United States
Environmental Protection
Agency

Air and Radiation

EPA420-R-01-007
April 2001

EPA's New Generation Mobile Source Emissions Model: Initial Proposal and Issues

Printed on Recycled Paper

EPA420-R-01-007
April 2001

EPA's New Generation Mobile Source Emissions Model: Initial Proposal and Issues

EPA Mobile Source Emissions Modeling Workgroup
Office of Air and Radiation
Office of Research and Development
Region 4

Table of Contents

I. Executive Summary

The purpose of this paper is to present issues and options regarding the future direction of EPA's mobile source emissions modeling program, and a proposed framework for the agency's future modeling work. We are applying the term "New Generation Model" to this effort because we believe that fundamental changes are required in order to meet the expanding challenge of mobile source emission estimation in a way which is comprehensive, thorough and quality-based. The concepts presented in this report reflect initial thinking; no decisions have been made regarding model scope, structure, content, data, platform, etc. Rather, these concepts are meant to elicit comment from users of EPA's mobile source emissions tools and stakeholders of the mobile source emissions estimation process.

Under the Clean Air Act, EPA is charged with developing emission factors for all emission sources. EPA's Office of Transportation and Air Quality (OTAQ) has been the source for emission factor development for on-road sources such as light and heavy-duty vehicles and trucks, and off-road sources such as construction and agricultural equipment. This has led to the development of a number of emission factor estimation tools such as MOBILE (for on-road VOC, CO and NOx), PART (on-road particulate matter and SOx), MOBTOX (on-road toxics), and NONROAD (all off-road pollutants). These tools have been focused on the estimation of mobile source emissions based on average operating characteristics over broad geographical areas. In recent years, however, analysis needs have expanded in response to statutory requirements that demand the development of finer-scale modeling approaches to support more localized emission assessments. The growing needs of model users and external recommendations from a variety of sources have indicated the need for more emission research and improved modeling methodologies.

A comprehensive review of EPA's mobile source modeling program was published by the National Research Council (NRC) in May 2000. It recommended that EPA develop a mobile source emission modeling system that is capable of supporting the expanding range of mobile source emissions analyses. EPA is in the process of releasing the updated on-road emission factor model MOBILE6, which represents a substantial improvement from MOBILE5, particularly for finer-scale modeling. We view the New Generation Model as a logical next step in the continual effort to improve mobile source emissions models to keep pace with new analysis needs, new modeling approaches, and new data.

The NRC recommendations also address the need for improved model science and improved model structure, two key objectives of the New Generation Model. An improved modeling structure will allow better responsiveness to new data and enable model validation, which in turn will facilitate improved science. Improved science is also a direct function of the quality of information feeding the model. We believe that the recent emergence of on-board emissions measurement devices will revolutionize how emissions data are collected for on-road and off-road mobile sources. We envision that this technology will become the focus of EPA's emissions factor testing program, and will provide the opportunity for a significant shift in how emissions modeling is approached.

The primary drivers for mobile source emissions analyses are a) statutory requirements, b) support for studies on emission trends, air quality and cross-media impacts, and c) support for EPA regulatory efforts. Borrowing from a similar breakdown in the NRC report, we have identified four fundamental analyses which a mobile source modeling system will need to perform in response to these three drivers:

- Large Area (e.g. National) Emissions Inventory Generation
- Local Area Emissions Inventory Generation
- Transportation Scenario Evaluation
- Corridor/Intersection Emissions Analysis

In order to address this range of analytical needs, we are proposing that the New Generation Model allow for analysis at different scales, depending on the desired application. The system as proposed would estimate the emissions inventory from the national level to the corridor level, for off-road and on-road sources, for all pollutants. In order to address this range of analyses, the system would employ three analysis scales termed macroscale, mesoscale and microscale, defined as follows:

- **Macroscale** analyses are appropriate for developing large-scale (e.g. national) inventories, and will likely continue to be the default choice for generating local inventories for use in SIP and conformity planning. The basic spatial unit for this scale would be the county. As envisioned, the macroscale level would be consistent in concept with the current applications of MOBILE (with inventory generation capability) and NONROAD.

- **Mesoscale** analyses are geared towards generating local inventories at a finer level of spatial and temporal resolution. The basic spatial unit for this scale would be the roadway link and traffic analysis zone, consistent with output from standard travel demand models. Three options are being proposed for the mesoscale level: “basic”, “modal”, and “advanced”, which incorporate increasing levels of resolution in vehicle activity and spatial characteristics.

- **Microscale** analyses allow the estimation of emissions for specific corridors and/or intersections, which is appropriate for assessing the impact of transportation scenarios and performing project-level analyses. As proposed, this scale would rely on modal emission rates from the mesoscale level, in conjunction with localized activity information.

The proposed modeling system would combine these three basic levels of analysis in a way that provides for consistency in the emission component between scales. The user would choose which scale to implement depending on the desired application and the availability of necessary input data. Clear guidance will be critical for determining the appropriate analysis level and setting the standard for adequate input data at a given level.

Ultimately, our goal is to use on-board emissions data as the basis for the New Generation Model.   The core of the proposed system would be an emission rate estimator, which processes instantaneous on-board emissions data into the modal and macroscale emissions rates used in the three analysis scales.   We have identified three fundamental approaches for developing an emission rate estimator which serves this function: a) develop a microscale emissions model; b) process instantaneous emissions measurements produced in a laboratory or in the field; or c) create a direct link between a database of raw instantaneous emissions measurements and the New Generation Model**.**

Scope considerations are an important aspect of the planning process for the New Generation Model. A reduced scope option is presented which focuses on the development of emission rates and software support for the macroscale level, while providing only guidance for developing microscale and mesoscale models. Other scope considerations are whether to pursue an interim product which focuses on improving the software structure, and whether the on-road component of the New Generation Model should be pursued before the off-road component.

The next step for the New Generation Model is the development of a comprehensive plan, slated for the Fall 2001. The main purpose of the comprehensive plan would be to provide concrete steps for the development of the New Generation Model and allow for more detailed determination of resource needs, data needs, and timing. In general, we see three major issues which need to be addressed as we work towards this plan: 1) further definition of the modeling system, including the underlying model theory and the input/outputs of each modeling component; 2) establishing a methodology for estimating emission rates, including an assessment of how on-board emissions would be analyzed and a sampling plan for populating the model with on-board emissions data; and 3) developing a software design.

II. Background on the New Generation Modeling System

A. Introduction

The purpose of this paper is to present issues and options regarding the future direction of EPA's mobile source emissions modeling program, and a proposed framework for the agency's future modeling work. We are applying the term "New Generation Model" (NGM) to this effort because we believe that fundamental changes are required in order to meet the expanding challenge of mobile source emission estimation in a way which is comprehensive, thorough and quality-based. This paper is intended to begin the dialogue between EPA and modeling stakeholders regarding the direction of the Agency's mobile source emissions modeling program. Ideas and options are presented in this paper in the spirit of eliciting stakeholder input.

It is important to reinforce that everything presented in this paper reflects initial thinking. No decisions have been made regarding model scope, structure, content, data, platform, etc. Rather, we are seeking input from modelers, users and stakeholders on these issues. The range of purposes the New Generation Model must serve will likely require more than a single piece of software. We therefore consider the New Generation Model as a modeling system, which includes underlying data structures, software platform, guidance, documentation, and model validation; we are interested in input on all of these issues.

The final version of MOBILE6 will be released in Summer 2001, and represents a significant improvement over MOBILE5. The New Generation Model is envisioned as eventually replacing both the MOBILE6 and NONROAD models. However, the potential magnitude of the NGM effort dictates that these models will continue to be our best tools for estimating mobile source emissions for some time to come. The need for model stability is an important factor in the timing and approach of the New Generation Model. As discussed in Section IV, we are considering an interim option for the New Generation Model which would focus on scope and software enhancements while maintaining the emission predictions of MOBILE6 and NONROAD as the core.

Comments, input and ideas are encouraged through writing or email. This is not a formal comment process, and hence there is no strict deadline. However, in order to ensure consideration of comments as we move forward in the planning process, it would be helpful to have comments by June 15, 2001. Written comments should be forwarded to John Koupal, Assessment and Standards Division, U.S. EPA Office of Transportation and Air Quality, 2000 Traverwood Drive, Ann Arbor, MI, 48105. Email should be sent to the New Generation Model email address, newgen@epa.gov.

B. The Need For a New Modeling Approach

Under the Clean Air Act, EPA is charged with developing emission factors for all emission sources. EPA's Office or Transportation and Air Quality (OTAQ) has been the source for emission factor development for on-road sources such as light and heavy-duty vehicles and trucks, and off-road sources such as construction and agricultural equipment. This has led to the development of a number of emission factor estimation tools such as MOBILE (for on-road VOC, CO and NOx), PART (on-road particulate matter and SOx), MOBTOX (on-road toxics), and NONROAD (all off-road pollutants). Information on these models is readily available through OTAQ's website (http://www.epa.gov/otaq).

EPA's mobile source emission estimation tools and underlying emission factors have been focused on the estimation of mobile source emissions based on average operating characteristics over broad geographical areas. Examples of this scale of analysis are the development of SIP inventories for urban nonattainment areas and the estimation of nationwide emissions to assess overall trends. In recent years, however, analysis needs have expanded in response to statutory requirements that demand the development of finer-scale modeling approaches to support more localized emission assessments. Examples include "hot-spot" analyses for transportation conformity, and the evaluation of the impact of specific changes in a transportation system (e.g. signalization, lane additions, etc.) on emissions. In response to the acknowledged shortcomings of the MOBILE5 model in addressing these modeling needs, separate modeling initiatives have been undertaken to develop tools which provide a better assessment of finer scale emissions. Three notable efforts are the Comprehensive Modal Emissions Model (CMEM) developed by UC Riverside under NCHRP Project 25-11; TRANSIMS, under development by Los Alamos National Laboratory through the U.S. Department of Transportation; and MEASURE, developed by Georgia Tech under cooperative agreement with EPA's Office of Research and Development.

The growing needs of model users and external recommendations from a variety of sources have indicated the need for more emission research and improved modeling methodologies. Recent reviews of EPA's mobile source modeling program have pointed to the limitations of MOBILE5 for finer-scale analyses. The most comprehensive of these reviews, entitled "Modeling Mobile Source Emissions," was published by the National Research Council (NRC) in May of 2000,[1] hereafter referred to as the NRC report. The NRC report recommended that EPA develop a mobile source emission modeling system that is capable of supporting the expanding range of mobile source emissions analyses and is based on more rigorous science and more extensive real-world data. Specific recommendations of this report are as follows:

- Develop a "toolkit" of models which allow for the prediction of emissions over a broad range of spatial and temporal scales
- Improve model evaluation, including sensitivity and uncertainty analysis
- Improve long-term planning and coordination
- Improve characterization of "real world" in-use emissions, particularly for high-emitters, heavy-duty vehicles and for PM and toxics
- Update the model more frequently
- Improve off-road emissions estimates

EPA is in the process of releasing the updated on-road emission factor model MOBILE6, which represents a substantial improvement from MOBILE5, particularly for finer-scale modeling. For example, MOBILE6 will allow a more disaggregated approach to emission estimation through the use of roadway facility-based emission factors, and the separation of start and running emissions. Finer levels of temporal disaggregation are possible through the modeling of emissions on an hourly basis, and the level of vehicle activity information which the model will accept has been greatly increased. Updates are currently planned for MOBILE6 to incorporate particulate matter, air toxics and greenhouse gases. As a result of these updates, we consider MOBILE6 to be a significant step towards addressing some of the acknowledged shortcomings of MOBILE5, as well as the NRC recommendations. We view the New Generation Model as a logical next step in the continual effort to improve mobile source emissions models to keep pace with new analysis needs, new modeling approaches, and new data.

The first NRC recommendation addresses the need for increased model scope, with a focus on modeling finer levels of resolution. Beyond this issue, we feel that a new modeling approach is needed in order to address the broadening range of pollutants and mobile sources in an integrated way. Increased interest in emissions of air toxics and greenhouse gases from mobile sources requires models which estimate these pollutants. A modeling system which can estimate all pollutants of concern in an integrated fashion is required to allow for the evaluation of pollutant trade-offs, which are of increasing import in the mobile source arena. For example, an emerging issue is the tradeoff between particulate matter and carbon dioxide resulting from the introduction of diesel technology to improve fuel economy. MOBILE6 in its full form will satisfy this need by including all pollutants; we envision that the New Generation Model will expand this capability to all analysis scales.

Off-road sources will likely become a larger issue for mobile source emission estimation, and evaluating tradeoffs between on-road and off-road sources will become more necessary. Having a modeling tool which integrates these two sources will allow this type of analysis to be performed more efficiently. Tradeoffs between mobile sources and other sources (e.g. point, area) are also an emerging issue, so it is important to consider a modeling system which can integrate well with modeling approaches for non-mobile sources as well.

The NRC recommendations listed above also address the need for improved model science and improved model structure. We consider these issues to be significant drivers for pursuing a new modeling approach. The physical design and coding of the current models do not facilitate straightforward updates and revisions; the MOBILE model is based on core computer code developed in the 1970s. An improved modeling system will enable easier updates, and improved responsiveness to new data. In this way, improved model structure can facilitate improved science through the more direct incorporation of new data. Improved science is also a direct function of the quality of information feeding the model, and we believe that the recent emergence of on-board emissions measurement devices will revolutionize how emissions data are collected for on-road and off-road mobile sources. Several commercial applications of this technology have begun to enter the marketplace, or are under development. In addition, EPA is undertaking a major effort to advance the development of on-board emissions analysis equipment, termed Portable Emissions Measurement System (PEMS). PEMS will ultimately allow the gathering of instantaneous exhaust emissions data for HC, CO, NOx, particulate matter, toxics and greenhouse gases. It will also include a global positioning system (GPS) to allow linkage of emission measurements with the location and speed of the vehicle. Vehicle operating information will be monitored, and for vehicles equipped with on-board diagnostics systems (OBD), the OBD stream will provide engine and vehicle operation information. This system will enable the gathering of in-situ emissions across all mobile sources in several geographic locations, for relatively low cost. It is envisioned on-board emissions measurement technology will become the focus of EPA’s emissions factor testing program, and will provide the opportunity for a significant shift in how emissions modeling is approached.

Software advances also provide new opportunities for improving the mobile source emissions models. Geographic Information Systems (GIS) provide the opportunity to improve the spatial element of emissions characterization. Object-oriented design methods provide the opportunity for software to correspond more directly to the subject matter being modeled. These approaches provide new opportunities for model development which can greatly enhance the efficiency, performance and reporting capability of the modeling systems.

C. What a New Generation Model Should Address

An important step in determining the scope of the New Generation Model is to identify what analyses a mobile source modeling system must perform. The primary drivers for mobile source emissions analyses are a) statutory requirements, b) support for studies on emission trends, air quality and cross-media impacts, and c) support for EPA regulatory efforts. Borrowing from a similar breakdown in the NRC report, we have identified four fundamental analyses which a mobile source modeling system will need to perform in response to these three drivers:

- Large Area (e.g. National) Emissions Inventory Generation
- Local Area Emissions Inventory Generation
- Transportation Scenario Evaluation
- Corridor/Intersection Emissions Analysis

While these analyses have some overlapping requirements, each requires specific functionality from a modeling system. Table 1 elaborates example applications and desired model attributes for each analysis area.

Table 1: Required Applications for a New Generation Model System

Analysis Category	Example Applications	Desired Model Attributes
Large Area Inventory Generation	• EPA Trends Report • EPA Greenhouse Gas Inventory Sources and Sinks Report • Regulatory support	• Complete inventory development for analysis area • Compatibility with air quality models • Ability to produce outputs at different scales (e.g. county, state, nation) • Default information for activity and fleet characteristics • Flexibility for incorporating local inputs
Local Area Inventory Generation	• SIP inventory development • Rate of Progress analyses • Conformity analyses • Support for emissions trading schemes	• Complete inventory development for analysis • Compatibility with air quality models • Ability to assess policy alternatives (e.g. fuel specifications, transportation control measures) at all scales • Ability to produce outputs at different scales (e.g. grid, county, region) • Flexibility for incorporating local inputs
Transportation Scenario Evaluation	• SIP inventory development • Conformity analyses • Transportation Control Measure evaluation • Congestion Mitigation & Air Quality (CMAQ) Funding	• Ability to estimate the emission effects of changes in microscale vehicle activity • Ability to relate emission impacts to urban inventory
Corridor/Intersection Analyses	• Project-level environmental assessments (e.g. NEPA) • "Hot-spot" analyses for conformity • Exposure assessments	• Ability to estimate the emission effects of changes in microscale vehicle activity • Compatibility with dispersion and/or exposure models • Flexibility for incorporating localized fleet and activity inputs

III. Four Objectives for the Proposed New Generation Modeling System

Our design of the new generation modeling system is guided by four broad objectives: a) encompassing all pollutants and all mobile sources at the levels of resolution needed for the diverse applications of the system (comprehensiveness); b) using principles of sound science (improving the science); c) using a software design and development method which is efficient and flexible (improving the software); and d) implementing the model in a coordinated, clear and consistent manner (process and implementation).

A. Comprehensiveness

This section discusses the first of these broad objectives, comprehensiveness. The remaining objectives are discussed in subsequent sections.

1. Defining Analysis Scales

An important objective for the New Generation Model is to develop a modeling system which integrates modeling scales, mobile sources, and pollutants. A fundamental issue with the New Generation Model is the definition of the modeling scales appropriate to address the analysis needs detailed in the previous section. For this purpose we are using the definitions of three analysis scales presented in Chapter 6 of the NRC report: macroscale, mesoscale, and microscale. Our definitions of these scales have been refined somewhat from the NRC report, as follows:

- **Macroscale** refers to analysis over a broad regional area (e.g., county, urban area, state, nation), for which emissions are estimated using aggregated analysis techniques. MOBILE and PART would be considered macroscale emission factor models; emission factors are based on average trips meant to reflect average travel over a large region. These emission factors are combined with aggregate estimates of vehicle activity (average speed, vehicle-miles traveled) to generate an emissions inventory. Likewise, the NONROAD model would be considered a macroscale inventory estimator for off-road emissions. This level of analysis is most common for national/regional inventory generation, and is also used for SIP inventory and conformity budget determination.

- **Mesoscale** refers to analysis at more localized levels, allowing estimation of emissions on specific roadway links and/or analysis zones. Within the mesoscale level there can be several approaches, at varying levels of detail. One end of the spectrum is to estimate emissions at the roadway link and analysis zone using macroscale emission factors (e.g. from MOBILE) - an approach used by many state and local areas in developing SIP inventories and conformity budgets. The other end of the spectrum is to refine vehicle emissions to account for specific driving modes (modal operation), and to refine spatial and temporal vehicle activity and fleet characteristics based on inputs such as census data, land use, vehicle address location and temporal activity profiles. This latter approach is employed in advanced modeling systems such as the Georgia Tech MEASURE and the DOT TRANSIMS models.

- **Microscale** refers to the estimation of vehicle emissions for a specific corridor or intersection. Within microscale modeling there can be several levels of differentiation. EPA/ORD's MicroFac model, for example, combines vehicle activity information gathered on a per-vehicle basis with aggregate emission factors derived from MOBILE.[2] Other approaches take into account vehicle operating modes or instantaneous activity (second-by-second vehicle driving trajectories), such as UC Riverside's Comprehensive Modal Emissions Model (CMEM).[3] As we are defining it, the emissions estimation process may not be different between the mesoscale level and microscale level; the primary difference between these two levels is that the microscale estimates emissions for a specific corridor or intersection , while the mesoscale develops area-wide emissions inventories by aggregating finer levels of spatial and temporal resolution.

In order to support the emissions analyses detailed in Table 1, the New Generation Model system will need to include each of these levels. One level cannot support all of the necessary analytical needs. A primary criticism of MOBILE is that it is not appropriate for mesoscale and microscale analyses, because aggregate emission factors are not well suited for evaluating changes in vehicle micro-activity on a specific corridor. While MOBILE6 will significantly improve this capability relative to MOBILE5, it is still fundamentally an area-wide inventory tool which bases emissions on average driving patterns. Likewise, a microscale model is not an appropriate tool for generating a nationwide emissions inventory, in no small part because doing so would require having instantaneous vehicle activity on every road in the nation, a prohibitive computational, analytical, and data-gathering burden. Recognizing the need for a multi-faceted approach, the NRC report recommended an integrated "toolkit" comprised of macroscale, mesoscale and microscale modeling capability, with each level applied as appropriate for a given analysis. We concur with this recommendation.

2. Application of Analysis Scales to the Analysis Needs

Based on the above definitions of macroscale, mesoscale and microscale, the application of these scales to the required analyses outlined in the previous section are shown in Table 2. This table shows shaded areas for analysis scales at which the four basic analysis types would be performed, as follows:

- Macroscale analyses are appropriate for developing large-scale (e.g. national) inventories, and will likely continue to be the default choice for generating local inventories for use in SIP and conformity planning.

- Mesoscale analyses are geared towards generating local inventories at a finer level of spatial and temporal resolution. This level is highlighted for assessing transportation scenarios as well, because, although the evaluation of these measures should be performed at the microscale level, there needs to be a link with mesoscale analyses to assess the impact of potential measures on SIP inventories, conformity budgets and conformity analyses.

- Microscale analyses allow the estimation of emissions for specific corridors and/or intersections, which is appropriate for assessing the impact of transportation scenarios and performing project-level analyses.

Table 2 - Intersection of Analysis Scales and Analysis Needs

	Large Area Inventory	Local Inventory	Transportation Scenario Evaluation	Corridor/ Intersection Analysis
Macroscale				
Mesoscale				
Microscale				

3. Proposed Approach for Integrating Analysis Scales

We have developed a proposed framework for the New Generation Modeling system based on our assessment of the needs presented in Section II.B and consideration of the overall objectives for the model. A graphical representation of this framework is presented in Figure 1. On this figure, rectangles represent processing steps, and parallelograms represent data elements. Arrows, which in some cases are bi-directional, reflect that one component receives input from another. The solid vertical lines in Figure 1 represent the boundaries of the NGM system.

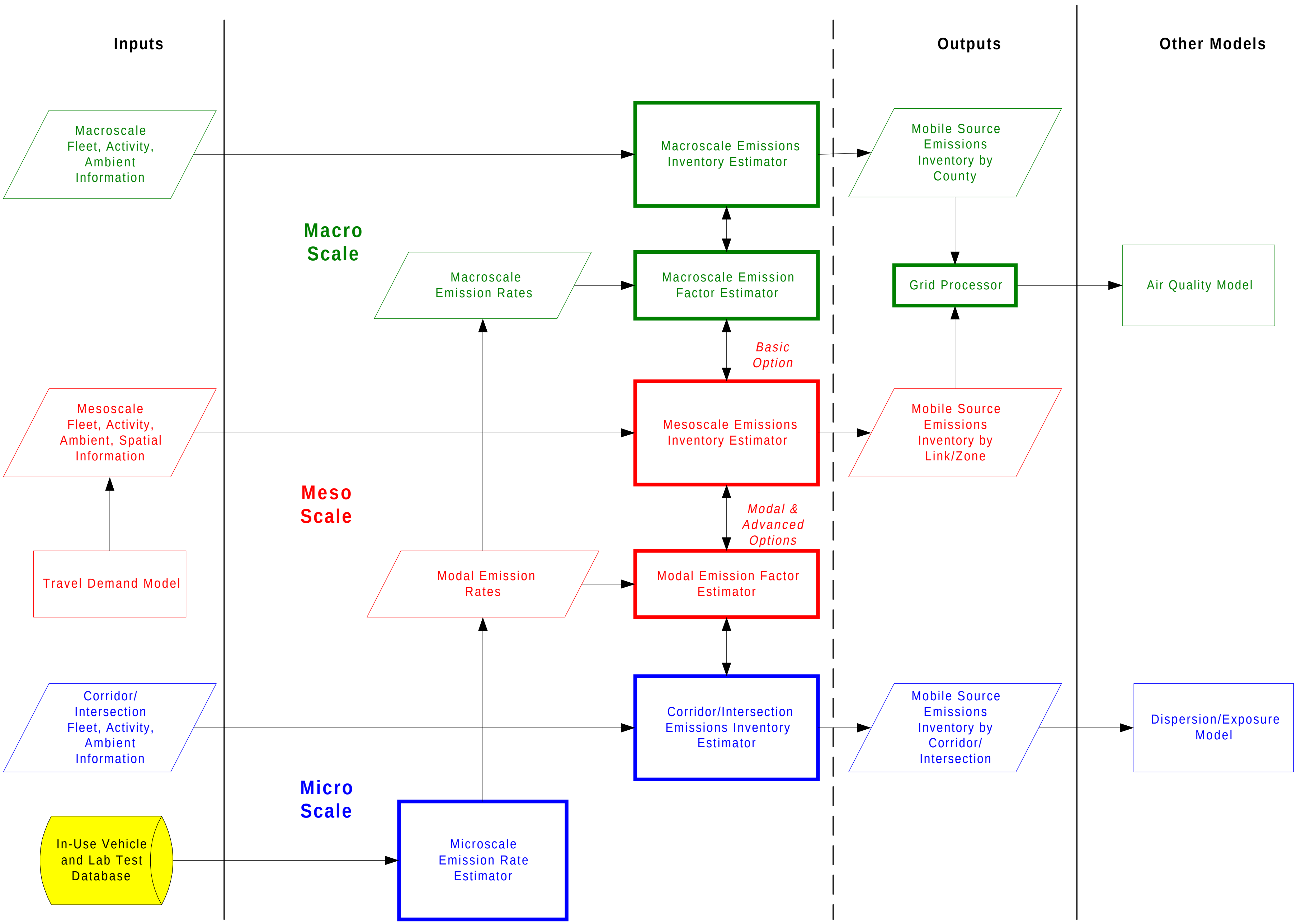
The proposed New Generation Model system combines the three basic levels of analysis - macroscale, mesoscale and microscale - in a way which allows for consistency in the emission component between scales. The scope of the framework is emissions inventory estimation from the national level to the corridor level, for off-road and on-road sources, for all pollutants. The user would choose which level to implement depending on the desired application and the availability of necessary input data. As discussed later in this paper, clear guidance will be critical for determining the appropriate analysis level and setting the standard for adequate input data at a given level. The following sections focus on the specifics of this multi-scale framework. Following these sections is a discussion of the inputs and outputs for the proposed modeling system.

a. Macroscale Level

The focus of the macroscale level is the generation of county-level emission inventories for areas up to the entire nation, using two basic components: a **Macroscale Emissions Factor Estimator** and a **Macroscale Emissions Inventory Estimator**. The emissions inventory estimator will calculate county-level emissions inventories for all pollutants and all sources based on activity, fleet, meteorological, fuel, emission control program and emission factor information. The emission factor information will come from the emissions factor estimator, which calculates aggregate emission factors at the level of detail (e.g. pollutant and vehicle/equipment class) required for the inventory analysis based on the activity, fleet, meteorological, fuel and control program information input into the inventory estimator. The inventory estimator could either execute the emissions factor estimator for each county (or set of unique county inputs), or use the emissions factor estimator to generate a look-up table of emissions factors (e.g. emissions factor by pollutant, vehicle/technology class, model year, speed, temperature etc.), which the inventory estimator would use for calculating emissions in each county. For on-road emissions, the function of these two components is consistent with the current application of MOBILE (as the emission factor estimator) and inventory post-processing applications (e.g. EMS-95 or SMOKE), which will either execute MOBILE repeatedly or use a MOBILE-generated look-up table to calculate emissions by county. For off-road emissions, the NONROAD model currently encompasses the functionality of both the emission factor and inventory estimator. As discussed below in the section addressing interface with air quality modeling (Section III.A.5.b), if the purpose of generating mobile source inventories is to combine them with other sources for air quality modeling, inventory estimation might be better handled by an external emissions processing application.

The macroscale level of the New Generation Model system would serve as the default level of operation, meaning that the system would include default information for activity and fleet, similar to MOBILE or NONROAD, to enable the generation of inventories with minimal user input. As is done with MOBILE, the user would be able to customize model inputs to reflect more accurate local information, following established guidance. As the macroscale level would serve low-end as well as high-end users, an objective of this level is that it would be platform independent. This is achieved with current macroscale emission factor models and inventory processing systems, and is considered a reasonable goal.

Figure 1 - Proposed Framework for New Generation Model System



b. Mesoscale Level

The focus of the mesoscale level is the generation of emission inventory results for a specific urban area, at a finer level of resolution than afforded by the macroscale level. The **Mesoscale Emissions Inventory Estimator** would be used in a similar fashion as the macroscale emissions inventory estimator, but would calculate emissions down to the roadway link and analysis zone; emissions factors would also be generated by an emissions factor estimator, although the aggregation of emissions factors will depend on the option used, as discussed below. The mesoscale level would be available to users who were able to generate the necessary input data, according to established guidance.

As discussed, many different levels of mesoscale analysis are currently performed by state and metropolitan planning organizations, according to the level of sophistication desired or access to appropriate input data and resources. In an attempt to capture the range of analyses currently performed at this level, we have defined three possible options available to the user interested in developing mesoscale emissions inventories. These options are focused on calculating on-road exhaust emissions. A discussion of evaporative and off-road emissions follows the discussion of options.

i. Basic Option

The objective of the "Basic" option is to provide "low-cost" mesoscale emission inventory capability, requiring the user to provide activity information from a traditional four-step travel demand model in order to generate emissions at the roadway link and travel analysis zone level. The function of this option would be to marry the output from travel demand models with aggregate emission estimates from macroscale emission factor models and county-level fleet, meteorology and control program information. In the current paradigm, this is similar to using MOBILE to generate emissions inventories at the link and travel analysis zone level, an approach used by many state and local areas in generating inventories for SIP and conformity purposes. Under this option, the **Mesoscale Emission Inventory Estimator** would calculate emissions for all pollutants at the roadway link and analysis zone level based on the input of speed and volume by roadway link, and trip starts per zone as calculated by a travel demand model. Fleet, meteorological, emission control program and emission factor information would be retained at the county level, and the **Macroscale Emissions Factor Estimator** would provide emission factors. By accessing the macroscale information, this option would allow default values to be used in the calculation of mesoscale inventories, making the mesoscale level more accessible to low-end users. Emissions at the link and zone level would be aggregated to a grid level for input into an air quality model by an external **Grid Processor**, a function currently served by existing applications such as SMOKE.

ii. Modal Option

Relative to the basic option, the modal option would add the capability to estimate emissions due to the change in vehicle operating modes, and would add road grade. This level would necessitate a shift from average cycle-based emission factors at the microscale to modal emission rates. This step will require modal emission rates generated by the **Modal Emission Factor Estimator**, which would generate emission factors by pollutant, vehicle class, etc. for specific operating modes (e.g. speed/load) instead of by average trip. Modal emission factors are required to better estimate the effects of changes in vehicle operating mode and/or grade, for which macroscale emissions rates are less appropriate. Using this option would require the development of a vehicle operating mode distribution at the roadway link level; default distributions (based the average driving cycles used in generating the macroscale emission factors) and road grade values (likely zero) would be available, but in order to make full use of this option the user would need to provide alternate operating mode distributions and/or road grade in addition to activity information from travel demand models.

iii. Advanced Option

Relative to the modal option, this option would add the capability to model activity and fleet information, and hence overall emissions, at finer spatial scales. This option would incorporate recent advances in emission modeling such as Georgia Tech's MEASURE model, which improves spatial resolution of vehicle activity using data sources such as census and land-use, and improves spatial resolution of fleet activity by assigning vehicles to their specific addresses.[4] At this level of analysis the need for the management of spatial data likely necessitates the use of GIS software packages, which may not be as essential for the basic and modal options. For this reason, we are considering two different approaches for supporting this option: a) include off-the-shelf GIS software in the New Generation Model system directly, or b) provide only guidance for implementing this approach, requiring the use of specific elements of the New Generation Model system within this option, such as the modal emission rates (this is discussed under Section IV.B, Reduced Scope Option). One consideration for this is that DOT's TRANSIMS model will likely provide a "high-end" alternative for modeling at the mesoscale and microscale levels. We are coordinating with DOT towards the longer-term goal of having TRANSIMS incorporate the modal emissions component of the New Generation Model. We will therefore have to consider whether supporting a separate "advanced" option for mesoscale modeling, particularly as a software application, would add value.

iv. Evaporative and Off-Road Emissions

A discussion of mesoscale for on-road evaporative emissions pertains primarily to the refinement of vehicle trip activity, rather than further refinement of emission modes. Evaporative emissions could be estimated at the mesoscale level through the information provided by trip origins and destinations from travel demand models. By combining this information with macroscale emission rates, a finer spatial and temporal resolution of evaporative emissions could be estimated. The "Basic" and "Modal" options would essentially be unchanged for evaporative emissions, but the "Advanced" option would provide improved resolution of these emissions through finer spatial allocation. For example, the approach proposed for TRANSIMS is to use the finer resolution of activity information available from the model, in conjunction with macroscale (MOBILE6) evaporative emission rates.

Similarly, the consideration of the mesoscale level for off-road emissions applies at this time primarily to the refinement of activity and population information. We have identified three approaches which could be used to generate off-road emissions at the mesoscale level: 1) starting with county or grid-level macroscale emissions, allocate down to finer spatial zones (e.g. land-use category, transportation analysis zone) based on surrogate information such as population or land-use type; 2) using localized activity and population studies to provide information directly for a specific area;. 3) using localized activity studies to develop relationships between equipment activity and surrogates (e.g. construction starts, crop production), and requiring the input of these surrogates. All three of these approaches would still rely on macroscale emission factors, derived over average duty cycles or from the combination of steady-state emission rates and load factors. We expect that modal emissions data for off-road equipment will become more prevalent with PEMS and similar devices. As this happens it is conceivable that modal emission rates could be generated for off-road as well as on-road sources. Using this information would require commensurate activity information which provides distributions of operating modes. An intermediate step towards this level of refinement might be to categorize emissions for idle and non-idle operation.

c. Microscale Level

The purpose of the microscale level is the estimation of emissions for specific corridors and intersections, as needed for project-level and hot-spot analyses. A **Corridor/Intersection Emissions Inventory Estimator** would combine modal emission rates with vehicle activity at a given location to generate emissions at that location. The user would input a vehicle operation distribution (or the model would assign one based on Level Of Service) and fleet information for the corridor/intersection being modeled. At the most basic level this could simply be applying the mesoscale “modal” option for a specific location, relying on defaults for fleet information, operating mode distribution and grade. Or, the user could input more detailed activity, fleet and grade information via advanced transportation models, or roadside observation. The output from this process would generally be fed into emissions dispersion models (e.g. CALINE, CAL3QHC) for the purpose of estimating pollutant concentration at desired receptor sites. The current application which encompasses this functionality is the MicroFac model under development by EPA’s Office of Research and Development.

d. Emission Rate Estimator

The Emission Rate Estimator processes instantaneous emissions data into the modal and macroscale emissions rates which form the foundation of the New Generation Model system, establishing a consistent foundation for modal and macroscale emission rates. We have identified three fundamental approaches for developing an emission rate estimator which serves this function:

1) Rely on microscale emissions models to generate modal and macroscale emission rates. This process has been used to develop modal emission rates for TRANSIMS using

the UC Riverside model;[5] modal emission rates are derived by using vehicle driving modes (e.g. speed and average acceleration) as input into the model, and using the resulting output to represent emission rates for that mode. Similarly, macroscale emission rates could be developed by using the microscale model to estimate emissions over a representative driving cycle, such as the facility-based driving cycles used as the basis for MOBILE6. This approach could rely on an established microscale emissions model, incorporating updated data as it becomes available. Under this scenario, the emission rates would be considered "base", with independently-derived correction factors applied to reflect the effects of fuel, temperature, etc. Alternately, a new model could be developed based on on-board emissions data (e.g. from PEMS), which would include the effects of fuel, temperature, etc.

2) Generate emissions rates at the higher levels of aggregation by processing instantaneous emissions measurements produced in a laboratory or in the field. This would likely be a pre-processing step, resulting in the modal emission rates included in the NGM. Modal emission rates would be generated by "binning" the instantaneous emissions using a pre-determined statistical process employed to ensure repeatability, consistency and objectivity. An example of one such approach could be Hierarchical Tree-Based Regression (HTBR) analysis, used in the development of modal emission rates for the MEASURE model.[6] Macroscale emission rates would then be derived from the modal emission rates by developing a modal operation distribution from the representative operating cycles and overlaying the modal emission rates, aggregating into a total emission rate over the cycle.

3) Maintain a direct link between a database of raw instantaneous emissions measurements and the New Generation Model. The "emissions rate estimator" would query the database, searching for conditions similar to those being modeled, and extract the relevant data for use in the emissions estimation. This method would require an extensive set of instantaneous emissions data, as well as a database software application to be included in the New Generation Modeling system.

In general, the use of on-board emissions measurement devices to generate emissions data could fundamentally change the approach to developing emission rates that account for the effects important to emission generation (e.g. fuel, temperature, air conditioning). The approach currently employed by MOBILE, necessitated by the reliance on laboratory-based testing data, is to develop "base" emission rates and apply independent emission correction factors to account for each effect. On-board emissions measurement systems will gather data in-situ, without controlling these effects. This should allow a more realistic characterization of in-use emissions, since these data will reflect emissions as they actually occur in the field, including all of the interactions between effects which are generally not accounted for by correction factors. The emission rates which would be derived from on-board data would be more comprehensive, since they would include all the effects which correction factors try to estimate individually. A challenge with basing emission rates on on-board emission data, however, is that it would be difficult to isolate individual effects for the purpose of assessing the emission impact of varying an individual emission effect (e.g. fuel sulfur). The correction factor approach lends itself well to this type of analysis.

e. Issues with Proposed System

- A fundamental issue in considering a model which includes multiple scales is consistency across the scales. First, it is important to recognize that consistency across scales does not mean the same answer would be produced if the same analysis (e.g. an emissions inventory for an urban area) were performed at the different scales. Even with complete consistency across the emissions, vehicle activity, and fleet components of two different scales, the aggregation and disaggregation from one scale to the next will introduce unavoidable differences. Thus, as discussed in the NRC report, clear guidance for which scale to apply for a given analysis will be essential (further discussed in Section III.D.1). The core of the new generation modeling system will be emission rates and factors, and consistency between the scales can be designed in for this component. As discussed in the previous section, the most appropriate way to do this is to ensure that the emission rates and factors for the macroscale, mesoscale and microscale levels are derived from the same data source. As envisioned, activity and fleet characteristics are inputs to the modeling system; for these elements, consistency between scales depends on the quality of the input data provided by the user.

- The application of the macroscale, mesoscale and microscale framework to evaporative on-road emissions and all off-road emissions poses several questions. For on-road evaporative emissions, current microscale and mesoscale emission models tend to rely on the more macroscale emission factors from MOBILE or EMFAC. Aside from running loss emissions, evaporative emissions are generated when the vehicle is parked. Modeling finer levels of resolution for evaporative emissions thus depends on where and when the vehicle is parked, rather than how the vehicle is operating.

- Current off-road modeling techniques (EPA's NONROAD model and ARB's OFFROAD model) would fall under the macroscale definition, in that they estimate overall emission levels for broader geographic areas using aggregate population and activity information. Considering mesoscale and microscale levels for off-road is a new issue. Some of the considerations for microscale or mesoscale modeling of off-road are likely similar to those for the finer levels of on-road emissions modeling, such as spatial and temporal allocation, and emissions by operating mode. Overall, on-road emission modeling methodologies have the benefit of decades of study on transportation activity and driving patterns, whereas the study of off-road activity patterns (compounded by the proliferation of equipment types to consider) is in its infancy.

- EPA's NONROAD model does not include the emissions from aircraft, commercial marine sources or locomotives. Emissions are estimated for these sources through separate models, such as the FAA's Emission Dispersion Modeling System (EDMS) model for aircraft. Integration of these sources into the NGM modeling system will present an additional challenge.

- A modeling system which integrates all scales, all pollutants and all mobile sources requires tremendous data to support it . Particular data voids are real-time emissions for particulate matter and toxics for all vehicles, all pollutants on heavy-duty vehicles, and activity information for off-road sources. EPA's investment in portable emission measurement devices (PEMS) is seen as a long term solution to supply this data. However, the absence of comprehensive data in the short term will necessitate a gradual approach to fulfilling the vision of the New Generation Modeling system.

4. Modeling System Input Data

a. Macroscale Level Input Data

In the interest of stability in transition from MOBILE6 to the New Generation Model, we propose that the content and structure of the input data required for the macroscale level (on-road) be comparable to the information used by MOBILE6. As users gain familiarity with MOBILE6 and the data input options, however, we will get a better sense of whether the MOBILE6 approach should be modified. As with MOBILE, the presence of default data for most of the input parameters would likely remain in the New Generation Model. Because the macroscale level will include the capability to calculate total inventory rather than just emission factors, estimates of vehicle miles traveled will also be a required user input.

The calculation of an emission inventory at the macroscale level has the following basic data elements:

- Emission rates
- Vehicle activity information
- Vehicle fleet characterization information
- Meteorological information
- Control program information
- Fuel specifications

Emissions rates would be considered a core part of the modeling system, and would not be subject to changes from users. However, it could be possible under the NGM system that new macroscale emission factors could be determined from local driving behavior information characterized as alternate driving cycles.

The vehicle fleet characterization encompasses two basic elements. The first is the distribution of vehicle attributes, including the following:

- Model year
- Fuel type
- Emission standards

- Technologies
- Vehicle class
- Vehicle age

This information would need to be provided by the local area or available as defaults for use when local data is not available.

The second element of fleet characterization is the distribution of emission levels within the fleet (and perhaps within the attributes listed above). These data are hard-coded into MOBILE6 as normal and high emitter rates, and are not subject to change. For the NGM, however, we would consider the option for user-supplied emission distribution. Approaches such as remote sensing, roadside pullover emission testing, and I/M data evaluation are possible for characterizing the emissions distribution of a local fleet. Emission distribution information is affected by local efforts to control emissions, such as vehicle inspection and maintenance (I/M) programs. In MOBILE6, emission distributions are adjusted for I/M programs internally by the model using program information supplied by the user. However, if local areas were able to characterize their own emissions distributions, the benefits of their specific I/M program could be estimated as well. Because accurate emission distributions are critical to improved estimates of local emissions inventory, we would like to allow user input to ensure that the best information possible could be used.

Vehicle activity information at the macroscale level includes parameters such as the following:

- Distribution of average speed
- Temporal distribution of engine starts
- Distributions of the time between engine stop and start (soak time)
- Distributions of the time between engine start and stop (trip duration)
- Distributions of the duration of hot soaks
- Temporal distributions of miles traveled
- Temporal distributions of trip ends

An additional element of vehicle activity in the generation of emissions inventories is vehicle miles traveled (VMT). This would be a new aspect of the NGM relative to MOBILE, which leaves the application of VMT to post-processing applications. Aggregate forecasts of vehicle miles of travel (VMT) used to developed county-level inventories of VMT are generally derived from the Highway Performance Monitoring System (HPMS), a nationwide inventory system that includes data for all of the nation's public roads. This data is collected by the Federal Highway Administration on an annual basis. Local areas may also estimate VMT by conducting their own traffic monitoring program, or aggregating results from a Travel Demand Model (TDM). The NGM should allow flexibility for the user to input the best available VMT data; in terms of defining VMT data input structure, however, we propose that the level of aggregation used by HPMS set the standard.

HPMS functional classification systems vary across states. The NGM should be designed to accept HPMS VMT and speed estimates that span the range of facility types represented at a national level. To interface with HPMS, various state facility and vehicle type classifications would have to be translated into those used within the NGM. In addition, counties outside nonattainment areas may have incomplete data, and a VMT scaling approach may be required to derive county level estimates from regional projections of VMT. Data preprocessing would likely be required to address these issues.

Calculations of emissions at the macroscale are relatively straightforward, generally requiring estimates of VMT and average travel speed by facility type, also reported in the HPMS data structure. The NGM should contain an interface that accepts VMT by facility type to support the development of emissions inventories. The NGM would produce emission estimates for each user-defined volume and speed category in an output format that could be used for generating county-wide emission inventories.

Meteorological and geographic data would include parameters such as the following:

- Altitude
- Temperature
- Humidity
- Cloud cover
- Latitude/Longitude (for sunrise, sunset, peak sun determinations)
- Road Grade

Fuel program information would define the fuel properties. Under the NGM system we would envision this input to be similar to EPA's Complex Model, which requires specifications for a number of fuel properties such as aromatic content, sulfur level, and oxygenate content. This level of detail goes beyond MOBILE, which has these properties hard-coded for specific cases (e.g. Phase II RFG). Since fuel specifications would be input separately, additional local control program information would focus mainly on I/M program parameters.

The basic data elements on input data discussed for on-road would apply largely to off-road as well. Meteorological and fuel information could for the most part be in common with on-road emissions, providing some opportunity for synergy between on-road and off-road. For activity and fleet information, we currently propose that the input structure be consistent with activity and "fleet" data used by the NONROAD model. NONROAD accounts for the activity of a given equipment type through the number of hours in use, and an average load factor (percentage of maximum), over the time period of interest. Fleet information is characterized by the total population for the analysis area, and the rated horsepower for the given equipment type. However, the current level of uncertainty in off-road activity and population information may require the development of new approaches for estimating these parameters. As mentioned in Section III.A.3.b, one such approach might be to develop relationships between off-road activity and surrogates, such as construction starts or number of agricultural acreage. If such a method were developed, the relevant surrogate information would be required as input for calculating off-road emissions.

i. Issues with Macroscale Inputs

- Aggregation of vehicle population and activity factors to an area wide basis for macroscale modeling will result in some loss of information. As a result, macroscale results from the NGM will not precisely match results from the mesoscale modeling which have been aggregated to an area wide basis. Specific guidance will be needed to help users decide when more resource intensive mesoscale modeling is needed and how to resolve conflicts between the two approaches.

- MOBILE6 will rely on fixed, national average default driving behavior estimates (driving cycles) to determine emission rates. The NGM will also have default emission rates which rely on national average driving behavior. We may consider allowing local areas to develop their own emission rates based on local driving behavior. However, clear guidance would be needed to assure that the development of local emission rates is conducted appropriately.

- MOBILE6 will not account for the effect of roadway grade on emissions. It is not clear how grade can be accounted for within macroscale modeling, since it affects vehicles on a roadway link basis. This is clearly one of the areas where mesoscale modeling emission results may differ from macroscale results. For many areas grade may not be an important emission estimate issue. However, it may be necessary to find a way to account for the effects of roadway grade on emissions for macroscale analysis.

- Average speed is used in current models as a surrogate for driving behavior. MOBILE5 used a single area wide average speed, but MOBILE6 requires a distribution of average speeds for macroscale modeling. At the macroscale level, an area wide average speed could be determined, but any aggregation of driving behavior will likely further differentiate the macroscale and mesoscale emission results. The current proposal would be to retain the average speed distribution approach used in MOBILE6 and not allow further aggregation of average speeds. Other methods should be investigated as alternatives.

- Different vehicle classes are operated on different roadways at different times of day. This affects their driving behavior (average speed) in relation to other vehicle classes. For this reason, vehicle population and activity information will need to be temporally distributed. MOBILE6 allows for hourly differences in many of the parameters.

- Many parameters will vary across the area (e.g., fleet age distributions, temperature). The spatial distribution of these parameters will be lost in macroscale modeling through aggregation. The differences within the macroscale area are probably best addressed by moving to a mesoscale modeling approach. Guidance will be needed as to when mesoscale modeling would be more appropriate.

b. Mesoscale Level Input Data

Mesoscale modeling involves estimating vehicle activity and emissions rates for specific roadway links and small areas such as traffic analysis zones (TAZ). The objective is to gain much better spatial resolution and some improvement in temporal resolution of emissions estimates. For advanced air quality dispersion modeling studies, human exposure studies, or transportation corridor assessments, the increased level of effort beyond a macroscale approach may be justified. Mesoscale modeling most effectively accomplished using Geographic Information System (GIS) software. This implies that only agencies having necessary personnel and GIS computer hardware and software are likely to implement a mesoscale approach. Many metropolitan planning organizations (MPO) have become enthusiastic GIS users and may be fully capable of implementing a GIS-based modeling application.

The usability concerns for mesoscale modeling are presented in the context of "basic," "modal," and "advanced" options. An overriding issue is the trade-off between resource requirements and the extent of spatial and temporal resolution that is achievable by these alternate approaches. Data requirements are lowest for the basic option and highest for the advanced option. The basic and modal options may not require a GIS but data could be output to a GIS for visualization purposes. The advanced option is similar to full implementation of the Mobile Emissions Assessment System for Urban and Regional Evaluation (MEASURE) model, while for other options resource needs are reduced by relying more on macroscale or fleet average data. An issue may be to what extent "mixing and matching" of the various approaches is acceptable, i.e., can an agency do a combination of basic, modal, and advanced methods.

i. Overview of data requirements

Data requirements for mesoscale emissions modeling for on-road mobile sources are summarized in Table 3.

Table 3 - Data Requirements for Mesoscale Level (On-Road Sources)

Data	Basic	Modal	Advanced
Major road network data	Yes	Yes	Yes
TAZ data	Yes	Yes	Yes
Engine start zone GIS coverages	No	No	Yes
Land-use and tax parcel data	No	No	Yes
Census data	Yes	Yes	Yes
Trip productions and attractions	By TAZ	By TAZ	By TAZ and start zone
Vehicle-miles of travel (VMT) by link	Yes	No	No
Average traffic speed data by link	Yes	Yes	Yes
Volume/capacity and speed/acceleration distributions by link	No	Yes	Yes
Vehicle fleet data	County level	County level	Link and start zone
Roadway grades by link	No	Yes	Yes
Temporal allocation factors	Yes	Yes	Yes
Exhaust emission factors	Macroscale	Modal	Modal
Evaporative emission factors	Macroscale	Macroscale	Macroscale
Meteorological data	Yes	Yes	Yes
Control program information	County level	County level	County level

ii. Data requirements discussion and issues

- Major road network data

  For all options, GIS coverages or database tables for major road networks are created from travel demand model (TDM) outputs. The NGM system would not require use of any particular product, but would define a file format interface for whatever product an MPO uses. Links in the major road network are the same as the major roads defined in a TDM. Spatial accuracy of the TDM network is a potential issue. A highly stylized network with relatively few links will not be as accurate as a network with more links, where all major road links are explicitly represented. For the advanced option, substantially more effort is devoted to developing a spatially accurate road network. At minimum, TIGER Census data should be used. If a local MPO or Department of Transportation has more accurate road network data, it should be used instead of (or to supplement) the TIGER data.

- Traffic analysis zone (TAZ) data

  Polygon GIS coverages or database tables are needed to account for minor road emissions and engine starts. For basic and modal options, the TAZ's an MPO has defined for TDM purposes are used. These typically correspond to US Census Block Groups, and thus may be of variable size. Whether these provide sufficient spatial resolution is an issue.

- Land-use and tax parcel data

  Start zones are created from detailed land-use and/or tax parcel data. Resources required to prepare GIS start zone polygon coverages will vary depending on the geographic completeness, temporal currency, and data elements available in existing databases.

- Census data

  Population and housing data are available in standard formats from the US Bureau of the Census. These data are used for all mesoscale modeling approaches.

- Trip productions and attractions

  Cold start emissions will be estimated by applying an emission factor to an estimated number of cold engine starts, equal to the number of trips predicted by a TDM. The TDM may not completely account for intra-zonal trips, so the actual number of starts may be underestimated. For basic and modal options, engine starts will be estimated by TAZ. For the advanced option, spatial resolution is increased by defining engine start zones. Also, more detailed technology groups of vehicles may be defined and spatially located.

- VMT and average traffic speed data by link

  These are output from the TDM process. For the basic option, both are used directly for MOBILE6-like emissions calculations at the link level. For modal and advanced options, average speed and volume/capacity ratio are needed to determine a level of service to which speed/acceleration profiles may be applied to estimate speed/acceleration distributions for specific road links.

- Volume/capacity and speed/acceleration distributions by link

  Volume/capacity ratios needed for modal and advanced options are a TDM output. Default speed/acceleration profiles would be provided, and would be based on the default driving cycles used to generate the macroscale emission rates. Field studies are needed to validate existing data and to develop profiles for other facility classes and for other cities. For best temporal resolution of trips, AM, PM, and off-peak network loading (volume and capacity) values are needed.

- Vehicle fleet information

  For basic and modal options, macroscale level fleet information would be used (e.g. county level registration). The spatial distribution of subfleets is not recognized with this approach. For the advanced option, address matching of motor vehicle registration data is performed to enable modeling of fleet characteristics for specific road links and start zones. For defining detailed technology groups, VIN decoding is required.

- Roadway grade by link

  Grade data are needed to account for engine load effects in modal emissions calculations for the modal and advanced options. In the absence of grade data, zero grade can be assumed, which will of course mean that grade effects were not accounted for. The least demanding approach for generating input data would be to obtain digital elevation maps (DEM) from the US Geological Survey. Alternate methods include ground surveys of roads, grade measurements with accurate Global Positioning System (GPS) equipment and engineering construction drawings. An issue is that modal emissions estimation approaches have not been adequately tested to determine their sensitivity to grade, so the actual benefit of the more sophisticated grade measurement approaches is not known.

- Temporal allocation factors

  Temporal allocation data may be the same for all options, and may be the same as for a macroscale approach. Temporal profiles are needed to allocate daily trips and starts to hours of the day. Profiles specific to individual start zones and major road links that reflect AM, PM, and off-peak traffic loads are ideal. Separate profiles are needed for commercial trucks and heavy duty trucks used for interstate commerce. Often, no TDM model outputs for predicting weekend traffic loads are available.

- Emission factors

  Macroscale exhaust emission factors for major roads for the basic option are the same as MOBILE6 emission factors, but applied at a link level. For the modal and advanced approaches, modal emission factors are applied for hot stabilized emissions on links and engine starts. For evaporative emissions, use of macroscale emission factors as in MOBILE6 is expected for all options. These would not be inputs per-se, since they are considered core elements of the NGM.

- Meteorological data

  Data such as temperature and humidity are the same for all approaches and would rely on macroscale values. Hence the same issues as noted in the macroscale discussion apply.

- Control program information

  Data inputs such as I/M program parameters and anti-tampering program data are the same for all approaches and rely on macroscale data in the same manner as these data inputs are treated in MOBILE6. The advanced mesoscale option adds value to control program information by linking it to vehicle subfleets by geocoding vehicle registrations, so that spatial differences in the impacts of control programs can be discerned.

As discussed under Section III.A.3.b, the consideration of the mesoscale level for off-road emissions applies at this time primarily to the refinement of activity and population information. The three approaches which could be used to generate off-road emissions at the mesoscale level are 1) allocating macroscale emissions down to finer spatial zones based on readily-available surrogate information such as population or land-use type, which would require input of the appropriate surrogate information; 2) using localized activity and population studies to provide information directly for a specific area, which would require the input of these data; and 3) using localized activity studies to develop improved relationships between equipment activity and surrogates (e.g. construction starts, crop production), which would require the input of these surrogates. Methods 2 and 3 in particular would require a significant effort to gather activity and emissions data at the local level. As on-board emission measurement technology develops, we would like to consider partnering with state and local air quality agencies or other groups to conduct off-road testing programs and surveys in order to support this level of off-road emissions modeling.

c. Microscale Input Data

As with the mesoscale approach, microscale modeling would likely entail "mixing and matching" of input data across different analysis scales. It would be conceivable to take most input parameters down to the microscale level. However, it is more likely that only activity and possibly fleet information would be obtained at the microscale level, while macroscale level estimates would be used for fuel, control program, and meteorological information. Alternately, a microscale analysis could simply be a mesoscale level analysis done for only a specific roadway link or intersection. This would entail the use of aggregate estimates of vehicle activity on that link, as defined under the mesoscale level.

Two approaches for generating activity and fleet data at the microscale level are roadside observation, and the use of advanced transportation models. Roadside observation, a method used to generate input information for the MicroFac model, employs license-plate identification in order to characterize of the fleet in terms of vehicle distribution. Remote sensing employed as part of the roadside observation is a possible method for characterizing the emission distribution of the fleet. Roadside observation using laser range-finder would allow the characterization of vehicle operating mode distribution.

A second method for generating input data is the use of transportation models which generate estimates of vehicle micro activity. Transportation micro-simulation models are commonly used to support the evaluation of transportation investments that improve corridor-level traffic flow and provide activity information for corridor/intersection inventory generation. The most basic micro-simulation models will produce vehicle delay, typically expressed as intersection level, link level, and stop delay. Estimates of delay are then used to refine estimates of link level speed, which are associated with macroscale emission rates to estimate emissions for that corridor/intersection. However, for effective use of the microscale level, we consider it necessary to use modal emissions rates. Advanced simulation models generally predict the percent of vehicle hours spent in each of four operating modes: cruise, acceleration, deceleration and idle. Emission estimates are also a function of the facility type, the physical characteristics of the specific segments of the facility, and the travel demand model predicted volume/capacity (v/c) ratio for each segment. The corridor/intersection inventory estimation component of the NGM could be designed to a) accept roadway segment characteristics from network simulation models through the incorporation of an internal processing function to calculate modal emissions, or b) generate a modal emissions rate input file stratified by the variables considered within network simulation models. The former approach would be difficult to adapt to the wide range of simulation models used in standard practice, while the latter approach would enable the NGM user to specify the range of network attributes to be considered in the preparation of modal emissions rates for microscale analysis.

Microscale analysis of off-road emissions would require location-specific information for activity and fleet characteristics. This could be obtained through observation, or through the use of surrogate activity information for the specific analysis area.

5. Modeling System Interface With Air Quality and Dispersion Models

As discussed in Section III.A.3, and shown in Figure 1, the output for each analysis level is emissions information at the appropriate scale: county-level for macroscale, roadway link and zone level for mesoscale, and specific corridor/intersection for microscale. An important consideration is that mobile source emission model outputs are used as inputs into air quality and/or dispersion modeling systems, which produce estimates of pollutant concentration at a given location and time of day. Air quality models such as the Urban Airshed Model (UAM) and REMSAD predict the dispersion of pollutants emitted from all sources (mobile, area, point, and biogenic) and model the complex photochemistry which leads to the formation of second-order air pollutants such as ozone and secondary particulate matter. Dispersion models such as CALINE predict the concentration of pollutants (primarily CO and PM) at a specific location, for use in estimating pollutant exposure at that location. The design of the NGM must consider how mobile sources information is input into these models. The following sections discuss issues related with the interface between the NGM and both air quality and dispersion models.

a. Interface with Air Quality Models

In general, interfacing the NGM with air quality models will require the development of gridded emissions estimates from either the macroscale or mesoscale analysis levels; this function is served by a **Grid Processor**, shown under the “output” column in Figure 1. We envision that the Grid Processor would not be a part of the NGM modeling system, since these applications already exist under air quality modeling frameworks developed and supported by EPA, such as MODELS3. However, the execution of a grid processor will require close coordination with the NGM development and execution, since there is the potential for significant overlap in the allocation of emissions and the use of meteorological and chemical speciation information.

Air quality models and the various emission processors which provide episodic emissions data to them require varying amounts of activity, emission factor, and source classification information. Outside of the emission estimation algorithms established by current on-road and off-road mobile source models, additional data related to the temporal, spatial, and speciation calculation needs to be provided before these emissions estimates can be used by air quality models. Mobile source emissions are just a part of the total emission estimates required to ensure proper air quality modeling results. The emission files needed by air quality models are currently generated by emission processors, which prepare input from all sources (area, point, mobile, biogenic) in a consistent spatial, temporal, and chemical format for use by the air quality model. For this reason, interface between the NGM, emissions processors and air quality models may best be met by defining those parameters and input data that are required to run the mobile source modules of current and expected emission processors and to let the processors generate the temporal, spatial, and speciated data needed for input to air quality models.

Currently, the on-road modules of emission processors used by the air quality modeling community generate emission data by running scripted versions of either the currently available on-road mobile model code (MOBILE5) and applying the resulting emission factors to an episode and geographically defined activity or directly processing pre-estimated emissions. The scripting method generates multidimensional emission factor matrices based on speed, temperature, vehicle class, and fuel parameters that are generally based on input data provided to MOBILE5 (and will be updated to incorporate MOBILE6). These emission factors are then applied to county, grid cell, or link-based activity input and processed by the emission processor for the specific episode and geography of study. These processors thus have the capability to generate emissions estimates at both the macroscale and mesoscale levels. Separate from the need to process emissions for air quality models, this functionality is desired within the NGM system, in the form of the Macroscale and Mesoscale Inventory Estimators. One issue to consider is whether relevant pieces of existing or future processing tools can be used to fill the function of these estimators within the NGM system.

For providing input to air quality models, one possible approach is to take advantage of the current capabilities for emissions processors to take pre-estimated source emissions; off-road emission estimates are currently handled in this way. In the NGM system, this approach would mean that the Grid Processor would accept complete emissions estimates from the Macroscale/Mesoscale Inventory Estimators. Emission processors would then merge these mobile-based emissions with pollutant data from other categories (point and area sources) and regrid this combination to the domain of study. Using chemical speciation assignments and profiles to further define chemical species, the processors classify each specie based on one of the simplified chemical mechanisms used for air chemistry by air quality models. It would be most appropriate for the processors to perform this speciation and classification, as well as combine the mobile source species with non-mobile based emissions, to generate the complete set of emissions data in a consistent manner.

Under this approach, it is up to the user of the NGM system to ensure consistency between the episode-specific inputs (e.g. meteorological data) used in generating the mobile source inventories, and all other sources.

In order to generate stand-alone mobile source emissions inventories, the NGM already will need to encompass the inventory-generation capabilities of current emissions processors. However, in cases where air quality modeling is specifically the goal, a second approach which ensures consistency across all sources may be necessary. This approach would be to have the NGM serve a function similar to the current use of MOBILE in emissions processors; that is, it would generate temporally and spatially resolved activity and emission factor data only, and allow the Grid Processor to assign meteorological and perhaps other spatial data for completing the mobile source emission estimates. Under this case, the Grid Processor might access the Macroscale/Mesoscale Emissions Factor Estimator directly. These emissions in combination with other source types, could then be chemically resolved into model-specific mechanisms and processed for air quality model inputs. Additionally, using the macroscale module of the NGM, county or grid-based emission estimates may be calculated using less spatially resolved temperature fields (similar to our existing Trends calculations) or by importing temperature data consistent with the emissions modeling episode.

i. Issues with air quality modeling interface

- Multiple air quality models, and therefore various emission processors, are currently used in air quality analyses. How best can the NGM interface with the various processors?

- Many existing emissions processors have portions of MOBILE model code built into their system for "on-the-fly" emissions estimation. Will the NGM's code be modular and flexible enough for an emissions processor to use internally, or is it expected that these processors be modified to run the NGM externally?

- Certain episode and geography-based elements (temperature, altitude, etc.) for the estimation of emissions from mobile sources are currently used by emission processors. To the extent that these data are required by the NGM to estimate emissions, how can they be made consistent with other sources' emission estimating elements from the same study domain?

- Mobile source emission factors defined as a function of speed, temperature, vehicle and fuel type are necessary parameters for on-road emission estimation. Are these factors best produced by the NGM modules and passed to the emissions processors as lookup tables or should the processors call the NGM directly?

- Meso- (and potentially macro-) scale NGM runs will provide valuable temporal and spatial resolution to emission generating activity. Should these user- or NGM-defined input data be output by the NGM for use as input in the various emission processors?

b. Interface with Dispersion Models

Among the suite of air quality models, the use of air dispersion models is widespread in evaluating the environmental impact of transportation programs. As part of their transportation conformity requirements in CO nonattainment and maintenance areas, state and local agencies are responsible for demonstrating that transportation projects will not cause or contribute to new violations of National Ambient Air Quality Standards (NAAQS), increase the frequency or severity of existing violations, or prevent attainment of the standards. Locations experiencing ambient concentrations of CO also must conduct air quality analyses at intersections and other local-scale "hot spot" sites which may be in violation of the NAAQS. The National Environmental Policy Act (NEPA) requires that the environmental impacts of a federally-funded project be assessed prior to its initiation, including changes made to the physical, social, and economic environment. Air quality impacts that result from transportation-related projects are often considered in such analyses. Human exposure assessment, conducted by various governmental entities, universities, industry, and non-governmental organizations, also depend on air quality models.

In conducting such air quality analyses, air dispersion models are used to predict the concentration of air pollution resulting from motor vehicle emissions. EPA provides guidance for the use of dispersion models in our *Guideline on Air Quality Models* and the *Guideline for Modeling Carbon Monoxide from Roadway Intersections*. These guidelines describe modeling procedures appropriate for assessing air quality impacts of motor vehicle transportation. CALINE3 is the EPA-recommended dispersion model for highway project-level modeling of potential "hot spots". CALINE3 is a simple Gaussian dispersion model that treats motor vehicle emissions in a simplified manner as "line sources." The model does not contain an emissions component, but requires inputs from a mobile source emission factor model, such as the MOBILE model. CAL3QHC is the EPA-recommended model for CO intersection modeling. It employs the CALINE3 dispersion modeling system, but employs procedures in the 1985 Highway Capacity Manual to calculate the expected length of vehicle queues. In some areas, the models TEXIN2 and CALINE4 are used in air quality assessments. These models are similar to CAL3QHC and CALINE3, respectively.

CALINE3 and CAL3QHC are used in modeling specific transportation projects, including free-flow roadways and intersections. In addition to emission rates (specific to each lane of travel), CALINE3 requires free-flow vehicle flow rates and roadway geometry, specified as model inputs. CAL3QHC requires additional information for calculating vehicle delay at intersections, including traffic light cycle time, red light time, vehicle delay (when accelerating from a stop), signal type, vehicle arrival rate at the intersection, and traffic volumes under both congested and free-flow traffic patterns. The specific output needed from the NGM by current dispersion models would be emission rates per lane of travel, provided in grams emitted per vehicle-mile. Each lane of travel may vary by speed, so the ability to accurately predict emissions by lane is required. As vehicle speeds and fleet characteristics may vary significantly between lanes, accurate emissions models will allow for specification of lane-specific vehicle and traffic parameters. Emission models should also take into account road grade and other parameters which may affect vehicle load. As proposed, the microscale level of analysis will be well-suited to provide this information, since it would be able to estimate corridor (and lane) specific emissions accounting for modal driving behavior and roadway grade.

EPA does not have near-term plans to replace CALINE3 and CAL3QHC as its standard dispersion models for project-level analysis of transportation-related air quality. However, some researchers are currently involved in the development of models able to estimate transportation-related air quality more realistically. In EPA's National Exposure Research Laboratory, computational fluid dynamics (CFD) simulation has been used in characterizing air flows near transportation facilities and may be used in improving dispersion algorithms in the CALINE series of dispersion models. Such simulation requires emissions which are resolved into small segments of time, on the order of 5-10 minutes and broken into individual vehicle lanes. Researchers at the University of Central Florida have developed the Traffic Air Quality SIMulation (TRAQSIM) which treats mobile sources as multiple moving point sources. It uses Gaussian Puff dispersion algorithms to predict air quality. TRAQSIM requires modal vehicle emissions, which it currently calculates from MOBILE emission factors. Researchers at the National Environmental Research Institute of Denmark have developed the Operational Street Pollution Model (OSPM). OSPM is a dispersion model using wind vortex meteorology in urban canyons. To date, OSPM assumes a uniform emission rate from roads, although more accurate model predictions could be gained with emissions predicted more accurately. Other researchers have investigated the use of Lagrangian dispersion models to predict air quality resulting from vehicular emissions in turbulent vehicle wakes.

All dispersion models currently under development would require emission rates resolved into space and time scales smaller than those currently available through the MOBILE model. Their input requirements are considerably greater than CALINE3 or CAL3QHC. In order to feed these models, the New Generation Model would have to estimate emissions at the space and time scales needed for accurate dispersion modeling.

B. Improving the Science

This section discusses the second of the four broad objectives in our design of the new generation modeling system - improving the science. We consider this to be two-pronged effort, involving the establishment of a quality process for model development, and the improvement of the underlying data upon which the model is based. Both issues are discussed in the following sections.

1. Establishing a Quality Assurance Process

a. Performance Criteria, Validation, Uncertainty

A primary issue with EPA's modeling program as identified in the NRC report was the lack of objective criteria upon which to measure the effectiveness of MOBILE or related models. This comment was made specifically in reference to the need for improved uncertainty assessment and validation effort, and the related determination of how accurate the model needs to be for the application it is being used for. The incorporation of these issues is crucial to the improvement of confidence in EPA's mobile source modeling tools. We believe these issues must be addressed in the model in a systematic fashion. As such, we plan to adopt a set of objective criteria which address a broad range of issues for model development, drawing from agency criteria for quality assurance, model evaluation, peer review and coding standards.

Guidance being developed under EPA's Quality System will be a useful roadmap for establishing these criteria. The EPA Quality System is administered under the Office for Environmental Information (OEI) and specifies compliance with ANSI standards for assuring quality implementation of environmental data collection and technology programs.[7] An important component of the EPA Quality System is the implementation of a Quality Assurance Project Plan. EPA's Office of Environmental Information is currently developing guidance for these plans specifically for model development. This guidance would require the establishment of a thorough process for model development which encompasses many of the NRC recommendations, covering the following elements:

- **Model needs and requirements analysis**: This process would entail determining what needs the model must address, defining how the modeling project will address these needs, and defining quality objectives and model performance criteria in meeting the established needs. Establishing objective criteria for the model requires determining the required accuracy of the model, and establishing a process to determine whether performance criteria are being met.

- **Model development process**, covering model design, coding and documentation. The design component includes developing the theoretical basis, mathematical formulation, necessary inputs, and methods for determining model uncertainty. The coding process involves establishing hardware/software and code performance requirements, and transforming the model design into software code. Peer review and documentation are fundamental to each stage of the development process.

- **Model application process**, focused on executing the model and evaluating model results against performance criteria through an established validation/calibration process.

The guidance for developing a Quality Assurance Project Plan will provide a useful blueprint for establishing objective criteria up-front for the development of both the science and software components of the New Generation Model. The guidance itself does not provide the criteria, but will require the establishment of criteria for model development and implementation drawing from sources within and outside the agency. One source for more concrete criteria could be an EPA initiative known as the Council for Regulatory Environmental Modeling (CREM).[8] This initiative is headed by EPA's Science Policy Council and is charged with developing criteria for model evaluation, including model applicability, uncertainty, peer review, documentation and calibration. However, this initiative is in its infancy and it is unclear whether concrete guidance will be available in the timeframe needed for the development of the New Generation Model.

b. Peer Review

EPA's Science Policy Council has developed peer review guidance for environmental regulatory modeling.[9] This guidance lays out a framework for stages of model development and implementation, including model objectives, theoretical basis, parameter estimation, data quantity/quality, key assumptions, model performance criteria, and model documentation. Acceptable peer review mechanisms include a) the establishment of an ad-hoc panel of at least three scientists, b) use of an established peer review mechanism such as the Science Advisory Board, or c) a technical workshop. This peer review guidance would provide a concrete process for implementation of the peer review elements of the Quality Assurance Project Plan.

2. Data

While a number of organizations are actively involved in generating and collecting mobile source emission test data, current mobile source emission models tend to be based upon relatively small data sets. There appears to be no widely established, systematic framework for sharing or pooling mobile source emission test data. It is often difficult to combine data from multiple studies and current models are frequently not updated as more data becomes available.

We realize that these problems will never be completely solved, but the New Generation Modeling system will include a data collection and storage framework to address them in a systematic way. Since it is becoming technically feasible to collect large amounts of data inexpensively by instrumenting in-use vehicles and recording information during their normal use, the data framework will be oriented toward this type of data. Until sufficient data is assembled from instrumentation of in-use vehicles, however, the modeling system will probably have to rely on laboratory test data (or a combination of lab data and on-board data), so this storage framework should also accommodate laboratory data. This includes summary emission test results based on driving cycles, and continuous measurements made during laboratory driving cycle testing. Because on-board measurement only gathers exhaust emissions, evaporative emissions will continue to rely on laboratory test methods.

The framework will include guidance describing what vehicle activity and emission test data measurements are needed to develop and maintain the NGM system and standard practices as to how they should be collected. The guidance will also define these data items and provide formats for their transmission and storage. We envision that data pertaining to the test studies, the test vehicles, vehicle trips, and fuels used will be needed in addition to instantaneous measurements of ambient conditions, vehicle operating parameters, and emissions.

The principal function of this database will be the derivation of modal emission rates for the NGM via the emission rate estimator discussed in Section III.A.3.d. It is envisioned that more aggregated emission factors in the model (e.g. over an operating mode or driving schedule) would be derived from the modal emission rates.

The NGM database format should also be suitable for validating aggregated emission factors and, by appropriately sampling or weighting of the data, to derive estimates of vehicle activity (e.g. starts per day, soak time length distributions, etc) that can be used for national level emission estimates and for local area estimates when local data are not available.

As early as possible in the implementation of the NGM the data design will be completed and mechanisms for the collection, transmission, combination, and storage of this information will be established. Actual loading of existing information and collection of new information will then begin. EPA will solicit the assistance of other organizations to provide this information.

C. Improving the Software

This section discusses the third of our four broad objectives in our design of the new generation modeling system, improving the software.

1. Producing Quality Software

A clear area for which the quality of EPA's mobile source emissions models can be improved is in software engineering. MOBILE code dates back to the 1970's, and the flexibility and maintainability of the code has been reduced with each update. The NGM provides a fundamental opportunity to create a new software framework. Bertrand Meyer, in *Object Oriented Software Construction*, describes software quality as follows:[10]

- Correctness - meets the specification.
- Robustness - handles cases that arise outside the specification.
- Extendability - the ease of adapting the software to changes of specification.
- Reusability - modules that perform a particular task (i/o for example) can be used by all portions of the program that require this task.
- Compatibility - ease of combining software elements.
- Efficiency - minimize the demands on hardware, time, storage.
- Portability - users will always want the program to run on their system.
- Ease of use - easy to install, easy to use. The program should be easy to use for the beginner, easy to automate for the expert. The dictum is, "Do not pretend you know the user." Programs have a way of growing past their initial audience.
- Functionality - Stick with a basic plan and get it working. If the design is modular, additional features can be added later.
- Timeliness - Produce it on schedule.

Borrowing on these points, the fundamental goals for the software structure of the NGM are that it be maintainable, amenable to frequent updates, flexible, developed expediently, and usable (as detailed in later section). The choice and skillful application of software engineering methods is central to achieving these goals. Ideally, model design methods, software design methods, and implementation methods should be chosen to enable a nearly seamless transition from model design to software. The software flowchart should look like the modeling flowchart contained in Figure 1. The issue is to properly choose modeling design methods, software design methods, and implementation language(s) to facilitate the transition from high level design to code. Software design experts should be consulted for model development as well as software development.

An example of the relationship between model design and software design is contained in documentation for TRANSIMS.[11] This documentation describes four principles of modern software engineering. These are layering, modularity, iteration, and object-orientation. Layering is a hierarchical arrangement with the user interface at the top, various system and subsystem components in intermediate layers, and technology components in the bottom layer. Each layer uses layers below it, but not above. Modularity keeps components separate so that maintenance, enhancements, and extensions are easier. Iteration is a development style that produces frequent interim versions of the model. Problems are then solved along the way and not accumulated at the end of the development process. Finally, object-orientation is the technology that makes the previous three principles possible. A similar architecture may or may not be appropriate for the NGM, but these issues must be addressed. For a model of this size and complexity, it is important to have software and model design inform each other.

Planning in the early stages of model development is crucial. The earlier in the development process that a problem is discovered and resolved, the cheaper it is to fix. Flaws discovered in the requirements or architectural phases of projects are 50 to 200 times cheaper to fix than those discovered in construction or maintenance.[12] This issue is addressed both by extensive early planning and by an iterative development cycle.

2. Usability

For the New Generation Model software to be accessible and easy to use a number of factors need to be considered, including how input data is provided to these components, their user interface, and the hardware and software needed to run them.

At the input level, we envision that NGM components requiring user input data would accept ASCII text files (and as discussed in Section III.A.4.a, would be based on the MOBILE6 input structure, possibly upward compatible if it makes sense). Ideally all NGM components which accept input data would do so in the same general way. As necessary, the NGM would also contain modules which would convert other commonly used formats to this common format.

Regarding the software interaction with the user, components which interface with the operator would offer both graphical and batch interfaces. Graphical interfaces are generally easier for beginning or low volume use, while batch interfaces, which allow for model runs to be specified from pre-stored files, are better for automation (voluminous or repetitive runs).

As regards computer hardware requirements, it is a goal of this project that all NGM components not require a high-end computing environment. They should operate on any EPA-standard desktop personal computer, or any widely-used, late-model computer system that is based on the WINDOWS operating system and Intel-compatible processor. That is, we want the system to run on the common computers that most people have. Components of the “advanced mesoscale modeling option”, to be described later in this paper, may necessitate some compromise of this objective.

We want the NGM software to be as platform-independent as possible and require no additional software licenses to use. In particular, the macroscale emissions component should require no additional software licenses. There are some areas, however, where compromise of this desire may be necessary. Some mesoscale components may be implemented in Geographic Information System (GIS) software such as ARCVIEW or ARC/INFO, which would constrain their portability and require the user to have a license to the GIS product. The database of emissions data which supports the NGM would probably also be implemented in a third party software product, in this case a Database Management Software (DBMS) tool. This component will not be part of the NGM itself, however, and can be made accessible using a wide variety of software tools.

3. Modularity

Software modularity involves a number of characteristics, some which are visible to the end user, and others which benefit the development and ongoing evolution of the NGM product. Perhaps the most important result of software modularity for the end user is that components of the NGM system should be usable individually or in combination with each other without the user having to modify them. For this to happen, and for these interfaces to be simple and easy to understand, the components need to have well-defined data interfaces, fit together into a logical overall design, and be documented with a common project methodology, terminology and style. The use of object-oriented software design and implementation methods will be strongly considered by this project. This project will also be undertaken with an awareness of the Multimedia Integrated Modeling System (MIMS) project already underway in EPA, which is designed to implement a framework for modeling components.

For those who will develop and maintain the NGM system, an important benefit of software modularity is that replacement of a model component with another which conforms to the same interface requirements should be possible without changing the way the program is used. Of course, if the new component produces different results, then the new results must be scientifically valid and be reviewed and approved.

Underlying any modular software structure must be a software design and development process to produce it. Given EPA’s capabilities and budget constraints, development of the NGM will undoubtedly involve contractor support to a significant degree. While contractors may support the software design process, EPA intends to functionally design the NGM software in considerable detail with in-house staff resources, and to determine the choice of tools used to develop the modeling system. EPA also intends to provide sufficient definition and oversight of any contracted efforts to insure that EPA retains an in-house understanding of how all model components are implemented, and not to become dependent on particular contractors for software maintenance and further development.

4. Coding Guidelines

The model should follow applicable EPA guidelines with respect to coding and software. Examples of EPA guidelines are those in preparation by the Council on Regulatory Environmental Modeling (CREM) and those promulgated by the Office of Environmental Information (OEI). Such guidelines apply to coding standards, quality assurance, documentation, and configuration management. In addition to Agency guidelines, we may wish to adopt additional guidelines to help assure that the NGM's objectives are met in a timely fashion.

Coding guidelines specify naming and usage conventions. They assist primarily in the readability and clarity of code. There are many ways of writing clear and readable code. By choosing a particular way as a standard, a reader who has mastered the coding conventions in one module will easily be able to read any module in the program. Coding guidelines thus promote quality control and ease of review. They also insure that programmers can work easily on multiple modules. As an example of coding guidelines, the document "TRANSIMS Software Architecture for IOC-1 (First Interim Operational Capability)" prescribes coding guidelines for TRANSIMS by reference to several generally available programming books plus a few additional recommendations. The exact choice of coding guidelines will depend on the specific language(s) that are employed in the model.

Architectural guidelines apply to a level of software structure above code and deal with the interrelationships between classes and packages. Their purpose is to make software that is easy to understand, maintain, enhance and extend. The architecture of the model needs to be inherently flexible, so that later changes can be readily accommodated. The architecture will likely be object-oriented, and standards of good object-oriented programming should be adhered to. Classes should be specified in a uniform manner. The use of assertions, contracts, and other program elements should be uniform throughout.

Many different people and groups may be working on this project at different times. Architectural and coding guidelines will insure compatibility and readability. To the degree possible, and certainly to the degree that contractors are writing code, formal rules of variable naming, commenting, and internal documentation should be adhered to. Someone reviewing the code should not have to adjust to a different set of conventions as they read the work of different software authors.

D. Process and Implementation

This section discusses the fourth of our four broad objectives in our design of the new generation modeling system: implementing the model in a coordinated, clear and consistent manner.

1. Guidance

The creation of clear, concise guidance for the proper use of the New Generation Model is a key component of the model's development. Without proper guidance providing non-ambiguous instructions for the use of the model, the practical implementation of the New Generation Model is impossible. Guidance for the New Generation Model will exist in two forms. One will be the New Generation Model User's Guide. The other form of guidance will be policy guidance. The goal of both forms of guidance will be the production of consistent results for all users that are as accurate as current science allows.

The User's Guide for the New Generation Model will be developed concurrently with the model itself and in coordination with selected members of the user community. The User's Guide will serve as the technical foundation for the actual running of the model. This will include detailed descriptions of data input formats, data sources, and modeling options. Furthermore, the User's Guide will provide specific, clear and easy-to-follow instructions on quality assurance, data sources, and modeling methodologies. This will ensure consistency between modeling results among various user groups. The User's Guide will enable the user to create an input file and to run the New Generation Model to obtain results. The goal of the User's Guide is to provide instructions that are clear and easy-to-follow for the various components of the New Generation Model. The use of each feature and option in the New Generation Model system will be discussed in the User's Guide. The User's Guide will also discuss the incorporation of modeling results into air quality models, such as the Urban Arced Model.

The second component of the New Generation Model guidance, the policy guidance, will also be developed with the cooperation of selected members of the user community. The policy guidance will serve as the policy foundation for using the model. The policy guidance will address the appropriate use of each component of the New Generation Modeling system, as well as the appropriate use of each control feature and option in the model. A key component of the policy guidance will be the appropriate use of the different scales available in the New Generation Model. Another very critical component of the guidance will be the issue of what type of data and what quality data is appropriate for each use and each different scale. This issue is much more critical than in the past due to the number of different levels of analyses that will be possible with the New Generation Model. Due to the mathematics of aggregation and disaggregation, even the use of the same data will yield different results when used at different scales. Additionally, use of the New Generation Model at increasingly fine scales will necessitate increasingly detailed data. These critical issues will need to be addressed when the model is released for use. The policy guidance will be updated when warranted. The policy guidance will also address the various uses for the modeling system, such as emission inventory generation, analysis of alternative control strategies, input for air quality models, and hot-spot or exposure analyses. Documentation and data requirements for each type of analysis as well as for the submission of State Implementation Plans (SIPs) using the New Generation Model will be incorporated into the policy guidance.

Both of the above documents will be developed in coordination between EPA's Office of Transportation and Air Quality and the Office of Air Quality Planning and Standards. This coordinated effort will ensure the consistent use of results obtained through the New Generation Model.

2. Coordination

There are many components of the New Generation Model that must be coordinated with other groups both within and outside of EPA. These components include the collection of data supporting the model, development of the model itself, guidance for the use of the model, training of personnel in the operation of the model, and testing the New Generation Model before its final release.

The development of a modeling system as complex as the New Generation Model requires the careful coordination of both data values and modeling algorithms. The most current and accurate data will serve as the basis of the model system. Additionally, the most current modeling algorithms will be used in the New Generation Model. This is to ensure that the final product achieves the goal of enabling more accurate modeling at different scales while also enabling users of a range of different capabilities and resources to use the model. Furthermore, the science behind the model must be validated. Therefore, the development of the New Generation Model will involve the on-going input of selected users as well as a peer review conducted by researchers and federal partners. It is important to note that the recommendations of both large and small organizations will be solicited to address the needs and abilities of diverse modeling groups. These groups will include selected state air agencies, transportation agencies, selected metropolitan planning organizations, other Federal partners, and other EPA offices. At key interim steps, the New Generation Model work product will be subject to peer review by researchers in the transportation and air quality fields. The peer review must be passed before the model is approved for regulatory use. Additionally, all of the EPA Regional Offices will be kept informed on the status of the development of the New Generation Model.

A critical component of the development of the New Generation Model is the close coordination with the federal Department of Transportation and the California Air Resources Board. This coordination was a specific recommendation of the NRC report and will serve to maximize the use and accuracy of the New Generation Model while minimizing resources required for its development. One key issue with the Department of Transportation is the application of the New Generation Model for transportation analyses from both a technical and policy perspective. Another key issue with the Department of Transportation is the coordination of the EPA's New Generation Model with DOT's TRANSIMS model. Ultimately, we would like to see both modeling systems using a common basis for emission estimation, a goal which is currently under discussion between EPA and DOT. Similarly, we would like to work with the California Air Resources Board towards a coordinated effort on emissions research and the development of emission estimates.

The development of guidance will be coordinated with EPA program and regional offices, as work on the New Generation Model progresses. As noted above, this guidance will consist of both a comprehensive User's Guide as well as policy guidance. These two documents will be provided for review to other organizations for comment before they are subject to final release.

The development of training for the use of the New Generation Model will be coordinated with other Federal partners, such as the Federal Highway Administration, and the national associations of potential users. These include the State and Territorial Air Pollution Program Administrators/Association of Local Air Pollution Control Official (STAPPA/ALAPCO), National Association of Regional Councils (NARC), and of state, regional, and local air planning agencies as well as transportation and planning agencies. Training may take several forms. Computer-based training may be developed to allow the individual user to progress through the training materials at their own rate. Training classes and programs may also be developed in response to requests. Satellite training would allow a large number of users to be trained simultaneously. These forms of training will be evaluated and recommendations will be made regarding the training of users in the New Generation Model.

IV. Scope Considerations and Next Steps

The New Generation Modeling system as laid out in this paper represents a significant increase in scope for EPA's mobile source emissions modeling program. Given that resource constraints may dictate how much of this scope could be accomplished, it is useful to bound the maximum and minimum scope which we would pursue under the NGM effort. A discussion of the "full scope" and "reduced scope" options are included below.

A. Full Scope Option

The full scope option incorporates the breadth of the proposal discussed throughout this paper. It entails developing emission rates, software, and guidance for each component in the NGM modeling system, for on-road and off-road.

B. Reduced Scope Option

The reduced scope option arises from the consideration of what would be the minimum necessary to meet the overall objectives for the project as stated in Section III and the analysis needs discussed in Section II.B. Along these lines, a fundamental question is what EPA's role should be in developing software, versus developing emissions estimates and guidance for using these estimates. Taking these issues into consideration, the reduced scope option is based on the following premise: a) the core of EPA's emissions modeling program should be the emission rates themselves; b) EPA must at a minimum continue providing software at the level of MOBILE and NONROAD, i.e. for macroscale emissions inventory estimation; c) many users of mesoscale and microscale modeling will develop their own software applications (or would use emerging transportation/emissions software such as TRANSIMS), and our primary concern is that consistent emission rates and input data be used; d) applications already exist for macroscale inventory generation, in the form of emissions processors discussed in Section III.A.5.a; and e) agency resources may be better focused improving emissions data and developing emission rates based on these data.

Based on these premises, a "reduced scope" option for the New Generation Model could consist of the following elements:

- **Emissions rate estimator** as discussed in Section III.A.3.d, the end product of which would be "agency-official" modal and macroscale emission rates
- **Macroscale emission factor estimator**, essentially updates of MOBILE and NONROAD but likely with new software
- **Guidance** for implementing the mesoscale and microscale levels. The guidance would require the use of EPA emission rates, and would specify acceptable input data for use at each level.

Another scope reduction to consider is the inclusion of off-road emissions as a longer-term goal. Ultimately, we think integrating on- and off-road sources into a single modeling system makes sense. However, given the data gaps for off-road emissions and activity, it may be more realistic to develop the on-road portion of the New Generation Model first, and bring off-road in when data gaps have been narrowed, particularly at the mesoscale and microscale levels. This notion is reinforced by the fact that aircraft, commercial marine and locomotives are sources which are not included in the current off-road modeling tools, and will require additional effort to integrate.

C. Interim Product

The New Generation Model system can be separated into two fundamental parts: the modeling system as a whole, and the individual components of the system. The intent of the overall modeling system is to establish an object-oriented design such that the individual components can be designed to fit into it without disrupting the rest of the system. As such, we are considering an interim product for the New Generation Model which would focus on developing the software framework and establishing the input/output criteria for each independent component, but rely as much as possible on existing applications to fill the individual processing components. For example, the macroscale emission factor estimator might be filled by MOBILE6, and the macroscale emissions inventory estimator might be filled by certain components of SMOKE. Many existing applications could be used or modified to fulfill the process functions of the modeling system. The purpose of such an interim product would be to a) develop the framework software architecture, and b) develop the New Generation Model at the proof- of-concept level. The modeling system would be in place without modification to the emissions result. The final product for the New Generation Model would then be focused on updating the software and science of the individual components, within the parameters of the established overall framework.

D. Next Steps

We plan to publish a comprehensive plan for the New Generation Model by Fall 2001. The main purpose of this plan would be to provide concrete steps for the development of the New Generation Model and allow for more detailed determination of resource needs, data needs, and timing. OEI's Quality Assurance Project Plan, discussed in Section III.B.1.a, may serve as a blueprint for the comprehensive plan. In general, we see three major issues which need to be addressed as we work towards this plan: 1) further definition of the modeling system, including the underlying model theory and the input/outputs of each modeling component; 2) establishing a methodology for estimating emission rates, including an assessment of how on-board emissions would be analyzed and a sampling plan for populating the model with on-board emissions data; and 3) developing a software design.

Appendix A: EPA Mobile Source Emissions Modeling Workgroup

EPA Office	Sub-Office	Workgroup Members
Office of Air and Radiation	Office of Transportation and Air Quality	John Koupal, Mitch Cumberworth, Dave Brzezinski, Harvey Michaels, Chad Bailey, Rich Wilcox, Gene Tierney
	Office of Air Quality and Planning Standards	Greg Stella, Bill Johnson, Gary Blaise
Office of Research and Development	NRMRL	Chuck Mann, Sue Kimbrough
	NERL	Alan Huber, Bill Benjey
Regions	Region 4	Dale Aspy, Alan Powell, Rob Goodwin

Appendix B: Acronym Glossary

ANSI	American National Standards Institute
ASCII	American Standard Code for Information Interchange
CAAA	Clean Air Act Amendments
CAL3QHC	A model for predicting pollutant concentrations near roadway intersections
CALINE	California Line Source Model (for Microcomputers)
CARB	California Air Resources Board
CMAQ	Congestion Mitigation & Air Quality
CMEM	Comprehensive Modal Emissions Model
CO	carbon monoxide
CORSIM	CORSIM is a microscopic, stochastic traffic simulation model.
CREM	Council for Regulatory Environmental Modeling
DBMS	database management system
DEM	digital elevation model
DOT	Department of Transportation
DQO	data quality objective
DTIM	Direct Travel Impact Model
EMFAC	Emission Factors (California Air Resources Board's Emissions Factor Model — e.g., EMFAC 2000)
EMME/2	System for planning the transportation of people on multi-modal networks
EPA	U.S. Environmental Protection Agency
EPS	Emissions Preprocessor System
GAO	General Accounting Office
GIS	geographic information system
GIT	Georgia Institute of Technology
GPS	global positioning system
HPMS	Highway Performance Management System
ITS	intelligent transportation system

MEASURE	Mobile Emissions Assessment System for Urban and Regional Evaluation
MicroFAC	Microscale Motor Vehicle Emission Factor Model
MIMS	Multimedia Integrated Modeling System
MOBILE	EPA's Emission Factor Model for Highway Vehicles
MOBTOX	EPA's Emission Factor Model for Toxics Emissions from Highway Vehicles
MODELS-3	EPA's Third Generation Air Quality Modeling System
MPO	Metropolitan Planning Organization
NAAQS	National Ambient Air Quality Standards
NARC	National Association of Regional Councils
NCHRP	National Cooperative Highway Research Program
NEPA	National Environmental Policy Act
NERL	National Exposure Research Laboratory (within EPA/ORD)
NGM	New Generation Model
NONROAD	EPA's Emission Factor Model for Nonroad Sources
NO_X	nitrogen oxides
NRC	National Research Council
NRMRL	National Risk Management Research Laboratory (within EPA/ORD)
OAQPS	Office of Air Quality and Planning Standards
OAR	Office of Air and Radiation
OBD	on-board diagnostics
OEI	Office of Environmental Information
OO	object-oriented
ORD	Office of Research and Development
OTAQ	Office of Transportation and Air Quality (formerly Office of Mobile Sources)
Paramics	Paramics is a suite of software tools for microscopic traffic simulation developed by Quadstone, Limited.
PART	EPA's Emission Factor Model for Particulate Emissions from Highway Vehicles
PDF	portable document format
PEMS	Portable Emissions Measurement Device

PM	particulate matter
PM-10	particulate matter < 10 µm
PPAQ	Post-Processor for Air Quality
QAPP	quality assurance project plan
SIP	state implementation plan
SMOKE	Sparse Matrix Operator Kernel Emissions
SO_X	sulfur oxides
STAPPA/ ALAPCO	State and Territorial Air Pollution Program Administrators/Association of Local Air Pollution Control Officials
TAZ	traffic analysis zone
TCM	transportation control measure
TDM	travel demand model
TEA-21	Transportation Equity Act for the 21st Century
TIGER	Topologically Integrated Geographic Encoding and Referencing System
TRANPLAN	**TRAN**sportation **PLAN**ning
TRANSIMS	Transportation Analysis Simulation System
UAM	Urban Airshed Model
UC Riverside	University of California - Riverside
USGS	U.S. Geological Survey
VHT	vehicle hours traveled
VIN	vehicle identification number
VMT	vehicle miles traveled
VOC	volatile organic compounds

References

1. National Research Council, *Modeling Mobile Source Emissions*, National Academy Press, May 2000

2. Singh, R., and Huber, A., "Development of a Microscale Emission Factor Model for CO (MicroFacCO) for Predicting Real-Time Motor Vehicle Emission", 2000

3. Barth, et al., "Development of a Comprehensive Modal Emissions Model", NCHRP Project 25-11 Final Report, April 2000

4. Bachman, et al., "Modeling regional mobile source emissions in a geographic information system framework", Transportation Research Part C 8 (2000) 205-229

5. Williams, et al., "The TRANSIMS Approach to Emissions Estimation", Los Alamos National Laboratory Report LA-UR 99-471, 1999

6. Wolf, et al., "High-Emitting Vehicle Characterization Using Regression Tree Analysis", Transportation Research Record 1641, 1998

7. EPA Office of Environmental Information (OEI) Quality System website: http://www.epa.gov/oei/quality.htm

8. EPA Council for Regulatory Environmental Modeling (CREM) website: http://www.epa.gov/ORD/spc/2crem.htm

9. EPA Science Policy Council Peer Review Guidance website: http://www.epa.gov/ORD/spc/modelpr.htm

10. Meyer, *Object-Oriented Software Construction*, 2nd Edition, 1997, Prentice Hall Professional Technical Reference.

11. Berkbigler, et al., "TRANSIMS Software Architecture for IOC-1", Los Alamos National Laboratory Report LA-UR-97-1242, 1997

12. Meyer, *Object-Oriented Software Construction*, 2nd Edition, 1997, Prentice Hall Professional Technical Reference.