Mesoscale Variations of Tropospheric Aerosols*

THEODORE L. ANDERSON AND ROBERT J. CHARLSON
University of Washington, Seattle, Washington

DAVID M. WINKER
NASA Langley Research Center, Hampton, Virginia

JOHN A. OGREN
NOAA/Climate Monitoring and Diagnostics Laboratory, Boulder, Colorado

KIM HOLMÉN
Stockholm University, Stockholm, Sweden

(Manuscript received 22 April 2002, in final form 15 July 2002)

ABSTRACT

Tropospheric aerosols are calculated to cause global-scale changes in the earth's heat balance, but these forcings are space/time integrals over highly variable quantities. Accurate quantification of these forcings will require an unprecedented synergy among satellite, airborne, and surface-based observations, as well as models. This study considers one aspect of achieving this synergy—the need to treat aerosol variability in a consistent and realistic way. This need creates a requirement to rationalize the differences in spatiotemporal resolution and coverage among the various observational and modeling approaches. It is shown, based on aerosol optical data from diverse regions, that mesoscale variability (specifically, for horizontal scales of 40–400 km and temporal scales of 2–48 h) is a common and perhaps universal feature of lower-tropospheric aerosol light extinction. Such variation is below the traditional synoptic or ''airmass'' scale (where the aerosol is often assumed to be essentially homogeneous except for plumes from point sources) and below the scales that are readily resolved by chemical transport models. The present study focuses on documenting this variability. Possible physical causes and practical implications for coordinated observational strategies are also discussed.

1. Introduction

a. Climate forcing problem requires a synergy among research methods

Three disparate approaches have heretofore been employed to study atmospheric aerosols: 1) in situ observations of microphysical, chemical, and optical properties of the aerosol and of thermodynamic state; 2) remote sensing of aerosol influence on the propagation of radiation in the atmosphere (including satellite observations); and 3) chemical mass balance, transport, and radiative modeling. Many lines of evidence to date provide compelling cases for regional- to global-scale aerosol effects including biogeochemical cycling of important species (Likens 1981), acidification of precipitation [both natural (Charlson and Rodhe 1982) and anthropogenic (Galloway et al. 1984)] and climate forcing by both direct and cloud-mediated mechanisms (Junge 1975; Charlson et al. 1992). The focus of this article is on climate forcing, although the results apply to research on large-scale aerosol effects in general.

Aerosol forcing estimates to date are highly uncertain (i.e., a factor of 2 or worse) and, indeed, are the dominant source of uncertainty in current estimates of industrial-era climate forcing (IPCC 2001, p. 8). A major cause of this uncertainty stems from aerosol variability that, in turn, arises from patchy sources, sinks, and the short residence time (days) of tropospheric aerosols. Thus, estimates of global (or even regional) climate forcings by aerosols must address the difficult issue of integrating in space and time over highly variable quantities. While only satellites can provide the needed global coverage, satellite sensors cannot determine chemical

* Joint Institute for the Study of Atmosphere and Ocean Contribution Number 915.

Corresponding author address: Dr. Theodore Anderson, Department of Atmospheric Sciences, Box 351640, University of Washington, Seattle, WA 98195-1640.
E-mail: tadand@u.washington.edu

composition; moreover, the aerosol properties that are retrieved from satellite-measured radiances are mathematically underdetermined [i.e., requiring assumptions about the chemical and microphysical nature of the aerosol; see, e.g., Mishchenko et al. (1999)]. Indeed, as recognized by the U.S. National Academy of Sciences Panel on Aerosol Radiative Forcing and Climate Change, none of the three approaches listed above is sufficient in itself, and overcoming their individual limitations depends on synergy among the three (Seinfeld et al. 1996).

Here we consider one aspect of achieving this synergy—the need to treat mesoscale aerosol variability (horizontal and temporal) in a consistent and realistic way. This need arises because 1) as shown herein, extensive[1] aerosol properties exhibit significant mesoscale variations, and 2) the three approaches vary widely in terms of their ability to resolve and survey these mesoscale variations. The term ''mesoscale'' refers in meteorology to horizontal scales from about 1 to 1000 km. Important mesoscale phenomena include gravity waves, diurnal cycles such as sea/land breezes, boundary layer rolls (Brown 1980), precipitation patterns (Austin and Houze 1972), as well as many aerosol sources such as dust storms, biomass fires, and megacity plumes. Here we focus specifically on scales of 10–600 km and 1–48 h.

b. *Variability data provide insights into physical processes*

The variability of an atmospheric trace substance contains information on the processes (sources, transformations, sinks) that determine its concentration. The pioneering work in this area was done by Junge (1974), who argued on both theoretical and observational grounds that the global spatial variability of a trace gas is inversely related to its atmospheric residence time. The range of residence times considered spanned many orders of magnitude from days (H_2O and Rn) to millennia (O_2). Subsequent theoretical (e.g., Hamrud 1983) and observational (e.g., Jobson et al. 1999) studies have refined knowledge of this relationship, demonstrating that variability is a function, not only of atmospheric residence time, but also of the space/time distribution of sources and sinks. This is especially true for short-lived constituents like aerosols.

Bolin and Rodhe (1973) distinguished between average residence time and average age of a substance moving through a reservoir, which are not necessarily equal even for the equilibrium case. The average age will be greater than the average residence time for the case where there are two classes of sinks: one close to the source that removes a large fraction of material and another distant from the source. This situation describes the cycling of most aerosol substances, which are emitted at the surface, removed largely by dry and wet deposition in the surface layer or boundary layer close to sources, but with some fraction surviving to be transported over longer distances.

Additional insight into physical processes comes from examining the frequency distribution of measured concentrations. For atmospheric pollutants, including aerosols, these are commonly found to be lognormal. This has generally been interpreted as the result of successive, random dilution events (Ott 1990) associated with atmospheric turbulence on a variety of scales. (Dilution is a multiplicative process, such that a series of random dilutions asymptotically approaches a lognormal probability distribution of concentration by the central limit theorem.) This analysis is intuitively attractive and should be amenable to quantitative diagnosis.

We defer discussion of variability in the vertical dimension to the time when adequate amounts of data become available, for example, from satellite-borne lidar. This is not to suggest that variability in the vertical is unimportant; indeed, it is essential to recognize and understand it. There are many facets of such variability that can be immediately recognized: layers of aerosol in the stratosphere from volcanic injection, visible and subvisible cirrus, the stability (on average) of the planetary boundary layer and the existence therein of higher aerosol concentrations than in the free troposphere, the observations of laminae of higher and lower aerosol content within and above the PBL (e.g., Kent et al. 1998; Hobbs 2002), the likely correlation of such layers with the thermodynamic state of the air (especially the relative humidity), and the complex effects of mixing processes in the vertical dimension that change both the aerosol extensive properties and the thermodynamic state. The present study does not address these issues except to comment, in the discussion section, on the consequences of horizontal variability on aircraft observations of vertical profiles.

c. *Challenging the ''airmass'' paradigm*

In the 1970s, pollution aerosols, generated in and near urban areas, were shown to be a regional phenomenon. That is, the concentration and composition at rural sites in the U.S. Midwest were observed (with extant technologies) to vary on timescales of days, and these variations could be related to airmass[2] character via back trajectories (Charlson et al. 1974; Vanderpol et al. 1975). The parameter used to identify airmass changes was the molecular form of sulfate, which varied from neutral to

[1] An *extensive* property depends on the amount of material present. Examples are mass concentration, light scattering coefficient, or optical depth. The variability of *intensive* properties, such as scattering-to-mass ratio, will be the subject of a future paper.

[2] The term ''air mass'' in meteorology refers to a widespread body of air, approximately homogeneous in the horizontal, with characteristic temperature and moisture properties established while that air was situated over a particular region of the earth's surface (Glickman 2000, p. 22).

highly acidic. Charlson et al. (1974) coined the term ''regional background'' to describe this phenomenon, meaning both natural and manmade aerosol material distant from its sources. The 1980s saw large networks of aerosol and rainwater monitoring stations operating in Europe and the United States that focused on understanding the acid rain phenomenon. Along with these observational efforts, chemical transport models were developed. Summarizing much of this new information, Schwartz (1989) showed that aerosol transport distances in the lower troposphere were several hundred to a few thousand kilometers associated with synoptic-scale meteorological processes (see also, Benkovitz et al. 1994).

The implicit notion emerging from these studies was that aerosol concentration and composition (away from the immediate vicinity of local sources) varies horizontally on the airmass scale (of order 1000 km) and is essentially constant within air masses. This notion remains to this day as a prevalent (usually implicit) assumption underlying many aerosol modeling and observational activities. (An example of this implicit assumption is the nearly universal use of 24-h chemical sampling in aerosol monitoring and field research, since 24-h samples can resolve synoptic scale but not mesoscale variations.) One purpose of this article is to challenge this notion of air mass by examining variability within the airmass scale.

Several previous studies have focused on issues associated with subsynoptic-scale aerosol variability. These include using aerosols to detect boundary layer rolls (Mourad and Walter 1996), health effects associated with short-term aerosol exposure (Michaels and Kleinman 2000), measurement discrepancies between instruments deployed on different platforms (Clarke et al. 2002; Masonis et al. 2002), analysis of subgrid-scale variability with respect to a sulfate chemical transport model (Benkovitz and Schwartz 1997), analysis of possible bias in polar orbiting satellites that measure aerosol optical depth at only one time of day (Kaufmann et al. 2000), and analysis of variability and interrelationships on diurnal to seasonal scales for aerosol optical properties measured from surface monitoring stations in North America (Delene and Ogren 2002). The latter paper examines a dataset that overlaps with the one analyzed herein. No previous study, to our knowledge, has attempted a general analysis of mesoscale aerosol variations. By examining high-resolution optical measurements from a variety of regions, the present study suggests that such a general analysis is both feasible and has important practical applications.

2. Aerosol variability

a. Dataset descriptions

Three of the four datasets consist of nephelometer measurements of low relative humidity (RH) light scattering at 550 nm, $\sigma_{sp,dry}$. This parameter is known to be highly correlated to low RH aerosol mass concentration ($r \sim 0.9$), with a mass scattering efficiency close to 3 $m^2\ g^{-1}$ for submicron-dominated industrial haze (Waggoner et al. 1981; White et al. 1994) and considerably lower and more variable efficiencies for supermicron-dominated dust (White 1990) and seasalt (Quinn et al. 1995) particles. Thus, within a given aerosol type, variation in $\sigma_{sp,dry}$ can be taken as an excellent proxy for variation in dry aerosol mass concentration.[3] The fourth dataset consists of aerosol optical depth (AOD) retrieved from spaceborne lidar. AOD is related to $\sigma_{sp,dry}$ since 1) AOD is the vertical integral of ambient light extinction σ_{ep} and 2) variations in σ_{ep} are dominated by variations in $\sigma_{sp,dry}$. However, AOD measurements differ in that 1) they include additional variation associated with changes in aerosol light absorption and in ambient RH, the latter being relevant whenever hygroscopic aerosol components are present; and 2) they involve considerable smoothing associated with integration in the vertical. All measurements apply to the lower troposphere. Measurements of $\sigma_{sp,dry}$ were obtained either at the surface (ground-based data) or within 500 m of the surface (airborne data). Lower tropospheric AOD, AOD_{lt}, is analyzed by integrating lidar-retrieved σ_{ep} from the surface to 5 km above sea level.

The four datasets taken together provide information on both temporal and spatial (horizontal) variations. To assess temporal variations at a fixed location, we examine multiyear data on $\sigma_{sp,dry}$ from two surface stations: one within a pollution source region (central United States) and one remote from aerosol sources (Arctic Ocean). Spatial variations are assessed with several thousand kilometers of 1-km-resolution boundary layer measurements of $\sigma_{sp,dry}$ acquired by airplane off the coast of China, and with several thousand kilometers of 7.4-km-resolution data on AOD_{lt} acquired by spaceborne lidar. While both the airborne and spaceborne measurements can provide vertical resolution, we restrict the present study to horizontal variations in the lower troposphere. Table 1 provides a general overview of each dataset. Below we provide further details.

The Bondville aerosol research station (BND) is located in a rural area of central Illinois (40.05°N, 88.37°W, 229 m) within a major source region for sulfate aerosol from coal-burning power plants. The nearest significant power plant is 65 km away, and there are approximately 80 significant power plants within a radius of 500 km. This station has operated since 1993 as the polluted, continental site in a network of midlatitude stations managed by the National Oceanic and Atmospheric Administration (NOAA) Climate Monitoring and Diagnostics Laboratory (CMDL). [Equivalent data are available from four other North American sites within this network; see Delene and Ogren (2002)

[3] In contrast, there is no general relationship between $\sigma_{sp,dry}$ and aerosol number concentration—indeed, the two can be correlated, uncorrelated, or anticorrelated.

TABLE 1. Description and overview of data.

(a) General description

	Location	Time	Platform	Measured quantity, wavelength	Inlet cut diameter	RH[a]
BND	Rural Illinois (polluted)	1997–2000	Ground	σ_{sp}, 550 nm	10 μm	Low
Spitz	Arctic Ocean (remote)	1991–95	Ground	σ_{sp}, 550 nm	2.5 μm	Low
ACE-Asia	Asia outflow region, mostly Yellow Sea and Sea of Japan	Spring 2001	Aircraft	σ_{sp}, 550 nm	10 μm	Low
LITE	Line from eastern United States across Atlantic to eastern Brazil	Sep 1994	Space shuttle STS-64	AOD (0–5 km), 532 nm	None	Ambient

(b) Size, coverage, and filtering

	Length of record	Resolution	N^b_{raw}	Coverage[c]	N^d_{filt}
BND	4 yr	1 h	30 449	87%	81
Spitz	4.33 yr	1 h	29 899	79%	9926
ACE-Asia	14 091 km	1 km (10 s)	14 091	n/a[e]	1194
LITE	7237 km	7.4 km (1 s)	943	96%	24

(c) Basic statistics

	N[f]	Arithmetic mean	Arithmetic std	Geometric mean[g]	Geometric std[h]
BND	30 368	55.7 Mm^{-1}	45.5 Mm^{-1}	40.9 Mm^{-1}	2.25
Spitz	19 973	4.67 Mm^{-1}	4.31 Mm^{-1}	3.15 Mm^{-1}	2.55
ACE-Asia	12 897	160 Mm^{-1}	134 Mm^{-1}	123 Mm^{-1}	2.02
LITE	919	0.159	0.099	0.124	2.17

[a] RH is the relative humidity at which the measurement is made.
[b] N_{raw} is the number of data points prior to filtering.
[c] Coverage is the ratio of data points to measurement opportunities over length of record.
[d] N_{filt} is the number of data points removed by filtering process (explained in text for each dataset).
[e] Measurements were acquired 100% of the time during all legs, but the airplane only flew along arbitrarily selected tracks. Thus, coverage is far less than 100% and cannot be quantified.
[f] N is the number of data points used in analysis (after filtering).
[g] Geometric mean = exp{mean[ln(x)]}.
[h] Geometric std = exp{std[ln(x)]}.

and Anderson et al. (1999).] The sampling system and instrumentation at BND have been described in previous publications (Koloutsou-Vakakis et al. 1998, 1999). For the present study, we examine 4 yr (1997–2000) of hourly averaged $\sigma_{sp,dry}$ data measured with a standard, high-sensitivity nephelometer (model 3563; TSI, Inc.). (Note that 1-min-resolution data are also available and are used herein to extend some of the analyses to shorter timescales.) All particles smaller than 10 μm (aerodynamic diameter at low RH) were sampled and measured. Note that at this station 80%–90% of the midvisible light scattering is due to particles smaller than 1 μm (Delene and Ogren 2002). The variability properties of the submicron aerosol, which is measured separately, are virtually indistinguishable from the properties reported herein.

The second surface site (Spitz) is the Zeppelinfjellet air chemistry station located on a small mountain on Spitzbergen[4] (79°N, 12°E, 474 m). This station has been operated since 1989 by the Department of Meteorology at Stockholm University for baseline aerosol and CO_2 measurements. Aerosol measurements are described by Heintzenberg and Leck (1994). We examine herein approximately 4 yr (1991–95) of hourly averaged measurements of $\sigma_{sp,dry}$, measured by a custom, high-sensitivity nephelometer (Heintzenberg and Bäcklin 1983). Particles larger than roughly 2 μm (aerodynamic diameter at ambient RH) were excluded via a cyclone placed at the sampling inlet. The station is sometimes engulfed in stratiform cloud, which prevents valid aerosol sampling since particles within cloud droplets will be excluded by the cyclone. To eliminate data from such periods, we adopt a simplified and conservative version of the cloud-clearing scheme developed by Heintzenberg and Leck (1994)—namely, data were eliminated whenever the light scattering fell below 0.5 times the monthly average. This procedure serves as well to reduce the noise component of variability (discussed in more detail below).

Airborne data from the Aerosol Characterization Experiment-Asia (ACE-Asia) are used to assess spatial variations. As part of ACE-Asia, the National Center for Atmospheric Research (NCAR) C-130 aircraft flew 19 flights out of Iwakuni, Japan, from 29 March–6 May 2001, mostly over the Yellow Sea and Sea of Japan. Two TSI nephelometers (model 3563) were used to measure $\sigma_{sp,dry}$, by submicron ($D_{aero,dry} < 1$ μm) and total ($D_{aero,dry} < 10$ μm) aerosol simultaneously, allowing a

[4] Also known as Svalbard.

real-time distinction between the two major aerosol types encountered: submicron-dominated pollution and supermicron-dominated mineral dust. A new, low-turbulence aerosol inlet (Lafleur 1998) was deployed on the C-130 and proved to have very high passing efficiency for supermicron particles up to roughly 10 μm. For example, the twin nephelometer data frequently showed that more than 80% of $\sigma_{sp,dry}$ was due to supermicron particles during dust-sampling events. For the present analysis, we examine the variability of scattering by the total aerosol only ($D_{aero,dry} < 10$ μm). In addition, for comparison to ground-based data we restrict the analysis to level flight legs of 4 min or longer flown in the boundary layer below 500 m. Ninety-seven flight legs, all flown over the ocean, met these criteria. Data are averaged to 10 s, resulting in 1-km spatial resolution for the C-130 flight speed of 100 m s^{-1}. The legs ranged from 4–108 min in duration (24–648 km in length). The data contain more information for assessing short spatial scales than long scales. For example, the dataset (12 897 valid measurements in all) includes 11 911 valid data pairs that are separated by 10 km, but only 7712 and 1470 valid data pairs at spatial separations of 60 and 240 km, respectively. The subset of data available for examining longer spatial separations is statistically different from the larger dataset (Table 1), requiring a careful approach to calculating the autocorrelation function. In addition, the larger dataset does not necessarily represent aerosol variability over the entire ACE-Asia study region, since the C-130 did not fly random flight tracks but, rather, sampled regions where significant aerosol events were anticipated. These issues are addressed further in the analysis section below.

Another means of assessing horizontal spatial variations is provided by spaceborne lidar data obtained from the Lidar In-Space Technology Experiment (LITE; Winker et al. 1996). The velocity of the LITE footprint at the earth's surface was about 7.4 km s^{-1}. With a laser pulse rate of 10 Hz and a footprint diameter of about 290 m, each laser pulse samples a different column of the atmosphere. For the present study, we examine a 7500-km portion of orbit 117, acquired during about 17 min of flight, that begins in the eastern United States (38.4°N, 78.0°W), crosses the Caribbean, and ends in eastern Brazil (15.1°S, 41.2°W). This transect was found to be fairly clear of dense clouds and an almost continuous record of lower tropospheric aerosol optical depth (AOD_{lt}) could be retrieved. The few clouds that were observed were cleared manually. This transect crossed a number of airmass regions dominated by various aerosol types: polluted continental over the eastern United States and off the U.S. coast, clean marine over the western Atlantic, Sahara dust over the Caribbean, and smoke from biomass burning near Brazil. Ten-shot averages were computed, giving a horizontal resolution of 7.4 km. A calibrated, elastic backscatter lidar provides a direct measurement of attenuated backscatter (aerosol plus molecular) as a function of height. A two-component forward retrieval scheme was used to derive vertical profiles of aerosol backscatter and extinction (the two are retrieved simultaneously). Extinction was then integrated from the surface to 5 km above sea level to obtain AOD_{lt} at a wavelength of 532 nm. The height of 5 km was chosen because the mixed layer, containing the vast majority of tropospheric aerosol in this transect, was always below 5 km. For this retrieval, the aerosol extinction-to-backscatter ratio (or lidar ratio) S_a was assumed to be constant at 30 sr. This assumption is unlikely to be valid for all the aerosol types encountered, but should not significantly affect the analysis of variability discussed herein. That is, the magnitude of the retrieved AOD_{lt} may be incorrect, but the relative variability of AOD_{lt} over short spatial scales should be retrieved accurately because the true value of the lidar ratio is not expected to change significantly over short spatial scales. Similarly, vertical variations of S_a (Ferrare et al. 2001) will affect AOD_{lt} absolute accuracy but should not affect the present analysis of horizontal variability. Put another way, our analysis of horizontal AOD_{lt} variability assumes that the observed mesoscale variations are caused by variations in aerosol concentration and not by local gradients in the lidar ratio. This assumption is plausible and is consistent with measurements of S_a variability (Anderson et al. 2000) but remains to be thoroughly tested. An alternate approach to variability analysis with lidar data (not pursued herein) would be to examine integrated, attenuated backscatter—a quantity that requires no assumption of S_a.

Sample selection is an important aspect of any statistical analysis. Two aspects relevant here are coverage and filtering (Table 1b). Coverage refers to the fraction of samples that were measured relative to the domain being investigated. (Unbiased estimates of domain properties require either complete coverage or random sampling.) Coverage is excellent at the ground stations (87% and 79%, respectively); data gaps reflect times when the instruments were not operating (e.g., due to maintenance or calibration) or were known to be operating improperly. Coverage of the LITE transect is also excellent (96%); the main data gap (see Fig. 2f) is due to a brief telemetry dropout. For the ACE-Asia dataset, it is not clear how to quantify the coverage. If the domain is defined as the actual flight tracks, then the coverage is 100%, since the instruments were always functioning. However, the domain thus defined is not a random sample of the atmosphere, as in the other three cases, because the flight tracks were deliberately chosen with the intention of finding and studying interesting aerosol phenomena. In terms of a more meaningful domain (e.g., significant aerosol features in the near-surface atmosphere in the ACE-Asia study region/time), it is clear that sample coverage was very small. Combined with the fact that the representativeness of the sample set is not known, this constitutes a weakness of our ACE-Asia dataset. (This weakness is typical of airborne data acquired during intensive campaigns of this nature.)

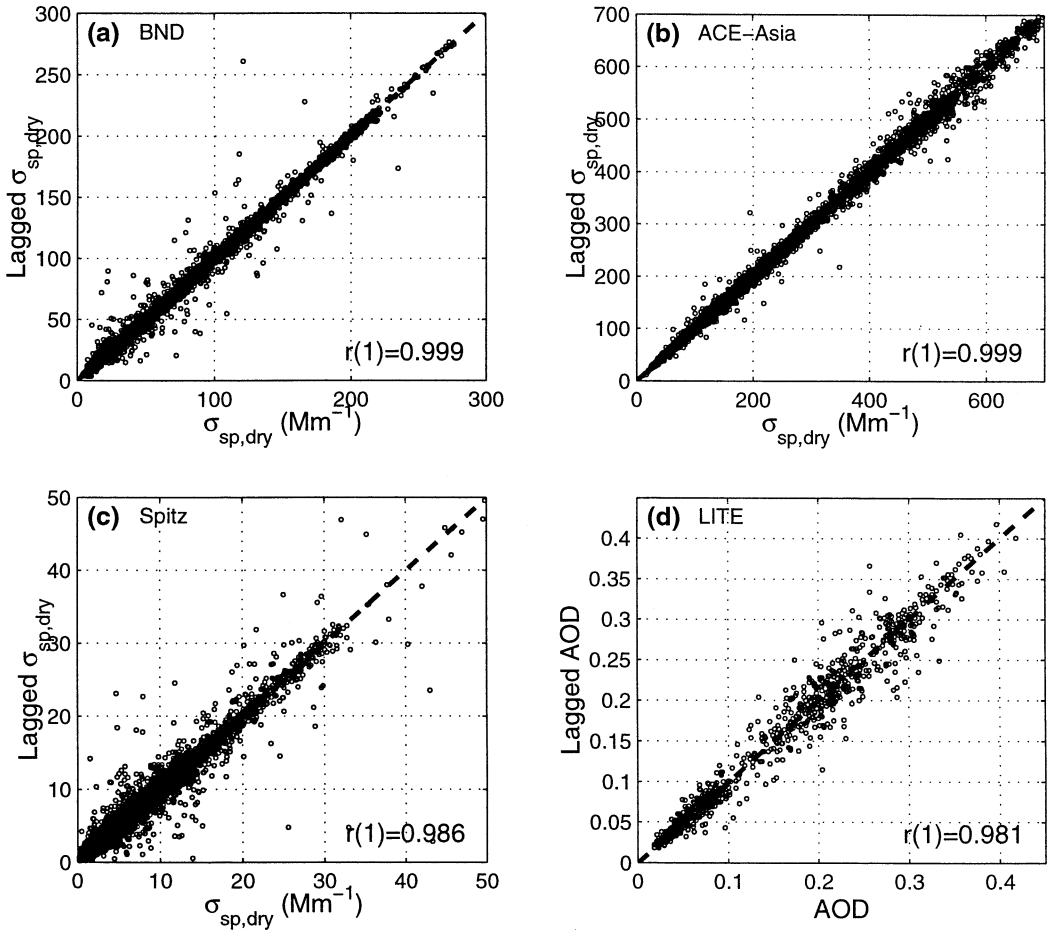

FIG. 1. Lag-1 correlation. The correlation coefficient shown on each panel is identical to the autocorrelation statistic [Eq. (1)] at a lag of one measurement interval. For the ground-based data, the highest resolution available is used. (a) Ground-based data from Bondville, IL, at 1-min resolution. (b) Airborne data from ACE-Asia at 10-s or 1-km resolution. (c) Ground-based data from Spitzbergen at 30-min resolution. (d) Spaceborne data from LITE at 1-s or 7.4-km resolution.

b. Sources of variation

1) NATURE VERSUS NOISE

Measured variability is the sum of ambient variations and measurement noise. Therefore, ambient variability can be studied only if 1) measurement noise is known to be negligible or 2) measurement noise can be quantified and a correction can be applied. As a robust method of evaluating measurement noise, we examine the lag-1 autocorrelation coefficient $r(1)$. For evenly spaced data, the autocorrelation function, $r(k)$, is simply the correlation coefficient among all data pairs x_i and x_{i+k} that exist at a separation, or lag, of k. That is,

$$r(k) = \frac{\sum_i [(x_i - m_1)(x_{i+k} - m_2)]}{s_1 s_2}, \qquad (1)$$

where m_1 and s_1 are the mean and standard deviation, respectively, of all data points that are located $-k$ away from another data point, and m_2 and s_2 are the corresponding mean and standard deviation of all data points that are located $+k$ away from another data point. This definition provides the least biased estimate because it explicitly takes into account that subsets of the data may have different means and standard deviations from the entire dataset (Isaaks and Srivastava 1989, p. 59). This is especially important for datasets that are relatively small and/or contain significant gaps, as herein.

Random instrumental noise will manifest itself by lowering the lag-1 autocorrelation $r(1)$. Note, however, that natural variability at the lag-1 scale will also act to lower $r(1)$. Therefore, if $r(1)$ is close to unity, it indicates that both measurement noise and ambient variability at the maximum instrumental resolution are small. This is the ideal case because it shows both that measurement noise does not contribute significantly to measured variability and that the instrumental method was successful in resolving all ambient variations on the scales of interest.

Figure 1 shows the $r(1)$ results for the four datasets.

For the ground-based datasets, we calculate $r(1)$ from the highest-resolution data available. This was 1-min data in the case of Bondville, Illinois, and 30-min data in the case of Spitzbergen. Although 2-s data are available from the ACE-Asia campaign, successive 2-s data points are not considered independent because the flush time of the nephelometer is about 6 s. For the LITE mission, 10-Hz data are available, but are known to contain significant instrumental noise. Therefore, $r(1)$ for the spatial datasets is analyzed at the resolutions used elsewhere in this paper: 10 s or 1 km for ACE-Asia, and 1 s or 7.4 km for LITE.

The resulting $r(1)$ values indicate that both the Bondville and the ACE-Asia data are very close to the ideal case described above—in each case, $r(1) = 0.999$. Instrumental noise is seen to be quite small for the Spitzbergen and LITE data as well—$r(1)$ values are 0.986 and 0.981, respectively. It is worth recalling at this point that the Spitzbergen and LITE data have been filtered to remove measurements that are thought to be cloud contaminated. This filtering removed a significant fraction of the Spitzbergen data (9926 of 29 899 points), yet this had only a small effect on the lag-1 autocorrelation—$r(1)$ for the noncloud-cleared data is 0.981. The explanation is that the data points removed by filtering occurred in contiguous chunks and were associated with very low values. In contrast, cloud clearing removed only a small fraction of the LITE data (24 of 943 points) yet had a significant effect on the lag-1 autocorrelation: $r(1)$ for the unfiltered data is 0.93. Here the explanation is that the data points removed by filtering occurred in small segments (one or two points each) and were associated with very high values.

In sum, the variations reported herein can confidently be ascribed to real variability in the atmosphere, not measurement noise. This is true absolutely for the BND and ACE-Asia datasets, which are virtually noise free, and close to true for the cloud-filtered versions of the Spitzbergen and LITE datasets.

2) TEMPORAL VERSUS SPATIAL

Ambient variation is driven both by temporal evolution (emissions, mixing, transformations, removal) and by advection. Any real measurement responds to a combination of these two, though their relative importance varies greatly with the method of sampling and with the spatiotemporal scale of the observations.

Sampling from fixed surface sites provides unambiguous data on temporal variation from an Eulerian perspective (i.e., variation at a fixed point). These same data can be related to spatial variations by making Taylor's frozen atmosphere assumption (see, e.g., Feijt and Jonker 2000), in which temporal evolution is neglected and the scales of turbulence are assumed to be constant. For short timescales (a few hours or less) this assumption is no doubt a good approximation to reality. It begins to break down at the diurnal scale and is completely invalid at the scale of synoptic disturbances. Since diurnal variations in $\sigma_{sp,dry}$ were modest at the BND site and negligible at the (Arctic) Spitzbergen site, and since we are considering here timescales from 1 to about 50 h, it seems reasonable to examine the surface data from a spatial perspective. Without formally transforming the data, we present the temporal and spatial data for a fixed set of scaling lengths that correspond to each other for a typical boundary layer advection velocity of 20 km h^{-1} (5.6 m s^{-1}).

An airplane traveling at 100 m s^{-1} is moving much faster than the boundary layer winds and this is even more true for the space shuttle for which the subsatellite point moves at a ground speed of 7400 m s^{-1}. For both these cases, the distances of interest to this study (10 to a few hundred kilometers) will be sampled in a short enough time that it is reasonable to neglect temporal evolution of the aerosol. Therefore, the ACE-Asia and LITE datasets examined can be taken as providing essentially direct information on spatial variations. For the airborne data, we examine measurements obtained near the surface (altitude $<$ 500 m) in order to study horizontal spatial variations that are most closely related to the temporal data obtained from surface stations. For the spaceborne data, we integrate vertically over the lowest 5 km in order to examine horizontal variations of aerosol optical depth for the lower troposphere. It is important to bear in mind that spatial variations in these datasets are an essentially random mixture of alongwind and crosswind variations. In contrast, subsynoptic-scale variations recorded at the fixed surface stations involve primarily alongwind variations as a body of air advects over the station.

3) DRY AEROSOL VERSUS AMBIENT

All nephelometer data analyzed (BND, Spitz, ACE-Asia) were obtained at low relative humidity (RH $<$ 50%) such that variations in ambient RH do not contribute substantially to the measured variability. Instead, the variability reflects changes in concentration of the dry aerosol. As discussed earlier, ''concentration,'' in this case, refers explicitly to light scattering but is closely related to aerosol mass concentration. The LITE data, on the other hand, reflect *ambient* aerosol extinction that varies not just with dry aerosol concentration but also with changes in ambient RH whenever hygroscopic aerosol components are present.

c. Results

1) DATA OVERVIEW AND DESCRIPTION

To illustrate the nature of mesoscale aerosol variability of interest to this analysis, we present example segments from all four datasets in Fig. 2. For the surface stations, 10-day time series typical of the more polluted times of year—late summer for BND (Fig. 2a) and late

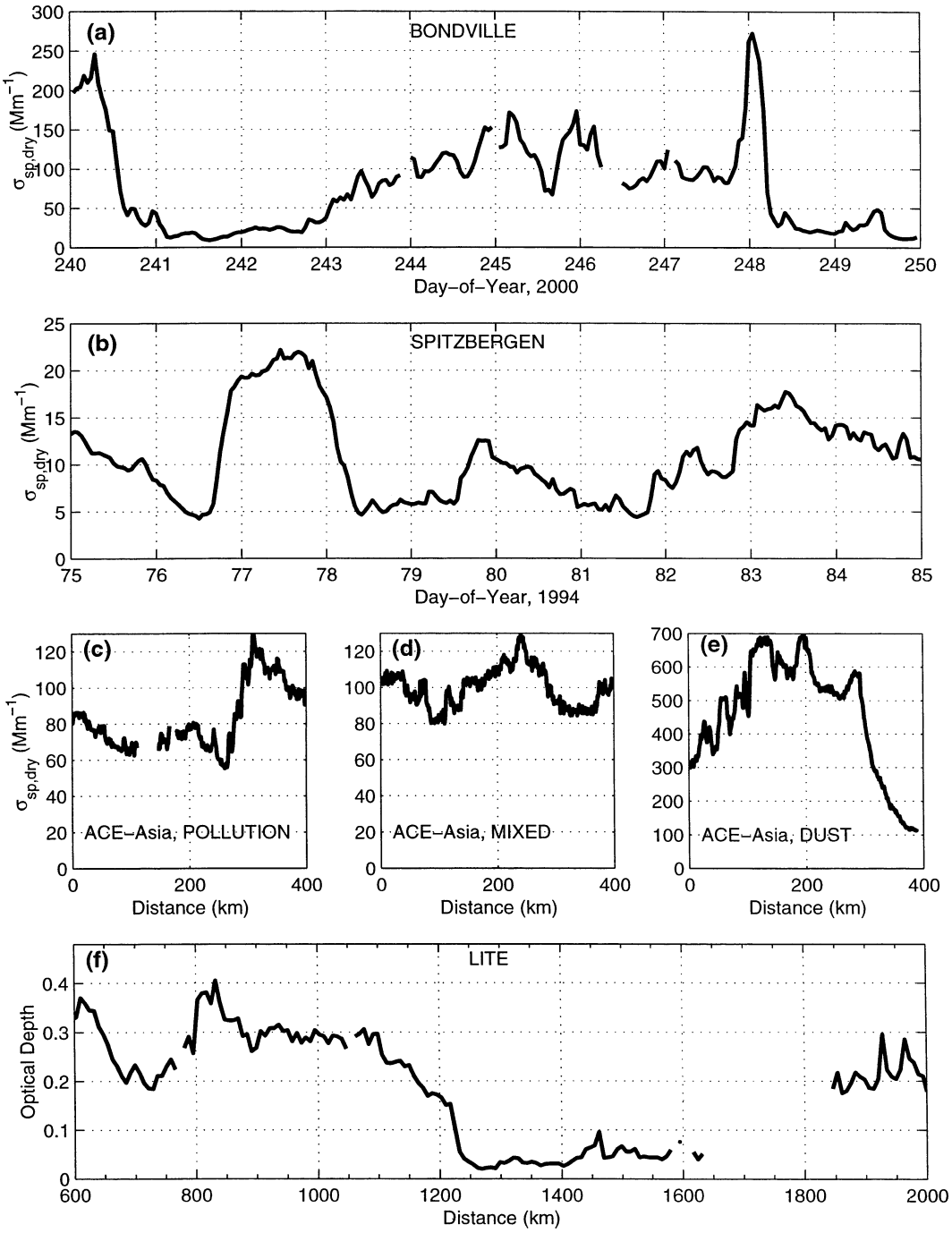

FIG. 2. Example data illustrating variability on scales of hours [(a), (b) temporal data] and tens of km [(c)–(f) spatial data]. For the LITE data, note that an optical depth of 0.2 corresponds to an extinction coefficient of 100 Mm^{-1}, if the particles are contained in a 2-km-thick layer.

winter for Spitzbergen (Fig. 2b)—were chosen. For ACE-Asia, only seven of the level flight legs stretched over 400 km or more. Three of these were selected for display: one from a pollution-dominated aerosol (Fig. 2c), one from a mixed aerosol (Fig. 2d), and one from a dust-dominated aerosol (Fig. 2e). The LITE dataset used here consists of one 7500-km orbital track. We select a 1400-km segment for display (Fig. 2f), which captures a pollution plume extending off the southeastern U.S. coast and out over the Atlantic. Figure 2 shows the potential for large changes over timescales or space scales of a few hours or a few tens of kilometers. In addition, it indicates that aerosol concentration is rarely homogeneous over temporal scales of more

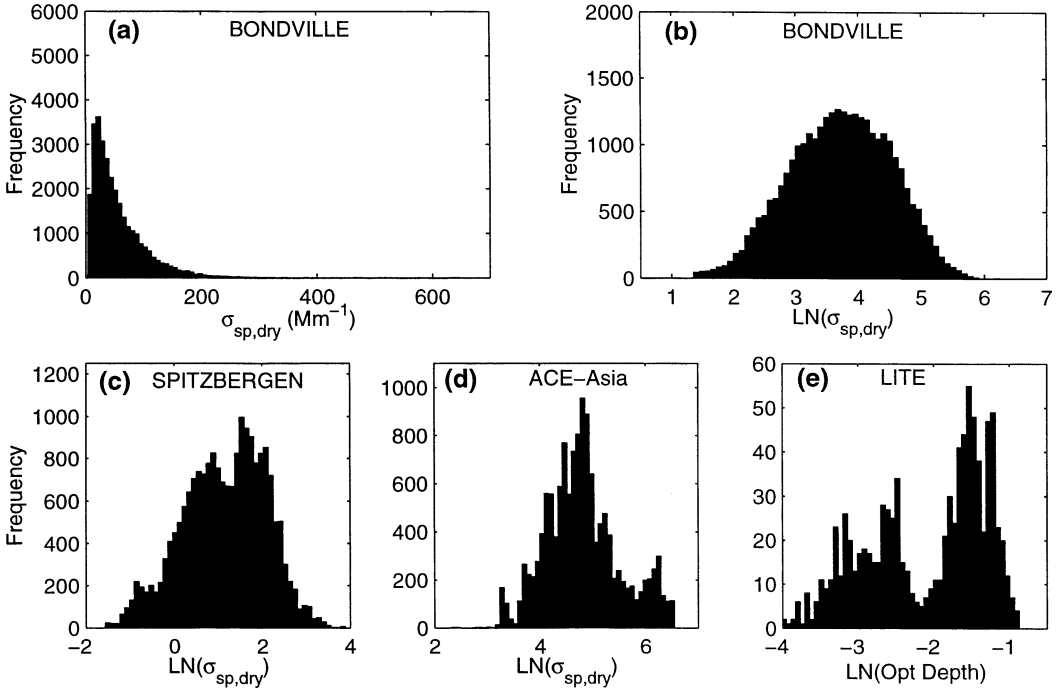

FIG. 3. Frequency-of-occurrence histograms. (a) Histogram of σ_{sp} for Bondville data, showing strongly skewed nature. (b)–(e) Note that histograms of ln(y), where y is σ_{sp} or AOD, are less skewed. Spaceborne data from (e) LITE are clearly bimodal and airborne data from (d) ACE-Asia show a hint of bimodality as well.

than 12 h or spatial scales of more than 200 km. These indications are examined comprehensively below.

Table 1 shows the basic statistical properties of the four datasets, and Fig. 3 presents frequency-of-occurrence histograms. None of the datasets are normally distributed and all are more closely represented by log-normal distributions. This is best displayed by the BND data (Figs. 3a and 3b), but it is also quite apparent in the Spitz data (Fig. 3c). The LITE data analyzed here (Fig. 3e) are clearly bimodal. As shown in the overall time series (Fig. 4), the LITE orbit segment crosses three distinct, synoptic-scale aerosol features with much cleaner regions in between. The bimodal structure of the distribution arises because the transitions between these features are narrow and, perhaps coincidentally, the mean optical depths of the turbid regions are similar. The distribution of the full set of LITE optical depths retrieved from all orbits is unimodal and quasi-lognor-

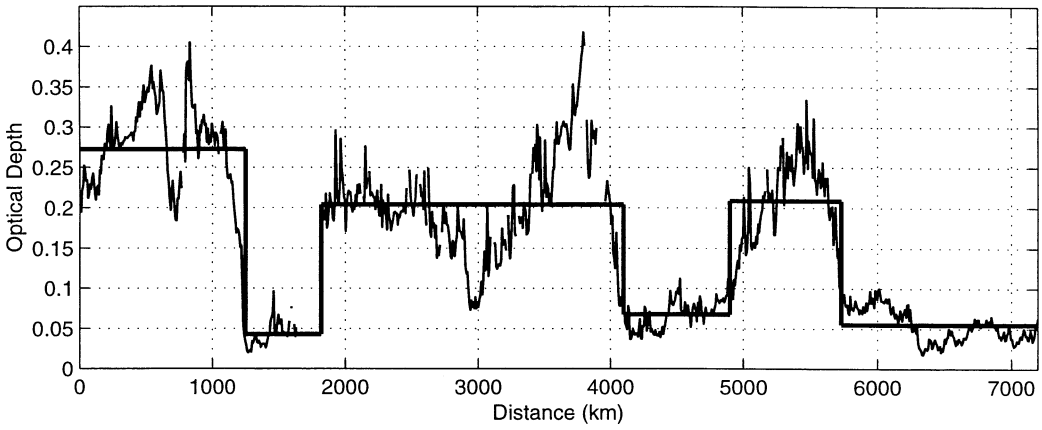

FIG. 4. Entire transect from LITE (described in text and Table 1). The thick line shows the average value of lower tropospheric AOD for six airmass scale features, manually identified in the data. High and low concentration regions tend to have similar concentrations, respectively, and rather sharp boundaries. This produces the bimodal frequency-of-occurrence histogram seen in Fig. 3e and the large-scale structure evident in Fig. 6b. Variability within air masses, the subject of this paper, is significant.

mal. The BND data, from a rural, farming area within a pollution source region, also contain a wide range of levels from highly turbid to quite clean. In contrast to the LITE data, the more extensive dataset from BND shows the full spectrum of intermediate levels and is unimodal (Fig. 3b). The ACE-Asia dataset (Fig. 3d) also shows some indication of bimodality, although in this case the two modes represent turbid and *very* turbid conditions. These data, it should be recalled, were not attained randomly, but are strongly influenced by the deliberate effort in ACE-Asia to sample high aerosol concentration events. Two factors, however, helped to impose a degree of randomness on the sampling: 1) attempts to locate aerosol plumes were sometimes unsuccessful, and 2) various logistical constraints (e.g., flight restrictions, flying transit legs to or from a desired target, or flying in coordination with other platforms) sometimes forced the airplane to sample along a certain track regardless of the amount of aerosol present. Despite these two factors, it is very likely that clean (less turbid) conditions are systematically underrepresented in the ACE-Asia dataset from the C-130 aircraft, such that Fig. 3d is likely to be a truncated version of the true frequency histogram of aerosol concentration for that region of the lower troposphere in spring.

Finally, it is worth noting the wide range of mean concentrations among the various datasets. The geometric mean value of σ_{sp} at Spitzbergen (3.15 Mm^{-1}) is 13 times lower than the value at BND and 39 times lower than the value for ACE-Asia. For the LITE data, the geometric mean value of AOD (0.124) corresponds to an extinction value σ_{ep} of 25 Mm^{-1}, averaged over the 5-km depth for which AOD was determined, or a value of about 60 Mm^{-1} for the typical case where the aerosol was concentrated into a roughly 2-km-deep layer. Despite the range of mean concentrations, the geometric standard deviations are similar for the four datasets—all falling in the range of 2.0 to 2.6. (For a lognormal distribution, a geometric standard deviation of 2 means that 67% of the values will fall within a factor of 2 of the geometric mean.) However, standard deviations reflect both large- and small-scale variability, for example, the Spitz values are dominated by seasonal variations.

2) PRACTICAL MOTIVE FOR ASSESSING MESOSCALE VARIABILITY

Field measurements and comparisons frequently rely on an assumption of aerosol homogeneity within the mesoscale domain (tens to hundreds of kilometers spatially; one to tens of hours temporally). For example, aerosol optical depth measurements by the Langley method (Thomason et al. 1983; Harrison and Michalsky 1994) assume temporal homogeneity of the atmospheric aerosol column over several hours so that the measurements will encompass a range of sun angles. Another example arises when in situ measurements are compared to predictions of chemical transport models (e.g., Benkovitz et al. 1994)—in this case, in situ sampling at one or a few points in space must be assumed to be representative of model grid boxes with horizontal dimensions of 100 km or more (an issue discussed by Benkovitz and Schwartz 1997). Figures 5–8 display the aerosol data in ways that allow the errors associated with assuming mesoscale homogeneity to be examined, pursuant to providing guidelines for defining sampling protocols.

3) POINT VERSUS POINT CORRELATIONS

Figures 5 and 6 address the question of how well a measurement at one point represents the value at another point, offset in time or space. Practical applications arise whenever comparisons or closure exercises involve measurements that are offset in time or space. The visual scatter, indicating a lack of correlation, can be quantified via the lagged correlation, or autocorrelation, coefficient [Eq. (1)]. Results shown in Fig. 5 correspond to r(3 h) for the surface data (Figs. 5a and 5c) and r(60 km) for the ACE-Asia (Fig. 5b) and LITE (Fig. 5d) data. Note that these space/time offsets would be equivalent for an advection velocity of 20 km h^{-1} (5.6 m s^{-1}), a typical boundary layer wind speed. Three-hour variability at BND is seen to be somewhat more significant ($r = 0.86$) than at Spitzbergen ($r = 0.92$). ACE-Asia variability for a 60-km offset is seen to fall between these two values ($r = 0.90$), consistent with the fact that the ACE-Asia sampling region is well downwind of aerosol sources but not as remote as the Spitzbergen station. LITE data are also well downwind of sources and exhibits variability on this scale ($r = 0.915$) similar to both ACE-Asia and Spitzbergen.

This lagged correlation analysis is generalized in Fig. 6 as the autocorrelation function $r(k)$. The surface data (Fig. 6a) can be analyzed over a large range of scales up to 1 yr. The curves are based on hourly data except that, for scales below 1 h at BND, 1-min-resolution data were used. A modest diurnal cycle is evident at BND by the slight flattening of the curve at a lag time of 1 day. Autocorrelation drops below 0.5 at about 1 day at BND and 1 week at Spitzbergen, indicating that most variation exists below these scales. However, a strong annual cycle is evident at Spitzbergen (hazy conditions occur almost exclusively during the winter and early spring). This is revealed by the peak in correlation at a lag of 1 yr and the mirror-image anticorrelation at a lag of 6 months. Large-scale structure, such as an annual cycle, influences the autocorrelation at smaller scales. In effect, data within any given season are more positively correlated if there is a large seasonal contrast. To study this effect, autocorrelation was calculated for both stations using data from just one season: July–September at BND (Fig. 6a, x's) and January–March at Spitzbergen (Fig. 6a, circles). The choice was arbitrary for BND since there is no significant seasonal cycle. Ac-

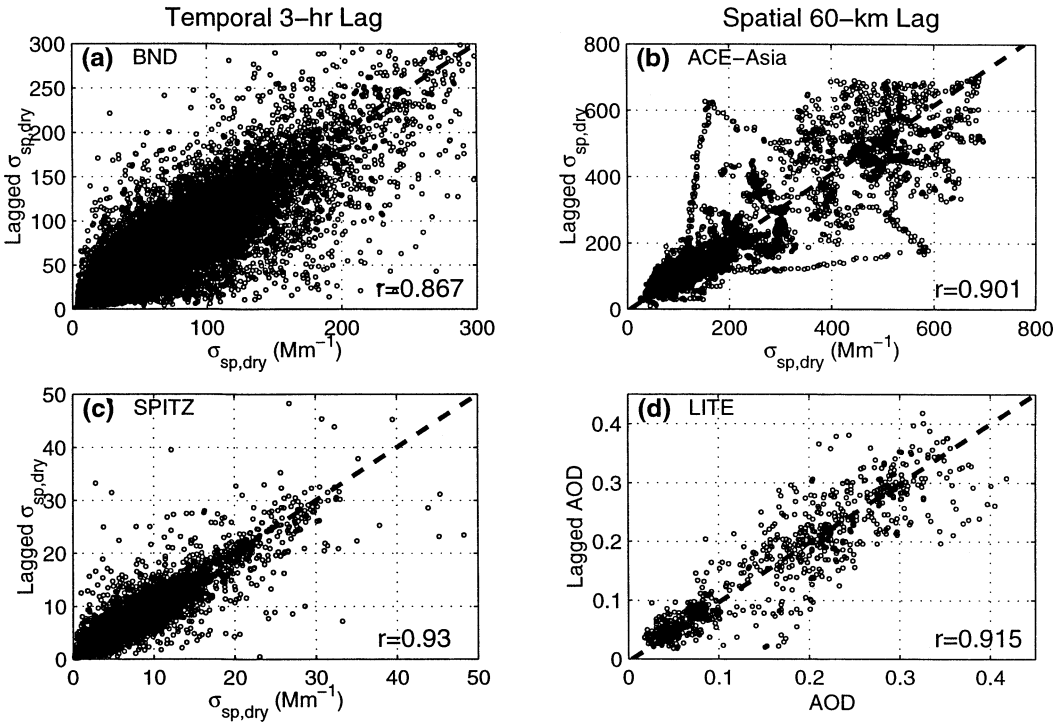

FIG. 5. Point vs point correlations for the four datasets at a lag offset of 3 h [(a), (c) ground-based data] or 60 km [(b) airborne data and (d) spaceborne data]. The correlation coefficient shown on each panel is identical to the autocorrelation statistic defined in Eq. (1). Note that the temporal and spatial lags would correspond for an advection speed of 20 km h^{-1}.

cordingly, no significant change in $r(k)$ was found. For Spitzbergen, we chose the season with the highest concentrations in order to minimize the influence of instrumental noise. Here the change is quite significant: $r(k)$ falls off much more rapidly and becomes rather similar to that for BND. For example, like BND, most of the variation is seen to exist at scales below 1 day in this analysis of winter-only data.

To facilitate comparison between the temporal and spatial datasets (Figs. 6a and 6b, respectively), the x axes have been scaled equivalently for an advection velocity of 20 km h^{-1}. (Recall the caveats in section 2b(2) regarding this comparison.) The spatial data allow analysis at somewhat smaller scales (down to 1 km in the case of ACE-Asia) but does not cover anywhere near the range of large scales encompassed by the temporal datasets. The LITE data exhibits synoptic-scale periodicity indicated by a strong peak in correlation at a lag of 1700 km and a corresponding dip at 850 km (note that the scale of the x axis in Fig. 6b requires multiplication by 20 to convert to kilometers). This large-scale structure is evident as well in the overall time series shown in Fig. 4 and corresponds to the fact that the satellite passed over three distinct, synoptic-scale aerosol plumes with relatively clean air in between. We have separately analyzed LITE variability using only data from within the synoptic-scale aerosol plumes (Fig. 6b, circles). Because the overall range of variation is significantly reduced, the autocorrelation falls off much more rapidly. This reduction of variation is somewhat artificial in that plume boundaries were selected based on aerosol concentration; however, the time series plot (Fig. 4) makes a convincing case for the reality of these synoptic-scale plumes.

Taken together, the various autocorrelation functions shown in Fig. 6 present a reasonably coherent picture of mesoscale aerosol variations. All the curves exhibit 1) a fairly flat region for scales below 3–6 h (temporal) or 40–160 km (spatial), and 2) a rapid falloff at larger scales out to lags of a few days (temporal) or several hundred kilometers (spatial). In other words, each dataset provides strong evidence for the presence of significant, mesoscale variations in aerosol concentration. This evidence is quantified in Table 2, which presents the offset times and distances corresponding to $r(k) = 0.9$ and $r(k) = 0.8$ for each of the seven curves shown in Fig. 6. The practical significance of these r values will be discussed in section 3b.

4) AVERAGE VERSUS POINT CORRELATIONS

Figures 7 and 8 address the question of how well a point measurement represents the average over some region of space or time. Practical applications arise whenever comparisons or closure exercises involve different time or space resolution among the various meth-

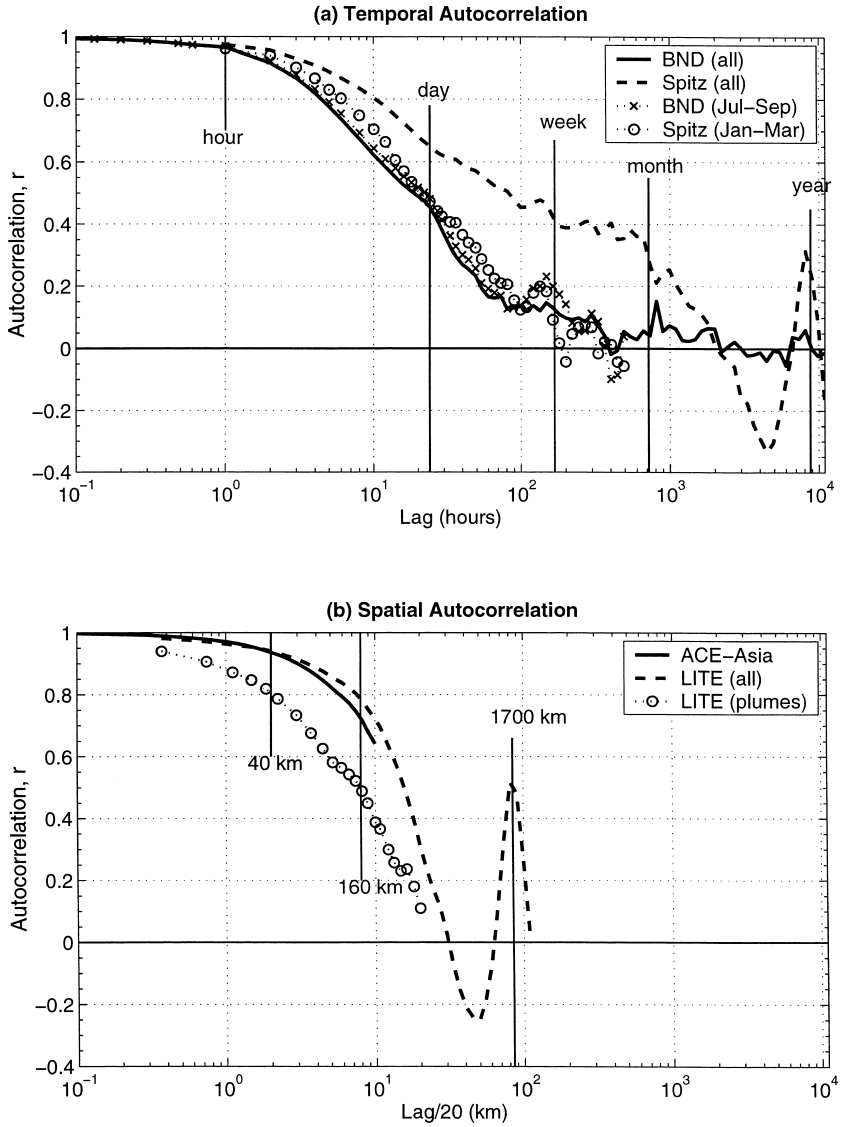

FIG. 6. Autocorrelation statistic [Eq. (1)] as a function of lag for the four datasets. (a) Temporal autocorrelation from surface stations for all data and for selected seasons—the latter calculations remove large-scale variation associated with the seasonal cycle. (b) Spatial autocorrelation from the airborne and satellite-borne (LITE) datasets. For LITE, results are shown for all data and for just the synoptic-scale plumes, where the latter calculation is intended to remove large-scale variation evident at a lag of 1700 km and evident in the time series (Fig. 4). The x axes in the two panels are scaled equivalently for an advection speed of 20 km h^{-1}.

ods. An important application (Fig. 7) is how well a filter sample obtained at a surface station over an extended period of time (typically, 24 h) represents the aerosol sensed at one instant of time by a polar-orbiting satellite (which covers several kilometers of ground per second.) Three cases are examined for each of the two surface datasets: 1) how well a 3-h average represents the midpoint of the 3-h period (Figs. 7a and 7d), 2) how well a 24-h average represents the midpoint of a 24-h period (Figs. 7b and 7e), and 3) how well a 24-h average represents the second hour of the 24-h period (Figs. 7c and 7f). The last case would represent, for example, a situation where 24-h filters were changed at local noon and related to a satellite overpass at 1400 LST. Figure 7 shows that a 3-h integrated sample is very well correlated to the midpoint, while correlations associated with 24-h sampling are significantly degraded, especially for the noncentered case.

Figure 8 generalizes the results of Fig. 7 and extends this analysis to the spatial datasets. First, we calculate the local coefficient of variation (LCOV) for data segments of various lengths, k,

$$\mathrm{LCOV}(k, i) = s\{x_i, \ldots, x_{i+k}\}/m\{x_i, \ldots, x_{i+k}\}, \quad (2)$$

TABLE 2. Autocorrelation summary. Temporal and spatial offsets associated with autocorrelation values r of 0.9 and 0.8 are shown. Parenthetical values are calculated assuming an advection velocity of 20 km h^{-1}.

	Temporal (h)		Spatial (km)	
	$r = 0.9$	$r = 0.8$	$r = 0.9$	$r = 0.8$
BND	2.3	4.4	(46)	(88)
BND (Jul–Sep)	2.5	4.8	(50)	(95)
Spitz	4.4	10.3	(88)	(207)
Spitz (Jan–Mar)	3.0	6.1	(60)	(122)
ACE-Asia	(3.2)	(5.8)	64	116
LITE	(3.5)	(7.6)	70	151
LITE (plumes)	(0.8)	(2.1)	16	41

where $m\{\ \}$ and $s\{\ \}$ are the mean and standard deviation, respectively, of the bracketed data segments. For selected segment lengths (see legend of Fig. 8) this statistic is calculated for every i from N to $N - k$, where N is the total number of data points. To avoid bias due to data gaps, we calculate LCOV only for segments that have at least half of the potential number of data points. Figure 8 displays the cumulative probability distributions of LCOV (k, i) for each segment length k. Each point along a given curve represents the probability, on the y axis, that LCOV for data segments of length k is less than the value indicated on the x axis. Symbols used for the curves follow a consistent convention among all four panels, based on space–time equivalency for an advection velocity of 20 km h^{-1}. For example, the heavy solid line with upward-pointing triangles represents a segment length of 6 h (temporal) and 120 km (spatial). For Fig. 8a (BND), the first three curves (segments lengths of 22, 45, and 90 min) are derived from the 1-min-resolution data. Note the smooth transition from minute to hourly data, indicating that 1-h averaging does not smooth out variations at the 3-h or larger timescales.

To illustrate the meaning and use of Fig. 8, each panel has a star at the location [$x = 0.1$, $y = 0.5$], the point representing 50% probability ($y = 0.5$) that data segments will have a COV of 0.1 or less. This location denotes a plausible criterion for designing a field experiment where an assumption of aerosol homogeneity must be made. (A COV of 0.1 for a random, normal distribution implies that any given point along the segment has a 70% chance of being within 10% of the mean and a 95% chance of being within 20% of the mean.) Each panel in Fig. 8 is annotated with the segment length corresponding to this criterion, derived by interpolation from the adjacent curves. This characteristic length is 3.4 h at the polluted BND station, and 6.5 h at the remote Spitzbergen station, reflecting greater small-scale variability at BND. For the spatial data, the characteristic length is 90 km for ACE-Asia and 57 km for LITE, indicating somewhat more small-scale variability in the AOD_{lt} measurements from LITE than

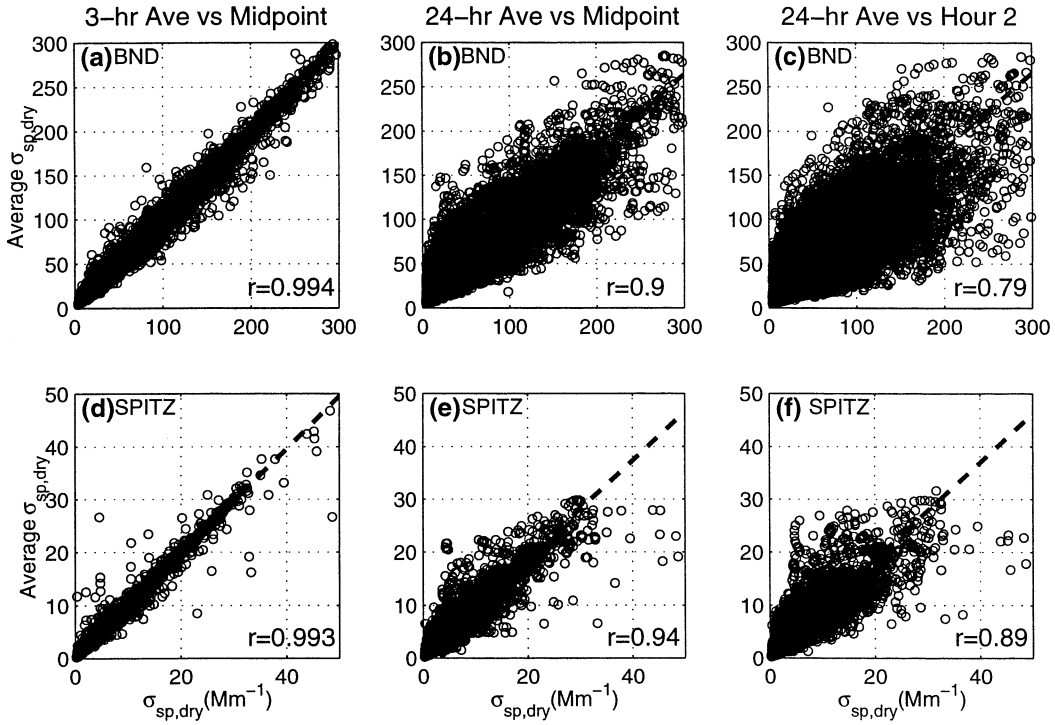

FIG. 7. Average vs point correlations for ground-based data. (a), (d) 3-h average σ_{sp} vs σ_{sp} for the middle hour of the 3-h period. (b), (e) Same except 24-h average vs midpoint. (c), (f) 24-h average σ_{sp} vs σ_{sp} for the second hour during the 24-h period.

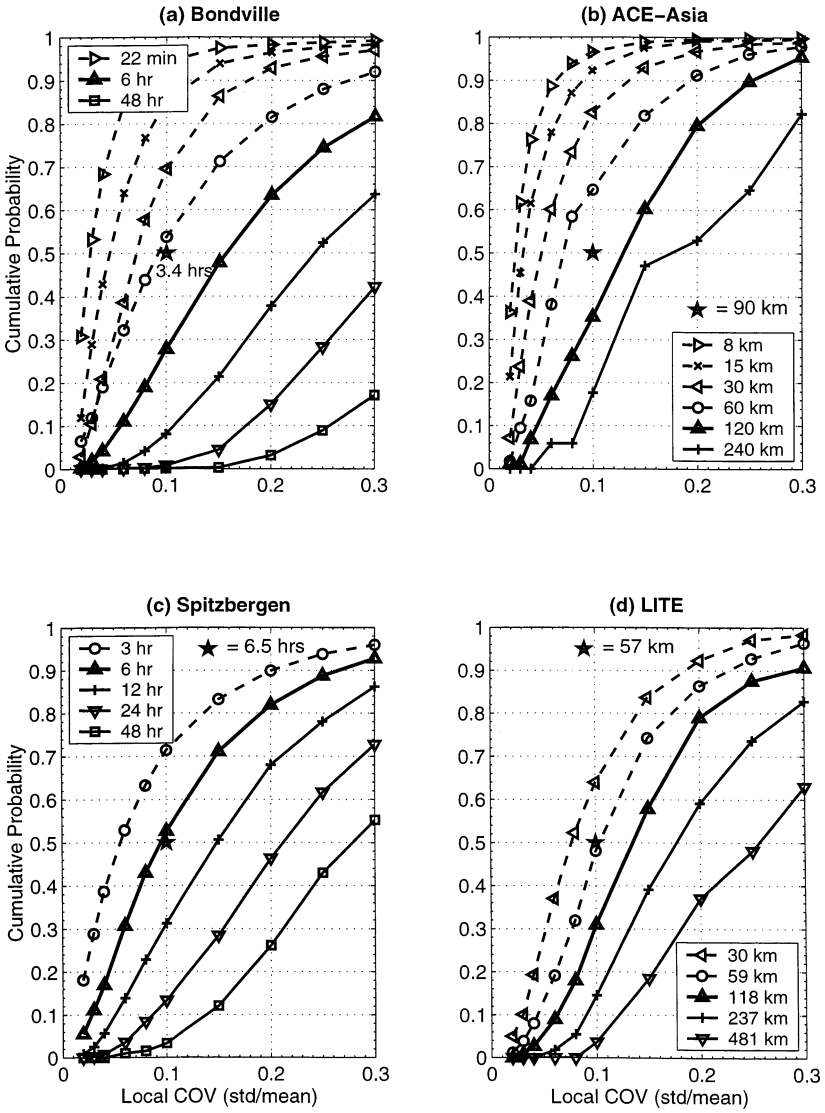

FIG. 8. Local coefficient of variation for the four datasets. Cumulative probability plots of LCOV(k, i)—Eq. (2)—for various segment lengths k. The heavy solid line with triangles on all panels corresponds to a 6-h segment length [(a), (c) temporal data] or 120 km [(b), (d) spatial data]. Line styles for all other segment lengths are matched as well, as shown in the legends. A truncated legend is shown in (a); the curves are right triangle (22 min); × (45 min); left triangle (90 min); circle (3 h); up triangle (6 h); + (12 h); down triangle (24 h); square (48 h). The first three curves on (a) are derived from 1-min-resolution data. The star on each panel at location [0.1, 0.5] indicates a 50% probability that data segments will have a COV of 0.1. Each panel is annotated with the segment length corresponding to this location, derived by interpolation from the adjacent curves. As explained in the text, this illustrates the use of the LCOV statistic for establishing experimental design criteria.

in the in situ, low-RH scattering measurements from ACE-Asia.

Note that the LCOV statistic is a pure measure of small-scale variability, unaffected by large-scale variations like the seasonal cycle. As such, it provides a less ambiguous means of comparing the four datasets than does the autocorrelation function. While the different datasets depicted in Fig. 8 contain information on different scales, these scales largely overlap. Overall, a remarkably coherent picture of mesoscale aerosol variation emerges from this analysis. For example, the probability that LCOV is 0.2 or less goes from very high to very low over temporal scales from 3–24 h and over the corresponding spatial scales of 60–480 km. Thus, all datasets indicate the presence of substantial variability associated with these scales.

3. Discussion

a. Nature of aerosol variability

Using high-resolution optical measurements, we have displayed mesoscale aerosol variability in several different ways in order to seek insights into its causes. Frequency-of-occurrence histograms (Fig. 3) reveal the lognormal nature of the data, which is consistent with concentration being controlled by multiple dilution events or by other multiplicative processes (Ott 1990). However, the histogram is clearly bimodal in the case of the spaceborne data, which reflects the small number and sudden transitions between airmass regions sampled during this transect (Fig. 4). The LCOV statistics (Fig. 8) show little variability below 20 km but a rapid increase in variability over scales from 20 to a few hundred kilometers. This could make sense in terms of the scales of major aerosol sources (e.g., dust storms, biomass fires, megacities) and the major sink for aerosol (precipitation), which all have physical dimensions around 20 to 200 km. [The mesoscale structure of precipitation is known from radar studies, beginning with Austin and Houze (1972).] Our data support the notion that aerosol plumes on the airmass scale (several hundred to a few thousand kilometers) do exist in the atmosphere, but are not internally homogeneous. This is illustrated, for example, by the LITE transect shown in Fig. 4. Mesoscale inhomogeneities, in turn, indicate that horizontal mixing processes are slow compared to aerosol residence times in the lower troposphere. Another factor may be vertical mixing from beyond the boundary layer, which promotes horizontal inhomogeneity due to large concentration gradients in the vertical. The data herein show that, on scales larger than a few hours or a few tens of kilometers, aerosol concentration is almost always rising or falling (or, in other words, rarely steady.) This is revealed qualitatively by the time series plots (Fig. 2). This effect is also revealed by the fact that the average versus point correlations fall off significantly when the point is not at the center of the averaging period (Figs. 7c and 7f).

b. Practical consequences for the integration of satellite retrievals, in situ measurements, and models

Based on data from diverse regions, we have shown that significant aerosol variability exists below the synoptic (or airmass) scales and, thus, below the scales that are readily resolved by chemical transport models and by measurements with long sample times. Further, the data analyzed here imply that the statistics of this variability may be largely independent of time and place. Figures 6 and 8 provide a preliminary, general model of mesoscale aerosol variability in the lower troposphere. They are intended to allow investigation of the consequences of aerosol variability for coordinated observations to assess regional-to global-scale aerosol effects. Figure 6 provides guidance relevant to point versus point comparisons, while Fig. 8 provides guidance relevant to average versus point comparisons.

Many existing and planned satellite sensors are capable of sensing aerosol properties and, thus, of contributing to a global assessment of atmospheric loadings and radiative effects. However, as mentioned in the introduction, aerosol properties retrieved from these instruments require careful validation, and these validation intercomparisons must consider the spatial and temporal variability of the aerosol. A critical, initial coordination activity for aerosol-sensing satellites is to demonstrate that they are indeed responding to the same phenomenon as can be measured with more accuracy and detail within the atmosphere. There are two essential aspects to this demonstration: 1) show correlation between measurements made within and beyond the atmosphere, and 2) show closure regarding the *magnitude* of the respective measurements. It is perhaps more compelling to state this in the negative. If we *cannot* demonstrate this fundamental, zeroeth-order coherence, we have no logical basis for integrating the two types of measurements. Such integration is essential, however, because neither the measurements made within the atmosphere nor those made by satellites can, by themselves, address the global and chemical nature of the aerosol problem.

As a specific example, consider the design of an experiment to test coherence between spaceborne lidar measurements and in situ optical/chemical measurements. To undertake this zero-order coupling, we must know what degree of correlation is possible for various space–time offsets. Autocorrelation of in situ optical measurements, as quantified herein, provides the limiting (best) case, since additional degradation will occur due to measurement noise and artifacts (e.g., satellite measurements may be affected by detector noise, thin cirrus, or surface reflectivity, while in situ measurements may be affected by particle losses at the inlet or particle volatilization due to sample heating). This limiting case is characterized as follows: good correlation ($r > 0.9$) occurs for time (space) offsets less than about 3 h (60 km); acceptable correlation ($r > 0.8$) occurs for time (space) offsets less than about 6 h (120 km). These results translate directly to required sampling protocols. Assuming that space/time offsets are the *only* source of error (i.e., neglecting measurement noise and artifacts), the number of random, independent samples required to establish a statistically significant correlation (i.e., 95% confidence) would be five in the former case ($r = 0.9$) and seven in the latter case ($r = 0.8$). Similarly, to determine the linear relationship between the two measurements such that the standard error of the slope is less than 10% would require six random, independent samples in the former case and 14 in the latter case. To determine the slope to within a standard error of 5% would require 20 and 54 samples, respectively. This example illustrates that closely collocated samples (e.g.,

such that $r = 0.9$) are potentially several times more valuable than samples that are only moderately well collocated (e.g., where $r = 0.8$).

As another example, variability on the scale of 100 km suggests that vertical profiles of aerosols by aircraft should be conducted so that the horizontal dimension of the flight path is at most a few tens of kilometers. This is necessary to ensure that observed variations are indeed in the vertical and not an expression of horizontal variability. Profile data acquired during take-offs and landings, for example, may require re-examination in this regard.

c. Comparison to previous studies

Benkovitz et al. (1994) used a chemical transport model to examine the temporal and spatial autocorrelation of instantaneous sulfate aerosol column burden over the northern Atlantic Ocean. The domain of investigation was well away from continental sources. Benkovitz et al. (1994) defined a characteristic scale as the time or distance required for the autocorrelation function to fall below 1/*e*. The average characteristic scales were found to be 10 h (temporal) and 900 km (spatial). Interestingly, both these values are at approximately 10 times the model resolution of 1 h (temporal) and 125 km (spatial). Compared to these model predictions, our analysis of autocorrelation reveals considerably longer characteristic times (30 h or more; Fig. 6a) and somewhat shorter characteristic distances (400 km or less; Fig. 6b). Many factors could contribute to these discrepancies: differences in the analyzed parameter, differences in location and time, lack of spatial resolution in the model, etc. Nevertheless, the discrepancies appear to be significant. They point to the importance (and the practicality) of testing the veracity of chemical transport models not just in terms of concentration but also in terms of scales of variability.

Benkovitz and Schwartz (1997) examined 24-h sulfate measurements from surface stations in North America and Europe with regard to ''subgrid scale'' spatial variability—that is, variability below the 125-km resolution of their chemical transport model. They defined a spread ratio $S_{o/o}$ as $\exp[2s(\ln(x)]$, where x represents a set of simultaneous measurements within a model grid box, $\ln(x)$ refers to the natural log, and s refers to the standard deviation. (This $S_{o/o}$ statistic is the square of the geometric standard deviation statistic that we report in Table 1c.) Over the 4-month study period, 503 calculations of $S_{o/o}$ could be made with the available data and these yielded a median value of 1.5. We can compare this median value to the same quantity calculated from our data over segments lengths of 120 km (spatial) and 6 h (temporal). The median values of $S_{o/o}$ for each dataset are as follows: 1.37 (BND), 1.21 (Spitz), 1.26 ACE-Asia, 1.31 (LITE). The systematically lower values (indicating less variability) that emerge from our datasets constitute a significant discrepancy, since the 24-h averaging used in the Benkovitz and Schwartz (1997) dataset ought to act as a smoothing function. A possible explanation is that the Benkovitz and Schwartz (1997) data were mostly from urban and suburban locations (mostly in the United States) and, as such, may be systematically much closer to local aerosol sources than any of the datasets examined herein. The BND data (from rural Illinois) provide the most legitimate comparison and give a roughly comparable median value.

Kaufman et al. (2000) examined ground-based AOD data from the Aerosol Robotic Network (AERONET) to determine if measurements at a single time of day are representative, in a climatological sense, of daily averaged AOD. (The motive was to assess the representativeness of AOD measurements by sun-synchronous, polar-orbiting satellites.) Included in their study (as Fig. 1) was a plot of the ratio of 90-min-averaged AOD to whole-day AOD for each individual observation. This ratio was centered very nearly on 1.0 (indicating little overall bias) with a standard deviation of about 0.20 over most of the range of measured values of AOD. This analysis is quite similar to the point versus average results shown herein (Fig. 7). However, to make a more direct comparison, we have calculated, for both surface stations, the ratio of 1-h measurements of $\sigma_{sp,dry}$ centered on 1030 LST to 12-h averages centered on local noon. The standard deviation of these daily ratios is 0.24 for BND and 0.17 for Spitzbergen, quite consistent with the findings of Kaufman et al. 2000. This consistency could be coincidental, since the respective measurements were made at entirely different locations. However, it could also be taken as additional evidence (beyond that provided by the LITE data analyzed herein) that the variability of low-RH aerosol light scattering at carefully selected surface stations is closely related to the variability of tropospheric AOD.

4. Conclusions and future work

We have analyzed mesoscale variations of aerosol optical properties (in situ light scattering and optical depth) in the lower troposphere. Using, for example, an autocorrelation criterion of 0.8, our data suggest that coherent timescales and space scales for aerosol concentration are less than 10 h and 200 km, respectively—in some cases much less (Fig. 6 and Table 2). Temporal variation, herein, refers to variation at a fixed point (not evolution in a Lagrangian sense) and is thus closely related to spatial variation through the mechanism of advection. This finding of significant mesoscale variability appears to apply generally in the lower troposphere, since it is observed in diverse datasets (surface polluted, surface remote, aircraft, and spaceborne). This result undermines the airmass paradigm (implicit in most attempts to understand regional- to global-scale aerosol effects) and calls for a physical explanation. We tentatively suggest that the explanation may lie in the mesoscale nature of both major aerosol sources (e.g.,

dust storms, biomass fires, megacity plumes) and the major sink [i.e., cellular precipitation; Austin and Houze (1972)], combined with the likelihood that aerosol residence times in the lower troposphere are too short for horizontal mixing processes to eliminate mesoscale inhomogeneities. Sporadic vertical transport (e.g., intrusions of clean air from the free troposphere into the planetary boundary layer) is likely to be another important factor.

We suggest that mesoscale aerosol variability imposes strong constraints on the design of field measurements, especially where multiple platforms or approaches with differing time/space resolution are being employed. Practicle implications include 1) in situ measurements intended to demonstrate coherence and quantiative closure with satellite measurements should be collocated within a few hours temporally and a few tens of kilometers horizontally, 2) in situ measurements used for validation of models should be appropriately matched to the timescales and space scales of the models, 3) models using horizontal grids of 100 km or more need to develop ways of accounting for subgrid scale variations and correlations, and 4) large-scale or long-term averages of either in situ or satellite data should be used with caution, not because it will necessarily be biased but because such averaging will tend to degrade correlations and mask important relationships.

The present study considers extensive aerosol properties only (i.e., properties that are proportional to aerosol concentration). Preliminary analysis suggests that intensive aerosol properties are more homogeneous at the mesoscale; however, this cannot be properly addressed with the methods and the datasets used herein. Thus, we defer this question to a future paper.

Several other recommendations for the future can be made. First, it is critical that other aerosol datasets be examined with an eye to confirming or refuting the generality of the mesoscale variations found herein. Second, to provide better insights into causes, it would be helpful to include in the analysis gas-phase constituents with relatively well-known lifetimes, such as CO, CO_2, SO_2, Rn, and CFCs. Third, it will be important to look not just for variations but also for correlations among variables. As an example, Charlson et al. (1999) have pointed out that a correlation between sulfate aerosol concentration and ambient relative humidity could substantially enhance radiative forcing by this hygroscopic component (and, conversely, an anticorrelation would diminish radiative forcing.) Fourth, there is a need for airborne experiments that are explicitly designed to study horizontal variability. Key design elements (which are not typically part of airborne observations) would include 2D surveys (x and y), repetitive sampling of the same location, and some portion of the flight hours devoted to random sampling. In addition, it is critical that statistically significant amounts of data be acquired. Finally, we have stressed the importance of taking mesoscale variations into account when designing in situ observations intended for coupling with either satellite sensors or chemical transport models. Substantial efforts will be required to devise optimal measurement strategies and to mesh these with practical (e.g., logistical and economic) constraints.

Acknowledgments. We thank Robert Wood and an anonymous reviewer for useful suggestions. This work was supported by the National Science Foundation (ATM-0002198), the National Aeronautic and Space Administration through its CALIPSO satellite mission and its Global Aerosol Climatology Project, and by the National Oceanic and Atmospheric Administration (NOAA) through the Joint Institute for the Study of the Atmosphere and Oceans (JISAO) under NOAA cooperative agreement NA17RJ1232.

REFERENCES

Anderson, T. L., D. S. Covert, J. D. Wheeler, J. M. Harris, K. D. Perry, B. E. Trost, and D. J. Jaffe, 1999: Aerosol backscatter fraction and single scattering albedo: Measured values and uncertainties at a coastal station in the Pacific NW. *J. Geophys. Res.,* **104,** 26 793–26 807.

——, S. J. Masonis, D. S. Covert, R. J. Charlson, and M. J. Rood, 2000: In-situ measurement of the aerosol extinction-to-backscatter ratio at a polluted, continental site. *J. Geophys. Res.,* **105,** 26 907–26 915.

Austin, P. M., and R. A. Houze, 1972: Analysis of the structure of precipitation patterns in New England. *J. Appl. Meteor.,* **11,** 926–935.

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 1986–1987. *J. Geophys. Res.,* **102,** 25 305–25 338.

——, R. C. Easter, S. Nemesure, R. Wagener, and S. E. Schwartz, 1994: Sulfate over the North Atlantic and adjacent continental regions: Evaluation for October and November 1986 using a three-dimensional model driven by observation-derived meteorology. *J. Geophys. Res.,* **99,** 20 725–20 756.

Bolin, B., and H. Rodhe, 1973: A note on the concepts of age distribution and transit time in natural reservoirs. *Tellus,* **25,** 58–62.

Brown, R. A., 1980: Longitudinal instabilities and secondary flows in the planetary boundary layer: A review. *Rev. Geophys.,* **18,** 683–697.

Charlson, R. J., and H. Rodhe, 1982: Factors controlling the acidity of natural rainwater. *Nature,* **295,** 683–685.

——, A. H. Vanderpol, D. S. Covert, A. P. Waggoner, and N. C. Ahlquist, 1974: $H_2SO_4/(NH_4)_2SO_4$ background aerosol: Optical detection in St. Louis region. *Atmos. Environ.,* **8,** 1257–1267.

——, S. E. Schwartz, J. M. Hales, R. D. Cess, J. A. Coakley, J. E. Hansen, and D. J. Hofmann, 1992: Climate forcing by anthropogenic aerosols. *Science,* **255,** 423–430.

——, T. L. Anderson, and H. Rodhe, 1999: Direct climate forcing by anthropogenic aerosols: Quantifying the link between radiation and sulfate models. *Beitr. Phys. Atmos.,* **72,** 79–94.

Clarke, A. D., S. Howell, P. K. Quinn, T. S. Bates, J. A. Ogren, E. Andrews, A. Jefferson, and A. Massling, 2002: The INDOEX aerosol: A comparison and summary of chemical, microphysical, and optical properties from land, ship, and aircraft. *J. Geophys. Res.,* in press.

Delene, D. J., and J. A. Ogren, 2002: Variability of aerosol optical properties at four North American surface monitoring sites. *J. Atmos. Sci.,* **59,** 1135–1150.

Feijt, A., and H. Jonker, 2000: Comparison of scaling parameters

from spatial and temporal distributions of cloud properties. *J. Geophys. Res.*, **105,** 29 089–29 097.
Ferrare, R. A., D. D. Turner, L. H. Brasseur, W. F. Feltz, O. Dubovik, and T. P. Tooman, 2001: Raman lidar measurements of the aerosol extinction-to-backscatter ratio over the Southern Great Plains. *J. Geophys. Res.*, **106,** 20 333–20 347.
Galloway, J. N., G. E. Likens, and M. Hawley, 1984: Acid precipitation: Natural versus anthropogenic components. *Science*, **226,** 829–831.
Glickman, T. S., Ed., 2000: *Glossary of Meteorology*. American Meteorological Society, 2d ed., 855 pp.
Hamrud, M., 1983: Residence time and spatial variability for gases in the atmosphere. *Tellus*, **35B,** 295–303.
Harrison, L., and J. Michalsky, 1994: Objective algorithms for the retrieval of optical depths from ground-based measurements. *Appl. Opt.*, **33,** 5126–5132.
Heintzenberg, J., and L. Bäcklin, 1983: A high sensitivity integrating nephelometer for airborne air pollution studies. *Atmos. Environ.*, **17,** 433–436.
——, and C. Leck, 1994: Seasonal variation of the atmospheric aerosol near the top of the marine boundary layer over Spitsbergen related to the Arctic sulphur cycle. *Tellus*, **46B,** 52–67.
Hobbs, P. V., 2002: Clean air slots amid atmospheric pollution. *Nature*, **415,** 861.
Huebert, B. J., G. L. Lee, and W. L. Warren, 1990: Aerosol inlet passing efficiency measurement. *J. Geophys. Res.*, **95,** 16 369–16 381.
IPCC, 2001: *Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the International Panel on Climate Change*. Cambridge University Press, 572 pp.
Isaaks, E. H., and R. M. Srivastava, 1989: *An Introduction to Applied Geostatistics.* Oxford University Press, 561 pp.
Jobson, B. T., S. A. McKeen, D. D. Parrish, F. C. Fehsenfeld, D. R. Blake, A. H. Goldstein, S. M. Schauffler, and J. W. Elkins, 1999: Trace gas mixing ratio variability versus lifetime in the troposphere and stratosphere: Observations. *J. Geophys. Res.*, **104,** 16 091–16 113.
Junge, C. E., 1974: Residence time and variability of tropospheric trace gases. *Tellus*, **26,** 477–487.
——, 1975: The possible influence of aerosols on the general circulation and climate and possible approaches for modeling. *The Physical Basis of Climate and Climate Modelling*, GARP Publication Series, Vol. 16, WMO–ICSU, 244–251.
Kaufman, Y. J., B. N. Holben, D. Tanre, I. Slutsker, A. Smirnov, and T. F. Eck, 2000: Will aerosol measurements from *Terra* and *Aqua* polar orbiting satellites represent daily aerosol abundance and properties? *Geophys. Res. Lett.*, **27,** 3861–3864.
Kent, G. S., C. R. Trepte, K. M. Skeens, and D. M. Winker, 1998: LITE and SAGE II measurements of aerosols in the Southern Hemisphere upper troposphere. *J. Geophys. Res.*, **103,** 19 111–19 127.
Koloutsou-Vakakis, S., M. J. Rood, A. Nenes, and C. Pilinis, 1998: Modeling of aerosol properties related to direct climate forcing. *J. Geophys. Res.*, **103,** 17 009–17 032.
——, C. M. Carrico, Z. Li, M. J. Rood, and J. A. Ogren, 1999: Aerosol properties and radiative forcing at an anthropogenically perturbed midlatitude Northern Hemisphere continental site. *Phys. Chem. Earth Pt. C.*, **24,** 541–546.
Lafleur, B. G., 1998: A low turbulence inlet for airborne aerosol sampling. M.S. thesis, University of Denver, 99 pp.
Likens, G. E., Ed., 1981: *Some Perspectives of the Major Biogeochemical Cycles*. Wiley and Sons, 170 pp.
Masonis, S. J., K. Franke, A. Ansmann, D. Mueller, D. Althausen, J. A. Ogren, A. Jefferson, and P. J. Sheridan, 2002: An intercomparison of aerosol light extinction and 180° backscatter as derived using in situ instruments and Raman lidar during the INDOEX field campaign. *J. Geophys. Res.*, **107** (D19), 8014, doi: 10.1029/2000JD000035.
Michaels, R. A., and M. T. Kleinman, 2000: Incidence and apparent health significance of brief airborne particle excursions. *Aerosol Sci. Technol.*, **32,** 93–105.
Mishchenko, M. I., I. V. Geogdzhayev, B. Cairns, W. B. Rossow, and A. A. Lacis, 1999: Aerosol retrievals over the ocean by use of channels 1 and 2 AVHRR data: Sensitivity analysis and preliminary results. *Appl. Optics*, **38,** 7325–7341.
Mourad, P. D., and B. A. Walter, 1996: Analysis of mesoscale linear features observed in the Arctic atmospheric boundary layer. *Mon. Wea. Rev.*, **124,** 1924–1940.
Ott, W. R., 1990: A physical explanation of the lognormality of pollutant concentrations. *J. Air Waste Manage. Assoc.*, **40,** 1378–1383.
Quinn, P. K., S. Marshall, T. S. Bates, D. S. Covert, and V. N. Kapustin, 1995: Comparison of measured and calculated aerosol properties relevant to the direct radiative forcing of tropospheric sulfate aerosol on climate. *J. Geophys. Res.*, **100,** 8977–8991.
Schwartz, S. E., 1989: Acid deposition: Unraveling a regional phenomenon. *Science*, **243,** 753–762.
Seinfeld, J. H., and Coauthors, 1996: *Aerosol Radiative Forcing and Climate Change*. National Research Council, National Academy Press, 161 pp.
Thomason, L. W., B. M. Herman, and J. A. Reagan, 1983: The effect of atmospheric attenuators with structured vertical distributions on air-mass determinations and Langley plot analysis. *J. Atmos. Sci.*, **40,** 1851–1858.
Vanderpol, A. H., F. D. Carsey, D. S. Covert, R. J. Charlson, and A. P. Waggoner, 1975: Aerosol chemical parameters and air mass character in the St. Louis region. *Science*, **190,** 570.
Waggoner, A. P., R. E. Weiss, N. C. Ahlquist, D. S. Covert, S. Will, and R. J. Charlson, 1981: Optical characteristics of atmospheric aerosols. *Atmos. Environ.*, **15,** 1891–1909.
White, W. H., 1990: The components of atmospheric light extinction: A survey of ground-level budgets. *Atmos. Environ.*, **24A,** 2673–2679.
——, E. S. Macias, R. C. Nininger, and D. Schorran, 1994: Size-resolved measurements of light scattering by ambient particles in the southwestern U.S.A. *Atmos. Environ.*, **28,** 909–921.
Winker, D. M., R. H. Couch, and P. McCormick, 1996: An overview of LITE: NASA's Lidar In-space Technology Experiment. *Proc. IEEE*, **84,** 164–180.