# Sampling Error in Aircraft Flux Measurements Based on a High-Resolution Large Eddy Simulation of the Marine Boundary Layer

Grant W. Petty[1]

[1]Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, 1225 W. Dayton St, Madison, WI, 53706, USA

**Correspondence:** Grant W. Petty (gwpetty@wisc.edu)

**Abstract.** A high-resolution (1.25 m) LES simulation of the nocturnal cloud-topped marine boundary layer is used to evaluate random error as a function of continuous track length $L$ for virtual aircraft measurements of turbulent fluxes of sensible heat, latent heat, and horizontal momentum. Results are compared with the theoretically derived formula of Lenschow and Stankov (1986). In support of these comparisons, we also evaluate and document the relevant integral length scales and correlations

5 and show that for heights up to approximately 100 m ($z/z_i$ = 0.12), the length scales are accurately predicted by empirical expressions of the form $I_f = Az^b$. The Lenschow and Stankov expression is found to be remarkably accurate at predicting the random error for shorter flight tracks, but our empirically determined errors decay more rapidly with $L$ than the $L^{-1/2}$ relationship predicted from theory. Consistent with earlier findings, required track lengths to obtain useful precision increase sharply with altitude.

## 1 Introduction

The eddy correlation (EC) method has long been the principal means of making field measurements of turbulent fluxes of energy, momentum, and trace gases in the planetary boundary layer (PBL). While towers are commonly used to measure fluxes over longer time periods (weeks to years) at fixed locations, aircraft are the preferred platform for obtaining estimates of fluxes

15 over larger areas, albeit for far shorter periods of time.

By their nature, EC estimates of fluxes are averages of quantities that randomly fluctuate on time scales ranging from fractions of a second to hours and over distance scales from centimeters to kilometers. Thus, in addition to the usual factors affecting measurement quality from any instrument, such as calibration, precision, time response, and site representativeness, the quality of eddy correlation estimates are critically subject to statistical sampling error as well as to potentially restrictive

20 assumptions about temporal stationarity and/or spatial homogeneity.

In this paper, we are concerned exclusively with the problem of random sampling error in the context of aircraft flux measurements relying on in situ (e.g., gust probe) measurements of fluctuating wind vector and scalar quantities directly along the

flight path. The key question that has been asked for decades is, how long of a flight path — whether consisting of a single extended leg or a series of shorter legs over a smaller region — is sufficient to obtain turbulent flux estimates of useful preci-

25  sion? This problem was examined theoretically, based on statistical models of turbulence, by Lenschow and Stankov (1986), Lenschow et al. (1994), and Mann and Lenschow (1994), with additional refinement of these ideas by Mahrt (1998). Others took an empirical approach, comparing aircraft flux estimates with those from nearby fixed towers (Desjardins et al., 1989; Grossman, 1992; Mahrt, 1998). Building on the theoretical arguments of Mahrt (1998), Hollinger and Richardson (2005) examined random errors in tower flux measurements; their findings have potential applicability to aircraft measurements as

30  well.

Finkelstein and Sims (2001) used what was characterized as a more general yet rigorous derivation to arrive at a different expression for the flux sampling error. Based on the properties of test data sets for several variables, their computed flux errors were typically about a factor of two larger than those predicted by the expressions of Mann and Lenschow (1994).

More recently, it has become possible to use large eddy simulation (LES) models to resolve turbulent fluctuations of mass,

35  energy, and motion in the planetary boundary layer at fairly fine scales. Virtual tower or aircraft can then sample these fluctuations in a manner consistent with the platform (Schröter et al., 2000; Sühring and Raasch, 2013; Sühring et al., 2019). An attractive feature of this approach is the unique ability to compare the resulting flux estimate with the known "true" domain-averaged value, in contrast to the case for real-world comparisons between inherently dissimilar and imperfect aircraft and tower measurements of the area-averaged flux. The LES approach is also gaining considerable interest for its ability to exam-

40  ine the measurement and interpretation of flux measurements in the presence of spatial inhomogeneities (Sühring and Raasch, 2013; Sühring et al., 2019).

The single most important limitation of the LES approach has been the difficulty of accommodating the full range of spatial scales relevant to turbulent flux measurements. It is therefore noteworthy that Matheou (2018) has undertaken an LES run of exceptional grid size ($4096{\times}4096{\times}1200$) and fine resolution (1.25 m) to simulate a nocturnal cloud-topped marine boundary

45  layer over a uniform $5.12{\times}5.12$ km domain.

The present paper examines the aircraft sampling problem utilizing the single final time step of that simulation. The nature of the simulation allows us to cleanly isolate and study the statistical sampling problem over a fairly wide range of spatial wavelengths without the need to consider complicating factors such as spatial inhomogeneity or subgrid-scale flux contributions.

50  Because the simulation utilized cyclic lateral boundary conditions, it is possible to construct simulated flight paths that are continuous over much longer distances and yet do not resample the same locations in the domain. We are thus able to focus narrowly on the precision of aircraft flux measurements as a function of flight track length in the particular environment represented by this simulation. Our methods and motivations are in many ways similar to those of Schröter et al. (2000) and Sühring and Raasch (2013). However, we place greater emphasis on documenting the specific properties of the environment

55  that most strongly affect sampling error; in particular, the integral length scale $I_f$ and the spatial correlation $\rho_{w,\psi}$ between vertical velocity $w$ and an arbitrary scalar variable $\psi$, both of which appear in theoretical expressions for sampling error derived by Lenschow and Stankov (1986) and Lenschow et al. (1994).

For ease of comparison with the earlier error studies using LES, Table 1 summarizes key features of those studies. The present study benefits from both the finest grid resolution and largest grid size, but the overall horizontal domain dimension is smaller than that used by the two most recent previous studies. Ours is also unique in using an LES to examine the aircraft flux sampling problem in a marine rather than continental environment.

## 2 Data

### 2.1 LES description

The LES utilizes a buoyancy-adjusted stretched-vortex model (BASVM) described in detail by Chung and Matheou (2014) and Matheou and Chung (2014). In what is possibly the largest LES model run ever published, Matheou (2018) used the BASVM to simulate a nocturnal stratocumulus case over the Pacific ocean southwest of Los Angles corresponding to the 10 July 2001 research flight (RF01) of the second Dynamics and Chemistry of Marine Stratocumulus (DYCOMS-II) field study (Stevens et al., 2003). The observed stratocumulus case was described by mission participants as "very homogeneous" (Earth Observing Laboratory, 2007).

The LES grid resolution is 1.25 m in $x$, $y$, and $z$, and the grid size is $4096 \times 4096 \times 1200$, corresponding to a $5.12 \times 5.12 \times 1.5$ km domain with cyclic lateral boundary conditions. Because of the considerable computing resources required, it was only possible to simulate the atmospheric boundary layer for 2 h of physical time. Nevertheless, the simulation required 36 days of wall-clock time on 4,096 CPU cores at the NASA Advanced Supercomputing (NAS) Division at Ames Research Center.

The single final time step of the simulation, comprising a total of $\sim$800 GB of gridded data, was utilized for this study. Output included all grid-resolved model variables, including temperature $T$, specific humidity $q$, wind velocity components $u$, $v$, and $w$, and cloud water content $q_l$.

### 2.2 Simulation properties

#### 2.2.1 Environment

Consistent with the observed environment on the flight date, the model was initialized with a boundary layer depth $z_i = 840$ m delineated by a sharp 8.5 K temperature inversion. Below $z_i$, the initial potential temperature $\theta$ was 289 K and the specific humidity $q$ was 9 g kg$^{-1}$. These values did not change significantly during the simulated evolution of the boundary layer (Fig. 1). A prescribed sea surface temperature of 292.5 K was used, leading to domain-averaged sensible and latent heat fluxes of 15 W m$^{-2}$ and 115 W m$^{-2}$, respectively. The wind was initialized as geostrophic; at the conclusion of the simulation, the mean horizontal wind components $\overline{u}$ and $\overline{v}$ below $z_i$ were nearly uniform at 6.7 m s$^{-1}$ and $-5.0$ m s$^{-1}$, respectively, corresponding to a wind direction of 307° (approximately northwesterly) and speed of 8.4 m s$^{-1}$.

The cloud base and top, defined here as the lowest and highest levels at which non-zero cloud liquid water occurred at any point in the horizontal domain, were found at 500 m and 885 m, respectively. We are concerned in this study primarily with the sub-cloud boundary layer at and below 400 m or $0.48 z_i$.

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU

### 2.2.2 Turbulent structure and domain-averaged fluxes

90 Many details concerning the model run and the overall turbulent structure of the simulated boundary layer are given by Matheou (2018). Here we focus on those grid-resolvable turbulent properties most directly relevant to an understanding of the simulated flux measurements.

Figure 2a shows that, at 400 m height ($0.48z_i$), the LES produces a vertical velocity spectrum approximately obeying a Kolmogorov power law for the vertical velocity field over more than two decades, from wavelengths greater than 1 km down 95 to $\sim$8 m or about three times the Nyquist wavelength of 2.5 m. At this height, peak energy in the spectrum is near 1 km wavelength. At the much lower height of 40 m ($0.05z_i$), peak energy is found near 150 m, and there is substantially less energy at longer wavelengths. In both cases, the peak energy in $w$ is thus found near $\sim$4 times the physical distance $z$ from the lower boundary, consistent with dominant contributions from circular eddies of radius $z$ centered at height $z$.

The spectra for temperature (Fig. 2b) and specific humidity (Fig. 2c) reveal considerably more energy at the longest wave- 100 lengths, even at 40 m height, while also rolling off earlier than $w$ at the short wavelength end. The spectrum for temperature is also somewhat flatter than for either $w$ or $q$. As expected, the spectrum for horizontal wind speed $U$ (Fig. 2d) is notable in having no significant dependence on height. Overall, the spectra of scalar variables give reason for some confidence in the realism of the LES for wavelengths spanning the range from $\sim$10 m to $\sim$1 km.

Throughout this paper, we will highlight results at four representative heights: 10 m, 40 m, 100 m, and 400 m. These 105 correspond to approximately $0.01z_i$, $0.05z_i$, $0.12z_i$, and $0.48z_i$, respectively.

Horizontal cross-sections of temperature $T'$, specific humidity $q'$, and vertical velocity $w'$ are presented for these heights in Figs. 3–5, where the primed quantities indicate deviations from the domain-averaged value at each height. These show the expected evolution from predominantly fine-scale structure near the surface to consolidated perturbations of order 1 km diameter at 400 m. In the $w'$ field near the surface, there is a visually obvious alignment of structures with the mean wind. This 110 directional anisotropy is far less apparent in $q'$ and is completely absent from $T'$ at low levels.

As shown in Fig. 6, the amplitudes of turbulent fluctuations also vary with height, with $\sigma_T$ decreasing from 0.12 K at 10 m to 0.04 K at 400 m and $\sigma_q$ decreasing from 0.36 g kg$^{-2}$ to 0.13 g kg$^{-2}$. $\sigma_w$, on the other hand, increases from 0.32 m s$^{-1}$ to 0.72 m s$^{-1}$.

The local contributions to the turbulent fluxes of sensible and latent heat are

115 $$H' = \rho C_\mathrm{p} T' w' \tag{1}$$

and

$$E' = \rho L q' w', \tag{2}$$

where $\rho$ is the air density, $C_\mathrm{p} = 1005$ J kg$^{-1}$ K$^{-1}$ is the specific heat capacity at constant pressure, and $L \approx 2.5 \times 10^6$ J kg$^{-1}$ is the latent heat of vaporization of water. Cross-sections of $H'$ and $E'$ are shown in Figs. 7 and 8.

120 We additionally compute the local $u$ and $v$ components of momentum flux as

$$(\tau'_u, \tau'_v) = \rho \cdot (u'w', v'w'). \tag{3}$$

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU

Where appropriate in the discussion below, the horizontal wind vector $(u, v)$ is rotated into a coordinate system aligned with the mean wind direction so as to isolate the along-wind and crosswind components $(U_\parallel, U_\perp)$.

Domain-averaged vertical fluxes and turbulent kinetic energy (TKE) are depicted in Fig. 9. These do not include parame-
125 terized subgrid-scale fluxes, which become negligible relative to resolved eddy fluxes above about the 5th grid level, or 8 m. These fluxes serve as the "truth" values against which simulated aircraft flux measurements will be evaluated. As expected, the profile of sensible heat flux $H$ (Fig. 9a) is quite linear except within cloud, where latent heating due to condensation sharply increases the positive correlation between $T$ and $w$.. It crosses through zero near 445 m $(0.53z_i)$. The profile of latent heat flux is noticeably less smooth, as are $\tau_u$ and $\tau_v$ (Fig. 9b). Turbulent kinetic energy varies only weakly below cloud and has a sharp
130 maximum at cloud top, as expected. (Fig. 9c).

Figure 10 presents ogive (cumulative distribution) plots of the spectral contributions to each flux. For all four flux variables, there is a pronounced shift from predominantly short-wavelength contributions at the lowest levels to much lower-frequency contributions at 400 m. In particular, we find that 90% of the total sensible heat flux contribution comes from the wavelength ranges [6 m, 146 m], [8 m, 0.51 km], [12 m, 1.0 km], and [17 m, 2.6 km] at heights of 10 m, 40 m, 100 m, and 400 m,
135 respectively, with qualitatively similar ranges for latent heat flux and momentum flux. Overall, these results suggest that 100 m is the height at which spectra of flux contributions for all variables are least likely to be biased by either the unphysical spectral roll-off below 10 m (see Fig. 2) or the LES domain size of 5.12 km.

## 3 Integral length scales

### 3.1 Significance

140 Central to the problem of estimating random error in aircraft measurements of turbulent fluxes is the integral length scale, also known as the autocorrelation length, of the fields being measured (Lenschow and Stankov, 1986; Lenschow et al., 1994; Sühring and Raasch, 2013):

$$I_f \equiv \int_0^{r_0} A_f(r) \, dr, \tag{4}$$

where $A_f(r)$ is the autocorrelation of generic spatial function $f(x, y)$ at displacement $r = \sqrt{x^2 + y^2}$ in any specified direction,
145 $A_f(0) \equiv 1$, and $r_0$ is the distance to the first zero crossing or one-half of the domain diameter (in this case 2.56 km), whichever is smaller. In the ideal "red noise" case, $A_f(r) = \exp(-r/I_f)$, and $A_f(I_f) = e^{-1}$. The length scale is thus useful for characterizing the minimum track length and/or spatial separation required in order to be able to treat two sets of measurements as statistically independent.

Viewed simplistically, for a track length $L = N I_f$, the random error in $\overline{f}$ as an estimate of the true ensemble value $\langle f \rangle$ should
150 decrease with $N^{-1/2}$, consistent with the calculation of standard error for any set of $N$ noisy measurements. We agree with Richardson et al. (2006) that it is more natural to focus on the absolute rather than relative random error, as the relative error becomes extremely large or even undefined in the common case that that the flux approaches or crosses zero; for example,

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

either at sunrise or sunset or near the altitude at which sensible heat flux switches signs. With that in mind, we rearrange equation (7) of of Lenschow and Stankov (1986) to give the following absolute random uncertainty

$$\sigma_\psi = \left[2(\rho_{w,\psi}^{-2} + 1)\frac{I_{w\psi}}{L}\right]^{1/2} |F_\psi| \tag{5}$$

where $F_\psi = \langle w'\psi' \rangle$ is the desired turbulent flux of scalar variable $\psi$, $\rho_{w,\psi}$ is the zero-lag Pearson correlation coefficient of $w'$ with $\psi'$, and $I_{w\psi}$ is the integral scale length of $w'\psi'$.

Lenschow et al. (1994) and Mann and Lenschow (1994) refer to the difficulty of obtaining estimates of $I_{w\psi}$ and propose theoretically derived substitutions based on the more commonly available $I_w$ and $I_\psi$. Subsequent authors, citing Mann and Lenschow (1994), variously utilize

$$I_{w\psi} \le \sqrt{I_w I_\psi} \qquad ; \qquad \text{e.g., Sühring and Raasch (2013)} \tag{6}$$

or

$$I_{w\psi} \le \frac{\sqrt{I_w I_\psi}}{\rho_{w,\psi}} \qquad ; \qquad \text{e.g., Bange et al. (2002),} \tag{7}$$

where $\rho_{w,\psi}$ is the spatial correlation between $w$ and $\psi$. Because we are able to compute all of the above length scales and correlations from the LES fields, it will be possible to empirically assess the validity of (5)–(7).

## 3.2 Determination from LES

The two-dimensional autocorrelation of the full model domain was obtained at a series of heights for each turbulent field of interest. $I_f$ was then determined in two orthogonal directions: one aligned with the mean wind ("parallel") and the other perpendicular ("crosswind"). Result are shown in Fig. 11. Note that for momentum fluxes, the wind vector $(u, v)$ was projected onto the parallel and crosswind directions to obtain the rotated vector $(U_\parallel, U_\perp)$. Thus, the integral length scale labeled "$U_\perp$ Parallel", for example, describes that computed for the turbulent flux of the momentum component perpendicular to mean wind measured along a track parallel to the mean wind.

A number of general observations can be made concerning the computed values of $I_f$:

- For every variable, $I_f$ at and below $z = 100 \text{ m} = 0.12z_i$, corresponding roughly to the surface layer, is well-approximated by a straight line on a log-log plot, implying an empirical relationship of the form $I_f \approx Az^b$. For some variables, such as the specific humidity (Fig. 11a) and the sensible and latent heat fluxes (Figs. 11f,g), the validity of such a relationship seems to extend considerably higher.

- For most variables, $I_f$ is larger for tracks parallel to the mean wind, as expected, than for crosswind tracks, often considerably larger. Two interesting exceptions are found in $I_f$ for temperature above 40 m (Fig. 11b) and for $U_\parallel$ above 100 m, where the crosswind value of $I_f$ is up to 50% greater than for the parallel direction.

- For many variables ($q$, $U_\perp$, $w$, $H$, and $E$), the difference between the parallel and crosswind values of $I_f$ is largest near the surface and decreases to near zero at higher altitudes. For temperature , however, the reverse is true (Fig. 11b).

- The three turbulent wind components behave quite differently, with $I_f$ for $U_\parallel$ and $U_\perp$ exhibiting a relatively weak and non-linear dependence on height, while for the vertical wind component $w$, the dependence is strong. Indeed for the crosswind case, $I_f$ for $w$ is almost perfectly proportional to $z$.

- Of all the variables considered, only the turbulent kinetic energy exhibits an integral length scale that is virtually constant with height throughout the full depth of the model domain (Fig. 11i).

The calculated $I_f$ results below 100 m were used to obtain coefficients $A$ and $b$ for the aforementioned power-law fit. These coefficients are provided in Table 2. Alternatively, $I_f = A_\star z_\star^b$, where $z_\star \equiv z/z_i$, $A_\star = A z_i^b$ and $z_i = 840$ m. It is not known to what degree the results found here are transferable to other environments.

It is noteworthy that the exponents $b$ found here for fluxes of sensible heat, latent heat, and horizontal momentum are quite different than the exponent of $-1/3$ or $-1/2$ obtained theoretically by Lenschow and Stankov (1986) for all of these (their equation 11), being closer to $-1$ for the crosswind length scales and $-2/3$ for the along-wind length scales. It is beyond the scope of this paper to investigate the reasons for the discrepancy.

## 3.3 Comparison with proxy estimates

Apparent in the above results is that (6) and (7) generally yield large overestimates of $I_{w\psi}$ for any scalar variable $\psi$. For example, at 100 m height and measured in the crosswind direction, $I_q = 220$ m and $I_w = 73$ m. $\sqrt{I_q I_w}$ is then 127 m, a factor of 2.5 larger than the directly computed $I_{wq} = 51$ m. In general, the minimum overestimate at or below 400 m was found to be about a factor of 2 or greater, with ratios in excess of 3–5 typical at lower heights. The difference increases further if one uses the correlation coefficient $\rho_{w,\psi} < 1$ in the denominator per (7). Specific combinations of variables and heights can be easily examined using the relationships in Table 2.

## 3.4 Correlations

Also required in the evaluation of random flux errors from (5) is the correlation $\rho_{w,\psi}$ between the vertical velocity $w'$ and the scalar variable $\psi'$ of interest. Computed profiles of $\rho_{w,\psi}$ are given in Fig. 12. For horizontal momentum, we computed $\rho_{w,U} = \sqrt{\rho_{w,u}^2 + \rho_{w,v}^2}$.

# 4 Simulated aircraft measurements

## 4.1 Track definitions

The central purpose of this paper is to investigate the empirical relationship between sampling error in turbulent flux measurements and the length of a continuous, ideal aircraft track through the virtual atmosphere represented by the LES domain. We require each track to be perfectly cyclic, albeit with a period $L$ substantially longer than the dimensions of the domain. Specifically, track angles relative the $x$-axis are defined by $\phi_n = \arctan(n)$. For $\phi_1 = 45°$, the track makes a single pass through the

Atmospheric
Measurement
Techniques

EGU

Open Access

Discussions

domain before returning to its starting point, and the track length is 7.2 km. A continuation of the track would simply resample the same grid points. For $\phi_2 = 63.4°$, the track makes two widely separated passes through the domain before returning to its starting point, and the non-repeating track length is 11.4 km. The maximum value of $n$ used was 20, for a 102.5 km continuous, non-overlapping path.

For any given $n$, starting points for different tracks of the same length are uniformly distributed along the boundary of the domain, maintaining prescribed minimum separation (measured perpendicular to the paths) intended to ensure a reasonable degree of statistical independence. These minimum separations were chosen to be greater than $I_f$ for all relevant variables at the level in question; we used 20 m, 50 m, 100 m, and 150 m at heights of 10 m, 40 m, 100 m, and 400 m, respectively.

The effect of the minimum separation requirement is to achieve a roughly constant overall sampling density for the ensemble of tracks of a given length. Thus, $N_n$ is small for long tracks and larger for short tracks. By taking mirror images and 90° rotations of the tracks defined in this way, the total number of distinct tracks available is actually $4N_n$, except for $n = 1$, in which case it is $2N_n$ owing to duplication when rotating and flipping the 45° track patterns. Figure 13 depicts an example track for $n = 3$ (top) and the full corresponding set of 32 distinct tracks (bottom) for a relatively wide minimum separation of 200 m used for visual clarity; the number of distinct tracks is significantly greater in the actual analysis.

Note that the algorithm described above is more elaborate than the fixed-angle track definitions utilized by Schröter et al. (2000) and Sühring et al. (2019). To reiterate, we are requiring each track of length $L_i$ to be cyclic with period $L_i$, and we are distributing the realizations of a given track length so as to both maximize and equalize their separation, subject to the aforementioned minimum separations. This allows us to obtain an optimal set of quasi-independent tracks. Constraining the tracks to be cyclic also means that there is by definition no trend or unsampled low frequency contribution to be concerned about—"measured" fluxes are the same regardless of the starting point along the track, as long as the total distance followed by the virtual aircraft is $L_i$.

### 4.2 Virtual flux observations

The "truth" value for each comparison is the domain-averaged flux value computed at each level (Fig. 9). These values utilized the deviations (primed quantities) from the true ensemble mean values for the entire domain. As measured against the true values, aircraft measurements are subject to two statistical sources of error: 1) error in the determination of the mean values relative to which primed quantities will be calculated , and 2) insufficient sampling of the fluctuating primed quantities themselves. Our virtual measurements simulate both effects by utilizing the track-determined mean values rather than the domain-averaged values. Both sources of error decrease asymptotically to zero in the limit of an infinitely long flight track, and the flux estimate will converge on the true ensemble value.

In the finite 5.12 km×5.12 km horizontal domain of our LES, there is the additional issue that a single long flight track can effectively sample the entire domain, so that the aircraft measurement can converge almost perfectly on the "true" value from measurements along a long but finite path. In principle, then, the results obtained herein for long flight tracks should be considered lower bounds on the expected random error. However, because of the minimum separation maintained between parallel segments of a given tracks, we believe this potential bias is unimportant for the flight track lengths considered here.

More important may be the inability of the finite LES domain to reproduce flux variations with wavelengths greater than the domain size. As previously discussed for Fig. 10, this seems likely to be an issue mainly at heights of $\sim$400 m and above.

Finally, a notable constraint rooted in the availability of only a single time step from the LES is that the virtual aircraft effectively traverses each path instantaneously. We are in effect relying on Taylor's "frozen turbulence" hypothesis and, following

250  Lenschow and Stankov (1986), postulate that this hypothesis is valid in the present case if $I_f < VT_f$, where $I_f$ is the integral length scale, $V$ is the speed of the aircraft, and $T_f$ is the autocorrelation time of the flux field being measured. Unfortunately, we are unable to determine $T_f$ from our single time-step, but for a real-world $V \approx 85$ m sec$^{-1}$ similar to that of the Univerity of Wyoming KingAir, we would only require $T_f > 1$ s to satisfy the above inequality for flux measurements at a height of 100 m, based on the values of $I_f$ seen in Fig. 11.

255  ## 5 Results

Raw sample results are presented in Figs. 14–16, with each blue dot representing the estimated flux from a single flight track of the indicated length. The true domain-averaged values are indicated as dashed black lines, and the root-mean-square (RMS) deviation from the true value is depicted as a red dashed line. For all flux variables, there is the expected decay in sampling error as the track length becomes longer.

260  Noteworthy features include the following:

- Overall scatter is largest at the highest altitudes, consistent with the longer integral length scales $I_f$ and with the common observation by previous authors that error tends to increase with height.

- For sensible heat flux at 400 m (Fig. 14d), the magnitude of the scatter is often far in excess of the very low (1.8 W m$^{-2}$) mean flux itself, so that shorter track lengths cannot even distinguish the sign of the mean flux.

265  - Relative to the mean values, scatter is particularly large for momentum flux. Because this quantity is positive-definite ($\tau = \sqrt{\tau_u^2 + \tau_v^2}$), the distribution of errors at 400 m is strongly skewed and positively biased.

In Figs. 17–19, we take a closer look at the dependence of sampling error on path length and compare our results with the random error predicted by (5). The black dots correspond to the empirical RMS error values previously represented in Figs. 14–16 as red dashed curves. The black dashed lines represent fits to those values. The accompanying power-law expression for

270  each fit appears in black. In red we have the minimum track length $L_{10}$ required for 10% relative accuracy. Finally, the blue line represents (5), with the parameter values required by that equation appearing in blue as well, where both $I_f$ and the correlation coefficients $\rho_{w,\psi}$ are computed directly from the LES-generated fields

Noteworthy findings include the following:

- Equation (5) due to Lenschow and Stankov (1986) and using the independently computed parameter values is remarkably

275  accurate at predicting the random error for shorter track lengths.

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

- Empirical fits generally show a steeper decrease with track length than the $L^{-1/2}$ relationship predicted by (5), with exponents ranging from $-0.64$ to $-0.70$ at 10 m and 40 m, to nearly $-1$ for sensible and latent heat fluxes measured at a height 400 m. It is not clear whether the more rapid decay in error is primarily related to the finite domain size—with this issue being most likely problematic at 400 m—or to other departures from the idealized statistical model of turbulence assumed by Lenschow et al. (1994).

- For flights at and below 100 m, track lengths of approximately 30 km or less are sufficient to achieve 10% precision in sensible and latent heat flux estimates in this particular environment. Required track lengths for momentum flux $\tau$ are considerably longer; nearly 90 km at 100 m height.

- At a height of 400 m, no reasonable track length is long enough to adequately measure the low sensible heat and momentum fluxes encountered there.

## 6  Conclusions

In this paper, we utilized the high-resolution large-eddy simulation of the marine boundary layer by Matheou (2018) to determine random flux sampling errors by a virtual aircraft flying tracks of various lengths at heights of 10 m, 40 m, 100 m, and 400 m. We eliminated the need for detrending by constructing these tracks to be cyclic with periods of the specified lengths.

We compared these empirical results with the theoretically derived expression of Lenschow and Stankov (1986). In support of these comparisons, we computed the required integral length scales $I_f$ and found that these are well described by expressions of the form $I_f = Az^b$ for $z \le 100$ m, with coefficents $A$ and $b$ given in Table 2. The empirical exponents $b$ are generally between about 2/3 and 1 for fluxes of sensible heat, latent heat, and momentum, in contrast to Lenschow and Stankov (1986), who found exponents of 1/3 to 1/2. The reason for this apparent discrepancy is unknown. We also evaluated the commonly cited approximation for the integral length scale for fluxes given by (6) or (7) and found that it consistently overestimated the true value by at least a factor of 2–5.

Using directly computed parameter values, we found that (5) is remarkably accurate at predicting random errors for short track lengths. However, it predicts a decay in error proportional to $L^{-1/2}$, whereas our empirical results show substantially more rapid decay in many cases. It is uncertain to what degree the difference is due to the finite domain size of the LES, to departures from the idealized statistical assumptions of Lenschow and Stankov (1986), or perhaps to a combination of both. It seems less likely that the first of these issues would affect the results for measurements at lower heights, as there is then very little contribution to total fluxes by wavelengths greater than 1 km, according to Fig. 10.

In general, we find that track lengths of $\sim$30 km or less are sufficient for measuring sensible and latent heat fluxes in this simulated environment to better than 10% precision, provided that flights are conducted at or below 100 m. Substantially longer flight tracks are required for momentum flux, a result traceable to the low correlation $\rho_{U,w}$ (and thus low corresponding fluxes) between the horizontal and vertical wind components even near the surface.

*Code and data availability.* Data sets and Jupyter notebook files utilized in the analysis may be requested by sending a 1 TB USB drive to the corresponding author.

*Author contributions.* The author is solely responsible for the analysis and discussion in this paper.

310 *Competing interests.* The author declares that he has no competing interests.

Atmospheric
Measurement
Techniques
Discussions

# References

Bange, J., Beyrich, F., and Engelbart, D. A.: Airborne measurements of turbulent fluxes during LITFASS-98: Comparison with ground measurements and remote sensing in a case study, Theoretical and applied climatology, 73, 35–51, 2002.

Chung, D. and Matheou, G.: Large-eddy simulation of stratified turbulence. Part I: A vortex-based subgrid-scale model, Journal of the
320    Atmospheric Sciences, 71, 1863–1879, 2014.

Desjardins, R., MacPherson, J., Schuepp, P., and Karanja, F.: An evaluation of aircraft flux measurements of CO 2, water vapor and sensible heat, in: Boundary Layer Studies and Applications, pp. 55–69, Springer, 1989.

Earth Observing Laboratory: DYCOMS-II C-130 Mission Summaries, http://catalog.eol.ucar.edu/dycoms/catalog/missions/, 2007.

Finkelstein, P. L. and Sims, P. F.: Sampling error in eddy correlation flux measurements, Journal of Geophysical Research: Atmospheres,
325    106, 3503–3509, 2001.

Grossman, R. L.: Sampling errors in the vertical fluxes of potential temperature and moisture measured by aircraft during FIFE, Journal of Geophysical Research: Atmospheres, 97, 18 439–18 443, 1992.

Hollinger, D. and Richardson, A.: Uncertainty in eddy covariance measurements and its application to physiological models, Tree physiology, 25, 873–885, 2005.

330    Lenschow, D., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, Journal of Atmospheric and Oceanic Technology, 11, 661–673, 1994.

Lenschow, D. H. and Stankov, B. B.: Length scales in the convective boundary layer, Journal of the Atmospheric Sciences, 43, 1198–1209, 1986.

Mahrt, L.: Flux sampling errors for aircraft and towers, Journal of Atmospheric and Oceanic technology, 15, 416–429, 1998.

335    Mann, J. and Lenschow, D. H.: Errors in airborne flux measurements, Journal of Geophysical Research: Atmospheres, 99, 14 519–14 526, 1994.

Matheou, G.: Turbulence structure in a stratocumulus cloud, Atmosphere, 9, 392, 2018.

Matheou, G. and Chung, D.: Large-eddy simulation of stratified turbulence. Part II: Application of the stretched-vortex model to the atmospheric boundary layer, Journal of the Atmospheric Sciences, 71, 4439–4460, 2014.

340    Richardson, A. D., Hollinger, D. Y., Burba, G. G., Davis, K. J., Flanagan, L. B., Katul, G. G., Munger, J. W., Ricciuto, D. M., Stoy, P. C., Suyker, A. E., et al.: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agricultural and Forest Meteorology, 136, 1–18, 2006.

Schröter, M., Bange, J., and Raasch, S.: Simulated airborne flux measurements in a LES generated convective boundary layer, Boundary-layer meteorology, 95, 437–456, 2000.

345    Stevens, B., Lenschow, D. H., Vali, G., Gerber, H., Bandy, A., Blomquist, B., Brenguier, J.-L., Bretherton, C., Burnet, F., Campos, T., et al.: Dynamics and chemistry of marine stratocumulus—DYCOMS-II, Bulletin of the American Meteorological Society, 84, 579–594, 2003.

Sühring, M. and Raasch, S.: Heterogeneity-induced heat-flux patterns in the convective boundary layer: can they be detected from observations and is there a blending height?—a large-eddy simulation study for the LITFASS-2003 experiment, Boundary-layer meteorology, 148, 309–331, 2013.

350    Sühring, M., Metzger, S., Xu, K., Durden, D., and Desai, A.: Trade-Offs in Flux Disaggregation: A Large-Eddy Simulation Study, Boundary-layer meteorology, 170, 69–93, 2019.

**Table 1.** Studies examining the sampling error problem in aircraft flux measurements with the aid of large eddy simulations (LES).

| Study | Model domain | Context | Applicable result or finding |
|---|---|---|---|
| Schröter et al. (2000) | $401 \times 401 \times 42$, $\Delta x = 50$ m, 2 km extent | Dry convective boundary layer | "flight duration of approximately 2000 s [at 100 m sec$^{-1}$] was necessary to obtain accuracy of 10% for the statistical error of the sensible heat flux in the lower half of the CBL..." |
| Sühring and Raasch (2013) | $1400 \times 1400 \times 100$, $\Delta x = 40$ m, 56 km extent | Heterogeneous land surface | included an examination of flux estimate errors from virtual flight legs through the LES domain |
| Sühring et al. (2019) | $1714 \times 2286 \times 500$, $\Delta x = 7$ m, 12 km $\times$ 16 km extent | Flux disaggregation over heterogeneous land surface | virtual flight segments of 200 m to 12.5 km at 49 m height |
| **This study** | $4096 \times 4096 \times 1200$, $\Delta x = 1.25$ m, 5.12 km extent | Cloud-topped marine boundary layer over homogeneous ocean surface | "instantaneous" (single time-step) virtual flight legs from 7.2 km to 102.5 km at heights of 10, 40, 100, and 400 m |

**Table 2.** Coefficients for the integral length scale $I_f \approx Az^b$, valid up to at least 100 m altitude in this simulation, depending on the variable (see Fig. 11), where both $I_L$ and $z$ are in meters. Also given is the approximation $I_f \approx \bar{A}$ for selected variables.

| | | $A$ | $b$ | $\bar{A}$ |
|---|---|---|---|---|
| $q$ | Crosswind: | 20.6 | 0.51 | |
| | Parallel: | 35.2 | 0.41 | |
| $T$ | Crosswind: | 20.5 | 0.56 | |
| | Parallel: | 33.1 | 0.36 | |
| $U_{\parallel}$ | Crosswind: | 309 | −0.06 | 278 |
| | Parallel: | 146 | 0.09 | 206 |
| $U_{\perp}$ | Crosswind: | 161 | 0.11 | 233 |
| | Parallel: | 387 | −0.03 | 371 |
| $w$ | Crosswind: | 0.63 | 1.03 | |
| | Parallel: | 2.95 | 0.82 | |
| $H$ | Crosswind: | 0.53 | 0.97 | |
| | Parallel: | 3.01 | 0.64 | |
| $E$ | Crosswind: | 0.50 | 1.00 | |
| | Parallel: | 2.85 | 0.69 | |
| $\tau_{\parallel}$ | Crosswind: | 0.91 | 0.81 | |
| | Parallel: | 2.88 | 0.70 | |
| $\tau_{\perp}$ | Crosswind: | 0.86 | 0.85 | |
| | Parallel: | 2.34 | 0.79 | |
| $TKE$ | Crosswind: | 77.3 | 0.06 | 96 |
| | Parallel: | 95.4 | 0.05 | 115 |

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU
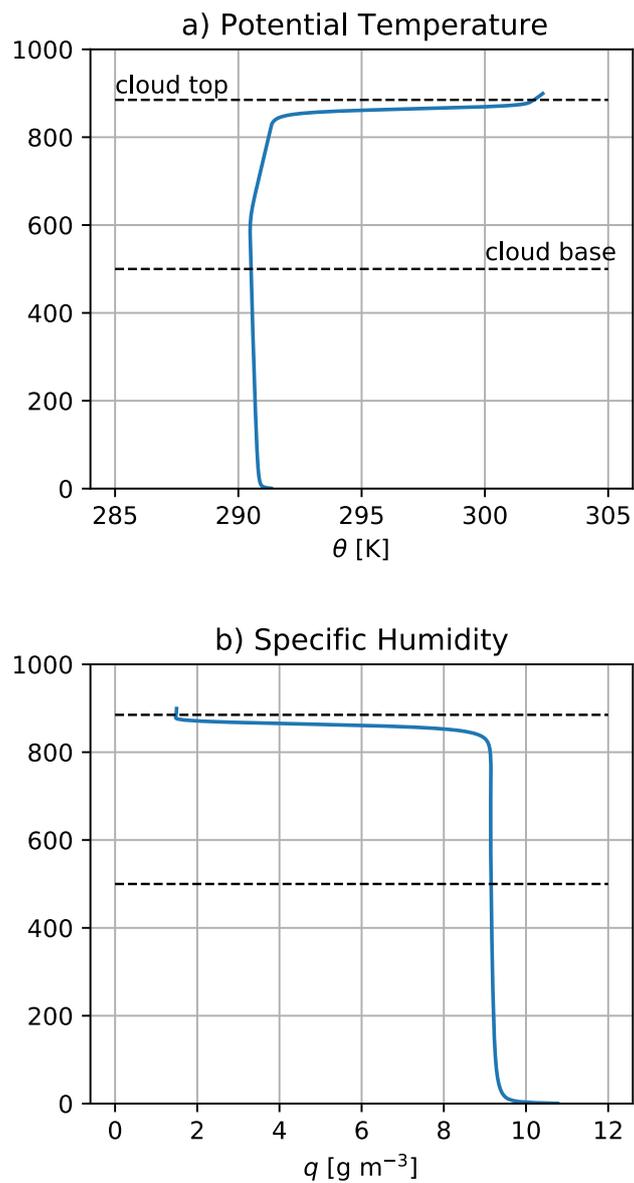


**Figure 1.** Domain-averaged profiles of a) potential temperature $\theta$ and b) specific humidity $q$ at the end of the LES run. Dashed lines indicate the heights of the minimum and maximum altitudes at which non-zero cloud water occurred at any point in the domain.
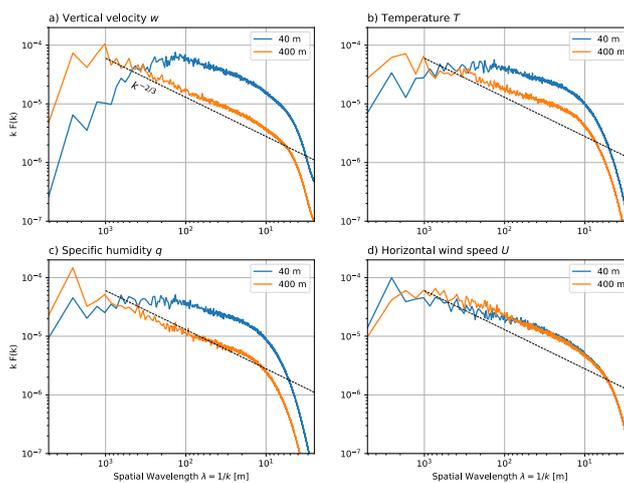
**Figure 2.** Radially averaged power spectra for selected scalar variables at two representative heights 40 m = $0.05z_i$ and 400 m = $0.48z_i$: a) vertical velocity, b) temperature, c) specific humidity, d) horizontal wind speed. For reference, black dashed lines depict the power-law exponent of $-2/3$ expected for a Kolmogorov spectrum in the inertial subrange.
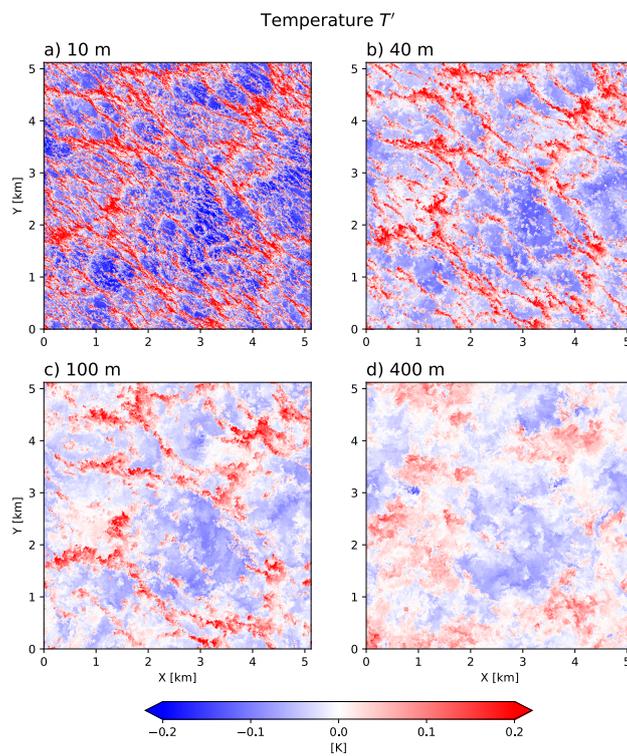
**Figure 3.** Horizontal cross-sections of $T'$ at a) 10 m, b) 40 m, c) 100 m, and d) 400 m heights.
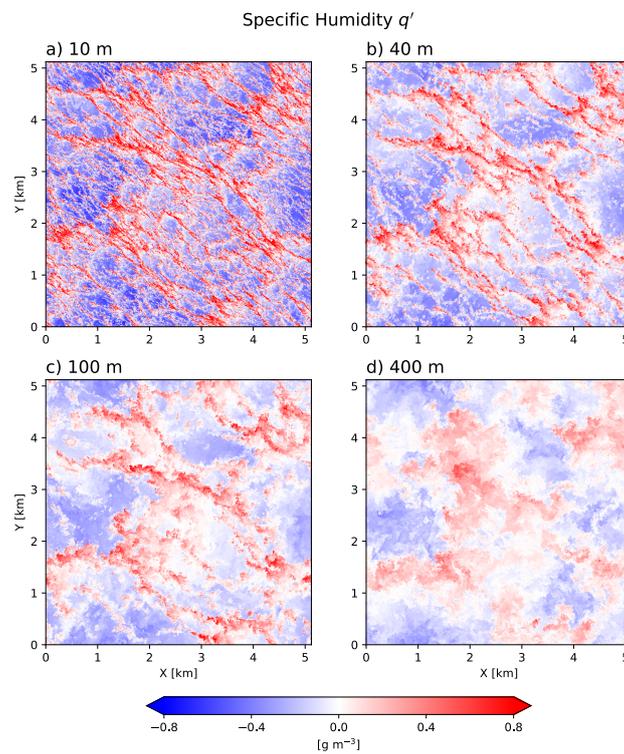
**Figure 4.** Horizontal cross-sections of $q'$ at a) 10 m, b) 40 m, c) 100 m, and d) 400 m heights.
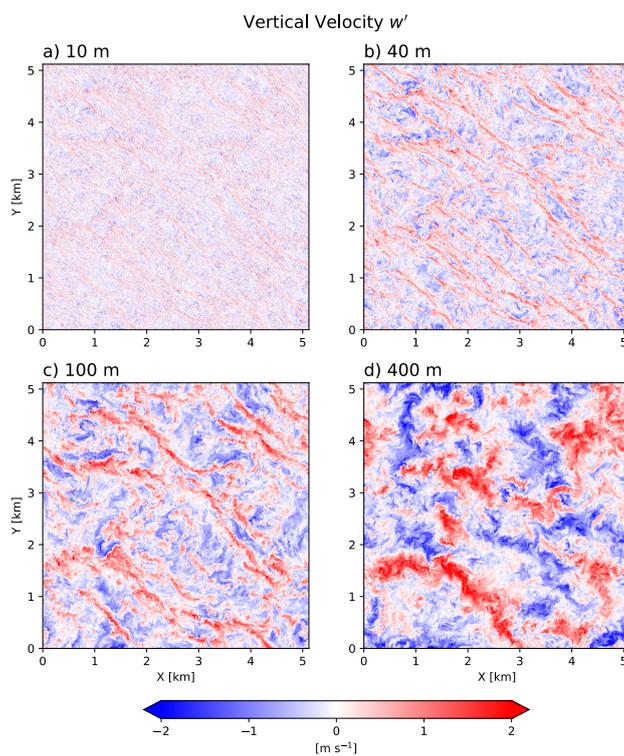
**Figure 5.** Horizontal cross-sections of $w'$ at a) 10 m, b) 40 m, c) 100 m, and d) 400 m heights.

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU
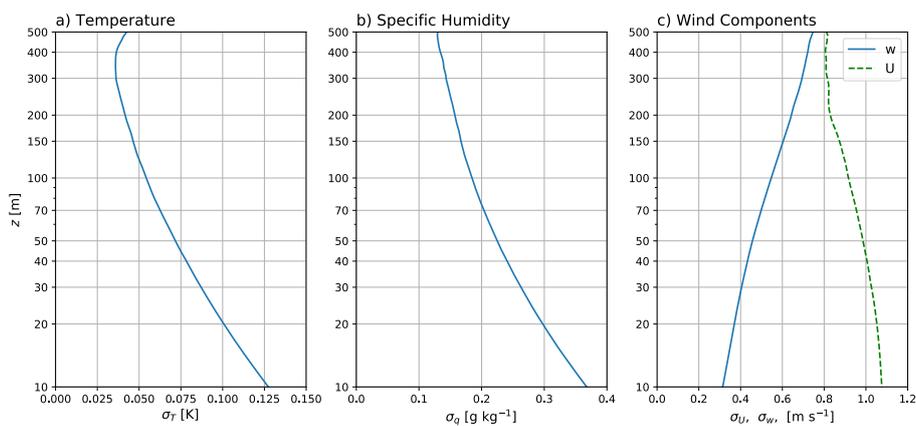
**Figure 6.** Sub-cloud profiles of the turbulent standard deviation of a) temperature, b) specific humidity, and c) vertical and horizontal wind speeds.
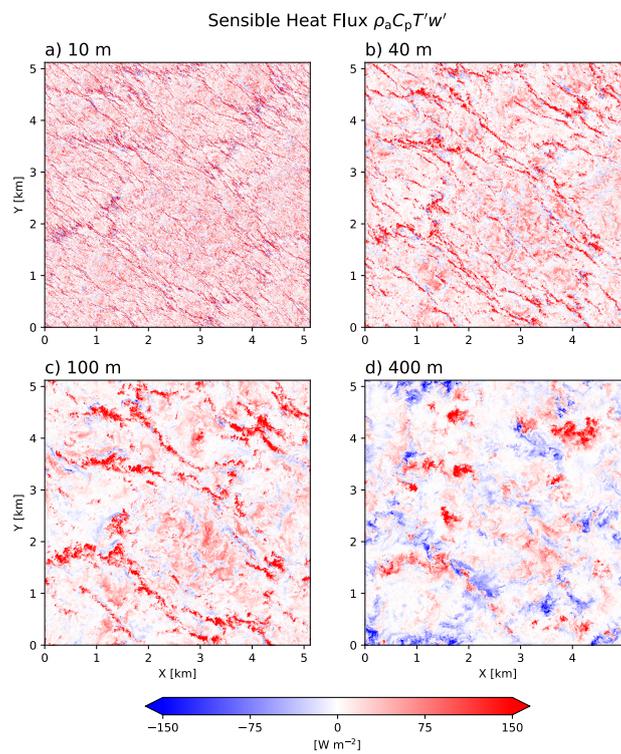
**Figure 7.** Horizontal cross-sections of $H' = \rho C_p T' w'$ at a) 10 m, b) 40 m, c) 100 m, and d) 400 m heights.
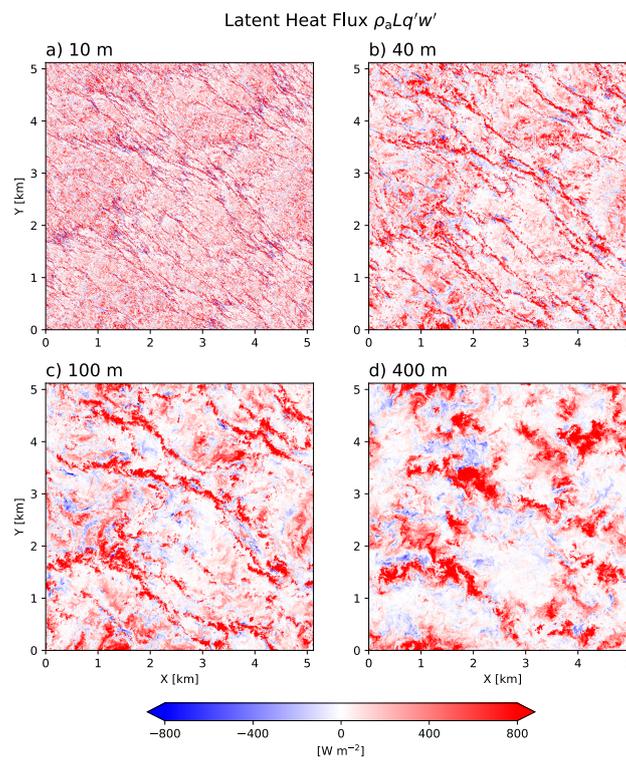
**Figure 8.** Horizontal cross-sections of $E' = \rho L q' w'$ at a) 10 m, b) 40 m, c) 100 m, and d) 400 m heights.
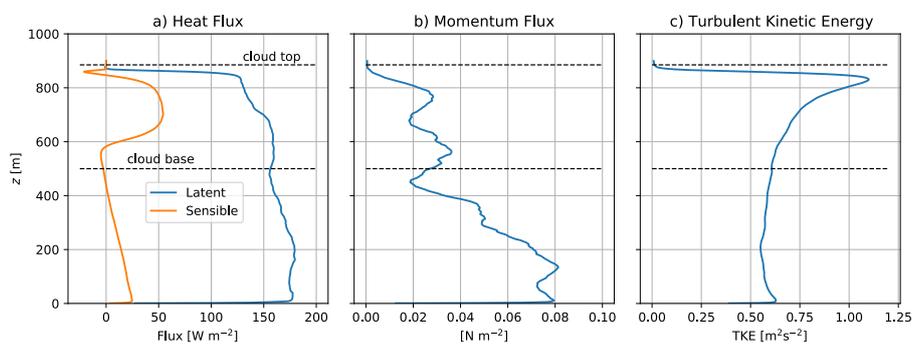
**Figure 9.** Domain-averaged turbulent quantities computed from grid-resolved $u'$, $v'$, $w'$, $T'$, and $q'$ in the LES simulation.
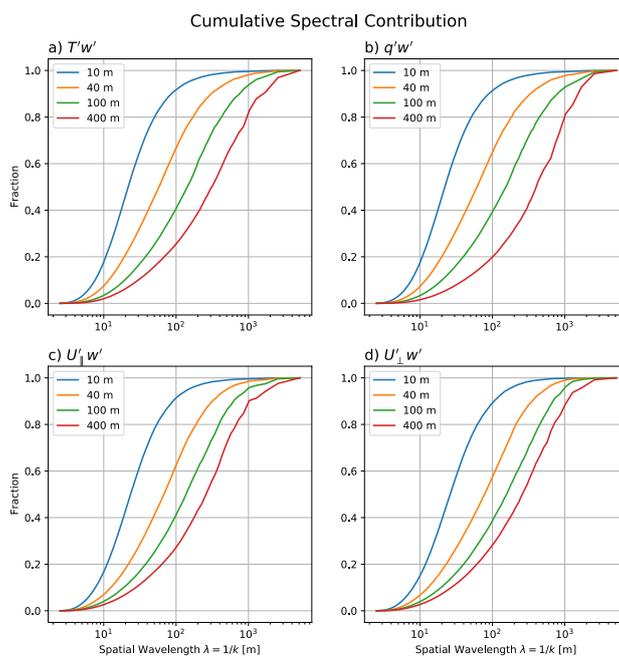
**Figure 10.** Ogive plots of the spectral contributions to the total eddy correlation fluxes for a) sensible heat, b) latent heat, c) horizontal momentum in the along-wind direction, d) horizontal momentum in the crosswind direction.

Atmospheric
Measurement
Techniques
Discussions

Open Access
EGU
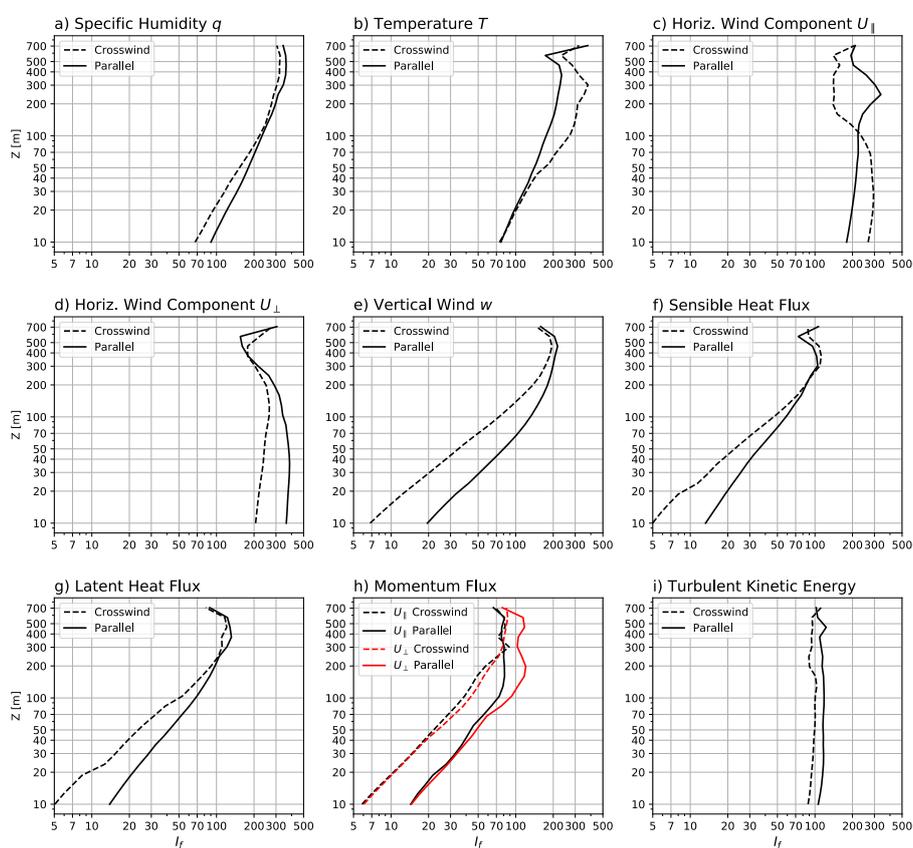
**Figure 11.** Integral length scales $I_f$ for the indicated quantities, measured along ("parallel") and perpendicular to ("crosswind") the mean wind direction.
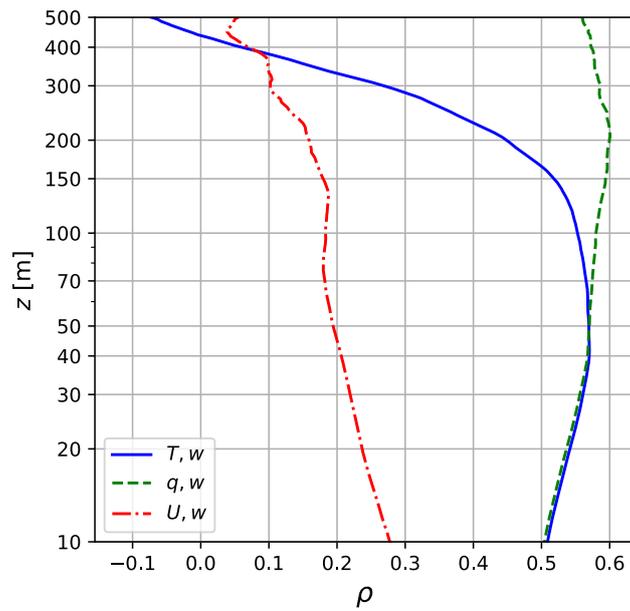
**Figure 12.** Sub-cloud profiles of correlations between vertical velocity $w$ and temperature $T$, specific humidity $q$, and horizontal wind speed $U$.
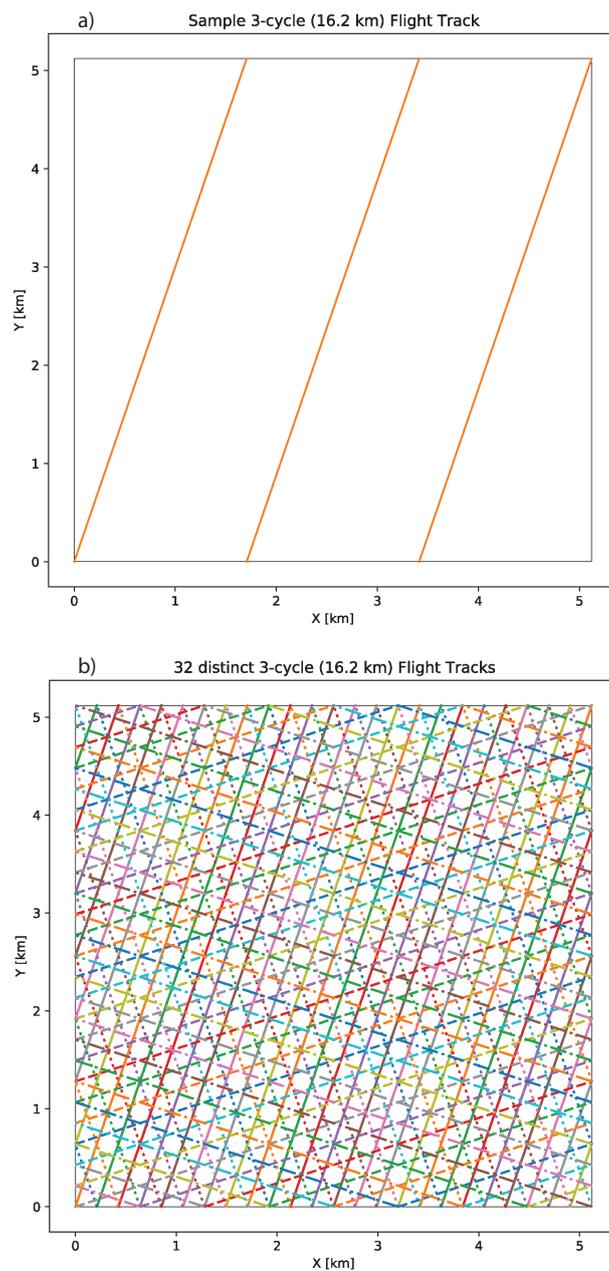
**Figure 13.** Example of flight tracks constructed to pass through the LES domain an integral number of cycles before exactly repeating. a) A single 3-cycle (16.2 km) track. b) All possible realizations of 3-cycle tracks such that a minimum parallel spacing of 200 m between track legs is maintained.

Atmospheric
Measurement
Techniques

Open Access

EGU

Discussions
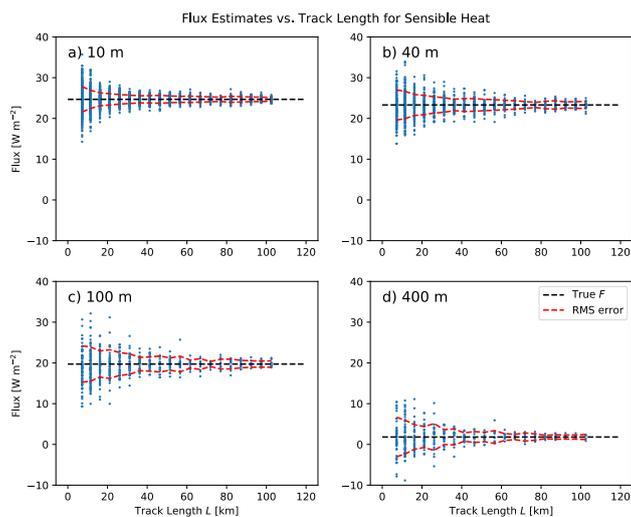
**Figure 14.** Flux estimates vs. track lengths for sensible heat $H$. Red dashed curves depict the root-mean-squared deviation of estimates from the true domain-averaged flux (horizontal dashed line).
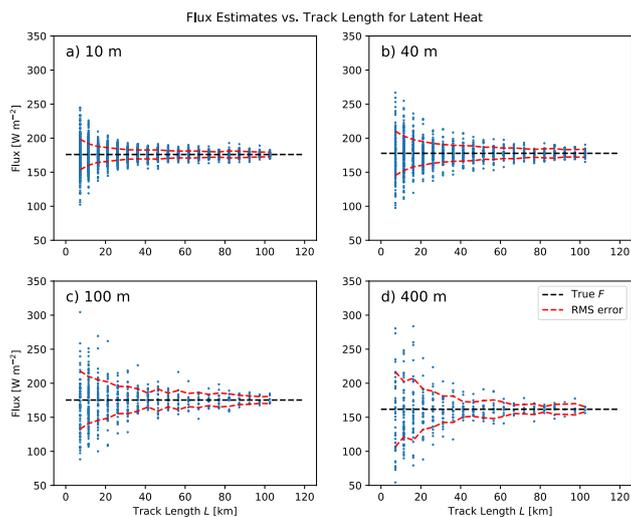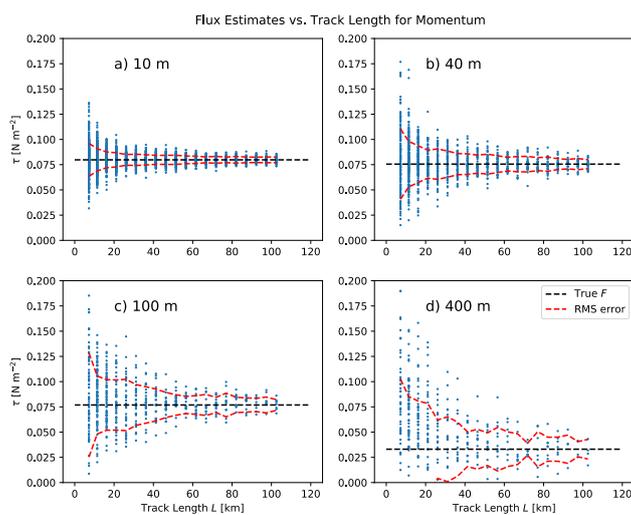
**Figure 15.** Flux estimates vs. track lengths for latent heat $E$. Red dashed curves depict the root-mean-squared deviation of estimates from the true domain-averaged flux (horizontal dashed line).

**Figure 16.** Flux estimates vs. track lengths for horizontal momentum $\tau$ (Reynolds stress). Red dashed curves depict the root-mean-squared deviation of estimates from the true domain-averaged flux (horizontal dashed line).
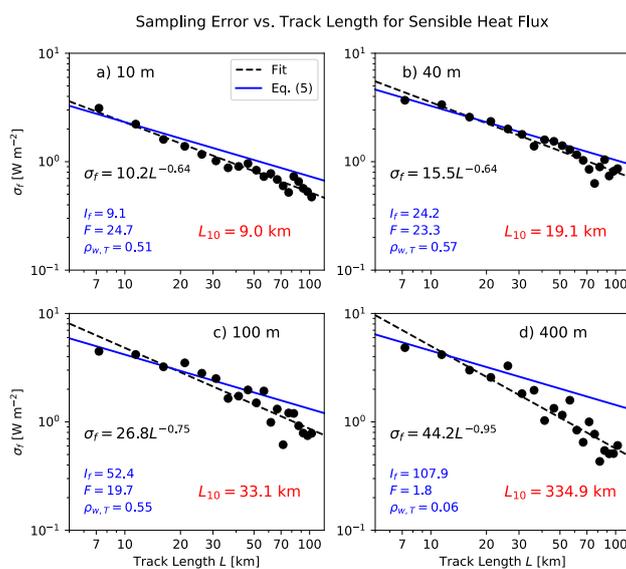
Atmospheric
Measurement
Techniques
Open Access

Discussions

EGU

**Figure 17.** Relationship between ensemble RMS error and track length for sensible heat flux $H$. Black dots correspond to the red curves in Fig.14. An empirical power-law fit is depicted by the dashed black line, with the coefficients of the fit indicated in the expression for $\sigma_f$. $L_{10}$ (in red) indicates the track length required to achieve a 10% relative error. The solid blue line represents (5) using the parameter values given in blue.
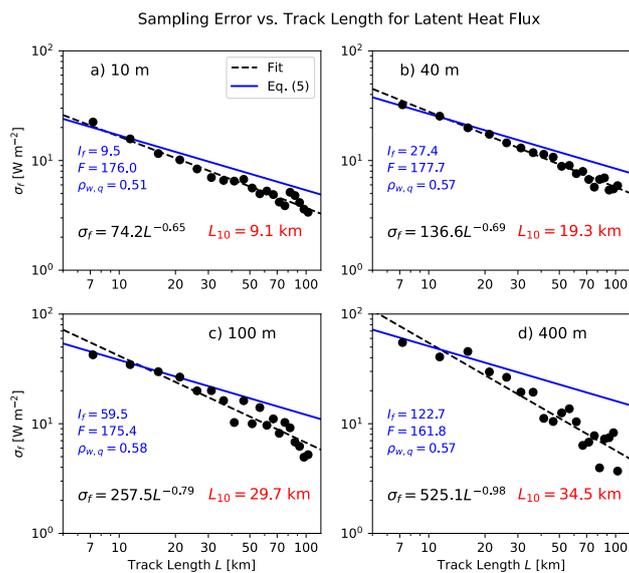
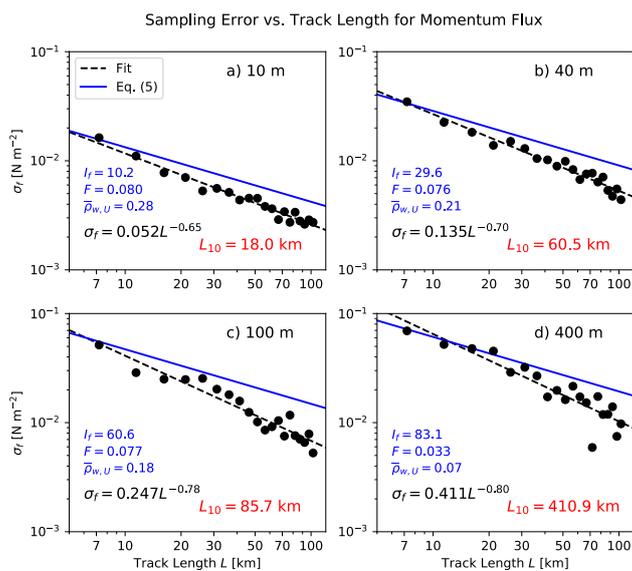**Figure 18.** Same as Fig. 17 but for latent heat flux $E$.

**Figure 19.** Same as Fig. 17 but for momentum flux $\tau$.