Integrating and Interpreting Aerosol Observations and Models within the PARAGON Framework

BY THOMAS P. ACKERMAN, AMY J. BRAVERMAN, DAVID J. DINER, THEODORE L. ANDERSON, RALPH A. KAHN, JOHN V. MARTONCHIK, JOYCE E. PENNER, PHILIP J. RASCH, BRUCE A. WIELICKI, AND BIN YU

A comprehensive, global aerosol picture can be generated using modern high-performance computing, geospatial statistics, assimilation, and data mining approaches.

The past decade has produced a remarkable increase in the amount of atmospheric data from observations and models, but these data are acquired or generated with nonuniform spatial and temporal sampling, scales, and coverage. The challenge for research into complex phenomena such as aerosol–climate interaction is to combine these disparate data into an integrated whole (Kahn and Braverman 1999; Huang et al. 2002). The ultimate goal is to establish a complete dataset that will effectively confront and constrain ever more realistic global three-dimensional models.

A first step in attaining this objective is to produce a measurement-based description of global tropospheric aerosols. Protocols are needed to organize this vast body of knowledge. For example, in considering aerosol optical depth (AOD), one of the most fundamental parameters of climate and air-quality significance, the Science Advisory Group of the Global Atmosphere Watch program (Baltensperger et al. 2003) comments "There is a need for a common strategy to merge the various network observations into a global dataset . . . and to develop with satellite agencies a system for integrating global AOD observations."

Achieving progress on the aerosol–climate problem requires applying this strategy to other fundamental parameters as well. Producing these integrated datasets will involve existing and new remote-sensing technologies, the expansion of observing systems in order to describe aerosol properties in increasing detail, and the use of new analysis techniques. This evolving level of sophistication is one element of the Progressive Aerosol Retrieval and Assimilation Global Observing Network (PARAGON) strategy (Diner et al. 2004a). This progressive approach will make it possible to tackle problems of increasing difficulty, such as the incorporation of cloud observations, process models, and parameterizations needed for understanding indirect effects, and the measurement of boundary-layer chemistry needed for understanding aerosol and gas interactions and their effects on air quality. Assembling the array of data is only one step;

AFFILIATIONS: ACKERMAN—Pacific Northwest National Laboratory, Richland, Washington; BRAVERMAN, DINER, KAHN, AND MARTONCHIK—Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California; ANDERSON—University of Washington, Seattle, Washington; PENNER—University of Michigan, Ann Arbor, Michigan; RASCH—National Center for Atmospheric Research, Boulder, Colorado; WIELICKI—NASA Langley Research Center, Hampton, Virginia; YU—University of California, Berkeley, Berkeley, California
CORRESPONDING AUTHOR: David J. Diner, JPL Mail Stop 169-237, 4800 Oak Grove Drive, Pasadena, CA 91109
E-mail: djd@jord.jpl.nasa.gov
DOI: 10.1175/BAMS-85-10-1523

In final form 28 July 2004

it is also necessary to establish interoperability, by which we mean the ability to process and exchange data from multiple platforms collected at disparate spatial and temporal scales. Inherent in this concept is the ability to interpret these disparate data correctly, and to take into account variability not only in sampling and coverage, but also in accuracy. Synthesis of the inputs will be required for rigorous evaluation and research. Algorithm development and validation will benefit from readily accessible aerosol information, and the synergistic use of multiple inputs can be explored. Data mining will be needed to sift through the data in an efficient, effective way. Geospatial statistics and data assimilation provide mechanisms for synthesizing the data.

While statistical and assimilation models serve a critically important role as information interpolators and integrators, scientific understanding requires the production of validated process models that incorporate our complete understanding of the relevant physics and chemistry. From a societal perspective, the ultimate goal is a global climate model (GCM) that contains the essential processes, but in a perhaps simpler and computationally less demanding framework than the process models. Assimilation acts as a large flywheel that keeps the model tuned to the observed atmospheric conditions. Climate models cannot have such a flywheel, but must maintain an appropriate climatology by the accurate specification of atmospheric (and oceanic) energetics through model equations. The very ability of the models to simulate current and future climate is intimately tied to these specifications. Generation of such a GCM is beyond the current scope of PARAGON; however, to make optimum use of both the physical insights and the data record generated, we suggest active coupling to the climate modeling community through the continued development and validation of chemical transport models (CTMs) that either use assimilated meteorological fields or are driven by meteorological fields from a GCM.

DATA INTEROPERABILITY AND COMPUTATIONAL INFRASTRUCTURE. We can capitalize on modern information technology to assemble a worldwide aerosol data system in order to promote the widespread exchange and use of data, and to enhance computational power for global modeling. The first step in making data widely accessible to the research community is to create an organizational infrastructure. Currently, the capability to assemble most of the existing aerosol data into a unified, interoperable whole is lacking. Establishing a distributed aerosol science information system will rectify this problem. The NASA Sensor Intercomparison and Merger for Biological and Interdisciplinary Oceanic Studies program (R. A. Barnes et al. 2001, personal communication; Fargion et al. 2003; Werdell et al. 2003) is an example of how such objectives have been met in ocean biology. NASA's distributed data archives, and initiatives within the Earth Science Enterprise (McDonald et al. 2002) are addressing ways of enhancing data access and exchange. The World Meteorological Organization World Data Center for Aerosols archives aerosol data from certain surface networks in a standardized format (Baltensperger et al. 2003). These elements are necessary parts of PARAGON; however, we need a broader approach to promote integrative and collaborative use of the diverse data.

In addition to making possible the global sharing of data, modern "grid computing" can help facilitate collaborative research. Grid architectures support distributed computing using high-speed networks, and can provide several advantages for PARAGON, such as the integration and testing of physics and chemistry modules developed at different institutions, fast data exchange, making it unnecessary to download large volumes of data from their archives, and aggregation of computing resources to create a virtual supercomputer. As a means of improving global model resolution, aggregation may be an alternative or complement to subgrid-scale parameterizations.

PARAGON can capitalize on advancing grid, massively parallel, and other high-performance computing and cyberinfrastructure initiatives (Foster 2002) to create an aerosol "virtual observatory" to enhance climate and environmental research. Precedents in other disciplines include NASA's Information Power Grid (Johnston et al. 1999; Hinke and Novotny 2000) and the National Science Foundation (NSF)'s National Virtual Observatory (NVO). The latter combines more than 100 TB of data from 50 ground- and space-based telescopes, with standards for services and data established by the astronomical community (Freeman 2003). NVO is part of a larger federation, the World-Wide Telescope (Szalay and Gray 2001), which is contending with the same technical and collaborative issues that we address here. NSF also supports physics and astronomy research through the international Virtual Data Grid Laboratory, which integrates heterogeneous computing, storage, and networking resources in the United States, Europe, Asia, and South America, and the Grid Physics Network, a collaboration of information technology researchers and experimental physicists. In high-energy and nuclear physics, the Department of Energy (DOE) (DOE 2001) supports a national collaboration known as the Particle Physics Data Grid. Research within the DOE High Performance Networks program (DOE 2002a,b) involves

- software services ("middleware") that mediate data access and exchange between archives and users to facilitate large-scale scientific applications;
- distributed computing, using advanced networking for terabits-per-second throughput;
- artificial intelligence and machine learning for better computational efficiency power and to guide large-scale analyses; and
- real-time visualizations of complex scientific simulation results in multiple remote locations.

DOE, NASA, and NSF plan to continue investing in advanced computing. PARAGON would be poised to take advantage of these advances, whether for increased computational throughput or for exciting new opportunities in data access and mining.

DATA INTEGRATION. Integrating observational and model data will require new methods of handling diverse spatial and temporal sampling, resolution, and coverage. These methods will include data- and model-driven approaches.

Geospatial statistics. Data-driven approaches to generating a global aerosol picture aim to integrate constraints from multiple surface and in situ measurements into a detailed description of regional atmospheric structure. The regional description can then be extrapolated to larger spatial and temporal domains based on satellite and surface network observations. The first steps in the direction of merging field data from multiple platforms into coherent "environmental snapshots" are being taken with data from recent field campaigns (Russell et al. 1999, 2002; Russell and Heintzenberg 2000; Clarke et al. 2002; Satheesh et al. 2002; Magi et al. 2004; Reid et al. 2003; Kahn et al. 2004a). Subsequent steps include 1) using such snapshots to assess satellite observations; 2) using these and other observations to characterize the three-dimensional spatial and temporal distributions of aerosol air masses; and 3) combining satellite and in situ measurements to assess aerosol variations within a specific air mass (e.g., Clarke et al. 2001; Rasch et al. 2001; Collins et al. 2001). Ascertaining the quality of individual datasets is a vital prerequisite to such efforts.

GEOSPATIAL STATISTICS

The most general statistical model for the space–time distribution of aerosols is the function $\mathbf{Z}(\boldsymbol{s},t)$, where $\boldsymbol{s}$ is a continuously valued three-dimensional spatial index [e.g., $\boldsymbol{s} = (x, y, z)$ with x latitude, y longitude, and z height] and t is time. Here, $\mathbf{Z}$ may be a vector with a first component giving the extinction coefficient, and the remaining components describing particle properties, for example; $\mathbf{Z}$ exists everywhere and at all times, but we only have access to observations $\mathbf{Y}_i\ (x_n, y_n, z_n, t_n)$, the measurement collected by instrument i at a subset of locations and times. As in the case of chemical samplers, $\mathbf{Y}$ may be acquired at a point location and aggregated over t, or it may be an instantaneously measured quantity acquired by a satellite and aggregated over x, y, and z. The spatial and temporal grids for instrument i may be different from that of instrument j, and the instruments most certainly have different measurement error characteristics. Making inferences about $\mathbf{Z}$, or aggregates of $\mathbf{Z}$, requires assumptions about the statistical relationships among the $\mathbf{Z}$'s, among the $\mathbf{Y}$'s, and between the $\mathbf{Z}$'s and $\mathbf{Y}$'s. Inference proceeds via the usual principles of Bayesian statistics (Berger 1985), which combines all this information to make optimal estimates of the unknown quantities.

Comprehensive texts on spatial statistics include Cressie (1993) and Chiles and Delfiner (1999). Cressie (1993) explicitly addresses geostatistical data obtained at irregular locations and lattice data obtained on a grid. Data from field experiments are of the first type, whereas satellite data are of the latter. Various geostatistical approaches to modeling space–time processes are also described in Kyriakidis and Journel (1999). Spatial statistical analysis of remote sensing data is closely related to other forms of digital image analysis, especially that of medical image data (NRC 1991). A fruitful approach to finding structure in spatial data is Bayesian hierarchical modeling (Wikle et al. 1998). Wikle et al. (2001) used Bayesian hierarchical models to combine satellite data at high resolution with model-generated data at coarse resolution to estimate tropical ocean surface winds. Huang et al. (2002) employed a Bayesian hierarchical model in the development of a multiresolution Kalman filter for producing statistically optimal estimates of ozone from Total Ozone Mapping Spectrometer data at multiple nested resolutions.

Various aspects of all these efforts apply to the data fusion problem posed by PARAGON, but none provides a complete solution. For that, we need to revisit the theoretical statistics underlying inferences required by PARAGON, and rebuild the model framework specifically for this context. The challenge is to formalize statistical relationships and assumptions by incorporating physics, chemistry, and other domain knowledge into the framework. For example, it is impossible to make inferences at every $(\boldsymbol{s}, t)$ from any one data source, and different measurements may represent averages over different extents of $\boldsymbol{s}$ and durations of t. Therefore, the integration approach will need to make use of spatiotemporal aggregation on varying scales.

A variety of approaches can deal with combining data from multiple sources. Engineers, image analysts, statisticians, and geoscientists have all attacked the problem. To realize the synergy among the diverse aerosol data sources, PARAGON needs a comprehensive, theoretically sound framework for combining information. There are both theoretical and applied studies in spatial statistics that are directly relevant. The sidebar "Geospatial Statistics" on the previous page provides further details.

Chemical transport and aerosol data assimilation models. In recent years, the distinction between observations and model output has become increasingly blurred by data assimilation. The weather forecasting community pioneered this approach to provide optimal initial conditions for forecast models. Their efforts were driven by the recognition that no observing system could ever hope to measure the properties of the global atmosphere at the requisite time and space scales for forecast initialization, and that interpolation of scattered measurements simply could not adequately represent atmospheric structure. Data assimilation techniques started with routine measurements of state variables from radiosondes and surface sites, but have now grown to include a bewildering array of measurements of state variables and radiances. It is widely recognized that assimilation models provide the best possible assessment of the atmospheric dynamic and thermodynamic state for comparison with climate model simulations (Trenberth et al. 2002). Due to the complexity of cloud properties and parameterizations, and a lack of measurements having good sampling on the right space–time scales, assimilation techniques have had little impact on cloud representations, but this is beginning to change. For example, the NASA Global Modeling and Assimilation Office is already applying assimilation techniques to cloud and precipitation observations. This same approach can and is being used to address the aerosol problem.

Aerosol modeling on the global scale has evolved in several directions. CTMs typically include the most advanced representations of the chemistry of the atmosphere and the aerosol physical state. These models are run using prescribed meteorological fields obtained from weather forecasting models in real time, meteorological assimilations, or GCMs. CTMs in this configuration are very useful for studying aerosol processes, but cannot be used to study aerosol impacts on climate because the meteorology is uncoupled with the chemistry. At the same time, highly simplified representations of aerosol physics were developed for GCMs in order to study aerosol–climate interactions. These models provide useful first-order estimates of aerosol–climate forcings but generally lack an adequate description of all of the competing direct and indirect effects. More recently, the field has been moving toward merging CTMs with complex aerosol physics representations with GCMs (or adding complex descriptions to existing GCMs). These coupled models can be used to study the effect of aerosols on the meteorological state, at least for short (several-year) simulations. Typically, advanced process representations are first developed in the CTM framework and then moved to assimilation models and, in simplified form, to GCMs.

CTMs have two different applications within PARAGON. CTM output can be directly compared to observations from both point measurements and satellites. If it is driven by meteorological fields from weather forecasting models or a meteorological assimilation, then CTM output can be compared to measurements for a particular time period. If they are driven by the meteorology from a GCM, then statistical comparisons with aggregates over longer time periods are appropriate. The differences between CTM predictions and observations often have led to significant improvements in the CTM representation of aerosol physics and chemistry and/or in source emissions specifications. They have also served to guide observational strategies, helped to pinpoint observational errors, and can be used to guide integrated strategies that will be of most benefit in decreasing uncertainties related to aerosol effects on climate.

A physically detailed CTM is also the fundamental tool for aerosol data assimilation. It includes a complete specification of aerosol sources by type and location, physical equations to describe modification by aging and humidity changes, and sink terms due to both wet and dry deposition. If coupled with a sophisticated radiative-transfer scheme to compute radiances at the top of the atmosphere and at the surface, the CTM then integrates datasets into a single, best-estimate aerosol product, just as current data-assimilation models do for meteorological fields.

The weather forecasting experience provides two important lessons for PARAGON. First, any aerosol assimilation model will be continually upgraded as understanding of aerosol physics grows, and it will have to be continually reconfigured to assimilate new data sources as they become available. Second, in order to produce assimilated datasets of consistent quality, the entire aerosol data record will have to be reprocessed periodically with the latest model version. Reprocessing is expensive but absolutely necessary in order to produce long-term records of known quality and uniformity.

The distinction between the spatial statistical approach and assimilation is becoming more and more blurred. In fact, for some time data-assimilation techniques based on Bayesian statistical models have been providing a theoretical foundation for combining physical models and observations. For example, Lorenc (1986) studied the intimate connections between a variety of forms of optimal estimation used for data assimilation in the context of weather forecasting. The underlying rationale for calculation of optimal estimates is the same whether the data are purely observational, model generated, or a combination of the two. PARAGON seeks to bring together and capitalize on recent advances in statistics, aerosol data assimilation, and transport modeling to produce a comprehensive representation of atmospheric aerosols.

MULTISENSOR ALGORITHM DEVELOPMENT. One motivation for establishing a data interoperability framework is to stimulate algorithm development and validation in ways that would not otherwise be possible, and to facilitate the complementary use of different data types. Following are two examples of the potential synergies of combining data from multiple sensors as part of the retrieval process.

Synergistic use of satellite- and surface-based multiangle data. Currently, surface-based measurements of directly and diffusely transmitted radiances are used primarily to validate satellite data. However, because retrievals of aerosol properties from such remote-sensing data (like those from satellites) are underconstrained, up- and down-looking data could be combined (where available) to more tightly constrain the description of the aerosol column and to account adequately for the effects of surface reflectance (Dubovik et al. 2003). For example, consider the synergy between multiangle Aerosol Robotic Network (AERONET) sky radiance measurements and multiangle satellite data from the Multiangle Imaging SpectroRadiometer (MISR) on *Terra.*

AERONET sun/sky photometer measurements are normally made in both the principal and almucantar planes (i.e., planes containing the sun's position perpendicular to the surface and parallel to the horizon, respectively). Sky radiance is, thus, sampled over a range of scattering angles, predominately in the forward-scattering direction. Optical depth can be determined in a straightforward manner from direct solar measurements. The sky radiance measurements consist of both a dominant radiance field, which diffusely scatters only within the atmosphere, and a field that also includes reflections from the surface. The current technique for retrieving aerosol optical properties and size distributions (Dubovik and King 2000; Nakajima et al. 1983) assumes a Lambertian surface (i.e., its radiance is the same from all viewing angles) with a predetermined albedo representative of the surface type at each site (Dubovik et al. 2000). For a dark surface, these assumptions do not appreciably affect the outcome of the sky radiance analysis, but for a bright surface, its actual reflectance properties can matter greatly.

To constrain ground-based sky photometer retrievals more effectively and account for the effects of surface reflectance, it is highly desirable to include simultaneous downward-looking multiangle measurements in the analysis (Dubovik et al. 2003). Although this is not currently being done for AERONET, these additional measurements offer two important advantages. First, the downward-looking multiangle measurements allow the retrieval of the actual surface directional reflectance properties, which can then be used in the sky photometer aerosol retrieval. Second, the downward-looking observations measure predominantly backscattered radiation, thus extending the radiance scattering angle range provided by the ground-based observations. This is important because the determination of aerosol particle shape and absorption is very sensitive to the angular form of the aerosol particle scattering pattern (the phase function) in some backscatter directions. Furthermore, a more complete characterization of the phase function with scattering angle results in a better determination of the aerosol single-scattering albedo, because it is obtained by integrating the phase function over angle. The complementarity in scattering angle coverage between a downward-looking multiangle instrument (e.g., MISR) and an upward-looking multiangle scanner (e.g., a CIMEL Electronique radiometer) is illustrated in Fig. 1. Aside from a geometric factor, the phase function is directly proportional to the single-scattered radiance field. Because a multiple-scattered radiance field is much less sensitive to the phase function shape, a robust retrieval of single-scattering albedo and particle shape via the phase function variability requires that the single-scattered radiance represents a significant component of the measured radiance field.

Combination of passive and active remote sensing. A second fruitful area of data synergy is the combination of passive and active remote sensing (Léon et al. 2003; Kaufman et al. 2003)—for example, when sat-

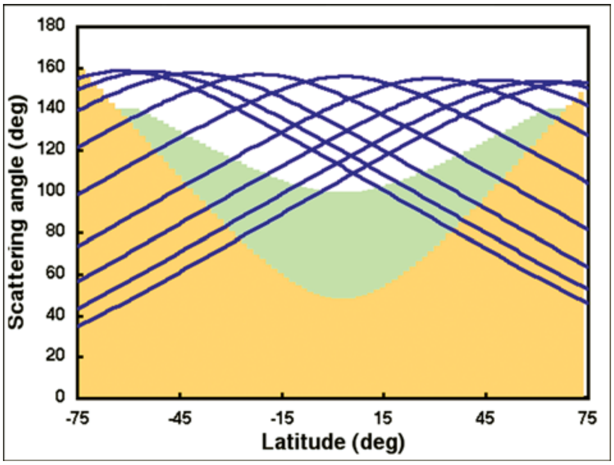

FIG. 1. Complementarity of space-based (blue: MISR nine cameras) and surface-based (tan: CIMEL principal plane; green: CIMEL almucantar) scattering angle coverage. Values for the center of the MISR swath are shown. The *Terra* spacecraft is in a sun-synchronous 10:30 A.M. equator-crossing orbit, and the scattering angles correspond to the illumination geometry obtained at the fall or spring equinox. A wider swath or an orbit closer to local noon could routinely achieve scattering angles that are even closer to 180°.

ellite spectroradiometry and laser ranging are used to derive vertically resolved aerosol optical depths and size fractions. Aerosol optical depth is one of the parameters required to assess direct radiative forcing, and great progress has been made in globally mapping this quantity from space (though further improvement is needed). One potential input that may help determine anthropogenic aerosol abundance is the fraction of aerosols that are less than about 1 μm in radius. This can be retrieved from space via measurements of the wavelength dependence of optical depth (Tanré et al. 2001; Kaufman et al. 2002) and/or from multiangle observations (Kahn et al. 1998; Martonchik et al. 2002), though the accuracy is not yet well understood. For the Moderate Resolution Imaging Spectroradiometer (MODIS), validation studies indicate that the wavelength dependence of optical depth for fine-mode-dominated aerosol is underestimated (Remer et al. 2002). Moreover, the retrieved product exhibits obvious discontinuities across many coastlines. These discontinuities could be real (e.g., coarse sea salt over ocean but not over land), or they could be an artifact of the switch between the MODIS land and ocean algorithms.

The afternoon MODIS sensor aboard *Aqua* will soon benefit from simultaneous lidar data from the Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite, and the latter will enhance the MODIS products in a number of ways. First, the lidar can identify cases of subvisible cirrus and allow tests of whether these cases are ever misinterpreted as coarse-mode aerosol by MODIS. Second, lidar retrievals of optical depth, unlike passive techniques, do not depend on surface reflectivity. Although its optical depth retrievals might be fundamentally less accurate than from MODIS, the lidar is likely to be much better at assessing the variability of optical depth in coastal regions. In addition, lidar can determine aerosol-layer elevation, which will be helpful in evaluating the possibility that sea-salt aerosol can cause land–ocean discontinuities. Equally important is that the MODIS measurement of optical depth will enhance the lidar product from CALIPSO. Aerosol extinction retrieval from a simple backscatter lidar is unconstrained without additional information. A priori knowledge of the extinction-to-backscatter ratio (e.g., based on climatology) is helpful, but, as shown by Stephens et al. (2001), the addition of aerosol optical depth information is required in most cases if the retrieval is to be constrained to better than about ±40%.

These two brief examples illustrate potential gains from data synergy. The idea of observational synergy was one of the core concepts in the development of the NASA Earth Observing System (Butler et al. 1984). It is also a key element of the ground-based sites operated by the DOE (Ackerman and Stokes 2003). Exploiting this synergy, however, requires dedicated research efforts, which we see as an integral part of PARAGON.

DATA SUMMARIZATION AND MINING.

Because integrated datasets are multidimensional (including three-dimensional space and time), multivariate (including aerosol physical, chemical, and radiative parameters), and massive (e.g., satellite-imaging radiometers generate many gigabytes per day), it is important to construct summaries in a way that preserves the information in the measurements while reducing size and complexity for science users. The NRC (1991) notes "a relatively new role for spatial statistics . . . is to synthesize and reduce large volumes of data into manageable pieces of information."

In light of the diversity, complexity, and large quantities of aerosol data at many different spatial and temporal resolutions, efficient interpretation will require modern statistical and data-mining techniques. The goal is to understand spatiotemporal, inter-resolution relationships between, for example, the so-called extensive properties of aerosols (i.e., quantities that are functions of the particle concentration); intensive quantities, which are typically thought of as being concentration independent (e.g., particle optical properties) (Ogren 1995); and other quantities of importance, such as relative humidity. More formally, we seek to model the joint (multivariate) probability distribution of the data in defined spatiotemporal regions and then model the relationships between these distributions as a function of space, time, and resolution. There are many techniques for doing this (e.g., Scott 1992), and the challenge here is to strike a balance between theoretical elegance and the practical requirements of working with large datasets.

One approach to the problem of estimating multivariate distributions from large datasets is to produce quantized approximations of the original data (Braverman and Di Girolamo 2002). This enhances the traditional approach of creating global, gridded maps of mean quantities for purposes of summarizing or studying global patterns. The multivariate mean is the best single statistic for describing a joint distribution; however, how many (and which) statistics are used should be decided on theoretically justifiable considerations. A well-developed framework for this problem already exists in signal processing and information theory (Chou et al. 1989; Gersho and Gray 1991; Shannon 1948) and statistics (MacQueen 1967; Zador 1964; Pollard 1982). We suggest using it as the basis for estimating distributions of aerosol properties and other related quantities.

We also envision using physical models to provide training data for supervised learning algorithms, which can extrapolate local results to larger scales. Regression is the simplest form of supervised learning, but many more sophisticated methods exist, including neural networks, support vector machines, and classification and regression trees (Hastie et al. 2001 provide a comprehensive review). Supervised learning techniques have been used successfully in the study of aerosol distributions (Stroud et al. 2000; Konovalov 2003) and the effect of tropospheric aerosols on global temperature (Walter et al. 1998). Garay et. al (2003) report the successful application of active learning based on support vector machine classifiers for cloud detection. It is logical to apply a similar methodology for the categorization of aerosol air masses. There is reason to believe that this may be a useful approach because seasonal and spatial distributions of aerosol amount and type exhibit patterns that repeat from year to year. Also, the microphysical properties of broad classes of mineral dust, sea salt, biomass burning, pollution, and background particles seem to vary little for many of the largest sources (e.g., Clarke and Kapustin 2002). Further testing of this hypothesis is a research topic for the data-mining component of PARAGON.

SYSTEMATIC APPROACHES TO MODEL EVALUATION. Determining the role of aerosols in past and future climate change ultimately requires the use of fully coupled climate and chemistry models, and evaluation of these models is required in order to trust their results. Verifying GCMs entails searching for climate change "fingerprints" and establishing relationships between GCM output and the data acquired over relatively short time frames (Goody et al. 1998). We confine our discussion to the need for critical examination of CTMs (for which the longest relevant time scale is on the order of months) and radiative models that relate particle physical and chemical attributes to optical properties and incoming and outgoing radiation.

Currently, discrepancies between models and measurements are among the main difficulties confronting the aerosol research community. For example, in models the average forcing over the oceans by aerosols is about 2 W m^{-2} smaller than that implied by measurements (Penner et al. 2002). Resolving such discrepancies requires establishing a data–model comparison strategy that can isolate the underlying cause(s). Methodical model and data evaluations (e.g., Balkanski et al. 1993; Benkovitz and Schwartz 1997; Rasch et al. 2000; Kinne 2003; Schulz and Kinne 2003) have progressed significantly. These evaluations are necessary to identify specific inconsistencies in either the observational data or the models.

Deficiencies might arise from the absence of key observables, measurement or retrieval errors, and/or misrepresentations of particular physical processes. Even "observations" require some degree of inference. For example, many assumptions about particle properties underlie the retrieval of optical depth from a remote sensor. Here we address how data integration, interpretation, summarization, and mining approaches discussed earlier can provide important tools for rigorously testing model assumptions.

Seinfeld et al. (2004) discuss the collection of processes involved in the life cycle of aerosol layers. Unlike clouds, where droplets or crystals grow and evaporate within hours or a day, aerosols can reside in the atmosphere from days to weeks. This suggests a strategy in which air masses are sorted by source type so that the aerosol processing along trajectories can be cataloged and compared. Such a catalog may be regarded as a statistical ensemble within which the evolution of observed and modeled particle properties can be evaluated. Spatiotemporal data mining (e.g., Vucetic and Obradovic 2000a,b), both supervised and unsupervised, could provide a means of rigorously partitioning massive observational and model-generated datasets into comparable subsets. The set of synoptic conditions need not be contiguous; for example, a single ensemble for the study of model performance for biomass burning might encompass areas of Indonesia, Africa, and South America during different months of different years.

We then need to consider which attributes of relevant variables need to be compared. Given that many variables of interest do not conform to Gaussian distributions (e.g., size distributions are multimodal and height distributions are multilayered), simple means and low-order moments such as variances are likely to be inadequate. Moreover, visual inspection of monthly means, for example, may be a necessary but wholly insufficient method of assessing the agreement between observations and models. Rigorous tests of models require capturing joint relationships between different variables, and the inclusion of covariances is adequate only when these relationships are linear. Writing on behalf of the Program for Climate Model Diagnostics and Intercomparisons, Gates (1992) observes that

> in most cases validation extends only to the average values of variables that are of particular interest. . . . A somewhat more systematic approach to model validation is needed if we are to identify and progressively reduce model errors. [A] comprehensive validation program of atmospheric models would include . . . the mean, variability, and complete frequency distribution (and corresponding error estimates) for the full suite of simulated variables and the associated fluxes, processes, and phenomena. . . . [R]eliable identification of model errors may require innovative techniques for time series analysis and pattern recognition.

Our earlier discussion on the summarization of observational data can also be applied to the outputs of models. Observed and modeled data distributions should be compared using formal hypothesis tests that establish whether or not both datasets derive from the same overall population. The fundamental presumption is that if the model produces results "close" to what is observed, then the model is "right." Of course, the challenge lies in defining the meaning of close. Statistical hypothesis testing provides a useful way of formalizing ideas about whether two datasets are close or a "match." We echo the words of Goody et al. (2002), written in the context of climate monitoring but appropriate for this discussion:

> [S]ome differences will be due to inadequacies in the model physics and dynamics. It is these latter differences that must be evaluated and refined by comparing model and observed statistics. Although this is a difficult task, we have no alternative but to undertake it . . . [S]uitable observational systems and assimilation techniques . . . to support tests of model predictions . . . would undoubtedly emerge if this line of investigation were supported by federal agencies, which is not the case at the present time.

Much of the model evaluation research to date has focused on limited datasets in time and space and on simple statistical tests, such as comparisons between models and observations of regional and monthly mean aerosol optical depths. While such comparisons are important, achieving climate-forcing accuracies on the order of a few watts per square meter requires integrated datasets, providing a much more comprehensive set of aerosol properties and sophisticated statistical tests that effectively describe model accuracy in a multidimensional statistical space. PARAGON seeks to provide the framework that supports and organizes research toward these ends.

CONCLUSIONS. Achieving the PARAGON vision requires establishing multidisciplinary, interagency, and international partnerships to advance fundamental scientific understanding of aerosol patterns and processes. Some portion of this vision will be fulfilled by the creation of shared observational systems and networks (Kahn et al. 2004b; Diner et al. 2004b). Another portion, however, must be fulfilled by combining many disparate observational and modeling components into an integrated whole. This second objective involves a complex interweaving of aerosol observations and geospatial statistics, assimilation and chemical transport modeling, information technology, and data-mining research. The resulting four-dimensional picture will provide the means to test and validate models of aerosol–climate interactions. Reaching this goal will not be easy; it entails planning and coordination that spans funding agencies and research institutions, program managers, and scientists. Without such a coordinated effort, however, the pace of progress will be slow as individual researchers struggle with disparate datasets, idiosyncratic models, and limited resources.

In considering the totality of this approach, we are led to several important conclusions. First, the full sweep of these integration activities is absolutely critical if we are to establish the quantitative role of aerosols in the earth system and achieve an understanding of how complex aerosol processes impact climate change and air quality. Second, none of these activities is beyond our reach—the scientific community has the requisite knowledge and experience to undertake them. Third, each of these activities has a research component. While we can see the broad outline of the path that needs to be taken, there are many issues that need to be resolved along the way. Finally, we can only reach our goal of understanding aerosol–climate interactions by a sustained research effort. We created the PARAGON vision as a blueprint for such a program.

ACKNOWLEDGMENTS. Support from the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration, the National Science Foundation, and the U.S. Department of Energy is gratefully acknowledged. The research of A. J. Braverman, D. J. Diner, R. A. Kahn, and J. V. Martonchik was conducted at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with NASA. The research of T. P. Ackerman was conducted at the Pacific Northwest National Laboratory under contract with JPL. T. L. Anderson acknowledges support from the National Science Foundation (Grant ATM-0138250). We thank E. Hansen of JPL for technical advice on information technologies.

REFERENCES

Ackerman, T. P., and G. M. Stokes, 2003: The Atmospheric Radiation Measurement Program. *Phys. Today,* **56,** 38–44.

Balkanski, Y., D. Jacob, G. M. Gardner, W. M. Graustein, and K. K. Turekian, 1993: Transport and residence times of continental aerosols inferred from a global 3-dimensional simulation of ^{210}Pb. *J. Geophys. Res.,* **98,** 20 573–20 586.

Baltensperger, U., and Coauthors, 2003: WMO/GAW aerosol measurement procedures: Guidelines and recommendations. WMO Tech. Doc. TD 1178, 72 pp.

Benkovitz, C. M., and S. E. Schwartz, 1997: Evaluation of modeled sulfate and SO_2 over North America and Europe for four seasonal months in 1986–87. *J. Geophys. Res.,* **102,** 25 305–25 338.

Berger, J. O., 1985: *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, 617 pp.

Braverman, A., and L. Di Girolamo, 2002: MISR Level 3 data products: A new approach. *IEEE Trans. Geosci. Remote Sens.,* **40,** 1626–1636.

Butler, D. M., and Coauthors, 1984: Earth Observing System science and mission requirements working group report. Vol. I, NASA Tech. Memo. 86129, 58 pp.

Chiles, J.-P., and P. Delfiner, 1999: *Geostatistics: Modeling Spatial Uncertainty.* Wiley, 720 pp.

Chou, P. A., T. Lookabaugh, and R. M. Gray, 1989: Entropy constrained vector quantization. *IEEE Trans. Acoust. Speech Signal Process.,* **37,** 31–42.

Clarke, A. D., and V. N. Kapustin, 2002: A Pacific aerosol survey. Part I: A decade of data on particle production, transport, evolution, and mixing in the troposphere. *J. Atmos. Sci.,* **59,** 363–382.

——, W. D. Collins, P. J. Rasch, V. Kapustin, K. Moore, and S. Howell, 2001: Dust and pollution transport on global scales: Aerosol measurement and model predictions. *J. Geophys. Res.,* **106,** 32 555–32 569.

——, and Coauthors, 2002: The INDOEX aerosol: A comparison and summary of microphysical, chemical, and optical properties observed from land, ship, and aircraft. *J. Geophys. Res.,* **107,** 8033, doi:10.1029/2001JD000572.

Collins, W. D., P. J. Rasch, B. E. Eaton, B. Khattatov, J.-F. Lamarque, and C. S. Zender, 2001: Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: Methodology for INDOEX. *J. Geophys. Res.,* **106,** 7313–7336.

Cressie, N. A. C., 1993: *Statistics for Spatial Data.* Wiley Series in Probability and Mathematical Statistics, Vol. XXII, Wiley, 900 pp.

Diner, D. J., and Coauthors, 2004a: PARAGON: An integrated approach for characterizing aerosol climate impacts and environmental interactions. *Bull. Amer. Meteor. Soc.,* **85,** 1491–1501.

——, and Coauthors, 2004b: Using the PARAGON framework to establish an accurate, consistent, and cohesive long-term aerosol record. *Bull. Amer. Meteor. Soc.,* **85,** 1535–1548.

DOE, cited 2001: Advanced computing and high energy physics for the 21st century. [Available online at www.eurekalert.org/features/doe/2001-12/dnal-aca061802.php.]

——, cited 2002a: MICS High Performance Networks FY 2002 HPN program highlights. [Available online at www.sc.doe.gov/ascr/mics/hpn/index.html.]

——, cited 2002b: MICS High Performance Networks program description. [Available online at www.sc.doe.gov/ascr/mics/hpn/HPN_Description.htm.]

Dubovik, O., and M. D. King, 2000: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements. *J. Geophys. Res.,* **105,** 20 673–20 696.

——, A. Smirnov, B. N. Holben, M. D. King, Y. J. Kaufman, T. F. Eck, and I. Slutsker, 2000: Accuracy assessments of aerosol optical properties retrieved from Aerosol Robotic Network (AERONET) Sun and sky radiance measurements. *J. Geophys. Res.,* **105,** 9791–9806.

——, C. Gatebe, A. Sinyuk, E. Vermote, M. King, and B. Holben, 2003: Aerosol and surface retrieval from a combination of up-looking and down-looking observations. *IGARSS Meeting 2003,* Toulouse, France, IGARSS. [Available online at http://car.gsfc.nasa.gov/_ppt/IGARSS_2003.ppt.]

Fargion, G. S., and Coauthors, 2003: SIMBIOS program in support of ocean color mission: 1997–2003. *Proc. SPIE,* **5155,** 49–60.

Foster, I., 2002: The grid: A new infrastructure for 21st century science. *Phys. Today,* **55,** 42–47.

Freeman, P. A., cited 2003: Revolutionizing science and engineering: The role of cyberinfrastructure. UCSD Jacobs School of Engineering Research Review, 28 pp. [Available online at www.cise.nsf.gov/oad/freeman_talk_page/freeman_jacobs_files/jacob_file.pdf.]

Garay, M. J., D. M. Mazzoni, R. Davies, D. Decoste, and A. Braverman, 2003: A hybrid global MISR cloud mask using support vector machines and active learning. *Eos Trans. Amer. Geophys. Union,* **84** (Fall Meeting Suppl. H11F-0913).

Gates, W. L., 1992: The validation of atmospheric models. Program for Climate Model Diagnosis and Intercomparison Rep. UCRL-ID-109810, 17 pp.

Gersho, A., and R. M. Gray, 1991: *Vector Quantization and Signal Compression.* Kluwer Academic, 732 pp.

Goody, R., J. Anderson, and G. North, 1998: Testing climate models: An approach. *Bull. Amer. Meteor. Soc.,* **79,** 2541–2549.

——, ——, T. Karl, R. B. Miller, G. North, J. Simpson, G. Stephens, and W. Washington, 2002: Why monitor the climate? *Bull. Amer. Meteor. Soc.,* **83,** 873–878.

Hastie, T., R. Tibshirani, and J. Friedman, 2001: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer Series in Statistics, Springer, 552 pp.

Hinke, T., and J. Novotny, 2000: Data mining on NASA's Information Power Grid. *Proc. Ninth IEEE Int. Symp. on High Performance Distributed Computing,* Pittsburgh, PA, IEEE, 292.

Huang, H.-C., N. Cressie, and J. Gabrosek, 2002: Fast, resolution-consistent spatial prediction of global processes from satellite data. *J. Comput. Graphic Stat.,* **11,** 63–88.

Johnston, W. E., D. Gannon, and B. Nitzberg, 1999: Grids as production computing environments: The engineering aspects of NASA's Information Power Grid. *Proc. Eighth IEEE Int. Symp. on High Performance Distributed Computing,* Redondo Beach, CA, IEEE, 34.

Kahn, R., and A. Braverman, 1999: What shall we do with the data we are expecting from upcoming Earth observation satellites? *J. Comput. Graphic Stat.,* **8,** 575–588.

——, P. Banerjee, D. McDonald, and D. Diner, 1998: Sensitivity of multiangle imaging to aerosol optical depth, and to pure-particle size distribution and composition over ocean. *J. Geophys. Res.,* **103,** 32 195–32 213.

——, and Coauthors, 2004a: Environmental snapshots from ACE-Asia. *J. Geophys. Res.,* in press.

——, and Coauthors, 2004b: Aerosol data sources and their roles within PARAGON. *Bull. Amer. Meteor. Soc.,* **85,** 1511–1522.

Kaufman, Y. J., D. Tanré, and O. Boucher, 2002: A satellite view of aerosols in the climate system. *Nature,* **419,** 215–223.

——, J. M. Haywood, P. V. Hobbs, W. Hart, R. Kleidman, and B. Schmid, 2003: Remote sensing of vertical distributions of smoke aerosol off the coast of Africa. *Geophys. Res. Lett.,* **30,** 1831, doi:10.1029/2003GL017068.

Kinne, S., 2003: Aerosol representation in global models: Discrepancies among models and to data. *Geophys. Res. Abstr.,* **5,** 9670.

Konovalov, I., 2003: Nonlinear relationships between atmospheric aerosol and its gaseous precursors: Analysis of long-term air quality monitoring data by means of neural networks. *Atmos. Chem. Phys. Discuss.,* **3,** 835–866.

Kyriakidis, P. C., and A. G. Journel, 1999: Geostatistical space-time models: A review. *Math. Geol.,* **31,** 651–684.

Léon, J.-F., D. Tanré, J. Pelon, Y. J. Kaufman, J. M. Haywood, and B. Chatenet, 2003: Profiling a Saharan dust outbreak based on a synergy between active and passive remote sensing. *J. Geophys. Res.,* **108,** 8575, doi:10.1029/2002JD002774.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

MacQueen, J. A., 1967: Some methods for classification and analysis of multivariate observations. Vol. 1, *Proc. Fifth Berkeley Symp. on Mathematical Statistics and Probability,* Univ. of California, Berkeley, CA, 281–296.

Magi, B. I., P. V. Hobbs, T. W. Kirchstetter, T. Novakov, D. A. Hegg, S. Gao, J. Redemann, and B. Schmid, 2004: Aerosol properties and chemical apportionment of aerosol optical depth at locations off the United States East coast in July and August 2001. *J. Atmos. Sci.,* in press.

Martonchik, J. V., D. J. Diner, K. A. Crean, and M. A. Bull, 2002: Regional aerosol retrieval results from MISR. *IEEE Trans. Geosci. Remote Sens.,* **40,** 1520–1531.

McDonald, K., L. Di, and S. W. Wharton, 2002: Strategic Evolution of NASA ESE Data Systems (SEEDS). *Int. Arch. Photogramm. Rem. Sens.,* **XXXIV,** 327–330.

Nakajima, T., M. Tanaka, and T. Yamauchi, 1983: Retrieval of the optical properties of aerosols from aureoles and extinction data. *Appl. Opt.,* **22,** 2951–2959.

NRC, 1991: *Spatial Statistics and Digital Image Analysis.* 244 pp.

Ogren, J. A., 1995: A systematic approach to in situ observations of aerosol properties. *Aerosol Forcing of Climate,* R. J. Charlson and J. Heintzenberg, Eds., John Wiley & Sons, 215–226.

Penner, J. E., and Coauthors, 2002: A comparison of model- and satellite-derived aerosol optical depth and reflectivity. *J. Atmos. Sci.,* **59,** 441–460.

Pollard, D., 1982: Quantization and the method of k-means. *IEEE Trans. Inf. Theory,* **28,** 199–205.

Rasch, P. J., and Coauthors, 2000: A comparison of scavenging and deposition processes in global models: Results from the WCRP Cambridge workshop of 1995. *Tellus,* **52B,** 1025–1056.

——, W. D. Collins, and B. E. Eaton, 2001: Understanding the Indian Ocean Experiment (INDOEX) aerosol distributions with an aerosol assimilation. *J. Geophys. Res.,* **106,** 7337–7345.

Reid, J. S., and Coauthors, 2003: Analysis of measurements of Saharan dust by airborne and ground-based remote sensing methods during the Puerto Rico Dust Experiment (PRIDE). *J. Geophys. Res.,* **108,** 8586, doi:10.1029/2002JD002493.

Remer, L. A., and Coauthors, 2002: Validation of MODIS aerosol retrieval over ocean. *Geophys. Res. Lett.,* **29,** 8008, doi:10.1029/2001GL013204.

Russell, P. B., and J. Heintzenberg, 2000: An overview of the ACE 2 clear sky column closure experiment (CLEARCOLUMN). *Tellus,* **52B,** 463–483.

——, J. M. Livingston, P. Hignett, S. Kinne, J. Wong, and P. V. Hobbs, 1999: Aerosol-induced radiative flux changes off the United States Mid-Atlantic coast: Comparison of values calculated from sun photometer and in situ data with those measured by airborne pyranometer. *J. Geophys. Res.,* **104,** 2289–2307.

——, and Coauthors, 2002: Comparison of aerosol single scattering albedos derived by diverse techniques in two North Atlantic experiments. *J. Atmos. Sci.,* **59,** 609–619.

Satheesh, S. K., V. Ramanathan, B. N. Holben, K. Krishna Moorthy, N. G. Loeb, H. Maring, J. M. Prospero, and D. Savoie, 2002: Chemical, microphysical, and radiative effects of Indian Ocean aerosols. *J. Geophys. Res.,* **107,** 4725, doi:10.1029/2002JD002463.

Schulz, M., and S. Kinne, 2003: AEROCOM: An intercomparison to improve global aerosol models. *Geophys. Res. Abstr.,* **5,** 11 640.

Scott, D. W., 1992: *Multivariate Density Estimation: Theory, Practice and Visualization.* John Wiley and Sons, 336 pp.

Seinfeld, J. H., and Coauthors, 2004: Scientific objectives, measurement needs, and challenges motivating the PARAGON aerosol initiative. *Bull. Amer. Meteor. Soc.,* **85,** 1503–1509.

Shannon, C. E., 1948: A mathematical theory of communication. *Bell Syst. Technol. J.,* **27,** 379.

Stephens, G. L., R. J. Engelen, M. Vaughan, and T. L. Anderson, 2001: Toward retrieving properties of the tenuous atmosphere using space-based lidar measurements. *J. Geophys. Res.,* **106,** 28 143–28 157.

Stroud, P. D., C. D. Cunningham, and G. Guethlein, 2000: Rapid detection and classification of aerosol events based on changes in particle size distribution. *Proc. SPIE,* **4036,** 42–53.

Szalay, A., and J. Gray, 2001: The World-Wide Telescope. *Science,* **293,** 2037–2040.

Tanré, D., F. M. Bréon, J. L. Deuzé, M. Herman, P. Goloub, F. Nadal, and A. Marchand, 2001: Global observation of anthropogenic aerosols from satellite. *Geophys. Res. Lett.,* **28,** 4555–4558.

Trenberth, K. E., T. R. Karl, and T. W. Spence, 2002: The need for a systems approach to climate observations. *Bull. Amer. Meteor. Soc.,* **83,** 1593–1602.

Vucetic, S., and Z. Obradovic, 2000a: Performance controlled data reduction for knowledge discovery in distributed databases. *Proc. Third Pacific-Asia Conf. on Knowledge Discovery and Data Mining (PAKDD'00),* Kyoto, Japan, 29–39.

——, and ——, 2000b: Discovering homogeneous regions in spatial data through competition. *Proc. 17th Int. Conf. on Machine Learning,* Stanford, CA, 1091–1098.

Walter, A., M. Denhard, and C. Schonwiese, 1998: Simulation of global temperature variations and signal detection using neural networks. *Meteor. Beit.,* **7,** 171–180.

Werdell, P. J., S. W. Bailey, G. S. Fargion, C. Pietras, K. D. Knobelspiesse, G. C. Feldman, and C. R. McClain, 2003: Unique data repository facilitates ocean color satellite validation. *Eos Trans. Amer. Geophys. Union,* **84,** 377–387.

Wikle, C. K., L. M. Berliner, and N. Cressie, 1998: Hierarchical Bayesian space–time models. *Environ. Ecol. Stat.,* **5,** 117–154.

——, R. F. Milliff, D. Nychka, and M. L. Berliner, 2001: Spatio-temporal hierarchical Bayesian modeling: Tropical ocean surface winds. *J. Amer. Stat. Assoc.,* **96,** 382–397.

Zador, P. L., 1964: Development and evaluation of procedures for quantizing multivariate distributions. Ph.D. thesis, Department of Statistics, Stanford University, Univ. Microfilm 64-9855.