CHAPTER 3 - Emission Inventories, Air Quality Measurements and Modeling: Guidance on Their Use for Air Quality Risk Management

Jeffrey R. Brook[1] * (measurements), William Pennell[2] (emissions), Michael D. Moran[1] (modeling)

[1] Air Quality Research Division, Atmospheric Science and Technology Branch, Science and Technology Directorate, Environment Canada
* Chapter Lead Author
[2] NARSTO (www.narsto.org)

KEY MESSAGES

- Three essential tools for managing the risk due to air pollution are multi-pollutant emission inventories, ambient measurements and air quality models. Tremendous advances have and continue to be made in each of these areas as well as in the analysis, interpretation and integration of the information they provide.
- Accurate emission inventories provide essential information to understand the effects of air pollutants on human and ecosystem health, to identify which sources need to be controlled in order to protect health and the environment, and to determine whether or not actions taken to reduce emissions have been effective.
- Air quality measurements are essential for public health protection and are the basis for determining the current level of population health risk and for prioritizing the need for reductions. They are also critical for evaluating the effectiveness of AQ management strategies and altering such strategies if the desired outcomes are not being achieved.
- Air quality models quantify the links between emissions of primary pollutants or precursors of secondary pollutants and ambient pollutant concentrations and other physiologically, environmentally, and optically important properties. They are the only tool available for detailed predictions of *future* air concentration and deposition patterns based on possible future emission levels and climate conditions.
- Air quality problems tend to become more difficult to address as the more obvious and less costly emission control strategies are implemented. This increases the demand for advanced scientific and technological tools that provide a more accurate understanding of the linkages between emission sources and ambient air quality.
- Despite scientific advancements, including improved understanding of the impacts of poor air quality, the pressure to identify cost-effective policies that provide the maximum benefit to public health push our current tools and knowledge to their limits and beyond.
- Due to scientific uncertainties, highly specific control options that target specific chemical compounds found on fine particles, specific sources or source sectors or that lead to subtle changes in the overall mix of chemicals in the air (gases and particles) remain extremely difficult to evaluate in terms of which options most benefit public health. Lack of a complete understanding of exposure and health impacts of the individual components in the mix and their additive or synergistic effects pose further challenges for health benefits evaluation. However, progress is being made and new ways of thinking about air quality and pollution sources, such as the concept of intake fraction, help to provide some perspective.
- A broader perspective, including consideration of environmental effects and the implications of climate change on air quality and on co-management of air pollutants and greenhouse gases, will be increasingly important to embrace.

3.1 Introduction

Emission inventories, air quality measurements and air quality modeling are scientific cornerstones supporting air quality risk management. Developing and applying these tools, along with source apportionment, which are depicted in Figure 3.1, are the key steps involved in understanding how chemistry, meteorology and natural and human emissions interact to produce observed levels of outdoor air pollution. In addition, a wide range of air quality (AQ) measurements and exposure analyses are essential for epidemiological research aimed at uncovering the current risks posed by air pollution and for subsequent risk assessment exercises. The purpose of this chapter is to provide an overview of the roles that emissions, measurements and models can play in air quality risk management and in understanding air quality issues. This information and the references therein are also intended to provide some insight into current capabilities and best practices associated with developing and applying these essential tools.

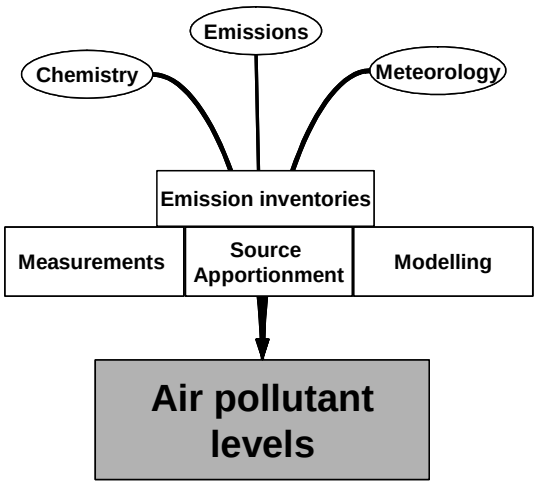

Figure 3.1 Emission inventories, ambient measurements and air quality models are the tools needed to understand the current air pollutant levels and predict future levels under various policy options.

Several valuable reports on air pollutant emissions, ambient measurements and air quality modeling have been published in the past. In particular, the NARSTO particulate matter assessment for policy makers (NARSTO, 2003) describes measurement methods, North American emissions and observations, receptor-based methods of data analysis and interpretation and the status of air quality models for particulate matter. The World Health Organization report on "Monitoring Ambient Air Quality for Health Impact Assessment" (WHO, 1999) outlines the principles underlying air quality monitoring networks and other related activities (e.g., modeling) that help insure they are of most use for supporting health impact assessment.

Figure 3.2 shows the basic steps of AQ risk management and specifies how scientific inputs from emissions, measurement and modeling play a direct role in the policy process. They enable the prediction of air quality improvements associated with emission reduction options, as well as the analysis of the costs and benefits of air quality management options. Although the figure depicts the process in a linear, sequential fashion, with science and policy proceeding separately, in practice the order of steps may be reversed or steps may occur in parallel. In addition, science plays a key role in identifying appropriate air quality goals and options for emission reductions. For example, in developing a conceptual model of the sources and atmospheric processes that lead

to current ambient pollutant concentrations, there may be a need to gather additional measurements to test and refine the model before one can thoroughly evaluate whether or not the tools are reliable. In addition, depending upon the maturity of air quality risk management in a particular location, not all steps may be required. Existing measurement programs may be fully adequate or the AQ models may have already been widely accepted for the intended use.

A crucial step in air quality management is to quantitatively link ambient pollutant concentrations at specific locations or within specific geographic regions to specific emissions (emissions to concentration relationship). This linkage is studied through both receptor and source-based AQ models. Source-based models are capable of predicting future ambient air quality concentrations and are applied to evaluate emission reduction scenarios in the context of Figure 3.2. Model estimates of concentration changes can then be integrated with concentration-response functions (CRFs) to estimate health benefits.

Table 3.1.1 summarizes the various ways in which emissions, measurements and models are applied directly in AQ risk management. Ideally, AQ management should strive to address problems from a multi-pollutant, risk-based perspective that emphasizes results over process, takes an airshed approach to controlling emissions, creates accountability for these results, and modifies air quality management actions as data on the effectiveness of these actions are obtained (NRC, 2004). Although improvements are needed, current emission inventories, measurement activities and modeling tools are consistent with this objective. To the extent that resources permit, they continually evolve attempting to incorporate the most up to date scientific thinking and technologies, which dictates that to understand and effectively address AQ problems a *one atmosphere* approach is necessary.

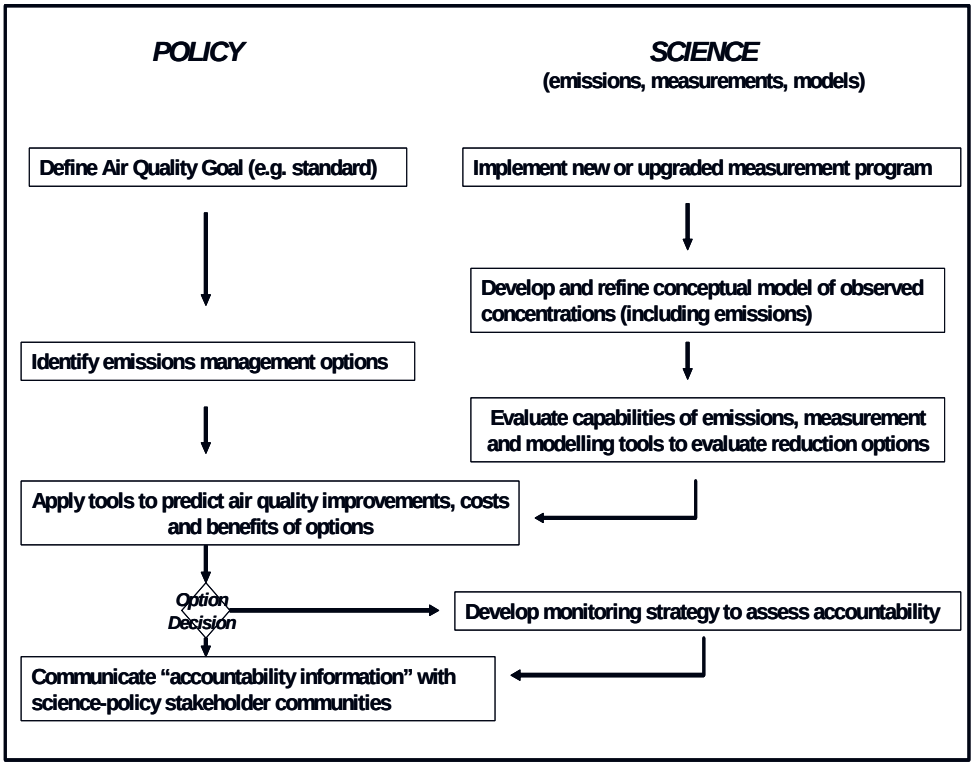

Figure 3.2. The role of emissions, measurement and modeling in local/regional air quality risk management.

Table 3.1.1: The application of emissions data, air quality measurements and air quality modeling in air quality risk management.

Tool	Area of Application in Air Quality Risk Management
Emissions	• Current emission rates for criteria gases and particles by source type and location serve as the starting point for assessing the need for and feasibility of reductions • Projected emission rates for criteria gases and particles by source type and location and detailed information on the causes of the future changes in emissions • Identification of broad based and detailed emission reduction strategies and technologies by source type and their effectiveness for the emissions of criteria gases and particles
Measurements	• Characterization of past and current pollutant levels and identification of exceedances of AQ standards, objectives, or targets • Time series of ambient concentrations at population-based monitoring sites for trend analysis in relation of emission reductions • Determining the relationship between ambient concentrations at population-based monitoring sites and a range of health endpoints (concentration-response function) • The relationship between ambient concentrations of primary and secondary pollutants and emission source categories (source apportionment or receptor models) • Development and evaluation of conceptual models and source-oriented models
Models	• Simulation of emission scenarios and quantification of resulting benefits and disbenefits by prediction of ambient concentrations at multiple time and space scales for: o Base case (e.g., current emissions) o Emission levels when policies currently "on-the-books" are fully implemented o New emission reduction scenarios • Estimation of emission changes required to attain AQ objectives or standards • Evaluation of emission estimates • Quantification of source-receptor relationships • Characterization of governing chemical regimes and limiting reactants for current and future conditions • Simulation and design of new or modified measurement systems (network optimization, site selection, input to data assimilation and analysis routines)

Emissions, measurements and models also play an indirect role in AQ management through the provision of information to the general public or specific stakeholder groups. This includes media reports conveying current pollutant levels or the air quality index (AQI), maps published on line (e.g., http://airnow.gov/) and AQ forecasts and/or smog advisories. An example of publicly available emissions information is the North American Commission for Environmental Cooperation (CEC) series of reports ranking major sources and assessing progress (www.cec.org/takingstock/index.cfm).

Right-to-know websites such as the Toxic Release Inventory in the U.S. (www.epa.gov/tri/) and the Canadian National Pollutant Release Inventory (www.ec.gc.ca/pdb/npri/npri_home_e.cfm) provide specific emissions information for local areas. Public access to emissions information is increasing worldwide (e.g., Mexico: http://app1.semarnat.gob.mx/retc/index.php and www.epa.gov/ttn/chief/net/mexico.html) and international standards for a Pollutant Release and Transfer Register (PRTR) have been established (www.epa.gov/tri/programs/prtrs.htm).

This information can be valuable for highlighting areas of large emissions and situations where there has been a lack of progress, as well as for assisting members of the general public in learning about emissions in their region. However, it should be viewed as a starting point for more detailed examination because the information may be outdated, incomplete and/or misinterpreted by the media or special interest groups.

This chapter elaborates on the functions of emissions, measurements and modeling in AQ risk management. Section 3.2 provides the basics of contemporary emissions inventory development, evaluation, and dissemination. Section 3.3 describes the various areas of application of measurement data in air quality risk management and provides guidance on technical issues to be considered in establishing a robust measurement program. Section 3.4 describes the application of models for air quality risk management, identifies key technical issues to consider in using air quality modeling systems and reviews eight steps for best practice in using models in air quality management. Section 3.5 describes current efforts and future directions to combine the capabilities of emissions and measurement information and air quality modeling to better support air quality management. Section 3.6 provides the chapter summary and recommendations.

3.2 Emissions Information for Air Quality Risk Management

3.2.1 Introduction

Without accurate information on the sources of air pollutants – what they are, where they are located, what they are emitting, and how much – it is impossible to identify which sources are most important to control, to predict the air-quality consequences of these emissions, or to monitor the effectiveness of emission reduction programs. Emissions information is provided, assessed and compiled at many different levels. These can include specific industries continuously measuring and reporting their emissions, national and local governments compiling information and running models to extrapolate the available information across sources, time and space, as well as air quality modelers, who process the information for specific applications. To obtain accurate information it is necessary for both government and industry to bear some responsibility, acting in partnership to ensure accuracy, proper interpretation and continual improvement. Arguably, the emissions information in the U.S. and perhaps North America (N.A.) represents the current state-of-the-art. Thus, this section of Chapter 3 discusses how emissions are determined in N.A. and the current strengths and weaknesses of the available information. Much of the material discusses is from a more extensive report: *Improving Emission Inventories for Effective Air Quality Management Across North America: A NARSTO Assessment* (NARSTO, 2005). This perspective should be relevant to all agencies/stakeholders, regardless of country, seeking to obtain and improve emissions inventories for AQ management purposes.

3.2.2 Emission Inventory Development

Emission inventories are usually developed using the following model

$$E = EF \cdot A \cdot (1\text{-}ER)$$

Where E is the emission rate (e.g., kg/hr or tonnes/yr) of a given pollutant, EF and A are the emission factor and activity factors, respectively, and ER ($0 < ER < 1$) an emission reduction factor, which accounts for any emission control devices that may be applied to the source. The emission factor, EF, is the mass of a given pollutant or chemical species emitted per unit process variable. The activity factor, A, is the related process variable such as mass of fuel consumed, vehicle kilometers traveled, etc. in a given amount of time. In reality, emission and emission reduction factors can vary from source to source as well as with the value of the activity factor, type of fuel, operating conditions, age of the source, geographical location, time of year, and so forth. Not all of these complexities can be accurately represented in such a simple relationship, and more sophisticated emission models have been developed for very complex categories such as mobile source emissions (Miller et al., 2006).

Except in rare instances, emission factors or their equivalents are based on measurements. The easiest source class to characterize is large point sources, such as electric generating units or stack emissions from large industrial operations. Emissions from these facilities can be measured by direct sampling of flue gases – as long as reliable sensors and methods are available (they can be in situ or remote) and appropriate sampling techniques are used.[1] Using these techniques, gaseous emissions from large point sources, such as CO_2, SO_2, and NO_x, can be estimated to better than +/- 20% over time periods as short as one hour (NARSTO, 2005).

Emissions from more dispersed and numerous individual sources are much more difficult to characterize and evaluate. Examples include fugitive (i.e., inadvertent) emissions from industrial sources, natural emissions from vegetation, agricultural emissions, emissions from small industrial or commercial sources, residential sources (e.g., particulate matter emissions from cooking or space heating), and large-scale biomass burning. For these types of emission sources, direct measurements may be difficult or they may be feasible for only a small sample of the sources in question. All of these factors lead to emission estimates that are more uncertain than for large point sources. These uncertainties can range from a factor of two to complete neglect of an unknown source or chemical precursor that upon later analysis turns out to be significant.

On-road and non-road mobile source (automobiles, trucks, aircraft, locomotives, construction equipment, ships, etc.) are a good example of an important, but widely dispersed and variable, component of pollutant emissions. Over the past 50 years, considerable effort and resources have been expended in several developed countries to develop procedures for estimating emissions from mobile sources. The traditional approach for estimating automobile and truck emissions has been to measure emissions from dynamometer tests of representative vehicles in the laboratory. The dynamometer tests are run to represent typical driving cycles, and the vehicle emissions are measured in real time. These measurements are used as input to complex mobile source emission models that attempt to simulate vehicle fleet operating conditions for a wide range of urban, suburban, and rural settings. The problems with dynamometer test are that number of sampled vehicles may be too small to represent a statistically valid sample, and they may not represent the range of fuels used, driving cycles or conditions, environmental factors, and states of repair of the actual vehicles in use. Consequently, field measurements using instrumented chase vehicles, road-side remote sensing of vehicle plumes, chemical sampling of the air in traffic tunnels, and other experimental set-ups are used to check and fine tune mobile source emission models. In the future low-cost portable emission measurement systems (PEMS) and on-board diagnostic sensors (OBDs) may allow cost-effective sampling of a much larger sample of in-service on-road and non-road vehicles under real-world operating conditions. These data could greatly improve the accuracy of mobile source emission estimates.

Because most sources are not equipped with continuous emission monitors to measure actual emissions, Equation 1 forms the basis of most data reported in inventories. Although the focus is often placed on the value of the emission or emission reduction factor, the activity factor is equally important. Activity factors can be developed from continuously monitored process data, but as with continuous emission monitors, these data are generally scarce. More frequently, activity factors are developed from economic activity data or activity surveys. Fuel consumption data are a good example of the use of economic activity data that are collected for reasons other than emissions, but can be used in the development of emission inventories. Data are usually available in the U.S. and Canada by type of fuel used, for various time periods (monthly or sometimes weekly), and for various geographical areas (counties or states/provinces). Information on construction activities can be used to develop emissions from

[1] For example, when measuring particulate matter emissions it is necessary to mimic the cooling and flue (or exhaust) gas dilution processes that occur immediately after the emissions enter the atmosphere, as many "primary" particles are formed in this near-source region.

off-road construction sources. Population densities coupled with activity surveys can provide inventory developers with information on emissions from residential woodstove, fireplace, and open burning. Land use maps and satellite data are useful for estimating the types and densities of vegetative cover, which in turn are used to estimate biogenic emissions. In each case, emission estimates can be developed using data that have been collected for other purposes, such as for tax estimation, economic development, or land-use planning.

Projections of future emissions also depend upon these data and estimates in their growth rates in future years. In the absence of significant technological change, past relationships between population growth and the types of activity factors noted above provide a good starting point for estimating future activity levels, and subsequently, future emission levels.

3.2.3 Evaluating Uncertainty in Emission Estimates

Uncertainties are introduced into emission inventories in a number of ways. Emission factors do not usually account for variability in emissions due to changes in source operating conditions, or across the individual sources that make up a source category. New technologies can change processes and emissions, and such changes will not be reflected in emission factors that were based on the original process. Emission factors that are based on idealized operations, such as use of vehicle operating cycles, do not accurately capture actual operation and therefore actual emissions. Measurement biases or errors introduce additional uncertainties into the reported inventory data. These differences can be associated with the location, time, or composition of emissions, leading to uncertainties in the spatial, temporal, or chemical data used in air quality models. Clearly, quantifying uncertainty is an essential 'best practice' in inventory development and it is more efficient to obtain the information needed to assess uncertainty at the time the emissions data are developed as opposed to afterwards.

Emission uncertainties tend to have a smaller impact on understanding and AQ management decisions when models are used to estimate changes in air quality over longer periods of time and across geographical areas that are of the same order of magnitude as the spatial scale covered by the inventories. In other words, national annual average pollutant concentration estimates are likely to have a lower uncertainty than concentration estimates for a specific urban area over the course of a single day, because the differences between actual and estimated emissions based on emission factors are more likely to average out over a longer period of time and over a broader area.

As air quality models become more sophisticated to meet the demand of more specific AQ management questions more detailed emissions information is needed. The ability to model atmospheric processes over the course of a single hour with more detailed chemical reaction mechanisms and in smaller areas means that the differences between the actual emissions within the modeled area and time and the estimate based on an annual average emission factor may be significantly different. Such discrepancies can result in estimates of pollutant concentrations that do not reflect actual conditions. These differences can lead to misidentification of the most important sources within a given area or erroneous estimates of the specific emissions that need to be controlled for a given source type, as well as incorrect AQ forecasts and population exposure estimates.

Characterizing the uncertainties in non point-source emission estimates is not a simple task. Two general approaches are usually taken – bottom-up and top-down. In the bottom-up approach, uncertainties (bias and random error) in the individual measurements or parameters that make up the emission model are estimated (e.g., from field measurements similar to those described in the previous section) and represented as a probability distribution function (pdf). These uncertainties are then propagated, often using a Monte-Carlo approach, through the model to provide an estimate of the uncertainty in the emission estimate. An alternative bottom-up approach is sensitivity analysis. A simple form of sensitivity analysis is to evaluate the sensitivity of emission-model output to its

various input values in terms of the partial derivative of the model output to the input parameter in question. This approach indicates the relative sensitivity of the emission model to its various inputs and enables a crude estimate of uncertainty by providing a measure of how much the emission estimate would change for a given change in an input parameter.

In top-down evaluations of emission inventories or emission models, ambient measurements or other independent data are used to evaluate the accuracy of the emission estimates. The most effective applications of top-down evaluations are those that are combined with concurrent examination of the original bottom-up inventory data, so that the source of the errors can be identified rather than simply stating that the inventory is in error (Miller et al., 2006). One top down method is to compare temporal trends in emission estimates with past trends or to compare trends in the ambient concentrations of a pollutant (or in ratios of pollutants) with the trend in estimated emissions under conditions in which the effects of transport, chemical transformation, and removal can be neglected or accounted for. Results from of this kind of analysis are reported in Parrish et al. (2002) and CRC (2004). A description of this approach is also provided in NARSTO (2005).

Other top-down methods for evaluating emission inventory uncertainties include using alternative methods for estimating emissions (such as comparing vehicular emission estimates based on vehicular distance traveled with those based on total fuel consumption), source apportionment techniques, and inversing modeling. Source apportionment (or receptor modeling) techniques use various multivariate statistical methods to infer source types, source location, and relative contribution from ambient measurements (Watson and Chow, 2005). These methods have been used to evaluate inventories of $PM_{2.5}$ and VOCs (McMurry, Shephard, and Vickery, 2004; Watson, Chow, and Fujita, 2001).

Inverse modeling involves reformulating source-based air quality models so that emission source strengths are expressed in terms of the observed concentrations. In other words, the model is used to deduce the temporal and spatial emissions that explain the observed concentration field. Because of the limited spatial resolution of air quality models, this technique is most commonly used to deduce area sources (Petron et al., 2002; Park et al., 2003; and Gilliand et al., 2003).

3.2.4 Weaknesses of Current State-of-the-Art Emission Inventories

Over the past 40 years there has been considerable improvement in the accuracy and completeness of emission inventories, but considerable challenges still remain. As of today, N.A. air quality managers have a good understanding of the emissions from major point sources, and they have used this knowledge in developing effective actions for reducing them. Models for estimating emissions from mobile sources have been continuously improved. The importance of natural and biogenic emissions has been recognized, and this knowledge has affected the design of air quality management strategies in regions where these emissions are significant. In Canada and the United States, emission inventories and models can provide quantitative estimates of emissions at national, state or provincial, and county (or their equivalent) levels for many source categories, and there is an improved understanding of the relative importance of various source categories to specific air quality problems. Air quality managers can use these inventories to track emission trends and to evaluate the effectiveness of measures designed to reduce these emissions. In Mexico, emission inventories have been completed for the Valley of Mexico and the states bordering the United States, and in September 2006 Mexico released its first National Emissions Inventory.

In spite of good progress, emission inventories in N.A. have significant weaknesses that will become increasingly important to address for continued success in dealing with future air quality problems. Recognizing these weaknesses helps provide guidance for improvements and for adopting best practices in inventory development programs in counties with less developed emissions databases. The main weaknesses identified by NARSTO (2005) are:

- Development of mobile source inventories, particularly regarding the speciation of volatile organic compounds, remains a challenging problem. National inventories in Canada and the United States have also indicated problems with the magnitude of carbon monoxide emissions and the temporal trend of nitrogen oxide emissions.
- Emissions for important categories such as biogenic emissions, ammonia, fugitive emissions, open biomass burning, and many other area sources are difficult to determine, and they remain inadequately characterized.
- Emission estimates for air toxics (e.g., the 188 hazardous air pollutants designated by the U.S. EPA) are particularly uncertain since there are so many of these compounds, so many potential sources (many of them area sources), and so little data for establishing emission factors or speciation profiles.
- Emissions of particulate matter and more-importantly its chemical constituents, size distribution and key volatile and semivolatile precursors are in need of improvement across many source categories. Carbonaceous particles (organic and elemental carbon) are a large contributor from many sources and there is limited information for several of them.
- Quality assurance and quality control procedures have not been strictly applied in the development of most North American emission models and inventories. In addition, the documentation of uncertainties and data sources in emission inventories has not been adequate to allow the uncertainties of the entire inventory, or of air quality models using the inventory, to be accurately estimated. These are issues that must be addressed in the initial design of a national inventory development program. Addressing them retroactively is expensive.
- Of necessity, emission estimates must be based on a limited number of emission measurements. If this number is not representative of real-world activity, the precision and accuracy of estimates developed from these measurements will be limited. More measurements are needed and the issue of representativeness needs to be examined closely.
- The process for developing information on emissions with the kinds of spatial and temporal resolution needed for location-specific air quality modeling and intra-urban scale exposure estimation is problematic and a source of unquantified uncertainty in model results.
- Methods used to estimate emissions of individual chemical species in emission models must be kept up-to-date if they are to provide accurate information.
- Emission inventories must be developed and updated in a timely manner.
- Differences in how emission inventories are developed in adjacent countries create difficulties for jointly managing air quality.

3.2.5 Actions for Addressing Weaknesses

Reducing known uncertainties in an inventory will provide a more accurate starting point for air quality management strategy development, which should result in more cost-effective approaches. Typically, management actions are initially focused on large point and mobile source emissions. Large point sources are the easiest to characterize and frequently the easiest to control. Mobile sources may be more difficult to completely characterize, but there are few points of manufacture. Thus, control devices can be readily installed during the manufacturing process. As emissions from these sources decline, however, remaining emissions are more evenly distributed across source categories that are even more difficult to characterize, model and control. These remaining sources will also grow as both population and economic activity increase and errors in emission estimates from smaller individual sources will have greater consequences. These consequences could range from wrongly identifying a pollutant that should be controlled to overlooking source categories whose control could result in more cost-effective emission reductions. As this situation is reached, it becomes increasingly important to address the kinds of emission inventory weaknesses described above. Clearly, if these pitfalls are

recognized in the initial phases of an inventory development program, it may be possible to avoid them or to address them in a more efficient manner.

The recent NARSTO (2005) assessment provided eight recommendations to the countries of N.A. on how to address the shortcomings of their national emission inventories. The first recommendation was judged to be the most important. The others were ranked as having somewhat lower priority, but in some cases they may also need to be addressed in the course of meeting the first objective. Given the different degrees of inventory development across countries, regions, and even pollutants, the secondary priorities will differ for each situation.

1. *Reduce uncertainties associated with emissions from key under-characterized sources.*

Comparisons of national emission inventories with ambient measurements and other independent measures should be used to indicate which source categories and pollutants are inadequately characterized and reported. Of particular concern are non-point sources including on-road and non-road mobile sources as well as fugitive emissions from industrial facilities, landfills, sewage disposal systems and feedlots. Sources of organic compounds, carbonaceous particulate matter, ammonia, and hazardous air pollutants are typically not well characterized. *Ideally resources should be targeted to reduce the greatest sources of uncertainty and focused on those source categories (or individual sources and conditions within these categories) whose control will be most effective in reducing costs and health risks while achieving air quality management goals.*

2. *Improve speciation estimates.*

Contemporary air quality issues, such as particulate matter and ozone and the identification of hazardous air pollutant "hot spots," require detailed information about the species being emitted from the sources. Contemporary emission inventories are weak in this regard. *It is essential that source speciation profiles be continually updated and assessed. In addition, the related activity data must be developed to estimate more accurately speciated emissions of particulate matter and precursors, volatile organic compounds, and toxic air pollutants.*

Speciation is most important to the management of ozone and fine PM, but it can also be important to air toxics and to some extent climate issues (black carbon vs. organic carbon; different GHGs). More accurate quantification of the species being emitted will result in better air quality modeling results, and in better identification of and/or discretion in which sources should be controlled in to gain the greatest improvements in air quality and in human health, assuming that some of the more toxic constituents and their sources can be identified.

3. *Improve existing and develop new emission inventory tools.*

Technical advances in instrumentation and computation are enabling emission measurements and analyses that have been previously impractical. Examples of these innovations include portable emission measurement systems for direct measurement of emissions under real-world conditions and the application of various remote-sensing techniques for measuring emissions or verifying emission estimates. Continuing development of these and other technologies as well as consideration of different approaches to deriving emission information, are likely to improve emission inventory measurements and analyses. *Funding agencies need to continue to support the development and application of new technologies for measurement of emissions and ambient concentrations of pollutants*. Application of these technologies will assist the development of emission models that more accurately represent emissions from real sources in time and space.

4. *Quantify and report uncertainty.*

Uncertainties in emission estimates and the parameters or data are used in emission models must be quantified. Uncertainties in emission inventories, processors, and models of Canada, the United States, and Mexico are poorly documented. As a result, the reliability of emission estimates cannot be quantified.

Quantitative measures of uncertainty and variability must be a standard part of reported emission inventory data. Agencies responsible for producing emission inventories must develop specific guidance on how to prepare and report information on emission uncertainties.

This recommendation is probably the most difficult to describe and implement. Quantifying the uncertainty of an inventory estimate, and subsequently the estimates of pollutant concentrations, will enable decision makers to assess the likelihood that the desired outcomes will be achieved. A highly uncertain inventory means that the desired air quality improvements are less likely to occur because there is less confidence that inventory estimates accurately describe the actual situation. The ability to assess the potential for strategies to achieve the desired results will help decision makers to determine what steps are most likely to yield improvements.

5. *Increase trans-national inventory compatibility and comparability.*

As air quality problems become increasingly global in nature, it is vital that emission inventory development and reporting programs be coordinated internationally. Although there have been substantial improvements in reporting national emission inventories in mutually consistent ways, further work is needed to make these diverse inventories more comparable across organizations, purposes, political boundaries, and time periods. *International standards for emission inventory structure, data documentation, and data reporting should be developed. Such standards are needed to facilitate management of long-range transport and trans-boundary air quality issues.*

Inventories that are compatible across boundaries, and with inventories of other pollutants (GHGs, toxics, etc.), allows air quality managers to account for emissions that occur outside their domain of responsibility. This reflects the physical reality that pollutants do not recognize political boundaries. Comparability with past inventories enables air quality managers to more accurately assess the degree to which previous air quality management strategies have been successful.

6. *Improve user accessibility.*

The accessibility of emission inventories and emission models is impeded by the sheer size of the files and the cumbersome manner in which the data are reported and archived. *As improved accessibility to emission data is critical to meeting the diverse needs of the user community, increased efforts should be made to facilitate user accessibility to emission inventory data and models through the Internet and other electronic formats.* It is also suggested that emission inventories be made more transparent and easy to update. National inventories do not always contain the most recent emission information. Methods for allowing continuous updating with appropriately validated data from a variety of sources (researchers, industry, government agencies, etc.) need to be developed.

As more groups are able to access inventory data quickly and easily, a better understanding of air quality problems can be expected and more alternative approaches to air quality management can be developed. The more alternative approaches that are available, the more likely it is that air quality will be improved in the most cost-effective manner.

7. *Improve timeliness.*

Timely and historically consistent emission inventories are essential for assessing the current emission environment (and for keeping abreast of economic conditions and changes in technology) and for tracking progress in improving air quality. *Plans and processes need to be put in place for preparing and reporting national emission inventory data on a yearly basis.*

Timely inventory data more accurately reflect the actual situation. Ideally, inventories would be updated almost continuously so that air quality managers could see how the mix of emissions are changing due to external changes (population growth, changes in technologies, economic forces, etc.) and to changes in air quality management strategies. Continuously updated inventories remain far off, but a realistic goal is to minimize the time between when the inventory data are submitted and when they are reported.

8. Assess and improve emission projections.

Realistic projections of future emissions are important for developing and assessing strategies for attaining air-quality standards and for evaluating future-year effects of new regulations. *Emission projection methodologies for all emission inventory sectors should be evaluated to determine the accuracy of past projections and to identify areas for improvement. Attention should be paid to assuring the compatibility of short-range projections that are more typical of air-quality related emissions with the long-range projections that are made for climate-change applications. Projections of future emissions are also dependent upon the quality of the base-year emissions. Therefore, realistic projections cannot be made unless these base-year emissions are as accurate as possible.*

Because most air quality management approaches estimate the future effects of air quality management actions, it is important to understand what changes are due to the air quality management strategy and what changes are due to changes in population, technology, etc. Improved projections also provide guidance regarding the level of emission reduction that may be needed to achieve and maintain the desired air quality over the long term.

3.2.6 Further Issues Regarding Emission Inventory Improvement

Scale: There is a growing demand for reliable small scale emissions data. They are needed to model urban scale air quality for the purpose of improving population exposure estimates used in health research and in risk assessment and for control scenarios focused on local emissions (e.g., mobile sources, urban planning). At the local level this issue of scale is significant and a detailed understanding of source variability and how sources differ from the national average need to be included. This involves not only the differences in technologies (processes, control technologies, etc.), but also how the sources operate. If there are significant differences between the local practices and national practices, the inventories will need to account for them. In addition, small sources that are not important at regional and national scales and hence may not be required to report their emissions can be important at the local scale.

Area and mobile sources are generally more important at the local scale and such changes in the mix of source types can result in air pollution issues at the local scale being different from those at a national or regional scale. A good example is $PM_{2.5}$ – In the eastern United States sulfates are the key to $PM_{2.5}$ reductions at the regional scale. At the local scale, however, carbonaceous PM tends to be the issue of most concern, because it is locally emitted (traffic, industries, open burning, and biogenic sources) and because the timescale of SO_2 chemistry results in sulfur dioxide being a gas-phase problem locally, but a PM problem farther down wind.

Clearly, developing a local scale inventory is very demanding. Often such efforts are conducted by local groups based upon their own techniques and/or assumptions. There is a need to consider how such parallel efforts that may involve differing approaches could be standardized and improved, including the assessment of hot spot emissions, interfacing with exposure models, such as the Regional Human Exposure (REHEX) Model, and intake fraction parameters. This latter parameter is the fraction of the emissions of a pollutant taken in by people (Marshall et al., 2003) and provides a means of weighting sources according to exposure potential. Primary emissions in a rural area and/or from high stacks have a smaller intake fraction than those emitted at ground level in cities and arguably, for health protection the latter types of sources should be more important to control.

Groundtruthing Efforts: In current practice, inventories are most effectively used as the starting point for understanding the contributors to air quality problems. They form the basis for air quality modeling and for identifying the sources that are most significant within an airshed. Additionally, best practice dictates that receptor models, fuel-based (vs. vehicle-travel based) inventories, inverse modeling, and other approaches be used, independently, for inventory verification. This is important to undertake to identify problem areas in the inventories in order to minimize the impacts of

erroneous inventory data. This intercomparisons or evaluations may also help identify when the other, perhaps newer approaches are out-performing the traditional approaches, thus leading to new models, different forms of input data and better emissions inventories.

Costs of Emission Inventory Development and Improvement: The U.S. federal government currently invests approximately $25 million per year to develop and update emission inventories (NARSTO, 2005). This does not include the amounts spent by state and local agencies (estimated at about $10 million per year) or the additional costs that would be required to address the shortcomings identified in the current inventories, which are estimated to be on the order of an additional $35 million per year (NARSTO, 2005). In Canada, about $6 million (U.S.) per year is invested in compiling its national inventory, not counting local and provincial efforts, and Mexico has spent about $600,000 (U.S.) per year in developing its National Emission Inventory. The cost of addressing emission inventory shortcomings in Canada and Mexico are proportionally similar to those of the United States – about $6 million and $1 million, respectively (NARSTO, 2005).

The cost of developing emission inventories is a function of their purpose. The relatively low cost of developing Mexico's national emission inventory is a consequence of its relatively low level of detail and the use of previously existing information. At the other end of the scale, the Electric Power Research Institute spent $50 million to quantify emissions of hazardous air pollutants from electric generation units (EPRI, 1994). The American Petroleum Institute, the U.S. Department of Energy, and others spent about $6 million to measure combustion emissions from refineries. These higher cost examples reflect the greater expense of obtaining highly detailed information, such as speciation of organic and metal compounds, that are present in flue gasses at very low concentrations (NARSTO, 2005). The additional annual costs of addressing emission inventory shortcomings in Canada and Mexico would be about $6 million and $7 million (U.S.), respectively, over a period of 3-5 years.

3.3 Measurement of Ambient Pollutant Concentrations

Along with emissions data, ambient air pollutant measurements are part of the foundation of effective air quality risk management. While all measurements have the potential to play a role, the degree of support varies depending upon a number of factors. The type of measurements and the design of the measurement program are critical factors, but just as important is the amount of time and effort dedicated to analysis and interpretation of the data. When ambient pollutant observations and related measurements are examined to the full extent possible they can support multiple objectives, which includes:

- Describing current risks and detecting potential future risks to human and environmental health;
- Documenting trends in order to demonstrate efficacy of past and present policies (e.g., emission reductions);
- Developing models capable of predicting ambient pollutant concentrations from knowledge of emissions and emission changes;
- Providing information needed to derive quantitative relationships between ambient concentrations and human health (or other adverse effects on the environment, climate or visibility).

Measurement data and most often routine monitoring data are central to AQ-health studies and the concentration-response functions (CRF) derived from this research. Ultimately, measurements can lead to new insight into the specific pollutants or sources posing the greatest risk to health, which can help in the development of more efficient risk management strategies. This section summarizes a range of issues related to how measurements best support air quality risk management. Protecting human health is one of the main motivations and thus, much of the discussion is geared towards that aspect of air quality measurements. However, other issues related to the measurement objectives listed above are also discussed including general guidance on technical issues to be considered in establishing a robust measurement program.

3.3.1 Application to Health Studies

Exposure to air pollutants has the potential to lead to a variety of adverse health impacts (see Chapter 2). Not all such impacts are likely to have been identified and/or adequately characterized given the diversity of exposure scenarios, the diversity of the population and the myriad of possible biological pathways. A range of health research approaches and detailed air quality and exposure measurements are thus needed to continue to advance knowledge.

The World Health Organization published a European Series Report on "Monitoring Ambient Air Quality for Health Impact Assessment" (WHO, 1999). The goal of that report was to describe strategies and methods for providing information on ambient air quality that would be adequate for health impact assessment. Many issues that are relevant to this chapter are discussed and therefore, it is recommended background reading.

Concentration Response Functions

For air quality risk management, CRFs based upon direct links between ambient observations of a range of air pollutants, as measured by standard monitoring networks, and acute and chronic impacts occurring within the general population are crucial. Demonstrating such 'real world' associations also establishes that air pollutant effects are relevant to actual conditions. This necessitates that ambient measurements continue to be obtained to support both acute and chronic exposure health studies. The level of detail that these measurements should provide, including direct personal exposure studies or other extrapolations related to exposure (e.g., intake fraction, population-weighted concentrations), will depend upon the type of study they are intended to support. There is also considerable demand for better information on the specific pollutants, mixtures and/or sources that have the greatest impact upon health. This can only be satisfied by enhancing ambient measurement and exposure research activities and fully capitalizing on the recent advances in measurement capabilities (Wexler and Johnston, 2006). In addition, advances in statistical methods that can simultaneously exploit geographic and temporal variations in the interrelationships among pollutants coupled with new monitoring strategies should provide new insights regarding the most harmful pollutants or mixtures.

One of the key challenges in working with ambient pollutant data is to derive CRFs for individual pollutants and for pollutant mixtures that are appropriate for risk management or cost-benefit analysis. They must be scientifically defensible and their uncertainties and/or their strengths and weaknesses need to be understood in detail. Significant challenges arise from confounding and differential exposure error among pollutants. Among other things, an association between a health endpoint and a given pollutant may be because that pollutant is acting as an indicator for an unmeasured pollutant (gas or particle) or mix of pollutants (Brook et al., 2007). In this situation the relevance of a CRF derived from these data must be questioned and applications (e.g., cost-benefit analysis) must be done with clearly stated caveats. Research is needed to determine the true implications of such caveats and the appropriate interpretation of predicted health benefits.

Conceptually, CRFs should be the same among different urban populations given similar distributions of time activity and susceptible individuals and, within the range of uncertainties, this is generally the case for PM. However, there remains some variability and the causes have not been resolved. Differences in the air quality data used to derive the CRFs likely play a role in that there are typically variations between how the monitoring sites relate to the population (i.e., their location(s)) and to their actual exposures (e.g., prevalence of air conditioning, which influences indoor penetration, differs geographically). It is also likely that the nature of the confounding among pollutants (measured and unmeasured) and their exposure errors differs among locations or cities. These issues lead to uncertainty in regards to which CRFs should be used to guide risk management. Heterogeneity among risk estimates and possible causes were discussed during the NERAM V Colloquium (Samet, 2006).

For cost-benefit analysis it is not possible or reasonable to have separate CRFs for every city

or population of concern. Meta-analyses have therefore been undertaken to derive 'generic' CRFs. This approach assumes, however, that the air pollutant for which the meta-analysis is undertaken is truly the cause of the adverse effects as opposed to being an indicator for some unmeasured component in the pollutant mixture. If the latter is the case then it is feasible that the relationship between the 'indicator' and the true causative pollutant(s) could change among study locations and hence the 'mean' effect from the meta-analysis would not be appropriate. This issue could be relevant with respect to nitrogen dioxide (NO_2), for which an association with mortality has been detected in a number of cities (Stieb et al., 2003, Burnett et al., 2004; Samoli et al., 2006).

Ambient Measurements to Indicate Exposure

No matter how detailed, ambient measurements cannot reflect what a person or most members of a population are truly exposed to. These measurements are an indicator for some aspect of the air pollutant stresses that the population is confronted with. In reality, the relationship between these indicators and the actual exposures of the population may differ from pollutant to pollutant and from one monitoring site to the next. Site location relative to the population and selection of the pollutants to measure at each site are therefore critical issues.

Acute exposure-effect studies and chronic exposure-effect studies have different requirements regarding site location. The former requires that the measurements accurately reflect temporal variations in the population's exposure while the latter is interested in how exposure levels vary across space. This could range from differences from one location in a large city to another (intra-urban) or differences between cities (inter-urban). For both types of studies the ideal measurement data are rarely if ever available for both financial and technical reasons. Therefore, compromises are necessary and it is important to understand the limitations of and implication in using the available measurements for health studies or risk assessment.

The best practice for both acute and chronic studies is to measure multiple pollutants at any site that is established and to operate more than one site in the region containing the population of interest. European criteria for site coverage and representativeness is discussed in Kuhlbusch et al. (2004). In their examination of the networks reporting to AirBase they presented an example for the Ruhr district in Germany and assessed differences among countries examining the extent to which monitoring networks were being operated in accordance to the air quality directives with respect to protection of human health. In terms of site locations some of the key criteria to consider in network evaluation are:

- Sites are established which provide data on the highest concentrations
- The network comprises both, hot spot and urban background sites
- Hot spot sites are representative for at least 200 m²
- Urban background sites are representative for several km²
- Urban background sites are representative for similar locations not in their vicinity
- Sites are established which are representative for the exposure of the general population

The spatial distribution of the population, local physical features (e.g., topography, shorelines) and prevailing meteorological conditions are all important to study when selecting site locations or choosing sites to be used in health studies. With this information a better understanding of the link between each set of measurements and the population will be realized and population-weighted concentrations can be derived with more confidence. This perspective also helps identify and address weaknesses in the monitoring network, via new sites or special studies (e.g., saturation monitoring, deployment of mobile labs).

Gaseous air pollutant time series developed for epidemiological research (acute effects) on the Toronto population have been based upon the average hourly concentrations from 3-4 different representative sites (e.g., Burnett et al., 1997). Their locations reflect some of the main outdoor environments people commonly encounter, such as residential and commercial areas and near roadway conditions. In many European cities it has been common practice to measure in the urban background and at

"kerbside" or traffic locations to obtain information on the range of concentrations experienced (Kuhlbusch et al., 2004). Combining data from multiple monitoring sites, potentially using population-based weighting factors, will typically provide a time series that is a better representation of the exposure variations experienced by the population.

It is important to consider how long of a time series will be needed when a new set of measurements are initiated in support of acute effects studies. A rule of thumb is that 2-3 years of daily air pollutant measurements for a population of about 1 million will provide sufficient power for reasonably precise determination of the magnitude of an acute effect on total non-accidental mortality. Obviously, the more data the better, as shown in Volume 2 of the 2004 USEPA Criteria Document for Particulate Matter (PM). The precision of the risk estimates tends to increase as the product of the daily death rate and number of PM measurement days increases (EPA, 2004). Constructing a time series of data with values every day is not an issue when hourly measurements are available. When integrated samples collected over 24 hr periods (daily samples) are all that is possible (e.g., for measurements of some PM chemical constituents) then common practice is to collect samples less frequently than every day. For many years TSP and PM_{10} samples were only collected every sixth day in most circumstances. Although an improvement, many new networks are only providing samples every third day, which means it will still take three times longer to have sufficient statistical power for an acute effects or time series study.

Judging Site Representativeness

For the most part, comparisons of monitoring site data within cities have demonstrated that day to day variations in outdoor pollutant concentrations (i.e., the acute exposure signal) are reasonably-well correlated across a city. However, this varies by pollutant and needs to be assessed before the data are used in time series studies (i.e., acute health effect studies). For example, average inter-site correlations (i.e., for all available pairs of sites in Toronto in 2000-2003) for CO, SO_2, NO_2, O_3, $PM_{2.5}$ and PM_{10} are: 0.41, 0.78, 0.77, 0.90, 0.96 and 0.84, respectively. Similar correlations are present among stations in Vancouver and Montreal. The poorer correlation for CO compared to NO_2, which are both predominantly vehicle-relate pollutants in the city, can be explained by the uncertainty in the measurements as opposed to being due to more heterogeneity in concentrations. In Canada at least, improvements in the monitoring of CO or the treatment of the available data are needed before it should be considered in health studies.

For typical monitoring networks the pair of sites with the minimum correlation likely indicates a lower limit of the representativeness of using a single site to estimate population exposure. A low minimum correlation may also indicate that a given site is not appropriate by itself for use in a time series study. Table 3.3.1 shows how the minimum correlation varied by pollutant and season within a smaller and larger region of Toronto. Not surprising, the representativeness of a given site's measurements decreased with distance and $PM_{2.5}$ varied most consistently across the region. However, the degree of correlation for $PM_{2.5}$ is not likely to be the case for the different PM chemical constituents. Three weeks of two per day sulphate, organic carbon and elemental carbon measurements across 3-7 separate sites in the downtown core of Toronto showed that sulphate was homogeneous while the carbonaceous constituents varied considerably (Brook et al., 2002). Organic carbon for which there are many urban sources, from traffic to cooking, was the most variable. The minimum correlations in Table 3.3.1 reveal that a single site's NO_2 measurements have a greater potential to misrepresent the temporal variations experienced by the population. In Table 3.3.1 the lowest correlation between two sites was 0.08 (NO_2 in the warm season), suggesting that there is good potential for local conditions to influence the actual exposure experienced by members of the population residing further from the city center. A closer look at the site in question would be needed before a single measurement time series could be used for the larger Toronto domain.

Table 3.3.1 Minimum site to site correlation in daily average concentration variations across Toronto and outlying suburbs during 2000-2003. The Toronto correlations are for the group of sites within the core of the city (approximately a 20x20 km area). The metro area correlations include these sites plus sites in the suburban area (approximately a 40x40 km area). Warm corresponds to the warm season (May-September) with the cold season being October-April).

		Toronto	Metro area
$PM_{2.5}$	**Annual**	0.91	0.90
	Warm	0.93	0.90
	Cold	0.92	0.86
NO_2	**A**	0.56	0.29
	W	0.43	0.08
	C	0.67	0.40
SO_2	**A**	0.62	0.43
	W	0.53	0.23
	C	0.77	0.42
O_3	**A**	0.83	0.76
	W	0.77	0.69
	C	0.77	0.62

In general, the correlation in day to day variability across Toronto is reasonable for NO_2, O_3, SO_2 and $PM_{2.5}$. As suggested above, combining the time series from a group of sites can improve the representativeness of a time series. For example, for NO_2 the average and minimum correlation between the downtown Toronto site and the sites within the city are 0.83 and 0.75. These values increase to 0.89 and 0.86 when each of the sites is compared to the average time series generated by a combination of the Toronto sites. Similarly, when considering the suburban sites (i.e., those that are further away) the minimum correlation increases from 0.58 to 0.66 when the day to day time series is represented by the combined sites as opposed to just the central downtown site.

Simple examination of site to site correlations provides insight regarding representativeness of a time series and can help explain differences among pollutants in observed health effects from time series studies or in measured personal exposures. However, it is difficult to specify criteria for a minimum correlation value, below which the sites and pollutants being considered would not be appropriate for deriving exposures to include in a time series study. This issue requires further attention, including closer examination of site representativeness and the sensitivity of epidemiological analyses, particularly those attempting to consider multiple pollutants, to this source of exposure error. Logically, this error can be expected to attenuate the significance of true associations, but in multi-pollutant analyses there could be the potential to "transfer" the association to the pollutant that can be monitored at fixed sites with greater spatial representativeness and/or with a stronger link to actual personal exposures.

Links to Population and Personal Exposure

Within limits, monitoring site locations can be optimized to best represent the potential exposure of the population. When combining information from sites the data from each site can also be weighted according to the size of the population within a certain radius. Measurements at sites with a greater surrounding population are thus counted more

heavily. A potentially stronger link to exposure can be derived by considering intake fraction (Marshall et al., 2003). This approach is more appropriately used for weighting the pollutant exposure risks posed by emissions from a variety of sources. Thus, assigning an intake fraction weighting to an ambient measurement also requires knowledge of the sources contributing to the observations and how the measurement, the source location and the population of interest are related, including estimates of pollutant concentrations in a range of microenvironments. Probabilistic exposure models, such as the REHEX model (Winer et al., 1989; Fruin et al., 2001), also provide an approach for estimating, from ambient concentrations, the range of exposures expected among the population. Among other things, these models require more detailed spatial information on concentrations, either by interpolating the observations or through the use of air quality models.

Short term field studies can help identify optimum site placement and the nature of the relationship with personal exposures. There have been several exposure studies examining how well outdoor measurements correspond to actual personal exposures. Brauer and Brook (1997) showed that day to day variations in outdoor central site O_3 measurements agreed reasonably well with the temporal variability in average personal exposures, but not necessary for all people. In addition, the amount of agreement has been found to vary from city to city and among pollutants (Sarnat et al, 2001; Kim et al., 2006). The latter represents one source of differential exposure error.

$PM_{2.5}$ has been the primary focus of most recent exposure studies and the results generally show that each individual's day to day personal exposure time series is correlated with the variations in the outdoor levels. However, there is wide variability from person to person depending upon their time activity, the predominant indoor microenvironments that they spend time in and the indoor sources of $PM_{2.5}$ in these locations. Nonetheless, the median personal to ambient correlation among a sufficiently large group tends to be positive and strong enough to indicate that, on average, the exposure variations experienced by the population are captured by the monitoring site measurements. Attempts have also been made to separate personal $PM_{2.5}$ exposures into particles of indoor and outdoor origin (Wilson and Brauer, 2006). This has helped, for example, to assess if there are differences in the impact on respiratory health of these two broad classes of $PM_{2.5}$ (Ebelt et al., 2005).

Ideally, when developing a time series for an acute effects study or quantifying the exposure gradient in a cross-sectional study (chronic exposure effects) the link to personal exposure should be tested. To date, there have been very few exposure studies geared towards comparing chronic personal exposure to the outdoor concentration "assigned" to that individual (Wheeler et al., 2006a,b), which is usually the concentration estimated at their home.

Limitations in Linking Air Quality Measurements and Health Data

In addition to the exposure errors, it is also important to consider the nature of the health effects being studied and the type of health data that are or will be available. True acute responses likely arise immediately after exposure (i.e., within minutes to hours), but may linger and/or continue to grow in severity for at least a few days. This time course is not well understood, but can be expected to differ depending upon the physiological functions being affected (e.g., respiratory, cardiovascular), from person to person and depending upon how the exposure conditions change during this period. Administrative health data do not provide the time resolution needed to improve understanding of this etiology. Emergency room visits, hospital admissions, deaths, absenteeism, etc., are recorded as daily counts corresponding to midnight to midnight. While at a minimum this implies that only daily pollutant averages and exposure metrics derived from these values are necessary, hourly measurements can be utilized in the form of daily maximum readings over shorter durations. This permits some assessment as to whether the acute responses are associated more strongly with short term peak levels or higher concentrations sustained over 24-72 hr periods. Confounding among these

different metrics will hinder definitive conclusions about what acute exposure period is of most importance to the population at large. However, it may be possible to gain some insight for more detailed follow-up using different study designs (e.g., prospective panel studies) and thus, hourly measurement data are valuable to obtain.

There are relatively few health and air quality databases for chronic exposure studies. Thus, when health data relevant for air pollutant effect studies are assembled, the exposure information is most often derived from whatever is available. In this case, prospective air quality measurements can be useful if they can be shown to correctly reflect the past and present gradient in exposure within the cohort. In the rare circumstance that a cohort is developed prospectively it is more likely that air quality measurements can be included in the program (e.g., Spengler et al., 1996; Peters et al., 1999). Even so, demonstrating that such measurements are representative of longer term or even lifetime exposure remains an issue (Brook and Spengler, 1997) and mobility of cohort members may need to be taken into consideration (Jerrett et al., 2006).

Improving Air Quality Information for Health Research

More targeted air quality risk management will ultimately require more knowledge regarding the specific pollutants or sources posing the greatest risk to human health. Presumably, standards for these pollutants (e.g., a chemical component found on $PM_{2.5}$) could be established or control measures could be implemented on specific sources. Acquiring the necessary knowledge requires highly specialized health studies. These need to be supported with much more detailed ambient air pollutant measurements and a better understanding of the relationship between ambient concentrations and actual exposures. These are likely to come from field studies as opposed to monitoring programs. The types of measurements needed and other health research related needs that the atmospheric science community can potentially satisfy were recently discussed by NARSTO (www.narsto.org).

There have been considerable advances in measurement capabilities due to extensive research on $PM_{2.5}$ during the past decade. Applying these capabilities for health research can be expected to lead to new understanding (Wexler and Johnston, 2006). However, even stronger collaboration among health, exposure and atmospheric scientists will be needed to take full advantage of these advances given the level of understanding required to operate equipment and design effective studies. At the same time there is increasing interest in studying a wider range of health endpoints (e.g., cancer risks, impacts on the fetus and trans-generational effects) and a wider range of air pollutants (e.g., toxics, radon, aeroallergens). This will hopefully lead to a more holistic understanding and approach to risk management. However, achieving this level of understanding will be challenging, requiring a sustained multi-disciplinary research effort in environmental health.

As monitoring programs are established or expanded, enhancing the spatial coverage of the sites could potentially improve the applicability of the data for future chronic health studies. However, even with close contact with health researchers it is difficult to anticipate all the areas that might need to be covered unless there are specific cohorts in mind. Even still, resource limitations usually make it impossible to monitor all locations, especially if only done on the speculation that a health study might use the information.

Intra-urban variations in exposure are now recognized as an important signal to exploit in chronic studies and again, it is not likely possible to operate enough monitors. Mobile measurement platforms (e.g., Bukowiecki et al., 2002; Westerdahl et al., 2005; Yli-Tuomi et al., 2005; Kolb et al., 2004; Polina et al., 2004; Guo et al., 2006; Xu et al., 2006) can assist in studying spatial patterns and in optimizing site placement. They are becoming more common as a facility for monitoring agencies and current advances in technology are allowing even more-sophisticated measurements to be obtained. However, measurements alone cannot provide all the spatial detail desired. Thus, alternate sources of information and/or a range of spatial

models, statistical or physical, are becoming increasingly important to develop. These will be discussed below in Section 3.5.

3.3.2 Tracking Progress

One of the primary objectives of monitoring networks is to track progress towards achieving standards and to insure that good air quality is maintained. Most mature networks have multiple sites that are in standardized locations, several if not all of the criteria pollutants are being monitored and long data records have been and are continuing to be collected. Thus, they are well-suited to satisfy this primary objective. Beyond selecting the right locations and pollutants to measure, best practice clearly dictates that for studying trends a long, unbroken record is important to maintain. This is because air quality changes can be small and gradual and are obscured by meteorological variability. Lack of continuity in measurements, lack of sensitive enough measurements and under-representation in some geographic areas hinders trend analysis. Consequently, best practices also dictates that closing or moving monitoring sites with long records should be considered very carefully. When a site is moved both the new and old site should be operated simultaneously for as long as possible in order to quantify concentration differences.

Demonstrating progress in direct response to implemented policies is now being referred to as 'accountability'. Ultimately, this should extend beyond just documenting air quality improvements to demonstrating that the desired benefits have been realized. This would include, for example, improved public health, recovering ecosystems and fewer poor visibility events. However, tracing back along the full accountability chain is very challenging (HEI, 2003) and the implications it might have for how air quality measurements are undertaken have not been fully resolved.

If accountability is limited to detecting the expected air quality improvements resulting from a new policy or a specific intervention then time series length may not be as important as for detecting gradual trends. However, sufficient, high quality baseline measurements in advance of the emission reduction(s) are critical as is continuation of these measurements after the reductions have been implemented. Existing sites providing the base line data need to be identified and, if necessary, improved (e.g., adding high sensitivity instruments or new measurements).

Air quality models and emissions inventories have improved significantly over the past 10-15 years. This includes them being more widely applied (i.e., more accessible to a large number of users) and availability of sufficient computing resources for longer term simulations. Thus, there is an opportunity for models to play a greater role in informing monitoring activities. They can be used to simulate the magnitude of the change at the sites being used for tracking progress or demonstrating accountability. This may indicate that more sensitive measurements will be needed given the anticipated changes or that additional measurements at existing sites or at new locations would help to more-effectively and more-rapidly demonstrate accountability. Models can also be applied to identify the most useful measurement locations for "filling in" information between sites and, as will be discussed below, they can provide a more continuous picture of spatial patterns by merging their output with the measurement data. In some countries or jurisdictions, previous data and models have provided knowledge to allow downsizing of networks with limited or no loss of information (e.g., leaving sufficient sites to monitor trends and model population exposures), which helps reduce costs. Thus, given their current level of development, best practice clearly dictates that models be used as much as possible to optimize and expand the usefulness of air quality measurement programs.

A weakness of many existing networks is that they have focused more-extensively on urban areas. Consequently, there are much fewer rural measurements and the length of their time series is shorter. For example, worldwide, there are few long term trends on rural NO_x levels, despite its importance to understanding regional O_3 and $PM_{2.5}$ and the impact nitrogen deposition can have on ecosystems. Furthermore, this limits our current ability to assess how the growth in the size and density urban areas (e.g., sprawl) is impacting proximate regional air quality. Urban-

rural pairs of measurements (Brook et al., 1999) are becoming more common. They are essential for untangling contributions and trends attributed to local/urban sources vs. upwind sources from regional scale transport. Conversely, some classes of pollutants have traditionally been measured in rural and remote areas (e.g., POPs, acid deposition, sulphate) and thus, understanding of the conditions in populated areas is more-limited.

More and more agencies responsible for air quality measurements are enhancing activities at selected sites (supersites or AQI sites or core sites) and potentially collecting fewer measurements at sites in between these locations. This approach supports a greater amount of air quality science at the supersites and can provide cost savings. The level of activity at the supersites often varies over time depending upon funding and the current "hot issues." However, stability for a core set of measurement is important since these supersites are likely to be ideal for studying long term trends and potentially for accountability purposes. The measurements obtained at the "in between sites" or "satellite sites" typically depends upon the objectives of the program funding the work. With respect to air quality and smog, such sites are most likely to monitor O_3 and $PM_{2.5}$ or PM_{10}. This is driven by the fact that these pollutants are closer to or over existing standards and there is a perception that they are most important with respect to health effects. While this latter point is valid, in many areas these pollutants exhibit less spatial variability implying that fewer sites are needed. This is, perhaps, less likely for $PM_{2.5}$ because although its total mass may vary relatively slowly over space some chemical constituents, particularly those related to primary emissions, will potentially vary much more rapidly. Thus, more knowledge is needed to in order to determine how best to optimize the number of $PM_{2.5}$ sites and $PM_{2.5}$ speciation sites.

3.3.3 Modeling, Process Studies and Source Apportionment

Identifying specific management strategies requires a good conceptual model of the causes of the air quality problems. Both monitoring and detailed field study data are necessary to develop this conceptual model and to obtain a greater understanding of atmospheric processes. Spatial and temporal coverage is the greatest asset of monitoring network data. Combined with meteorological information (e.g., trajectories) relatively complete conceptual models can be devised and considerable information about contributing sources areas can be obtained (e.g., Brook et al., 2005). As highlighted by discussion and examples in Chapter 6 of the NARSTO PM Assessment, all forms of data analysis, from the simplest to the most complex, will provide insight. However, this requires that resources be continually dedicated to this purpose and that experienced analysts, with backgrounds in atmospheric chemistry, meteorology and statistics, are employed.

Monitoring network data can also play a valuable role in model evaluation. This role is becoming more important because many models are being run continuously for AQ forecasting and detailed field studies cannot be undertaken continuously. These new long term modeled datasets are offering new opportunities to learn a great deal about how the models perform and the quality of the emissions information. While the network data tend to lend themselves more to operational evaluations of the model, there are opportunities for diagnostic evaluations as well (see Section 3.3). Increased use of continuous or semi-continuous particle composition instruments for monitoring can be expected to provide greater opportunities for such evaluations. Network data are also critical to define the model's initial conditions. Advances in rapid data assimilation (Menard and Robichaud, 2005) have been occurring for this purpose. These need to be continued since there are many potential applications of these assimilated datasets.

More detailed measurements, which can only be sustained for relatively short field studies are ultimately needed to study atmospheric processes (dynamical, chemical and physical) and for more-detailed diagnostic model evaluations. The scope of these studies can vary greatly from a small team collecting measurements to study one process (e.g., Makar et al., 1998) related to one model module to

large collaborations across institutions (e.g., EMEFS, ICARTT; Dennis et al., 1993; Frost et al., 2006). Just as there has been advances in what can be reliably measured at monitoring sites there have been significant gains in what can be measured during field studies. Methods have improved and some advanced technologies are commercially available and are reasonably reliable and straightforward to operate with highly qualified personnel. This means that the ease with which a field study with highly comprehensive and technical measurements can be launched is now much greater than 5-10 years ago. This is beneficial for obtaining, quickly, much more data at more locations from which one can study, in detail, source apportionment and atmospheric processes. However, the risk has never been greater for valuable data to be under-interpreted. Best practice obviously dictates that this be avoided. Thus, with these new capabilities comes an increased need for highly qualified and creative experts in data analysis along with careful advance thinking regarding the underlying hypotheses motivating any venture into the field.

3.3.4 Public Information

In the long run informing the public is critical to the process of air quality management. They are the ultimate decision-maker since when a large enough majority decide that an issue is important elected officials are more likely to respond. This can create opportunity for progress on air quality. Science/environmental advisors must be ready with the right advice, based upon emissions, measurements and modeling information, when government leaders are prepared to make decisions. In recent years, health research results have garnered considerable attention in the media and, as discussed above, air quality measurements are critical to this research.

In terms of public information, the Air Quality Index (AQI) has been utilized for many years. In cities or regions where there are more frequent bad air days the public is generally more aware of AQ issues, at least partly due to the publicity of the AQI. The form of the AQI is similar in many countries, reporting air quality using descriptive terms such as good, moderate, poor, very poor, unhealthy, etc., (www.msc-smc.ec.gc.ca/aq_smog/aqcurrent_e.cfm; http://airnow.gov/index.cfm?action=static.aqi). This approach is easier for the public to understand and act upon as opposed to reporting actual pollutant concentrations. It has been designed to identify the worst effects that may result from the mixture of pollutants currently being measured and to describe the prevailing air quality. However, ozone and particulate matter are more often the driving pollutants (i.e., leading to an AQI other than very good or good). The AirNow website (http://airnow.gov/), which reports actual ozone and $PM_{2.5}$ concentrations all across the U.S. and Canada along with colour codes indicating the AQI, represents a significant advance in the information being publicly provided. It allows for the spatial extent of elevated air pollutant levels to be visualized and for users to quickly compare their region to others. The systems developed to obtain and synthesize this information and present it in near real time have only recently become possible and this infrastructure is also critical for improved AQ forecasts. Similar systems for providing realtime AQ data and/or AQI values exist for many countries (e.g., the Netherlands: www.lml.rivm.nl/data/smog/index.html; Mexico: www.ine.gob.mx/). These are continually evolving and being integrated into multi-national systems.

The increase in knowledge regarding air pollutant health effects has been leading to growing interest in upgrading or modernizing the AQI. One such program is the Air Quality Health Index (AQHI) being developed in Canada. This is currently being pilot tested in the Province of British Columbia (http://www.airplaytoday.org/). The goal of this pilot is to introduce the AQHI to the public and gather feedback, especially from people who are sensitive to air pollution. The unique feature of the AQHI is that it is based upon recent epidemiological results from across Canada. In addition, it considers multiple air pollutants simultaneously and they all contribute to the index value in every case.

The AQI has typically been for reporting current conditions so that the public can respond

immediately. In recent years, however, the capacity to predict future O_3 and $PM_{2.5}$ concentrations, using physical and/or statistical models, has improved. Thus, the public can be informed in advance. There are also efforts being planned to predict more of the pollutants considered in the AQI or AQHI (e.g., SO_2, NO_2). Consequently, air quality advisories are being supplemented with daily air quality forecasts (www.msc-smc.ec.gc.ca/aq_smog/aqforecasts_e.cfm; www.msc-smc.ec.gc.ca/aq_smog/chronos_e.cfm), but the information is not usually widely distributed (e.g., radio, television) until high levels, warranting an advisory, are predicted. Thus, despite the value of daily information for susceptible individuals, it is unclear if it is influencing the general public. Daily forecasts are also available in the U.S. (www.arl.noaa.gov/ready/ozone/), where some media are publishing maps every day. Countries outside of North America are also providing forecasts (e.g., www.epa.vic.gov.au/Air/AAQFS/default.asp).

Regular air quality forecasts, made possible through real time reporting and assimilation of measurements and advances in air quality modeling, clearly does present an additional opportunity, beyond advisories, to inform the public of air quality issues (i.e., increase public awareness). A telephone-based survey undertaken after the original air quality advisory program for Canada had started indicated there was partial success in achieving this goal (Stieb et al., 1996), but few actually changed their behaviour. Ideally, when public information programs are planned there should be some collection of baseline data, as opposed to retrospectively initiating such activities. This should give a truer picture of how the public's awareness and/or behaviour changed.

Ultimately, poor air quality situations need to be minimized through preventative measures (i.e., new policies on emissions or activities producing emissions) and increased public awareness help create the climate for political action. However, providing routine, reliable and understandable air quality information to susceptible members of the population allows them to reduce their own exposure. This should not be underestimated as an important component of air quality risk management. Therefore, communication plans and health messages require careful consideration and regular evaluation for effectiveness.

3.3.5 Technical Issues in Establishing a Measurement Program

Air pollutant measurements should only be taken if there is an ongoing commitment to a recognized standard of quality, and a plan for data archival and for interpretation. Poor quality and/or incomplete data or data of unknown quality have limited usefulness. It must also be recognized that knowledge and technology are continually improving and thus, to the extent possible, new measurements should seek to use the most current, accepted methods. This will increase the probability that the data are acceptable far into the future. In the long run, a small amount of high quality measurements will be of more value than many measurements collected with insufficient documentation, quality assurance and interpretation.

Air quality measurement activities generally fit into one of two categories:

1. Monitoring - A core set of systematic measurements at well-selected locations that are maintained indefinitely for trend analyses (i.e., evaluate effectiveness of current policies), to determine if an area is complying with or achieving an official air quality standard or guideline and to identify emerging problems as soon as possible, which may involve ongoing environmental health studies (e.g., epidemiological studies).
2. Field studies - A relatively short period (<2 years) of more-detailed or more-specific measurements collected within a well-defined geographic area or at a given location or for a given population. These data are essential for development of conceptual models, source-oriented models, more-refined source apportionment studies and for understanding the relationship between emissions, ambient concentrations and personal and/or population exposure. A variety of prospective health studies may also derive their exposure information from air quality field studies.

Data from monitoring programs lend themselves to a standard set of reports documenting current conditions, trends and comparison with other geographic areas. Quick release of such information into the 'right hands' helps to keep air quality management issues in the forefront. A range of user-friendly software tools that can process the air quality data along with meteorological data are becoming common. This enables air quality scientists and managers to examine some of the causes of pollution events in near real time, providing information that, in the past, could take a year or more to obtain.

Field study data are usually more complex and less standardized. Sufficient time and resources need to be dedicated to working with the data after the study. At a minimum, 1-2 years is likely necessary. To guide the planning of the study and subsequent data analysis there needs to be a set of testable hypotheses in place well before the study begins. Generation of new scientific knowledge is likely part of the study objectives and thus the first official reporting of the data tends to be in peer-reviewed journals. This is the only approach for ensuring credibility of the data, which is ultimately necessary for air quality managers to be confident with subsequent policy decisions. Prior to journal publication, preliminary results often appear at conferences and before that formal workshops help maintain momentum and offer an opportunity to combine data and build consensus. Air quality managers or their advisors should participate at this stage in order to stress the policy issues they are expecting the results will help inform.

Air quality measurement programs are expensive and so in advance, must have clear short and long-term objectives. Ideally, the program will be sufficiently flexible and broad to support multiple objectives, some foreseen and others not yet appreciated. Additional resources/expenses to insure data completeness, quality, analysis, interpretation and reporting should not be overlooked since they are likely to be incremental (i.e., a small cost relative to the overall cost of obtaining and maintaining the data). The key technical issues to consider when establishing a measurement program are:

- What to measure and how often
  - Ideally, multiple pollutants should be measured at the same site to assist in interpretation and to serve more than one objective
  - Temporal resolution; could range from seconds to days
  - Measurement methods to be utilized
  - Length of time series to be collected
  - Personnel needs in the field and lab/office
  - Criteria for introduction of new technology or additional pollutant measurements if the measurements are part of a longer term program.
- Siting criteria and where to measure
  - Impact of local sources - may be desired or important to avoid
  - Type of sites include: source-oriented, such as curbside or other high impact areas (e.g., hot spots); local or neighborhood scale; urban background; regional background (indicative of long range transport)
  - Representativeness to population and/or to region needs to be assessed
  - Geographic coverage and spatial density of sites if the program involves a network
  - Site access, serviceability and security
  - Documentation of site meta-data
- Quality assurance (QA)
  - Required level of accuracy, precision and data completeness
  - Frequency of collection of specific QA measures (e.g., duplicates, blanks, zeros, spans, calibrations, external audits)
- Data archiving and reporting
  - Data turn-around time and policy for data exchange and criteria for permitting use in publications
  - Provision of data to national or international public archives
- Consistency with methods used at other sites in the same and different networks and between countries

Addressing these technical issues, which will be expanded upon below, before measurements

start helps ensure that the data are of greatest value. In establishing a measurement program it may also be relevant to consider the potential applicability of the data for evaluating models and also for integrating the data with model output and other information to improve the detail and coverage of ambient concentration information (see below). Linkages to personal exposures and the ability to quantify the degree of exposure error associated with using the measurements for health research may also be necessary to consider.

What to measure and how often

For a wide range of both gaseous and particulate pollutants Chapter 5 of the assessment published by NARSTO (McMurry, Shephard, and Vickery, 2004) provides considerable detail on what can be measured, the methods available, how reliable they are and reasons such measurements might be needed (e.g., for health effects studies, compliance monitoring, visibility, scientific understanding, etc.).

When feasible, greater frequency of measurement (i.e., finer time resolution) is preferred because this permits a much better understanding of source contributions and atmospheric processes (Wexler and Johnston, 2006). If a standard exists then its 'form' or 'metric' (e.g., hourly maximum, 8 hour maximum, 24 hour, annual) will dictate that a certain resolution be achieved. Inclusion in a real time reporting program, such as may be needed for air quality index and air quality advisory purposes or for air quality forecasting, will also likely demand that data be available on a frequent basis (e.g., hourly). Choice of resolution also has an impact upon the resources needed for QA, data storage, as well as data analysis and interpretation activities. Ultimately, the time resolution that is measured is dictated by instrument capabilities.

Although air quality standards or other types of regulations/guidelines require that several common pollutants are monitored indefinitely, several other pollutants or trace gases are important to measure in support of air quality management. This wide range of trace atmospheric chemicals can be classified in a variety of ways. Here we choose to consider four classifications, however with any such attempt, the distinctions are blurred. These are:

- Pollutants formed during combustion
- Pollutants released from the surface or fugitive releases
- Volatile organic compounds (VOC)
- Secondary pollutants.

Table 3.3.2 provides a summary of pollutants under each category and highlights important requirements and considerations for their measurement. Arguably the largest group, in terms of quantity of emissions, contain pollutants associated with combustion emissions. Many of these pollutants are produced and emitted simultaneously, which presents opportunities for co-management. This includes nitrogen oxides (NO, NO_2 or NO_x), carbon monoxide (CO), fine particles ($PM_{2.5}$), ultrafine particles ('ultrafines' or $PM_{0.1}$) and, depending upon the presence of sulphur in the fuel, sulphur dioxide (SO_2). The first two, plus SO_2 and some form of particles (e.g., total suspended particulate-TSP, particulate matter less than 10 μm in diameter-PM_{10}) are generally referred to as criteria pollutants.

Particles are monitored according to total mass below a specific size, but in the case of ultrafines measurements are based upon total number. Combustion particles are also composed of many different chemical compounds, which will be discussed below. However, one important particle phase constituent that should be explicitly included in this group is black carbon (BC). Also, referred to as elemental carbon or soot, BC is being measured more frequently due to its strong link to traffic emissions, especially diesel particulate matter, which is becoming more and more recognized as posing a risk to health. In Germany, the National Environment Agency has implemented a BC ambient concentration limit of 8 $\mu g/m^3$ (arithmetic annual average value) and regulations in other jurisdictions (e.g., California) are being considered or are in place).

Table 3.3.2 Air pollutant classes, measurement capabilities and other issues to consider.

Pollutant Class	Examples	Measurement Capabilities	Comments
Combustion Emissions	Nitrogen oxides (NO_x or NO_2)	Can be measured with instruments that provide hourly or better time resolution	• Off the shelf use of most instrumentation will not correctly measure NO_2 due to interference from other forms of oxidized nitrogen such as nitric acid, particle nitrate and peroxyacetyle nitrate (PAN). • Unknown inlet losses from some of these species leads to additional uncertainty • Relative size of this interference increases further away from high NO emissions areas (e.g. large cities) and when atmosphere is more photochemically active (i.e. summertime) • The nature of this interference should be understood before reporting and using NO_2 concentrations.
	Carbon monoxide (CO) Sulphur dioxide (SO_2)	Can be measured with instruments that provide hourly or better time resolution	• Measuring the low concentrations present in many areas requires higher sensitivity instrumentation. • Since these low levels are below standards network managers are tempted to stop measurements or pay less attention to the quality of the low concentration valuests. This greatly hinders the use of these data in studying atmospheric processes, source apportionment and health effects.
	Carbon dioxide (CO_2)	Can be measured with instruments that provide hourly or better time resolution	• Links climate issue (i.e., greenhouse gas emissions) and air quality issue
	Fine particles ($PM_{2.5}$)	Total mass can be measured with hourly or better resolution, but the techniques available have limitations. Integrated sampling on pre-weighed filters is the most widely accepted approach, but is not without uncertainties. Instrumentation now exists for automated, semi-continuous measurement of the main chemical constituents.	• A portion of $PM_{2.5}$ is semivolatile and this is the main cause of the measurement uncertainty • Main contributors to $PM_{2.5}$ mass are organic carbon compounds (40-60%), sulphate (20-50%), nitrate (0-50%) and ammonium (5-15%). • The latter three have been successfully measured in many locations and countries via filter sampling and laboratory analysis. • Measurement of the carbonaceous material, which is typically separated into elemental (EC) and organic carbon (OC), is more uncertain. A significant difference exists among the different approaches and among different laboratories determining OC and EC • Sample collection artifacts caused by semivolatile OC and relatively high blank filter concentrations for OC lead to considerable uncertainty.

Pollutant Class	Examples	Measurement Capabilities	Comments
	Black carbon (BC)	Semi-continuous measurement approaches are commonly used. The simplest measurement technique is based upon light attenuation once a filter collects sample over a pre-determined length of time. British smoke, coefficient of haze and soiling index are predecessors to the present day instruments measuring BC via light absorbtion.	• When BC is measured from filter samples using thermal techniques it is typically referred to as EC and OC is often measured at the same time. • The distinction between OC and EC is operationally defined and method specific. • For BC or EC there is strong link to traffic emissions especially diesel PM
	Ultrafine particles (UFP)	Measurement is based upon the total number of particle counts (per cubic centimeter) for all sizes below 0.1 μm in diameter. A condensation particle counter (CPC) operated with no specific size separation at the inlet essentially measures UFP because particles below 0.1 μm completely overwhelm the remaining counts above this size (i.e., for all the rest of the particles from 0.1 to ~100 μm) One minute or better time resolution is possible	• The smallest particle size present varies between 0.005 μm (5 nanometers) and 0.02 μm (20 nano-meters) and so it is important to know the smallest detectable size and corresponding count efficiency for each CPC utilized for UFP measurement. • Electrostatic classifiers upstream from the CPC yields information on numbers of particles within many size ranges (i.e., particle size distribution). • Typical systems (scanning mobility particle sizers – SMPS) can discriminate 32 or more size ranges or bins from 0.01 μm to 0.9 μm, providing a size distribution every 10 minutes or better.
Non-combustion surface or fugitive releases	A ammonia (NH_3) Methane (CH_4)	Continuous or semi-continuous methods for NH3 are available.	• Ammonia plays a role in $PM_{2.5}$ formation therefore improved knowledge of its behaviour and emissions is needed for $PM_{2.5}$ management.
	Pesticides (persistent organic pollutants – POPs)	Low frequency filter and PUF based samplers are commonly used. Daily samples are very rare.	• Monitoring requirements may be country-specific and/or specified in international agreements (e.g. Stockholm Convention requirement for POP monitoring).
	Resuspended dust ($PM_{10-2.5}$ or $>PM_{10}$)	$PM_{10-2.5}$ is traditionally included as part of PM_{10} or TSP measurements. Filter-based technique are typically used, but continuous $PM_{2.5}$ instruments can be adapted for $PM_{10-2.5}$	• Data limitations hinder determination of the health risk posed by $PM_{10-2.5}$ and feasibility of managing ambient concentrations • Sources include road dust, wind blown dust from agricultural practices and from industrial facilities (e.g. mining and smelting), bioaerosols (e.g., pollen, spores)
	Total reduced sulphur compounds (TRS) or H_2S	Very few of these pollutants are measured routinely and the techniques are widely variable depending upon the compound	• Releases from industrial processes and waste management • Examples of sources are pulp mills, sour gas flaring and certain agricultural practices • Data are also limited due to a lack of regulatory requirements for monitoring and/or the localized nature of the emissions

Pollutant Class	Examples	Measurement Capabilities	Comments
Secondary pollutants	O_3	Can be measured with instruments that provide hourly or better time resolution	
	~ 60-70% of $PM_{2.5}$	Instrumentation now exists for automated, semi-continuous measurement of the main chemical constituents.	• Sulfate, nitrate and ammonium is essentially all secondary. • Secondary organic aerosols varies in amount from near zero to 50% or more of the OC.
	Some VOCs and SVOCs	Semi-continuous to nearly continuous measurements are possible for some compounds using research grade measurement methods. Integrated samples with canisters or traps are common.	• Even the more routine approaches for measurement for compounds such as formaldehyde and PANrequire highly trained operators and/or capable analytical laboratories.
	Other oxidants (e.g., H_2O_2, OH)	Research grade measurement methods are possible	• Measurement can be very important for understanding atmospheric chemistry. • They are only undertaken during detailed research studies. • Measurement is difficult as methods are expensive, experimental and challenging to implement and some of the compounds of interest are very short-lived in the atmosphere.
Volatile and semi-volatile organic compounds (VOCs and SVOCs)	100s of individual compounds Examples: Benzene Tolulene Xylene 1,3, Butadiene Isoprene	Total non-methane hydrocarbons (NMHC) can be measured with hourly or better resolution. Semi-continuous measurement methods are possible for some individual VOC and SVOC compounds. Air samples can be collected in the field and analyzed with a good degree of accuracy and precision for the lower molecular weight compounds (i.e., fewer than about 10 carbon atoms). Larger molecules and the more-oxygenated species are more difficult to measure with confidence. Sorbant traps are more commonly used.	NMHC provides limited information since the compound(s) responsible for the higher concentrations (i.e., the dominant fraction of the total VOC) and/or the temporal variation will vary and cannot be discerned. In addition, inlet losses, which are not necessarily the same for all the VOC and SVOC species, adds uncertainty to the measurement. VOC and SVOC measurement is a labour-intensive and expensive. In many cases the dominant species are not those of greatest interest from the standpoint of either atmospheric chemistry (i.e., O3 or PM2.5 formation) or toxicity. Commercially-available systems for time-resolved, real-time measurement of some compounds require highly trained operators and/or very capable analytical laboratories

Carbon dioxide (CO_2) is typically the most abundant trace gas in combustion emissions thereby linking, at the source, the issues of air quality and climate change (note that BC and aerosols are also important pollutants regarding climate). All of the pollutants mentioned above can be measured with instruments that provide hourly or better time resolution and that are capable of reliable, realtime data storage and transmission. With the exception of the particle-related measures, regular, automated QA or calibration can be included in monitoring routines. Calibration of the particle instruments is generally not possible because absolute standards are not available. Other approaches to insure data quality are thus, necessary.

Another relatively distinct class of pollutants of importance to measure can be characterized as being associated with non-combustion surface or fugitive releases. This includes both those related to human activities and natural emissions. For many of these compounds current capabilities do not permit high frequency measurements. Agricultural emissions such as ammonia (NH_3) and methane (CH_4), as well as dust, pesticides (persistent organic pollutants - POPs) and bioaerosols fit in this class. Releases from a range of industrial processes and waste management are also important sources. In addition to the pollutants listed just above, reduced sulphur compounds (e.g., H_2S) best fit into this class of pollutants. Similarly, resuspended dust, which is typically in the coarse ($PM_{10-2.5}$ or PM_{coarse}) particle and giant particle (i.e., $>PM_{10}$) size ranges, best fits in this class. This includes road dust, wind blown dust from agricultural practices and from certain types of industrial facilities (e.g., mining and smelting).

Very few of these pollutants are measured routinely due to the lack of reliable, cost-effective techniques, because there are no air quality standards that necessitate monitoring and/or because they are only a problem in localized areas. For example, total reduced sulphur compounds (TRS) or H_2S can be serious issues in the vicinity of pulp mills, sour gas flaring and certain agricultural practices. In addition to country-specific air quality standards, international agreements often necessitate some level of monitoring. For example, POPs are routinely measured at several locations in support of the Stockholm Convention (www.pops.int/).

With respect to current air quality risk management issues, ammonia (NH_3) and coarse particles ($PM_{2.5-10}$ or PM_{coarse}) are considered to be the most important to monitor or otherwise gain more information on their levels and spatial and temporal variation. Ammonia plays a role in $PM_{2.5}$ formation and thus, improved knowledge of its behaviour and emissions is needed to manage $PM_{2.5}$. PM_{coarse} has traditionally been included as part of PM_{10} or TSP measurements. However, as the monitoring focus shifts to $PM_{2.5}$ the need to continue to manage the coarse particle fraction is becoming an independent issue. At present, data limitations hinder determination of the health risk posed by PM_{coarse} and assessment of the feasibility of managing its level, especially given the range of natural and anthropogenic sources that contribute.

Secondary compounds, which are distinguished by the fact that they form in the atmosphere, represent an important class of air pollutants to measure. A large fraction of the chemical constituents found on $PM_{2.5}$ (fine particles) are secondary. This includes sulphate, nitrate, ammonium and some organic species. In the gas phase, ozone (O_3) is the most well-known and commonly measured pollutant. However, there are several other secondary oxidants, acids or VOCs of importance, either because of their potential to have health or environmental effects or because of their role in atmospheric chemistry. Examples are hydrogen peroxide, the hydroxyl radical, PAN, nitric acid, nitrous acid, hydrochloric acid, formic acid, acetic acid, formaldehyde, acetaldehyde, 1-3 butadiene and acrolein. Some of these may be emitted directly (i.e., primary pollutant), but atmospheric formation is likely the most important source.

Although O_3, $PM_{2.5}$ and some specific secondary constituents of $PM_{2.5}$ (e.g., sulphate and nitrate) can be measured with relative ease, it is much more difficult for most of the other secondary compounds. The methods available are expensive, experimental and challenging to

implement and some of the compounds of interest are very short-lived in the atmosphere. Yet their measurement can be very important for understanding atmospheric chemistry. Consequently, when possible, measurements are undertaken during detailed research studies. Even the more routine approaches for measurement, such as exist for compounds such as PAN and formaldehyde, requires highly trained operators and/or very capable analytical laboratories.

Volatile and semivolatile organic compounds (VOCs and SVOCs) are associated with both combustion and fugitive emissions as well as secondary formation and are thus part of all of the three classes above. SVOCs are also found in both gas and particle phase depending upon their properties and ambient conditions. However, due to the large number of compounds, their complexity, the challenging nature of their measurements and the tendency for many to be toxic, VOCs and SVOCs are considered here to be a separate class of air pollutants.

Total gas phase non-methane hydrocarbons are often measured with relatively simple instrumentation at routine monitoring sites. However, such measurements provide a limited amount of information since the compound(s) responsible for the higher concentrations (i.e., the dominant fraction of the total VOC) and/or the temporal variation will vary and cannot be discerned. In many cases the dominant species are not those of greatest interest from the standpoint of either atmospheric chemistry (i.e., O_3 or $PM_{2.5}$ formation) or toxicity. In addition, inlet losses, which are not necessarily the same for all the VOC and SVOC species, adds uncertainty to the measurement. It is important to distinguish between CH_4 and the remaining VOCs because the former is often found in much higher concentrations and behaves differently in the atmosphere (e.g., has a much longer lifetime).

Measurement of individual VOC and SVOC compounds is necessary to provide insight into their contributions to O_3 and $PM_{2.5}$ formation. In addition, they are useful for source apportionment and in order to characterize 'hot spots' of high exposure to toxics. Air samples can be collected in the field and analyzed with a good degree of accuracy and precision for the lower molecular weight compounds (i.e., fewer than about 10 carbon atoms). Larger molecules and the more-oxygenated species are more difficult to measure with confidence. In all cases, VOC and SVOC measurement is a labour-intensive and expensive. Although there are commercially-available systems capable of time-resolved, real-time measurement of some compounds, these require highly trained operators and/or very capable analytical laboratories and careful consideration of the uses of the data and their subsequent storage is necessary.

Particle Composition

A large amount of information on $PM_{2.5}$ mass measurement and $PM_{2.5}$ sampling and chemical analysis is provided by Chow (1995). Since that publication there have been significant advances in semi-continuous measurement of nitrate and sulphate, OC and EC. These measurements, although more challenging than semi-continuous mass measurement and than the traditional approach of using filters and laboratory methods, are providing new insights in the sources of $PM_{2.5}$. Some discussion on the various technologies being developed and applied is included in the NARSTO Assessment (McMurray, Shepherd and Vickery, 2004) and an overview of some of the new insights these instruments have enabled can be found in Wexler et al. (2006).

Organic carbon is one of the most challenging aspects of $PM_{2.5}$ measurement. As indicated above, sampling leads to uncertainties. The other main difficulty is that not all of the specific chemical compounds contributing to the total OC are known. In general, only about 20% of the OC can be consistently identified. These are chemical species such as polycyclic aromatic hydrocarbons (PAHs), alkanes and a large variety of organic acids. There is a considerable amount of ongoing research on the chemical speciation of specific organic compounds and much more data, although far from routine, now exist. There is a growing body of evidence of particle surface chemistry contributing to the uptake of organic mass and altering the chemical nature of the compounds present. Such processes need to be understood much better, as does the

emissions of OC and the contribution for natural sources (primary and secondary OC) before OC can be understood sufficiently to model specific control strategies. Thus, organic material on $PM_{2.5}$ currently represents one of the greatest and most-important challenges to the scientific community.

Siting criteria and where to measure

Location of measurement and the immediate surroundings have a large impact upon the concentrations observed. The monitoring objectives play a large part in dictating the types of locations that are desirable. This typically leads to more than one network being necessary to support all issues. The USEPA describes four categories of networks and sites and lays out the general purposes of the overall program and of each network on their website: www.epa.gov/air/oaqps/qa/monprog.html. Similarly, a range of site categories, their purposes and criteria for inclusion in EuroAirNet are described at: http://air-climate.eionet.europa.eu/databases/EuroAirnet/euroairnet_criteria.html.

Standardization is important to consider for monitoring networks and this because more challenging when they span multiple countries. Clearly, strict siting criteria need to be adhered to. For example, criteria for particulate monitoring in British Columbia are detailed at: www.env.gov.bc.ca/air/particulates/amgv1pnc.html.

On a broad geographic scale sites can be classified as either being remote, regional, urban background, urban exposure hot spot or industrial. The spatial scale or 'footprint' that each type of site can represent clearly varies from being nearly continental for remote sites to less than neighbourhood scale for industrial sites. Kuhlbusch et al. (2004) show a breakdown, by country, of what type of station - traffic, industrial or background; monitoring sites contributing to the European AirBase dataset fall into. There is considerable variability among the 21 countries examined in the distribution of station type and with respect to what fraction is urban, suburban or rural. Sites supporting air quality risk management typically fall within the urban exposure hot spot to regional scale. The latter type of site is best situated to provide information on transboundary transport as well as the general conditions in the rural area surrounding, but upwind of nearby urban areas (i.e., what is blowing into our cities from sources upwind). Ideally, such a site will be representative of at least a 20,000 km^2 area. The former type of site is best typified by measurements in high traffic areas. These urban exposure hot spot sites should be representative of conditions that the general population is exposed to on a regular basis and/or of the conditions over a neighbourhood experiencing higher concentrations due to the amount or type of emissions in the vicinity. Environmental justice issues are leading to an increase in the interest in studying high exposure neighbourhoods. With the exception of industrial sites, the specific location of measurement should not be directly impacted upon by local emissions or by nearby obstructions to wind flow. The definition of local varies from the urban background to the regional sites ranging from about 2 km to 25 km, respectively. Obstructions could be adjacent buildings in an urban area or the edge of a forest clearing at a rural or regional site.

The height of measurement is also an important siting criteria. Closer to the surface the measurements can be impacted upon by local dry deposition or local surface emissions or resuspension of dust. At regional and remote locations, 10 m is often the standard height as long as there are no nearby (~50 m) obstructions. In urban areas the concept of breathing zone enters into consideration, but there is a much greater risk that a very local emission will influence such a low measurement. Furthermore, low nighttime mixing heights tend to enhance this effect. Breathing zone measurements may be desired for specific exposure field studies, but are less likely to be representative of the urban background and, as well, not likely of the conditions occurring within a hot spot. Alternatively, in cities higher measurement heights such as rooftop locations tend to provide a better indication of the urban background and its temporal variation (Brook et al., 1999). Rooftops offer additional security for the

equipment because they are difficult to access. This raises a key point about finding secure, accessible sites with power, communications and a high likelihood of long term stability, especially in urban areas and given budget limitations. The point is that compromise is sometimes necessary when selecting a site because few locations are "perfect" and opportunity to have access to a reasonable location may be difficult to find given time constraints.

Existing measurements and/or past experiences can provide considerable insights in selecting measurement sites or designing networks. Air quality models can provide guidance and are also valuable for data analysis and for expanding spatial detail (i.e., conditions between sites). As many of these sources of information as possible should be used to optimize and enhance sampling strategies.

Quality Assurance

Documentation on the quality assurance (QA) measures and expectations that are or will be followed throughout a measurement program helps ensure the data's value and that they are not misinterpreted by other users. It is entirely the responsibility of the data generator(s) to initiate and uphold this plan and to make users of the data aware of the QA details. Providers of the funding for measurements should demand that evidence of a QA plan be available before a program proceeds. Conversely, QA details and other data limitations are important for data users to understand to ensure that correct conclusions are drawn.

Measurement methods and types and model numbers of the instrumentation used should be recorded, as well as any changes during the program. Details of the measurement site, such as latitude, longitude, elevation (above MSL), inlet height and design and proximate emissions and obstructions (photos), are also necessary to document. This is referred to as site "meta data." Most commercial instruments have known detection limits and levels of precision and accuracy as well as information on interferences. Nonetheless, precision and accuracy targets for the measurements need to be quantified and the actual values being achieved should be determined routinely to ensure a failing piece of equipment is identified and replaced quickly to avoid data loss. Duplicate or repeated measurements of the same samples and analysis of standards with known, traceable concentrations are, therefore, critical to undertake routinely to track precision and accuracy, respectively. The amount of resources needed for a measurement program is impacted upon by the frequency of QA measurements (e.g., number of site visits, amount of standard gases used), but, as stated above, this aspect should not be underappreciated. On the other hand, when QA measurements are being made (e.g., 'zero' or 'span' readings) actual measurements are being missed and thus the appropriate balance needs to be established.

AQ measurements that involve sample collection in the field followed by chemical or gravimetric analysis in the laboratory require QA in both the field and the lab. With respect to the field, one of the most important QA measures is the collection of field blanks. Whatever the approach to capturing the sample (e.g., filters, denuders, canisters, passive sorbants, traps or cartridges), at least 10% of the samples analyzed in the lab should be field blanks. These procedures should be clearly stated in the QA plan and the resulting data need to be rigorously analyzed to ensure data quality objectives are upheld and/or adjusted.

QA measurements should be based upon concentration levels that are typical for the site and that are within the measurement range of interest to the program. QA samples should also be introduced into the instrument or the analytical procedure (i.e., for laboratory analysis) in a manner that mimics the real measurement process and the conditions during measurement as much as possible. In general, the more experimental the measurement method the more challenging the QA and more frequent QA checks are usually necessary. As discussed in the NARSTO Assessment (McMurray, Shepherd and Vickery, 2004), for some measurements appropriate standards are not available. In this case, method intercomparison can provide information to judge the level of confidence in the data. With respect to current AQ health issues of concern, BC and UFP

measurements are hindered by the lack of standards.

Data archiving and reporting

Long term storage of final measurement data, including the QA data and site meta-data, is of utmost importance. At the same time data accessibility throughout the future needs to be simplified and rules for providing the data to all users also need to be considered. Whenever possible, data should be archived at their native temporal resolution since averaging to longer time periods will likely lead to unrecoverable loss of information that may be of value in the future. Given the present costs of storage media (hard drives, DVDs, etc.) archiving one-minute data is no longer an issue.

A key issue with AQ measurement programs is 'data turn-around time'. Objectives for this and rules for how other users can or cannot report the data need to be established. Faster turn around can be expected to increase the value of the data assuming that the more current the information, the greater the number of interested users and the greater the impact of publicizing what is being observed. There is an increasing demand for real-time data and data products reported via websites and the media (see Section 3.2.3.4 below). As indicated above, this is possible for a growing number of pollutants because of improved instrumentation, automated QA and communications. Continued improvements in these areas can be expected to reduce subsequent QA work and to increase data usage and publicity.

There are a number of recognized national or international data archives or portals for data access. Each has their own criteria for accepting and then documenting and preparing or formatting the data for storage and exchange. Examples of national archives of standard monitoring data are the National Air Pollutant Surveillance Network (NAPS) maintained by Environment Canada's Environmental Technology Centre (www.etcentre.org/NAPS/index_e.html), the Air Quality System (AQS), which is the USEPA's repository of ambient air quality data (www.epa.gov/ttn/airs/airsaqs/), Instituto Nacional de Ecología's archive of Mexico's air quality data (www.ine.gob.mx/) and the Air Quality Archive (AQA) for data for the United Kingdom (www.airquality.co.uk/archive/data_and_statistics_home.php). Data from many European countries are also available from AirBase, which is under the European Topic Centre on Air and Climate Change (http://etc-acc.eionet.europa.eu/databases/airbase/airbasexml/index_html).

A growing amount of North American AQ research data, are being kept in the NARSTO archive. These data are available from the NASA Langley Atmospheric Science Data Center as ASCII data files most of which are in the NARSTO Data Exchange Standard (DES) format. This format is described on the NARSTO Quality Systems Science Center site (http://cdiac.ornl.gov/programs/NARSTO/). Another common format in which data are provided is 'NASA Ames' (http://cloud1.arc.nasa.gov/solve/archiv/archive.tutorial.html). Similar to NARSTO, the header section of a file contains important metadata, including instrument type, instrument name, data resolution, and units. The Joint Research Centre of the European Commission uses the NASA Ames Data Exchange Standard (http://airispra.jrc.it/Start.cfm). The netCDF (network Common Data Form) library also defines a machine-independent format for representing scientific data (www.unidata.ucar.edu/software/netcdf/). Together, the interface, library, and format support the creation, access, and sharing of scientific data.

Consistency

Proper QA and data archiving significantly increases the likelihood that measurement data are consistent among countries. Utilization of instrumentation that has been approved by national or international standards organizations (e.g., USEPA reference or equivalent methods, NIST) also helps ensure consistency. This is critical for successful AQ management and thus, the more there can be international consensus on QA requirements and data storage and sharing protocols to better. Clearly, due to individual needs of each country and the fact that resources for measurements within a country will, by in large, be provided by that country implies that there will be differences. The people implement-

ing the measurement programs will differ among countries. Establishing a small number of internationally supported, master stations where a wide range of ongoing measurement comparisons can be undertaken will provide valuable insight regarding consistency.

3.4 Air Quality Modeling for Risk Management

3.4.1 Introduction

As noted at the beginning of this chapter, a crucial component in understanding and managing atmospheric pollution is our ability to quantify the links between emissions of primary pollutants or precursors of secondary pollutants on the one hand and ambient pollutant concentrations and other physiologically, environmentally, and optically important properties on the other. Air quality (AQ) models provide this capability. Such models consist of mathematical representations of the relevant physical and chemical atmospheric processes that are then solved by means of numerical algorithms to obtain pollutant concentration fields as functions of space and time for a given set of pollutant emissions and meteorological conditions (e.g., Peters et al., 1995; Seinfeld and Pandis, 1998; Jacobson, 1999; Russell and Dennis, 2000; Reid et al., 2003). Figure 3.4.1 shows a schematic of the atmospheric "process web," including the numerous links and interconnections between different atmospheric components, that should be represented in an AQ model.

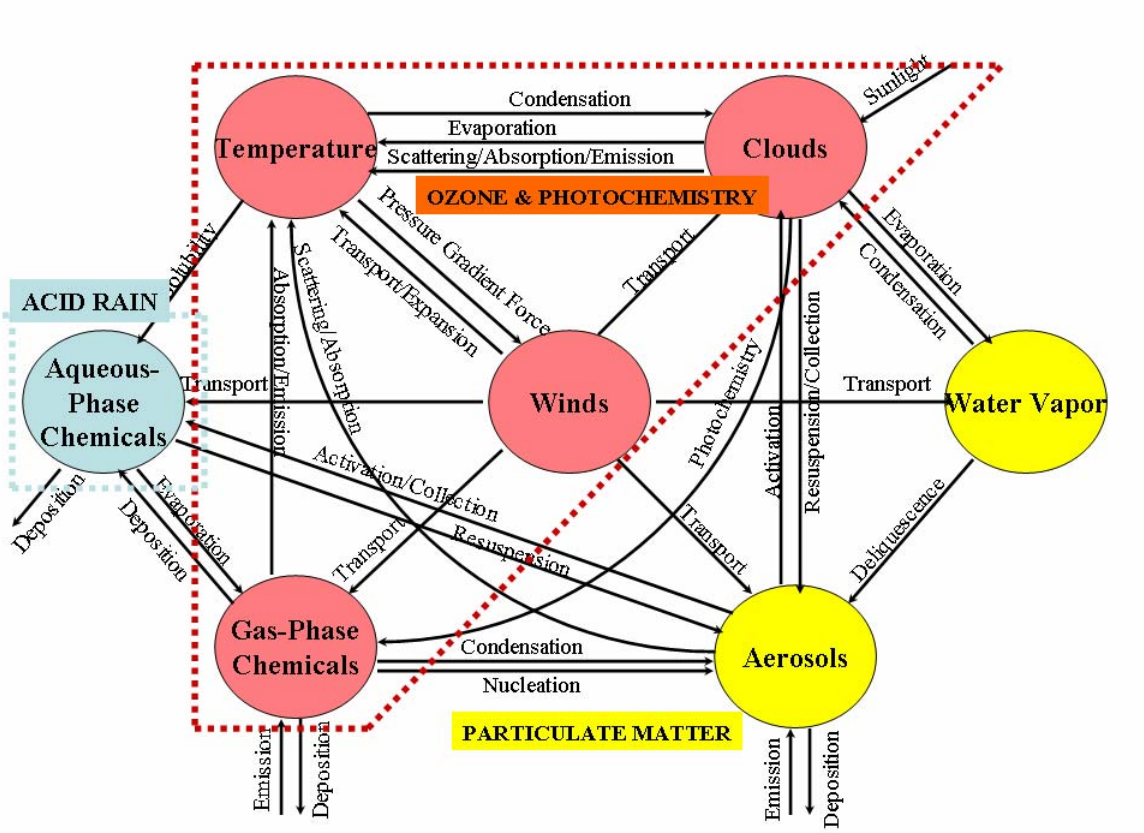

Figure 3.4.1. Schematic diagram of atmospheric physical and chemical components and their interactions (adapted from Peters et al., 1995).

AQ models are also referred to by other names, including chemical transport models, long-range-transport models, emissions-based models, source-based models, source-oriented models, and source models. Depending on the particular set of atmospheric processes included in such models, they can be classified into various categories such as photochemical models, acid deposition models, and particulate-matter or aerosol models (e.g., Seigneur and Moran, 2004). All of these models, however, include a representation of some atmospheric *chemical* transformations along with representations of emissions, transport, diffusion, and removal processes. The inclusion of chemistry typically requires consideration of time scales ranging from fractions of seconds to days in order to account for many important chemical reactions, and hence AQ model domains need to extend at least several hundreds of kilometers in the horizontal and up to at least the middle of the troposphere in the vertical for compatibility with the transport that can occur during a multiple-day simulation. Models of air pollutants that do not consider chemistry, on the other hand, are generally referred to as "dispersion models."

Figure 3.4.2 shows a flowchart of the data flow required to apply an AQ model. In fact, as this figure makes clear, it is more accurate to refer to this as an AQ *modeling system* since the emissions files and meteorological files that are needed to drive an AQ model are provided by two other complex computer programs, namely an emissions processing system (e.g., Dickson and Oliver, 1991; Houyoux et al., 2000) and a numerical weather prediction model (e.g., Seaman, 2000). An emissions processing system in turn requires one or more emission inventories as its primary input plus such ancillary information as population, socioeconomic, and geophysical data. A numerical weather prediction model needs meteorological observations from a variety of observational platforms (e.g., surface instruments, rawinsondes, aircraft, satellites) as its primary inputs plus various geophysical data sets.

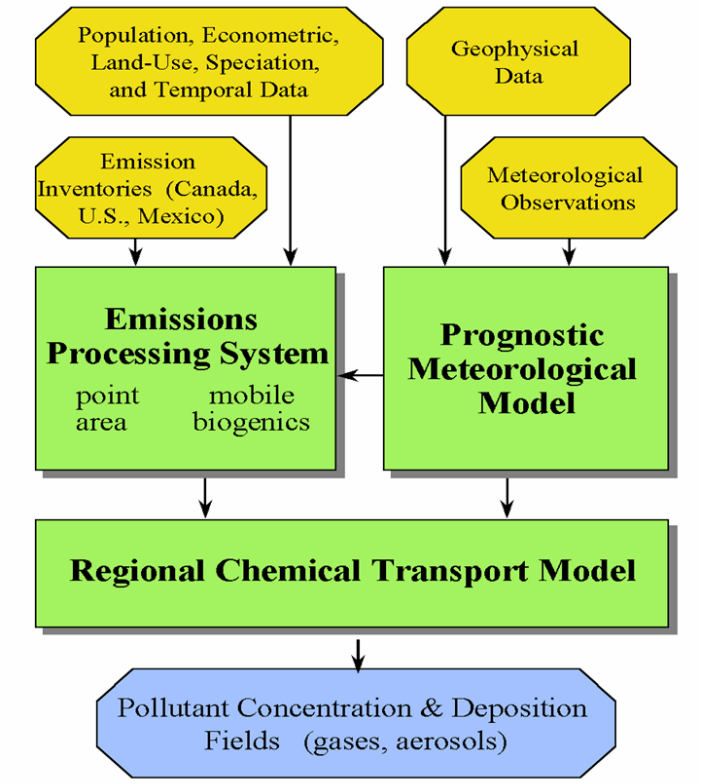

Figure 3.4.2. Schematic description of the components of an AQ modeling system (from Seigneur and Moran, 2004).

AQ modeling systems can be used to quantify source-receptor relationships for a wide range of air pollutants. They are also the only tool available to predict *future* air concentration and deposition patterns based on possible future emission levels. That is, they are prognostic in nature, unlike receptor models, which depend upon ambient measurements and hence are applicable only to periods for which measurements exist. For AQ models to be useful, however, there must already be information available about emissions and atmospheric measurements. If such data are not available for a region, then AQ model applications for that region can seldom provide much useful guidance for policymakers due to the much greater uncertainties associated with model predictions due to the uncertainties in model inputs. (The need for emissions data is obvious as illustrated by Figure 3.4.2. AQ measurement data are needed to specify chemical initial conditions and boundary conditions as well as to evaluate model performance.)

AQ models have been used for decades to support AQ management, but at the same time they have also undergone continued and rapid development. For example, the first meeting in the long-running NATO-CCMS series of international technical meetings on air pollution modeling was held in Eindhoven, The Netherlands in 1971. Most air-pollution models at that time were short-range, single-source dispersion models, and multiple-source models for modeling primary pollutants in urban settings were just being developed. Some of the earliest air-pollution models with parameterizations of chemistry were developed in the 1970s to simulate either the formation of photochemical smog in the Los Angeles basin or the long-range transport and transformation of air pollutants contributing to acid deposition in Europe and in North America. Given this considerable history, there have been a number of overviews of AQ models and AQ modeling over the years, including textbooks such as Jacobson (1999) and review articles such as Peters et al. (1995), Russell and Dennis (2000), Seigneur (2001), and Seigneur and Moran (2004).

The discussion in this section is not intended to provide a comprehensive review of AQ models and modeling practices. Instead, it builds upon the earlier NERAM paper on AQ modeling by Reid et al. (2003) and focuses on AQ model capabilities and uncertainties and on the management of these uncertainties in AQ model applications. Section 3.4.2 summarizes the variety of ways in which AQ models can contribute to AQ risk management. Section 3.4.3 reviews key technical choices and issues related to AQ model applications, especially those factors contributing to model uncertainty. Next, Section 3.4.4 provides an overview of "best practices" for using AQ models and their results and for managing associated uncertainties. Finally, Section 3.4.5 presents conclusions and recommendations.

3.4.2 Applications of Models for AQ Risk Management

AQ models can be applied in a number of ways, both directly and indirectly, to support AQ management and policy formulation:

- evaluation of impact of emissions changes, including proposed control measures;
- source apportionment and source attribution;
- input to conceptual model development;
- emission inventory evaluation;
- measurement network and field experiment design;
- AQ forecasting;
- testing current understanding of science.

Let us consider each of these applications in turn.

Evaluation of impact of emissions changes. The use of AQ models to assess the impact on air quality of emission changes due to pollutant abatement strategies, new pollution sources, population and economic growth, etc. has probably been the most common application of AQ models. Reid et al. (2003) gave three examples of such applications for Spain, Australia, and Canada, respectively: (a) the assessment of ozone abatement strategies for the Greater Madrid area (Palacios et al., 2002); (b) the impact of alternative urban forms of the city of Melbourne on urban air pollution levels (Manins et al., 1998); and (c) the sensitivity of PM concentrations in Ontario to changes in

emissions of primary PM and PM precursor gases. Four more examples include (d) the generation of intra-European source-receptor matrices ("blame matrices") by individual country for oxidized sulphur, oxidized nitrogen, and reduced nitrogen species for 2003 emissions and projected 2010 emissions (EMEP, 2005), (e) the assessment of the impact of possible emission control strategies in the Pearl River Delta region of southern China (Streets et al., 2006), (f) the estimation of AQ benefits from implementation of SO_2 and NO_x emission reductions under the 2005 U.S. Clean Air Interstate Rule (U.S. EPA, 2005a), and (g) the evaluation of the potential impacts of proposed SO_2 and NO_x emission control measures in Canada and in the U.S. on acid deposition in Canada (Moran, 2005).

Many of these studies follow a similar approach. The AQ model is first run for a "base-case" simulation, for which the emissions used are either historical or current and for which AQ measurements are available with which to evaluate model performance, and then again for one or more emission "scenarios," in which the assumed emissions correspond to a possible future state. For the simplest type of emission scenario, a "roll-back" scenario, emissions of one or more species may be changed by a fixed percentage for all source types across either the entire model domain or a selected subdomain. In more realistic emission scenarios, selected source types such as on-road motor vehicles or coal-fired electrical generating stations may be targeted. Sometimes one of the scenarios corresponds to a "business-as-usual" (BAU) scenario, in which emissions from the base case have been projected forward in time to account for population and economic growth and the implementation of any scheduled control measures expected by the scenario year. A companion future-year emission scenario may then be run that is identical to the BAU scenario except for the addition of a new candidate abatement strategy. Comparison of the base case, the BAU future scenario, and the companion future scenario then allows the impact of the abatement strategy to be estimated as well as any changes expected in future-year AQ relative to current conditions if the abatement strategy were or were not implemented. As one example, Figure 3.4.3 shows predictions of effective acidity wet deposition from an acid deposition model for two cases: a 1989 base case and a realistic 2020 emissions scenario.

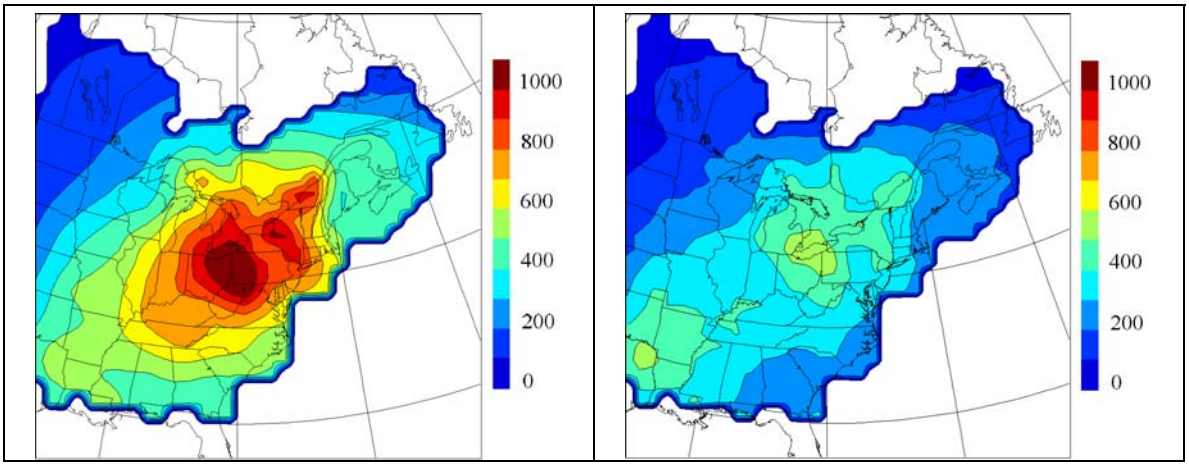

Figure 3.4.3. Plots of annual effective acidity wet deposition (units of eq/ha/yr) predicted by the ADOM acid deposition model for (left) 1989 base case and (right) 2020 SO_2 and NO_x emission scenario. Effective acidity is defined to be the sum of sulphate and nitrate wet deposition. See Moran (2005) for details.

Source apportionment and source attribution. AQ models can be used to investigate and quantify source-receptor relationships in a manner that is complementary to the use of receptor models, that is, based on a source-oriented framework rather than on a receptor-oriented framework (Blanchard, 1999). The simplest approach is the so-called "zero-out" approach, in which emissions from a particular source sector (e.g., petroleum refineries) or from a particular jurisdiction (e.g., a city, a province/state, or a country) are set to zero while leaving emissions from other source sectors or jurisdictions unchanged. Predictions from this case can then be subtracted from predictions from a base run in which all emission sources are considered to estimate the impact of the targeted source sector or jurisdiction. Two more sophisticated approaches are "source tagging" and inverse methods. In the former, pollutant emissions from particular source sectors or geographic locations are tracked in the model as separate ("tagged") species (e.g., Kleinman, 1987; McHenry et al., 1992; Kleeman and Cass, 1999a,b; Zhang et al., 2005). In the latter, the adjoint of the AQ model can be constructed, used to quantify the sensitivity of the model to emission inputs, and then combined with ambient measurements, or else initial attribution results can be refined based on the synthesis inversion technique (e.g., Uliasz, 1993; Pudykiewicz, 1998; Mendoza-Dominguez and Russell, 2001; Mallet and Sportisse, 2005; Knipping et al., 2006).

Input to conceptual model development. An AQ conceptual model is a qualitative mental model for a geographic region that is based on a synthesis and simplification of available AQ information obtained from analysis of emissions, measurements, and AQ model results to distill the primary contributing factors, including key emission sources, terrain characteristics, and local weather and climate. Useful information from AQ models can include results from a suite of emissions scenarios, from source apportionment studies, and from sensitivity studies (see Section 3.4.3). One good example of the development of such a conceptual model is the analysis described by Pun and Seigneur (1999) for PM pollution in California's San Joaquin Valley. A number of the regional conceptual models for PM presented in the *2003 NARSTO PM Science Assessment* are based in part on AQ model results (NARSTO, 2003). One report (U.S. EPA, 2005b) gives a useful list of questions and points to consider in constructing a conceptual model of ozone pollution, including the relative contribution of local and distant sources, the role of certain weather patterns, and the nature of the local chemical regime. Answering some of these questions requires the application of AQ models. And in the paper by Zunckel et al. (2006), AQ model results contribute significantly to a conceptual model for surface ozone in southern Africa even with very limited availability of emissions data and AQ measurements.

Emission inventory evaluation. Since AQ model predictions depend directly on the input emission fields, the comparison of AQ model predictions with AQ measurements can give some indication of the accuracy of the input emissions. For example, if AQ model predictions are biased significantly high or low in a certain region as compared to measurements, one possible cause could be a corresponding high or low bias in the input emissions for that region. Or if measurements are stratified into rural and urban sites, then AQ model predictions for rural and urban locations can be checked independently to see if any systematic difference exists for rural vs. urban emissions (e.g., Yu et al., 2004). Inverse modeling analyses, in which enhanced AQ models are combined with ambient measurements, provide another, more quantitative, approach to estimating emissions strengths on either a regional or a global basis (e.g., Pudykiewicz, 1998; Mendoza-Dominguez and Russell, 2001; Martin et al., 2002; Palmer et al., 2003; Gilliland et al., 2006).

Measurement network and field experiment design. AQ models can also be employed as sophisticated interpolation schemes since their results are based on consistent and comprehensive representations of physical and chemical laws. For example, an analyst could use modelled pollutant fields as "true" patterns and then sample the model fields for different numbers of grid cells in order to investigate the

impact of adding or removing stations to a network on the estimation of the actual pollutant spatial pattern from measurements, including the identification or non-identification of strong gradients and "hot spots." Similarly, scientists planning and designing a field experiment could use model predictions to help them choose measurement site locations, aircraft flight tracks, and so on.

AQ forecasting. When AQ models are used to provide policy guidance or to interpret field-experiment measurements, historical periods are typically considered. Such applications are referred to as retrospective runs or "hindcasts." If, however, an AQ model can be run in "real time" quickly enough (i.e., in a few hours), then it can provide an AQ forecast for the next day or two, that is, a prospective simulation. Both the Canadian and the U.S. national weather services now run AQ models in order to issue public regional AQ forecasts (e.g., www.msc-smc.ec.gc.ca/aq_smog/chronos_e.cfm and www.nws.noaa.gov/aq/) and the Australian national weather service issues AQ forecasts for two large urban areas, Melbourne and Sydney (Cope et al., 2004). Besides providing useful guidance to the public about next-day AQ, such forecast programs have the added benefits of (a) maintaining or raising public awareness about AQ and (b) challenging the AQ models with a broader range of weather conditions than they are typically subjected to in scenario modeling (e.g., photochemical pollutant scenarios are almost always summer cases). Performance evaluations for AQ forecasts can then provide additional insights into model skill and reliability and identify model weaknesses (e.g., Eder et al., 2006).

Testing current understanding of science. Finally, AQ models provide a means to represent and link in a single package our best understanding of all of the chemical and physical processes relevant to AQ. This knowledge synthesis can then be evaluated by comparing model predictions with enhanced measurement data sets obtained from dedicated AQ field campaigns such as SCAQS (1987), EMEFS (1988-90), NARE (1993), NARSTO-NE (1995), ESQUIF (1998-99), BRAVO (1999), TexAQS (2000), ACE-ASIA (2001), ESCOMPTE (2001), Pacific 2001 (2001), TRACE-P (2001), and ICARTT (2004) (e.g., Dennis et al., 1993; Berkowitz et al., 1998; Heald et al., 2003, 2005; Frost et al., 2006; Hodzic et al., 2006; Pun et al., 2006; Smyth et al., 2006a). Conversely, an AQ model may also be used to help interpret field-campaign measurements, which can be difficult for a set of measurements of limited duration and restricted to a small number of locations due to the complexity of geography, meteorology, and interconnected chemical and physical processes. An AQ model can also be used as a test bed to test a new parameterization of a key chemical or physical process (e.g., Padro et al., 1993; Pierce et al., 1998). Such activities probe both our current scientific understanding and our representation of it in AQ models, often leading to improvements to both.

3.4.3 Key Technical Issues to Consider in AQ Modeling Programs

Worldwide, there are a number of AQ modeling systems available, and each is typically composed of a set of large, complex computer programs. As a consequence, there are many choices to be made and issues to be considered by a modeller when using an AQ modeling system for any application. Such choices and issues, however, also need to be taken into account by users of model results when judging the robustness and reliability of those results. Consider the following.

Choice of model. Some race-track devotees offer the advice that there are "different horses for different courses." The same is true of AQ models. The first step in applying an AQ model is to define the questions that need to be answered, and then, if possible, identify or develop an appropriate conceptual model. Doing so should immediately narrow down the set of AQ models that might be used to answer the question. For example, a very detailed but computationally expensive AQ model might not be the best choice for performing a multi-year AQ simulation, if such is called for. Another consideration is that a model designed to address one AQ issue (e.g., photochemical smog) may not include representations of all of the processes necessary to address another issue

(e.g., deposition of acidic species – see Figure 3.4.1). And a model designed for highly polluted atmospheres may not be appropriate to model a clean atmosphere and vice versa (e.g., regional atmospheric chemistry in source regions vs. background global chemistry).

Model configuration. There are many choices to be made in configuring (i.e., setting up) an AQ model run. These include (a) the location and (b) the size, in both the horizontal and vertical, of the model domain, (c) the map projection to be used, (d) the grid spacing in both the horizontal and vertical, (e) the integration time step, (f) the simulation period, including any required "spin-up" time (the time for atmospheric concentration fields to reach an equilibrium between emissions and removal processes), (g) the "refresh" rate (the length of time that the *meteorological* model will be run before being re-initialized using a new set of meteorological analyses), (h) the set of chemical species to be considered, (i) the choice (in some cases) of parameterizations to be used for different chemical and physical processes, (j) the specification of initial chemical conditions, (k) the treatment of chemical lateral and upper boundary conditions, and so on. Each choice has implications. For example, the use of large horizontal grid spacing may "average out" a suspected hot spot or not represent small-scale meteorological circulations forced by local terrain features. See U.S. EPA (2005b) for a discussion about the choice of horizontal grid spacing, Berge et al. (2001) for a discussion about the specification of chemical initial conditions, and Brost (1988) for a discussion about the specification of chemical lateral boundary conditions.

Science considerations. One key limitation of AQ models is in fact gaps in our scientific understanding of the pollutants of interest. For example, it is well known that the sources of much of the carbonaceous component of atmospheric $PM_{2.5}$ are not presently known in spite of the fact that this component typically contributes ~40-50% of total $PM_{2.5}$ mass. Another example is our limited understanding of nighttime NOx chemistry (e.g., Brown et al., 2006). A second limitation is the use in current AQ models of process parameterizations of limited fidelity to the real atmosphere. For example, Dabberdt et al. (2004) recently identified the need for improved treatments of the influence on AQ of (i) the planetary boundary layer and (ii) clouds and cloud processes. Another important consideration is the presence of nonlinear effects in the chemical reactions that generate some pollutants of interest. For example, the possibility of a nonlinear response in sulphate deposition to SO_2 emission reductions due to oxidant limitations was identified in the 1980s as a potential concern for managing acid deposition (e.g., Misra et al., 1989). Nonlinearities in ozone photochemistry are also well known (e.g., Seinfeld and Pandis, 1998), but PM chemistry possesses even more nonlinearities. For example, Meng et al. (1997) presented AQ model predictions for two simple ozone control scenarios run for a Los Angeles smog episode in which VOC emission reductions reduced ozone levels but caused increases in $PM_{2.5}$ mass. And West et al. (1999) presented AQ model results in which reductions in SO_2 emissions in eastern North America increased $PM_{2.5}$ concentrations due to so-called "nitrate substitution." Such nonlinear responses can further complicate the interpretation of model predictions and the formulation of possible emission control measures.

Model parameterization and algorithmic limitations. Even when AQ processes are well understood scientifically, they must still be represented mathematically in AQ models by so-called process parameterizations, and then the complex, coupled system of governing equations that comprise the AQ model must be solved numerically. Both steps have limitations and can introduce errors. For example, in many cases a number of different parameterizations have been developed to describe the same chemical or physical process, and these different parameterizations will produce different results (e.g., Kuhn et al., 1998; Zhang et al., 2000, 2001; Mallet and Sportisse, 2006). Typically, more sophisticated (and complex) parameterizations have a greater number of parameters and coefficients that must be specified, but measurements to do so may be scarce or lacking completely. This implies that

while a more sophisticated scheme may have the potential to do a better job in describing a process, there are no guarantees that it will actually do a better job in practice. For example, some gas-phase chemistry mechanisms consider a few dozen species whereas others consider hundreds or even thousands of species. But besides reaction rates for all of the reactions that these species participate in, emissions must also be specified for each species as must a number of chemical and physical properties such as molecular diffusivity and Henry's Law constant that are needed to estimate dry and wet removal rates. Such physico-chemical data may not be available for every species (e.g., Zhang et al., 2002).

The numerical integration of the AQ model also introduces errors since it usually requires the solution of large coupled systems of both ordinary and partial differential equations. As discussed by Pielke (1984) and Jacobson (1999) among others, finite-difference methods are usually employed in both time and space. Any time-stepping scheme used to integrate the AQ model in time will have truncation errors that depend upon both the order of the scheme and the chosen time step. Operator splitting is usually employed on the right-hand side of the governing equations to allow each process parameterization to be calculated separately, but operator splitting also introduces errors that depend upon the order of the splitting and the overall time step. Advection is well known to be a difficult process to solve, and literally hundreds of numerical schemes have been developed for advection. All suffer to varying degrees from some or all of truncation errors, numerical diffusion, phase errors, lack of positive definiteness, and violation of mass conservation.

Model "resolution" is another important consideration. The choice of a discrete model time step and grid-cell size implicitly imposes numerical filtering on the model solution. In essence, no temporal feature shorter than $2\Delta t$ and no spatial feature smaller than $2\Delta x$ can be predicted by the model, and $4\Delta t$ and $4\Delta x$ are probably a more realistic threshold (e.g., Pielke, 1984; Grasso, 2000). This has important implications for processes operating at smaller temporal and spatial scales (e.g., Uliasz and Pielke, 1998). As a consequence, many parameterization schemes to represent the influence of subgrid-scale processes at grid scale have been developed. One obvious example is the representation of point source emissions. In any Eulerian (grid) model, all or most point sources will be represented as volume sources since the emissions are assumed to be well-mixed throughout at least one grid cell, thus introducing large numerical (i.e., artificial) diffusion in the vicinity of major point sources. To address this problem, which will be most pronounced for isolated sources, so-called plume-in-grid schemes have been developed to represent near-source diffusion and chemistry in plumes from major point sources more realistically. The treatment of vertical diffusion in any AQ model is also intrinsically a subgrid-scale parameterization since it must represent the impact of a spectrum of atmospheric motions that cannot be resolved by the model. Cumulus parameterizations that represent the impact of unresolved clouds on model fields are another important class of subgrid-scale parameterizations that are employed in weather and AQ models (e.g., Haltiner and Williams, 1980).

Input data. AQ models require a number of input data sets in order to run. First are emission input files. Emission rates of a number of gaseous and particulate species must be specified for each model time step at each model grid cell at all levels. As discussed in Section 3.2, there are significant uncertainties associated with such emission files due both to errors in the emission inventories themselves and to additional uncertainties introduced by the emissions processing systems that perform the speciation and spatial and temporal disaggregation steps needed to create model-ready emission files (e.g., Hogrefe et al., 2003). For large point sources, ancillary information about smokestack characteristics such as stack height, stack diameter, stack-gas exit velocity, and stack-gas exit temperature is also needed. And if a future-year scenario is being considered, current emission inventories must be manipulated and modified to account for all assumptions built into the scenario.

Second, meteorological input files are needed

if an "off-line" AQ model is being used (i.e., meteorological model and AQ model are separate), which is the case for most current AQ modeling systems. Meteorology is very important as it influences every aspect of the AQ system, including (a) natural sources of PM such as wind-blown dust and sea salt, (b) plume rise, (c) transport and diffusion, (d) gas-phase and heterogeneous-phase chemistry (via temperature and humidity effects), (e) cloud shading, (f) aqueous-phase chemistry, (g) dry removal, and (h) wet removal. Third, chemical initial conditions must be supplied for all model species for every grid cell, chemical upper boundary conditions must specified for all model species at the top model level, and, for a limited-area AQ model, chemical lateral boundary conditions must also be specified for all model species. And fourth, a number of geophysical fields may also be required, including terrain height, land-use type, vegetation type, aerodynamic surface roughness, albedo, sea surface temperature, and soil texture. The accuracy and representativeness of all of these input files are a key concern, since even for a perfect model, the well-known aphorism "garbage in, garbage out" still holds.

Infrastructure considerations. As discussed by Reid et al. (2003), AQ modeling is typically resource-intensive in terms of model input data, people, calendar time, and computer power. In order to apply an AQ model for a particular case, the input data sets described above must be prepared for the model configuration selected, including emissions, meteorological, and geophysical files, the model must be run, and then the model results must be processed, analyzed, and interpreted. Typically, a minimum of three highly-trained modellers would be required to contribute, namely an emissions-processing specialist, a meteorological-modeling specialist, and an AQ-modeling specialist. The required calendar time from start to finish, including configuring and testing the model for the application, would likely be a minimum of weeks but more likely months. The minimum computer resources needed would be a high-end PC with multiple processors, large internal memory and disk space, and off-line archiving hardware to save numerous large model output files. Access to emissions data, meteorological data, geophysical data, and AQ measurement data is of course assumed as well.

Model accuracy, sensitivity, and uncertainty. For an AQ model's predictions to be used by policymakers, the model and its predictions should be credible. To be credible, the model should give the right answers for the right reasons. The determination of whether a model is giving the right answers is addressed by model performance evaluations, in which model predictions are compared to measurements. Model performance evaluation is discussed in more detail in the next section. However, there are some fundamental issues related to model accuracy, sensitivity, and uncertainty that need to be kept in mind. For one thing, how is the "right answer" determined? For another, how can model uncertainty be determined?

There are actually a surprising number of issues that arise in comparing AQ model predictions with ambient measurements. The biggest one is incommensurability, which arises due to the fact that model predictions correspond to grid-volume averages whereas measurements are typically made at points in space or along lines (e.g., aircraft flight tracks, DIAL). For example, for a regional-scale AQ model whose smallest grid volume is 20 km by 20 km by 50 meter, how representative would a single point measurement be of the 20 km^3 of air contained in that grid volume? One effort to address this question was made during the 1988-90 EMEFS field experiment in eastern North America. Surface measurements were taken of 24-hour SO_2, particulate SO_4, total NO_3 and hourly O_3 concentrations for two years at six clusters of 3 to 5 stations that fell within the confines of model 80-km by 80-km horizontal grid cells. Cluster or subgrid daily variability was found to be approximately linearly related to mean concentration, with the largest variability associated with SO_2 and minimum O_3 (~ ±70%), intermediate variability associated with p-SO_4 (~ ±30%) and t-NO_3 (~ ±40%), and the smallest variability associated with maximum O_3 (~ ±20%) (Seilkop, 1995a,b). Significantly, uncertainties due to this subgrid-scale variability overwhelmed uncertainties associated with instrument error. McNair et al. (1996) performed

a somewhat related analysis in the Los Angeles basin for O_3, NO_2, and CO for two 1987 SCAQS cases. For circular areas with 25-km radius, smaller than those considered during EMEFS, they found that local inhomogeneities for these three species had normalized gross errors in the 25-45% range.

A second issue related to the comparison of model predictions and ambient measurements is the need to compare "apples with apples." For example, for gas-phase species, AQ model predictions correspond to ambient conditions whereas some networks report measurements at STP. For PM comparisons, model PM predictions are calculated based on Stokes diameter whereas PM measurements are reported using aerodynamic diameter, PM measurements unlike models can suffer from artifacts related to volatile species such as nitrate, some organic compounds, and aerosol-bound water, and the distinction between elemental carbon (EC) and organic carbon (OC) is analysis-method-based and can vary from network to network (Seigneur and Moran, 2004). Note that the *definition* of the EC and OC variables predicted by an AQ model also depends on the EC-OC analysis method used to speciate primary PM emissions.

Turning to model uncertainty, the discussions and conclusions of a 1982 workshop on AQ model uncertainty have been described by Fox (1984) and Venkatram (1988). The latter identified three main sources of model uncertainty as "(1) errors in model inputs, (2) errors in model formulation, and (3) inherent uncertainty associated with the stochastic nature of turbulence." The last source constitutes a lower limit on model uncertainty since it cannot be reduced even if all model-related errors are corrected. One aspect of this inherent uncertainty is related to the time and space averaging used in measurements vs. the ensemble averaging that is used to describe atmospheric turbulence. That is, atmospheric measurements correspond to samples from a single flow realization whereas AQ model parameterizations related to diffusion and mixing are based on ensemble averages for (theoretically) an infinite number of flow realizations with identical external conditions (e.g., Moran, 2000).

Reid et al. (2003) have noted that it is not possible to quantify overall model uncertainty because it is dependent on so many factors, some of them dependent on the particular application being considered, but also on the *interactions* of these factors. As already discussed, these contributing factors include errors and uncertainties in input data such as emissions, meteorology, and boundary conditions, uncertainties in our scientific understanding and in process parameterizations, errors associated with numerical methods, and uncertainties associated with required parameters like reaction rates. It is, however, possible to quantify some individual sources of uncertainty, particularly for numerical methods, to identify model sensitivity to various inputs and parameters, and finally to compare results from parameterizations and even entire models in order to try to characterize the range of uncertainty.

Error characterization is generally reported as part of the description of new numerical methods and parameterization techniques. A wide range of sensitivity analysis techniques exist, including DDM, ADIFOR, FAST, variational techniques, perturbation theory techniques, Green's function techniques, and stochastic techniques, that can be used to understand which model parameters and input variables most influence selected model outputs (see Zhang et al. (2005) for a useful literature review). Besides being compared side by side outside of models, the impact of different parameterization schemes can also be compared when embedded in a host model (e.g., Padro et al., 1993; Mallet and Sportisse, 2006). And some studies have compared differences in AQ modeling system results due to the use of different component models. For example, Hogrefe et al. (2003) compared the impact of using emissions files constructed by two different emissions processing systems from an identical emission inventory on the predictions of one AQ model. They found differences on the order of ±20 ppb in predicted daily maximum 1-hour ozone concentration. Another source of uncertainty is meteorological inputs. Smyth et al. (2006b) compared the outputs from one

emissions processing system and one regional PM model for two sets of meteorological input files for the same period that were provided by two different meteorological models. An operational evaluation of the two meteorological models suggested that their performance was essentially equivalent, as was the performance of the AQ model for the two sets of meteorological files, but when grid cells were matched for the same time, large variability was observed, particularly in aerosol quantities influenced by relative humidity. And recently, the performance of seven AQ models in predicting ozone was compared for the same period (summer 2004) and region (eastern U.S.). The range of model predictions generally bracketed the measurements, and interestingly none of the models individually could match the skill of a weighted average of the seven forecasts (McKeen et al., 2005).

One other approach to assessing uncertainty is to synthesize expert opinion. Seigneur and Moran (2004) prepared a table that presented qualitative ratings of PM modeller's level of confidence in major aspects of the predictions of current PM AQ models. Only a few model aspects (SO_2, NO_x, and p-SO_4 air concentrations) were judged to have a "high" level of confidence. Most aspects were assigned "medium" or "low" ratings, and a few aspects, such as secondary OC and PM ultrafine mass and number concentrations, were assigned "very low" ratings. These ratings were based on an assessment of all contributing uncertainties, including uncertainties associated with the emissions of different pollutants and with scientific understanding.

3.4.4 Review of Best Practice for Using Models for AQ Management

Let us now build upon the previous sections and consider that wry but wise epigram by Box (1979): "All models are wrong, but some are useful." That is, in applying models for AQ management, we must accept from the start that no model is perfect. Instead, as discussed in the previous section, AQ model predictions can be affected by numerous sources of error and uncertainty. How then can we account for the resulting uncertainty and apply models in a reasonable and defensible way in order to inform AQ management?

To start, how can we judge whether an AQ model will in fact be useful? For a model to be useful, presumably it must be credible. That is, it must have demonstrated sufficient skill and reliability that its predictions can be used with some confidence by analysts and policymakers in the formulation of AQ management strategies. Confidence can in turn be built in two ways: first, through *model verification* to assess the consistency, completeness, and correctness of the model and through *model performance evaluations* to characterize its performance and quantify its errors; and second, by applying the model in as appropriate, transparent, and defensible a manner as possible for the AQ issues being considered.

Model verification and model evaluation. Model verification and model performance evaluations should always be a required step before a model is applied in the policy arena. According to Fox (1981) and Russell and Dennis (2000), model verification is an assessment of the accuracy, reality, or truth of a model. It does not require a model to be run. Rather, model verification is a "desk check" in which the consistency, completeness, and correctness of a model's design, science, process representations, algorithms and numerical methods, inputs, and source code are examined and assessed. Peer reviewers should be involved in such an examination, and, ideally, any interested party should have unrestricted access to the model source code for this purpose.

Model performance evaluation is the process of examining and appraising model performance through comparison of model predictions with measured AQ data and/or predictions from other models (e.g., Fox, 1981; Dennis et al., 1990; Russell and Dennis, 2003). There are four main types of model performance evaluation: (i) operational; (ii) diagnostic; (iii) mechanistic; and (iv) comparative (Seigneur and Moran, 2004).

- An *operational* evaluation requires the statistical evaluation of model predictions of a few key pollutants of interest with atmospheric measurements over time and space scales consistent with the intended applications of the model. An operational

evaluation is intended to answer the basic question: "Are we getting the right answers?" Examples of operational evaluations include EMEP (2003), Eder and Yu (2006), and Eder et al. (2006). A paper by Fox (1981) reviews a wide range of statistical measures that have been used in operational evaluations, but a recent U.S. EPA report (U.S. EPA, 2005b; Section 15.2) recommends a small number of statistical measures that have been found to be representative and useful in evaluating the performance of photochemical AQ models.

- A *diagnostic* evaluation is more of a research-level evaluation and involves an examination of model performance at the process level for all relevant species. A diagnostic evaluation addresses the basic question: "Are we getting the right answers for the right reasons?" Because diagnostic evaluations are more wide-ranging and comprehensive than operational evaluations and generally make use of non-routine measurement data sets such as those from specialized field campaigns, they can identify the presence of compensating errors or excessive "tuning." Examples of diagnostic evaluations include Dennis et al. (1993), Karamchandani and Venkatram (1992), Sillman et al. (1998), Hogrefe et al. (2001a,b) and Biswas et al. (2001), Heald et al. (2005), and Yu et al. (2005). Also, Seigneur et al. (2000) have described how to optimize the design of field studies that will be used in the evaluation of PM AQ models, a recent U.S. EPA report (U.S. EPA, 2005b; Section 15.3) lists some diagnostic analyses that have been found useful in assessing the ability of photochemical AQ models to predict changes in ozone due to changes in emissions of ozone precursors, and a recent paper by Zhang et al. (2005) examines three diagnostic probing tools that have been used to examine photochemical AQ model performance.
- A *mechanistic* evaluation involves testing individual model components (i.e., process representations) in isolation against field or laboratory measurement. Such evaluations address the question: "Are we using good parameterizations?" Some examples of mechanistic evaluations include Pleim and Xiu (1995), Odum et al. (1996), Geron et al. (1997), and Zhang et al. (2001).
- Finally, a *comparative* evaluation involves a side-by-side comparison with another model or model component for identical or similar inputs. A comparative evaluation addresses the basic question: "Are we getting comparable answers from comparable models?" Examples of comparative evaluations include Alapaty et al. (1997), Hass et al. (1997), Kuhn et al. (1998); Ansari and Pandis (1999), Zhang et al. (2000), Hogrefe et al. (2001a,b), and McKeen et al. (2005).

Note that the term "model evaluation" denotes a process rather than an outcome or conclusion. For the terms "model verification" and "model validation," on the other hand, Oreskes et al. (1994) argued that numerical models of natural systems can never truly be verified or validated, since these terms imply the absolute correctness of a model. Fox (1981) and Russell and Dennis (2000) were careful to restrict their definitions of these terms. Model verification, as described above, refers to an examination process that at best leads to a provisional conclusion. And model validation is a process leading to a judgement on the quality, suitability, and usefulness of a model for a particular application that should be based on evidence from both model verification and multiple model performance evaluations. Such a judgment, however, must always be viewed as provisional, since additional information such as the results of a new evaluation may change the balance of evidence.

It is also important to consider which aspects of model performance need to be evaluated. Most AQ model evaluations involve case studies in which a model is run for a particular period using input emissions and meteorology suitable for that period and then model performance is examined using measurements from that same period. However, as already discussed, the most common AQ model application is to evaluate the impact of emissions *changes* on AQ. The key aspect of model performance in this instance is how well the model predicts the atmospheric *response* to the change in input emissions, and the approach to the corresponding performance evaluation is necessarily somewhat different.

For a direct evaluation of model response, AQ measurements are required for *two* different periods so that an atmospheric response can be calculated, which means that the AQ model must be run for the same two periods using different input emissions corresponding to each of the two periods. Obviously, such a model-response evaluation is more demanding than the usual single-period evaluation since considerably more data and more modeling effort are required. Confounding issues include (i) the need to use emissions for two different periods estimated using a consistent methodology and (ii) the additional variability introduced by interannual meteorological variability. As a consequence, published model-response evaluations are uncommon, but a few are available (e.g., Moran and Zheng, 2006).

Note that so-called "accountability" studies, in which the emissions changes that have occurred are due primarily to legislated control measures and the study goal is to assess the performance of the AQ models used to predict the benefits of those control measures *before* the control measures were enacted and implemented, are also model-response studies. Given the considerable time that will have elapsed, however, between the time the original AQ modeling runs were performed and the time that the AQ measurements were made following implementation of the control measures of interest, it is not likely that the particular version of the AQ model (or even the model itself!) is still being used. On the other hand, current AQ models can also be evaluated in a retrospective mode for the same legislated emission changes (e.g., 1985 Eastern Canada Acid Rain Program, 1990 U.S. Clean Air Act Amendments, 1998 U.S. NO_x SIP Call). Note also that in terms of U.S. regulatory modeling terminology (e.g., U.S. EPA, 2001, 2005b), a model-response evaluation is equivalent to the evaluation of model-predicted *relative reduction factors*.

Finally, in considering the question "How accurate does a model need to be," Reid et al. (2003) suggested that the general answer is that "… the model predictions should be good enough that model uncertainty does not affect the decisions that are based on the predictions." In the real world, of course, this may not always be the case. How then should models be used given such uncertainties?

Model applications. In their review of photochemical models and modeling, Russell and Dennis (2000) discussed the *modeling process* as a separate topic. By this they meant the set of steps required to apply a model, including selection of model domain, grid resolution, and model configuration, preparation of model input files, model execution, and postprocessing and analysis of model predictions. The modeling process itself is worthy of individual attention because, as discussed in the previous section, all of these steps may influence the results provided by the modeling system. It is thus important to work through the modeling process in as reasonable and defensible a way as possible. Some limited guidance on how to do this does exist. For example, the U.S. EPA has prepared several documents to help modellers follow "best practice" when using regional AQ modeling systems for certain regulatory applications (U.S. EPA, 2001, 2005b).

Best practice basically boils down to thoughtful and careful selection, set up, and application of an AQ modeling system accompanied by careful scrutiny and consistency checking of the results by various means, including the use of measurements and results from both alternate configurations of the selected AQ model and from other AQ models. Figure 3.4.4 describes eight steps of best practice for AQ modeling based on guidance from two EPA reports (U.S. EPA, 2001, 2005b). Most of the following steps will be relevant to any AQ model application.

Let us consider each step in turn. Some relevant background material has already been discussed in Section 3.4.3.

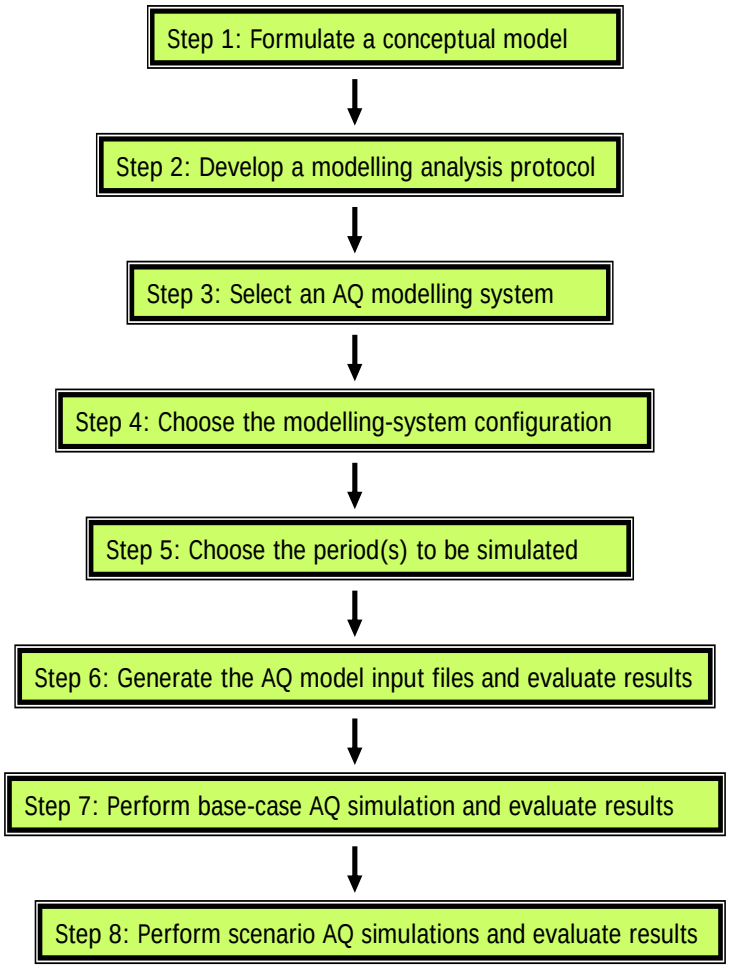

Figure 3.4.4. Eight-step set of best practices for air quality modeling applications for emission control strategies.

1. Formulate a conceptual model.

Both modeling specialists and modeling "clients" should have a conceptual understanding of the AQ issue to be considered with an AQ model. A conceptual model will provide useful guidance for all of the remaining modeling-process steps, including the identification of stakeholders, the selection and configuration of the AQ model, the development of candidate emission control scenarios, and the assessment of model results. As an example of how to formulate a conceptual model, Section 8 of U.S. EPA (2005b) lists a large number of questions about and analyses of emissions data, measurement data, and AQ model results that could be considered in developing a conceptual model for the occurrence of high annual or daily ozone levels in a particular locale or region.

2. Develop a modeling/analysis protocol.

This step may be handled either formally or informally, but if model results are to be provided to interested parties, it is desirable early on to identify all interested parties and to obtain agreement on (a) which questions should be addressed, (b) what assumptions are reasonable to make (e.g., What processes can be neglected? How large does the model domain need to be?), (c) how the modeling work should

be performed, (d) what sorts of results should be generated, (e) who should review them and how, (f) what the timetable should be, and so on. The conceptual model from Step 1 can guide this development, for example, by suggesting whether long-range transport is likely to be important or not, which will help to define the geographic extent of the "community" of stakeholders. Section 9.2 of U.S. EPA (2005b) gives a detailed suggested outline of what such a protocol might look like.

3. Select an AQ modeling system.

Once the questions of interest have been identified in Step 2, the next step is to select an AQ modeling system that is capable of answering those questions and that has been judged to be credible and skillful through peer review and performance evaluations. Availability of model source code, previous successful applications to similar problems, and advanced science features and tools are useful additional selection criteria. As well, the time and resources available for the model application are practical considerations that may also affect the choice of modeling system. [McKeen et al. (2005) describe seven current PM AQ modeling systems. Other AQ modeling systems are described by EMEP (2005), Heald et al. (2005), Hodzic et al. (2006), Mallet and Sportisse (2005), Meng et al. (1997), and Zunckel et al. (2006)]

4. Choose the modeling-system configuration.

Selection of an AQ modeling system is not sufficient by itself. It is also necessary to choose a modeling domain and map projection, horizontal and vertical grid resolution, level of nesting if any, an integration time step, a "spin-up" or "ramp-up" time, a "refresh" rate for the meteorological model, methods to specify chemical initial and boundary conditions, and, where choices are available, the particular physics and chemistry process parameterizations to be used in the meteorological and AQ models (e.g., convective parameterization, gas-phase chemistry mechanism, secondary organic aerosol scheme). In the case of an "off-line" AQ model, it is also desirable to harmonize to the extent possible the AQ model domain, map projection, and horizontal and vertical resolution with those of the companion meteorological model. Many of these choices will be guided by the conceptual model from Step 1 and the question(s) to be answered from Step 2. For example, the relative importance of long-range transport and the role, if any, of local terrain-forced meteorological circulations such as sea-land breezes will need to be considered. Sections 12 and 13.2 of U.S. EPA (2005b) provide some useful discussions about some of these choices.

5. Choose the period(s) to be simulated.

This is one of the most open-ended steps, but it will be strongly constrained by the question(s) to be answered and, if relevant, by the form of the AQ standard (e.g., daily or annual, average or maximum) or the exact wording of the legislation of interest. In the case of short-term effects or AQ standards, the conceptual model should provide useful guidance, particularly related to the meteorological conditions that are associated with AQ exceedances. When choosing short-term simulation periods, Section 11 of U.S. EPA (2005b) recommends choosing a set of periods (a) for which extensive emissions, meteorological, and AQ data sets exist, (b) that correspond to a variety of relevant synoptic conditions, and (c) that provide enough samples to have statistical significance, (d) where each period is long enough to span a full synoptic cycle (~5–15 days) and includes a relevant exceedance. By considering full synoptic cycles, the model is forced to simulate the conditions before and after an exceedance as well, allowing confirmation that the model can forecast non-exceedances as well as exceedances (e.g., Biswas et al., 2001). Additional reasons for choosing specific periods include (e) periods during intensive AQ field experiments, for which more detailed diagnostic evaluations can be performed, and (f) periods that have already been modeled, so that either model performance is already known to be satisfactory or else comparable results are available for comparison from a peer AQ model.

In the case of long-term effects or AQ standards, correspondingly longer simulation periods will be required. Continuing advances in computer technology have meant that running AQ models for periods as long as a year or more has become feasible (e.g., Eder and Yu, 2006),

but even so there are still likely to be representativeness issues due to interannual meteorological variability (e.g., Brook and Johnson, 2000). Choosing periods that satisfy short-term selection criteria (a), (e), and (f) is desirable.

6. Generate the AQ model input files and evaluate results.

This step builds upon the previous four steps and will usually require (a) preparing geophysical fields for the selected domain and grid, (b) running a prognostic meteorological model with some type of data assimilation for the simulation periods selected in Step 5 to prepare meteorological input files, and (c) running an emissions processing system for the same simulation periods to prepare emissions input files for a base case and often a number of emission scenarios as well. For any regional (i.e., limited-area) AQ modeling system, it may also be necessary (d) to run both global meteorological models and AQ models or to analyze available chemical climatologies (e.g., Logan, 1999) in order to provide chemical boundary conditions. In preparing the input emission files, the size of the model domain will dictate how many emission inventories will need to be processed. For many North American model domains, it will be necessary to process both Canadian and U.S. or both U.S. and Mexican inventories, or in some cases, all three. For AQ modeling elsewhere in the world, such as East Asia or Europe, it is also likely that multiple national inventories will need to be combined.

As part of this step, it is also important to check the input files produced so as to ensure that the inputs provided to the AQ model are as accurate and credible as possible. As discussed before, meteorology drives the AQ simulation and the AQ model results are very sensitive to the meteorological inputs in complex and nonlinear ways. At a minimum, an operational evaluation should be performed against meteorological measurements: the suite of meteorological parameters considered should include temperature, humidity, wind speed, wind direction, cloud-related fields, precipitation, and, if possible, planetary-boundary-layer depth (e.g., Hogrefe et al., 2001a, Smyth et al., 2006). Evaluation of the processed emissions is not as straightforward, but current emissions processsing systems produce a range of log files and summary tables that can be checked for warning and error messages and for continuity, consistency, and plausibility, particularly when data from more than one country or jurisdiction are being combined. Visualization tools can also be applied to check the spatial and temporal patterns contained in the processed emission files. The emission files for various emissions scenarios should probably receive even greater scrutiny since extensive manipulations were likely required to transform current inventories to account for various socio-economic projections and control measures. The inclusion or exclusion (depending upon the modeling/analysis protocol) and the treatment of natural emissions such as wildfires and windblown dust should also be checked. Sections 13 and 14 of U.S. EPA (2005b) provide useful and detailed discussions concerning this step.

7. Perform base-case AQ simulation and evaluate results.

The selected AQ model should have already undergone performance evaluations, but these may have been for other time periods. In this step the AQ model is run for the base case for the time periods selected in Step 5 and its performance is evaluated so as to characterize and quantify the overall modeling system's performance (i.e., including the treatment of emissions and meteorology) and to determine whether that performance is acceptable. Given known model limitations, errors, and uncertainties, Russell and Dennis (2000), Reid et al. (2003), Seigneur and Moran (2004), and the U.S. EPA (2001, 2005b) have all argued that this performance evaluation for the base case should not be restricted to just a basic operational evaluation against surface measurements of one or two pollutants, but instead should include a broader set of analyses that all feed into a "weight-of-evidence" judgement. Clearly, such an evaluation is somewhat open-ended and not prescriptive, but it should be more likely to lead to a correct judgement.

This broader set of analyses, many of them independent tests, could include any of the

following possibilities:

- a more comprehensive operational evaluation, including consideration of a suite of ozone and PM precursors and other related gas-phase species (e.g., NO_x, NO_y, CO, NH_3, H_2O_2, HNO_3, individual VOC species) and PM chemical components both at the surface and aloft (e.g., Biswas et al., 2001);
- sensitivity tests based on alternate configurations of the AQ model, including the use of a different emissions processing system or meteorological model, different rate constants and other model parameters, different grid resolutions, different chemistry mechanisms, and different boundary conditions (e.g., Mallet and Sportisse, 2006);
- bounding tests in which emissions inputs are either increased or decreased to reflect the magnitude of uncertainties related to those inputs;
- comparisons with results from peer AQ models, including operational AQ forecast models, if these have been run for the same region and time period(s) (e.g., Hogrefe et al., 2001b; Biswas et al., 2001; McKeen et al., 2005);
- if appropriate, comparison with receptor-based model results (e.g., Marmur et al., 2006);
- comparison with observation-based models or analyses for chemical regimes, including indicator species ratios and gas ratio (e.g., Sillman et al., 1997; Stein and Lamb, 2002; Martin et al., 2004); and
- use of model probing techniques, including process analysis and direct decoupled method (e.g., Zhang et al., 2005).

8. Perform scenario AQ simulations and evaluate results.

In this last step, once the scenario simulations have been performed and the results analyzed, several additional diagnostic or comparative evaluations can be carried out to examine the reasonableness of the AQ model's response to specified emission changes, particularly if disbenefits as well as benefits are predicted to occur. These include (a) applying observation-based models for chemical regime, model probing techniques, and sensitivity/bounding tests to the scenarios, (b) comparison with the relative response functions predicted by peer AQ models for the same set of scenarios, and (c) retrospective analyses of model response to historical emission changes.

Seigneur and Moran (2004) described one comparative evaluation in which the predictions of two different PM AQ models were compared for the same three emission-change scenarios. Although the magnitudes of the responses for ozone, particulate nitrate, and $PM_{2.5}$ mass were all different between the two models for the three scenarios, the directions (i.e., sign) of the responses were the same, providing support for the general conclusions about atmospheric response. This directional consistency was particularly important in the scenario in which VOC emissions were reduced by half: both models predicted a ~30% decrease in ozone levels (at one station) but an increase in both particulate nitrate and $PM_{2.5}$ mass levels, that is, a $PM_{2.5}$ disbenefit.

The effort required to follow best AQ modeling practice and to carry out each of the above eight steps for a model application may seem overwhelming. It is worth noting, however, that this is the worst case. For a jurisdiction with a history of AQ problems, a conceptual model (Step 1) likely already exists, and some AQ modeling may have already been performed. If an active in-house or external AQ modeling team with past experience for that jurisdiction can be accessed and the AQ modeling system that they use is credible, then Steps 2, 3, and 4 may not be needed and the modelers can begin at Step 5. If a new scenario is similar to a past scenario in terms of the periods to be simulated (Step 5) or most assumptions about emissions (Step 6), then the generation of input data sets likely will not require as much effort as a completely new scenario for a new period and/or new model domain. And if the base scenario has been considered before, then Step 7 may not be required either, so that the completion of Step 8 is effectively the minimum requirement for a new modeling study.

Furthermore, given the open-endedness of some of the above steps and the reality of limited resources, it may not be possible to do as

thorough a job as policymakers and modellers would like to do. The penalty for "cutting corners" could but may not be incorrect predictions, but at a minimum it will be a greater degree of uncertainty and lower confidence in those predictions. Application of an AQ modeling system always entails many compromises, and the work that can be performed for the resources that are available is just one more compromise. However, the 8-step set of modeling best practices described above should be viewed as a goal to be approached as closely as possible if AQ modelers are to provide their clients with the best possible guidance.

3.5 Combining Measurements, Emissions and Model Output

Independently, emission inventories, measurement programs and models are essential tools for AQ risk management and for describing the state of the atmosphere. A range of new methods are being explored that combine emissions and measured and modeled concentration fields to expand the capability of daily, routine AQ forecasts and improve estimates of intra-urban and inter-urban variation in long-term or chronic exposure. 'Fusing' these diverse information sources together to support a wide range of health and air quality studies, as well as real time data reporting and analysis, holds considerable promise. Figure 3.5.1 presents a conceptual picture of the types of multi-scale information that can potentially be assimilated or 'fused' into a complete picture of the spatial variation in air pollutant concentrations. Although they are a source of input to the AQ models, emission inventories may also represent an independent source of spatial information and/or a predictor for use in empirical models.

Data assimilation routines using real-time data and model output are now being applied on a continuous basis to characterize large scale patterns across North America. The amount and quality of information available varies from pollutant to pollutant and geographically. At present, in North America, ozone is the most advanced, while routines for $PM_{2.5}$ are being developed. Figure 3.5.2 presents ozone concentrations across eastern North America derived from the Canadian Meteorological Centre air quality forecasting system and ozone data compiled under the AirNow program (Menard and Robichaud, 2005). This image is derived from gridded ozone concentrations that are produced by combining observations with model predicted concentrations to 'interpolate' and prepare the data for computing future concentrations using an air quality forecasting model. Hourly concentrations at each grid point were used to compute the maximum 8 hr average concentration at each point on each day. These concentrations were then combined for the five month period known as the "ozone season."

At present AQ models only assimilate surface observations, but approaches for "chemical data assimilation" are undergoing considerable research and development (Menard, 2006). The long term goal is to begin utilizing observations from satellites and possibly other irregular sources of information (e.g., aircraft). The most advanced satellite instrument is OMI (Ozone Monitoring Instrument) on the Aura spacecraft, which was launched in 2004 (Schoeberl et al., 2004). In terms of the common air pollutants, daily, 13x24 km resolution observations for O_3, NO_2, SO_2 and aerosols are being measured. Devising the appropriate procedures for assimilating and/or interpreting such data presents a significant scientific challenge. Even with a satellite such as Aura, observations are, at best, once per day if no clouds obscure the measurements.

The map of NO_2 over the northeast of North America shown in Figure 3.5.3 provides an indication of the capabilities of OMI. It is important to note that satellite observations of trace gases and aerosols are 'total vertical column' amounts (i.e., not necessarily surface conditions). Research is needed to further improve the processing of satellite data from the raw signals and from other supporting data (e.g., correcting for clouds and variations in surface albedo) and then in deriving boundary-layer and/or surface concentrations, as well as vertical profiles. The latter of which requires, in itself, the combined use of AQ and meteorological models and surface observations.

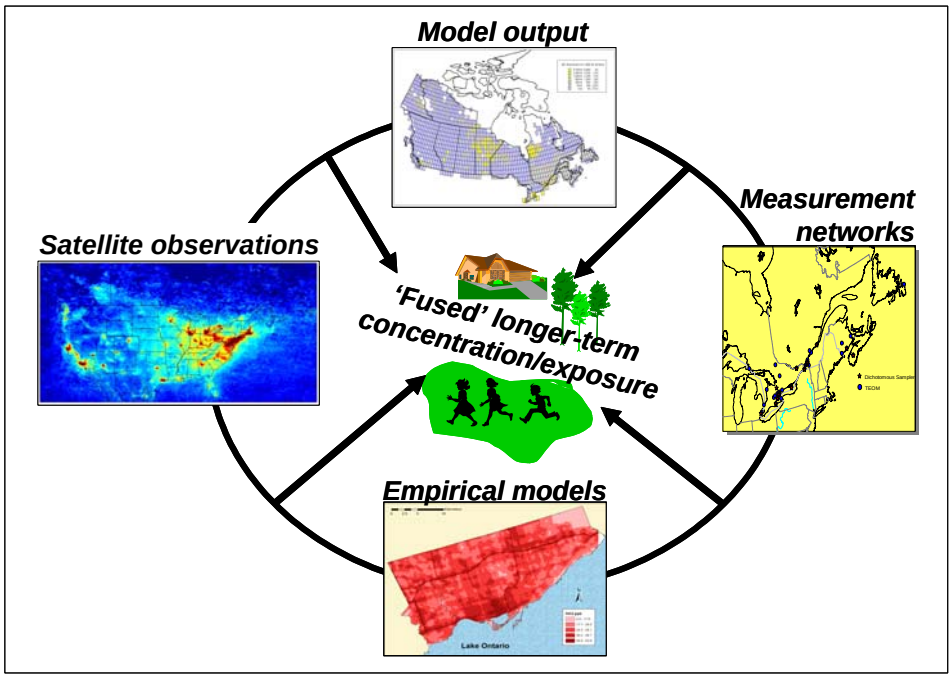

Figure 3.5.1. Conceptual system of fused or assimilated data for estimating longer-term ambient concentration patterns or chronic air pollutant exposure at any geographic location and at any scale, from regional to neighbourhood. The system needs to be flexible in terms of which source(s) of information the estimate relies most heavily upon at each scale of interest and must be capable of providing estimates of location-dependent uncertainty. This uncertainty will vary geographically due to inconsistencies in the amount of information that is available at the finer scales.

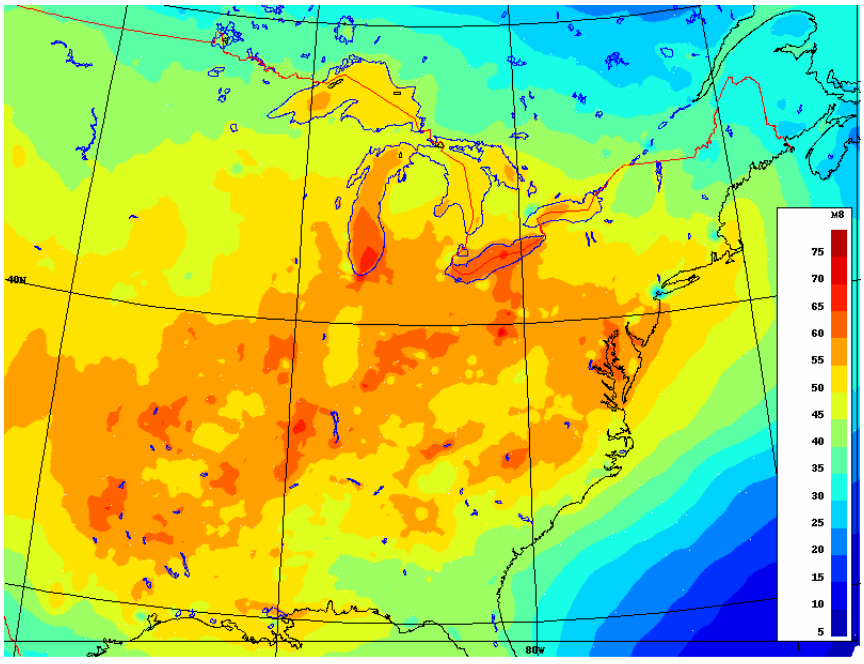

Figure 3.5.2. Average daily 8 hour maximum ozone (ppb) for summer (May-Sept.) 2005.

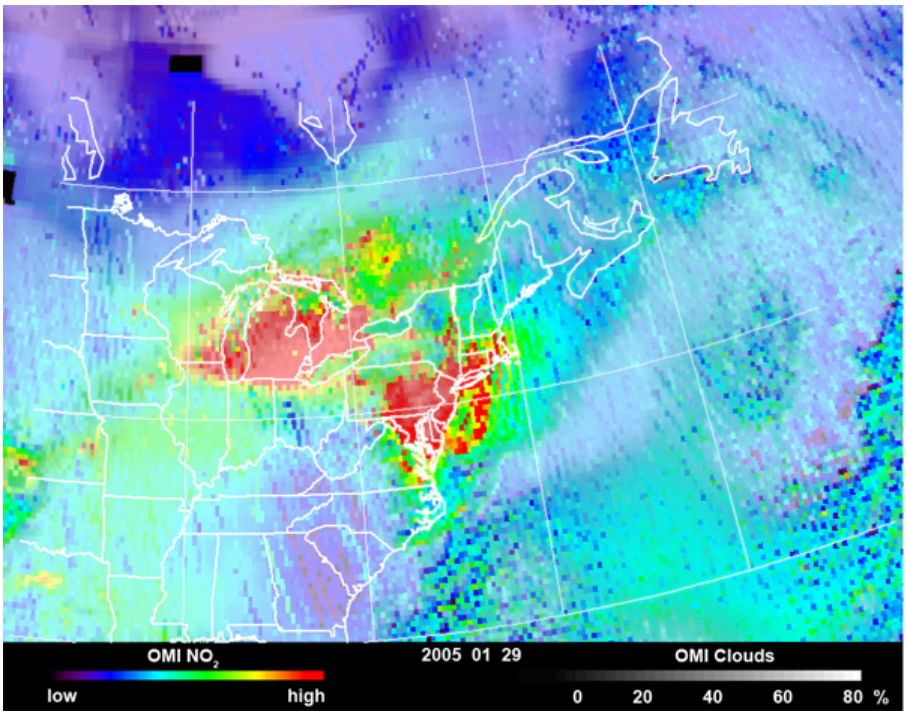

Figure 3.5.3. Nitrogen dioxide (NO_2) observations from the ozone monitoring instrument (OMI) on the AURA satellite. The red areas show high amounts, the purple areas show regions of clouds which shield the sensor from nitrogen dioxide below the clouds. On Jan. 29, 2005, and subsequent days there was a large $PM_{2.5}$ event due to high particle nitrate concentrations. This period represents the first wintertime air quality advisory ever issued in the Province of Ontario. Air quality alerts were also issued for Michigan during this period. The large concentrations over southern Michigan and southwestern Ontario are consistent with the surface conditions observed during the period. The OMI instrument was provided to the Aura Mission by the Dutch and Finnish space and meteorological agencies. Image generated by OMI NO_2 team. (contact James F. Gleason, NASA/GSFC).

Nonetheless, satellite data represents a valuable source of information because it is freely available and provides global coverage - air pollutant information can be obtained where no monitoring exists. In addition to the initialization of AQ models, the spatial patterns derived from satellite observations (i.e., across days, weeks or months) are well-suited to determining, in an internally-consistent manner, gradients in chronic exposure across large regions and among different countries. Thus far, aerosol observations ($PM_{2.5}$) have received the most attention for this purpose (e.g., Liu et al., 2005).

Population exposure to ambient air pollution occurs at neighbourhood scales. This is beyond the resolution of all the sources of information discussed above (i.e., Figures 3.5.1, 3.5.2, and 3.5.3). Furthermore, it is unlikely that any of these will be able to resolve such scales in the future and in terms of the types of deterministic AQ models discussed above, it is not reasonable to expect the meteorology to be modeled or for emissions information to provided at a fine enough scale. At best, new parameterization schemes or independent emissions models will be developed to treat sub-grid scale features to enable AQ models to reliably predict at the 1 to 5 km resolution. However, research is needed to determine how best to use such models to predict population exposure changes and their uncertainties within these grid sizes so that the costs and benefits of local scale air quality risk management strategies can be evaluated.

To resolve urban to neighbourhood scale exposure patterns for health studies a variety of approaches are currently being applied. These range from interpolation of monitoring site data (e.g., Jerrett et al., 2005) to exposure surrogates such as distance-to-roadway and traffic counts (e.g. Hoek et al., 2002) to small scale dispersion models and/or combinations of both (Wu et al., 2005; Cyrys et al., 2005). The local scale exposures of interest have generally been associated with traffic since data on road networks are readily available. However, a wider range of emission sources have been included in some ambient air pollutant exposure modeling efforts (e.g., Gram et al., 2003).

Geographic Information Systems (GIS) have proven to be useful for mapping exposure patterns, integrating different sources of information and in developing land-use regression models (Brauer 2006a). Intra-urban chronic exposure estimates have been derived using LUR for several cities (Brauer et al., 2003; Kanaroglou et al., 2005; Sahsuvaroglu et al., 2006; Gilbert et al., 2005; Silva et al., 2006; Setton et al., 2006; Luginaah et al., 2006; Brauer 2006b). Cyrys et al. (2005) compared both LUR and dispersion model estimates for NO_2 and $PM_{2.5}$ and reported that for their cohort of interest in Munich, Germany, the two approaches led to similar exposure classifications. These results and most other LUR efforts have focused on traffic-related pollutants (e.g., NO_2 and sometimes $PM_{2.5}$). However, recent studies in Windsor, Ontario, have expanded the dependent variables in LUR to include SO_2, benzene and toluene (Wheeler et al., 2006).

The empirical model image in Figure 3.5.1 is an example of the NO_2 surface predicted by LUR for Toronto, Ontario (Kanaroglou et al., 2005). The small scale spatial variability (i.e., neighbourhood scale or better) produced by applying the LUR for all points in a GIS database appears more-realistic compared to the pattern obtained using interpolation, with respect to the known distribution of traffic (Jerrett et al., 2004). A LUR model, once developed for the area of interest, also provides the capability of estimating chronic exposures for each member of a health study cohort if their addresses are known. Ideally, such estimates should be spot-checked with independent measurements within residential areas, inside a variety of homes and also in comparison to personal exposure measurements. This could potentially lead to the coupling of LUR models for outdoor, at home, concentrations with individual

exposure models that consider home characteristics and time activity.

3.6 Conclusions

Improving or maintaining air quality is a science and technology based activity, requiring governmental commitment to invest in the tools needed to reach informed decisions. In jurisdictions where AQ management has not been a priority, solutions to the current problems may be relatively straightforward, such as eliminating local/residential burning for cooking and heating fuels. However, even when all the obvious and/or cost-effective measures have been implemented AQ problems can still persist and even become worse due to economic and/or population growth. In this situation, which may be the case for many developed countries, the best approach(es) to improve AQ are not as easy to identify. Accurate, comprehensive emission inventories, AQ measurements and models are therefore essential to make headway. However, they need to be applied intelligently following, as much as possible, best practices as informed by experience. This includes a well-developed conceptual model of the relationship between emissions and observations. To make headway air quality targets (or standards or objectives) are also needed and clearly, public health protection is one of the main motivations behind such targets. This necessitates establishing a quantitative link between air pollutant levels and health impacts, such as a concentration response relationship, which is dependent upon the availability of measurements. However, there is a need to develop air quality risk management methods that integrate multiple environmental issues, not just human health concerns.

This chapter reviewed several key issues related to the development, use and improvement of emissions, measurements and models for AQ management. Each can provide useful information for AQ risk management, and when they are considered together they can provide additional insights and guidance. However, AQ models depend upon the availability of information about measurements, emissions, and meteorology whereas the converse is not true. AQ modeling should thus follow and not precede the development of measurement and emissions information.

Accurate emission inventories are the foundation of all air quality management programs. They provide the essential information needed to understand the effects of air pollutants on human and ecosystem health, to identify which sources need to be controlled in order to protect health and the environment, and they provide the information needed to determine whether or not actions taken to reduce emissions have been effective.

In principle, the development of emission inventories would seem to be a relatively straightforward process, but in practice their production is found to be a very complex and demanding task. As initial actions to reduce emissions from large point sources find success, understanding and addressing residual air quality problems requires greater effort and emission inventories of increasing sophistication. Fortunately, much has been learned over the past 40 years that can make the development of new inventories a more systematic process. New measurement technologies and better understanding of the chemistry and physics of pollutant formation will continue to multiply the number of sources that can be measured directly and assure that these measurements reflect what is actually entering the atmosphere. Likewise, new methods for deducing and characterizing uncertainty will result in better understanding of the accuracy with which we know primary and precursor emissions. Finally, better data management software, the ubiquitous availability of low-cost, high-end computing, and growing availability of high bandwidth communications have made the development, maintenance, dissemination and use of large data sets practical for nearly everyone.

Air quality measurements are essential for

public health protection. They define the problem to be managed, serve as the basis for determining the current level of health risk a given population is experiencing and consequently for prioritizing the need for reductions. Measurements are also critical for evaluating the effectiveness of AQ management strategies and altering such strategies if the desired outcomes are not being achieved.

Detailed analysis of measurement data can help target the most effective approaches to reduce ambient concentrations and, hopefully in the future, the optimum (i.e., cost effective) approach to protect public health. Full understanding of particulate matter, in terms of impacts, formation processes and optimal control strategies remains the air pollutant requiring the greatest attention in terms of detailed measurement studies and data interpretation efforts.

When measurement programs are forward-looking, pushing the limits of what can be routinely monitored, they can provide new insights regarding additional air pollutants of concern and can support future epidemiological studies to uncover new risks to the population. Ultimately, the availability of air quality measurements dictates what can be studied and thus, there is a continual need to expand the pollutants measured and the location and temporal resolution of such measurements. Combining the data from a variety of measurement approaches, including remote sensing, with the data from both physical and empirical models provides an improved picture of spatial and temporal patterns. These improvements are providing better AQ information to scientists, the public and decision-makers and ultimately can be expected to lead to a better understanding of AQ impacts and of possible approaches to protect public health.

AQ models are able to quantify the links between emissions of primary pollutants or precursors of secondary pollutants and ambient pollutant concentrations and other physiologically, environmentally, and optically important properties. They are the only tool available that can predict, based on possible future emission levels, spatially and temporally resolved air concentration and deposition patterns and that can address multiple pollutants simultaneously and quantify possible co-benefits. AQ models can also account for the impacts of nonlinear processes and are able to predict whether a candidate abatement strategy will lead to benefits or disbenefits or both.

There are a large number of possible sources of AQ model error and uncertainty ranging from not understanding the underlying science and truncation errors intrinsic to the numerical techniques employed by the model to uncertainties in the input emissions and model-measurement incommensurability. There are even more ways for these numerous sources of error and uncertainty to interact, often nonlinearly and sometimes cancelling out (so-called "compensating errors"). As a consequence, AQ model uncertainty is impossible to quantify but it is possible to characterize through model performance evaluations, model intercomparisons, and sensitivity and bounding tests. There are also varying degrees of uncertainties across pollutants and their components.

Box (1979) wrote that "All models are wrong, but some are useful." For an AQ model to be "useful," it should be credible. For it to be credible, it should give the right answers for the right reasons. AQ model credibility is established through model review, model performance evaluations, and successful model applications. However, model credibility is always provisional, so model evaluation (and model improvement) should be an ongoing process.

A key question in applying AQ models is: "How accurate does a model need to be?" Reid et al. (2003) suggested that the general answer is that "model predictions should be good enough that model uncertainty does not affect the decisions that are based on the predictions." AQ modeling uncertainty can be managed and limited by following "best practice" at all stages of the modeling

process. Best practice basically boils down to thoughtful and careful selection, set up, and application of a credible AQ modeling system accompanied by careful scrutiny and consistency checking of the results by various means, including measurements and results from both alternate configurations of the selected AQ model and from other AQ models. The credibility of the model predictions for a given application is then determined based on a weight-of-evidence judgment that considers all of the evaluation results. This process is not at all "cut and dried" — it is much more in the nature of applied research than a routine activity.

Acknowledgement

The authors would like to acknowledge the contribution of Michael Brauer of the School of Occupational and Environmental Hygiene, & Department of Medicine, & Atmospheric Science Programme at the University of British Columbia in providing insightful review comments on a draft version of this paper.

3.7 References

Alapaty, K., Pleim, J.E., Raman, S., Niyogi, D.S., and Byun, D.W. 1997. Simulation of atmospheric boundary layer processes using local- and non-local closure schemes. *J. Appl. Meteor.* 36:214-233.

Ansari, A.S., and Pandis, S.N. 1999. An analysis of four models predicting the partitioning of semi-volatile inorganic aerosol components. *Aerosol Sci. Technol.* 31:129-153.

Bell M.L. 2006. The use of ambient air quality modeling to estimate individual and population exposure for human health research: A case study of ozone in the Northern Georgia Region of the United States. *Environ. Int.* 32:586-593.

Berge, E., Huang, H.-C., Chang, J., and Liu, T.-H. 2001. A study of the importance of initial conditions for photochemical oxidant modeling. *J. Geophys. Res.* 106:1347-1363.

Berkowitz, C.M., Fast, J.D., Springston, S.R., Larsen, R.J., Spicer, C.W., Doskey, P.V., Hubbe, J.M., and Plastridge, R. 1998. Formation mechanisms and chemical characteristics of elevated photochemical layers over the northeast United States. *J. Geophys. Res.* 103:10631-10647.

Biswas, J., Hogrefe, C., Rao, S.T., Hao, W., and Sistla, G. 2001. Evaluating the performance of regional-scale photochemical modeling systems: Part 3. Precursor predictions. *Atmos. Environ.* 35:4175-4188.

Blanchard, C.L. 1999. Methods for attributing ambient air pollutants to emission sources. *Annu. Rev. Energy Env.* 24:329-365.

Box, G.E.P. 1979. Robustness in the strategy of scientific model building. In Robustness in Statistics, R.L. Launer and G.N. Wilkinson (eds). New York NY: Academic Press, p. 202.

Brauer, M. 2006a. Health and air quality: Border air quality study update. Presented at the International Airshed Strategy Coordinating Committee. Parksville, BC May 31, 2006 www.cher.ubc.ca/UBCBAQS/Presentations/BAQS-IAS%20May31%202006.pdf.

Brauer, M., and Brook, J.R. 1997. Ozone personal exposures and health effects for selected groups residing in the Fraser Valley. *Atmos. Envir.* 31:2113-2121.

Brauer, M., Henderson, S.B., Jerrett, M., and Beckerman, B. 2006b. Land Use Regression Modeling of Nitrogen Oxides and Fine Particulate Matter in the Greater Vancouver Regional District, presented at the Border Air Quality Strategy Meeting. www.cher.ubc.ca/UBCBAQS/BAQS.

Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., Heinrich, J., Cyrys, J., Bellander, T., Lewne, M., and Brunekreef, B. 2003. Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology* 14:228-239.

Brook, J.R., Dann, T.F., and Bonvalot, Y. 1999. Observations and interpretations from the Canadian fine particle monitoring

program. *J. Air Waste Manage. Assoc.* 49:35-44.

Brook, J.R., and Johnson, D. 2000. Identification of representative warm season periods for regional air quality (ozone) model simulations. *Atmos. Environ.* 34:1591-1599.

Brook, J.R., and Spengler, J.D. 1995. Exposure to fine particle acidity and sulfate in 24 North American communities: The relationship between single-year observations and long-term exposures. *J. Air Waste Manage. Assoc.* 45:709-721.

Brook J., Dann T., Mabury S., and Burnett R. 2002. Health Effects Study with Enhanced Characterization of Urban Air Pollutant Mix. Executive Summary available at: www.hc-sc.gc.ca/sr-sr/finance/tsri-irst/proj/urb-air/tsri-41_e.html

Brook J.R., Burnett R.T., Dann T.F., Cakmak S., Goldberg M.S., and Fan X. 2007. Further interpretation of the acute effect of nitrogen dioxide observed in Canadian time series studies. *J Expos. Anal. Environ. Epi.* (submitted).

Brost, R.A. 1988. The sensitivity of input parameters of atmospheric concentrations simulated by a regional chemical model. *J. Geophys. Res.* 93:2371-2387.

Brown, S.S., Ryerson, T.B., Wollny, A.G., Brock, C.A., Peltier, R., Sullivan, A.P., Weber, R.J., Dubé, W.P., Trainer, M., Meagher, J.F., Fehsenfeld, F.C., and Ravishankara, A.R. 2006. Variability in nocturnal nitrogen oxide processing and its role in regional air quality. *Science* 311:67-70.

Bukowiecki, N., Dommen, J., Prévôt, A.S.H., Richter, R., Weingartner, E., and Baltensperger, U. 2002. A mobile pollutant measurement laboratory - measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution. *Atmos. Environ.* 36:5569-5579.

Burnett R.T., Cakmak S., Brook J.R., and Krewski D. 1997. The role of particulate size and chemistry in the association between summertime ambient air pollution and hospitalization for cardio-respiratory diseases. *Environ. Health Persp.* 105:614-620.

Burnett, R.T, Stieb, D., Brook, J.R., Cakmak, S., Dales, R., Raizenne, M., Vincent, R., and Dann, T. 2004. The short-term effects of nitrogen dioxide on mortality in Canadian cities. *Arch. Environ. Health* 59:228-237.

Chow, J. 1995. Measurement methods to determine compliance with ambient air quality standards for suspended particles. *J. Air Waste Manage. Assoc.* 45:320-382.

Chow, J. 2006. Introduction to the A&WMA Critical Review. Health effects of fine particulate air pollution: Lines that connect. *J. Air Waste Manage. Assoc.* 56:707-708.

CRC. 2004. Evaluation of the U.S. EPA MOBILE6 Highway Vehicle Emission Factor Model, Final Report CRC Project E-64, ENVIRON International Corp., Novato, CA.

Cope, M.E., Hess, G.D., Lee, S., Tory, K., Azzi, M., Carras, J., Lilley, W., Manins, P.C., Nelson, P., Ng, L., Puri, K., Wong, N., Walsh, S., and Young, M. 2004. The Australian Air Quality Forecasting System. Part I., Project description and early outcomes. *J. Appl. Meteor.* 43:649-662.

Cyrys, J., Hochadel, M., Gehring, U., Hoek, G., Diegmann, V., Brunekreef, B., and Heinrich, J. 2005. GIS-based estimation of exposure to particulate matter and NO_2 in an urban area: Stochastic versus dispersion modeling. *Environ. Health Persp.* 113:987-992.

Dabberdt, W.F., Carroll, M.A., Baumgardner, D., Carmichael, G., Cohen, R., Dye, T., Ellis, J., Grell, G., Grimmond, S., Hanna, S., Irwin, J., Lamb, B., Madronich, S., McQueen, J., Meagher, J., Odman, T., Pleim, J., Schmid, H.P., and Westphal, D. 2004. Meteorological research needs for improved air quality forecasting. *Bull. Amer. Meteor. Soc.* 85:563-586.

DEFRA. Department for Environment, Food and Rural Affairs 2006. The Air Quality Strategy for England, Scotland, Wales and Northern Ireland. A consultation document

on options for further improvements in air quality. Volume 1. London. www.defra.gov.uk/corporate/consult/airqualstrat-review/index.htm.

Dennis, R.L., Binkowski, F.S., Clark, T.L., McHenry, J.N., Reynolds, S., and Seilkop, S.K. 1990. Selected applications of RADM (Part II). In NAPAP SOS/T Volume 1, Report 5 (App. 5F), National Acid Precipitation Assessment Program: State of Science and Technology, P.M. Irving (ed). Washington DC: National Acid Precipitation Assessment Program, 37 pp.

Dennis, R.L., McHenry, J.N., Barchet, W.R., Binkowski, F.S., and Byun, D.W. 1993. Correcting RADM's sulfate underprediction: discovery and correction of model errors and testing the corrections through comparisons against field data. *Atmos. Environ*. 27A:975-997.

Dickson, R.J., and Oliver, W.R. 1991. Emissions models for regional air quality studies. *Environ. Sci. Technol*. 25:1533-1535.

Dockery, D.W., Cunningham, J., Damokosh, A.I., Neas, L.M., Spengler, J.D., Koutrakis, P., Ware, J.H., Raizenne, M., and Speizer, F. 1996. Health effects of acid aerosols on North American children: Respiratory Symptoms. *Environ. Health Persp.* 104:500-505.

Ebelt S., Wilson W.E., and Brauer M. 2005. Exposure to ambient and nonambient components of particulate matter: a comparison of health effects. *Epidemiology* 16:396-405.

Eder, B., Kang, D., Mathur, R., Yu, S., and Schere, K. 2006. An operational evaluation of the Eta-CMAQ air quality forecast model. *Atmos. Environ*. 40:4894-4905.

Eder, B., and Yu, S. 2006. A performance evaluation of the 2004 release of Models-3 CMAQ. *Atmos. Environ*. 40:4811-4824.

EMEP 2003. Transboundary acidification, eutrophication and ground level ozone in Europe Status report 1/2003. Part II: Unified EMEP model performance. ISSN 0806-4520, July, Norwegian Meteorological Institute, Oslo, Norway, 169 pp. www.emep.int//publ/common_publications.html

EMEP 2005. Transboundary acidification, eutrophication and ground level ozone in Europe in 2003. Status report 1/2005, ISSN 0806-4520, August, Norwegian Meteorological Institute, Oslo, Norway, 229 pp.
www.emep.int//publ/common_publications.html

EPRI 1994. *Electric Utility Trace Substances Synthesis Report*, EPRI TR-104614, Electric Power Research Institute, Palo Alto, CA.

Fan, X., Brook, J.R., and Mabury, S. 2003. Sampling semivolatile organic species associated with $PM_{2.5}$ using an integrated organic gas and particle sampler. *Envir. Sci. Technol.* 37:3145-3151.

Fox, D.G. 1981. Judging air quality model performance. *Bull. Amer. Meteor. Soc*. 62:599-609.

Fox, D.G. 1984. Uncertainty in air quality modeling. *Bull. Amer. Meteor. Soc*. 65:27-36.

Frost, G.J., McKeen, S.A., Trainer, M., Ryerson, T.B., Neuman, J.A., Roberts, J.M., Swanson, A., Holloway, J.S., Sueper, D.T., Fortin, T., Parrish, D.D., Fehsenfeld, F.C., Flocke, F., Peckham, S.E., Grell, G.A., Kowal, D., Cartwright, J., Auerbach, N., and Habermann, T. 2006. Effects of changing power plant NOx emissions on ozone in the eastern United States: Proof of concept. *J. Geophys. Res*. 111, D12306, doi:10.1029/2005JD006354.

Fruin, S.A., St Denis, M.J., Winer, A.M., Colome, S.D., and Lurmann, F.W. 2001. Reductions in human benzene exposure in the California South Coast Basin. *Atmos. Environ.* 35, 1069–1077.

Geron, C.D., Nie, D., Arnts, R.R., Sharkey, T.D., Singsaas, E.L., Vanderveer, P.J., Guenther, A., Sickles, J.E., and Kleindienst, T.E. 1997. Biogenic isoprene emission: model evaluation in a southeastern United States bottomland deciduous forest. *J. Geophys. Res*. 102:18,889-18,901.

Gilbert, N., Goldberg, M.S., Beckerman, B., Brook, J.R., and Jerrett, M. 2005. Assessing spatial variability of ambient nitrogen dioxide in Montreal, Canada, with a land-use regression model. *J. Air Waste Manage. Assoc.* 55:1059-1063.

Gilliland, A.B., Appel, K.W., Pinder, R.W., and Dennis, R.L. 2006. Seasonal NH_3 emissions for the continental United States: Inverse model estimation and evaluation. *Atmos. Environ*. 40:4986-4998.

Gilliand, A.B., Dennis, R.L., Roselle, S.J., and Pierce, T.E. 2003. Seasonal NH_3 emission estimates for the Eastern United States using ammonium wet concentrations and an inverse modeling method. *J. Geophys. Res.* 108:4477.

Gram, F., Nafstadb, P., and Haeheim, L.L. 2003. Estimating residential air pollution exposure among citizens in Oslo 1974-1998 using a geographical information system. *J. Environ. Monit*. 5:541–546.

Grasso, L.D. 2000. The differentiation between grid spacing and resolution and their application to numerical modeling. *Bull. Amer. Meteor. Soc.* 81:579-580.

Guo P., Xu X., and Brook J.R. 2006. Development of an air quality sampling strategy for mobile measurements. Paper #522. Air & Waste Management Association Annual Meeting, New Orleans, LA., June 2006.

Haltiner, G.J., and Williams, R.T. 1980. Numerical Prediction and Dynamic Meteorology (Second Edition). New York: John Wiley & Sons, 477 pp.

Hass, H., Builtjes, P.J.H., Simpson, D., and Stern, R. 1997. Comparison of model results obtained with several European regional air quality models. *Atmos. Environ*. 31:3259-3279.

Heald, C.L., Jacob, D.J., Fiore, A.M., Emmons, L.K., Gille, J.C., Deeter, M.N., Warner, J., Edwards, D.P., Crawford, J.H., Hamlin, A.J., Sachse, G.W., Browell, E.V., Avery, M.A., Vay, S.A., Westberg, D.J., Blake, D.R., Singh, H.B., Sandholm, S.T., Talbot, R.W., and Fuelberg, H.E. 2003. Asian outflow and trans-Pacific transport of carbon monoxide and ozone pollution: an integrated satellite, aircraft, and model perspective. *J. Geophys. Res*. 108-D24, 4804, doi:10.1029/2003JD003507.

Heald, C.L., Jacob, D.J., Park, R.J., Russell, L.M., Huebert, B.J., Seinfeld, J.H., Liao, H., and Weber, R.J. 2005. A large organic aerosol source in the free troposphere missing from current models. *Geophys. Res. Lett*. 32, L18809, doi:10.1029/2005GL023831.

HEI 2003. Assessing the Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research. Report#: Communi-cation 11. Health Effects Institute, Sept. 2003. http://pubs.healtheffects.org/

Hodzic, A., Vautard, R., Chazette, P., Menut, L., and Bessagnet, B. 2006. Aerosol chemical and optical properties over the Paris area within ESQUIF project. *Atmos. Chem. Phys. Discuss*. 6:401-454.

Hoek, G., Brunekreef, B., Goldbohm, S., Fischer, P., and van den Brandt, P.A. 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. *Lancet* 360:1203–1209.

Hogrefe, C., Rao, S.T., Kasibhatla, P., Kallos, G., Tremback, C.J., Hao, W., Olerud, D., Xiu, A., McHenry, J., and Alapaty, K. 2001a. Evaluating the performance of regional-scale photochemical modeling systems: Part 1. Meteorological predictions. *Atmos. Environ*. 35:4159-4174.

Hogrefe, C., Rao, S.T., Kasibhatla, P., Hao, W., Sistla, G., Mathur, R., and McHenry, J. 2001b. Evaluating the performance of regional-scale photochemical modeling systems: Part 2. Ozone predictions. *Atmos. Environ*. 35:4175-4188.

Hogrefe, C., Sistla, G., Zalewsky, E., Hao, W., and Ku, J.-Y. 2003. An assessment of the emissions inventory processing systems EMS-2001 and SMOKE in grid-based air quality models. *J. Air Waste Manage. Assoc*. 53:1121-1129.

Houyoux, M.R., Vukovich, J.M., Coats, Jr.,

C.J., and Wheeler, N.J.M. 2000. Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. *J. Geophys. Res.* 105:9079-9090.

Jacobson, M.Z. 1999. *Fundamentals of Atmospheric Modeling*. Cambridge UK: Cambridge University Press, 656 pp.

Jerrett, M., Arain, M.A., Kanaroglou, P., Beckerman, B., Crouse, D., Gilbert, N.L., Brook, J.R., and Finkelstein, N. 2005. Modelling the intra-urban variability of ambient traffic pollution in Toronto, Canada. *J. Toxicol. Environ. Health* (in press).

Jerrett, M., Burnett, R.T., Ma, R., Pope, C.A., Krewski, D., Newbold, K.B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E.E., and Thun, M.J. 2005. Spatial analysis of air pollution and mortality in Los Angeles. *Epidemiolo*gy 16:727-736.

Jerrett, M., Newbold K. B., Burnett, R.T., Thurston, G., Lall, R., Pope, C. A. III, Ma, R., De Luca, P., Thun, M., Calle, J., and Krewski, D. Geographies of uncertainty in the health benefits of air quality improvements. (submitted).

Kanoraglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., and Brook, J.R. 2005. Establishing an air pollution monitoring network for intra-urban population exposure assessment: a location-allocation approach. *Atmos. Environ.* 39:2399-2409.

Karamchandani, P., and Venkatram, A. 1992. The role of non-precipitating clouds in producing ambient sulfate during summer: Results from simulations with the Acid Deposition and Oxidant Model (ADOM). *Atmos. Environ*. 26A:1041-1052.

Kim, D., Sass-Kortsak, A., Purdham, J.T., Dales, R.E., and Brook, J.R. 2006. Associations between personal exposures and fixed-site ambient measurements of fine particulate matter, nitrogen dioxide and carbon monoxide in Toronto, Canada. *J. Expos. Anal. Environ. Epi.* 15:172-183.

Kleeman, M.J., and Cass, G.R. 1999a. Identifying the effect of individual emissions sources on particulate air quality within a photochemical aerosol processes trajectory model. *Atmos. Environ*. 33:4597-4613.

Kleeman, M.J., and Cass, G.R. 1999b. Effects of emissions control strategies on the size and composition distribution of urban particulate air pollution. *Environ. Sci. Technol*. 33:177-189.

Kleinman L.I. 1987. Source-receptor relations from nonlinear atmospheric models, *Atmos. Environ*. 21:1219-1225.

Knipping, E.M., Kumar, N., Pun, B.K., Seigneur, C., Wu, S.-Y., and Schichtel, B.A. 2006. Modeling regional haze during the BRAVO study using CMAQ-MADRID: 2. Source region attribution of particulate sulfate compounds. *J. Geophys. Res*. 111, D06303, doi:10.1029/2004JD005609.

Kolb, C.E., Herndon, S.C., McManus, J.B., Shorter, J.H., Zahniser, M.S. Nelson, D.D. Jayne, J.T. Canagaratna, M.R., and Worsnop, D.R. 2004. Mobile laboratory with rapid res-ponse instruments for real-time measurements of urban and regional trace gas and particulate distributions and emission source characteristics. *Environ. Sci. Technol.* 38:5694-5703.

Krupnick, A.J. 2003. Challenges to managing air pollution. NERAM document: AIRNET-NERAM Rome Conference on Strategies for Clean Air and Health, November 5-7, 2003. www.irr-neram.ca/rome/rome.html#proceed.

Kuhlbusch T., Astrid J., Hugo A., Peters A. von Klot S., Cyrys J., Wichmann H.- E., Quass U., and Bruckmann P. 2004. Analysis and design of local air quality measurements: Towards European Air Quality Health Effect Monitoring. Service Contract 070501/2004/389487/MAR/C1, EURAQHEM Final Report. http://ec.europa.eu/ environment/air/cafe/activities/pdf/euraqhem_final_report.pdf.

Kuhn, M., Builtjes, P.J.H., Poppe, D., Simpson, D., Stockwell, W.R., Andersson-Sköld, Y., Baart, A., Das, M., Fiedler, F., Hov, Ø, Kirchner, F., Makar, P.A., Milford,

J.B., Roemer, M.G.M., Ruhnke, R., Strand, A., Vogel, B., and Vogel, H. 1998. Intercomparison of the gas-phase chemistry in several chemistry and transport models. *Atmos. Environ*. 32:693-709.

Lee, P.K.H., Brook, J.R., Lu, G., Mihele, C.M., and Aklilu, Y.A. 2006. A Mobile Air-Monitoring Laboratory (CRUISER) in Action Across Canada. (submitted).

Liu, Y., Jeremy, A., Sarnat, J.A., Coull, B.A., Koutrakis, P., and Jacob, D.J. 2004. Validation of Multiangle Imaging Spectroradiometer (MISR) aerosol optical thickness measurements using Aerosol Robotic Network (AERONET) observations over the contiguous United States. *J. Geophy. Res*. 109, D06205, doi:10.1029/2003JD003981.

Liu, Y., Park, R.J., Jacob, D.J., Li, Q., Kilaru, V., and Sarnat, J.A. 2004. Mapping annual mean ground-level $PM_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *J. Geophy. Res*. 109, D22206, doi:10.1029/2004JD005025.

Liu, Y., Sarnat, J.A., Kilaru, V., Jacob, D.J., Li, Q., and Koutrakis, P. 2005. Estimating Ground-Level $PM_{2.5}$ in the Eastern United States using satellite remote sensing. *Environ. Sci. Technol*. 39:3269-3278.

Logan, J.A. 1999. An analysis of ozonesonde data for the troposphere: Recommendations for testing 3-D models and development of a gridded climatology for tropospheric ozone. *J. Geophys. Res.* 104:16,115-16,149.

Lu, G., Brook, J.R., Alfarra, M.R., Anlauf, K., Leaitch, W.R., Sharma, S., Wang, D., Worsnop, D.R., and Phinney L. 2006. Identification and characterization of inland ship plumes over Vancouver, BC. *Atmos. Environ.* 40:2767-2782.

Luginaah, I., Xu, X., Fung, K.Y., Grgicakmannion, A., Wintermute, J., Wheeler, A.A., and Brook, J. 2006. Establishing the spatial variability of ambient nitrogen dioxide in Windsor, Ontario. *Int. J. Environ. Studies,* 63:487-500.

Makar, P.M., Wiebe, H.A., Staebler, R.M., Li S.-M., and Anlauf, K.G. 1998. Measurement and modeling of nighttime particle nitrate formation. *J. Geophys. Res.*, 103:13,095-13,110.

Mallet, V., and Sportisse, B. 2005. A comprehensive study of ozone sensitivity with respect to emissions over Europe with a chemistry-transport model. *J. Geophys. Res*. 110, D022302, doi:10.1029/2005JD006234.

Mallet, V., and Sportisse, B. 2006. Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations: an ensemble approach applied to ozone modeling. *J. Geophys. Res*. 111, D01302, doi:10.1029/2005JD006149.

Manins, P.C., Cope, M.E., Hurley, P.J., Newton, P.W., Smith, N.C., and Marquez, L.O. 1998. the impact of urban development on air quality and energy use. Proc. 14th International Clean Air & Environment conf., Melbourne, Australia.

Marmur, A., Park, S.-K., Mulholland, J.A., Tolbert, P.E., and Russell, A.G. 2006. source apportionment of $PM_{2.5}$ in the southeastern United States using receptor and emissions-based models: Conceptual differences and implications for time-series health studies. *Atmos. Environ*., 40:2533-2551.

Marshall, J.D., Riley, W.J., McKone, T.E., and Nazaroff, W.W. 2003. Intake fraction of primary pollutants: motor vehicle emissions in the South Coast Air Basin. *Atmos. Environ* 37:3455-3468.

Martin, R.V., Fiore, A.M., and Van Donkelaar, A. 2004. Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions. *Geophys. Res. Lett*. 31, L06120, doi:10.1029/2004 GL019416.

Martin, R.V., Jacob, D.J., Chance, K., Kurosu, T.P., Palmer, P.I., and Evans, M.J. 2002. Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns. *J. Geophys. Res*. 108.17, 4537, doi:10.1029/2003JD003453.

McHenry, J.N., Binkowski, F.S., Dennis, R.L., Chang, J.S., and Hopkins, D. 1992. The Tagged Species Engineering Model (TSEM). *Atmos. Environ*. 26A:1041-1052.

McKeen, S., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Hsie, E.-Y., Gong, W., Bouchet, V., Ménard, S., Moffet, R., McHenry, J., McQueen, J., Tang, Y., Carmichael, G.R., Pagowski, M., Chan, A., and Dye, T. 2005. Assessment of an ensemble of seven real-time ozone forecasts over Eastern North America during the summer of 2004. *J. Geophys. Res*. 110, D21307, doi:10.1029/2005JD005858.

McMurry, P.H., Shephard, M., and Vickery, J. (eds) 2004. Particulate Matter Science for Policy Makers: A NARSTO Assessment, New York, NY: Cambridge University Press.

McNair, L.A., Harley, R.A., and Russell, A.G. 1996. Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluations. *Atmos. Environ*. 30:4291-4301.

Ménard, R., 2006. Chemical-dynamical coupling in data assimilation. Presented at the Air Quality 2006, Banff, AB, Feb. 2006. www.ec.gc.ca/aqw-aqa/program_e.html.

Ménard, R., and Robichaud, A. 2005. The Chemistry-Forecast System at the Meteorologi-cal Service of Canada. Proceedings of ECMWF Seminar on Global earth-system monitoring 5-9 Sept. 2005. ECMWF 2005- Annual Seminar Proceedings. www.ecmwf.int/publications/library/do/references/show?id=86891, www.ecmwf.int/publications/library/ecpublications/_pdf/seminar/2005/sem2005_menard.pdf.

Mendoza-Dominguez, A., and Russell, A.G. 2001. Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmos. Environ*. 35:2879-2894.

Meng, Z., Dabdub, D., and Seinfeld, J.H. 1997. Chemical coupling between atmospheric ozone and particulate matter. *Science* 277:116-119.

Miller, C.A., Hidy, G., Hales, J., Kolb, C.E., Werner, A.S., Haneke, B., Parrish, D., Frey, C., Rojas-Bracho, L., Deslauriers, M., Pennell, B., and Mobley, J.D. 2006. Air Emission Inventories in North America: A Critical Assessment, *J. Air Waste Manage. Assoc.* 56:1115-1129.

Misra, P.K., Bloxam, R., Fung, C., and Wong, S. 1989. Non-linear response of wet deposition to emissions reduction: A model study. *Atmos. Environ*. 23:671-687.

Moran, M.D. 2000. Basic aspects of mesoscale atmospheric dispersion. In Mesoscale Atmospheric Dispersion, Z. Boybeyi (ed). Southampton: WIT Press, 27-119.

Moran, M.D., and Zheng, Q. 2006. Modelling long-term sulphur and nitrogen deposition using Lagrangian chemical transport models, Proc. 14th Joint AMS/A&WMA Conf. on the Applications of Air Pollution Meteorology, Atlanta, January 30 - February 2, American Meteorological Society, Boston, 23 pp. http://ams.confex.com/ams/pdfpapers/101260.pdf.

Moran, M.D. 2005. Current and proposed emission control programs: How will acid deposition be affected? Chapter 4 in 2004 Canadian Acid Deposition Science Assessment, ISBN 0-662-68662-4, Environment Canada, Downsview, 99-162. www.msc-smc.gc.ca/saib/acid/acid_e.html.

NARSTO 2003. Particulate Matter Science for Policy Makers: A NARSTO Assessment, P. McMurry, M. Shepherd, and J. Vickery (eds). Cambridge: Cambridge University Press, 510 pp.

NARSTO 2005. Improving Emission Inventories for Effective Air Quality Management Across North America: A NARSTO Assessment, NARSTO-05-001, Oak Ridge, TN.

NRC 2004. Air Quality Management in the United States. Washington, DC: The National Academies Press.

Odum, J.R., Hoffmann, T., Bowman, F., Collins, D., Flagan, R.C., and Seinfeld, J.H. 1996. Gas/particle partitioning and

secondary organic aerosol yields. *Environ. Sci. Technol*. 30:2580-2585.

Oreskes, N., Shrader-Frechette, K., and Belitz, K. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263:641-646.

Padro, J., Puckett, K.J., and Woolridge, D.N. 1993. The sensitivity of regionally averaged O_3 and SO_2 concentrations to ADOM dry deposition velocity parameterizations. *Atmos. Environ*. 27A:2239-2242.

Palacios, M., Kirchner, F., Martilli, A., Clappier, A., Martín, F., and Rodríguez, M.E. 2002. Summer ozone episodes in the Greater Madrid area. Analyzing the ozone response to abatement strategies by modelling. *Atmos. Environ*. 36:5323-5333.

Palmer, P.I., Jacob, D.J., Jones, D.B.A., Heald, C.L., Yantosca, R.M., Logan, J.A., Sachse, G.W., and Streets, D.G. 2003. Inverting for emissions of carbon monoxide from Asia using aircraft observations over the western Pacific. *J. Geophys. Res*. 108:D21, doi:10.1029/2003JD003397.

Park, R.J., Jacob D.J., Chin, M., and Martin, R.V. 2003. Sources of carbonaceous aerosols over the United States and implications for natural visibility, *J. Geophys. Res.* 108:4355.

Parrish, D.D., Trainer, M., Hereid, D., Williams, E.J., Olszyna, K.J., Harley, R.A., Meagher, J.F., and Fehsenfeld, F.C. 2002. Decadal change in carbon monoxide to nitrogen oxide ratio in U.S. vehicular emissions. *J. Geophys. Res.* 107:4140.

Peters, L.K., Berkowitz, C.M., Carmichael, G.R., Easter, R.C., Fairweather, G., Ghan, S.J., Hales, J.M., Leung, L.R., Pennell, W.R., Potra, F.A., Saylor, R.D., and Tsang, T.T. 1995. The current state and future direction of Eulerian models in simulating the tropospheric chemistry and transport of trace species: A review. *Atmos. Environ*. 29:189-222.

Peters J.M., Avol E., Navidi W., London S.J., Gauderman W.J., Lurmann F., Linn W., Margolis H., Rappaport E. Gong Jr. H., and Thomas D.C. 1999. A study of twelve Southern California communities with differing levels and types of air pollution. *Am. J. Respir. Crit. Care Med.*, 159:760-767.

Petron, G., Granier, C., Khattatov, B., Lamarque, J.-F., Yudin, V., Muller, J.-F., and Gille, J. 2002. Inverse modeling of carbon monoxide surface emissions using CMDI network observations, *J. Geophys. Res.* 107:4761.

Pielke, R.A. 1984. Mesoscale Meteorological Modeling. Orlando: Academic Press, 612 pp.

Pielke, R.A. and Uliasz, M. 1998. Use of meteorological models as input to regional and mesoscale air quality models – limitations and strengths. *Atmos. Environ*. 32:1455-1466.

Pierce, T., Geron, C., Bender, L., Dennis, R., Tonnesen, G., and Guenther, A. 1998. Influence of increased isoprene emissions on regional ozone modeling. *J. Geophys. Res*. 103:25,611-25,629.

Pleim, J.E., and Xiu, A. 1995. Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. *J. Appl. Meteor*. 34:16-32.

Polina, B.M., Offenberg, J.H., Clemente, J., Blaustein, M., Thurston, G.D., and Chen, L.C. 2004. Ambient pollutant concentrations measured by a mobile laboratory in South Bronx, NY. *Atmos. Environ.* 38:5283-5294.

Pudykiewicz, J.A. 1998. Application of adjoint tracer transport equations for evaluating source parameters. *Atmos. Environ*. 32:3039-3050.

Pun, B.K., and Seigneur, C. 1999. Understanding particulate matter formation in the California San Joaquin Valley: Conceptual model and data needs. *Atmos. Environ*. 33:4865-4875.

Pun, B.K., Seigneur, C., Vijayaraghavan, K., Wu, S.-Y., Chen, S.-Y., Knipping, E.M., and Kumar, N. 2006. Modeling regional haze in the BRAVO study using CMAQ-MADRID: 1. Model evaluation. *J. Geophys. Res*. 111, D06302, doi:10.1029/2004JD005608.

Reid, N., Misra, P.K., Amman, M., and

Hales, J. 2003. Air quality modelling. Paper presented at Third NERAM International Colloquium on Health and Air Quality. 5-7 Nov. 2003, Santo Spirito Hospital, Rome, Italy. www.irr-neram.ca/rome/rome.html#proceed.

Russell, A., and Dennis, R. 2000. NARSTO critical review of photochemical models and modeling. *Atmos. Environ*. 34:2283-2324.

Sahsuvaroglu, T., Arain, A., Kanaroglou, P., Finkelstein, N., Newbold, B., Jerrett, M., Beckerman, B., Brook, J., Finkelstein, M., and Gilbert, N.L. 2006. A land-use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Canada. *J. Air Waste Manage. Assoc.* 56:1059-1069.

Samet J. 2006. Impact of Air Pollution on Public Health: Transportability of Risk Estimates. www.irr-neram.ca/Vancouver/Samet1.ppt

Samoli E., Aga E., Toulouni G., Nisictes K., Forsberg, B., Lefranc A., Pekhanen J., Wojtynich B., Schindler C., Niciu E., Brunstein R., Dodic Fihfch M., Schwartz J., and Katsoujanni K. 2006. Short term effects of nitrogen dioxide and mortality: an analysis within the APHEA project. *Eur. Respir. J.* 27:1129-1137.

Sarnat, J.A., Schwartz, J., Catalano, P.J., and Suh, H.H. 2001. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? *Environ. Health Persp*. 10:1053-1061.

Schoeberl, M.R., Douglass, A.R., Hilsenrath, E., Bhartia, P.K., Barnett, J., Gille, J., Beer, R., Gunson, M., Waters, J., Levelt, P.F., and DeCola, P. 2004. Earth observing system missions benefit atmospheric research. *EOS* 85:177-178.

Seaman, N.L. 2000. Meteorological modeling for air-quality assessments. *Atmos. Environ*. 34:2231-2259.

Seigneur, C. 2001. Current status of air quality models for particulate matter. *J. Air Waste Manage. Assoc*. 51:1508-1521.

Seigneur, C., Pun, B., Pai, P., Louis, J.-F., Solomon, P., Emery, C., Morris, R., Zahniser, M., Worsnop, D., Koutrakis, P., White, W., and Tombach, I. 2000. Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility. *J. Air Waste Manage. Assoc*. 50:588-599.

Seigneur, C., and Moran, M.D. 2004. Using models to estimate particle concentration. Chapter 8 in Particulate Matter Science for Policy Makers: A NARSTO Assessment, P. McMurry, M. Shepherd, and J. Vickery (eds), Cambridge: Cambridge University Press, 42 pp.

Seilkop, S.K. 1995a. Representativeness of surface site air concentrations relative to an 80 km grid. In Regional Photochemical Measurement and Modeling Studies. Vol. 1. Results and Interpretation of Field Measurements, A.J. Ranzieri and P.A. Solomon (eds), VIP-48, Pittsburgh, PA: Air & Waste Management Assoc., pp. 197-210.

Seilkop, S.K. 1995b. Characterization and usage of subgrid scale variability in model evaluation. Report prepared by Analytical Services, Inc., Durham, North Carolina, for U.S. Environmental Protection Agency, Research Triangle Park, NC, November, 57 pp.

Seinfeld, J.H., and Pandis, S.N. 1998. Atmospheric Chemistry and Physics – From Air Pollution to Climate Change. New York, NY: John Wiley & Sons, 1326 pp.

Setton, E., Hystad, P., Keller, C.P., Cloutier-Fisher, D., Foster, L., Copes, R., and Brauer, M. 2006. Simulating risk of exposure to traffic-related air pollution in urban and suburban working and non-working populations. Presented at the Air Quality 2006, Banff, AB, Feb. 2006. www.ec.gc.ca/aqw-aqa/program_e.html.

Sillman, S., He, D., Pippin, M., Daum, P., Kleinman, L., Lee, J.H., and Weinstein-Lloyd, J. 1998. Model correlations for ozone, reactive nitrogen and peroxides for Nashville in comparison with measurements: implications for VOC-NO_x sensitivity. *J. Geophys. Res*. 103:22629-22644.

Silva, C., Gould, T., Larson, T., Henderson, S., Allen, R., Cohen, M., and Buzzelli, M.

2006. A tale of two cities: Comparing land use & traffic related pollution exposure models in Seattle and Vancouver. Presented at the Air Quality 2006, Banff, AB, Feb. 2006. www.ec.gc.ca/aqw-aqa/program_e.html.

Smyth, S.C., Jiang, W., Yin, D., Roth, H., and Giroux, E. 2006a. Evaluation of CMAQ O_3 and $PM_{2.5}$ performance using Pacific 2001 measurement data. *Atmos. Environ*. 40:2735-2749.

Smyth, S., Yin, D., Roth, H., Jiang, W., Moran, M.D., and Crevier, L.-P. 2006b. The impact of GEM and MM5 meteorology on CMAQ air quality modeling results in eastern Canada and the northeastern United States. *J. Appl. Meteorol*. 45:1525-1541.

Spengler, J.D., Koutrakis, P., Dockery, D.W., Raizenne, M., and Speizer, F.E. 1996. Health effects of acid aerosols on North American children: Air pollution exposures. *Environ. Health Persp.* 104:492-499.

Stein, A.F., and Lamb, D. 2002. Chemical indicators of sulfate sensitivity to nitrogen oxides and volatile organic compounds. *J. Geophys. Res*., 107-D20, 4449, doi:10.1029/2001JD001088.

Stieb, D.M., Paola, J., and Neuman, K. 1996. Do smog advisories work? Results of an evaluation of the Canadian Smog Advisory Program. *Can. J. Public Health* 87:166-169.

Stieb, D.M., Judek S., and Burnett R.T. 2003. Meta-analysis of time-series studies of air pollution and mortality: update in relation to the use of generalized additive models. *J. Air Waste Manage. Assoc*. 53:258-261.

Streets, D.G., Yu, C., Bergin, M.H., Wang, X., and Carmichael, G.R. 2006. Modeling study of air pollution due to the manufacture of export goods in China's Pearl River Delta. *Environ. Sci. Technol*. 40:2099-2107.

Uliasz, M. 1993. The Atmospheric Mesoscale Dispersion Modeling System. *J. Appl. Meteor*. 32:139-149.

US EPA 2001. Guidance for Demonstrating Attainment of Air Quality Goals for $PM_{2.5}$ and Regional Haze. U.S. Environmental Protection Agency Report Draft 2.1, Jan., 273 pp. www.epa.gov/scram001/guidance/guide/draft_pm.pdf.

US EPA 2004. Air Quality Criteria Document for Particulate Matter. Volume II. EPA/600/P-99/002bF. Office of Research and Development. USEPA. Research Triangle Park, NC. http://cfpub2.epa.gov/ncea/cfm/recordisplay.cfm?deid=87903.

US EPA 2005a. Technical Support Document for the Final Clean Air Interstate Rule Air Quality Modeling. U.S. Environmental Protection Agency, March, 285 pp. www.epa.gov/cair/pdfs/finaltech02.pdf.

US EPA 2005b. Guidance On The Use Of Models And Other Analyses In Attainment Demonstrations For The 8-Hour Ozone NAAQS. Report No. EPA-454/R-05-002, U.S. Environmental Protection Agency, October, 128 pp. www.epa.gov/scram001/guidance/guide/8-hour-o3-guidance-final-version.pdf.

Venkatram, A. 1988. Inherent uncertainty in air quality modeling. *Atmos. Environ*. 22:1221-1227.

Watson, J.G. and Chow, J.C. 2005. Chapter 16B Receptor Models, Air Quality Modeling: Theories, Methodologies, Computational Techniques, and Available Databases and Software, Vol. II, ed. P. Zannetti. EnviroComp Institute and Air & Waste Management Assoc. pp. 455-501.

Watson, J.G., Chow, J.C., and Fujita, E.M. 2001. Review of volatile organic compound source apportionment by chemical mass balance. *Atmos. Environ.* 35:1567-1584.

West, J.J., Ansari, A.S., and Pandis, S.N. 1999. Marginal $PM_{2.5}$: nonlinear aerosol mass response to sulfate reductions in the Eastern United States. *J. Air Waste Manage. Assoc*. 49:1415-1424.

Westerdahl, D., Fruin, S., Sax, T., Fine, P.M., and Sioutas C. 2005. Mobile platform measurements of ultrafine particles and associated pollutant concentrations on freeways and residential streets in Los

Angeles. *Atmos. Environ*. 39:3597–3610.

Wexler A.S., and Johnson M.V. 2006. What have we learned from highly time resolved measurements during the EPA supersite program and related studies? *J. Air Waste Manage. Assoc*. (accepted).

Wheeler, A.J., Smith-Doiron, M., Xu, X., Gilbert, N.L., and Brook, J.R. 2006a. Intra-urban variability of air pollution in Windsor, Ontario - Measurement and modeling for human exposure assessment. *Environ. Res.* (submitted).

Wheeler, A.J., Xu, X., Smith-Doiron, M., Gilbert, N.L., and Brook J.R. 2006b. Intra-urban variability of air pollution - monitoring and modelling of $PM_{2.5}$, $PM_{2.5-10}$ and NO_2. Poster presentation at: ISEE ISEA 2006, Paris, Sept.

WHO 1999. Monitoring ambient air quality for health impact assessment. WHO Regional Publications, European Series, No. 85, ISBN 92 890 1351 6 / ISSN 0378-2255.

WHO 2004. Health aspects of air pollution: Results from the WHO project Systematic Review of Health Aspects of Air Pollution in Europe. Publications WHO Regional Office for Europe Scherfigsvej 8 DK-2100 Copenhagen, Denmark.

Winer, A.M., Lurmann, F.W., Coyner, L.A., Colome, and Poe, M.P. 1989. Characterization of air pollutant in the California South Coast Air Basin: application new regional human exposure (REHEX) model. Report to the South Coast Air Quality Management District and the California State University Foundation, Contract Number TSA 106-01-88.

Wu, J., Lurmann, F., Winer, A., Lu, R., Turco, R., and Funk, T. 2005. Development of an individual exposure model for application to the Southern California children's health study. *Atmos. Environ*. 39:259-273.

Xu, X., Brook, J.R., and Guo, P. 2006. A statistical assessment of mobile sampling strategy. *J. Air Waste Manage. Assoc*. (Submitted).

Yli-Tuomi, T., Aarnio, P., Pirjola, L., Mäkelä, T., Hillamo, R., and Jantunen, M. 2005. Emissions of fine particles, NO_x, and CO from on-road vehicles in Finland. *Atmos. Environ*. 39:6696–6667.

Yu, S., Dennis, R., Roselle, S., Nenes, A., Walker, J., Eder, B., Schere, K., Swall, J., and Robarge, W. 2005. An assessment of the ability of three-dimensional air quality models with current thermodynamic equilibrium models to predict aerosol NO_3^-. *J. Geophys. Res*. 110, D07S13, doi:10.1029/2004JD004718.

Yu, S., Dennis, R., Bhave, P.V., and Eder, B.K. 2004. Primary and secondary organic aerosols over the United States: estimates on the basis of observed organic carbon (OC) and elemental carbon (EC), and air quality modeled primary OC/EC ratios. *Atmos. Environ*. 38:5257-5268.

Zhang, Y., Seigneur, C., Seinfeld, J.H., Jacobson, M.Z., Clegg, S.L., and Binkowski, F.S. 2000. A comparative review of inorganic aerosol thermodynamic equilibrium modules: Similarities, differences, and their likely causes. *Atmos. Environ*. 34:117-137.

Zhang, L., Moran, M.D., and Brook, J.R. 2001. A comparison of models to estimate in-canopy photosynthetically active radiation and their influence on canopy stomatal resistance. *Atmos. Environ*. 35:4463-4470.

Zhang, L., Moran, M.D., Makar, P.A., Brook, J.R., and Gong, S. 2002. Modelling gaseous dry deposition in AURAMS - A Unified Regional Air-quality Modelling System. *Atmos. Environ*. 36:537-560.

Zhang, Y., Vijayaraghavan, K., and Seigneur, C. 2005. Evaluation of three probing techniques in a three-dimensional air quality model. *J. Geophys. Res*. 110, D02305, doi:10.1029/2004JD005248.

Zunckel, M., Koosailee, A., Yarwood, G., Maure, G., Venjonoka, K., van Tienhoven, A.M., and Otter, L. 2006. Modelled surface ozone over southern Africa during the Cross Border Air Pollution Impact Assessment Project. *Environ. Model. Software* 21:911-924.