data are correct, consistent relationships between the spatial, temporal, and meteorological variability

in emissions and ambient measurements should be found. The tasks associated with this objective will examine these relationships and will explain why they are observed or are not observed.

Approaches for evaluation of emission inventories include: 1) performance evaluations of air quality simulation models; 2) source apportionment by receptor modeling; 3) spatial and temporal comparisons of ambient and emission inventory non-methane organic gas speciation profiles and pollutant ratios (e.g., CO/NO_x and VOC/NO_x); 4) comparisons of long-term trends in ambient pollutant concentrations and concentration ratios with emission inventory trends; 5) comparisons of on-road measurements with motor vehicle emission models; and 6) fuel-based inventory based on regional gasoline sales and fleet-averaged, fuel-based emission factors from remote sensing measurements. These approaches are described below.

Compare Proportions of Species

Comparisons of the ambient pollutant ratios with corresponding ratios derived from emission inventory estimates have been widely used to determine consistency between emission inventory estimates and ambient measurements. This approach has been used in the Southern California Air Quality Study (Fujita et al., 1992), the San Joaquin Valley Air Quality Study (Fujita et al., 1994), the Lake Michigan Ozone Study (Korc et al., 1993) and by the ARB in response to requirements of SB 2174 (ARB, 1997). The premise of this evaluation approach is that in areas with high emission rates, early morning NMOG/NO_x and CO/NO_x ratios and speciation profiles in emission estimates should compare well with observed ambient ratios and speciation profiles for the same time period. Issues that are important to address in this type of comparison include: 1) the importance of ground level versus elevated plumes and their influence on ground level measurements; 2) the location of ambient monitoring sites (at urban locations, nearly all sites are located adjacent to roadways); 3) pollution carry-over from the previous day(s); 4) the reliance upon NO_x as the reference species for the "top-down" ratio comparison, including the assumption that NO_x emissions are reasonably accurate; 5) hydrocarbon speciation differences between ambient data and the inventory; and 6) representativeness of the monitoring site in characterizing a grid cell's emissions and ambient concentrations.

Calculate the average VOC/NO_x and CO/NO_x ratios and ratios of selected species (e.g., acetylene/benzene, ethylene/acetylene, benzene/xylenes, carbon monoxide/selected VOCs, total aromatics/total VOC, total biogenic species/total VOC, >C₁₀/total VOC, benzene/VOC, MTBE/VOC) in source-dominated areas for samples which are expected to be dominated by fresh, local emissions (e.g., morning samples). Compare these ratios with those in emission inventory grid squares in the vicinity of the sampling site and with ratios in speciated profiles.

Calculate the average ratios of VOC/NO_x and ratios of selected species (e.g. homologous groups/total VOC, oxygenates/total VOC, methyl vinyl ketone and methyl acrolein/isoprene, other daughter products/precursors) for samples that are expected to contain end products of photochemical reactions. Compare these ratios with those in emissions inventory grid squares in the vicinity of the sampling site and with ratios in speciated profiles to determine the extent to which primary emissions and secondary products contribute to the entire VOC loading. Explain

differences between inventoried and measured VOC/NO_x ratios in terms of chemical reactions, if possible.

Source Apportionment by Chemical Mass Balance Receptor Modeling

The Chemical Mass Balance (CMB) model (Friedlander, 1973; Cooper and Watson, 1980; Gordon, 1980, 1988; Watson, 1984; Watson et al., 1984; 1990; 1991; Hidy and Venkataraman, 1996) consists of a least-squares solution to a set of linear equations which expresses each receptor concentration of a chemical species as a linear sum of products of source profile species and source contributions. The source profile species (the fractional amount of the species in the NMOG emissions from each source type) and the receptor concentrations, each with uncertainty estimates, serve as input data to the CMB model. The output consists of the contributions for each source type to the total ambient NMOG as well as to individual NMOG concentrations. The model calculates values for contributions from each source and the uncertainties of those values. Input data uncertainties are used both to weight the relative importance of the input data to the model solution and to estimate uncertainties of the source contributions. CMB software currently in use (Watson et al., 1990) applies the effective variance solution developed and tested by Watson et al. (1984) because: 1) it calculates realistic uncertainties of source contributions from both the source and receptor uncertainties; and 2) chemical species measured more precisely in both source and receptor samples are given greater influence in the solution than are less precisely measured species. The software also incorporates collinearity measures (Henry, 1982, 1992) to assess the effects of source profile similarity on source contribution estimates and their standard errors. The software is interactive, allowing many sensitivity and assumptions-testing sets to be performed rapidly.

In addition to developing the CMB software, DRI investigators have developed and formalized the protocol for applying and validating the CMB model (Pace and Watson, 1987; Watson et al., 1991; 1998) for apportioning particles (Watson et al., 1994) and gaseous organic compounds (Fujita et al., 1994). In the past six years, DRI investigators have applied the CMB receptor model to PAMS or PAMS-type speciated NMOG data from southern California (Fujita et al., 1994), San Joaquin Valley, CA (Fujita et al., 1995a), southeast Texas (Fujita et al., 1995b), northeastern U.S. (Fujita and Lu, 1998a), Phoenix, AZ (Fujita and Lu, 1997a), western Washington (Fujita et al., 1997b), El Paso, TX (Fujita, 1998), and Austin, TX (Fujita et al., 1999).

The application of continuous speciated VOC data in source apportionment offers additional insights regarding the temporal variations in source contributions that are difficult to discern from a limited number of canister samples that are integrated over a period of 3 hours or more. The following graphical display show examples of CMB results obtained from hourly speciated hydrocarbon data.

- Average source contribution estimates of ambient hydrocarbons by hour of day (e.g., Figure 3.2-1)
- Diurnal plots of the average CMB source contribution estimates by site for each day of the week (e.g., Figure 3.2-2).

- CMB source contributions and residual hydrocarbon concentrations and graphical displays of residuals by wind direction and time of day (e.g., Figure 3.2-3).

Another useful graphical display that can also be applied to time-integrated canister samples show relationships between source contribution estimates (in particular the residual unexplained mass) and extent of reaction of the ambient air sample (estimated by a ratio of reactive to unreactive hydrocarbon species). Scatterplots of CMB-predicted versus measured concentrations for reactive species by site and time of day are also useful in examining the photochemical age of the ambient hydrocarbons.

For CCOS, we propose comparing the ambient source apportionment at selected measurement sites with a model and emission inventory based CMB analysis. A dispersion model could be used to estimate the concentrations of the various species at measurement sites based upon the meteorology and emission inventory. This calculation should be made for specific chemical species and not lumped species. First order loss chemistry based upon rate constants and concentrations of HO and O₃ could be included in the dispersion models for each reactive VOC. The HO concentrations would be calculated from 3-d model such as CAMx. The dispersion model estimated concentrations at a specific measurement site would be subjected to CMB analysis. Comparison of the observed source apportionment with the model derived source apportionment will provide an excellent test of the emission inventory. Comparison of the model derived source profiles with and without chemistry should provide a useful check of the CMB approach. Specifically compare ozone precursor, NO_x, CO and VOC on a diurnal and weekday basis as estimated by model with the measurements should also be included.

Remote Sensing of Vehicle Exhaust Emissions and Fuel-Based Inventories

With initial support from the Colorado Office of Energy Conservation in 1987, the University of Denver (DU) developed an infra-red (IR) remote monitoring system for automobile CO exhaust emissions. The current instrument measures CO, HC, NO and smoke opacity. The additional capability to simultaneously monitor aromatic hydrocarbons and low molecular weight aldehydes is currently under development. The remote sensing detector (RSD) has been widely used to identify high-emitting vehicles and to characterize the emissions distributions for large fleets of on-road motor vehicles.

The instrument was designed to emulate the results one would obtain using a conventional non-dispersive infra-red (NDIR) exhaust gas analyzer. Thus, the RSD is also based on NDIR (NDUV for NO). An interference filter that transmits infra-red (IR) light of a wavelength known to be absorbed by the molecule of interest is placed in front of a detector. Reduction in the signal caused by absorption of light by the molecules of interest produces a reduction in the detector's voltage output. One way of conceptualizing the instrument is to imagine a typical garage exhaust NDIR instrument in which the separation of the IR source and detector is increased from 10 cm to 20-40 feet. Instead of pumping exhaust gas through a flow cell, a car now drives between source and detector. The light source, across the road, now contains a deuterium lamp, which is mounted in such a manner that the net result from the source is a collimated beam of UV and IR light. Because the effective plume path length and amount of plume seen depends on turbulence and wind, one can only look at ratios of CO, HC, of NO to CO₂. These ratios are termed Q for CO/CO₂, Q' for HC/CO₂, and Q'' for NO/CO₂ and are

constant for a given exhaust plume. By themselves, Q and Q' are useful parameters with which to describe the combustion system.

With the aid of a fundamental knowledge of combustion chemistry, many parameters of the vehicle's operating characteristics can be determined including the instantaneous air/fuel ratio, the %CO, %HC, or %NO which would be read by a tailpipe probe, and the grams CO, HC, or NO emitted per gallon of gasoline (gCO/gallon or gHC/gallon) emissions (Bishop and Stedman 1996). The following formulas assume a fuel density of 750 g/L and 0.86 g carbon per gram of fuel (Stedman 1998).

$$\text{gCO/gal} = 5671 * Q / (1 + Q + 3*Q')$$

$$\text{gHC/gal} = 8911 * Q' / (1 + Q + 3*Q')$$

$$\text{gNO/gal} = 6076 * Q'' / (1 + Q + 3*Q')$$

Remote sensing has been shown to give accurate readings for CO by means of double-blind studies of vehicles both on the road and on dynamometers (Lawson et al., 1990; Stedman and Bishop, 1990). EPA has shown that the readings are closely comparable to laboratory readings from a vehicle on a dynamometer (Stedman and Bishop, 1990). Lawson and coworkers used a vehicle with variable emissions under passenger control to show the correctness of the on-road readings (Lawson et al., 1990). Independent studies (Ashbaugh et al., 1992) show that the CO readings are correct within $\pm 5\%$ and HC within $\pm 15\%$. Recent work (Stedman et al., 1997) has shown that fleet average on-road emissions by model year correlate with IM240 readings with r^2 greater than 0.95 for CO, HC and NO.

The fact that remote on-road readings are well correlated with more complex tests is also illustrated by a study in California (CARB, 1994 and Knapp 1992) in which the remote sensing readings were used in California immediately to pull over apparently gross polluting vehicles which were then tested by a team of Smog-Check engineers, and then brought to a dynamometer and subjected to the EPA IM240 test. Of 79 vehicles tested on IM240, 76 failed and the three which passed had all failed the previous Smog-Check.

The remote sensor is accompanied by a video system when vehicle identification information is required. The video camera is coupled directly into the data analysis computer so that the image of each passing vehicle is frozen onto the video screen. The computer writes the date, time, and the CO, NO, HC and CO₂ concentrations at the bottom of the image. These images are then stored on videotape. The digital database also contains uncertainty limits and opacity data. The height of the sensing beam is typically set at 20-30 cm above the road surface optimally to observe exhaust plumes from light duty vehicles, including gasoline and diesel-powered vehicles, as long as the exhaust plume exits the vehicle within a few feet of the ground.

The remote sensor is effective across single or multiple traffic lanes of up to 40 feet in width. However, if one wishes to positively identify and video each vehicle with its exhaust, it is optimally efficient when used across a single lane of traffic. FEAT operates most effectively on dry pavement. Rain, snow, and vehicle spray from very wet pavement cause interferences with the optical beam. These interferences do not cause incorrect readings, rather they cause the frequency of invalid readings to increase, ultimately to the point that all data are rejected as

being contaminated by too much "noise". At suitable locations exhaust can be monitored from over one thousand vehicles per hour.

Determine the Effects of Meteorological Variables on Emissions Rates

Select representative sampling sites from major source regions (e.g., urban, industrial, biogenic, etc.) and examine concentrations of "marker" species for sources as functions of temperature, relative humidity, wind speed, and other environmental variables which are used to adjust emission factors. If possible, draw conclusions concerning the efficacy of current emission factors to respond accurately to changes in meteorological variables. Compare VOC source contributions when intermittent sources are known to operate and when they are known not to operate. Attempt to detect effects of different wind speeds, temperatures, solar radiation levels, and relative humidities on biogenic and industrial emissions.

Determine the Detectability of Day-Specific Emissions at Receptors

Examine day-specific emissions and identify sampling locations and times that correspond to fires, pesticide applications, spills, etc. From previous measurements of emission compositions, identify chemical species that are likely to be contributed by that source. Examine transport and dispersion patterns to determine the likelihood of influence at nearby sites. Compare the ambient concentrations at likely impact sites with concentrations at that site when day-specific emissions do not exist. Examine nearby sites and draw conclusions about the region of influence of intermittent emitters. Draw conclusions regarding the importance of intermittent emissions for ozone formation.

3.2.3 Air Quality Model Evaluation

The reliability of model outputs is assessed through operational and diagnostic evaluations and application of alternative diagnostic tools. Operational evaluations consist of comparing concentration estimates from the model to ambient measurements. The key question in an operational model evaluation is to determine the extent of agreement between simulated and measured concentrations of ozone and its precursors. The measured and simulated ozone and its precursor concentrations should agree in their spatial extent and in their timing. Typical statistics for model evaluation: the comparisons of predicted and observed daily maximum 1-hr observed ozone concentrations, comparison of 90th percentile concentrations, mean bias (ppb) and mean normalized bias (%) and the mean error (ppb) and mean normalized error (%) (Lurmann et al., 1998).

Examples of other operational evaluation include comparison of model predictions with airborne measurements of VOCs (total VOC, homologous groups, and lumped VOC classes) along boundaries and above the mixed layer and the magnitude and constancy of their concentrations in space and time. List assumptions and input data requirements for VOC boundary conditions, and estimate the effects of uncertainties caused by insufficient data on the ability of the model to represent reality. Nitrogenous species concentrations are usually assumed to be constant or negligible at the western and northern boundaries and at the top of the study domain. Plot airborne measurements of NO, and NO_y along boundaries and above the mixed layer, then examine the magnitude and constancy of their concentrations in space and time.

Specify the model assumptions for boundary conditions of nitrogenous species, and estimate the effect if deviations from those assumptions on chemical concentrations.

Diagnostic evaluations are required to determine if the model is estimating ozone concentrations for the right reasons. The emissions, chemistry and transport are assessed to determine if these are treated correctly within the model. The emissions may be assessed through process analysis and mass balance analysis. The chemistry may be assessed through the measurement of predicted secondary chemical products. The photolytic rate parameters used in models should be evaluated by comparing parameters estimated by measurements versus model default values. The transport might be evaluated through comparisons of the spatial and vertical distributions of ozone and its precursors. This broad task also involves reconciliation of data analysis and observation-based results with modeling results. It includes the evaluation of modeling uncertainties, processes, and the assumptions, and their effect on observed differences among model results, measurements, and data analysis results. A key finding of a diagnostic evaluation should be to determine the physical and chemical reasons for the concentration differences between those predicted by models and the measurements. It would also be highly desirable if the likely model bias induced by a model's deficiencies could be identified. Examples of diagnostic tests include examining ratios of chemical species that are sensitive to specific processes within the model as as O₃/NO_y, O₃/NO_z, examining the flux of ozone and ozone precursor across interbasin transport corridors, and comparing concentration changes from weekdays to weekends.

Jeffries and Tonnesen (1994) have developed methods to modify air quality models so that the rate of each chemical and physical process that affects a chemical concentration is integrated and saved in external files. Post processors are used to analyze the integrated rates (masses determined for each process), to make a detailed mass balance for a simulation. From the mass balance calculations it is possible to determine quantitatively the effect of each process on the concentrations of ozone or any air pollutant. Although this process analysis works well in models it is difficult to apply directly to field measurements due to the lack of data in any reasonable field campaign.

Due to inherent limitations of measurement techniques and budget considerations, it is not possible to develop an independent, comprehensive mass balance throughout the study region. However, an alternative that should be applied to CCOS is a "model assisted mass balance." In this approach, measurements are used to check key parameters within simulations made for process analysis without the assistance of a model. For example, the concentration ratio of NO to NO₂ or the concentration ratio of a highly reactive VOC to a less reactive VOC could be compared with the model results. Alternatively model calculated parameters such as HO could be used along with measured concentrations of VOC to estimate VOC to NO_x ratios. Both approaches should provide strong tests of the conceptual model.

3.3 Data Analysis Methods

As discussed in Section 1.2, data analysis is an essential part of any large field study and will complement and enhance the model development component of CCOS. The major goals of CCOS data analysis are to:

- Evaluate field measurements for applicability for model input, parameterization, evaluation, and verification.
- Describe the air quality and meteorology during the field study period and determine the degree to which these measurements represent other summertime pollution levels.
- Further develop conceptual models of physical and chemical processes which affect ozone formation and transport in southern California.

While these data analysis goals provide guidance, a critical outcome of the CCOS planning process is to specify the:

- Data analysis objectives and hypotheses to be tested.
- Data analysis methods and work elements by which objectives can be met and hypotheses can be tested.
- Required field measurements to collect sufficient data to apply specified data analysis methods and work elements.

With these three planning carried out and integrated into publications and reports to meet data analysis objectives. Data analysis activities are defined here within the following topic areas: 1) measurement evaluation; 2) spatial, temporal, and statistical descriptions; 3) qualitative transport characterization; 4) dispersion characterization; 5) emission characterization; 6) quantitative transport characterization (pollutant fluxes); 7) ozone chemistry; 8) episode characterization; and 9) refinement of conceptual models. Activities are detailed in the following sub-sections.

3.3.1 Accuracy, Precision, Validity, and Equivalence of Field Measurements

There has never been a field study to date that did not require substantial examination and investigation of the measurements prior to their further use in data analysis and modeling. This first topic area is essential to all subsequent data analysis efforts, and needs to be performed as data are received into the database. Major concerns focus on: 1) accuracy of VOC peak identification, and proper, consistent quantification of total nonmethane hydrocarbons; 2) representativeness of meteorological measurements, especially those drawn from existing networks; 3) accuracy and precision of low-concentration measurements, especially NO_y and PAN; and 4) comparability of measurements taken from different networks with different procedures.

Evaluate the Precision, Accuracy, and Validity of Light and Heavy Hydrocarbon and Aldehyde Measurements

For sites and sampling periods which represent different expected mixtures of VOCs (e.g., background, fresh urban, aged urban, regional, forested, industrial), create scatterplots of the sum of measured species vs. total nonmethane VOC. Calculate slopes, intercepts and correlations and compare these among different sites and sampling times. Estimate the range of values for unaccounted hydrocarbons and the degree to which these differ at different sites.

Compare the sum of identified VOC concentrations to total nonmethane VOCs and identify significant differences among sampling sites and sampling periods. Plot hydrocarbon profiles (bar charts of percent composition for each carbon number and for the lumped parameter classes used in photochemical models) and compare these among different sampling periods and sampling sites.

Compare the detailed speciation of quality assurance samples with the routine speciation from normal network samples. Ascertain which species are not typically identified in normal network samples. Estimate the feasibility and effort required to reduce existing chromatograms with unidentified peaks for more complete speciation of normal network samples. Examine a selection of chromatograms from routine analyses and determine the feasibility and value of re-processing these chromatograms for more complete speciation. Use collocated and replicate analysis results to determine an overall precision for hydrocarbon measurements. Examine the discrepancies between different analyses as a function of aging time in the canisters and attempt to quantify the extent to which the gaseous contents change with time.

Evaluate sampling and analysis methods for aldehyde measurements by comparing collocated and replicate measurements. Determine the extent to which reactions take place in sampling bags used in aircraft from collocated measurements of bag and cartridge sampling. See Fung et al. (1993) for examples.

Evaluate the Precision, Accuracy, and Validity of Nitrogenous Species Measurements

Create scatterplots and calculate slopes, intercepts, and correlation coefficients for normal sensitivity chemiluminescent NO_x, high sensitivity chemiluminescent NO_y measurements sites, and spectroscopic measurements (e.g., TDLAS). Determine the equivalency of these different measurement methods by evaluating these plots and statistics. Examine differences among sites (along with calibration and performance test data) to attribute differences to instrument differences or to interferents in the sampled air streams (e.g. HNO₃ detected by chemiluminescence). Create scatterplots and statistics of collocated PANalyzer values. Explain differences in terms of measurement methods or environmental variables (e.g., interferents). Compare differences to propagated uncertainty intervals derived from performance tests and extrapolate the collocated uncertainties to other luminol PAN sampling sites in the network. Reconcile laboratory comparison data with field comparison data.

Evaluate the Precision, Accuracy, Validity, and Equivalence of Meteorological Data

Devise methods to compare upper air measurements of wind speed and direction from profilers, acoustic sounders, and surface wind towers and with corresponding measurements from radiosondes. Determine equivalent averaging layers and averaging times. Determine times of day (e.g., early morning, mid-day, and late-day) when these measurements are similar and when they are not. Draw conclusions regarding the equivalence of these methods. Compare surface measurements of wind speed and direction with the lowest elevation values from collocated profile, sonde, and sounder measurements. Stratify comparisons by time of day to obtain well-mixed and layered vertical structures. Determine when surface measurements are an adequate estimate of upper air winds and when they are not, both with respect to elevation above ground level and time of day.

Evaluate the Precision, Accuracy, Validity, and Equivalence of Solar Radiation Data

Compare nearby measurements from the spectral radiation, ultraviolet radiation, and total solar radiation instruments. Determine the extent to which total solar radiation and spectral radiation are correlated with the ultraviolet wavelengths which are most influential in photochemistry. Determine the equivalence or difference between measurements taken in the CCOS and existing networks.

3.3.2 Spatial, Temporal, and Statistical Distributions of Air Quality Measurements

More data will be produced by field monitoring than can be examined by any data analysis plan. Summaries need to be created which can serve as a guide to the database and for the formulation of hypotheses to be tested by more detailed analyses. These summaries may be examples drawn from a data display package (such as that which was developed to display data in real-time during the study). The database and display software could then be used in other data interpretation projects to provide support for their findings.

Examine Average Diurnal Changes of Surface Concentration Data

Create diurnal box plots (which graphically show quartiles, median, mean, and extrema) for the entire sampling period and for each hour at each site of ozone, NO, NO₂, and PAN concentrations. Note differences in the timing and intensity of peak values as a function of sampling site, episode, and stratification. Group sites for which diurnal variations behave in a similar manner. Compare these with plots from selected sites in prior summers where ozone and NO_x data are available. Evaluate the extent to which the summer of 2000 is similar to or different from prior summers.

Create diurnal box plots for each hour at each site of ozone, NO₂, and PAN (where available) for each episode and the stratified intensive sampling days. Superimpose diurnal average total VOC, total aldehyde, nitric acid, and particulate nitrate boxes for specified sampling periods. Note differences in the timing and intensity of peak values as a function of sampling site and how these are similar to or different from those of the longer-term averages.

Create diurnal box plots of surface temperature, relative humidity, insolation, sigma theta, and scalar wind speed for each hour at each site for the entire monitoring period and for the intensive episodes. Identify similarities and differences among sites and between the intensive episodes and the entire study period.

Examine Spatial Distributions of Surface Concentration Data

Plot spatial isopleths corresponding to each intensive episode sample for total VOCs, selected VOC surrogates (selected to represent reactivity class, sources of precursors, or end products), NO_x, NO₂, PAN, HCHO, NO/NO₂, VOC/NO_x, and O₃. Note similarities and differences in patterns with time of day, pollutant, and from episode to episode.

Examine Statistical Distributions and Relationships Among Surface Air Quality Measurements

Combine hourly averages of air quality measurements into sample averages corresponding to VOC and aerosol samples. For each site, calculate averages, standard deviations, first, second, third and fourth maxima (with date and time of occurrence), and minima (with date and time of occurrence) concentrations for each species measured. Identify differences between morning, afternoon and nighttime, sampling locations, episodes, and chemical observable. Give special attention to differences between rural vs. urban areas, central California vs. other air basins, morning vs. afternoon vs. nighttime.

Combine hourly averages of continuous measurements into sampling periods corresponding to VOC and aerosol samples. For each site, calculate the temporal correlation coefficients for all measured variables. Identify those variables that are highly correlated (negatively or positively) with each other and identify observables which might be represented by a single surrogate at each site. Combine these surrogates with meteorological variables and note positive or negative correlations among them.

Perform Principal Component Analysis (PCA) on VOC and nitrogenous species. Calculate eigenvectors of the correlation matrix and perform a varimax rotation to identify empirical factors that explain the variability in the data. Describe these factors in terms of physical phenomena, and examine factor scores to determine when each factor has much greater than or much less than average influence at each site.

Using selected species concentrations, calculate spatial correlations for the intensive episode samples. Examine correlations and identify which sites are highly correlated (positively or negatively) with each other. Calculate eigenvectors of this correlation matrix, perform a varimax rotation, and plot the empirical orthogonal functions that are deemed significant. Use these to select surrogate sites which can be used to represent neighboring sites for different observables (the surrogates and the area which they represent may not be the same for all observables).

Plot surface wind roses for all data at selected times of the day on a single map. Identify when flow reversals take place and when wind speeds increase or decrease. Note which areas have the highest frequencies of calms and when they occur. Perform these analyses for meteorologically stratified cases. Perform analyses for stations near sea-level as well as those located on higher terrain, and compare and explain differences.

Examine Vertical Distribution of Concentrations from Airborne Measurements

Plot VOC (total and selected species, aldehydes (total and selected species), NO₂, NO, O₃, and bscat as a function of altitude for each spiral. Note similarities and differences with respect to location, time of day, and chemical species. Give special attention to VOC speciation in morning samples (i.e., local emissions) over urban (motor vehicle), non-urban (biogenic), and industrial (oil-field) areas. Compare canister and cartridge samples of VOC and aldehydes taken in spirals with those taken in circles at the top of the spiral.

Plot spatial distributions of VOC (total and selected species, aldehydes (total and selected species), NO₂, NO, O₃, and bscat along aircraft traverses (within the mixed layer). Note similarities and differences with respect to location, time of day, chemical species, and the

spatial distributions derived from surface-based measurements. Compare concentrations along boundaries (long-range aircraft) with those measured at locations within the study domain and note similarities and differences.

Examine the Spatial and Temporal Distribution of Solar Radiation

From the radiometer data, estimate the photosynthetically active radiation (needed for deposition and biogenic emissions), direct beam solar radiation, diffuse solar radiation, visible radiation (for light extinction), incident flux, and actinic flux (at frequencies relevant to photochemical reactions). Create spatial and temporal plots of these observables. Describe differences between sites and time-of-day in terms of measurement uncertainty, meteorological and air quality parameters. Note the effects of clouds and smoke on different types of radiation.

3.3.3 Meteorological Transport Phenomena

This topic addresses the major mechanisms for the movement of air into, out of, and between the different air basins in both horizontal and vertical directions. This requires an examination of conditions near sea level and also just below, within, and above the inversion layer.

Determine Horizontal Transport Patterns and Intensities Into, Out of, and Within the Air Basins

Plot 0400, 1000, 1600, and 2200 PST horizontal windfields at three different heights (surface, within the mixed layer, and above the mixed layer) using continuous and radiosonde data. Examine the consistency of these flow vectors with those predicted from synoptic weather maps and pressure gradients. Note similarities and differences with respect to time of day, elevation, and episode. Associate the directions with the expected phenomena of low-level jet, slope flows, eddies, coastal meteorology and bifurcation. Examine aircraft data to further describe the evolution of these phenomena. Determine the time of occurrence, spatial extent, intensity, and variability of these phenomena.

Plot detailed horizontal wind vectors as a function of height for the Carquinez, Altamont Pass, Pacheco Pass. Examine these to determine the intensity and duration of transport from the the Bay Area to the Sacramento and San Joaquin Valleys. Note similarities and differences with respect to time of day, elevation, and episode.

Plot back-trajectories for critical receptors (e.g., ozone hot-spots, forested sites, and transport corridors) for each intensive episode at three elevations. Start trajectories at the time of maximum ozone concentration. Identify general areas over which air masses might have passed to reach these receptors. Classify sampling periods into categories that are likely to be influenced by different types of source areas (e.g., industry, traffic, forest). Track horizontal pollution movements (e.g., visibility analyses) and determine if ozone values at stations downwind can be traced to photochemical changes on precursors upwind.

Determine Vertical Transport Patterns and Intensities within the Modeling Domain

Examine wind flow patterns to identify convergence and divergence zones. Examine acoustic sounder, profiler, and aircraft meteorological data for evidence of vertical exchanges in

these regions. Determine the extent to which surface air is transported above the mixed layer, or air above is transported into the mixed layer at these locations. Verify this with vertical profiles of pollutant concentrations from onboard aircraft measurements. Examine and describe the intensity and duration of upslope flows to estimate the amount and frequency with which pollutants might be transported above the mixed layer. Verify this with onboard aircraft pollutant measurements. Plot the vertical velocity structure as a function of time. Examine monostatic sounder data and profiler data to determine the degree of layering in the atmosphere, especially during the morning. Identify the locations of wind shears and their effect on layering. Note the differences in layers at different locations throughout the study area.

3.3.4 Meteorological Dispersion Processes

Dispersion processes address the mixing of pollutants within the mixing layer, especially elevated and ground-level emissions, dispersion within and between modeling grid cells, and transport to the surface where deposition of pollutants may occur.

Characterize the Depth, Intensity, and Temporal Changes of the Mixed Layer and Characterize Mixing of Elevated and Surface Emissions

Plot the spatial distribution of expected mixing depths derived from temperature soundings at 0400, 1000, 1600, and 2200 PST. Examine sounding, aircraft, and profiler data to determine the accuracy of interpolations of mixing depths between the 4 per day soundings. Examine aircraft and profiler data for evidence of other layers within the mixed layer, their time of formation and dissipation, and their typical duration and intensity. Describe the changes in layers as a function of time, especially during the morning when rapid changes are taking place with heating. Associate changes in layers with changes in surface temperature and solar radiation.

Examine vertical mixing as a function of location and time of day using aircraft data and continuous profiler and acoustic sounder measurements. Estimate the times of day on which pollutants emitted from stacks, and pollutants carried over from the previous day above the mixed layer, will combine with pollutants emitted at the surface. Verify this by examining aloft and surface level concentrations which are associated with aged emissions.

3.3.5 Characterize Pollutant Fluxes

A "flux plane" is a rectangular cross-section that is perpendicular to the prevailing horizontal wind direction at a location between major emissions areas. The major transport pathways that are suspected of passing through these flux planes were specified in Section 2.

Define the Orientations, Dimensions, and Locations of Flux Planes

Examine wind fields to identify areas in air that is transported across a boundary at vertical levels, especially at entry and exit points to the Central Valley. Specify the horizontal and vertical coordinates of these flux planes. Examine the usefulness of different conceptual definitions of flux (i.e., mass/unit area/time, mass/time, upwind/downwind concentrations along a wind vector). Estimate uncertainties due to: 1) mis-specifying the portion of ozone flux attributable to background vs. that generated in the upwind source area; 2) the effects of vertical

and horizontal wind shears or reversals on the definition of the flux plane; 3) variations in wind speed and direction between measurements; 4) mis-specification of the boundary plane height; and 5) effects of wind speed and direction variability on flux estimates.

Estimate the Fluxes and Total Quantities of Selected Pollutants Transported Across Flux Planes

Using aircraft spiral and traverse data, lidar data, and ground-based concentration data for VOCs, NO_x, and O₃ coupled with average wind speeds that are perpendicular to the chosen flux planes, calculate the mg/m²-sec of each pollutant which crosses each plane as a function of time of day. Compare the fluxes for the different planes and assign downwind fluxes to a combination of fresh pollutant generation and contributions from the upwind flux plane. Examine fluxes at different layers, especially at night, if major differences are observed in vertical concentrations and wind speeds. Plot vertical cross-sections of concentrations, wind speed, and direction.

Compare the magnitudes of inflow to and outflow from regions which are bordered by flux planes. Advance explanations for major differences between inflow and outflow. Using all relevant field study data, test the hypotheses that: 1) there is significant local generation of pollutants; 2) there is significant venting through the mixed layer; and 3) there is substantial reverse or lateral transport owing to eddies, nocturnal jets, and upslope/downslope flows.

3.3.6 Characterize Chemical and Physical Interactions

Ozone, several nitrogenous species, and significant portions of the VOCs found in the study domain are not emitted directly from sources, but form from precursors. In particular, it is necessary to determine where or when ozone concentrations are limited by the availability of NO_x or VOC. These issues are addressed within this topic.

State the equations, assumptions, and input data for ozone formation. Identify those species in these models which were measured during CCOS. Evaluate the uncertainties introduced by non-continuous, 2-hour VOC measurements, variations in solar radiation, and uncertainties in boundary and initial conditions.

VOC and Nitrogen Budgets

Plot pie charts of VOC speciation (with species expressed as mg/m³ carbon categorized into homologous groupings such as paraffins, olefins, aromatics, formaldehyde, other aldehydes, heavy hydrocarbons, unidentified peaks, and total nonmethane VOC minus sum of peaks and as average carbon number for <C₁₀ and >C₁₀) maps for morning, afternoon, and evening sampling periods, with the size of the pie proportional to the total VOC. Select relatively inert compounds (e.g., ethane and acetylene) as well as reactive species to help identify sources. Determine the extent to which total VOCs are accounted for as a function of location and time of day. Advance hypotheses regarding the compositions of the unidentified peaks and the unaccounted-for VOC. Test these hypotheses by examination of example chromatograms from source-oriented and receptor-oriented sampling sites and from direct samples of source emissions. Draw conclusions regarding the effects of these unidentified and unknown fractions on chemical mechanisms used in air quality models.

Plot pie charts of the gaseous and particulate nitrogen (NO, NO₂, HNO₃, PAN, HONO, and NO₃-), and carbon (heavy and light VOCs, aldehydes, and organic aerosol carbon) as a function of location. Make the radius of each pie proportional to the total number of N, S, or C atoms, and make each wedge proportional to the number of N, S, or C atoms contributed by each species. Examine the distributions of these species among gaseous and particulate phases as a function of time and location.

Reconcile Spatial, Temporal, and Chemical Variations in Ozone, Precursor, and End-Product Concentrations with Expectations from Chemical Theory

Calculate pseudo-steady-state values of ozone using NO/NO₂ ratios (from maps generated in work element 2.2.1) and assumed photolysis rates. Compare the spatial distribution of these calculated values with the measured values plotted in earlier data analysis activities. Explain possible reasons for differences. Compare NO₂ plots to PAN plots to determine when NO₂ is less than PAN.

Calculate average VOC-OH reactivity and determine VOC vs. NO_x limitations. Estimate (from literature) typical VOC-OH rate constants for individual VOC species. Calculate an average VOC reactivity by weighting each rate constant by the proportion of total VOC represented by its corresponding species. Compare the resulting average rate constants at different types of sites (urban, rural, oilfield) and sampling times (morning, day, night) with the VOC-OH reactivity of a standard urban VOC mixture (~3500 ppm-1min⁻¹). Stratify data near elevated point sources to separate periods when plumes are above the mixed layer from periods for which elevated plumes are within the mixed layer. Note similarities and differences with respect to site-type and time of day. Calculate adjusted VOC/NO_x ratios by multiplying a reference value of VOC/NO_x (for typical urban areas) by the ratio of the adjusted reference rate constant. Relate the results to bands of demarcation on an EKMA-type diagram between regions of VOC, combined, and NO_x limitation in producing ozone. Use the results to classify areas of the study domain into ones for which the ozone-forming potential is limited by VOC or NO_x.

Simulate ozone-producing potential of ambient VOC samples using an accepted chemical mechanism (e.g., CBM-IV, RADM, SAPRC) in a simple box using a range of diurnal radiation profiles from the radiometers. Artificially add NO_x concentrations and calculate time to formation of peak ozone and the amount of ambient carbon carried over or available for reaction the next day for an assumed radiation intensity and time. Define an "ozone-forming-potential" figure of merit and map these as a function of location, time, and VOC level. Estimate the effects of VOC, radiation, and other measurement uncertainties on ozone-forming potential. Identify differences in potential owing to changes in the model mechanisms (e.g., compare results from CBM-IV with SAPRC and RADM).

Stratify episodes by high and low photochemical potential days, and compare photochemical products along trajectories estimated by other data analyses. Further stratify these episodes by VOC/NO_x ratios in the western and northern parts of the study domain and estimate the extent to which this ratio affects the maximum ozone levels in the eastern and southern portions of the study domain. Performing the same analyses as above for ozone levels in central California air basins. Recalculate VOC/NO_x ratios for specific reactivity classes, especially aromatics, and examine receptor area ozone concentrations for cases of high and low

ratios in the source areas. Examine the emission maps to compare the quantity of fresh hydrocarbon and NO_x injected along the trajectories, and determine the degree to which this might interfere with the conclusions drawn from the VOC/NO_x ratios in the source areas.

Examine aircraft traverses and compare ozone, NO_x, and VOC levels to simple equilibrium calculations. Examine nighttime concentrations of O₃ and precursors above the mixed layer and off the coastline to determine the degree of carryover from the previous day. Compare VOC/NO_x ratios above the mixed layer with those calculated from localized emissions grid squares near the northern and western boundaries of the study domain. Identify potential causes of discrepancies, if they are found.

Compare day-to-day changes in emission patterns (using day-specific inventories) with O₃ and VOC concentrations for otherwise similar meteorological conditions. Compare oxidant values in cases with high ambient aromatic VOC concentrations to values obtained when aromatic VOCs are low.

Apply and Evaluate Ozone Receptor Models to Determine VOC and NO_x Limitations

Several receptor-oriented models have been developed and applied to determining relationships between oxides of nitrogen, VOC, and ozone levels. In each case, evaluate and determine the applicability of these models to CCOS by evaluating their fundamental assumptions as part of applying them to appropriate measurements from the database.

Compute correlations and regression relationships between ozone, NO_x and NO_y for measurement locations with fresh and aged emissions. Determine the extent to which ozone increases in photochemically aged air when NO_y is less than 1 ppb and when it is higher than 10 ppb (e.g., Trainer et al., 1993; Jacob et al., 1995). Calculate Integrated Empirical Rates (e.g., Johnson, 1984, Blanchard et al., 1994; Chang et al., 1995) to determine at what downwind distances from major source regions NO_x reductions or VOC reductions would most affect ozone concentrations. Compare these distances with those determined by other methods. Examine ratios among ozone and nitrogenous species (e.g., Sillman et al., 1990; Sillman, 1995; Milford et al., 1994; Jacob et al., 1995; Watkins et al., 1995) to estimate when VOC and NO_x limitations might apply. Apply the observation-based model (OBM) of Cardelino and Chameides (1995).

3.3.7 Characterize Episodes

Each of the episodes of two- to four-day duration has similarities and differences with respect to emissions, meteorology, transformation, deposition, and air quality levels. These episodes may be high for similar or for different reasons. Information derived the proceeding activities is synthesized to provide an anatomy of each episode. Conclusions are drawn with respect to which episodes are, for all practical purposes, the same, and which ones are substantially different.

Describe Each Intensive Episode in Terms of Emissions, Meteorology, and Air Quality

Prepare written overviews of each intensive sampling period. Describe the synoptic meteorology leading up the episode and summarize the forecasting rationale. Illustrate, with

plots generated in other work elements, the general wind flows for the duration of the period and any deviations from these generalizations. Identify major emissions events, identified as significant in other work elements, which affected pollutant concentrations. Summarize the key pollutant concentrations at key times and key locations in the study domain. Summarize the completeness and validity of the data set from each episode with respect to modeling of ozone. Identify the transport and transformation mechanisms that are likely to be dominant in each episode. Evaluate each episode for its potential use in model testing and control strategy evaluation.

Determine the Degree to which Each Intensive Episode is a Valid Representation of Commonly Occurring Conditions and its Suitability for Control Strategy Development

Examine continuous meteorological and air quality data acquired for the entire study period, and determine the frequency of occurrence of days which have transport and transformation potential similar to those of the intensive study days. Generalize this frequency to previous years, using existing information for those years.

3.3.8 Observation-Driven Methods

Observation-based approaches use high-quality ambient measurements of O₃, its precursors, and/or secondary products of the photochemical mechanism to diagnose the underlying relationship between O₃ production and sources of O₃ precursors. The methods range from those that rely solely on analysis of chemical measurements to more complex methods that rely on photochemical models as well as observations to diagnose VOC- or NO_x-limitation.

The following are four such methods.

1. The correlation between ozone and NO_y or NO_z (e.g., Trainer et al., 1993; Jacob et al., 1995)
2. The Integrated Empirical Rate (IER) model of Johnson (1984), as revised by Blanchard et al., (1994) and Chang et al., (1995). This method is also known as Smog Production (SP) algorithms.
3. The use of indicator species and ratios (e.g., NO_y, NO_z, O₃/NO_y, O₃/HNO₃, HCHO/NO_y, H₂O₂/HNO₃, H₂O₂/NO_y, and H₂O₂/NO_z) (Sillman et al., 1990; Milford et al., 1994; Sillman, 1995; Watkins et al., 1995; Jacob et al., 1995), and
4. The observation-based model (OBM) of Cardelino and Chameides (1995).

Observation-based methods provide procedures for deriving O₃ precursor relationships that are independent of emission inventories and other inherently uncertain inputs. However, these methods are not without their own limitations and uncertainties, and are best used in conjunction with air quality simulation models.

3.3.9 Contribution of Transported Pollutants to Ozone Violations in Downwind Areas

Although past transport studies have documented pollutant transport on specific days, they have not always quantified the contribution of transported pollutants to ozone violations in the downwind area. Quantitative estimates of the contribution of transported pollutants to ozone violations in the downwind area can be accomplished by photochemical grid modeling and by advanced data analysis techniques such as “flux planes” measured by aircraft which traverse a vertical plane perpendicular to a suspected transport corridor at different elevations.

In principle, well-performing grid models have the ability to quantify transport contributions. However, many of the interbasin transport problems involve complex flow patterns with strong terrain influences that are difficult and expensive to model. Upper-air meteorological and air quality data in critical transport locations is generally required in order to properly evaluate and use grid models for quantifying transport contributions. In combination with modeling, data analyses can improve the evaluation of modeling results and provide additional quantification of transport contributions.

In order to quantify pollutant transport and to provide data for modeling and data analyses, surface and aloft measurements are needed at locations where transport can occur and at the times when transport is occurring. These monitoring locations include in and near mountain passes, along coastlines, offshore, and at various locations in the downwind air basin.

Previous studies (e.g., Roberts et al., 1993) have used aircraft measurements to calculate transport across flux planes. Vertical planes, intersecting the profiler sites downwind of and perpendicular to the transport path, can be defined and provide estimates of transport through these passes using surface and aircraft measurements of pollutant concentrations and surface and wind profiler data for volume flux estimations.

3.3.10 Contributions of Elevated NO_x Sources to Downwind O₃

Power plants and other sources with tall stacks are significant sources of NO_x, which in the presence of NMHC can lead to catalytic formation of ozone downwind of the source. However, close to the stack there is a temporary decrease in ozone levels due to “titration” by high levels of NO in the near field of the plume. Further downstream, ozone levels above the local background indicate net ozone production due to the reaction of plume NO_x with NMHC that are entrained into the plume in the dilution process. However, questions remain as to how much ozone is actually produced in the plume, how the ozone production efficiency depends on the chemical composition of the plume, and what the relative contributions of power plants are to high ozone episodes ozone in downwind areas.

Models use plume-in-grid to represent the chemistry and dispersion of large elevated point source plumes. Typically the individual sub-grid scale plumes are simulated in a Lagrangian mode. The plumes are assumed to have a Gaussian distribution that can be treated analytically and these plumes may extend 10 to 20 km from the point source. The chemistry is often simplified to account only for the high NO_x concentrations found within the plumes and the effect of NO_x on ozone concentrations while ignoring VOC chemistry. The plumes disperse and undergo chemical reaction until the spatial extent of the plume, and the pollutant concentrations

reach levels that can be adequately represented within a 3-d model grid. Criteria for terminating the plume, i.e., mixing its contents into the regular grid, include plume size relative to the size of the grid cell and the age of the plume.

It is not clear that the treatment of plumes by state-of-the-art models is adequate. Vertical transport (e.g., plume rise and fall the plume during downwind transport) may not be adequately described. Recent power plant plumes studies (Senff et al., 1998) utilized airborne ozone and aerosol lidar in conjunction with other instrumented aircraft. Because of its ability to characterize the two dimensional structure of ozone and aerosols below the aircraft, the airborne lidar is well suited to document the evolution of the size and shape of the power plant plume as well as its impact on ozone concentration levels as the plume is advected downwind. This aircraft was considered as a study option, but budget constraints will prevent its use during CCOS. However, aircraft measurements of NO_x, ozone and VOC concentrations made in plumes are required to test the validity of the treatment of plume dispersion and chemistry and the procedures for terminating the plume into the regular model grid by plume-in-grid parameterizations.

3.3.11 Deposition Studies

During the California Ozone Deposition Experiment (CODE) in 1991, aircraft and tower-based flux measurements were taken over different types of San Joaquin crops, irrigated and non-irrigated fields, and over dry grass. Results are briefly summarized in Pun et al. (1998) and include estimates of ozone deposition velocities of 0.7-1.0 cm/s (Pederson et al. 1995). Vertical fluxes (deposition rates) can be calculated if a vertical gradient is known (assumed or measured). Order of magnitude calculations by Pun et al. show that dry deposition can be a few percent (~3-5%) of the total ozone budget in the San Joaquin Valley.

Reasons for further flux and deposition measurements during the CCOS study, with its expanded geographic scope, are at least three-fold:

1. Further consideration of NO sources. While NO was not considered to be a problem at three sites reported on by Pederson et al. (1995), Mahrt et al. (1995) found that rapid titration of O₃ by NO did affect aircraft-based flux O₃ surface flux measurements. Controlled experiments that employ a spatially diverse array of NO_x monitors upwind or surrounding a flux measurement site could help quantify this effect.
2. Further consideration of relative humidity effects. The Sacramento River Delta region and coastal regions are part of the CCOS domain. Ozone is not highly water soluble, but McLaughlin and Taylor (1981) report that ozone deposition to plants can increase by a factor of 2-3 when relative humidity changes from 35% to 75%. Based on this, Mahrt et al. (1995) suggest that ozone deposition may show a more complex spatial and/or temporal pattern than heat and moisture fluxes. Deposition studies for CCOS could include measurement of ozone fluxes over more varied types of terrain within the CCOS study area, including the higher humidity areas of delta or the coastal areas. Desjardins et al. (1995) report that aircraft flux measurements compared well with tower-based flux measurement at two instrumented vineyard sites during CODE. Aircraft could be used to expand the diversity of sites measured.

3. Precursor deposition rates. Fluxes of NO_x and VOC were not measured during CODE.

Specify the equations, assumptions, input data, and uncertainties for the deposition model. From the examination of micrometeorological data and vertical flux measurements, determine the extent to which these equations represent reality, and the degree to which assumptions are complied with. Evaluate the effects of input data uncertainties on deposition estimates.

3.3.12 Reformulate the Conceptual Model

The conceptual model described in Section 2 must be revisited and refined using the results yielded by the foregoing data analyses. New phenomena, if they are observed, must be conceptualized so that a mathematical model to describe them may be formulated and tested. The formulation, assumptions, and parameters in mathematical modules which will be included in the integrated air quality model must be examined with respect to their consistency with reality.

3.3.12.1 Refine Conceptual Models of Pollutant Emissions

Specify motor vehicle emission model equations, assumptions, input data, and uncertainties. Reconcile the ambient species ratios found in ambient data (as studied in prior work elements) with ratios determined from emissions models in terms of model or measurement biases. By stratifying samples, estimate the effects of different meteorological variables on emission rates. Pay special attention to the validity of models regarding vehicle age, maintenance, effects of hot and cold operating conditions, and vehicle-type distributions. Recommend improvements to emission models based on these observations.

Specify biogenic emission model assumptions, input data, and uncertainties. Reconcile ambient hydrocarbon species ratios at sites located in agricultural and forested areas with those determined from emission models in terms of model or measurement biases. By stratifying samples, estimate the effects of different meteorological variables, especially wind speed, on emission rates. Examine chemical speciation as a function of vegetation type. Examine total ammonia concentrations as a function of nearby soil types and fertilizer applications. Recommend improvements to emission models based on these observations.

Specify oilfield and refinery emission model assumptions, input data, and uncertainties for cogeneration systems, diesel-powered internal combustion engines associated with pumps, natural gas plants, drilling rigs, remote operations, spills, leaks from valves and flanges, evaporation from storage tanks, cyclic and non-cyclic well head vents, sumps, measuring stations, evaporation from tanker trucks and loading racks, pumping stations, and vacuum trucks, and gasoline stations. Evaluate how well these emissions estimates relate to reality, and which variables are not included in current methodology.

Reconcile ambient hydrocarbon species ratios at sites located near sewage plants and feedlots with those determined from emission models in terms of model or measurement biases. By stratifying samples, estimate the effects of different meteorological variables, especially temperature and relative humidity, on emission rates. Recommend improvements to emission models based on these observations.

Intermittent events include fires, entertainment and sporting events, and industrial upsets. Specify the models that treat intermittent events, their assumptions, and input data. Examine the magnitude of emissions from intermittent events with respect to other emissions to determine whether or not these emissions are significant. Recommend improvements to emission models based on these observations.

Examine the variability in emissions for intensive analysis days. Compare this variability to that assumed by point source models. Recommend improvements in emission models based on these observations.

3.3.12.2 Refine Models of Pollutant Transport and Dispersion

Thermal lows are caused by surface heating over the Mojave Desert and the Central Valley and are one of the major causes of flow between the Bay Area and the central coast to the Central Valley. Specify meteorological model assumptions and input data relevant to calculating the vertical profile of this transport pathway. For unidirectional flows, the mathematical formulation should show a minimum in turbulence or laminar flow occurring at the height of the wind maximum, and the width of the jet should be adequately estimated. The model formulation should allow this region of minimal turbulence to intensify the nighttime inversion in the vicinity of the flow and to inhibit vertical transport. Reconcile the model formulation with the location, dimensions, intensity, duration, and frequency of the thermal low transport. Quantify and compare measurement and model uncertainties.

Upslope flows in the Sierra Nevada result from heating of the mountain sidewalls by the afternoon. The downslope flows commence after sunset when the slopes cool. When the intensity of these up-slope winds is large, pollutants are vented from the Central Valley into the air above the mixed layer, and possibly into neighboring air basins. Examine the assumptions and input data of the meteorological model, which relate to slope flows. Determine whether or not the intensity and timing of these flows corresponds to those observed in other data analysis work elements. Identify those cases in which upslope flows vent pollutants above the mixed layer or over the summit, and determine the extent to which the mathematical formulation can represent these cases. Identify areas of uncertainty in the modeling and measurement processes and attempt to quantify these uncertainties.

Marine airflows mix, age, and transport pollutants from coastal air basin to the Central Valley. These airflows develop primarily from strong coast-to-inland pressure gradients. The most visible artifact of these flows is the marine stratus that forms along the coastal areas. A stronger inversion, a feature of the subtropical high, is usually present above the marine stratus layer and may extend to heights of 100 to 1000 m ASL, sufficiently deep to extend over the coastal mountains. Examine the results of windflow analysis to determine where and when this transport phenomenon occurs. Determine the presence or absence of fogs and high humidity at night, with the intent of understanding whether or not NO_x transported off the coast can be rapidly transformed to particulate nitrate that could be advected on-shore during the next day. Examine meteorological model formulation for those features that will describe these flows. Identify uncertainty in the modeling and measurement processes and attempt to quantify these uncertainties. Compare different pollutants with diurnal/spatial variations in humidity, visibility,

cloud cover, and solar radiation for the same air mass. Evaluate episodes using Leipper Inversion Based Statistics (LIBS) discussed in Section 2.6.

Identify the occurrence and thickness of different atmospheric layers from other data analysis work elements. Determine the modeled layer structure. Examine the extent to which layers can be assumed to be uniform, or must vary in depth as a function of location and time. Evaluate the uncertainty introduced to the modeling process by anticipated deviations from layering assumptions.

Table 3.1-1. Emission Inventory Roles and Responsibilities

Agency	Program Responsibilities						
	Stationary Sources	Area Sources	Area Sources - Agricultural and Controlled Burns	On-road Mobile Sources	Non-Anthropogenic Sources – Wildland Fires	Non-Anthropogenic Sources – Biogenic	QC & QA Checks
Air Resources Board		✓		✓	✓	✓	✓
Local Air Districts	✓	✓	✓			✓	✓
Councils of Government				✓			
Contractor	✓	✓	✓	✓		✓	
State and Federal Agencies			✓	✓	✓		

Table 3.1-2. Required meteorological files.

Mixing depth field
Three dimensional wind field
Surface temperature field
Absolute humidity field
Solar radiation scaling field
Ultraviolet radiation scaling field

Table 3.1-3. Typical chemical species in an emission inventory for modeling.

1.	Nitrogen dioxide
2.	Nitric oxide
3.	Sulfur dioxide
4.	Carbon monoxide
5.	Methane
6.	Ethane
7.	Higher alkanes
8.	Ethene
9.	Terminal alkenes
10.	Internal alkenes
11.	Isoprene
12.	Benzene
13.	Toluene
14.	Xylene and more reactive aromatics
15.	Carbonyls
16.	Formaldehyde
17.	Acetaldehyde and higher aldehydes
18.	Ketones

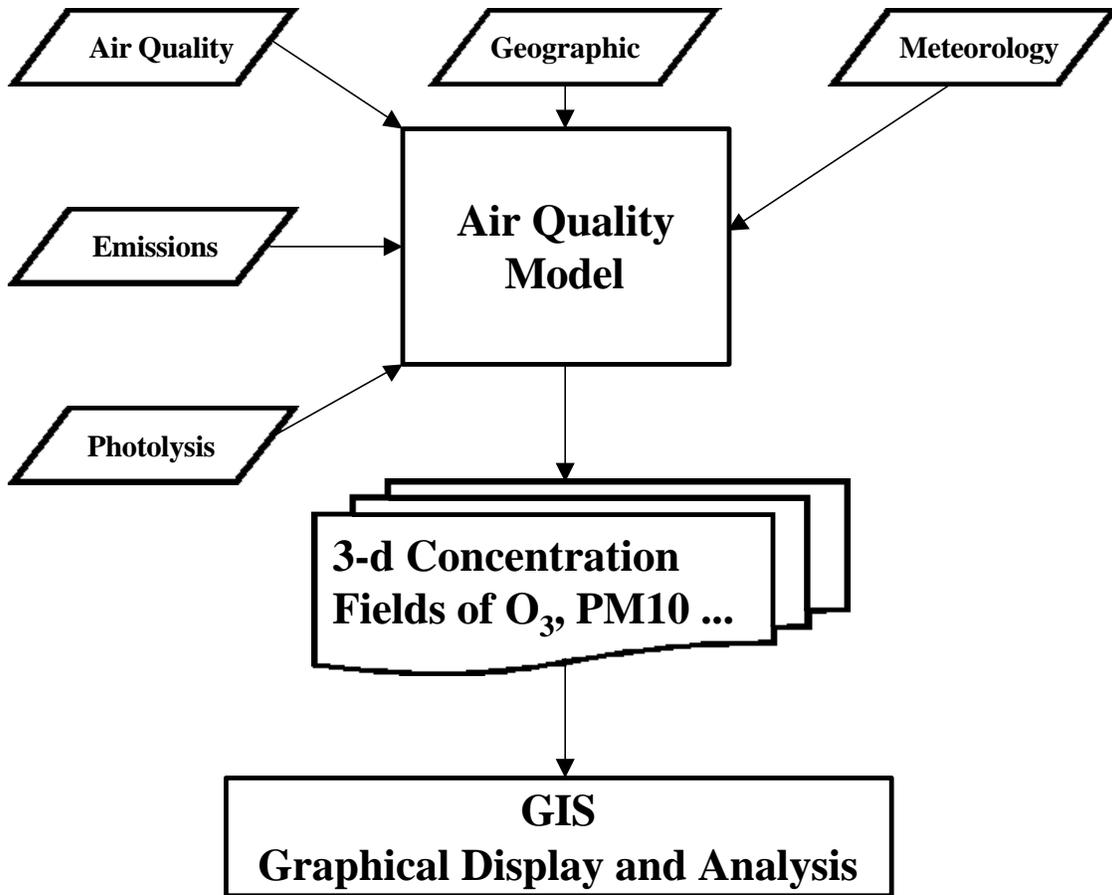


Figure 3.1-1. File typically required by photochemical air quality model.

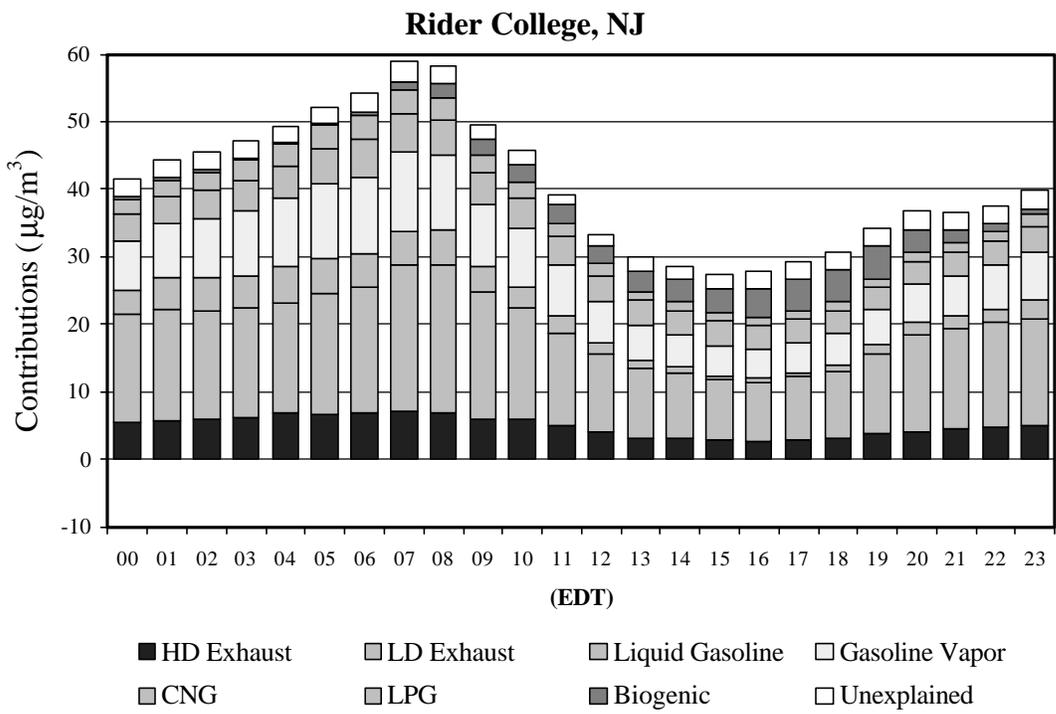


Figure 3.2-1. Average source contribution estimates of ambient hydrocarbons at Rider College, NJ during summer, 1995 by time of day. Source: Fujita and Lu, 1998.

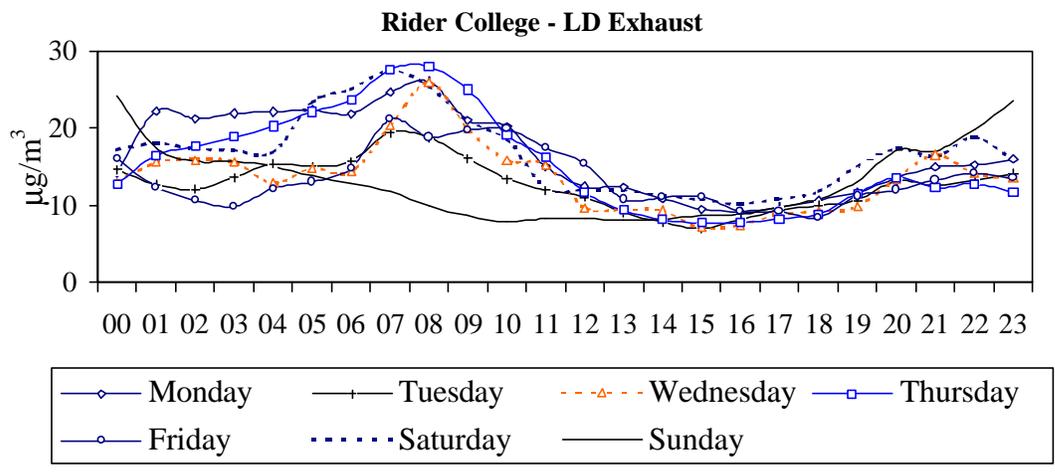


Figure 3.2-2. Average source contribution estimates of hydrocarbons at Rider College, NJ during summer, 1995 (EDT) by day of the week. Source: Fujita and Lu, 1998.

