CRC Report No. A-61

FEASIBILITY OF USING SATELLITE DATA IN AIR QUALITY MODELING

FINAL REPORT

JUNE 2007

COORDINATING RESEARCH COUNCIL, INC.
3650 MANSELL ROAD·SUITE 140·ALPHARETTA, GA 30022

LEGAL NOTICE

FEASIBILITY OF USING SATELLITE DATA IN AIR QUALITY MODELING

Krish Vijayaraghavan, Hilary E. Snell and Christian Seigneur

Atmospheric & Environmental Research, Inc.
2682 Bishop Drive, Suite 120
San Ramon, CA 94583
and
131 Hartwell Ave
Lexington, MA 02421

Final Report

CRC Project A-61

Prepared for

Coordinating Research Council
3650 Mansell Road, Suite 140
Alpharetta, GA 30022

Document CP221-07-01

June 2007

TABLE OF CONTENTS

LIST` OF TABLES

LIST OF FIGURES

Executive Summary

The results of computer simulations of air pollution are usually evaluated by comparison with measurements near the ground and, occasionally, from aircraft and balloons. Over the past decade, several satellites have been launched that measure atmospheric constituents over the Earth using remote sensing devices. This report investigates the feasibility of using such data to evaluate air quality modeling results and to improve the quality of future simulations.

The key findings of this study are as follows:

- Validated tropospheric satellite data are commonly available only for the following species or physical quantity: ozone (O_3), carbon monoxide (CO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), formaldehyde (HCHO) and aerosol optical depth (AOD). So it is currently feasible to evaluate tropospheric air quality models typically for these species only. Satellite data for other species such as bromine oxide (BrO), nitric acid (HNO_3) and glyoxal (OCHCHO) are becoming available but are sometimes limited by the time period of availability.

- Satellite data provide the spatial coverage needed for the boundary and initial conditions (of O_3 and CO, for example) of regional air quality models, particularly, aloft and over the oceans and other areas where other data may not be available.

- Satellite data offer significant potential for inverse modeling and data assimilation in air quality models for species such as NO_2, HCHO and O_3 to estimate emissions of the precursors of these species or use those data to improve the performance of air quality simulations for these species.

- Satellite data are also useful to improve concentration maps of air pollutants such as O_3, NO_2 and particulate matter (PM) in a combined post-processing step of data fusion between model simulation results and measurements.

We elaborate on our findings below and also highlight the limitations of satellite measurements.

The primary organizations involved in the launch and operation of satellites used in the remote sensing of air pollution are the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), the Canadian Space Agency, the European Space Agency, and the European Organization for the Exploitation of Meteorological Satellites. Most of the satellites launched by these agencies orbit near the Earth's surface at an altitude of about 800 km and hence offer data with high vertical and horizontal resolutions. The sensors aboard these satellites detect the scattering or emission of radiation from the constituents of the Earth's atmosphere such as O_3 and PM. Geophysical quantities of interest such as O_3 concentrations are then extracted from the measured radiances through a process known as the "retrieval". The retrieval team uses a priori information from aircraft data or global modeling to constrain the retrieval solution. The retrieved products are validated and finally released to the public.

The primary location for data from NASA sensors and satellites is the NASA Earth Observing System Data Gateway internet site. This site contains links to relevant validated data sets and also lists other locations for obtaining data such as those for European sensors. Satellite data are subject to limitations such as limited temporal, horizontal and vertical resolutions, and the effect of clouds and ground albedo (reflection). These uncertainties are typically minimized or resolved by the retrieval team before releasing the data.

Following the download of the "retrieved data", air quality modelers need to process the data to match the model temporal/spatial sampling to that of the satellite measurement. They also need to match the vertical resolutions of the model and satellite data. The proper way to treat these differences in vertical resolution is through the use of the sensor 'averaging kernel' which represents the way in which the vertical structure of the atmospheric profile is mapped into the radiances measured by the sensor. The model output must be "degraded" in vertical resolution before comparison with the vertical profile retrieved from the satellite measurements using an equation that is a function of the averaging kernel and the a priori profile. Under certain circumstances, such as in the vicinity of large biomass fires or large urban areas, the retrieved profile may be unduly influenced by the a priori profile and the model vertical profile may then be a better approximation to the true profile than the profile retrieved from the satellite data.

After due processing, the satellite data may be used to evaluate, initialize, constrain and/or improve the performance of air quality models. Satellite data provide two important sources of information compared to surface and aircraft monitoring data: more complete spatial coverage and a vertically-integrated measure of air quality. A large number of chemical species may be planned for retrieval during a satellite mission but many of these are either not retrieved due to instrument/algorithm issues or are not validated and quickly made available for public dissemination. Validated tropospheric satellite data are commonly available for O_3, CO, NO_2, SO_2, HCHO and AOD. So it is currently feasible to evaluate tropospheric air quality models typically for these species only. Data for halogens such as bromine oxide may become more widely used as halogen chemistry starts to be taken into account in air quality models (for example, in the case of atmospheric mercury deposition). Satellite data for other species such as nitric acid and glyoxal are becoming available but tropospheric measurements are currently available only for limited time periods.

Satellite data are useful for specifying boundary and initial conditions for regional air quality models due to their large spatial coverage particularly, aloft and over the oceans and other areas where data are limited. Satellite measurements, when used as boundary conditions, can be used to account for the contributions of pollutants such as O_3 and CO transported over long distances, for example, from Asia over the Pacific Ocean to the United States.

Satellite data can also be used for inverse modeling and data assimilation in air quality models. Emissions of nitrogen oxides (NO_x) and isoprene have been estimated from the column densities of NO_2 and HCHO, respectively, by assuming linear chemistry relationships between the emissions and the satellite column data. Satellite measurements of CO and NO_2 have also been used to estimate the contributory emissions by taking into account both transport and chemistry using a variational approach. Assimilation of satellite data for O_3, NO_2 and AOD (as a surrogate for PM) has been shown to directly improve the performance of air quality simulations for these pollutants. There are some specific areas of data assimilation where satellite data provide information that is not directly available from other sources. For example, AOD measurements can provide valuable information on the magnitude and extent of biomass fires, and SO_2 measurements can help characterize volcanic eruption plumes. The use of satellite data, along with surface ambient air quality measurements, to improve air quality forecasting is being planned in the United States and Europe. Satellite data may also be used to improve air concentration maps of pollutants such as O_3, NO_2 and PM in a post-processing step of data fusion between modeling results and measurements.

There may be several sources of error in the satellite data such as instrument issues, the choice of a priori constraints, etc. Due to these errors and the other limitations described above, satellite data should be used as a quantitative bench mark in the performance evaluation of air quality models only after the satellite retrievals have been independently validated against other data such as those from aircraft, sondes and ground-based measurements. Such validation is routinely performed by the satellite data retrieval teams, usually using aircraft/sonde data, before releasing the measurements to the public.

1. Introduction

The evaluation of regional air quality models is typically conducted using ground-level measurements. There are only a few instances where data from aircraft, helicopters, balloons or towers have been used to evaluate the third dimension of simulated air quality. Satellite data provide two important sources of information compared to surface monitoring data: more complete spatial coverage and a vertically-integrated measure of air quality (e.g., Engel-Cox et al., 2004; Edwards et al., 2006). Some applications to date have focused on identifying specific events such as forest fires or desert dust plumes (e.g., Spichtinger et al., 2001; Falke et al., 2001), characterizing the long-range transport of some pollutants in combination with global-scale modeling (e.g., Heald et al., 2006), augmenting the spatial coverage of surface monitoring data (e.g., Wang and Christopher, 2003; Liu et al., 2005a) and evaluating regional air quality model simulations (e.g., Hodzic et al., 2006; Vijayaraghavan et al., 2006; Kondragunta et al., 2006; Byun et al., 2006). There is clearly an enormous potential for the use of satellite data to improve our capabilities in air quality modeling and it is, therefore, of particular interest to investigate how satellite measurements can be used to improve our characterization of the atmosphere and evaluate air quality models in a more comprehensive manner.

We investigate here the feasibility of using satellite data to evaluate air quality models and improve their performance via data assimilation. Section 2 presents an overview of satellite remote sensing data relevant to air quality. In Section 3, we explain the acquisition and processing of satellite data for air quality applications. Section 4 discusses the major air quality applications for satellite data.

2. Satellite Remote Sensing Data Relevant to Air Quality

In this section, we present some basic ideas on remote sensing followed by an overview of satellites and sensors. Satellites and sensors currently used for the remote sensing of tropospheric air quality are listed. We then discuss the chemical species that are measured by these sensors. Finally, we discuss the limitations and uncertainties in current satellite retrievals.

2.1 Essentials of remote sensing

Remote sensing is defined as the science by which the characteristics of an object of interest can be identified, measured or analyzed without direct contact (Japan Assoc. Remote Sensing, 1996). Electro-magnetic radiation that is reflected or emitted from an object is the usual source of remote sensing data. In our study, the object of interest is the Earth's atmosphere and its constituents. Each atmospheric constituent such as ozone, water vapor, etc. has its own unique spectral characteristic of emission and absorption. A device to detect the electro-magnetic radiation reflected or emitted from an object is called a "remote sensor" or "sensor". These sensors are carried aboard platforms such as aircraft and satellites. The process of extracting the geophysical information of interest from the measured radiances is known as the "retrieval".

2.2 Overview of satellites

Based on their orbital distance, satellites launched by humans can be classified into Low Earth Orbit (LEO), Mid-Earth Orbit (MEO), Geostationary Orbit (GEO), and Lagrangian Point (L-1) satellites. LEO satellites have elliptical or circular orbits typically at a height of less than 1,000 km above the surface of the earth. Polar-orbiting satellites are a subset of LEO satellites that can be used to view only the poles or to view the same place on earth at the same time each 24-hr day; these are usually at an altitude of 700-800 km. A sun-synchronous orbit is a special case of a polar orbit that crosses the equator at the same time each orbit. MEO satellites have circular orbits at an altitude of 1,000-10,000 km with an orbital period of 6 hours or less. GEO satellites have circular orbits oriented in the plane of the earth's equator and at an altitude of 35,800 km. Because its orbital period of 24 hours is equal to that of the rotational period of the Earth, a GEO satellite will appear fixed above the surface of the earth, i.e., at a fixed latitude and longitude. L-1 satellites orbit around the Sun such that they are continuously between the Sun and the Earth, at a distance of 1.5 million km from the Earth.

Of these satellite types, the ones commonly used for the remote sensing of air quality are LEO polar-orbiting sun-synchronous satellites and to a lesser extent, GEO satellites. MEO and L-1 satellites are not used for remote sensing currently but have been proposed for future missions. A MEO satellite constellation has been proposed to the National Oceanic and Atmospheric Administration (NOAA) for consideration for a post-2012 mission. An L-1 mission called "Janus", which will provide the first comprehensive and continuous observation of the Earth's whole dayside atmosphere, the solar wind, and the Sun, has been proposed to the National Aeronautics and Space Administration (NASA).

Polar-orbiting and GEO satellites provide complementary information about the state of the atmosphere. The GEO satellites are able to provide measurements with high temporal resolution (by remaining fixed in the equatorial plane over a given point) but at (typically) reduced horizontal and vertical resolution compared to polar-orbiting satellites. Polar-orbiting satellites have a poor temporal resolution (at best, a 12-hour measurement repeat cycle for a given geographic location) but tend to offer higher vertical and horizontal resolution than the GEO satellites due to greater proximity to the Earth's surface. However, polar-orbiting satellites can have a fixed equator crossing time and thus will measure the same geographic location at the same time each day. (This can be a benefit or detriment depending upon the desired use of the data.)

Satellites can also be classified into operational and research satellites based on their mission and available products. Operational satellites typically have a clearly defined mission with specific, well-defined products, while research satellites tend to have more research-grade products that are state-of-the-art in terms of sensor and data retrieval technology. Operational satellites are intended to provide stable, long-term inputs into numerical weather prediction models, but have the ability to obtain some information relevant to air quality models. In contrast, research satellites tend to be designed for very specific purposes, such as measurement of aerosols or ozone. The primary difference with the operational satellites lies in the availability of data products, both in terms of the ability to obtain the data and the timeliness of the data distribution after the measurement. The "operational products" are generally available and have well-documented error characteristics, while "research grade" products tend to contain less documentation about errors and may only be available directly from the scientists working on the product.

The primary agencies/organizations involved in the launch and/or operation of satellites and sensors used in the remote sensing of air quality are NASA, NOAA, the European Space Agency (ESA), the Canadian Space Agency (CSA), and the Japan Aerospace Exploration Agency (JAXA). One of the first remote sensing satellites launched was the Earth Resources Technology Satellite or Landsat, launched by NASA in 1972, which provided multi-spectral data related to crops, minerals, soils, urban growth, and many other Earth features. Table 1 lists the key satellites currently used for the remote sensing of the chemical constituents of the Earth's atmosphere.

Table 1. Satellites currently used for the remote sensing of air quality.[a,b]

Satellite	Organization	Launch Date	Orbit
ERS-2	ESA	Apr 21, 1995	Sun-synchronous
TOMS-EP	NASA	Jul 2, 1996	Sun-synchronous
Terra	NASA	Dec 18, 1999	Sun-synchronous
ODIN	SSC/CSA/CNES/TEKES	Feb 23, 2001	Sun-synchronous
GOES-M	NOAA/NASA	Jul 23, 2001	Geo-synchronous
ENVISAT	ESA	Mar 1, 2002	Sun-synchronous
Aqua	NASA	May 8, 2002	Sun-synchronous
ACE/SCISAT	CSA/NASA	Aug 12, 2003	Sun-synchronous
Aura	NASA	Jul 15, 2004	Sun-synchronous
NOAA-N POES	NOAA/NASA	May 20, 2005	Sun-synchronous
CALIPSO	NASA	Apr 28, 2006	Sun-synchronous
CloudSat	NASA	Apr 28, 2006	Sun-synchronous
GOES-N	NOAA/NASA	May 24, 2006	Geo-synchronous
MetOp-A	ESA	Oct 19, 2006	Sun-synchronous

[a] This list includes satellites with widely disseminated data and is not comprehensive.
[b] See Appendix A for a list of acronyms used here.

A list of the acronyms used in Table 1 and elsewhere in this document is presented in Appendix A.

NASA's air quality remote sensing satellites include Aqua, Aura, CALIPSO, and CloudSat (which are all part of the "A-Train", a series of six satellites that fly in close proximity to one another thus allowing for coordinated measurements), and Terra and TOMS-EP. All of these satellites have LEO sun-synchronous orbits.

NOAA's environmental satellite system is composed of two types of satellites: Geostationary Operational Environmental Satellites (GOES) for national, regional, short-range warning and "now-casting"; and the Polar-orbiting Operational Environmental Satellites (POES) for global forecasting and environmental monitoring. The current operational version in the GOES series is GOES-M (known as GOES-12 once on-orbit) while that in the POES series is NOAA-N (known as NOAA-18 once on-orbit). Note that GOES-N was launched recently (May 24, 2006). The POES spacecraft serve as complementary satellites to the GOES system. Where the GOES satellites provide near-term data from the continental United States and Hawaii, the POES spacecraft provide full global data for short- and long-range forecast models, climate modeling, and various other secondary missions. NASA is responsible for the launch and testing of the spacecraft, instruments and unique ground equipment and turns operational control of the spacecraft over to NOAA after 21 days of comprehensive subsystem checkout. Most of the visible satellite images seen currently on television weather forecasts use data from the GOES satellites.

The Atmospheric Chemistry Experiment (ACE) satellite, also known as SCISAT-1, was primarily funded by the Canadian Space Agency (CSA) and launched by NASA. CSA also manages a sensor on the Odin satellite in collaboration with the Swedish Space Corporation (SSC), the Finnish National Technological Agency (TEKES), and the French National Space Study Center (CNES). The European Space Agency (ESA) has launched three satellites that are currently used for the remote sensing of air quality, the Environmental Satellite (ENVISAT), the European Remote Sensing Satellite (ERS-2, the second in the ERS series), and the Meteorological Operational Satellite Programme (MetOp-A). ERS-2 continues to provide good data far beyond its nominal lifetime of 10 years. JAXA operates weather monitoring satellite missions jointly with NASA and also manages some sensors in satellites such as Aqua.

Most current air quality remote sensing satellites are research satellites. The POES series of satellites are the exception; these are operational polar-orbiting satellites. These will be replaced in the future with the National Polar-orbiting Operational Environmental Satellite System (NPOESS) series of satellites. The NPOESS sensors provide enhanced spectral and spatial coverage over the existing systems and will be of increased utility to the air quality community.

2.3 Overview of sensors

Remote sensors can be categorized into passive and active sensors. Passive sensors take advantage of the interaction between naturally occurring radiation (such as sunlight or infrared radiation emitted by objects) and atmospheric matter. Most current satellites use passive sensors. Active remote sensors work by emitting a signal and then processing the backscattering (return) of the emitted signal. Satellite-based active remote sensing systems include radars and lidars that emit radio waves and laser beams, respectively, in the direction of the object to be sampled, and then utilize a parabolic dish to collect the backscattered radiation. For example, the CALIPSO satellite will use lidar to create high-resolution vertical profiles of clouds and aerosols.

The electro-magnetic radiation regions used in remote sensing (Japan Assoc. Remote Sensing, 1996) are near UV (ultra-violet) (0.3-0.4 μm), visible light (0.4-0.7 μm), near shortwave and thermal infrared (0.7-14 μm) and microwave (1 mm - 1 m). Thus, remote sensors can be classified into three types with respect to the wavelength regions; (1) Visible and Reflective Infrared Remote Sensors, (2) Thermal Infrared Remote Sensors and (3) Microwave Remote Sensors. The energy source used in passive visible and reflective infrared remote sensing is the sun. The source of radiant energy used in thermal infrared remote sensing is the object itself. The energy source for microwave remote sensing could be the object itself or another source depending on whether the sensing is passive or active, respectively; lasers are the active source in other regions of the spectrum.

Table 2 lists the sensors on the satellites shown earlier in Table 1. The acronyms used for the sensors are expanded in Appendix A.

Table 2. Sensors currently used for the remote sensing of air quality.[a]

Satellite	Sensors used for remote sensing air quality
Aqua	AIRS, MODIS
Aura	HIRDLS, MLS, OMI, TES
CALIPSO	Lidar
CloudSat	CPR
Terra	MISR, MODIS, MOPITT
TOMS-EP	TOMS
GOES-M, GOES-N	Weather monitoring instruments
NOAA-N POES	SBUV/2
ACE/SCISAT	ACE-FTS, MAESTRO
ENVISAT	MIPAS, SCIAMACHY
ERS-2	GOME
MetOp-A	IASI, ASCAT, GOME-2
ODIN	OSIRIS, SMR

[a] This list is not comprehensive

2.4 Air quality related data

We describe below the air quality products determined from satellite measurements and retrievals along with characteristics such as detection limit, spatial and temporal characteristics. We also describe products that may be derived from combinations of measurement parameters. We focus our discussion on tropospheric air quality products; we list only those stratospheric products that are directly relevant to the troposphere or are used to compute tropospheric columns (e.g., MLS measures a stratospheric OH concentration, but it is not listed in the table).

The emphasis is on standard data products that are readily available from sources such as NASA, but the discussion also includes specialized data products developed for various sensors by the science team investigators.

Table 3 lists the species available from NASA's Earth Observing System (EOS) platform, which is currently among the most widely available in terms of data and documentation. The EOS includes the following satellites: Aqua, Aura, CALIPSO, CloudSat, and Terra (and several others not measuring air quality).

Table 3. Air quality related species and sensors from the NASA-EOS satellites.

Product	**Sensor(s)**
Aerosol properties	HIRDLS, Lidar, MISR, MODIS, OMI
Bromine oxide (BrO)	OMI
Carbon monoxide (CO)	AIRS, MLS, MOPITT, TES
Chlorine dioxide (OClO)	OMI
Chlorine bitrate ($ClONO_2$)	HIRDLS
Cloud properties	AIRS, CPR, HIRDLS, Lidar, MLS, MODIS, OMI
Dinitrogen pentoxide (N_2O_5)	HIRDLS
Formaldehyde (HCHO)	OMI
Glyoxal (OCHCHO)	OMI
Methane (CH_4)	HIRDLS, MOPITT, TES
Nitric acid (HNO_3)	MLS
Nitrogen dioxide (NO_2)	HIRDLS, OMI
Nitrous oxide (N_2O)	HIRDLS, MLS
Ozone (O_3)	AIRS, HIRDLS, MLS, MODIS, OMI, TES
Sulfur dioxide (SO_2)	AIRS, MLS, OMI

Table 4 lists the characteristics of sensors providing air quality products on the NASA-EOS platforms. Ozone has been the most commonly retrieved air quality product.

In addition to these sensors, the other NASA/NOAA sensors that measure air quality include the TOMS and SBUV/2 instruments aboard the TOMS-EP and NOAA-N POES satellites, respectively. The TOMS sensor provides global measurements of total column ozone. It also measures SO_2 released in volcanic eruptions. Ozone is measured by observing both incoming solar energy (solar irradiance) and backscattered ultraviolet (UV) radiation at six wavelengths. "Backscattered" radiation is solar radiation that has penetrated to the Earth's atmosphere and is then scattered by air molecules, clouds and the surface back through the atmosphere to the satellite sensor. Along that path, a fraction of the UV is absorbed by ozone. By comparing the amount of backscattered radiation to observations of incoming solar energy at identical wavelengths, one can calculate the Earth's albedo, the ratio of light reflected by Earth compared to what it receives. Changes in albedo at the selected wavelengths are used to derive the amount of ozone above the surface (NASA Science Missions web-site, 2006). The SBUV/2 instrument measures solar irradiance and backscattered solar energy in the near ultraviolet spectrum (160 to 400 nm). These are used to retrieve the global ozone concentration in the stratosphere to an absolute accuracy of 1 percent and the vertical distribution of atmospheric ozone to an absolute accuracy of 5 percent. The measurements of total ozone column from TOMS and stratospheric ozone column from SBUV/2 have been frequently used (e.g., Fishman et al., 2003, 2005) to derive the tropospheric ozone residual (TOR).

Table 4. Spatial resolution, coverage and accuracy of sensors providing air quality related products on the NASA-EOS platforms.

Sensor	**Product**	**Spatial Resolution**	**Spatial Coverage**	**Accuracy**	**Notes**
OMI	Total Column Ozone (O_3)	13 x 24 km, total column	Global, Day, once/day	3% absolute, 1.5% relative	
	Ozone (O_3) Profile	20 x 45 km horizontal, 6 km vertical (20 – 45 km)	Global, Day, once/day	10% absolute, 1.5% relative	
	Tropospheric Ozone (O_3)	52 x 48 km, 60N – 60S latitude, total (tropospheric) column	Global, Day, once/day	25% absolute, 10% relative	
	Cloud Scattering Layer Pressure	13 x 24 km	Global, Day, once/day	100mb absolute, 30 mb relative	Limited to optically thick clouds
	Aerosol Optical Thickness	13 x 24 km, total column	Global, Day, once/day	0.1 or 30% absolute, 0.05 or 10% relative	Cloud-free pixels only
	Aerosol Single Scattering Albedo (SSA)	13 x 24 km, total column	Global, Day, once/day	0.1 absolute, 0.05 relative	Cloud-free pixels only
	Sulfur Dioxide (SO_2)	13 x 24 km, total column	Global, Day, once/day	Non-volcanic: 50% absolute, 20% relative; Volcanic: 30% absolute, 20% relative	
	Nitrogen Dioxide (NO_2)	26 x 48 km, total column	Global, Day, once/day	Background: 2E+14 molecules/cm2; Polluted: 30% absolute, 20% relative	

Table 4. Spatial resolution, coverage and accuracy of sensors providing air quality related products on the NASA-EOS platforms (continued).

Sensor	Product	Spatial Resolution	Spatial Coverage	Accuracy	Notes
OMI	Formaldehyde (HCHO)	13 x 24 km, total column	Global, Day, once/day	35% absolute, 25% relative	
	Bromine Oxide (BrO)	13 x 24 km, total column	Global, Day, once/day	25% absolute, 25% relative	
	Chlorine Dioxide (OClO)	26 x 48 km, slant column	Global, Day, once/day	15% absolute, 10% relative	
HIRDLS	Aerosol Extinction Coefficient	500 km horizontal, 1.25 km vertical (10-30 km)	Twice/day	2-10% absolute, 2-10% relative	Units: km^{-1}
	CFC-11 ($CFCl_3$)	500 km horizontal, 1.25 km vertical (7-28 km)	Twice/day	4-8% absolute, 2-10% relative	
	CFC-12 (CF_2Cl_2)	500 km horizontal, 1.25 km vertical (7-30 km)	Twice/day	4-8% absolute, 1-10% relative	
	Methane (CH_4)	500 km horizontal, 1.25 km vertical (10-65 km)	Twice/day	3-5% absolute, 3-10% relative	
	Chlorine Nitrate ($ClONO_2$)	500 km horizontal, 1.25 km vertical (17-40 km)	Twice/day	5-10% absolute, 8-15% relative	
	Cloud Top Height	500 km horizontal, 250 m vertical (7-24 km)	Twice/day	250 m absolute, 125 m relative	

Table 4. Spatial resolution, coverage and accuracy of sensors providing air quality related products on the NASA-EOS platforms (continued).

Sensor	Product	Spatial Resolution	Spatial Coverage	Accuracy	Notes
HIRDLS	Nitric Acid (HNO_3)	500 km horizontal, 1.25 km vertical (10-40 km)	Twice/day	3-5% absolute, 2-10% relative	
	Nitrous Oxide (N_2O)	500 km horizontal, 1.25 km vertical (10-55 km)	Twice/day	3-5% absolute, 3-10% relative	
	Dinitrogen Pentoxide (N_2O_5)	500 km horizontal, 1.25 km vertical (20-45 km)	Twice/day	5-10% absolute, 2-10% relative	
	Nitrogen dioxide (NO_2)	500 km horizontal, 1.25 km vertical (20-60 km)	Twice/day	3-5% absolute, 3-10% relative	
	Ozone (O_3)	500 km horizontal, 1.25 km vertical (10-80 km)	Twice/day	3-5% absolute, 1-10% relative	
MLS	Cirrus Ice Content	200 km horizontal, 3 km vertical (10-20 km)	Monthly global map	TBD	
	Carbon Monoxide (CO)	500 km horizontal, 3 km vertical (8-90 km)	Monthly global map	5-10% absolute, 3-10 ppbv relative	
	Stratospheric Ozone (O_3)	500 km horizontal, 3 km vertical (15 – 90 km)	Daily global map	5-10% absolute, 2-10% relative (15-50km)	
	Tropospheric Ozone (O_3)	500 km horizontal, 3 km vertical (8-15 km)	Monthly global map	5-10% absolute, 2-10 ppbv relative (8-15 km)	

Table 4. Spatial resolution, coverage and accuracy of sensors providing air quality related products on the NASA-EOS platforms (continued).

Sensor	Product	Spatial Resolution	Spatial Coverage	Accuracy	Notes
MLS	Sulfur Dioxide (SO_2)	500 km horizontal, 3 km vertical (10-40 km)	Daily global map	5-10% absolute, 1-2 ppbv relative (10-30 km)	
TES	Ozone (O_3)	Nadir: 5.3 x 8.5 km horizontal, 2-6 km vertical (0-34 km)	Every other day	3% absolute, 3-20 ppbv relative	
	Carbon Monoxide (CO)	Nadir: 5.3 x 8.5 km horizontal, 2-6 km vertical (0-34 km)	Every other day	3% absolute, 10 ppbv relative	
	Methane (CH_4)	Nadir: 5.3 x 8.5 km horizontal, 2-6 km vertical (0-34 km)	Every other day	3% absolute, 14 ppbv relative	
	Carbon Monoxide (CO)	Nadir: 5.3 x 8.5 km horizontal, 2-6 km vertical (0-34 km)	Every other day	3% absolute, 10 ppbv relative	
AIRS	Cloud Mask	40.6 km horizontal	Twice per day	TBD	Cloud/no-cloud
	Total Column Ozone (O_3)	40.6 km horizontal, total column	Global, twice per day	~10% total column	
MODIS	Aerosol Product	10 km horizontal, global over oceans	Daily	0.05	Problems over land
	Cloud Products	1 – 5 km, global	1-2 per day	2-5% absolute, 10% relative	Characteristics vary depending upon product

Table 4. Spatial resolution, coverage and accuracy of sensors providing air quality related products on the NASA-EOS platforms (continued).

Sensor	Product	Spatial Resolution	Spatial Coverage	Accuracy	Notes
MODIS	Aerosol Optical Depth (AOD)	1, 4.6, 36 km, global over ocean	Daily	??	Clear sky only
	Ozone (O_3)				
MISR	Aerosol Optical Thickness (AOT)	17.6 km horizontal	Global coverage once per week	0.05 or 20%	Better over ocean than over land or dusty sites
MOPITT	Total Column Carbon Monoxide (CO)	22 km horizontal resolution		10% precision	Also gridded in 1 degree x 1 degree daily and monthly averages
	Profile Carbon Monoxide (CO)	22 km horizontal resolution, 4 km vertical resolution		10% precision	Also gridded in 1 degree x 1 degree daily and monthly averages
	Total Column Methane (CH_4)	22 km horizontal resolution, 4 km vertical resolution		1% precision	Also gridded in 1 degree x 1 degree daily and monthly averages

Table 5 lists the air quality products measured by the sensors on the Canadian and European satellites listed earlier (ACE/SCISAT, ENVISAT, ERS-2, and ODIN).

Table 5. Available air quality related products from the sensors on Canadian and European satellites.

Sensor	**Air Quality Products***	**Spatial resolution**
ACE-FTS	O_3, NO, NO_2, HNO_3, N_2O, N_2O_5, CO, CH_4, HCl	4 km vertical
MAESTRO	O_3, NO_2	1-2 km vertical
MIPAS	O_3, NO_2, HNO_3, N_2O, CH_4	3 km x 30 km horizontal, 3 km vertical
SCIAMACHY	O_3, NO_2, N_2O, SO_2, CO CH_4, HCHO, BrO, OClO glyoxal (CHOCHO), AOD	32 km x 215 km horizontal, 3 km vertical
GOME	O_3, NO_2, HCHO, AOD	40 km x 40 km to 40 km x 320 km horizontal, 5 km vertical
OSIRIS	O_3, NO_2, BrO, OClO, aerosol properties	~ 2 km vertical
SMR	O_3, NO, N_2O, HNO_3, HO_2, CO, chlorine monoxide (ClO)	~ 2 km vertical

* Not comprehensive

While satellite sensors measure the aerosol optical depth (AOD), several properties of the aerosol (e.g., $PM_{2.5}$ contribution to AOD, mean effective diameter) can be deduced from measurements made at several wavelengths (e.g., Remer et al., 2005).

In addition to the direct comparison of air quality model outputs with measurement data, additional parameters not necessarily computed by the model may be used to improve the understanding of differences between the models and the measurements. For example, an examination of cloud parameters can be used as a quality control mechanism when evaluating daily measurements, since the amount of cloud cover can impact both the measurement and the chemistry. Further, one could use averages of cloud and aerosol information when comparing monthly data to see regions where the model/measurement bias might be impacted by cloud/aerosol. Also, stratospheric information, while not directly relevant to most air quality model calculations, can be used to diagnose problems with tropospheric column comparisons.

2.5 Limitations of satellite data

The main advantage of satellites is the ability to provide global measurements of a particular quantity with moderate spatial and temporal resolution. Satellites in geostationary orbit are able to provide high temporal resolution with reduced spatial coverage, while satellites in polar orbits provide global coverage with less frequent temporal coverage. There are, however, a number of limitations compared to in-situ sensors. These limitations include the following:

(1) Many chemical species of relevance to air quality are present in trace amounts that cannot be measured by satellite sensors and must be inferred from those that can be measured.
(2) Temporal sampling is limited (sensors on sun-synchronous platforms sample a given location at the same time each day).
(3) The horizontal spatial resolution may be coarse compared to the air quality model resolution, particularly in the case of regional/urban scale modeling.
(4) The vertical resolution is usually limited to total atmospheric column (or tropospheric column) or at most very limited information about the vertical distribution of the chemical species.
(5) Cloud cover will limit coverage, making regions with pervasive cloud cover difficult to sample.
(6) The spatial and temporal co-location of data from several sensors is usually required to maximize the number of parameters available for model validation. This can be difficult due to different spatial and temporal measurement scales for the different sensors. The complexity of this process further increases when not all of the sensors are available on the same satellite platform.
(7) There can be day/night and land/ocean differences in the measurement errors depending upon the type of sensor and the methods used to extract the geophysical data from the radiometric measurements.
(8) Ground albedo (reflection) can result in a lack of contrast between the atmosphere and the surface, making it difficult to obtain the atmospheric quantity from the measurement. This also pertains to low clouds, as are often found over the ocean just off the coast.

We elaborate on these uncertainties and how some of them are resolved, using the satellite retrieval of tropospheric columns of NO_2 and HCHO from the GOME sensor as an example below.

1. The largest uncertainties are due to clouds, as they will shield near-surface NO_2 and HCHO from the view of the satellite. The retrieval is sensitive to the presence of clouds, and even small cloud fractions (between 5 to 20%) have a major impact. High quality observations of the cloud properties (at least cloud fraction and cloud top height) are necessary for a quantitative retrieval.

2. The surface albedo directly influences the sensitivity of GOME for boundary layer NO_2 and HCHO (the spectral signature of changes in surface albedo can be confused with changes in the gas concentration). High quality albedo maps in the relevant spectral range are essential.

3. Profiles of NO_2 are characterized by a large range of spatial and temporal variability. In emission areas, the NO_2 concentration will peak at the surface, while, downstream of such areas the pollution plume will peak at higher altitudes. Aspects such as the distribution of emission sources, the stability and height of the boundary layer, wet removal of nitric acid, deep convection and long-range transport by the wind will determine the NO_2 profile, all of which are strongly varying in time and space. This information will not necessarily be captured in the tropospheric column measurement.

4. The NO_2 columns measured by GOME consist of comparable stratospheric and tropospheric contributions. The stratospheric background has to be quantified carefully in order to derive the tropospheric column. Atmospheric dynamics is well known to generate significant variability in stratospheric tracer amounts, consistent with for instance HALOE observations of NO_2. A standard approach applied to GOME is based on the assumption that stratospheric NO_2 is zonally uniform, or at least has only a small longitudinal variation. This simplification can introduce errors larger in magnitude than the small tropospheric NO_2 column amount, making the retrieval of such amounts practically impossible.

5. Aerosols constitute another source of uncertainty. Thick aerosol layers influence the radiation field and the sensitivity of GOME for near-surface NO_2 and HCHO. Under high-aerosol conditions, the measurement may not be representative of the actual column abundance.

One important improvement of SCIAMACHY as compared to GOME is the smaller ground pixel size. In this way, the variability of NO_2 and HCHO can be better resolved, and the fraction of cloud-free pixels will be larger, improving the quality of the retrieval.

Typically, the uncertainties due to clouds and the surface albedo are handled by using coincident information about the cloud and surface properties, such as land-type maps which provide a “typical” albedo for a given location, and by tailoring the geophysical data retrieval algorithms to provide optimal results under specific measurement conditions. The limitations in temporal and spatial resolution can be overcome, in part, by using the results of retrievals of the same species from different sensors (with careful consideration to the error sources and magnitude for each set of data).

3. Acquisition and Processing of Satellite Data for Air Quality Applications

3.1 Acquisition of satellite data

The acquisition of satellite data is a relatively straightforward process once the data have been identified. Identification of relevant data sets, however, depends upon a number of factors relating to the type of product desired.

Satellite data was traditionally processed by a team lead by the Principal Investigator (PI) of the sensor. Often the data was held until the team had a chance to thoroughly evaluate (and publish) results from the data, and one still needed to contact the PI in order to obtain the data. The process has changed somewhat in that data is now often available in a number of different archive locations (such as the web sites of NASA data processing centers). In these cases, the data is usually cataloged in such a way as to be searchable, allowing the user to easily download the desired data. However, products that are relatively new or un-validated are often delayed for posting to these web sites. Further, there can be multiple teams working on different ways to retrieve a particular product, leading to multiple sources of data for non-standard products. In most cases, the data is available only to the science team until the products have been validated. (After this validation, however, the processed data are typically uploaded quickly to NASA websites.)

The data available often comes in a variety of spatial and temporal domains. The initial data sets are, of course, at the native temporal and spatial resolution of the measurement. However, post-processing of the data is often done to reduce the impact of sensor noise or to provide the consistent treatment of spatial/temporal averaging of the data to all users of the data. (Consistency is important because it requires a thorough understanding of the data to know how and when to include data within an average – often some data must be rejected from the average due to factors such as cloud contamination of the pixel.)

The primary location for data from NASA sensors and satellites is the NASA Earth Observing System (EOS) Data Gateway (http://redhook.gsfc.nasa.gov/%7Eimswww/pub/imswelcome/). This site contains links to relevant data sets and also lists other locations for obtaining data (such as those for European sensors). The data are generally in a standard "hdf" format, with various readers available for ease of data processing. Information is also available regarding the processing of the data – this is very important as it often tells about quality control of the data and how any spatial and temporal averaging was performed. The satellite data are available at different "levels" or stages of data processing. Level-1 data are usually raw radiances measured by satellites after some calibration. Level-2 data refer to chemical species and other parameters at the finest space-time resolution; these are processed ("retrieved") from the corresponding level-1 data. Level-3 data are time and space averaged quantities of the level-2 satellite data. The data nomenclature can vary slightly depending upon the organization responsible for the data. The two primary definitions, used by NASA and NOAA, are summarized in Table 6.

Table 6. Satellite data nomenclature.

Data Name	Description
RDR ("Raw Data Record")	Raw data downloaded from the sensor to the ground processing station. Usually in detector "counts" without any calibration or processing applied. Usually no geospatial information, just some sort of reference time tag of when the data was collected (usually seconds from a particular date). Of little use to "scientists".
SDR ("Sensor Data Record")	Sensor data that has been transformed from "counts" to engineering units. Calibration, bias correction, etc., has been applied. Data is directly linked to geospatial information such as latitude, longitude, date, time. Some data quality flags may be added to the header.
EDR ("Environmental Data Record")	Geophysical quantities derived from the SDRs, such as temperature profiles or total column ozone. Contains the necessary information about date, time and Earth location. Usually has quality-control parameters listed (e.g., did the algorithm converge or is this potentially a bad data point?).
Level 1A	Raw data from the spacecraft are uncompressed and the actual measured "counts" are reconstructed. File headers typically contain important ancillary data such as time, date, spacecraft and target location, and instrument point angle.
Level 1B	Sensor data that has been transformed from "counts" to engineering units. Calibration, bias correction, etc., has been applied. Certain data quality flags are added to the header.
Level 2	Geophysical quantities derived from the SDRs, such as temperature profiles or total column ozone. Contains the necessary information about date, time and Earth location. Usually has quality-control parameters listed (e.g., did the algorithm converge or is this potentially a bad data point?).
Level 3	Data has undergone one or more post-processing steps such as re-gridding to a standard spatial grid, spatial and/or temporal averaging and/or subsetting for certain conditions (e.g. clear/cloudy).

In addition to products distributed by government agencies (primarily NASA and NOAA in the United States), there are a number of groups involved in determining geophysical parameters from satellite data. The list of institutions includes, but is certainly not limited to, AER, Harvard University, the Harvard-Smithsonian Center for Astrophysics (HS-CfA), the University of Maryland, Hampton University, the University of Bremen, the University of Wisconsin, and other governmental organizations such as the Royal Netherlands Meteorological Institute (KNMI), the French meteorological administration (Météo France), the United Kingdom Meteorology Office (UKMO) and the German Aerospace Center (DLR).

3.2 Processing of satellite data

Multiple steps are involved in order to process the satellite data into a form that can be used to validate, constrain or initialize models. The overall scope of this task will vary depending upon the characteristics of the sensor, knowledge of system parameters (such as knowing errors in where the satellite is pointing, and thus where the measurement occurs), and the form in which one wants to use the data (e.g. specific spatial/temporal points or mean values, and geophysical parameters or raw radiances). This section provides an overview of this process and outlines some of the key factors that must be considered when deciding to work with satellite data. The overall process is summarized in Figure 1. Note that not all of these steps need to be done by the air quality modeler, but one should understand the overall process in order to appropriately use the level 2 or level 3 products.

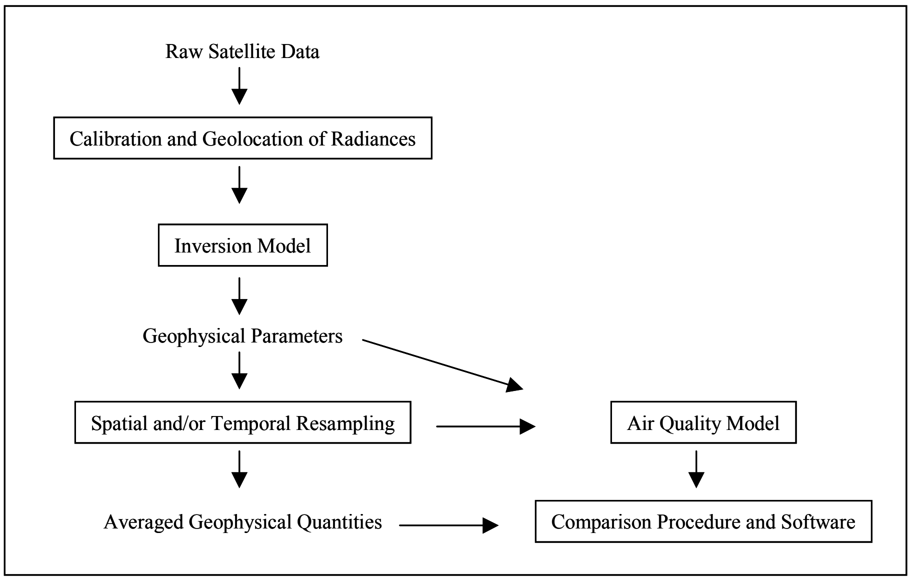

Figure 1: Flowchart of the top level processes that occur when comparing satellite measurements with model outputs (quality control checks that must occur at each step are not shown).

The initial set of processing relates to the determination of geophysical quantities from the measured radiances. This process can be quite complex and requires a number of different steps, from determining calibrated radiances to verifying geo-location information. The calibrated radiance data for a given measurement can be considered as a matrix that is a function of the wavelength of the measurement. The process of determining the geophysical information from the radiance data, the "inversion", involves a matrix inversion:

$$y = Kx \implies x = K^{-1} y \qquad (1)$$

where y is the observation vector (e.g., a set of radiance measurements), x is the set of geophysical quantities of interest (e.g. an ozone profile) and the matrix K represents the translation of the geophysical parameters into radiances. Of course there are a number of pitfalls involved with a non-linear, noisy system, and the science of the inversion process has been thoroughly discussed by a number of authors (Menke, 1984; Rodgers, 1976, 1996, 2000) and is beyond the scope of this document. However, one should realize that there are a number of different approaches to the inversion and the resulting geophysical data can have a wide range of characteristics depending upon temporal and spatial averaging (used to reduce measurement noise), constraints added to the inversion to provide stability, and how well the signal of interest can be recovered from the measurement itself (due to factors such as low concentrations of the quantity of interest as well as interference by other gases, clouds, and the surface).

The use of a priori information is required to constrain the retrieval solution, which is inherently an ill-posed problem with more unknowns (information about the atmosphere) than knowns (the radiance measurement). This information takes the form of both the retrieval quantity itself (e.g. the ozone profile) and a covariance matrix of the uncertainty of this quantity. Thus the a priori information could represent, for example, the climatological mean profile shape along with the inherent variability of this shape. The maximum likelihood retrieval method is one example of this type of retrieval algorithm (Rodgers, 2000).

After the initial processing has been performed, the data may be analyzed in the context of the problem at hand. Depending upon the way in which the data will be used with the model, some additional processing may be required. It is important to match the model temporal/spatial sampling to that of the measurement, realizing that care must be taken in constructing averages of the satellite data. In particular, one must be certain to minimize the influence of external factors that may affect the averaging. The main factor for consideration is clouds, which can obscure the measurement footprint and lead to a misleading characterization of the data of interest. One must also consider temporal issues related to the measurement and the calculation. For example, the time-step for the model represents some sort of averaging in time, and the instantaneous measurement from the satellite may not be directly comparable.

Another factor to consider in the processing of satellite data is what is actually represented by the vertical profile of the data. The comparison of model data with measurements is a relatively straightforward process for total column quantities – one needs only to be sure that the sensor measures the total column directly, or, in the case of a profile measurement that is integrated to provide the total column, that the mass of the species not measured is small compared to the total column or can be quantified and removed from model-measurement differences. The comparison of vertical profiles is more complex.

Remote sensing measurements have an inherent vertical resolution that depends upon the nature of the parameter being measured (i.e. the vertical profile and how changes in the concentration manifest themselves as changes in radiance) and the spectral characteristics of the measurement (i.e., to what degree can the measurement identify the relevant spectral characteristics of the parameter). Thus a vertical profile given for a set of levels above the ground really represents some sort of vertically averaged quantity of the true profile. Models, on the other hand, have their own set of vertical averaging characteristics, dictated by the way in which the model was built. In order to use satellite data in conjunction with model data, it is necessary to match these vertical resolutions. The proper way to treat these differences in vertical resolution is through the use of the sensor "averaging kernel". The following discussion summarizes Deeter (2002).

The averaging kernel represents the way in which the vertical structure of the atmospheric profile is mapped into the radiances measured by the sensor. It is expressed mathematically as a matrix where each row defines the averaging kernel for a particular retrieval level within the measured profile, and each element in this row represents the contribution of other levels in the atmospheric profile to the retrieved profile value. The value of the averaging kernel is a function of sensor parameters (such as the field-of-view) as well as those parameters input to the forward radiative transfer model (such as the temperature profile and the species of interest itself). It will also be a function of the a priori (guess) profile shape if that information is used to constrain the retrieval.

As an example of the use of averaging kernels, consider data from MOPITT and how it must be transformed for comparison with a model. Because the MOPITT retrieval algorithm incorporates a priori information about the profile to constrain the retrieval, the profile retrieved from the measurement is actually a linear combination of the true atmospheric profile and the a priori profile:

$$\begin{aligned} x_{retrieval} &\approx x_{apriori} + A\left(x_{true} - x_{apriori}\right) \\ &\approx Ax_{true} + \left(I - A\right)x_{apriori} \end{aligned} \tag{2}$$

In this equation the x vectors are a function of altitude (or pressure). The averaging kernel is given by "A" and "I" is the identity matrix. As the vertical resolution of the measurement becomes higher, A tends toward the identity matrix and the retrieved profile will match the true profile exactly. Both the averaging kernel and the a priori are provided along with the satellite retrieval and depend on time and location (sometimes the a priori is a constant profile). A is dependent on the sensor used and the radiative transfer model used for the satellite retrieval.

In order to compare the finite-resolution MOPITT profiles with the model output, the model results are first interpolated to the same vertical profile as the a priori and the satellite retrieval and then the averaging kernel matrix is applied to the model output. In other words, the model data is "degraded" in vertical resolution to be directly comparable to the profile retrieved from the MOPITT measurements. This is done directly from equation (2):

$$x'_{MODEL} \approx x_{apriori} + A\left(x_{MODEL} - x_{apriori}\right) \tag{3}$$

The x' quantity for the model can now be compared directly with the MOPITT retrieval.

When comparing regional-scale models to satellite retrieval values, there may be some instances where the model vertical profile is a better approximation to the true profile than the profile retrieved from the satellite data. This may occur because the retrieved profile is strongly influenced by the a priori profile, particularly when deviations from the a priori profile occur near the surface or in the upper troposphere (for example, the satellite sensors are most sensitive to CO in the mid-troposphere). Examples include urban air pollution (i.e., deviation from the a priori profile near the surface, Vijayaraghavan et al., 2006b) and biomass fires with high plume rise (i.e., deviation from the a priori profile in the upper troposphere, Gevaerd et al., 2006). Then, it would be more appropriate to use the model vertical profile as the a priori profile, to retrieve the satellite data using this location/time specific profile, and to compare the new retrieved profile to the model results (note that if the model profile is used as the a priori profile, there is no need according to Equation 2 to process the model output through the averaging kernel). However, because the satellite data must be reprocessed, this approach is more resource-intensive than that where the model output is processed via Equation 2.

4. Use of Satellite Data for Air Quality

4.1 Evaluation of air quality models

The evaluation of regional air quality models is typically conducted using ground-level measurements and occasionally using data aloft from aircraft, helicopters, sondes or towers. Satellite remote sensing data provide two important sources of information compared to surface monitoring data: more complete spatial coverage and a vertically-integrated measure of air quality (e.g., Engel-Cox et al., 2004; Al-Saadi et al., 2005; Edwards et al., 2006). Satellites in polar orbits provide good spatial (typically global) coverage albeit at a low temporal resolution. However, satellite measurements have several limitations as discussed earlier. These must be taken into account before air quality model outputs are compared with satellite data for the purpose of model performance evaluation.

Many of the chemical species that are simulated by an air quality model are present in trace amounts that cannot be measured by satellite sensors. Also, a large number of chemical species may be planned for retrieval during a satellite mission (as listed in Tables 3, 4, and 5) but many of these are either not retrieved due to instrument/algorithm issues or are not validated and thus not quickly made available for public dissemination. Validated tropospheric satellite data are commonly available only for the following species or physical quantity: O_3, CO, NO_2, SO_2, HCHO, and AOD. So it is currently feasible to evaluate tropospheric air quality models typically for these species only. Thus, model evaluation studies have tended to focus on one or more of these species (e.g., Chin et al. 2002; Boersma et al. 2006; Byun et al. 2006; Fishman et al. 2006a, 2006b; Hodzic et al. 2006; Jing et al. 2006; Kondragunta et al. 2006, Lyon et al. 2006; Rao et al. 2006; Pickering et al. 2006; Pierce 2006; Szykman et al. 2006; Vijayaraghavan et al., 2006a, 2006b; Zhang et al. 2005; Ziemke et al. 2006). Satellite data for other species such as HNO_3 are becoming available (e.g., Santee, 2006) but are sometimes limited by the time period of availability.

Currently, air quality models are typically evaluated using point surface measurements with various averaging times (ranging from 1 hour to 24 hours). The 1-hour temporal resolution of the model is, therefore, consistent with or finer than that of the measurement. However, the spatial resolution of the model which is of several kilometers provides a volume-average result that is generally not consistent with the point measurement. When using satellite data, we will typically compare the model to data that offer a range of temporal and spatial resolutions. As discussed earlier, satellites in geostationary orbit are able to provide high temporal resolution with reduced spatial coverage, while satellites in polar orbits provide global coverage with less frequent temporal coverage. Because most satellites involved in the remote sensing of air quality have a sun-synchronous polar orbit (see Table 1), poor temporal resolution is usually more of a concern than spatial coverage. For example, sensors such as OMI typically provide global coverage at 13 x 24 km horizontal spatial resolution (see Table 4) which is comparable to the typical horizontal resolution of regional-scale air quality models (12 to

36 km) but provide data only once a day. Thus, weekly, monthly, seasonal, and annual averages (rather than hourly or daily values), at the particular time of day of the satellite observation, are more appropriate for comparison between the model and satellite data due to the latter's poor temporal resolution.

To ensure horizontal spatial compatibility between the modeling results and the satellite measurements, we need to either re-grid the satellite data to the air quality model grid or the model results to the satellite grid before comparing the two. Although recent sensors such as OMI have a relatively fine spatial resolution, poor temporal resolution and other confounding factors could necessitate a comparison at a coarser horizontal resolution. For example, consider the AOD product derived from MODIS data. The inherent spatial resolution is quite high, though the temporal coverage is limited. Moreover, one wants to separate cloudy measurements (which are not representative of the total AOD) from those of "clear sky" aerosol measurements. This selective sampling of the data can introduce biases and thus the result should be compared to a similar set of data points from the model. Selecting only the high resolution pixels that meet this sort of quality control requirement will result in a sparse dataset. Also, the model run itself is not exactly on the same time grid as each of the selected measurement points, making it difficult to make a meaningful comparison of the measurement and the model. To mitigate these effects, and ensure that all researchers are using a consistent, validated product, the MODIS team has developed a data product consisting of a uniform 1^o x 1^o spatial grid that contains AOD measurements. This Level-3 product is available for data averaged over various timescales and can be compared with air quality model outputs (e.g., Matsui et al., 2004). The inputs to the average are only the measurement points that meet the quality control criteria. Thus each averaged point consists of a temporal and spatial average of measurements. For comparison purposes with models that have a finer spatial and/or temporal scale than this measurement grid one should perform a similar averaging procedure for the model output prior to the comparison with the measurement product. Thus, the comparison of models and data requires a careful understanding of the data and the spatial and temporal scales associated with both the quantity measured and the measurement process itself.

One of the key advantages of using satellite data for evaluation of air quality models is the availability of data aloft. Satellite retrievals may be available as a total atmospheric column (e.g., for ozone, AOD), tropospheric column (e.g., for NO_2, ozone, AOD, HCHO), and/or vertical profiles (e.g., for CO). Several studies have compared air quality model simulations with satellite retrievals of ozone column (e.g., Fishman et al. 2006b; Liu et al., 2006; Ziemke et al., 2006), tropospheric NO_2 column and profiles (e.g., Boersma et al. 2006; Eskes and Levelt, 2006; Kunhikrishnan et al., 2006), CO column and profiles (e.g., Allen et al., 2004; Vijayaraghavan et al., 2006b), total and tropospheric column AOD (e.g., Chin et al. 2002; Hodzic et al. 2006; Yu et al., 2003), and HCHO column (e.g., Wittrock et al., 2006; Vijayaraghavan et al., 2006a).

A complicating factor in the evaluation of model results aloft is the representation of the vertical grid for both the model and the measurement. As discussed above in terms of "averaging kernels", the vertical resolution of the satellite measurement is limited by characteristics of the sensor itself (such as the vertical field-of-view for a limb-viewing sensor) and the spectral/radiometric properties of the atmosphere itself. For example, it is very difficult to measure the tropospheric column of ozone from a satellite because of interference due to the large stratospheric column that exists between the troposphere and the measurement (thus one is trying to measure very small changes, the tropospheric column amount, to a fairly large number, the total atmospheric ozone column amount). Even for measurements of the total atmospheric column, the impact of clouds must be considered within the spatial averaging domain as for these points the vertical column is no longer from space to the surface. As discussed earlier, the total ozone column and stratospheric ozone column are usually used to derive the tropospheric ozone residual (TOR). In contrast, Liu et al. (2005b, 2006) demonstrated that the global distribution of Tropospheric Column Ozone (TCO) could be directly retrieved from GOME data. They followed the methodology of Chance et al. (1997) and used observations of backscattered radiance spectra with moderate spectral resolution of 0.2-0.4 nm and the presence of a high signal to noise ratio in the ultraviolet ozone absorption bands to retrieve the vertical distribution of ozone down through the troposphere. In both cases, the selection of the tropopause and additional processing are important during the comparison of model-derived TCO with satellite-retrieved TCO/TOR. For example, Pickering et al. (2006) took the following steps when comparing the ozone column amount from the Eta/CMAQ model with the TCO calculated from OMI total column ozone minus MLS SCO over the United States. They regridded only the Eta/CMAQ results at 1900 UTC, i.e., the closest to the OMI overpass time in the Eastern United States, to the OMI 1^{o} x 1.25^{o} grid and integrated model data from the surface to the NCEP tropopause. They then filtered the OMI-MLS and Eta/CMAQ gridded data to remove regions where the tropopause pressure (as determined by NCEP) exceeded 170 hPa (this eliminated strong stratospheric ozone gradients). The OMI Level-2 averaging kernels were then applied to the Eta/CMAQ tropospheric ozone profiles before comparing with OMI-MLS TCO data.

Another important consideration in the comparison of model simulation results with satellite data is the uncertainty in the satellite measurements. Clearly, there are uncertainties associated with ground-based air quality measurements (e.g., Bhave, 2004; Seigneur, 2004); but in many cases, these uncertainties can be minimized and/or estimated. Satellite air quality measurements are subject to several limitations as explained earlier. There may also be errors in satellite data due to factors such as sensor radiometric calibration, bias in the spectroscopic parameters used in the retrieval algorithm, and/or the choice of a priori constraints used to stabilize the retrieval algorithm. The magnitude and characteristics of these errors is determined during the post-launch calibration and validation period. The nature of these error sources is such that the data are often found to have a slight systematic bias, yet are still able to capture the spatial variations of the quantities of interest. Thus, there can often be good qualitative agreement between the model and the measurement even with not so good quantitative agreement. This is not a limitation of satellite data, but rather a warning that it is important to understand the influence of the retrieval process on the output geophysical parameters when performing model/measurement comparisons. In general, satellite data should be used as a quantitative benchmark in the performance evaluation of air quality models only after the satellite retrievals have been independently validated against other data such as from aircraft, sondes and ground-based measurements.

Satellite measurements are routinely validated by the retrieval team before being released to the public. For example, Emmons et al. (2004) validated MOPITT CO retrievals with aircraft profiles at several locations around the Earth during 2000-2002. NASA and other organizations also conduct special aircraft campaigns, such as the Intercontinental Chemical Transport Experiment (INTEX-A) in 2004, which provide additional data for validation. Additional examples of satellite data verification include the validation of ACE-SCISAT retrievals of O_3 (Walker et al., 2005), TES retrievals of CO, O_3 and water vapor (Osterman et al., 2005) and MODIS retrievals of AOD (Kleidman et al., 2005). The data used to validate the satellite retrievals are typically different from the land-based data used for air quality model evaluations. The satellite data validations are, however, often done for specific time periods corresponding to the aircraft campaign flights. It is not reasonable to expect the satellite retrievals to be validated for all time periods that are of interest to the air quality modeler. Thus, satellite data should be used in conjunction with other measurements, and not as a sole test, for the evaluation of air quality models.

4.2 Boundary and initial conditions

Boundary conditions can have a significant influence on the simulated pollutant concentrations, particularly for pollutants that have a long atmospheric lifetime, such as $PM_{2.5}$ in the absence of precipitation. In-situ measurements available to provide those boundary conditions are generally sparse and limited to surface locations. Global-scale models are typically used to provide concentrations at the boundaries of a regional model; however, there may be significant uncertainties associated with their simulated concentrations. Satellite data provide the spatial coverage needed for the boundary and initial conditions of regional air quality models, particularly, aloft and over the oceans and other areas where other data may not be available.

The IDEA (Infusing satellite Data into Environmental Applications) project, an EPA/NASA/NOAA partnership, aims to improve the results of regional air quality model simulations when upper and lateral boundary air pollutant data from satellites are used to describe the influx of pollutants (Neil et al., 2004). Satellite measurements, when used as boundary conditions, can be used to account for the contributions of pollutants transported over long distances, for example, from Asia to the United States over the Pacific Ocean. This phenomenon of long-range transport has been demonstrated in several studies using CO data from MOPITT (Heald et al. 2003), O_3 data from TES (Zhang et al., 2006), AOD data from MODIS (Heald et al. 2006), etc. A combination of satellite measurements and air quality modeling of O_3 and CO can be used to quantify the continental outflow of these pollutants from the United States (e.g., Zhang et al., 2006).

Initial concentrations of chemical species such as O_3 and CO may have a non-negligible effect on simulations that cover a limited time period (say, a week or less). A "spin-up" simulation period is typically used to minimize the effect of the initial conditions but using available data can nevertheless improve the initialization of an air quality model, particularly when conducting air quality forecasting (Goldberg and Kondragunta, 2006).

The incorporation of satellite data as boundary and initial conditions for air quality models will be subject to some of the disadvantages of satellite data discussed earlier such as limited temporal and/or spatial resolution, gaps in data due to clouds, and availability of column information rather than vertical profiles. Nonetheless, this area promises to be a fruitful application of satellite measurements for air quality modeling.

4.3 Inverse modeling and data assimilation

There has been an increasing effort over the past few years to use experimental data to improve the performance of air quality models. This process involves inverse modeling and data assimilation. Inverse modeling consists in using the observations to derive optimized model inputs that are most consistent with those observations. Variational methods (also referred to as adjoint methods because they use an adjoint model of the air quality model), sequential methods (based on the use of Kalman filters) and iterative methods (minimization of the modeling error by iteration) are typically used to perform the inverse modeling. Inverse modeling has been used in air quality modeling to optimize emission inventories (e.g., Mendoza-Dominguez and Russell, 2000; Gilliland et al., 2003) and boundary conditions (e.g., Roustan and Bocquet, 2006). The optimized model inputs may then be used to generate a model simulation that is in better agreement with the data. It is also possible to assimilate concentration data directly into the simulation to force the model simulation toward the observations as the simulation progresses. Finally, model simulation results and measurements can be combined to create air pollutant concentration (or deposition flux) fields that leverage the best information of both data sets; this type of data assimilation, which is conducted after completion of a model simulation, is typically referred to as data fusion. Most of these applications so far have used surface measurements of ambient pollutant concentrations or wet deposition fluxes. There is obviously some interest and potential benefits in using satellite measurements as well for inverse modeling and data assimilation. In particular, this is an area of current research for air quality forecasting (Goldberg and Kondragunta, 2006; Hollingsworth, 2004).

4.3.1 Basics of inverse modeling and data assimilation

Chemical transport models (CTMs) of air quality simulate atmospheric chemical concentrations (and deposition fluxes) using initial and boundary conditions, meteorology and emissions as inputs. In inverse modeling, the inverse problem is solved: given measurements/estimates of chemical concentrations (or deposition fluxes), one calculates some of those inputs (e.g., emissions or boundary conditions). The set of concentrations defines the model state and these concentrations are referred to as the state variables; the inputs are referred to as the control variables. Although there exist a large number of techniques that have been applied to conduct inverse modeling in atmospheric science (Enting, 2002), the two major approaches are the sequential methods (which include various Kalman filter techniques) and the variational methods (also referred to as adjoint techniques).

4.3.1.1 Brief overview of inverse modeling, sensitivity analysis and data assimilation

Figure 2 depicts schematically the overall process of inverse modeling and data assimilation. In this conceptual example, we can consider the model input to be the emission inventory of the precursor species (e.g., NO_x), E, and the model output to be the atmospheric concentrations of the chemical species of interest (e.g., NO_2), C. In the case of the use of satellite data, these concentrations will be vertically integrated to provide the column density. The model is first applied with an initial guess of the input, which is referred to as the a priori input, E_a. There is an error (uncertainty), ε_E, associated with this a priori input. Observations, O, such as column densities from satellite data, are available with an associated error, ε_O. (In theory, an error due to the model can also be associated with the model output and added to the observation error; since it is generally not invoked in most current applications, it is not included in this example. It seems counterintuitive to associate the observation error and model error; however, the model and the observation are used here together to obtain the input and our ability to estimate the "true" input is limited by both the error in the observation and the error in the model.) The objective is to optimize the model input. The optimized model input is referred to as the a posteriori input. (Representativeness error refers to scales that are included in the observation that are not "represented" by the model. Representativeness errors are often treated together with the observation error because, as discussed above, they are model errors that, together with the observation error, affect our ability to estimate the "true" input.)

The general procedure for obtaining an optimized model input involves inverse modeling. First, a cost (or objective) function is defined based on Bayesian and maximum likelihood arguments. With the assumption of normal (i.e., Gaussian) errors, the cost function is found to be the sum of two terms: (1) the first term represents the deviation of the model output from the observation (we can refer to this term as the performance term) and (2) the second term represents the deviation of the input from the a priori input (generally referred to as the penalty term). Each term is weighted by its associated error such that if the error is large the term contributes less to the cost function. For example, if the satellite data have large associated errors (e.g., due to presence of clouds, assumption of vertical profile, interference from aerosols), the first term will have little influence on the cost function and the a priori model input will be heavily weighted. On the other hand, if there are huge uncertainties in the a priori input, the second term will have little weight and the optimized input will depend almost entirely on the first term, i.e., the observation.

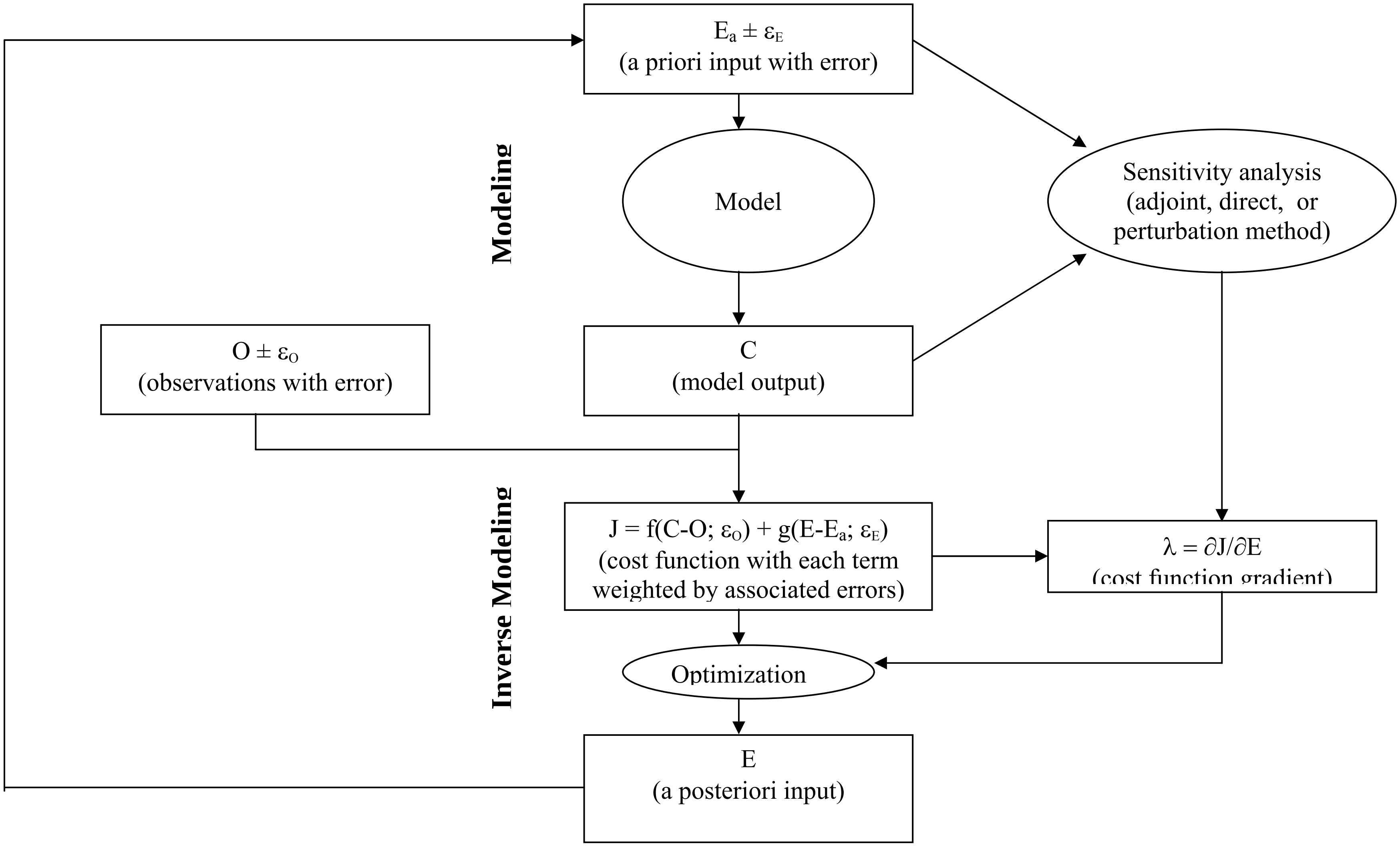

Figure 2. Schematic representation of the sensitivity analysis, inverse modeling and data assimilation steps with respect to modeling.

The cost function will be minimized when its derivative with respect to the input becomes zero. Therefore, the optimization procedure requires obtaining the first-order sensitivity (or gradient) of the cost function to the input. A standard approach to obtain the model output sensitivity is to calculate the derivative of the model output with respect to a model input, which can be done by solving a set of differential equations for this derivative (the direct method) or by applying a small perturbation to the model input (perturbation or indirect method, also referred to as "brute force" method) and comparing the results to the original simulation. This approach is sometimes referred to as "forward sensitivity" because the sensitivity equations are solved forward in time. Note that the result ($\partial C(x,y,z,t)/\partial E$) provides the sensitivity of the model output that is a function of space and time with respect to a model input that is not resolved in space nor time (typically a perturbation across the board or averaged over a specific domain and time period). Sequential methods such as the Kalman filter use a forward sensitivity analysis. On the other hand, variational methods use the adjoint sensitivity, which is obtained by means of an adjoint model. The adjoint model reverses the order of the computations of the forward model. The variable solved for in the adjoint model is the sensitivity of the output of the original model (here, C) to the input (here, E). In the case of the adjoint sensitivity, the model output is averaged over some spatial domain and time period and the adjoint sensitivity equations provide its first-order derivative with respect to a model input that is resolved in space and time ($\partial C/\partial E(x,y,z,t)$. Thus, the forward and adjoint sensitivities provide distinct types of information.

If the final objective is to obtain a model simulation that is in better agreement with the observations than the original simulation, the a posteriori input can then be used in the model. This forward model step is, of course, part of the iterative adjoint solution. In inverse modeling only the model input is affected, and the model is considered to be a perfect representation of the atmosphere. It is also possible to perform the assimilation of the observational data directly (i.e., without any inverse modeling) by forcing the model output toward those data. (This is analogous to the data assimilation process used in numerical weather prediction.) However, the result is then a combination of the model output and observations and it does not correspond to a solution of the model (this includes techniques referred to as nudging or data fusion).

We provide some additional descriptions of the sequential and variational methods below, highlighting their respective advantages and shortcomings.

4.3.1.2 Sequential methods

Sequential methods use available measurements to correct the prediction of the model for the next time step. The Kalman filter is the best known sequential method. In the standard application of the Kalman filter, the output of the model is compared to measurements (or estimates derived from measurements). The sensitivity of the model input to the variable being measured is obtained, for example, by conducting two separate simulations with slightly different initial (a priori) inputs and relating the change in the output variable to the difference in the input or by using a forward sensitivity method such as the decoupled direct method (DDM). Then, a new (a posteriori) value of the input is calculated by using the measured value and the sensitivity of the input to the measured variable. If several measurements are available, then a least-square minimization can be conducted to obtain the solution. An advantage of the Kalman filter is that it evolves the error covariance matrix of the model state using the model dynamics. At the same time, it takes into account the error in the measurements and the error in the model input to calculate the optimized model input. As discussed above, if large errors are associated with the measurements, the initial (a priori) value of the model input will carry more weight than the measurement. Conversely, if the measurements are accurate, the a posteriori value of the model input will be strongly influenced by the measurements.

The advantage of this method is that it is relatively easy to implement. At each model time step, available measurements are used to minimize the error in the model and the next model time step is calculated using the improved model state estimate. The disadvantage of this method is that it only uses information at a given time step. Consequently, applying a Kalman filter will not provide spatial or temporal information on the optimization of the inputs. Furthermore, if the differences between the model outputs and the measurements are significant and vary widely, the successive corrections may lead to a result that shows discontinuities as a function of time. Such discontinuities can be minimized by applying the Kalman filter over several previous time steps, thereby smoothing the error minimization over several time steps. However, the application of the Kalman smoother requires inverse modeling via a variational approach (see below). Kalman filters have been applied with global models such as GEOS-Chem and regional/urban models such as CMAQ.

4.3.1.3 Variational methods

Variational methods are optimization techniques that also provide sensitivity analysis information. They are based on the definition of an objective (or cost) function that is to be minimized. As mentioned above, the cost function includes two terms. In the variational method, the second term is weighted by a regularization parameter that is defined empirically. A first-order derivative of the cost function is calculated to carry out the minimization process. As discussed above, the variational analysis requires the development of the adjoint model of the CTM. In theory, the equations for the adjoint model can be developed analytically and subsequently solved numerically. However, it is typically best to develop the adjoint from the numerical code of the CTM to eliminate inconsistencies between calculations of the objective function and calculations of its gradient. The validity of the adjoint code should be checked for accuracy by carrying out numerical tests of input perturbations.

The advantage of the variational method is that it uses information from previous time steps and it provides sensitivity information that is spatially-distributed and temporally-resolved with respect to the model input. It is, therefore, useful if one wants to optimize a model input (e.g., emission inventory) as a function of location and time, rather than across-the-board (see example below). The disadvantage is that it requires the development of the adjoint model, which can be a significant effort for a CTM. Note, however, that adjoint models have been developed for the global CTM GEOS-Chem (Henze and Seinfeld, 2006) and the regional/urban CTM CMAQ (Hakami et al., 2006).

4.3.2 Applications of inverse modeling/data assimilation to satellite data

There are two major categories of applications of inverse modeling techniques using satellite data: (1) the estimation of a set of inputs by minimizing the difference between the model output and the satellite data and (2) the assimilation of satellite data in the model simulation. Both approaches aim at improving model performance by either refining the model inputs or correcting the model output. We describe some practical applications of those two major categories.

It should be noted that in all these examples a CTM is used and the input meteorology is held fixed. In fact, observations of chemical species may also be useful to refine estimates of the meteorology. For example, a passive tracer provides information on advecting winds. Coupled chemistry meteorology models may one day make optimal use of both types of observations and for many years ECMWF has included ozone in meteorological models.

4.3.2.1 Estimation of input data

The most common application of inverse modeling to CTMs has been to estimate emission data. For example, Gilliland et al. (2003) applied a Kalman filter with CMAQ to estimate ammonia emissions based on ammonium wet deposition data, Mendoza-Dominguez and Russell (2000) applied the Direct Decoupled Method (DDM) of sensitivity analysis combined with an iterative optimization technique with a CTM to estimate ozone precursor emissions and Pison et al. (2006) applied the adjoint of a CTM to optimize NO_x emissions in an ozone simulation over the Paris region. These examples used surface data with regional-scale CTMs. To date, the use of satellite data for inverse modeling has mostly been limited to global-scale CTMs, although some recent work has been done with regional-scale models such as CMAQ as well. We summarize some recent examples below.

Martin et al. (2003) performed inverse modeling with GEOS-Chem using NO_2 column densities retrieved from the GOME satellite to estimate global NO_x emissions. The estimation of NO_2 column densities was improved by (1) taking into account the effect of atmospheric particulate matter on the signal and (2) by using the vertical profile simulated by GEOS-Chem instead of assuming a universal profile. This analysis ignored atmospheric transport and the NO_2 column density was assumed to be related to the NO_x emission rate in that column through linear chemistry relationships only. Those relationships (NO_2 fraction of NO_x and loss of NO_x via chemical reactions) were obtained from the GEOS-Chem simulation (and, therefore, varied among the columns). The spatial resolution of the analysis was 2^o latitude and 2.5^o longitude. The analysis took into account the errors present in the a priori and estimated emissions simply by calculating the a posteriori emission as the geometric average of the a priori and estimated emissions weighted by their respective errors. The results showed (1) improvement in the GEOS-Chem simulation when the results are compared to the GOME-derived NO_2 column densities (the error was halved, an expected result since the a posteriori inventory includes information from the GOME column densities), (2) good agreement on average with two global emission inventories, (3) but significant differences with those inventories in specific regions of the globe. Martin et al. (2006) repeated the same analysis using SCIAMACHY NO_2 data and adding NO_x emissions from lightning. Their results highlighted the rapid increase of NO_x emissions in Asia.

Palmer et al. (2006) performed a similar analysis to estimate isoprene emissions from the HCHO column density from the GOME satellite. Errors in the retrieval of the HCHO column result in part from interferences from particulate matter, the assumed HCHO vertical profile and clouds, with clouds being the major source of error (Millet et al., 2006). The estimation of HCHO column densities used the vertical profiles simulated by GEOS-Chem. The error in the retrieved HCHO column density was estimated to be about 40%. The HCHO concentrations were related to the isoprene emissions via a linear relationship that accounted for the oxidation of isoprene to HCHO, the photolysis of HCHO and the oxidation of other VOC to HCHO (considered a background HCHO concentration due to the lower oxidation rate of most VOC compared to isoprene). As for NO_x, transport was ignored (the oxidation of isoprene to HCHO was considered to be faster than the advection of isoprene at the GEOS-Chem spatial resolution). In this application, no a posteriori emissions were calculated; instead, the GOME-derived isoprene emissions (so-called top-down, literally) were compared to values from an emission model (so-called bottom-up) and from measurements. This approach seems appropriate for estimating isoprene emissions with a coarse spatial resolution; at a finer resolution, emissions of anthropogenic VOC, which are sources of HCHO, will interfere with the inverse modeling of isoprene emissions. Therefore, at regional/urban scales, it will be necessary to include some treatment of transport processes and to use a variational approach.

Kopacz et al. (2006) have used a variational approach to optimize CO emissions with good spatial resolution using CO column densities from the MOPITT satellite with GEOS-Chem. As mentioned above, the objective of estimating emissions with good spatial resolution requires the use of a variational method because all relevant atmospheric processes can then be taken into account (i.e., chemistry and transport).

Hakami et al. (2006) used NO_2 column densities from the SCIAMACHY satellite to optimize the NO_x emissions using a variational approach with CMAQ. Because the emission input file of CMAQ is a 3-D file, the inverse modeling was spatially resolved for the 3-D field of 47,610 grid cells (some of the non-surface grid cells have zero emissions). No time dependence of the scaling of the emissions was used. The inverse modeling was conducted for a modeling domain that covers the southeastern United States with a 36 km resolution for a three-day period (this period corresponds to the time needed for the satellite to cover the full domain). The results showed considerable improvements in CMAQ's ability to reproduce the satellite NO_2 column densities. However, the NO_x emissions needed to be scaled up with factors in the range of 1.5 to 3 with some values up to 5. These results suggest a very significant underestimation of the NO_x emission inventory. It is possible that NO_x emissions from lightning, which are missing from the original CMAQ emission inventory, are a significant cause of this discrepancy.

In addition to emission estimation, inverse modeling can also be used to estimate boundary conditions of regional-scale CTMs. For example, Roustan and Bocquet (2006) used mercury surface data with a variational approach to estimate mercury boundary concentrations for a CTM domain covering Europe. Their results properly reflected the influence of the northern boundary during spring when mercury depletion events lead to increased reactive gaseous mercury concentrations. Also, Vautard et al. (2000) applied a CTM with its adjoint to obtain the sensitivity of simulated O_3 concentrations to upwind boundary conditions and to improve model performance using the optimized boundary conditions. However, to our knowledge, satellite data have not been used to estimate boundary conditions for a regional-scale model. Since satellite data have a large spatial coverage, they can be used directly to provide information on boundary conditions and, therefore, their use in an inverse modeling exercise would seem superfluous. Nevertheless, Hakami et al. (2006) listed the estimation of boundary conditions with a variational method as a possible future task.

4.3.2.2 Data assimilation

Data assimilation is primarily of interest in two areas: (1) air quality forecasting (or hindcasting) and (2) data fusion. Air quality forecasting is typically conducted over a period of one day to a few days. Although the atmospheric diffusion equation, which governs the air quality system, unlike the equations governing meteorology, is not chaotic, it is useful to have initial conditions that are as accurate and consistent as possible. Thus, data assimilation can be useful to improve the initial chemical concentration fields. Data fusion refers to the combination of model simulations results and measurements. Spatial interpolation techniques are used to develop fields of concentrations and atmospheric deposition fluxes that combine model outputs and measurements. Assimilating satellite (and surface) data into the results of a model simulation is similarly a data fusion approach. Data assimilation for air quality forecasting is a dynamic process because the data affect the model simulation as it progresses, whereas data fusion can be seen as a static assimilation process because the assimilation step does not feed back into the model simulation.

Jones et al. (2006) used a Kalman filter to assimilate CO and O_3 data from the Aura Tropospheric Emission Spectrometer (TES) in two global-scale models, GEOS-Chem and AM2, for a two-week period in November 2004. The assimilation led to significant increases in CO throughout the southern hemisphere and significant increases (20 to 50%) in O_3 over the Indian Ocean and the Indonesian/Australian region. The model results with data assimilation showed much better agreement with the satellite data, as expected, and GEOS-Chem and AM2 were in much better agreement after data assimilation. The revised O_3 simulation affected the NO_x chemistry, which would affect the NO_x/NO_2 relationship that is used in the retrieval of NO_2 satellite data to estimate NO_x emissions (see above). This result points out the interrelationship between various chemical species in a non-linear system and the benefits (i.e., lower errors) that can be gained by conducting satellite data retrieval jointly for several chemical species.

Sarigiannis et al. (2006) assimilated AOD data from the Earth Observation (EO) satellite into REMSAD using a Kalman filter to develop maps of PM_{10} concentrations over southern Europe. The objective was to combine the model simulation output results with satellite data to obtain improved PM_{10} concentrations. Yu et al. (2003) presented global monthly distributions of AOD after integrating AOD from MODIS retrievals with that from Georgia Tech/Goddard Global Ozone Chemistry Aerosol Radiation and Transport (GOCART) simulations to account for gaps in the MODIS data due to highly reflective arid and snow-covered lands. These analyses are examples of the use of satellite data for data fusion.

The use of satellite data, along with surface ambient air quality measurements, to improve air quality forecasting is planned in the United States (e.g., Hoff et al., 2006; Kondragunta et al., 2006). These plans include assimilation of AOD and O_3 column densities into CMAQ simulations for air quality forecasting. In Europe, Eskes and Levelt (2006) have presented plans to use a combination of data (O_3, CO, NO_2, SO_2, HCHO and CH_4) from the OMI (Aura) satellite and surface data to improve model performance for air quality forecasting using a variational approach.

NOAA has deployed a smoke forecast tool which integrates GOES satellite information on the location of wildfires with NOAA National Weather Service weather inputs from the North American Mesoscale model and smoke dispersion simulations from the NOAA Research HYSPLIT model to produce a daily updated 48-hour prediction of surface $PM_{2.5}$ concentrations due to wildfire smoke transport in the US (http://www.arl.noaa.gov/smoke/).

4.3.3 Future prospects

Information from satellites on concentrations of chemical species and particulate matter in the atmosphere can be very valuable to improve the performance of air quality models. We can distinguish three major categories of procedures to use satellite data in air quality modeling:

- *Input data optimization* pertains to the development of optimized emission inventories (and possibly boundary conditions) that provide better agreement with the data.

- *Data assimilation* involves the use of the data in an air quality simulation, either directly (concentrations) or after processing (optimized emissions), to obtain better performance.
- *Data fusion* combines the results of an air quality simulation with data to develop fields of air concentrations or atmospheric deposition fluxes that leverage the best aspects of model results and data.

Table 7 presents an overview of the possible applications of satellite data to those various aspects of inverse modeling and data assimilation.

For the optimization of input data, satellite data can be very useful for the improvement of the emission inventories of some species and source categories (see Section 4.3.2 for some recent examples at the global scale and the regional scale). For global scale applications where the spatial resolution allows one to neglect the effect of atmospheric transport, the use of a Kalman filter or similar technique is sufficient. However, at the regional scale where both chemistry and transport are important in the inverse modeling process, a variational approach is needed. As mentioned above, we do not see the optimization of boundary conditions as a major application of satellite data via inverse modeling because the satellite data can be used directly to specify the boundary conditions (see Section 4.2).

For data assimilation into an ongoing air quality simulation, both concentrations and optimized emissions can be assimilated to improve model performance. In the case of concentrations, assimilation of satellite data for O_3, NO_2 and AOD (surrogate for PM) will directly improve the performance of air quality simulations for these pollutants. In the case of emissions, the ability to perform the inverse modeling in a real-time manner will be the limiting step for air quality forecasting (this is not an issue for hindcasting). There are some specific areas where such data assimilation will be key because satellite data provide information that is not directly available from other sources: for example, AOD measurements can provide valuable information on biomass fires (Kondragunta et al., 2006) and SO_2 measurements can help characterize volcanic eruptions.

For data fusion, satellite data will be most useful to improve concentration maps of air pollutants such as O_3, NO_2 and PM (using AOD). Some applications have already been performed (see Section 4.3.2) or are ongoing for the display of air pollutant concentration maps or their use in epidemiological studies.

Table 7. Overview of the possible applications of satellite data in air quality modeling with inverse modeling and data assimilation.

Procedure	Specific Application	Use of Satellite Data
Input data optimization	Emission inventories	NO_2 data for NO_x emissions HCHO data for isoprene emissions CO data for CO emissions AOD data for $PM_{2.5}$ biomass fire emissions SO_2 data for volcanic emissions
	Boundary conditions	Not likely because data can be used directly as model input (see Section 4.2)
Data assimilation	Emissions	Same as above for input data optimization but with near-real time (nrt) processing
	Concentrations	O_3, AOD, NO_2, SO_2, HCHO, CO
Data fusion	Concentrations	O_3, NO_2 and AOD (for PM)
	Deposition fluxes	No direct application because satellite data do not provide quantitative information of atmospheric deposition; use of ground-level monitoring data instead.

5. Conclusion

Satellite remote sensing measurements of some chemical species offer more complete spatial coverage and integrated vertical column/profile information compared to the in-situ data typically used to evaluate, constrain or initialize regional air quality models. Validated tropospheric satellite data are currently available typically only for the following species and physical quantities: O_3, CO, NO_2, SO_2, HCHO, and AOD. So most current evaluations of tropospheric air quality models focus on these species. As new validated data for species such as HNO_3 become available, these could be used for model evaluation and data assimilation.

Multiple steps are involved in retrieving air quality data from the observed radiances in the atmosphere and an a priori or background profile. The air quality modeler would typically be interested in the final processed product that provides the temporally- and spatially-averaged geophysical quantity of interest. The satellite retrieval and the air quality simulation results should be mapped to the same spatial and temporal grid and processed with the satellite averaging kernel matrix before comparison with each other.

Most satellites currently used for the remote sensing of air quality are sun-synchronous polar-orbiting and provide data at a higher vertical and horizontal resolution than geo-synchronous satellites but at a poor temporal resolution, once or twice daily. Satellite data may suffer from other limitations such as a coarse horizontal spatial resolution compared to the air quality model resolution and uncertainties in retrieval due to cloud cover, ground albedo, and day/night and land/ocean differences in the sensor measurement errors. There may also be errors in satellite data due to factors such as sensor calibration, bias in the spectroscopic parameters used in the retrieval algorithm, and/or the choice of a priori constraints used to stabilize the retrieval algorithm. In general, satellite data should be used as a quantitative bench mark in the performance evaluation of air quality models only after the satellite retrievals have been independently validated against other data such as from aircraft, sondes and ground-based measurements. Such validation is routinely done by the retrieval team, usually using aircraft/sonde data which are different from the surface data typically used for air quality model evaluations.

Satellite data provide the spatial coverage needed for the boundary and initial conditions of regional air quality models, particularly, aloft and over the oceans and other areas where other data may not be available. Satellite measurements, when used as boundary conditions, can be used to account for the contributions of pollutants transported over long distances, for example, from Asia to the United States over the Pacific Ocean. This phenomenon of long-range transport has been demonstrated in several studies using CO, O_3 and AOD data. A combination of satellite measurements and the air quality modeling of species with long residence times such as O_3 and CO can be used to quantify the continental outflow of these pollutants from the United States and other countries. Satellite data can also be used to initialize an air quality model, particularly when conducting air quality forecasting.

Another promising application of satellite data in air quality modeling lies in the areas of inverse modeling/data assimilation. Inverse modeling consists in using the observations to derive new model inputs that are optimized with respect to those observations using variational (adjoint) methods, sequential (e.g., Kalman filter) methods, and iterative methods. It is also possible to assimilate concentration data directly into the simulation to force the model simulation toward the satellite observations as the simulation progresses or to combine simulation results and satellite measurements to create air pollutant concentration or deposition flux fields after completion of a model simulation (data fusion). These are all rapidly evolving areas of research and each have advantages and disadvantages. Sequential methods are easier to implement than variational methods. For global scale applications where the spatial resolution allows one to neglect the effect of atmospheric transport, the use of a Kalman filter or similar technique is sufficient. However, at the regional scale where both chemistry and transport are important in the inverse modeling process, a variational approach is needed. But variational methods require the development of the adjoint of the air quality model, a non-trivial process. For the optimization of input data, satellite data can be very useful for the improvement of the emission inventories of some species and source categories. For data assimilation into an ongoing air quality simulation, both concentrations and optimized emissions can be assimilated to improve model performance, for example, in air quality forecasting. Satellite data are useful to provide information at inaccessible areas: for example, AOD measurements from biomass fires and SO_2 measurements from volcanic eruptions. For data fusion, satellite data will be most useful to improve concentration maps of air pollutants such as O_3, NO_2 and PM (using AOD).

References

Al-Saadi, J., J. Szykman, R.B. Pierce, C. Kittaka, D. Neil, D.A. Chu, L. Remer, L. Gumley, E. Prins, L. Weinstock, C. MacDonald, R. Wayland, F. Dimmick and J. Fishman, 2005. Improving national air quality forecasts with satellite aerosol observations, *Bull. Amer. Meteor. Soc*., 1249-1261.

Allen, D., K. Pickering, and M. Fox-Rabinovitz, 2004. Evaluation of pollutant outflow and CO sources during TRACE-P using model-calculated, aircraft-based, and Measurements of Pollution in the Troposphere (MOPITT)-derived CO concentrations, *J. Geophys. Res.*, **109**, D15S03, doi:10.1029/2003JD004250.

Bhave, P., 2004. PM Model Performance Workshop, 10 February 2004, Chapel Hill, NC; http://www.cleanairinfo.com/PMModelPerformanceWorkshop2004/presentations/Bhave_Postprocessing_model_output.ppt.

Boersma, D. Jacob, R. Park, R. Hudman, H. Eskes, P. Veefkind, R. van der A, P. Levelt, E. Brinksma, A. Perring, R. Cohen, T. Bertram, P. Wooldridge, E. Bucsela, J. Gleason, 2006. Validation and interpretation of OMI tropospheric NO_2 observations during INTEX-B, Aura Science Team Meeting, 11-15 September 2006, Boulder, CO; http://avdc.gsfc.nasa.gov/ST&WG/ST/ST_200609/Presentations/01Plenary_MondayMorning/Boersma.pdf.

Byun, D.W., C-K. Song, H.C. Kim, S. Kim, B. Cheng, 2006. Utilization of satellite measurements at various stages of atmospheric modeling to improve air quality simulations, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/Air_Quality_Remote_Sensing/Presentations/3.6.Byun.pdf.

Chance, K.V., J.P. Burrows, D. Perner, W. Schneider, 1997. Satellite measurements of atmospheric ozone profiles, including tropospheric ozone, from ultraviolet/visible measurements in the nadir geometry: a potential method to retrieve tropospheric ozone, *J. Quant. Spectrosc. Radiat. Transfer*, **57**, 467-476.

Chin, M., P. Ginoux, S. Kinne, O. Torres, B. N. Holben, B. N. Duncan, R. V. Martin, J. A. Logan, A. Higurashi, and T. Nakajima, 2002: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, *J. Atmos. Sciences*, **59**, 461-483.

Deeter, M.N., Calculation and Application of MOPITT Averaging Kernels, July 2002, available from http://www.eos.ucar.edu/mopitt/data/avg_krnls_app.pdf.

Edwards, D., P. DeCola, J. Fishman, D. Jacob, P. Bhartia, D. Diner, J. Burrows and M. Goldberg, 2006. Community input to the NRC decadal survey from the NCAR workshop on air quality remote sensing from space: defining an optimum observing strategy, National Center for Atmospheric Research, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/Air_Quality_Remote_Sensing/Reports/AQRSinputDS.pdf.

Engel-Cox, J., R. Hoff and A. Haymet, 2004. Recommendations on the use of satellite remote-sensing data for urban air quality, *J. Air Waste Manage. Assoc*., **54**, 1360-1371.

Enting, I.G., 2002. Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, New York.

Eskes, H. and P. Levelt, 2006. Use of PMI data for assimilation and validation of air quality model in Europe, Aura Science Team Meeting, 11-15 September 2006, Boulder, CO; http://avdc.gsfc.nasa.gov/ST&WG/ST/ST_200609/Presentations/WorkingGroups_MondayAfternoon/TropAirQuality_Pickering/EskesLevelt.pdf.

Falke, S.R., R.B. Husar and B.A. Schichtel, 2001. Fusion of SeaWIFS and TOMS satellite data with surface observations and topographic data during extreme aerosol events, *J. Air Waste Manage. Assoc*., **51**, 1579-1585.

Fishman, J., A.E. Wozniak, and J.K. Creilson, 2003. Global distribution of tropospheric ozone from satellite measurements using the empirically corrected tropospheric ozone residual technique: Identification of the regional aspects of air pollution, *Atmos. Chem. Phys*., **3**, 893-907.

Fishman, J., J.K. Creilson, A.E. Wozniak, and P.J. Crutzen, 2005. Interannual variability of stratospheric and tropospheric ozone determined from satellite measurements, *J. Geophys. Res*., **110**, D20306, doi:10.1029/2005JD005868.

Fishman, J., J.K. Creilson, A.E. Wozniak, D.W. Byun, M. Jang, J.J. Szykman, C. Kittaka, D.O. Neil, R.B. Pierce, 2006a. Challenges of Characterizing and Forecasting the Outbreak of Pollution Episodes: Defining Requirements for Spaceborne Air Quality Observations, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/Air_Quality_Remote_Sensing/Presentations/4.6.Fishman.pdf.

Fishman, J., R.B. Pierce, A.E. Wozniak, J.K. Creilson, J. Al-Saadi, K. Bowman, G. Osterman, A. Thompson, 2006b. Characterization and Validation of Ozone Pollution Episodes Using Aura Measurements and Assimilated Models, Aura Science Team Meeting, 11-15 September 2006, Boulder, CO; http://avdc.gsfc.nasa.gov/ST&WG/ST/ST_200609/Presentations/WorkingGroups_MondayAfternoon/TropAirQuality_Pickering/Fishman.ppt.

Gevaerd, R., S.R. Freitas and K. Longo, 2006. Numerical simulation of biomass burning fires and transportation during 1998 RORAIMA fires, Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 883-889.

Gilliland, A.B., R.L. Dennis, S.J. Roselle and T.E. Pierce, 2003. Seasonal NH_3 emission estimates for the eastern United States on ammonium wet concentrations and an inverse modeling approach, *J. Geophys. Res*., **108**, doi:10.1029/2002JD003063.

Goldberg, M. and S. Kondragunta, 2006. Current and Future Air Quality Applications of NOAA Operational Satellite Data, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/ Events/Meetings/Air_Quality_Remote_Sensing/Presentations/1.3.Goldberg.pdf.

Hakami, A., J.H. Seinfeld, Q. Li, et al., 2006. Satellite based inversion of NO_x emission using the adjoint of CMAQ, Fifth Annual CMAS Conference, 16-18 October 2006, Chapel Hill, NC; http://www.cmascenter.org/conference/2006/ppt/ session7/hakami.ppt.

Heald, C.L., D.J. Jacob, A.M. Fiore et al., 2003. "Asian outflow and trans-Pacific transport of carbon monoxide and ozone pollution: An integrated satellite, aircraft, and model perspective," *J. Geophys. Res*., **108**, doi:10.1029/2003JD003507.

Heald, C. L., D. J. Jacob, R. J. Park, B. Alexander, T. D. Fairlie, R. M. Yantosca, D. A. Chu, 2006. Transpacific transport of Asian anthropogenic aerosols and its impact on surface air quality in the United States, J. Geophys. Res., **111**, D14310, doi:10.1029/2005JD006847.

Henze, D.K. and J.H.. Seinfeld, 2006. Development of the adjoint of GEOS-Chem, *Atmos. Chem. Phys. Discuss*., **6**, 10591-10648.

Hodzic, A., Vautard, R., Chepfer, H., Goloub, P., Chazette, P., Menut, L., 2006. Evolution of aerosol optical thickness over Europe during the August 2003 heat wave as seen from POLDER data and CHIMERE model simulations, *Atmos. Chem. Phys.*, 6, 1853-1864.

Hoff, R.M., J.A. Engel-Cox, F. Dimmick, et al., 2006. 3D-AQS: a three dimensional air quality system, *Proc. SPIE,* **6299***, Remote Sensing of Aerosol and Chemical Gases, Model simulation/Assimilation and Application to Air Quality*, A. Chu, J. Szykman, S. Kondragunta, Eds.

Hollingsworth, A., Research Goals of GEMS: Global Earth-system Monitoring using Space and in-situ data, presentation at ENVISAT Summer School, ESRIN, August 2004; available from http://envisat.esa.int/envschool_2004/presentations/ th1_hollingsworth.pdf.

Japan Assoc. of Remote Sensing, 1996: Remote Sensing, http://www.profc.udec.cl/ ~gabriel/tutoriales/rsnote/contents.htm.

Jing, P., D. Cunnold, Y. Choi, and Y. Wang, 2006. Summertime tropospheric ozone columns from Aura OMI/MLS measurements versus regional model results over the United States, *Geophys. Res. Lett.*, **33**, L17817, doi:10.1029/2006GL026473.

Jones, D., M. Parrington, D. MacKenzie, et al., 2006. Chemical data assimilation of observations from TES, TES Science Team Meeting, 12-14 November 2006, Pasadena, CA.

Kondragunta, S., P. Lee, J. McQueen, C. Kittaka, A. Prados, P. ciren, I. Laszlo, B. Pierce, R. Hoff and J. Szykman, 2006. Air quality forecast verification using satellite data, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO.

Kopacz, M., D.J. Jacob, D. Henze, C.L. Heald, D.G. Streets, Q. Zhang, 2006. A comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of CO using satellite (MOPITT) measurements of CO columns, Eos Trans. AGU, 87(52), Fall Meet. Suppl., Abstract A31B-0875.

Kunhikrishnan, T., M. G. Lawrence, R. von Kuhlmann, M. O. Wenig, W. A. H. Asman, A. Richter, and J. P. Burrows, 2006. Regional NOx emission strength for the Indian subcontinent and the impact of emissions from India and neighboring countries on regional O3 chemistry, *J. Geophys. Res.*, **111**, D15301, doi:10.1029/2005JD006036.

Liu, Y., J. Sarnat, V. Kilaru, D. Jacob and P. Koutrakis, 2005a. Estimating ground-level $PM_{2.5}$ in the eastern United States using satellite remote sensing, *Environ. Sci. Technol.*, **39**, 3269-3278.

Liu, X., K. Chance, C.E. Sioris, R.J.D. Spurr, T.P. Kurosu, R.V. Martin and M.J. Newchurch, 2005b. Ozone profile and tropospheric ozone retrievals from the Global Ozone Monitoring Experiment: Algorithm description and validation, *J. Geophys. Res.*, **110**, doi:10.1029/2005JD006240.

Liu, X., et al., 2006. First directly retrieved global distribution of tropospheric column ozone from GOME: Comparison with the GEOS-CHEM model, *J. Geophys. Res.*, **111**, D02308, doi:10.1029/2005JD006564.

Lyon, J.G. and J. Szykman, 2006. EPA: Requirements for AQ Monitoring and Regulation, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/Air_Quality_Remote_Sensing/Presentations/1.5.Lyon.pdf.

Kleidman, R. G., N. T. O'Neill, L. A. Remer, Y. J. Kaufman, T. F. Eck, D. Tanré, O. Dubovik, and B. N. Holben, 2005, Comparison of Moderate Resolution Imaging Spectroradiometer (MODIS) and Aerosol Robotic Network (AERONET) remote-sensing retrievals of aerosol fine mode fraction over ocean, *J. Geophys. Res.*, 110, D22205, doi:10.1029/2005JD005760.

Matsui, T., S. M. Kreidenweis, R. A. Pielke Sr., B. Schichtel, H. Yu, M. Chin, D. A. Chu, and D. Niyogi, 2004. Regional comparison and assimilation of GOCART and MODIS aerosol optical depth across the eastern U.S., *Geophys. Res. Lett.*, **31**, L21101, doi:10.1029/2004GL021017.

Martin, R.V., D.J. Jacob, K. Chance, et al., 2003. Global inventory of nitrogen oxide emissions constrained by space-based observations of NO_2 columns, *J. Geophys. Res.*, **108**, D17, doi:10.1029/2003JD003453.

Martin, R.V., C.E. Sioris, K. Chance, et al., 2006. Evaluation of space-based constraints on global nitrogen oxide emissions with regional aircraft measurements over and downwind of eastern North America, *J. Geophys. Res.*, **111**, D15308, doi:10.1029/2005JD006680.

Menke, W., Geophysical data analysis: discrete inverse theory, Academic Press, ISBN: 0-12-490921-3, 1984.

Mendoza-Dominguez, A. and A.G. Russell, 2000. Iterative inverse modeling and direct sensitivity analysis of a photochemical air quality model, *Environ. Sci. Technol.*, **34**, 4974-4981.

Millet, D.B., D.J. Jacob, S. Turquety, et al., 2006. Formaldehyde distribution over North America: implications for satellite retrievals of formaldehyde columns and isoprene emission, *J. Geophys. Res.*, **111**, D24802, doi:1029/2005JD006853.

NASA Science Missions web-site, 2006. http://science.hq.nasa.gov/missions/ satellite_27.htm.

Neil, D., J. Szykman, J. Fishman, R. B. Pierce, J. Al-Saadi, and C. Kittaka, 2004. A good IDEA (Infusing satellite Data into Environmental Applications, 13th Conference on Satellite Meteorology and Oceanography, 20-23 September 2004, Norfolk, VA.

Osterman, G, 2005. Tropospheric Emission Spectrometer (TES) Validation Report, Jet Propulsion Laboratory, California Institute of Technology, August 2005, Pasadena, California,.

Palmer, P.I., D.S. Abbot, T.-M. Fu, et al., 2006. Quantifying the seasonal and interannual variability of North American isoprene emissions using satellite observations of the formaldehyde column, *J. Geophys. Res.*, **111**, D12315, doi:10.1029/2005JD006689

Pickering, K.E., J. Ziemke, C. Ahn, P.K. Bhartia, L. Marufu, R. Dickerson, R. Mathur, P. Davidson, 2006. Comparison of Eta/CMAQ and OMI –MLS Tropospheric Ozone, Aura Science Team Meeting, 11-15 September 2006, Boulder, CO; http://avdc.gsfc.nasa.gov/ST&WG/ST/ST_200609/Presentations/WorkingGroups _MondayAfternoon/TropAirQuality_Pickering/Pickering.ppt.

Pierce, R.B., 2006. Rapid synthesis of NRT satellite, airborne, surface and assimilated data sets during TEXAQS II to support the 2006 Texas Air Quality State Implementation Plan (SIP), TES Science Team Meeting, 12-14 November, 2006, Pasadena, CA.

Pison, I., L. Menut and N. Blond, 2006. Inverse modeling of emissions for local photooxidant pollution: testing a new methodology with kriging constraints, *Ann. Geophys*., **24**, 1523-1535.

Rao, S.T., Mathur, R., R. Pinder, J. Pleim, S. Roselle, D. Roy, 2006. Use of Remote Sensing Air Quality Information in Regional Scale Air Pollution Modeling: Current Use and Requirements, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/ Air_Quality_Remote_Sensing/Presentations/1.6.Mathur.pdf.

Remer, L., Y.J. Kaufman, D. Tanré, S. Mattoo, D.A. Chu, J.V. Martins, R.-R. Li, C. Ichoku, R.C. Levy, R.G. Kleidman, T.F. Eck, E. Vermote and B.N. Holben, 2005. The MODIS aerosol algorithm, products and validation, *J. Atmos. Sci*., **62**, 947-973.

Rodgers, C.D., 1976. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys.,* **14**, 609-624.

Rodgers, C.D., 1996. Information Content and Optimization of High-Spectral-Resolution Measurements, *Proceedings of SPIE*, **2830**, 136-147.

Rodgers, C.D., 2000. Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Press, Singapore.

Roustan, Y. and M. Bocquet, 2006. Inverse modelling for mercury over Europe, *Atmos. Chem. Phys*., **6**, 3085-3098.

Santee, M., 2006. Validation of the Aura Microwave Limb Sounder HNO_3 Measurements, Aura Science Team Meeting, 11-15 September 2006, Boulder, CO; http://avdc.gsfc.nasa.gov/ST&WG/ST/ST_200609/Presentations/WorkingGroups_MondayAfternoon/HNO3_Santee/Santee_MLS_HNO3Val.ppt.

Sarigiannis, D.A., A. Gotti, N.I. Sifakis, et al., 2006. Satellite-aided computational assessment of air quality and associated health effects, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/Air_Quality_Remote_Sensing/Presentations/3.10.Schafer.pdf.

Seigneur, C., 2004. Model evaluation – comparing model output to ambient data, PM Model Performance Workshop, 10 February 2004, Chapel Hill, NC;

http://www.cleanairinfo.com/PMModelPerformanceWorkshop2004/presentations/Seigneur_Model_Evaluation_Ambient.ppt.

Spichtinger, N., M. Wenig, P. James, T. Wagner, U. Platt and A. Stohl, 2001. Satellite detection of a continental-scale plume of nitrogen oxides from boreal forest fires, *Geophys. Res. Lett*., **28**, 4579-4582.

Szykman, J., 2006. TES Science Team Meeting, 12-14 November, 2006, Pasadena, CA.

Vautard, R., M. Beekman and L. Menut, 2000. Applications of adjoint modeling in atmospheric chemistry: sensitivity and inverse modeling, *Environ. Modelling Software*, **15**, 703-709.

Vijayaraghavan, K., Y. Zhang, H.E. Snell and J.-P. Huang, 2006a. Evaluation of a regional air pollution model with satellite measurements, Community Workshop on Air Quality Remote Sensing from Space, 21-23 February 2006, Boulder, CO; http://www.acd.ucar.edu/Events/Meetings/Air_Quality_Remote_Sensing/Presentations/Posters/krish.pdf.

Vijayaraghavan, K., J. Hu, Y. Zhang, X. Lui, K. Chance and H. Snell, 2006b. Evaluation of CMAQ using Satellite Remote Sensing Observations, Fifth Annual CMAS User's Conference, 16-18 October 2006, Chapel Hill, NC; http://www.cmascenter.org/conference/2006/abstracts/vijayaraghavan_session7.pdf.

Walker, K. A., C. E. Randall, C. R. Trepte, C. D. Boone, and P. F. Bernath, 2005, Initial validation comparisons for the Atmospheric Chemistry Experiment (ACE-FTS), *Geophys. Res. Lett.*, 32, L16S04, doi:10.1029/2005GL022388.

Wang, J. and S.A. Christopher, 2003. Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: implications for air quality studies, *Geophys. Res. Lett*., **30**, 2095.

Wittrock, F., A. Richter, H. Oetjen, J. P. Burrows, M. Kanakidou, S. Myriokefalitakis, R. Volkamer, S. Beirle, U. Platt, and T. Wagner, 2006. Simultaneous global observations of glyoxal and formaldehyde from space, *Geophys. Res. Lett.*, **33**, L16804, doi:10.1029/2006GL026310.

Yu, H., R. E. Dickinson, M. Chin, Y. J. Kaufman, B. N. Holben, I. V. Geogdzhayev, and M. I. Mishchenko, 2003. Annual cycle of global distributions of aerosol optical depth from integration of MODIS retrievals and GOCART model simulations, *J. Geophys. Res.*, **108**, 4128, doi:10.1029/2002JD002717.

Zhang, Y., H.E. Snell, K. Vijayaraghavan and M.Z. Jacobson, 2005. Evaluation of Regional PM Predictions with Satellite and Surface Measurements, AAAR 2005 Annual Conference, October 17-21, 2005, Austin, Texas.

Zhang et al., 2006. Continental outflow and intercontinental transport of ozone pollution as determined from O_3-CO correlations from TES, Aura Science Team Meeting, 11-15 September 2006, Boulder, CO; http://avdc.gsfc.nasa.gov/ST&WG/ST/ST_200609/ Presentations/01Plenary_MondayMorning/Zhang.ppt.

Ziemke, J. R., S. Chandra, B. N. Duncan, L. Froidevaux, P. K. Bhartia, P. F. Levelt, and J. W. Waters (2006), Tropospheric ozone determined from Aura OMI and MLS: Evaluation of measurements and comparison with the Global Modeling Initiative's Chemical Transport Model, *J. Geophys. Res.*, **111**, D19303, doi:10.1029/2006JD007089.

Appendix A: Acronyms

Organizations
CNES - Centre national d'études spatiales (French National Space Study Center)
CSA - Canadian Space Agency
ESA - European Space Agency
JAXA - Japan Aerospace Exploration Agency
NASA - National Aeronautics and Space Administration
NOAA - National Oceanic and Atmospheric Administration
SSC – Swedish Space Corporation
TEKES – Finland National Technological Agency

Satellites
ACE – Atmospheric Chemistry Experiment Satellite (also called SCISAT-1)
CALIPSO – Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
ENVISAT – Environmental Satellite
ERS – European Remote Sensing Satellites
GOES – Geostationary Operational Environmental Satellites
MetOp – Meteorological Operational Satellite Programme
NPOESS – National Polar-orbiting Operational Environmental Satellite System
POES – Polar-orbiting Operational Environmental Satellites
SCISAT-1 – Another name for the Atmospheric Chemistry Experiment (ACE) Satellite
TOMS EP – Total Ozone Mapping Spectrometer Earth Probe

Sensors
ACE-FTS - Atmospheric Chemistry Experiment Fourier Transform Spectrometer
AIRS – Atmospheric Infrared Sounder
ASCAT – Advanced Scatterometer
AVHRR – Advanced Very High Resolution Radiometer
CERES - Cloud's and the Earth's Radiant Energy System
CPR – Cloud Profiling Radar
GOME – Global Ozone Monitoring Experiment
HIRDLS – High Resolution Dynamics Limb Sounder
HIRS – High Resolution Infrared Radiation Sounder
IASI – Infrared Atmospheric Sounding Interferometer
IMG – Interferometric Monitor for Greenhouse Gases
ILAS – Improved Limb Atmospheric Spectrometer
LIDAR - Light Detection And Ranging
MAESTRO - Measurement of Aerosol Extinction in the Stratosphere and Troposphere Retrieved by Occultation
MIPAS - Michelson Interferometer for Passive Atmospheric Sounding
MISR – Multi-angle Imaging SpectroRadiometer
MLS – Microwave Limb Sounder
MODIS – Moderate Resolution Imaging Spectroradiometer
MOPITT – Measurements of Pollution in the Troposphere

OMI – Ozone Monitoring Instrument
OSIRIS - Optical Spectrometer and InfraRed Imager System
SBUV – Solar Backscatter Ultraviolet Radiometer
SCIAMACHY – Scanning Imaging Absorption Spectrometer for Atmospheric Chartography
SeaWiFS – Sea-viewing Wide Field-of-view Sensor
SMR - Sub-millimeter Microwave Radiometer
SOFIS – Solar Occultation FTS for Inclined-Orbit Satellite
TES – Tropospheric Emission Spectrometer
TOMS – Total Ozone Mapping Spectrometer
TOVS – TIROS Operational Vertical Sounder

Other Acronyms
GEO – Geostationary Earth Orbit
L-1 – Lagrangian Point 1
LEO – Low Earth Orbit
MEO – Mid-Earth Orbit