Estimates of Turbulence from Numerical Weather Prediction Model Output with Applications to Turbulence Diagnosis and Data Assimilation

Rod Frehlich

Research Application Program, National Center for Atmospheric Research,* and CIRES, University of Colorado, Boulder, Colorado

Robert Sharman

Research Application Program, National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT

Estimates of small-scale turbulence from numerical model output are produced from local estimates of the spatial structure functions of model variables such as the velocity and temperature. The key assumptions used are the existence of a *universal* statistical description of small-scale turbulence and a locally *universal* spatial filter for the model variables. Under these assumptions, spatial structure functions of the model variables can be related to the structure functions of the corresponding atmospheric variables. The shape of the model spatial filter is determined by comparisons with the spatial structure functions from aircraft data collected at cruising altitudes. This universal filter is used to estimate the magnitude of the small-scale turbulence, that is, scales smaller than the filter scale. A simple yet universal description of the basic statistics (such as the probability density function and the spatial correlation) of these small-scale turbulence levels in the upper troposphere and lower stratosphere is proposed. Various applications are presented including 1) predicting the statistics of turbulence experienced by aircraft at upper levels, 2) diagnosing and forecasting turbulence for aviation safety, and 3) estimating the total observation error for optimal data assimilation and for improving operational weather prediction models. It is determined that the total observation error for typical rawinsonde measurements of velocity are dominated by the sampling error or ''error of representativeness'' resulting from the effects of small-scale turbulence.

1. Introduction

Fluid dynamical simulation models typically suffer from underrepresentation of the smallest scales of motion resolvable by finite-difference numerical models. This underrepresentation comes from two sources (e.g., Ferziger and Perić 2002): 1) truncation error (i.e., errors due to the discrete representation of continuous processes) and 2) error introduced by explicit and implicit model smoothing and filtering effects. We assume here that both of these effects can be represented through a universal spatial filter; that is, the functional form of the filter is independent of horizontal location for a given model and altitude region (Daley 1993; Pielke 2002). This spatial filtering is well characterized in large eddy simulation (LES) spectral models by the use of well-defined explicit filters and careful definitions of the resolved and the unresolved scales (e.g., Sullivan et al. 2003). But in global and mesoscale atmospheric models these effects are mostly undocumented since emphasis has traditionally been on accurate representations of the larger-scale motions that contribute most significantly to meteorological phenomena, and to some extent, the need to produce numerically stable solutions rather than to describe the details of the turbulence field. Since most parameterizations of subgrid turbulent processes are a function of the locally resolved fields, the effects of model filtering should be included in the subgrid turbulence parameterization. In this paper we develop methods to quantify, for certain atmospheric regimes, the effects of spatial filtering in mesoscale models, and consequently produce more accurate estimates of the small-scale turbulence. Here the term ''small scale'' is used to refer to the smallest resolvable scales of an atmospheric dynamical model (i.e., scales smaller than the filter scale), which with suitable extrapolation, may be used to infer the magnitude of ''subgrid'' scale motions as well.

This study is motivated by the realization that statistically, the atmospheric wind and temperature spatial

---

Corresponding author address: Dr. Rod Frehlich, CIRES, 216 UCB, University of Colorado, Boulder, CO 80309.
E-mail: rgf@cires.colorado.edu

spectra at middle latitudes and upper levels exhibit a $k^{-5/3}$ behavior, where k is the horizontal wavenumber (or equivalently exhibit an $s^{+2/3}$ behavior for the second-order structure function, where s is the separation) from scales ranging from about 400 km down to 1 km. This was demonstrated rather convincingly by the Global Atmospheric Sampling Program (GASP) and Measurement of Ozone by Airbus In-Service Aircraft (MOZAIC) data analyses (Nastrom and Gage 1985; Lindborg 1999; Cho and Lindborg 2001). These datasets were obtained by specially instrumented commercial aircraft collecting wind and temperature data at aircraft cruise levels (approximately 8–10 km MSL) over several thousand flight legs.

Individual flights from National Aeronautics and Space Administration (NASA) and National Center for Atmospheric Research (NCAR) research aircraft also display the $k^{-5/3}$ behavior. For example, Sharman and Frehlich (2003) consistently obtained a $k^{-5/3}$ behavior on scales from 600 km down to about 40 m from various research aircraft field campaigns. In the Sharman and Frehlich paper some examples were provided from short flight legs of the NASA B-757 research aircraft. The spectra in the Sharman and Frehlich study always exhibited a $k^{-5/3}$ behavior for all 67 flight segments studied, in both low- and high-turbulence regions. We did not observe steeper spectral slopes as reported by Vinnichenko (1970). The average of 40 nonoverlapping spectra produced from a particularly long (600 km) leg of measurements from the NCAR EC-130 aircraft flying above the planetary boundary layer at 4806-m elevation as part of the Indian Ocean Experiment (INDOEX) is shown in Fig. 1. Details of the NCAR aircraft measurement systems can be found in Lenschow (1986), Lenschow and Spyers-Duran (1989), and Khelif et al. (1999). As can be seen, the $-5/3$ spectral slope is a good fit to all three velocity components and temperature spectra for wavelengths from about 6 km to about 30 m. At scales smaller than about 30 m, the velocity spectra deviate from the $-5/3$ slope because of measuring difficulties of the differential pressure sensors and possibly other effects. Also note that for scales less than about 600 m ($k > 0.01$ m^{-1}), the spectra of all three velocity components have nearly the same magnitude, but the outer length scale for the vertical velocity component is 1–2 km, while the horizontal velocity components, consistent with the GASP and MOZAIC data, exhibit no identifiable length scale in the range of wavelengths shown.

The cause of the observed $k^{-5/3}$ statistical behavior is at the moment not completely understood. Discussions of competing theories are provided, for example by Tung and Orlando (2003) and Koshyk and Hamilton (2001) within the context of recent simulation results that seem to point to a downscale cascade. For our purposes, it is sufficient merely to accept the $k^{-5/3}$ spectral behavior as a universal statistical description of midlatitude upper-level "turbulence" so that this behavior
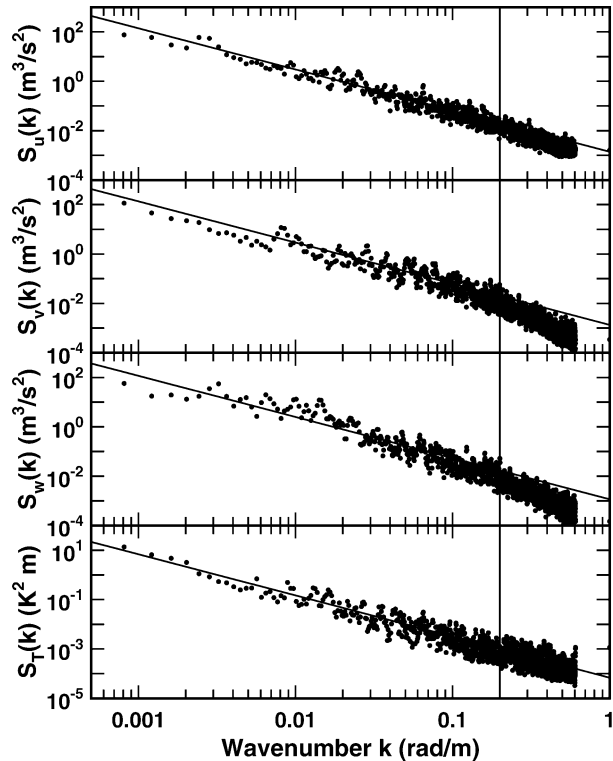

FIG. 1. Longitudinal, transverse, vertical velocity, and temperature spectra, denoted $S_u(k)$, $S_v(k)$, $S_w(k)$, and $S_T(k)$, respectively, as derived from INDOEX field campaign data and the best-fit Kolmogorov slope of $-5/3$ over the interval $k = (0.0005–0.2$ rad m$^{-1})$ as shown by the vertical line.

can be used as "truth" to quantify the underrepresentation of the small-scale motions in mesoscale models. Depending on the model numerics and filtering used, a mesoscale model will in general show an energy deficit from this assumed universal statistical behavior, and generally this deficit will be largest for the smallest scales resolved by the model. This model-dependent deficit can be estimated from the s dependence of model-derived second-order structure functions compared to the expected $s^{2/3}$ behavior. For this comparison the use of structure functions is preferred since they permit accurate measurements of turbulence over small measurement domains and are more robust compared with spectral methods, which suffer from windowing and aliasing effects (Frehlich 1997; Frehlich et al. 1998, 2001). Once the spatial filter of the model has been determined, information about turbulence levels can be extrapolated to scales smaller than the model resolution to provide information about the energy content of the grid scale and subgrid-scale turbulent motions.

The layout of the remainder of the paper is as follows. In section 2 results from aircraft measurements of the spatial statistics of upper-level turbulence are reviewed. Using these results and comparing to numerical weather prediction (NWP) models, the National Centers for Environmental Prediction (NCEP) Rapid Update Cycle

(RUC) (Benjamin et al. 2004) and the fifth-generation (Pennsylvania State University) PSU–NCAR Mesoscale Model (MM5) (Grell et al. 1994; Davis et al. 1999) in particular, universal filter functions for these models are derived in section 3. Similar methods are used to derive a universal filter function for the Clark–Hall cloud-resolving model (Clark 1977; Clark and Hall 1991). A procedure to estimate small-scale turbulence intensities is outlined in section 4 along with some results using this procedure. The remainder of the paper presents some applications of the technique. Section 5 provides an application of the estimated local turbulence intensities to produce estimates of sampling error or "error of representativeness" of rawinsonde wind and temperature measurements. This provides a more accurate description of total observation error statistics (section 6), which can be used for optimal data assimilation and for improving ensemble forecasts based on more realistic perturbations of the initial state. In section 7 the technique is used to provide model-derived turbulence estimates as a diagnostic for aircraft-scale turbulence. The diagnostic is then used to derive a climatology of aircraft-scale upper-level turbulence in section 8. A summary and conclusions are provided in section 9.

2. Statistical description of upper-level turbulence

The most reliable statistical descriptions of turbulence on spatial scales relevant to current mesoscale NWP models (1–200 km) have been produced by specially instrumented commercial aircraft, in particular GASP (Nastrom and Gage 1985) in the United States and MOZAIC in Europe (Lindborg 1999; Cho and Lindborg 2001). Since most of these flights were of fairly long duration, most of the measurements were taken at cruise altitudes, that is, in the upper troposphere and lower stratosphere.

The spatial statistics of upper-level turbulence may be described by spatial structure functions. Longitudinal and transverse second-order structure functions are defined respectively by

$$D_{LL}(\mathbf{s}) = \langle [v_L(\mathbf{x}) - v_L(\mathbf{x} + \mathbf{s})]^2 \rangle,$$

$$D_{NN}(\mathbf{s}) = \langle [v_N(\mathbf{x}) - v_N(\mathbf{x} + \mathbf{s})]^2 \rangle, \quad (2.1)$$

where $v_L(\mathbf{x})$ and $v_N(\mathbf{x})$ are the velocity components along and transverse to the displacement vector $\mathbf{s} = (x, y, z)$, respectively, and $\langle \rangle$ denotes an ensemble average. Following Lindborg (1999) and Cho and Lindborg (2001) we assume the statistics of the structure functions follow simple scaling laws for homogeneous turbulence, valid for either 2D or 3D isotropic conditions. For 3D isotropic turbulence, $D_{NN}(s)$ is related to $D_{LL}(s)$ by [Monin and Yaglom (1975), Eq. (13.70)]

$$D_{NN}(s) = D_{LL}(s) + \frac{s}{2}\frac{d}{ds}D_{LL}(s) \quad (2.2)$$

and for 2D isotropic turbulence by (Lindborg 1999)

$$D_{NN}(s) = \frac{d}{ds}[sD_{LL}(s)], \quad (2.3)$$

where $s = |\mathbf{s}|$.

The statistics of fully developed 3D isotropic turbulence are well described by the Kolmogorov model, in which case the turbulence intensity, as measured by the eddy dissipation rate ε, is related to $D_{LL}(s)$ and $D_{NN}(s)$ [Monin and Yaglom (1975), Eq. (2.17)] through

$$D_{LL}(s) = C_K \varepsilon^{2/3} s^{2/3} \approx 2\varepsilon^{2/3} s^{2/3},$$

$$D_{NN}(s) = \frac{4}{3}C_K \varepsilon^{2/3} s^{2/3} \approx \frac{8}{3}\varepsilon^{2/3} s^{2/3}, \quad (2.4)$$

where $C_K \approx 2$ is the Kolmogorov constant. For fully developed 2D isotropic turbulence analogous relations can be written as (Lindborg 1999)

$$D_{LL}(s) = C_K \varepsilon^{2/3} s^{2/3} \approx 2\varepsilon^{2/3} s^{2/3},$$

$$D_{NN}(s) = \frac{5}{3}C_K \varepsilon^{2/3} s^{2/3} \approx \frac{10}{3}\varepsilon^{2/3} s^{2/3}, \quad (2.5)$$

realizing that ε as used here may not strictly satisfy the definition of 3D turbulence intensity.

Both the GASP and MOZAIC campaigns produced essentially the same average statistics for midlatitude upper-level turbulence. The best-fit models to the combined datasets were derived by Lindborg (1999) as

$$D_{LL}(s) = a_1 s^{2/3} + b_1 s^2 - c_1 s^2 \ln s, \quad (2.6)$$

$$D_{NN}(s) = a_2 s^{2/3} + b_2 s^2 - c_2 s^2 \ln s, \quad (2.7)$$

where $a_1 \approx 0.0036$ m$^{4/3}$ s^{-2}, $a_2 \approx 0.004$ m$^{4/3}$ s^{-2}, $b_1 \approx 2.4 \times 10^{-9}$ s^{-2}, $b_2 \approx 6.5 \times 10^{-9}$ s^{-2}, $c_1 \approx 0.16 \times 10^{-9}$ s^{-2}, and $c_2 \approx 0.43 \times 10^{-9}$ s^{-2}. Note that for small spacings s, the first term dominates, and the structure functions recover the Kolmogorov form (2.4). For larger spacings, the second and third terms recover the so-called enstrophy cascade (k^{-3}) region observed at larger scales in both the GASP and MOZAIC data. Note the coefficients in (2.6) and (2.7) are the original values from Lindborg (1999) and include measurements in both the upper troposphere and lower stratosphere from low latitudes to high latitudes. Cho and Lindborg (2001) document a weak latitude dependence, but for the purposes of comparison to numerical model output we choose to use the average values.

If (2.6) is used with the 2D isotropic assumption (2.3), the same form as Lindborg's best-fit transverse structure function model (2.7) is recovered but with slightly different coefficients

$$D_{NN}(s) = a_3 s^{2/3} + b_3 s^2 - c_3 s^2 \ln s, \quad (2.8)$$

where $a_3 = 5/3\, a_1 \approx 0.006$ m$^{4/3}$ s^{-2}, $b_3 = 3b_1 - c_1 \approx 7.04 \times 10^{-9}$ s^{-2}, and $c_3 = 3c_1 \approx 0.48 \times 10^{-9}$ s^{-2}.

The structure function of temperature D_T is more complicated because of the average temperature variation with latitude. However, the structure function in the zonal direction (E–W) is approximately proportional to D_{LL}:

$$D_T(s) = a_4 s^{2/3} + b_4 s^2 - c_4 s^2 \ln s. \quad (2.9)$$

From Fig. 5 of Nastrom and Gage (1985), assuming the integral of the temperature and velocity spectra are equal to their respective variances, the constants evaluate to $a_4 \approx 6.36 \times 10^{-4}$ K^2 m$^{-2/3}$, $b_4 \approx 4.24 \times 10^{-10}$ K^2 m^{-2}, and $c_4 \approx 2.83 \times 10^{-11}$ K^2 m^{-2}.

3. Statistical description of mesoscale model output

A finite-difference numerical simulation model of the atmosphere consists of a discrete representation of the atmospheric-state variables on a finite-difference grid and a discrete representation of the governing equations of motion. It is common practice to represent the desired discrete model variables ("truth") as a spatial average of the random atmospheric fields centered in a model grid cell (e.g., Daley 1993),

$$v_{\text{truth}}(\mathbf{r}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\mathbf{s} - \mathbf{r}) v(\mathbf{s}) \, d\mathbf{s}, \quad (3.1)$$

where $\mathbf{r} = (x, y, z)$ denotes the center of the grid cell, $W(\mathbf{s})$ is the effective weighting function [filter function with normalization $\iiint W(\mathbf{s}) \, d\mathbf{s} = 1$] of the model, and $d\mathbf{s}$ denotes 3D integration. Assuming the filter function is a spatial average over the grid cell (e.g., Deardorff 1970; Pielke 2002), that is, $W(\mathbf{s}) = 1/(L_x L_y L_z)$ inside the grid cell, 0 outside, (3.1) becomes

$$v_{\text{truth}}(\mathbf{r}) = \frac{1}{L_x L_y L_z} \int_{-L_x/2}^{L_x/2} \int_{-L_y/2}^{L_y/2} \int_{-L_z/2}^{L_z/2} v(\mathbf{r} + \mathbf{s}) \, d\mathbf{s}, \quad (3.2)$$

where L_x, L_y, L_z denote the dimensions (spacings) of the model grid cell. We will call a model that satisfies this filter function in a statistical sense a "perfect" model.

For a homogeneous field the local horizontal spatial statistics of a model output variable $\tilde{u}$ is well described by the spatial structure function defined by

$$D_{\tilde{u}}(r, s) = \langle [\tilde{u}(x, y, z) - \tilde{u}(x + r, y + s, z)]^2 \rangle. \quad (3.3)$$

Assume the model values $\tilde{u}$ are given by the general form (3.1) and the filter function $W(r)$ is a universal function *locally* (since the filter function may also depend on other conditions, e.g., stability and the subgrid parameterization). Then unbiased estimates of the local structure function in the (x, y) plane of $\tilde{u}$ for a small region centered at coordinate $(x, y, z) = (iL_x, jL_y, H)$ are given by

$$D_{\tilde{u}}(lL_x, mL_y, H) = \frac{1}{(2N_i - l)(2N_j - m)} \sum_{p=i-N_i}^{i+N_i} \sum_{q=j-N_j}^{j+N_j}$$
$$\times \{ \tilde{u}(pL_x, qL_y, H)$$
$$- \tilde{u}[(p + l)L_x, (q + m)L_y, H] \}^2, \quad (3.4)$$

where H denotes the altitude or pressure level of the model grid cell, and N_i and N_j define the spatial domain in the x and y directions, respectively. The size of the measurement domain is determined by the required statistical accuracy. Note that the model structure function is a filtered version of the actual structure function of the continuous random process for variable u because of the spatial filter of the model.

The ensemble average of the model structure function is given by (Papoulis 1965; Frehlich 1997)

$$D_{\text{model}}(\mathbf{s}) = \langle D_{\tilde{u}}(\mathbf{s}) \rangle = \int_{-\infty}^{\infty} D_u(\mathbf{s} - \mathbf{r}) V(\mathbf{r}) \, d\mathbf{r}$$
$$- \int_{-\infty}^{\infty} D_u(\mathbf{r}) V(\mathbf{r}) \, d\mathbf{r}, \quad (3.5)$$

where

$$V(\mathbf{s}) = \int_{-\infty}^{\infty} W(\mathbf{r} - \mathbf{s}) W(\mathbf{r}) \, d\mathbf{r} \quad (3.6)$$

is the autocorrelation of the filter function W.

Using these definitions, spatial statistics from NWP model output can be developed and compared to Lindborg's results to derive statistical metrics for NWP model performance. These metrics have been derived for three NWP models: NCEP's RUC (Benjamin et al. 2004) 20-km horizontal resolution model (RUC20) as well as an earlier 40-km version (RUC40), and a 30-km horizontal resolution version of MM5 (Grell et al. 1994; Davis et al. 1999) run operationally at the NCAR Mesoscale and Microscale Meteorology (MMM) Division. Examples of the average longitudinal and transverse model-derived structure functions using (3.4) from 1 yr of RUC20, RUC40, and MM5 0000 UTC analyses are shown in Figs. 2 to 4, respectively. Also plotted are the $s^{2/3}$ scaling, Lindborg's (1999) best-fit models (2.6) and (2.7), along with the prediction of a perfect model (3.5), assuming the spatial filter of a perfect model is the average of the velocity field over the grid cell (3.2). As expected, the agreement with Lindborg's (1999) best-fit models at large scales (~500 km) for all the three NWP models is excellent, and the effective model filter is more pronounced than the perfect model filter. This comparison promotes the use of second-order spatial structure functions as a diagnostic for the spatial filter of the model. Note that the absolute levels of the structure functions derived from the various NWP models agree very well.

An empirical function $D_{\text{model}}(s)$ for the model-derived structure functions can be determined to explicitly account for the model-dependent filter function. The simplest such empirical function has the form

$$D_{\text{model}}(s) = K D_{\text{cor}}(s) D_{\text{ref}}(s), \quad (3.7)$$

where K is a constant, $D_{\text{cor}}(s)$ describes the correction produced by the model filter, and $D_{\text{ref}}(s)$ is a normalized
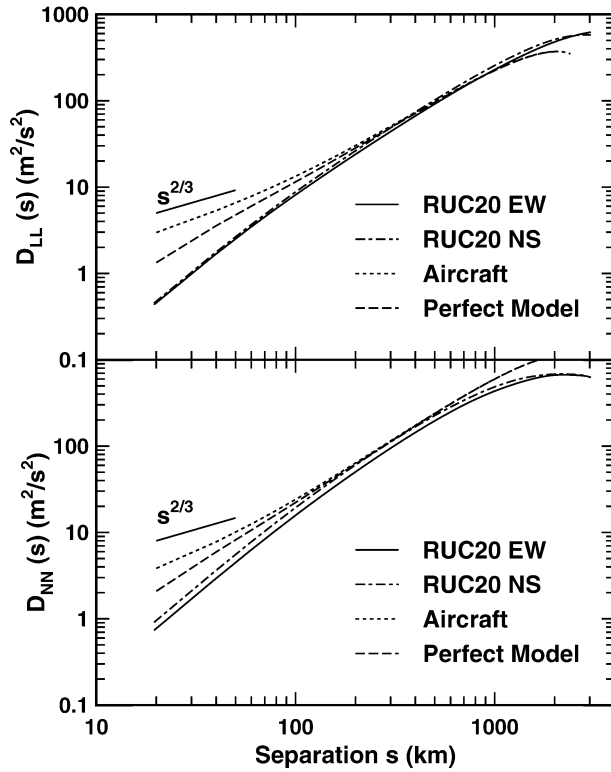
FIG. 2. Average structure functions from RUC20 model east–west (E–W) direction (solid line) and north-south (N–S) direction (dash–dot line), in situ measured best-fit average structure function (dotted line) [(2.6) and (2.7)] from aircraft data, prediction of perfect model (dashed line) [(3.5)], and the $s^{2/3}$ scaling.
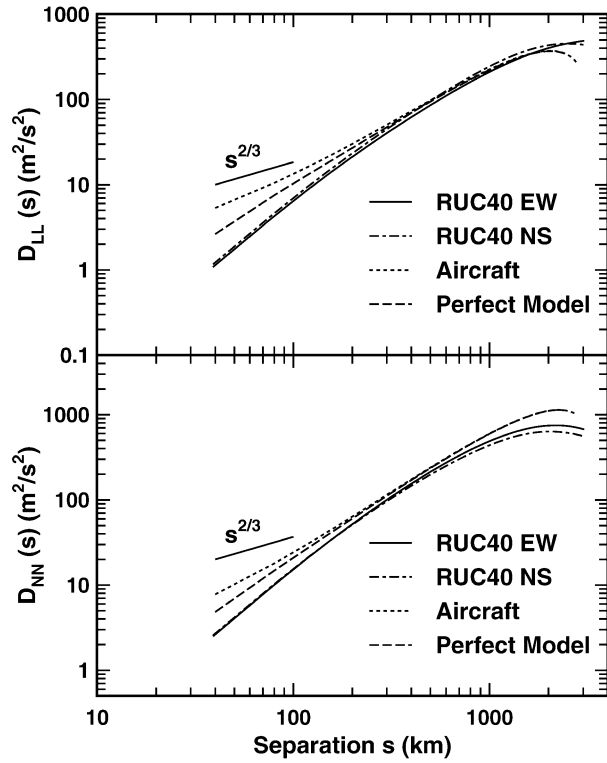

FIG. 3. As in Fig. 2, but for average structure functions from RUC40 model.

FIG. 4. As in Fig. 2, but for average structure functions from MM5.

form of the in situ measured structure functions, for example, the Lindborg model. Note that according to the analysis of the GASP data by Nastrom and Gage (1985) and of the MOZAIC data by Cho and Lindborg (2001), the reference structure function model depends weakly on the measurement region, that is, altitude, latitude, and longitude, but this effect will be ignored here.

If the reference normalized longitudinal structure function is taken as the Lindborg best-fit model (2.6)

$$D_{\mathrm{LLref}}(s) = s^{2/3} + \frac{b_1}{a_1}s^2 - \frac{c_1}{a_1}s^2 \ln s, \qquad (3.8)$$

the correction function $D_{\mathrm{cor}}(s)$ should approach unity for large separations where the effects of the spatial filter are negligible. For small spacings, the model structure function $D_{\mathrm{model}}(s)$ should approach s^2 to reflect the spatial smoothing of the fields at the smallest scales, that is, an effective inner scale (Monin and Yaglom 1975, p. 353). Therefore, we select the following simple empirical function which satisfies both these requirements:

$$D_{\mathrm{LLcor}}(s) = \frac{(s/p_1)^{4/3}}{1 + (s/p_1)^{4/3} + p_2(s/p_1)^{2/3}}, \qquad (3.9)$$

where p_1 is the length scale of the model filter, and p_2 is a fitting parameter. The best-fit constant
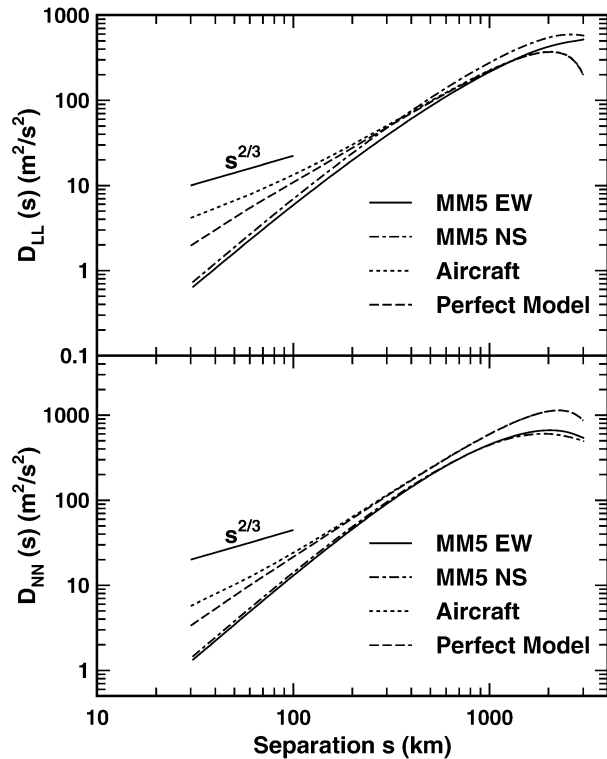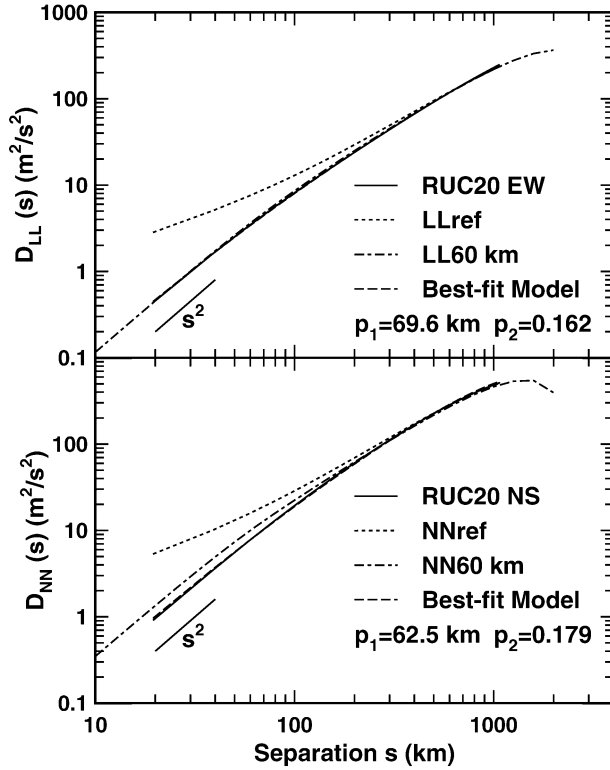
FIG. 5. Average velocity structure functions from RUC20 (solid lines) for the longitudinal velocity in the east–west (E–W) direction and the transverse velocity in the north–south (N–S) direction, the best-fit model [(3.7)] (dashed line) with parameters p_1 and p_2, the best-fit reference models [(3.8) and (3.11)] (dotted line), the theoretical predictions for a 60-km grid cell [(3.5)] (dash–dot line), and the s^2 scaling.
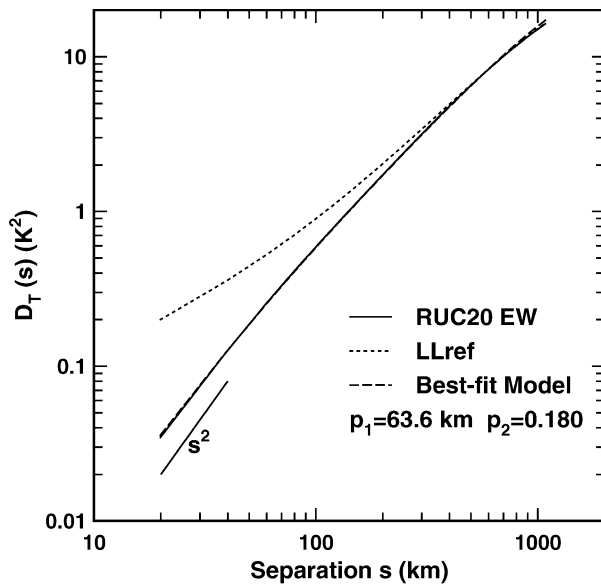

FIG. 6. Average temperature structure functions from RUC20 (solid lines) in the east–west (EW) direction, best-fit model [(3.7)] (dashed line) with parameters p_1 and p_2, best-fit reference model [(3.8)] (dotted line), and the s^2 scaling.

$$K = C_K \varepsilon^{2/3} \approx 2\varepsilon^{2/3} \qquad (3.10)$$

provides an estimate of ε.

Alternatively, if the reference normalized transverse structure function is taken as the Lindborg 2D isotropic model (2.8)

$$D_{\mathrm{NNref}}(s) = s^{2/3} + \frac{9b_1 - 3c_1}{5a_1}s^2 - \frac{9c_1}{5a_1}s^2 \ln s, \quad (3.11)$$

then the best-fit constant

$$K = \frac{5}{3}C_K \varepsilon^{2/3} \approx \frac{10}{3}\varepsilon^{2/3} \qquad (3.12)$$

provides an estimate of ε. Similarly, for estimates of the temperature structure function, the reference structure function model is the same as (3.8), and the best-fit constant $K = C_T^2$ provides an estimate of the temperature turbulence level.

The unknown parameters p_1 and p_2 are determined by minimizing the chi-squared error (Press et al. 1986) between the empirical model and the average structure functions derived from the NWP model. Examples from the RUC20 analysis of Fig. 2 are shown in Fig. 5 for the velocity structure functions and in Fig. 6 for the temperature structure function in the east–west direction. (The structure functions of temperature in the north–south direction were not computed because the strong latitude dependence significantly alters the structure function shape.) Note that the best-fit model structure function is almost identical to the measured structure function. Similar results were produced for the RUC40 and MM5 models. Also shown in Fig. 5 is the theoretical calculation (3.5) of the longitudinal (LL60 km) and transverse (NN60 km) structure functions for a 60-km grid cell and the 2D isotropic Lindborg model (2.6) and (2.8). The RUC20 analysis is well approximated by the 60-km grid cell filtering for the longitudinal structure function while there is a small error for the transverse structure function. This heavy filtering is typical of NWP models.

A similar procedure was also applied to output from the higher-resolution Clark–Hall anelastic cloud-resolving model (e.g., Clark 1977; Clark and Hall 1991). Figure 7 shows the best-fit global structure functions obtained from an analysis of a deep convective cloud simulation from the Clark–Hall model (Lane et al. 2003) using a horizontal resolution of 25 m. Here the reference model is chosen as the Kolmogorov model (2.3) because of the clear $s^{2/3}$ power-law region produced in the range of 500–2000-m separations. Thus the fitting procedure appears to work equally well for the microbeta scales of the cloud-resolving model as for the mesoscale NWP models.

4. Estimates of small-scale turbulence

a. Estimates of local turbulence intensities

Various methods have been proposed to estimate turbulence associated with unresolved scales [i.e., subgrid
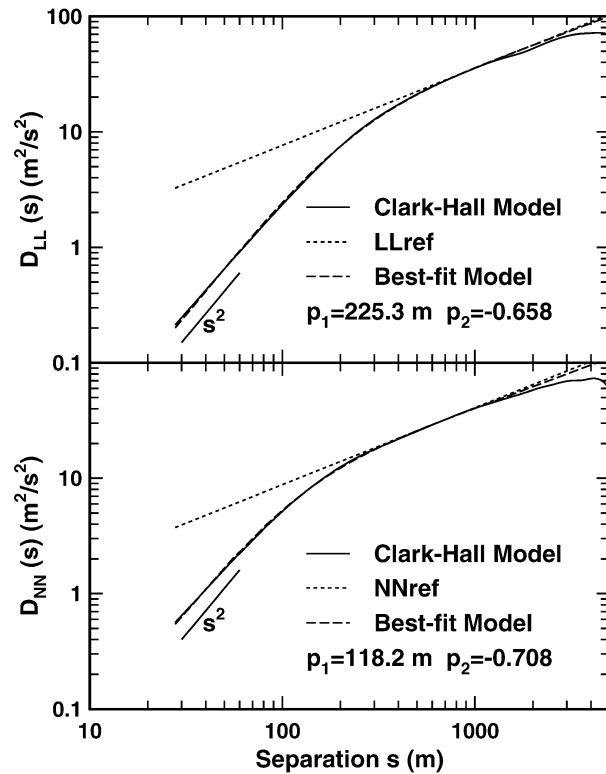
FIG. 7. Average structure functions from the Clark–Hall cloud resolving model (solid lines), best-fit model [(3.7)] (dashed line) with parameters p_1 and p_2, best-fit reference model [(2.4)] (dotted line), and the s^2 scaling.
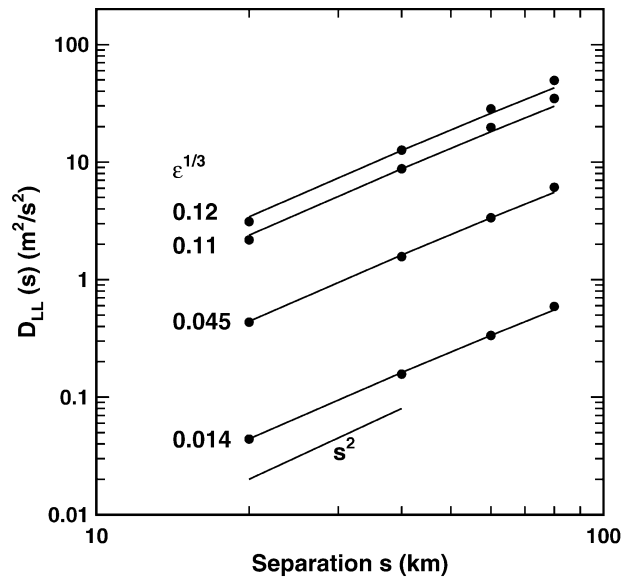

FIG. 8. Example estimates (●) of longitudinal velocity structure functions $D_{LL}(s)$ from the RUC20 model using 100 km $\times$ 100 km subdomains, the best-fit model [(3.7)] (lines) with turbulence estimates $\varepsilon^{1/3}$ (m$^{2/3}$ s^{-1}) given to the left of each line, and the s^2 scaling for reference.

scale (sgs)] from resolved-scale model fields. The most common are based on prognostic equations of the sgs turbulent kinetic energy (TKE) (e.g., Deardorff 1970; Moeng 1984; Moeng and Sullivan 1994; Pielke 2002). This requires assumptions about the local conditions to connect resolved-scale gradients to sgs TKE. Since the horizontal statistics of the velocity and temperature fields have a robust description, we propose instead to estimate local (i.e., all scales less than the model filter cutoff, including subgrid scales) turbulence intensities based on the assumption of a universal spatial filter for the model as outlined in the previous section. A similar approach has been used to estimate the statistics of small-scale turbulence from lidar measurements using a correction for the spatial average of the lidar pulse (Frehlich 1997; Frehlich et al. 1998; Frehlich and Cornman 2002). If the model has a universal filter function—that is, the functional form of the structure functions is independent of location—then local estimates of the level K (or equivalently of $\varepsilon^{2/3}$ or C_T^2) can be produced by a best fit to the shape of the structure functions [see (3.7)] where the reference models are given by (3.8) and (3.9) for the longitudinal and transverse structure functions, respectively. Local estimates of ε can be produced from the best-fit levels of the two longitudinal and two transverse structure functions using (3.10) and

(3.12). Since the longitudinal structure functions seem to have a slightly better defined $s^{2/3}$ region (cf. Fig. 5), we use the average of the two longitudinal structure functions (east–west and north–south) with (3.10) to produce local estimates of ε. In Fig. 8 plots are provided of average longitudinal structure functions at various locations in a RUC20 model field at an altitude of 10 km and the corresponding estimates of $\varepsilon^{1/3}$ derived from a 5 $\times$ 5 point horizontal domain (100 km $\times$ 100 km). Note that at each location the fit is quite good, and the shape of the structure functions is approximately the same, but the levels $\varepsilon^{1/3}$ can differ considerably.

b. Estimates of the probability density function

The statistical description of the structure function estimates of the local turbulence levels ε is fully described by the probability density function (PDF). The PDF of the variations in the spectral level of the GASP data is well described by a lognormal distribution (Nastrom and Gage 1985). This is equivalent to a lognormal PDF for ε, which is completely defined by two parameters: $\langle \log \varepsilon \rangle$ or the median value ε_{50} and the standard deviation $\sigma_{\log \varepsilon}$ of $\log_{10} \varepsilon$. From the GASP data, Frehlich (2001) derived these parameters as $\varepsilon_{50} = 2.66 \times 10^{-5}$ m^2 s^{-3}, $\sigma_{\log \varepsilon} = 0.63$, and in addition, $\langle \varepsilon \rangle = 7.64 \times 10^{-5}$ m^2 s^{-3}. The PDF of the RUC20 estimates of ε from local structure function estimates over an $L \times L$ subdomain for two values of L are compared to the lognormal model with these parameters in Fig. 9. Similarly, the PDF of the RUC20 estimates of C_T^2 is shown in Fig. 10. It should be pointed out that the estimates of ε and
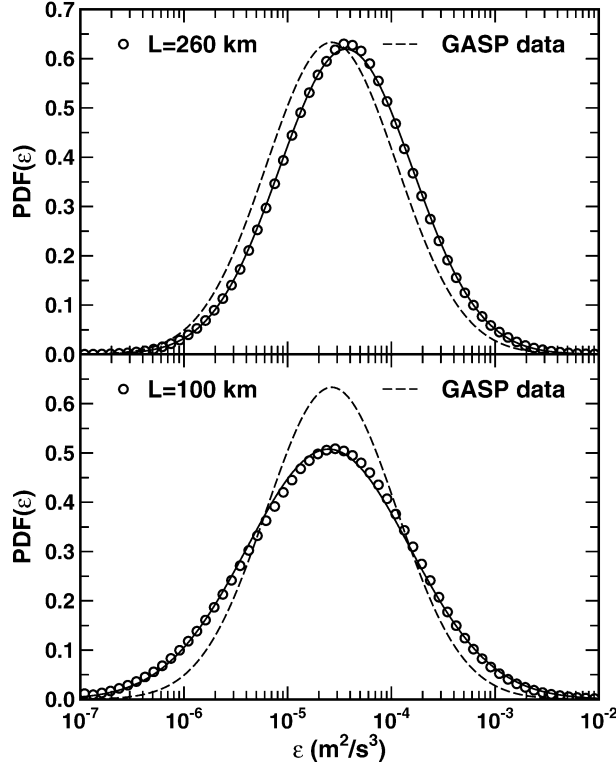
FIG. 9. PDF of RUC20 estimates of ε(o) from two different square subdomains $L \times L$, the lognormal model prediction based on the mean and standard deviation of $\log\varepsilon$ values (solid line), and the lognormal model prediction from the GASP data (dashed line) derived by Frehlich (2001).
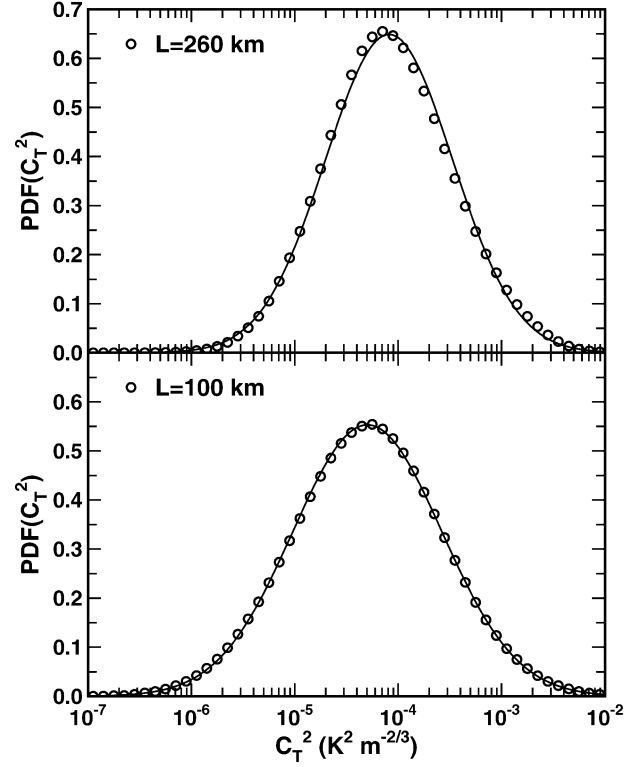

FIG. 10. PDF of RUC20 estimates of C_T^2(o) from two different square subdomains $L \times L$ and the lognormal model prediction based on the mean and standard deviation of $\log C_T^2$ (solid line).

C_T^2 from the GASP data were produced from the spectral level at a wavelength of 400 km for flight legs longer than 2400 km (i.e., *along a line*), which will have different statistics than the RUC20 estimates, which were produced *over a square domain*. However, all the PDFs agree well with the lognormal model, although the RUC20-derived estimates depend on the averaging length L.

c. Scaling laws as a function of the averaging dimensions

As shown above, the statistics of ε and $\log\varepsilon$ depend on the averaging length L. To appreciate this dependence, the statistics of the lognormal PDF of ε from RUC20 analyses at 10-km altitude are plotted in Fig. 11 as a function of L, as well as the best-fit power law from the first (i.e., smallest L) four data points, namely,

$$\langle\log\varepsilon\rangle = 0.41782 \log L - 6.6956, \quad (4.1)$$

$$\sigma_{\log\varepsilon} = 7.5019 L^{-0.19622}. \quad (4.2)$$

These relations provide a simple scaling of the statistics of ε to any square $L \times L$ averaging domain. Note that for L larger than about 200 km, the fit deviates slightly from the best-fit lines, probably due to the inclusion of

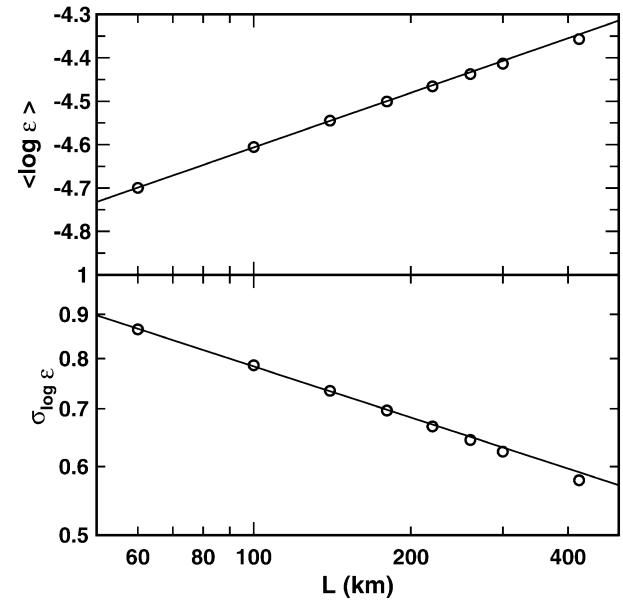FIG. 11. Parameters (o) of the lognormal PDF of ε vs the averaging length L. (top) The average of $\log\varepsilon$, and (bottom) the standard deviation of $\log\varepsilon$. The best-fit straight line based on the first four data points is also shown.
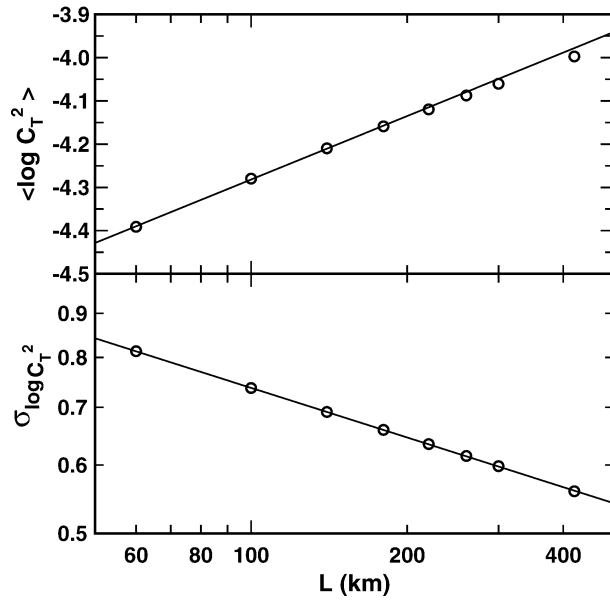
FIG. 12. Parameters of the lognormal PDF of C_T^2 vs the averaging length L(o). (top) The average of $\log C_T^2$, and (bottom) the standard deviation of $\log C_T^2$. The best-fit straight line based on the first four data points is also shown.

the enstrophy cascade region at larger scales as observed in the GASP data. Other simple scaling laws are derived from the RUC20 data similarly:

$$\langle \varepsilon \rangle = 2.5848 \times 10^{-4} L^{-0.068755}, \quad (4.3)$$

$$\langle \varepsilon^{2/3} \rangle = 1.04162 \times 10^{-3} L^{+0.043359}, \quad (4.4)$$

$$SD[\varepsilon] = 5.04535 \times 10^{-2} L^{-0.41150}, \quad (4.5)$$

$$VAR[\varepsilon^{2/3}] = 5.8795 \times 10^{-4} L^{-0.37429}. \quad (4.6)$$

Note that, consistent with refined similarity theory, the variance of small-scale turbulence intensity increases with a decrease in the averaging domain L. Also, as expected, the exponent of L in the $\langle \varepsilon \rangle$ and $\langle \varepsilon^{2/3} \rangle$ expressions is very small, indicating only a weak dependence on L.

The average longitudinal structure function derived in Lindborg (1999) for all the cruising level aircraft data produced $\langle \varepsilon \rangle = 7.64 \times 10^{-5}$ m^2 s^{-3} and similarly, for 30°–50°-latitude tropospheric data Cho and Lindborg (2001, Table 1) found $\langle \varepsilon \rangle = 7.0 \times 10^{-5}$ m^2 s^{-3}. Note that these estimates are determined from the average structure function level in the $s^{2/3}$ power-law region, that is, $\langle \varepsilon \rangle = (a_1/C_K)^{3/2} = \langle \varepsilon^{2/3} \rangle^{3/2}$ since $a_1 = C_K \langle \varepsilon^{2/3} \rangle$ [see (2.5) and (2.6)]. Using (4.4) the RUC20 value is $\langle \varepsilon \rangle = 7.11 \times 10^{-5}$ m^2 s^{-3} for $L = 100$ km, and $\langle \varepsilon \rangle = 8.26 \times 10^{-5}$ m^2 s^{-3} for $L = 1000$ km, in good agreement with Lindborg's estimate.

Similarly, the scaling of the parameters of the lognormal distribution of C_T^2 can be derived from the RUC20 data, and the results are shown in Fig. 12. Well-defined power-law scalings are produced for averaging lengths less than about 200 km:

$$\langle \log C_T^2 \rangle = -6.7195 + 0.48755 \log L \quad (4.7)$$

$$SD[\log C_T^2] = 6.6724 L^{-0.19131}, \quad (4.8)$$

$$\langle C_T^2 \rangle = 1.2493 \times 10^{-4} L^{0.046511}, \quad (4.9)$$

$$VAR[C_T^2] = 5.1552 \times 10^{-4} L^{-0.61839}. \quad (4.10)$$

Equivalent expressions for the statistics of ε and C_T^2 can be derived from the output of *any* NWP model provided the spatial resolution is adequate to resolve at least part of the $s^{2/3}$ region. Ideally, different NWP models would produce the same climatology and statistical description of the small-scale turbulence in concert with the GASP and MOZAIC data.

The smallest scales for which these power-law scalings are valid is unclear. However, the GASP data indicate the spatial spectrum of velocity follows the $-5/3$ power law to scales of a few kilometers, and aircraft data (e.g., Fig. 1) show this trend continues down to at least a few tens of meters, implying a universal behavior. It is possible that the scaling for the variations in turbulence levels ε and C_T^2 also follows a universal power law down to very small scales. But this can only be verified by many high-resolution measurements of upper-level turbulence, and at the moment these are not available.

The well-behaved power-law scalings of the statistics of small-scale turbulence metrics ε and C_T^2 are indications of an underlying simple description of the spatial statistics of these fields analogous to high Reynolds number homogeneous and isotropic turbulence and Kolmogorov's refined similarity theory (Monin and Yaglom 1975). Any scalar field with homogeneous statistics and a Gaussian or lognormal distribution function is completely described by its covariance function, related to the structure function through

$$C_q(s) = \langle [q(s_0) - \langle q \rangle][q(s_0 + s) - \langle q \rangle] \rangle$$
$$= \sigma_q^2 - D_q(s)/2. \quad (4.11)$$

A simple model for C_q that produces the power-law scalings (4.3)–(4.10) is given by [e.g., Monin and Yaglom 1975, Eq. (25.28)]

$$C_q(s) = \frac{K_q}{(L_0^2 + s^2)^{\beta/2}}, \quad (4.12)$$

where L_0 is the turbulence length scale for the variations of q, and β is the intermittency exponent, which may depend on conditions (e.g., altitude). As shown in the appendix, the universal constants K_q and β for the scalar field $\varepsilon^{2/3}(x, y)$ can be estimated from the scaling laws produced from the RUC20 analysis as $\beta = 0.37429$ and $K_q = 4.21737 \times 10^{-4}$ m$^{8/3-\beta}$ s^{-4}.

5. Application to estimates of sampling error or error of representativeness

The correct statistical description of small-scale turbulence is critical for defining sensor data requirements

(Frehlich 2001) and for performing optimal data assimilation. The magnitude of local turbulence levels ε is typically highly variable (e.g., Monin and Yaglom 1975; Frehlich 1992; section 4). This variability will impact the sampling error associated with atmospheric measurement systems. In many cases the sampling error may be significantly larger than the instrument error of the observation. For example, the newer rawinsonde designs, namely, Vaisala's RS80-15G based on GPS positioning, have a stated accuracy of better than 0.5 m s^{-1} for each wind component (Jaatinen and Elms 2000). The automated weather reports from commercial aircraft outfitted with the Aircraft Communication, Addressing, and Reporting System (ACARS) have an estimated wind component error of 1.1 m s^{-1} (Benjamin et al. 1999). For comparison, research aircraft have an estimated accuracy of about 0.4 m s^{-1} (Lenschow 1986; Khelif et al. 1999). As will be presently shown, sampling errors for wind observations can be much larger than these instrument errors, as much as several meters per second.

a. Measurement errors

The error budget of NWP models can be separated into analysis errors, forecast errors, model errors, and measurement errors. A measurement $\hat{v}$ can be decomposed as

$$\hat{v} = v_{\text{truth}} + e_v + \text{BIAS}_v, \quad (5.1)$$

where v_{truth} is the desired measurement or truth, e_v is the total zero-mean random observation error, and BIAS_v is the total measurement bias. The magnitude of the random error e_v is defined as the standard deviation Σ_v of the measurement where

$$\Sigma_v^2 = \langle [\hat{v} - \langle \hat{v} \rangle]^2 \rangle = \langle e_v^2 \rangle, \quad (5.2)$$

and the bias is defined as

$$\text{BIAS}_v = \langle \hat{v} - v_{\text{truth}} \rangle, \quad (5.3)$$

where $\langle \ \rangle$ denotes an ensemble average over many realizations of a statistically similar atmosphere. A critical component for the definition of error is the choice of the desired measurement v_{truth}, which will depend on the application.

Any measurement can also be represented by

$$\hat{v} = v_S + i_v + b_v, \quad (5.4)$$

where v_S denotes the effective spatial sampling of the continuous random variable $v(\mathbf{r})$, b_v is the instrument bias, and i_v is the zero-mean random instrument error with magnitude $\sigma_v^2 = \langle i_v^2 \rangle$. The total measurement bias is then

$$\text{BIAS}_v = \langle v_S + b_v - v_{\text{truth}} \rangle = \text{BIAS}_{\text{atm}} + b_v, \quad (5.5)$$

where $\text{BIAS}_{\text{atm}} = \langle v_S - v_{\text{truth}} \rangle$ is the bias from the sampling pattern of the atmosphere. For many applications

$\text{BIAS}_{\text{atm}} = 0$ and $\text{BIAS}_v = b_v$. If the bias is zero, the magnitude of the total random error for the measurement of v is given by

$$\Sigma_v^2 = \langle [\hat{v} - v_{\text{truth}}]^2 \rangle = \langle [v_s + i_v - v_{\text{truth}}]^2 \rangle = \sigma_v^2 + \delta_v^2, \quad (5.6)$$

where

$$\delta_v^2 = \langle [v_S - v_{\text{truth}}]^2 \rangle \quad (5.7)$$

is the sampling error and we have assumed that the instrument error i_v is uncorrelated with $v_S - v_{\text{truth}}$; that is, the instrument error is uncorrelated with the local turbulence intensities.

An important factor for future measurement systems is the trade-off between the random error from the instrument σ_v and the sampling error δ_v. For example, for Doppler lidar space-based wind measurements, the magnitude of both of these errors is a function of the lidar shot pattern and the statistical description of the turbulent velocity field (Frehlich 2001). Note that the sampling error is defined only by the sampling pattern of the instrument and the statistics of the random atmosphere. This is slightly different from past definitions that essentially define v_{truth} in terms of model interpolation filters (Lorenc 1986; Daley 1993; Cohn 1997). However, the concept of sampling error is the same; that is, it defines the ability of the measurement to represent the intended spatial average or interpolation of the model variable.

b. Sampling error

Rigorous numerical and theoretical analysis of the sampling error requires a complete statistical description of the 3D random velocity field, that is, the small-scale turbulence. Simple analytic results for sampling errors associated with random processes can be produced by integrating over the z coordinate to produce a 2D random process:

$$\overline{v}(x, y, H) = \frac{1}{L_z} \int_{-L_z/2}^{L_z/2} v(x, y, H + z) \, dz, \quad (5.8)$$

where H denotes the height of the center of the grid cell. The random field is now a two-dimensional field, which is more tractable for numerical (Frehlich 2000) and analytic analysis. For example, the definition of "truth" as the average over the model grid cell becomes

$$v_{\text{truth}}(p, q, H) = \frac{1}{L_x L_y} \int_{-L_x/2}^{L_x/2} \int_{-L_y/2}^{L_y/2} \overline{v}(p + x, q + y, H) \, dx \, dy, \quad (5.9)$$

which is more convenient for numerical analysis. Although the exact statistical description of the two-dimensional random fields is not available, it is reasonable

to assume that the statistics of $\overline{v}(x, y, H)$ are homogeneous in the horizontal plane and will be similar to aircraft measurements because the vertical dimension of the grid cell is typically much less than the horizontal dimensions ($L_z \ll L_x, L_y$) and the integration over z will have only a small effect.

Assuming the 2D approximation for the random field, the sampling error for a rawinsonde measurement randomly located in the measurement volume is given by (Frehlich 2001)

$$\delta_v^2 = \frac{2}{L_x L_y} \int_0^{L_y} \int_0^{L_x} \left(1 - \frac{x}{L_x}\right)\left(1 - \frac{y}{L_y}\right) D_v(x, y, 0) \, dx \, dy,$$
(5.10)

where

$$D_v(x, y, z) = \langle [v(x_0, y_0, z_0) \\ - v(x_0 + x, y_0 + y, z_0 + z)]^2 \rangle$$
(5.11)

is the structure function of any variable v over a locally homogeneous domain centered at (x_0, y_0, z_0). Note that the variable v is general and the sampling error is completely defined by its structure function $D_v(x, y, 0)$.

Calculations of the sampling error component of the total observation error require the dimensions (L_x, L_y, L_z) of the chosen measurement volume, the sampling pattern of the measurement, and the spatial statistics of the small-scale turbulence. For measurements of winds by aircraft or by space-based lidar using linear tracks we define the horizontal velocity vector by its along- (u) and cross-track (v) components. Assuming the 3D isotropic Kolmogorov model for the horizontal velocity field, the sampling errors for a square grid ($L_x = L_y = L$) become (Frehlich 2001, Table 1)

$$\delta_u = \delta_v = 0.856835(\varepsilon L)^{1/3} \quad (5.12)$$

for a random rawinsonde measurement inside the measurement volume, and

$$\delta_{urc} = \delta_{vrc} = 0.687592(\varepsilon L)^{1/3} \quad (5.13)$$

for a rawinsonde measurement located at the center of a measurement cell. For an observation that samples a line along the total length of the measurement cell in the direction of u (profiler or lidar from space)

$$\delta_u = 0.350945(\varepsilon L)^{1/3}, \quad (5.14)$$

$$\delta_v = 0.211040(\varepsilon L)^{1/3}. \quad (5.15)$$

Similarly, the sampling error δ_T for temperature measurements with a rawinsonde randomly located inside a square measurement cell of length L is

$$\delta_T = 0.560930 C_T L^{1/3}. \quad (5.16)$$

The sampling error can be determined for any observation system given information about the spatial statistics of the measured variable. Therefore, the spatial statistics of the total observation error can also be determined. This includes the variance and the error covariance matrix of a collection of multiple observations. The local estimates of the observation error statistics can then be used to improve data assimilation techniques. This requires estimates of the local turbulence levels such as those provided here.

As an example, the magnitude of the sampling error for measurements of one component of horizontal velocity and temperature with a rawinsonde randomly located in a measurement cell is given by (5.12) and (5.16), respectively. Estimates of the sampling error using (5.12) and (5.16) for a measurement of a single component of the horizontal velocity and temperature from a rawinsonde randomly located in a 100 km $\times$ 100 km measurement cell based on the estimates of $\varepsilon^{1/3}$ and C_T^2 for the RUC20 domain for a particular case are shown in Fig. 13. Note the large variability in the sampling error (from near 0 to 8 m s^{-1} for a single component of velocity and from near 0 to 5 K for temperature) and therefore the total observation error. In contrast, current operational data assimilation schemes (Shaw et al. 1987; Courtier et al. 1998) assume the observation error is independent of location (i.e., is constant) at a given altitude. Obviously this large variability in sampling error is a critical issue for optimal data assimilation.

The recent development of ensemble forecasting techniques requires an appropriate selection of the members of the ensemble, either by selecting ensembles of forecasting algorithms, or of ensembles of initial states, or both (Toth and Kalnay 1997; Kalnay 2003). The local estimates of the observation error statistics can be used to generate realistic ensembles of observations for data assimilation and for ensemble forecasts based solely on the initial state of the atmosphere, thus providing a methodology to separate model errors from initial-state errors.

6. Comparison of sampling errors to operational observation errors

One of the standard input parameters for operational data assimilation is an estimate of the total observation error. This is typically determined from short-term forecast errors (Hollingsworth and Lonnberg 1986; Lonnberg and Hollingsworth 1986; Dee and Da Silva 1999; Dee et al. 1999). But an alternative, and probably more accurate, estimate of total observation error can be produced directly from (5.6) where the instrument error component is given by the instrument specifications, and the sampling error component is determined from the sampling pattern of the observation and the climatology of the small-scale turbulence. For a rawinsonde observation that randomly ascends through a square model grid cell of length L, the total sampling error for each velocity component is from (5.12),
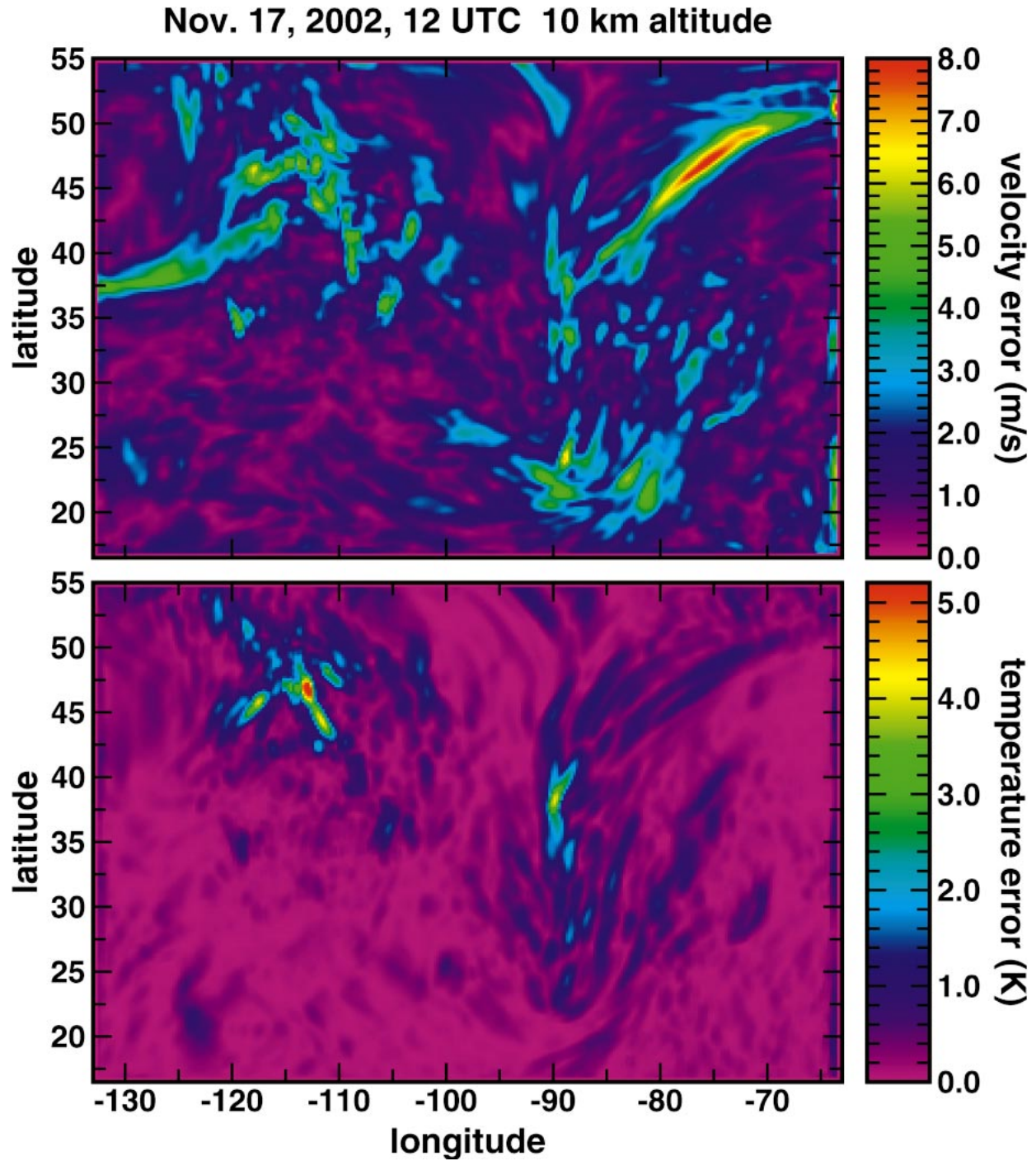
Nov. 17, 2002, 12 UTC 10 km altitude

FIG. 13. Estimates of the sampling error using (5.12) and (5.16) for a measurement of a single component of the horizontal velocity and temperature from a rawinsonde randomly located in a 100 km $\times$ 100 km measurement cell based on the estimates of $\varepsilon^{1/3}$ and C_T^2 for the RUC20 domain.

$$\langle \delta_u^2 \rangle = \langle \delta_v^2 \rangle = 0.734166 \langle \varepsilon^{2/3} \rangle L^{2/3}. \quad (6.1)$$

Substituting (4.4) for the scaling law derived from the RUC20 analysis gives

$$\langle \delta_u^2 \rangle = \langle \delta_v^2 \rangle = 0.76472 \times 10^{-3} \, L^{+0.7100} \quad (6.2)$$

at 10-km altitude ($\sim$250 hPa).

Note that (6.2) is a *universal* scaling based on the climatology of small-scale turbulence extracted from RUC20 analyses and therefore can be used to determine the contribution of the sampling error to the total observational error for *any* NWP model. For example, the NCEP (http://wwwt.emc.ncep.noaa.gov/gmb/bkistler/

oberr) and European Centre for Medium-Range Weather Forecasts (ECMWF) (Courtier et al. 1998) assimilation algorithms assign a total rawinsonde observation error in the east–west and north–south velocity components at 250 hPa of 3.2 and 3.5 m s^{-1}, respectively. For comparison, the prediction of the average sampling error from (6.2) for a T63 spectral code ($L \sim 210$ km) is 2.1 m s^{-1}, and 1.8 m s^{-1} for a T254 code ($L \sim 55$ km). It is difficult to quantify the effective sampling pattern and the correct definition of truth for NWP models; however, it is clear that the total observation error is dominated by the sampling error.

In a similar manner, the average sampling error for measurements of temperature from a rawinsonde ascending randomly through a grid cell is from (5.16),

$$\langle \delta_T^2 \rangle = 0.31464 \langle C_T^2 \rangle L^{2/3}. \qquad (6.3)$$

Substituting (4.9) for the scaling law of $\langle C_T^2 \rangle$ from the RUC20 analysis produces

$$\langle \delta_T^2 \rangle = 3.9308 \times 10^{-5} L^{+0.7132} \qquad (6.4)$$

at 10-km altitude (~ 250 hPa). The total rawinsonde observation error for temperature at 250 hPa used by NCEP is 2.3 K (http://wwwt.emc.ncep.noaa.gov/gmb/ bkistler/oberr), which using (6.4), compares with a predicted sampling error of 0.496 K for a T63 spectral code ($L \sim 210$ km) and 0.307 K for a T254 code ($L \sim 55$ km). Since the reported instrument error σ_T for the Vaisala rawinsonde is 0.5 K (Luers 1997), the total observation error would be $(\sigma_T^2 + \delta_T^2)^{1/2}$ or only 0.7 and 0.58 K for T63 and T254 spectral codes, respectively. This large disagreement may be due to the effects of global model errors in the NCEP estimation of the total observation error and larger instrument error for the global sounding network.

7. Application to estimates of upper-level turbulence climatology

A reasonable metric for the magnitude of turbulence encountered by aircraft is

$$g_{rms} = \left[\frac{1}{L} \int_0^L \varepsilon^{2/3}(x, 0) \, dx \right]^{1/2} = \bar{q}^{1/2}, \qquad (7.1)$$

which is proportional to the rms vertical acceleration experienced by an aircraft (Cornman et al. 1995). Here, $\bar{q}$ is the average of $q(x, 0) = \varepsilon^{2/3}(x, 0)$ over the linear dimension L [see (A. 1) in the appendix].

A climatology of g_{rms} can be inferred from the following arguments. From section 4b, g_{rms} is expected to obey a lognormal distribution with parameters determined from the scaling laws of small-scale turbulence derived from mesoscale model output. For RUC20, the two parameters of the lognormal distribution of g_{rms} are derived in the appendix as $\langle \log g_{rms} \rangle = -1.6737$ and SD[$\log g_{rms}$] = 0.34198.

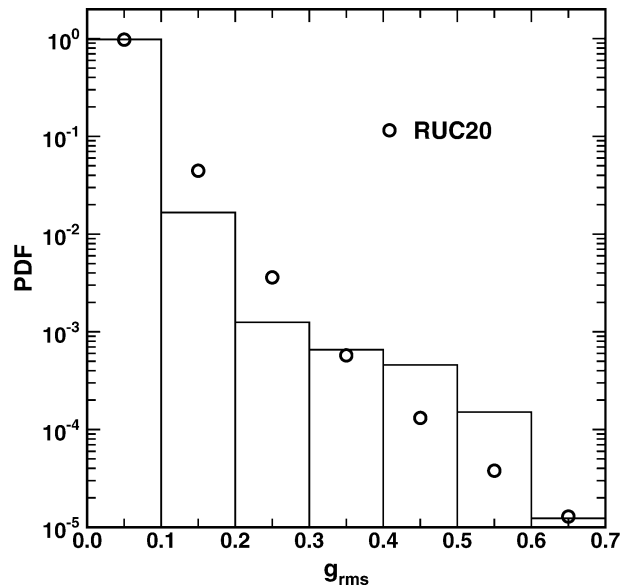The parameters for the RUC20-derived climatological

FIG. 14. Histogram of aircraft turbulence level g_{rms} and the prediction (o) from the climatology determined from RUC20 turbulence statistics.

distribution of g_{rms} can be compared to atmospheric measurements produced from routine in situ measurements of g_{rms} as described in Cornman et al. (1995). Currently, about 100 United Airlines 737-300 aircraft are equipped with the Cornman algorithm, and the data is recorded as 1-min (thus $L \approx 10$ km) median and peak values. For the calendar year 2002, slightly more than 4.4 million median and peak values were recorded. In Fig. 14 the distribution of the binned median values is plotted as a histogram and compared to the predictions of RUC20-derived lognormal g_{rms} climatology. Because the operational binning of the in situ data is very coarse, most of the data are contained in the first bin, but it does agree very well with climatological prediction. The agreement is not as good for higher-turbulence values. However, it is difficult to remove random outliers in the data, which would have a large impact on the small number of occurrences in the higher-turbulence regions. Also, the data are probably biased low because commercial aircraft will attempt to avoid known regions of strong turbulence, an effect that is difficult to quantify.

8. Application to the diagnosis of upper-level turbulence conditions

Over the years there have been continuous efforts to nowcast and forecast the occurrence upper-level aircraft-scale turbulence through the use of various manual and automated turbulence diagnostics derived from NWP model output (e.g., Ellrod and Knapp 1992; Marroquin 1998; Tebaldi et al. 2002). Most of these turbulence diagnostics do not have a rigorous basis from turbulence theory, but rather relate expected turbulence intensities to larger-scale features in the atmosphere, such as upper-
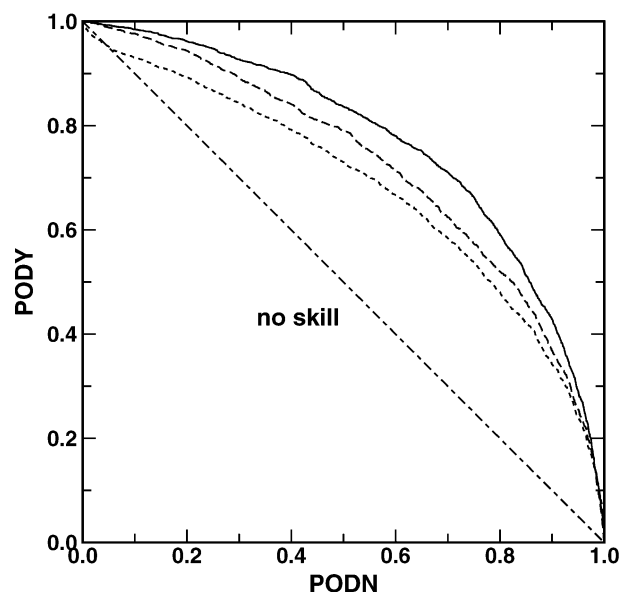
FIG. 15. The probability of moderate-or-greater turbulence detection PODY vs the probability of null turbulence detection PODN based on comparison of turbulence pilot reports to RUC20-based structure function turbulence estimates (solid line), TKE diagnostic estimates (dashed line), and Ellrod–Knapp diagnostic estimates (dotted line). The case of no skill is also shown (dot–dash line).

level fronts and jet streams. However, we can apply the ideas presented in the previous sections to use NWP model-derived structure functions to produce patterns of turbulence intensities $\varepsilon^{1/3}$ as shown in the example of Fig. 13.

The accuracy of these derived values of $\varepsilon^{1/3}$ can be assessed from the only routine observations of upper-level turbulence available, reports of encounters with turbulence by commercial airline pilots. Pilot reports are semiautomated and give information about a turbulence encounter (time, latitude, longitude, altitude, severity). There is some subjectivity associated with these reports, especially with regard to severity (reported on a five-point scale: null, light, moderate, severe, or extreme), and it must be realized that the report is based on a turbulence experience along a flight path, that is, along a line. If the model-derived diagnostics are supposed to be a gridpoint average, the correspondence to a line is not necessarily direct. Nevertheless, the relative performance of various diagnostics can be evaluated by comparisons to turbulence pilot reports as in Tebaldi et al. (2002). In that study the metric used to evaluate the performance of various turbulence diagnostics was the area contained under probability of detection (POD) curves, similar to radar operating characteristic curves. Figure 15 shows POD curves for two of the more robust diagnostics determined in the Tebaldi et al. study [the Ellrod–Knapp (1992) frontogenesis-related diagnostic and the Marroquin (1998) TKE diagnostic] compared to the $\varepsilon^{1/3}$ diagnostic based on structure function computations. The Marroquin (1998) ap-

proach uses a steady-state approximation to the TKE prognostic equation and should therefore provide similar estimates of TKE to those produced by most subgrid parameterizations used in current NWP models. All diagnostics were computed from RUC20 0000 UTC analyses at upper levels over a 3-month period from November 2002 to January 2003. A set of thresholds was assumed for each diagnostic, and given that threshold, the diagnostic performance based on comparisons to available turbulence pilot reports was evaluated for both null (as measured by PODN, the fraction of null events correctly detected) and moderate or greater turbulence reports (as measured by PODY, the fraction of moderate or greater turbulence events correctly detected). In Fig. 15 PODY is plotted against PODN for various thresholds. For small values of the chosen threshold, PODY will obviously be high, near unity, while PODN will be low, near 0, and vice versa for large values of the chosen threshold. For the range of thresholds selected, higher combinations of PODY and PODN, and therefore larger areas under the PODY–PODN curves, imply greater skill in discriminating between null and moderate-or-greater turbulence events. Based on the area under the curves, the structure-function-derived $\varepsilon^{1/3}$ diagnostic is superior to the better performers identified in the Tebaldi et al. (2002) study. This algorithm is therefore potentially quite useful for operational forecasting of upper-level and midlevel turbulence, and its utility in that regard is currently being evaluated.

9. Summary and discussion

A procedure has been presented to estimate the small-scale turbulence levels from any mesoscale NWP model using estimates of the local structure functions of model variables, provided adequate spatial resolution is available to resolve at least part of the $s^{2/3}$ power-law region. The key assumptions for a given altitude of interest are the existence of 1) a universal statistical description of the small-scale turbulence, and 2) a universal representation of the spatial filter for the model. The extensive archive of in situ aircraft measurements provided by the GASP and MOZAIC datasets has produced a robust universal description of the velocity and temperature fields at typical aircraft cruising altitudes of 10 km. This description is duplicated by the average structure functions derived from three different mesoscale NWP models (Figs. 2–4). An accurate estimate of the spatial filter of mesoscale models such as RUC20 and MM5 is produced from the average structure functions of the model variables and the universal in situ model (Figs. 5 and 6). The average velocity structure functions derived from RUC20 analyses agree well with the theoretical calculation assuming the model values are produced as an average over a 60 km × 60 km grid cell (Fig. 5), thus indicating the true spatial filtering of the model. For a high-resolution cloud-resolving model or LES model, the universal filter is obtained by comparisons

of the average structure function in convective regions to the Kolmogorov in situ model (Fig. 7). In either case, the estimates of small-scale turbulence are produced from the scaling constant of the best-fit universal structure function to the local structure function estimates (e.g., Fig. 8). These techniques could be applied to other levels (i.e., lower to mid-troposphere and stratosphere) if suitable estimates of the spatial statistics become available, either from data or reliable numerical simulation models.

The climatology of the small-scale turbulence is well defined by the PDF and the spatial statistics of the estimates. The PDF of small-scale turbulence levels for the velocity field described by ε (Fig. 9) and the temperature field described by C_T^2 (Fig. 10) is approximately lognormal; that is, the $\log_{10}$ of ε and C_T^2 have a normal or Gaussian distribution.

The parameters of the lognormal distribution depend on the averaging length L and have simple scaling laws as shown in Figs. 11 and 12. Similar results have been produced for surface-layer measurements of turbulence. Note that these scaling laws may depend slightly on the averaging time, for example winter versus summer, although the overall shape is expected to be preserved, that is, expected to maintain a lognormal distribution and have a power-law dependence on the averaging domain. The variability of the small-scale metrics for turbulence intensity is the essential component of Kolmogorov's refined similarity theory for velocity fields and the extension to the temperature field. An accurate statistical description of atmospheric turbulence must include this variability as outlined in the appendix.

There are many applications for these local turbulence estimates. These include incorporating the turbulence-related variability into optimal data assimilation methods, improving subgrid-scale turbulence parameterizations, forecasting turbulence for aviation safety, extracting turbulence climatologies from archived model fields, and calculating accurate error bars for various statistics. With regards to optimal data assimilation, in many cases the total observation error is dominated by sampling errors produced by the measurement sampling pattern. This is particularly important for rawinsonde data as shown in Fig. 13. *A point measurement is a poor representation of the average over the measurement volume, especially in regions of high turbulence. There may be little benefit in improving the instrument error of a rawinsonde measurement if the sampling error is larger than the instrument error for a large fraction of the observation space.* Next-generation forecast models will have higher resolution and therefore may require higher density and more accurate observations. More spatial averaging may be required to reduce the sampling error (e.g., line averages using Doppler lidar on multiple satellites) as an alternative to increasing the density of the rawinsonde network. A rigorous evaluation of the optimal data density and measurement accuracy will require Observing System Simulation Experiments (OSS-

Es) with the correct description of the sampling error and turbulence statistics.

The climatology extracted from 1 yr of RUC20 data has produced an estimate of the spatial distribution of turbulence that might be encountered by aircraft at cruising altitudes (Fig. 14). In addition, local estimates of ε allow useful predictions of small-scale turbulence (Fig. 15). As future operational models increase the spatial resolution and assimilate more accurate data, the true statistical description of the atmospheric fields should be revealed. This will permit still more accurate forecasts of turbulence for aviation safety as well as provide an improved understanding of turbulence processes provided the universal description of turbulence can be determined globally.

Finally, it has been shown that the climatology of small-scale turbulence is an ideal diagnostic of the total observation error used in NWP models. The sampling error for rawinsonde velocity measurements is the dominant contribution to observation error at 250 hPa; however, the sampling error for temperature at this level is much smaller than the current observation error used at NCEP. In principle, the sampling error, the total observation error, and the climatology of the errors can be produced using the local estimates of small-scale turbulence provided rigorous definitions of the errors are used (section 5). These estimates of total observation error are useful diagnostics of the performance of data assimilation algorithms since a perfect data assimilation scheme should produce an analysis error equal to the total observation error. In addition, the spatial statistics of the sampling error and the total observation error can be used for optimal data assimilation based on the model-derived local estimates of turbulence (Daley 1991; Kalnay 2003). The random spatial variations of the sampling errors should also be included in OSSEs to correctly represent the observation error statistics (Atlas 1997; Rohaly and Krishnamurti 1993). The statistical variability of small-scale turbulence and its lognormal PDF should also be included in the calculation of error bars for atmospheric statistics (Lenschow et al. 1994).

APPENDIX

Spatial Statistical Description of Small-Scale Turbulence

The universal description of the spatial statistics of small-scale turbulence can be defined by the spatial covariance $C_q(s)$ of the turbulence statistic q, which can by approximated in the upper troposphere and lower stratosphere by Eq. (4.12). The spatial covariance $C_q(s)$ can be related to the statistics of q as follows:

The spatial average of q over a line of length L is

$$\overline{q} = \frac{1}{L} \int_0^L q(x,\ 0)\ dx. \qquad (A.1)$$

If q is a homogeneous isotropic random field, then $\langle \overline{q} \rangle = \langle q \rangle$ and the variance of $\overline{q}$ is

$$\text{VAR}[\overline{q}] = \frac{2}{L} \int_0^L C_q(x,\ 0)(1 - x/L)\ dx. \qquad (A.2)$$

For the covariance model (4.12)

$$\text{VAR}[\overline{q}] = 2K_q L^{-\beta}\ f(L_0/L,\ \beta), \qquad (A.3)$$

where

$$f(\alpha,\ \beta) = \int_0^1 \frac{(1 - x)\ dx}{(\alpha^2 + x^2)^{\beta/2}}. \qquad (A.4)$$

Similarly, in 2D, the spatial average of q over a square with side of length L is

$$\overline{q} = \frac{1}{L^2} \int_0^L \int_0^L q(x,\ y)\ dx\ dy, \qquad (A.5)$$

and if $q(x,\ y)$ is a homogeneous isotropic random field, the variance of $\overline{q}$ is

$$\text{VAR}[\overline{q}] = \frac{4}{L^2} \int_0^L \int_0^L C_q(x,\ y)\left(1 - \frac{x}{L}\right)\left(1 - \frac{y}{L}\right) dx\ dy. \qquad (A.6)$$

For the covariance model (4.12)

$$\text{VAR}[\overline{q}] = 4K_q L^{-\beta} g(L_0/L,\ \beta), \qquad (A.7)$$

where

$$g(\alpha,\ \beta) = \int_0^1 \int_0^1 \frac{(1 - x)(1 - y)\ dx\ dy}{(\alpha^2 + x^2 + y^2)^{\beta/2}}. \qquad (A.8)$$

Thus the covariance $C_q(s)$ of small-scale turbulence metric q is the essential component for connecting a one-dimensional average to a two-dimensional average.

The universal constants K_q and β for the scalar field $\varepsilon^{2/3}(x,\ y)$ can be estimated from the scaling laws above. For example, the aircraft turbulence metric g_{rms} is a function of $\overline{q}$, the average of the scalar field $q(x,\ y) = \varepsilon^{2/3}(x,\ y)$ along a line of length L [(A.1)]. The ensemble average of $\overline{q}$ is approximated as the average of $\varepsilon^{2/3}(x,\ y)$ over a $L \times L$ square domain using the scaling of (4.4).

This is a good approximation because of the weak dependence on L in (4.4). If $L = 10$ km, from RUC20 analyses, $\langle \overline{q} \rangle = 1.5529 \times 10^{-3}$ m$^{4/3}$ s^{-2}. The universal scaling constants K_q and β for the field $\varepsilon^{2/3}(x,\ y)$ are determined by equating (A.7) and (4.6) assuming $L_0/L = 0$, to give $\beta = 0.37429$, $g(0,\ \beta) = 0.34853$, and $K_q = 4.21737 \times 10^{-4}$ m$^{8/3-0.37429}$ s^{-4}. The variance of $\overline{q}$ [Eq. (A.2)] then becomes $\text{VAR}[\overline{q}] = 2.6393 \times 10^{-5}$ m$^{8/3}$ s^{-4}, where $L = 10$ km and $f(0,\ \beta) = 0.98307$. The parameters of the lognormal distribution of $\overline{q}$ are given by (Monin and Yaglom 1975, p. 618; Frehlich 1992)

$$\text{VAR}[\ln \overline{q}] = \ln(1 + \text{VAR}[\overline{q}]/\langle \overline{q} \rangle^2) = 2.4803, \qquad (A.9)$$

$$\langle \ln \overline{q} \rangle = \ln \langle \overline{q} \rangle - \text{VAR}[\overline{q}]/2 = -7.7078. \qquad (A.10)$$

Since $g_{\text{rms}} = \overline{q}^{1/2}$, the parameters of the lognormal distribution of g_{rms} are

$$\langle \ln g_{\text{rms}} \rangle = \langle \ln \overline{q} \rangle/2 = -3.8539 \quad \text{and} \qquad (A.11)$$

$$\text{SD}[\ln g_{\text{rms}}] = \text{SD}[\ln \overline{q}]/2 = \text{VAR}[\ln \overline{q}]^{1/2}/2 = 0.78745. \qquad (A.12)$$

The lognormal parameters for $\log_{10} g_{\text{rms}}$ ($\log g_{\text{rms}}$) are given by $\langle \log_{10} g_{\text{rms}} \rangle = \langle \ln g_{\text{rms}} \rangle/\ln 10 = -1.6737$ and $\text{SD}[\log_{10} g_{\text{rms}}] = \text{SD}[\ln g_{\text{rms}}]/\ln 10 = 0.34198$.

REFERENCES

Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan,* **75B,** 111–130.

Benjamin, S. G., B. E. Schwartz, and R. E. Cole, 1999: Accuracy of ACARS wind and temperature observations determined by collocation. *Wea. Forecasting,* **14,** 1032–1038.

——, G. A. Grell, J. M. Brown, and T. G. Smirnova, 2004: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.,* **132,** 473–494.

Cho, J. Y. N., and E. Lindborg, 2001: Horizontal velocity structure functions in the upper troposphere and lower stratosphere. 1. Observations. *J. Geophys. Res.,* **106,** 10 223–10 232.

Clark, T. L., 1977: A small-scale dynamic model using a terrain-following coordinate system. *J. Comput. Phys.,* **24,** 186–215.

——, and W. D. Hall, 1991: Multi-domain simulations of the time-dependent Navier–Stokes equations: Benchmark error analysis of some nesting procedures. *J. Comput. Phys.,* **92,** 456–481.

Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan,* **75,** 257–288.

Cornman, L. B., C. S. Morse, and G. Cunning, 1995: Real-time estimation of atmospheric turbulence severity from in-situ aircraft measurements. *J. Aircr.,* **32,** 171–177.

Courtier, P., and Coauthors, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.,* **124,** 1783–1807.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

——, 1993: Estimating observation error statistics for atmospheric data assimilation. *Ann. Geophys.,* **11,** 634–647.

Davis, C., T. Warner, E. Astling, and J. Bowers, 1999: Development and application of an operational, relocatable, mesogamma-scale weather analysis and forecasting system. *Tellus,* **51A,** 710–727.

Deardorff, J. W., 1970: A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *J. Fluid Mech.,* **41,** 453–480.

Dee, D. P., and A. M. Da Silva, 1999: Maximum-likelihood estimation

of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Wea. Rev.,* **127,** 1822–1834.

——, G. Gaspari, C. Redder, L. Rukhovets, and A. M. Da Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part II: Applications. *Mon. Wea. Rev.,* **127,** 1835–1849.

Ellrod, G. P., and D. L. Knapp, 1992: An objective clear-air turbulence forecasting technique: Verification and operational use. *Wea. Forecasting,* **7,** 150–165.

Ferziger, J. H., and M. Perić, 2002: *Computation Methods for Fluid Dynamics.* Spinger-Verlag, 423 pp.

Frehlich, R., 1992: Laser scintillation measurements of the temperature spectrum in the atmospheric surface layer. *J. Atmos. Sci.,* **49,** 1494–1509.

——, 1997: Effects of wind turbulence on coherent Doppler lidar performance. *J. Atmos. Oceanic Technol.,* **14,** 54–75.

——, 2000: Simulation of coherent Doppler lidar performance for space-based platforms. *J. Appl. Meteor.,* **39,** 245–262.

——, 2001: Errors for space-based Doppler lidar wind measurements: Definition, performance, and verification. *J. Atmos. Oceanic Technol.,* **18,** 1749–1771.

——, and L. Cornman, 2002: Estimating spatial velocity statistics with coherent Doppler lidar. *J. Atmos. Oceanic Technol.,* **19,** 355–366.

——, S. Hannon, and S. Henderson, 1998: Coherent Doppler lidar measurements of wind field statistics. *Bound.-Layer Meteor.,* **86,** 233–256.

——, L. Cornman, and R. Sharman, 2001: Simulation of three dimensional turbulent velocity fields. *J. Appl. Meteor.,* **40,** 246–258.

Grell, G. A., J. Dudhia, and D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 138 pp.

Hollingsworth, A., and P. Lonnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus,* **38A,** 111–136.

Jaatinen, J., and J. B. Elms, 2000: On the windfinding accuracy of Loran-C, GPS and radar. *Vaisala News,* **152,** 30–33.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, 342 pp.

Khelif, D., S. P. Burns, and C. A. Friehe, 1999: Improved wind measurements on research aircraft. *J. Atmos. Oceanic Technol.,* **16,** 860–875.

Koshyk, J. N., and K. Hamilton, 2001: The horizontal energy spectrum and spectral budget simulated by a high-resolution troposphere–stratosphere–mesosphere GCM. *J. Atmos. Sci.,* **58,** 329–348.

Lane, T. P., R. D. Sharman, T. L. Clark, and H.-M. Hsu, 2003: An investigation of turbulence generation mechanisms above deep convection. *J. Atmos. Sci.,* **60,** 1297–1321.

Lenschow, D. H., 1986: Aircraft measurements in the boundary layer. *Probing the Atmospheric Boundary Layer,* D. H. Lenschow, Ed., Amer. Meteor. Soc., 39–55.

——, and P. Spyers-Duran, 1989: Measurement techniques: Air motion sensing. NCAR Research Aviation Facility Bulletin 23, NCAR, Boulder, CO, 36 pp.

——, J. Mann, and L. Kristensen, 1994: How long is long enough when measuring fluxes and other turbulence statistics? *J. Atmos. Oceanic Technol.,* **11,** 661–673.

Lindborg, E., 1999: Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *J. Fluid Mech.,* **388,** 259–288.

Lonnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus,* **38A,** 137–161.

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

Luers, J. K., 1997: Temperature error of the Vaisala RS90 radiosonde. *J. Atmos. Oceanic Technol.,* **14,** 1520–1532.

Marroquin, A., 1998: An advanced algorithm to diagnose atmospheric turbulence using numerical model output. Preprints, *16th Conf. on Weather Analysis and Forecasting,* Phoenix, AZ, Amer. Meteor. Soc., 79–81.

Moeng, C.-H., 1984: A large-eddy-simulation model for the study of planetary boundary-layer turbulence. *J. Atmos. Sci.,* **41,** 2052–2062.

——, and P. P. Sullivan, 1994: A comparison of shear- and buoyancy-driven planetary boundary layer flows. *J. Atmos. Sci.,* **51,** 999–1022.

Monin, A. S., and A. M. Yaglom, 1975: *Statistical Fluid Mechanics: Mechanics of Turbulence.* Vol. 2, MIT Press, 874 pp.

Nastrom, G. D., and K. S. Gage, 1985: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.,* **42,** 950–960.

Papoulis, A., 1965: *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 583 pp.

Pielke, Roger A., Sr., 2002: *Mesoscale Meteorological Modeling.* Academic Press, 676 pp.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1986: *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 963 pp.

Rohaly, G. D., and T. N. Krishnamurti, 1993: An observing system simulation experiment for the laser atmospheric wind sounder (LAWS). *J. Appl. Meteor.,* **32,** 1452–1471.

Sharman, R., and R. Frehlich, 2003: Aircraft scale turbulence isotropy derived from measurements and simulations. *Proc. AIAA 41st Aerospace Science Meeting and Exhibit,* Reno, NV, AIAA, 2003-194.

Shaw, D. B., P. Lonnberg, A. Hollingsworth, and P. Unden, 1987: Data assimilation: The 1984/85 revision of the ECMWF mass and wind analysis. *Quart. J. Roy. Meteor. Soc.,* **113,** 533–566.

Sullivan, P. P., T. W. Horst, D. H. Lenschow, C.-H. Moeng, and J. C. Weil, 2003: Structure of subfilter-scale fluxes in the atmospheric surface layer with application to large-eddy simulation modelling. *J. Fluid Mech.,* **482,** 101–139.

Tebaldi, C., D. Nychka, B. G. Brown, and R. Sharman, 2002: Flexible discriminant techniques for forecasting clear-air turbulence. *Environmetrics 2002,* **13,** 859–878.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

Tung, K. K., and W. W. Orlando, 2003: The k^{-3} and $k^{-5/3}$ energy spectrum of atmospheric turbulence: Quasigeostrophic two-level model simulation. *J. Atmos. Sci.,* **60,** 824–835.

Vinnichenko, N. K., 1970: The kinetic energy spectrum in the free atmosphere—1 second to 5 years. *Tellus,* **22,** 159–166.