JGR Atmospheres

RESEARCH ARTICLE
10.1029/2022JD036654

Special Section:
SOUTHTRAC-GW: An airborne field campaign to explore gravity wave dynamics at the world's strongest hotspot

High-Resolution Aircraft Observations of Turbulence and Waves in the Free Atmosphere and Comparison With Global Model Predictions

Andreas Dörnbrack[1], Peter Bechtold[2], and Ulrich Schumann[1]

[1]DLR Oberpfaffenhofen, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany, [2]ECMWF, Reading, UK, Bologna, Italy, Bonn, Germany

Key Points:

- Small eddy dissipation rates were observed in the free atmosphere along extended research flights during Southern Hemisphere Transport, Dynamics, and Chemistry in austral winter 2019
- Stronger turbulence events are rare and are mostly correlated with enhanced vertical shear of the horizontal wind
- EDR predictions of a 15-member ensemble shows higher correlation with research aircraft observations than with those by commercial aircraft

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Dörnbrack,
andreas.doernbrack@dlr.de

Citation:

Dörnbrack, A., Bechtold, P., & Schumann, U. (2022). High-resolution aircraft observations of turbulence and waves in the free atmosphere and comparison with global model predictions. *Journal of Geophysical Research: Atmospheres*, *127*, e2022JD036654. https://doi.org/10.1029/2022JD036654

Received 17 FEB 2022
Accepted 19 JUL 2022

Author Contributions:

Conceptualization: Andreas Dörnbrack, Ulrich Schumann

Abstract High-resolution flight data obtained from in situ measurements in the free atmosphere aboard the High Altitude and Long Range Research Aircraft (HALO) are used to determine eddy dissipation rates along extended flights during the recent Southern Hemisphere Transport, Dynamics, and Chemistry aircraft campaign (SOUTHTRAC) in the 2019 austral winter. These data are analyzed and correlated with quantities characterizing the ambient airflow and the magnitudes of vertical energy propagation through internal gravity waves. The observed turbulence events are strongly correlated with elevated shear values, and overturning gravity waves do not appear to play a role. A highlight of the analysis is the validation of a recently implemented Clear Air Turbulence (CAT) forecast index in the European Centre for Medium-Range Weather Forecast integrated forecast system. Here we find a slightly better correlation of the CAT prediction with the HALO research aircraft observations compared to those of commercial aircraft. The observed turbulence during SOUTHTRAC was never stronger than moderate, as EDR values remained below 0.3 $m^{2/3}$ s^{-1}. In general, light and light-to-moderate turbulence events were extremely rare, occurring in only about 5% of the flight time, and stronger events in less than 0.2%. These results are also reflected in the local atmospheric conditions, which were dominated by a thermally very stable airflow with low vertical shear and large Richardson numbers.

Plain Language Summary This study analyzes high-resolution data of velocity components in the upper troposphere and lower stratosphere collected with the German research aircraft High Altitude and Long Range Research Aircraft during the Southern Hemisphere Transport, Dynamics, and Chemistry (SOUTHTRAC) campaign in September–November 2019. Flights were conducted predominantly over the southern part of South America, the Drake Passage, and the Antarctic Peninsula. The objective of the analysis was to determine the eddy dissipation rates during the 22 flights. The cubic root of eddy dissipation rates is a common measure used to characterize turbulent regions in the atmosphere. High quality observations with a very accurately calibrated sensor are rare, especially in the remote areas of the SOUTHTRAC campaign. Observed eddy dissipation rates have been correlated with gravity wave activity, but these correlations are very small. A much stronger dependence of the eddy dissipation rates exists on the vertical shear of the horizontal wind. Thus, mechanical generation of turbulence appears to dominate in the observed cases. Overall, the observed turbulence was never stronger than moderate. Turbulence events were extremely rare, occurring in only about 5% of the flight time, and stronger events less than 0.2%. Finally, the observed eddy dissipation rates were compared with weather model forecasts, demonstrating their reliability in predicting turbulent regions.

1. Introduction

Turbulence in the free atmosphere—or Clear Air Turbulence (CAT)—is generated by a variety of processes: shear instabilities, breaking gravity waves, inertial instabilities as often described, for example, Dutton (1967), Goldburg and Pao (1969), Dutton and Panofsky (1970), Vinnichenko and Dutton (1969), Knox (1997). CAT poses a potential hazard to flight operations, for example, Sharman, Trier, et al. (2012), Sharman et al. (2014), Sharman and Lane (2016), S.-H. Kim et al. (2022), but it is also a fascinating topic for basic research because CAT regions are places where energy dissipates and momentum is deposited and where tracers can be irreversibly exchanged between different layers of the free atmosphere, for example, Holton et al. (1995), Gettelman et al. (2011). Improving reliable CAT predictions depends on observations made particularly in the upper troposphere and lower stratosphere. There is often a lack of thermodynamic and wind observations at the horizontal

Formal analysis: Andreas Dörnbrack, Peter Bechtold, Ulrich Schumann
Methodology: Andreas Dörnbrack, Peter Bechtold, Ulrich Schumann
Software: Andreas Dörnbrack, Peter Bechtold, Ulrich Schumann
Validation: Andreas Dörnbrack, Peter Bechtold, Ulrich Schumann
Writing – original draft: Andreas Dörnbrack
Writing – review & editing: Peter Bechtold, Ulrich Schumann

scale of numerical simulations currently used to study individual cases of turbulence generation and to test CAT predictions, for example, Trier et al. (2020) as an illustrative and instructive example of a recent study.

The earliest records of turbulence from research aircraft in the free atmosphere date back to the flights over the United Kingdom (Briggs & Roach, 1963) and the high-altitude CAT (HICAT) program in the United States (Lilly et al., 1974). During the multi-year flight program HICAT, a series of 285 flights totaling 800000 km were conducted, covering a wide range of latitudes, longitudes, seasons, and terrains below the flight paths. These observations yielded mean dissipation rates ϵ (i.e., an ϵ averaged from the contributions of all velocity components) between 6 and $17 \cdot 10^{-4}$ m^2 s^{-3}, where the individual ϵ based on the vertical velocity component was up to about a factor of 5 smaller than that calculated based on the lateral or longitudinal velocity components. At this time, the small ϵ values (and the derived turbulent diffusivity) from the airborne measurements were considered the most striking result of the study, since earlier observations indicated values that were an order of magnitude larger. In addition, the much higher turbulence intensity over high mountains was assumed to be related to the breaking of vertically propagating mountain waves, as demonstrated in some case studies already conducted at that time, while the reason for the lower ϵ values over lowlands and sea was less clear.

Turbulence in breaking mountain waves has been investigated in a number of single case studies, for example, Whiteway et al. (2003), Pavelin et al. (2002). For these events, eddy dissipation rates reached relatively large values. For example, using the formula $K_H = \epsilon/(3\ N^2)$ and the values of K_H and N from the case study of an overturning mountain wave near the tropopause as given by Whiteway et al. (2003), yields $\epsilon \approx 1.2 \cdot 10^{-3}$ m^2 s^{-3}. On the other hand, Schumann et al. (1995), examining about a dozen observations near the tropopause over the North Atlantic, estimated that ϵ is in the range between 1 and $20 \cdot 10^{-8}$ m^2 s^{-3}, that is, at least four orders of magnitude smaller than the mentioned previous estimates. According to the authors, these really small values represent atmospheric conditions heavily affected by the preponderance of strong thermal stratification (N usually greater than 0.015s^{-1}) during the research flights. Other measurements found similarly low dissipation values (Cho et al., 2003; Podglajen et al., 2017).

Besides airborne measurements, there exist different observational methods of determining ϵ in the free atmosphere. A traditional data source with a long history is radar observations, see Hocking (1985) for an early overview. Recently, to cite just one example, Li et al. (2016) estimated eddy dissipation rates based on spectral width data collected by the Middle Atmosphere Alomar Radar SYstem (MAARSY). The altitude-resolved ϵ data reveal values larger than 10^{-5} m^2 s^{-3}, with mean values in the order of 10^{-4} m^2 s^{-3} confirming earlier reports from VHF radars, for example, by Pepler et al. (1998). Usually, radar profiles are validated with radiosonde data by applying a so-called Thorpe analysis (Thorpe, 1977), see eg. Kantha and Hocking (2011), Luce et al. (2014); Li et al. (2016), Kohma et al. (2019).

Generally, radiosonde data with high vertical resolution are another valuable source for estimating turbulence in the free atmosphere, for example, Clayson and Kantha (2008), Ko et al. (2019), Geller et al. (2021). In the recent work by Ko et al. (2019), eddy dissipation rates throughout the troposphere and lower stratosphere are estimated using 4 years of US high vertical-resolution radiosonde data from 68 operational stations. The mean ϵ values in the turbulent layers of the troposphere and stratosphere are 1.84 and $1.37 \cdot 10^{-4}$ m^2 s^{-3}, respectively. Large ϵ values are found mainly over mountainous regions in the troposphere, but this pattern is not observed in the stratosphere. A slightly different result was obtained by Zhang et al. (2019) who calculated mean ϵ-values of 2.7 and $2.9 \cdot 10^{-4}$ m^2 s^{-3} for the troposphere and stratosphere, respectively. Their analysis uses vertical ascents with the same resolution but from 90 stations distributed globally.

In the past, mainly reports of atmospheric turbulence intensity from commercial aircraft have been employed to verify turbulence forecasts of numerical weather prediction (NWP) models. Traditionally, pilot reports were taken to associate the felt shaking of the aircraft to turbulence intensity which was then put into categories of CAT strength as smooth, light, moderate, severe, or extreme. Sharman et al. (2014) presented results from an automated algorithm that has been implemented on a huge fleet of commercial airliners. The turbulence intensity, represented as the cube root of the eddy dissipation rate and denoted as EDR $= \epsilon^{1/3}$, is based on estimated vertical wind velocity or aircraft vertical acceleration. These methods and data are widely used for validating CAT predictions of NWP centers, for example, J.-H. Kim et al. (2018), Storer et al. (2019), Goecke and Machulskaya (2021). Recently, the European Centre for Medium-Range Weather Forecasts (ECMWF) developed predictions of CAT based on the projections of the model turbulent tendencies onto the EDR Bechtold, Bramberger, Dörnbrack,

Leutbecher, and Isaksen (2021), Bechtold, Bramberger, Dörnbrack, Isaksen, and Leutbecher (2021). We utilize this new implementation to validate the global CAT predictions of the ECMWF's Integrated Forecasting System (IFS) with measurements from a research aircraft, the German High Altitude and Long Range Research Aircraft (HALO).

High-resolution, flight-level observations of wind and pressure in the free atmosphere aboard HALO are used to characterize the airflow at flight level, the turbulence intensity, expressed by the eddy dissipation rate ϵ, the wave energy fluxes EF_j ($j = x, y, z$ as the zonal, meridional, and vertical coordinates), and the momentum fluxes MF_x and MF_y from 22 extended research flights during the recent Southern Hemisphere Transport, Dynamics, and Chemistry (SOUTHTRAC) field campaign, for example, Rapp et al. (2021), Dörnbrack et al. (2020), Wildmann et al. (2021), Jesswein et al. (2021). More than 120 flight hours of straight and constant flight level legs are available for SOUTHTRAC. The research flights are distributed over the southern Andes, the Drake Passage and the Antarctic Peninsula and—during the transfer flights—over the Atlantic Ocean. The majority of the analyzed flights took place at altitudes of about 12 km. With the exception of the mountain crossings, the other research flights were not specifically designed to sample regions of predicted CAT.

The SOUTHTRAC research flights are sorted by whether they took place over land or over water, whether significant topography was below the flight path, and whether they occurred in the troposphere or stratosphere. Thus, the first objective of this study is to present probability density functions of ϵ for the different segments of the research flights and to ask how large the eddy dissipation rate can be and what are its typical values in the free atmosphere. A number of studies dealing with radiosonde and airborne turbulence observations near the tropopause mention the coexistence with internal gravity waves, for example, Turner (1955), Bellenger et al. (2017), their possible role as triggers of instabilities in layers of the atmosphere (Panofsky et al., 1968), or even the potential of breaking gravity waves in generating turbulent patches, see for example, Dörnbrack (1998), Bramberger et al. (2020), Wilms et al. (2020). Therefore, the second objective is to provide wave energy and momentum fluxes from the flight level observations to investigate if increased gravity wave fluxes also imply larger ϵ values.

The background flow conditions along the flight paths are investigated using ECMWF's operational analyses and short-term forecasts to identify dynamically or convectively unstable regions. As mentioned above, the final objective of our study is to compare the observed eddy dissipation rates with IFS ensemble predictions of a specially designed CAT index recently implemented in the operational suite of the ECMWF's IFS, see Bechtold, Bramberger, Dörnbrack, Leutbecher, and Isaksen (2021), Bechtold, Bramberger, Dörnbrack, Isaksen, and Leutbecher (2021). Here, as a completely new data set for the validation, we use high-resolution in situ observations of turbulence collected by the Basic HALO Measurement and Sensor System (BAHAMAS) aboard HALO. BAHAMAS consists of a nose boom probe with a 5-hole sensor and provides in situ measurements of horizontal and vertical wind components as well as temperatures, pressures, and water vapor mixing ratios at flight altitude with high temporal resolution, that is, up to 100 Hz (Giez et al., 2017, 2019, 2021). This data set from the SOUTHTRAC campaign provides reliable estimates of ϵ for the upper troposphere/lower stratosphere because ϵ is calculated from the spectra of the complete 10 Hz wind data for successive short portions of the research flights. For selected flight legs these estimates are compared with 100 Hz.

In this study, we pose the following research questions:

Q1: Along the SOUTHTRAC research flights, what is the probability of encountering regions with EDR values above a certain threshold for moderate or moderate-to-severe turbulence?

Q2: Is it possible to infer the causes of turbulence from the calculation of wave energy and momentum fluxes determined from BAHAMAS data? Is there any evidence of local instabilities that may be responsible for the turbulence?

Q3: How well do IFS predictions of EDR, calculated using a specially developed CAT index, compare with accurate EDR estimates based on high-resolution in situ measurements aboard a research aircraft?

The remainder of the paper is organized as follows. Section 2 describes the aircraft data and the methodology for analyzing them. Section 3 introduces the calculation of the CAT index implemented in the IFS. The corresponding Appendix A briefly presents a first verification using data from commercial airliners. Sections 4 and 5 present the results of this study, with Section 5 focusing on the comparison with the CAT predictions of the IFS. In the final Section 6, we discuss our findings and conclude the study by answering the above research questions.

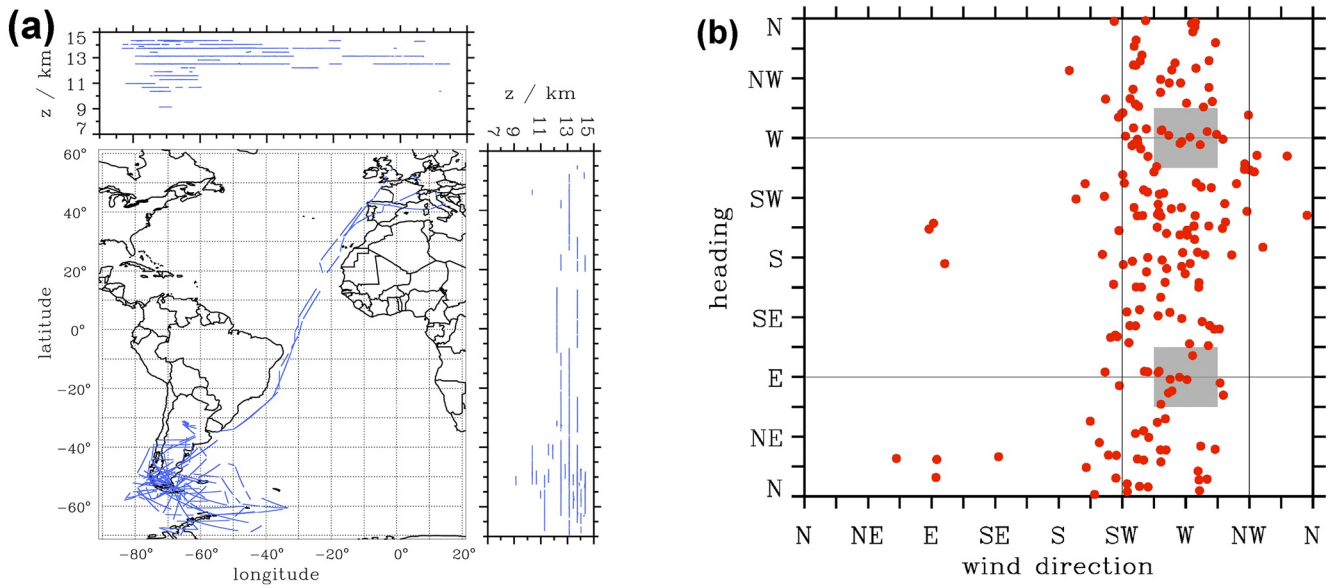

Figure 1. (a) Geographical locations of the 10 Hz straight and level flight legs together with HALO's altitude as function of latitude and longitude. (b) HALO's heading versus horizontal wind direction for all 188 straight and constant altitude legs of the SOUTHTRAC research flights as shown in Figure S1 in Supporting Information S1. The values are averages over the individual legs. The gray-shaded squares mark the legs where the heading is either along or against the westerly cross-mountain wind direction.

2. Research Aircraft Data and Analysis Methodology

The entire SOUTHTRAC data set of HALO data—available on the HALO database—includes 28 flights, hereafter referred to as STIJ, where IJ stands for research flight number. The first three SOUTHTRAC flights ST01 to ST03 were instrument tests over Germany, and the last two (ST27, ST28) were short transfers at low altitude from Rio Grande, Argentina, to Punta Arenas, Chile. Research flight ST05 was a short overflight of a convective system over the mid-Atlantic to test a new dropsonde system. Here, a total of 22 research flights, ST04 and ST06 through ST26, are evaluated and presented and will be characterized in the following subsection. Their geographical distribution is shown in Figure 1a. All SOUTHTRAC observational data employed in this study are from BAHAMAS at temporal resolutions of 1, 10, or 100 Hz. Giez et al. (2021) presented a detailed and complete description of the different calibration steps of the BAHAMAS sensor for the German research aircraft HALO. There, they summarize: "… we can state that HALO is capable of acquiring high-quality 3d wind data up to 100 Hz during straight and leveled flight". Comparison with other turbulence measurements exist for the HALO-FAAM formation flight on 13 July 2017 during the EMeRGe European intensive operational period (Andrés Hernández et al., 2022); details of this comparison are documented in Schumann (2020). As discussed below, the low frequency velocity spectra measured by HALO agree well with other observations (Schumann, 2019).

In addition, 6-hourly operational IFS analyses and 1-hourly short-term IFS forecasts, that fill the gaps between the analysis times, are used to generate two-dimensional curtains of meteorological variables that are interpolated to the latitude-longitude positions and times of the aircraft along the flight tracks during the measurement flights.

2.1. HALO Flights During the SOUTHTRAC Campaign

For all 22 research flights, 188 legs were identified that were straight (nearly constant flight direction) and that were performed at a certain flight level, that is, at nearly straight and at constant flight levels, see Figure 1a. In total, this gives a flight time of about 123 h and a horizontal distance of 103718 km (about 2.5 times the circumference of the Earth, a little more than a quarter of the distance to the Moon) traveled by HALO. The average ratio of time during such straight routes to total flight time of research flights is 68.5%, with a maximum of 82.4% (ST22) and a minimum of 35.4% (ST16).

Figure S1 in Supporting Information S1 displays four variables that characterize these 188 legs based on the 1Ḣz BAHAMAS data for the considered research flights. In addition, Table S1 in Supporting Information S1 list a more complete set of variables describing the physical properties of the airflow along these legs. Figure S1a in Supporting Information S1 shows that about 80% of all legs were in the altitude range of 12–14 km, corresponding to the dominance of flight levels greater than FL390 or higher than 180 hPa. Consequently, the potential temperatures Θ at these flight tracks were predominantly (83%) in the range of 340–380 K, only 10% of all flights were in the so-called overworld ($\Theta > 380$ K, see Holton et al. (1995)). Absolute temperatures at flight altitude were rather low, mostly between 210 and 220 K (75%), and it is interesting to note that the minimum leg-averaged temperature fell once to 198 K = −75°C (ST07, leg 5, Table S1 in Supporting Information S1), causing icing on the nose boom and laser window. In general, the relative humidity over water was very low, and the averaged values of relative humidity in the legs never exceeded 60%, see Table S1 in Supporting Information S1. There are few "wet legs" where local values of relative humidity exceeded 80% (leg 7 of ST19, leg 1 of ST22, and leg 1 of ST25). Figure S1b in Supporting Information S1 shows that most legs were less than 600 km long (about 68%), 24% were between 600 and 1200 km, and 8% were longer than 1200 km.

Horizontal winds ranged mostly between 10 and 50 m s^{-1} (84%), with stronger winds (30 … 50 m s^{-1}) dominating the distribution (47%), see Figure S1c in Supporting Information S1. Wind direction was dominated by the 225°(SW) to 315°(NW) sector; about 80% of all straight-line flights at nearly constant altitude were conducted under these conditions, see Figure S1d in Supporting Information S1. Figure S1b in Supporting Information S1 shows that only a few flights were made along and against the westerly winds dominating the stratospheric airflow. The chosen flight strategy during SOUTHTRAC differed from the predominantly mountain-crossing research flights oriented in the direction of the mean wind at flight level during the T-REX (Grubišić et al., 2008) and DEEPWAVE (Fritts et al., 2016) campaigns. Here, due to the particular flight strategy imposed by the remote-sensing instruments Gimbaled Limb Observer for Radiance Imaging of the Atmosphere (GLORIA, see e.g., Friedl-Vallon et al. (2014), Riese et al. (2014)) and ALIMA (Airborne Lidar for Middle Atmosphere research, see Rapp et al. (2021)) and the rather limited number of flight hours, only a few of the research flights were aligned with the stratospheric winds at flight levels over the southern Andes.

2.2. Gravity Wave Energy and Momentum Fluxes

Of the 188 straight-level legs, 177 legs were selected that were longer than about 100 km, they are listed in Table S2 in Supporting Information S1. Since these 177 legs were mostly longer than 500 km, they were divided into equal sections by simply dividing the total leg length by 2, 3, …. The resulting 357 segments were selected subjectively in such a way that the length of each section should be in the range of 150–350 km. The resulting average length of all 357 segments is 287 km.

We restrict ourselves here to the presentation of the gravity wave energy fluxes $\mathbf{EF} = \overline{\mathbf{u}' p_c'}$, where $\mathbf{u} = (u, v, w)$ are the zonal, meridional and vertical components of the wind vector and p_c is the hydrostatically corrected static pressure, and the vertical fluxes of horizontal momentum $\mathbf{MF} = (\overline{\rho}\,\overline{u'w'}, \overline{\rho}\,\overline{v'w'})$ in terms of probability density functions (PDFs) from the 357 segments. Here, the overbar denoted averages over the individual segments. From the flight level in-situ observations, **EF** and **MF** are calculated as follows: For each of the segments, all observed quantities are de-trended (linearly) and de-meaned to compute perturbation variables $\Psi' = (u', v', w', p_c')$. Before de-trending and de-meaning, the static pressure p has to be corrected to a constant reference altitude z_{ref} assuming hydrostatic balance $p_c(x) = p(x) + \overline{\rho}\, g(z - z_{ref})$, where z is the aircraft geometric altitude determined from the global positioning system (GPS), g the gravitational acceleration, and $\overline{\rho}$ the leg-averaged density, for example, Smith et al. (2008, 2016). From these variables, the energy and momentum fluxes **EF** and **MF** of the gravity waves are calculated by averaging the point values over the respective lengths of the 357 segments. All the data for this analysis are based on the 10 Hz data as provided by BAHAMAS via the HALO data base. A graphical presentation of the energy fluxes **EF** and the negative scalar product of the vectors of the horizontal wind **U** and the momentum flux **MF** for all 177 legs are shown by red dots in Figure S2 in Supporting Information S1 and listed in Table S2 in Supporting Information S1. In addition, the corresponding values for the 357 segments in Figure S2 are symbolized by black dots.

2.3. Determination of the Eddy Dissipation Rate

Eddy dissipation rates ϵ_i (i stands for the foreward, sideward, and upward velocity components u_b, v_b, w, respectively) are calculated along all SOUTHTRAC flight tracks. There are two data sets employed to determine ϵ_i: one based on the 10 Hz BAHAMAS data available for the entire flight set (their geographical locations at nearly constant altitude and flight directions are sketched in Figure 1a), and another, shorter set for selected 20 flight hours based on the 100 Hz turbulence analysis. To determine ϵ_i, we assume that the airflow is turbulent and has an inertial subrange. Therefore, the spectral energy density S_i for the respective component of the wind velocity vector in aircraft coordinates $u_i = (u_b, v_b, w)$ is

$$S_i = \alpha_i \, \epsilon_i^{2/3} \, k^{5/3}, \tag{1}$$

where k is the wavenumber and $\alpha_i = (0.53, 0.707, 0.707)$ are the Kolmogorov constants for u_b, v_b, and w, respectively (Oncley et al., 1996; Piper & Lundquist, 2004; Strauss et al., 2015).

Each flight is divided into 10 s segments, each 5 s apart (i.e., there are 1000 or 10000 individual data points), for which the velocity components u_b, v_b, and w are determined. Altogether, the 10 Hz data set comprises 88340 individual periods for which the variance spectra are calculated using a Tukey filter to mitigate the effects of jumps due to the aperiodicity of the data (Sharman et al., 2014, Section 2c).

The resulting spectra are fitted between 0.4 and 4 Hz with the assumed −5/3-Kolmogorov spectrum, and the respective ϵ_i are calculated from these fits (Smalikho, 1997). The limits of 0.4 and 4 Hz are deliberately selected to avoid, on the one hand, peaks caused by the oscillations of the nose boom and, on the other hand, the noise of the data, that is, the true airspeed for u_b, side slip angle β for v_b and the angle of attack α for w, as far as possible. In this way, the spectrum represents the average over 10 s or about 2.4 km flight distance and resolves motion scales down to about 4 Hz or 60 m horizontal distance. In addition, mean slopes of the log-log w-spectrums are calculated in the same frequency range. The method was tested by comparing independently computed EDR results from Bramberger et al. (2020) and Rodriguez Imazio et al. (2022).

To compare the observed turbulence intensities, as represented by ϵ_i, along the research flight tracks with the IFS predictions, the quantity

$$\mathrm{EDR}_i = \epsilon_i^{1/3} = \left(\frac{S_i \, k^{5/3}}{\alpha_i} \right)^{\frac{1}{2}}. \tag{2}$$

is calculated. EDR is a common measure of turbulence intensity in aviation, and the next section will explain how it is calculated from the parameterized quantities of the IFS.

There are two fundamental length scales for stratified turbulence. The Kolmogorov microscale

$$\eta = \left(\frac{\nu^3}{\epsilon} \right)^{1/4} \tag{3}$$

is the inner scale of velocity fluctuations where kinetic energy is viscously dissipated into heat (Kolmogorov, 1941). Here, the kinematic viscosity $\nu = \mu/\rho$ can be calculated from the temperature dependent dynamic viscosity μ using the Sutherland formula (White, 1991, Equation 1–36) and the air density ρ. The Ozmidov scale

$$L_O = \left(\frac{\epsilon}{N^3} \right)^{1/2} \tag{4}$$

is the largest scale where the flow resembles a three-dimensional isotropic turbulence (Ozmidov, 1965). Here, the buoyancy frequency

$$N = \sqrt{\frac{g}{\Theta} \frac{\partial \Theta}{\partial z}}, \tag{5}$$

can be calculated from the potential temperature Θ.

An airflow can be considered to be actively turbulent when the ratio L_O/η is greater than about 10, otherwise the flow is dominated by wave-like motions with little energy cascading to smaller scales, for example, Schumann et al. (1995); Riley and Lindborg (2008). This means, for a fixed ratio $L_O/\eta = 10$, a minimum eddy dissipation rate can be computed according to:

$$\epsilon_{\mathrm{MIN}} = \left(\frac{L_O}{\eta}\right)^{4/3} \nu\, N^2 \approx 21.5\, \nu\, N^2 \tag{6}$$

This simple formula can be used to estimate the lower bound of eddy dissipation rates to be expected under the atmospheric conditions represented by $\nu(T)$ and N. For all the SOUTHTRAC flights, the average ϵ_{MIN} is $4.1{\cdot}10^{-7}$ m^2 s^{-3}, see Appendix A. A similar ϵ_{MIN}-dependence of ν and N^2 was derived by Lübken (1997), making an assumption about the minimum turbulent diffusivity without explicitly referring to L_O/η, see also Lübken (1993, Eq. 3.146).

3. CAT Indices as Computed by the IFS

Recently, ECMWF has developed and tested a set of quantities for predicting CAT in the operational suite of IFS Bechtold, Bramberger, Dörnbrack, Isaksen, and Leutbecher (2021), Bechtold, Bramberger, Dörnbrack, Leutbecher, and Isaksen (2021). An operationally available and regularly archived measure of turbulence intensity is useful for research purposes, whether for developing turbulence parameterizations applicable in the free atmosphere or for interpreting historical cases. In this study, the finally selected and operationally implemented CAT index CAT2 is used. It is based on projections of IFS tendencies of parameterized turbulent fluxes onto EDR. Originally, the positive definite Ellrod index (Ellrod & Knapp, 1992) was considered as another CAT index, but for efficiency reasons it is not part of the current operational IFS.

The CAT index CAT2 combines two different parameters. First, the parameterized drag contribution from the breaking of convectively generated gravity waves over deep convective precipitating clouds. A simple approach used here is to scale ϵ from the non-orographic gravity wave scheme (Orr et al., 2010) with the normalized vertically integrated convective heating between 500 hPa and cloud top. This parameter is denoted by GWD:

$$\mathrm{GWD} = \left[\left|\left(u\frac{\partial u}{\partial t}|_{\mathrm{gwd}} + v\frac{\partial v}{\partial t}|_{\mathrm{gwd}}\right)\right| \hat{T}_{\mathrm{conv}}\right]^{1/3} \tag{7}$$

with

$$\hat{T}_{\mathrm{conv}} = \frac{c_p}{\hat{T}_0}\int_{p=500\mathrm{hPa}}^{\text{cloud top}} \frac{\partial T}{\partial t}|_{\mathrm{conv}}\frac{dp}{g}, \tag{8}$$

where the subscript "gwd" denotes the tendencies (wave drag) from the non-orographic wave scheme and the subscript "conv" the temperature tendency from the convection parameterization, c_p the specific heat at constant pressure, and $\hat{T}_{\mathrm{conv}}$ is the integrated convective heating between 500 hPa and the convective top normalized by $\hat{T}_0 = 1$ W m^{-2}. The integration limits in Equation 8 assure that only heating from penetrative convection contributes to the generation of turbulence. For non-orographic gravity waves a globally uniform launch spectrum is prescribed. Equation 8 therefore assumes that convection modulates only the amplitude of the gravity waves. Due to the used scaling, GWD and EDR have the same unit.

The second contribution to the CAT2 index comes from the dissipation due to turbulent mixing, including the contribution due to orographic wave drag and orographic blocking (Beljaars et al., 2004) plus the convective momentum transport. This parameter is denoted by DISS:

$$\mathrm{DISS} = \left|\left(u\frac{\partial u}{\partial t}|_{\mathrm{diff}} + v\frac{\partial v}{\partial t}|_{\mathrm{diff}}\right)\right|^{1/3} + \left|\left(u\frac{\partial u}{\partial t}|_{\mathrm{conv}} + v\frac{\partial v}{\partial t}|_{\mathrm{conv}}\right)\right|^{1/3} \tag{9}$$

where subscript "diff" denotes the tendencies for horizontal momentum from the vertical diffusion scheme and the subscript "conv" the tendencies due to convective momentum transport. DISS, of course, has the unit of EDR, but strictly speaking it is defined only as the vertical integral of Equation 9. To obtain positively defined values

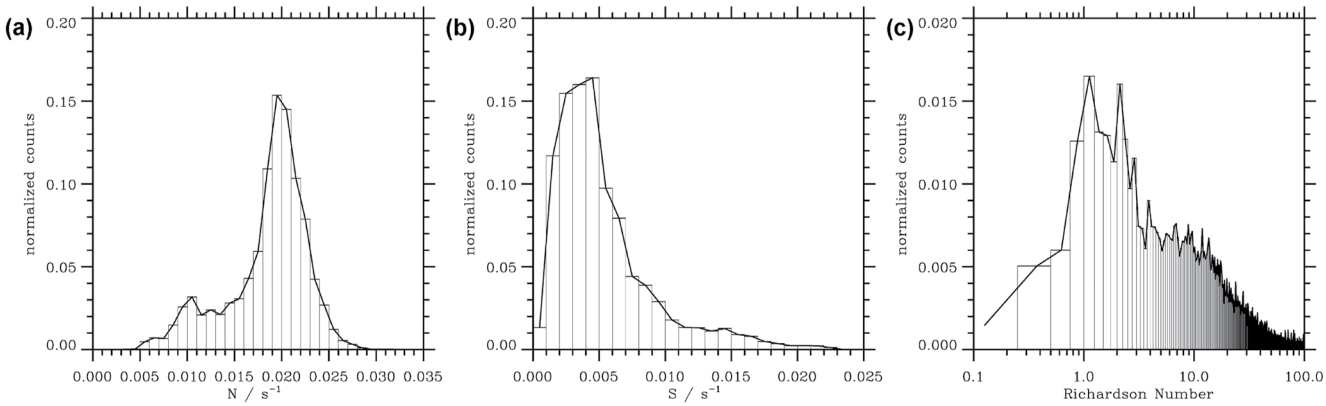

Figure 2. Computed probability density functions of the buoyancy frequency N (a), the vertical shear S (b) and the Richardson number (c) along HALO's straight and constant altitude flight legs during SOUTHTRAC. There are 88340 values every 5 s taken from the IFS data interpolated onto the HALO flight tracks in space and time to compute the probability density functions. The vertical lines represent the data binned according to the width of the bars, and the black curves connect the values in the middle of the bars.

for each level, the absolute value of the kinetic energy tendency is kept. Note that Shutts (2015) used a different formulation of the convective dissipation rate, which is proportional to the product of the square of the vertical updraft velocity and the convective detrainment rate. However, we did not retain this definition because tests showed that it did not conform to a log-normal distribution and degraded correlations with observations when used in Equation 9.

The total dissipation rate projected onto EDR units is estimated as the weighted sum of the turbulent dissipation rate and contributions from convective momentum transport and convective gravity wave drag

$$\mathrm{CAT2} = 0.66\ (\mathrm{DISS} + \mathrm{GWD})\,, \tag{10}$$

where the coefficient of 0.66 is a tuning parameter as derived by Bechtold, Bramberger, Dörnbrack, Leutbecher, and Isaksen (2021). The contribution from the turbulent dissipation rate DISS is by far the dominant term of the above CAT index. Appendix A provides details on the verification of the CAT2 index using EDR reports from commercial aircraft over the continental U.S. and places the results in context with those of other CAT indices tested at ECMWF.

It should be noted that the proposed approach is unique and differs from the Graphical Turbulence Guidance (GTG) system, which uses resolved quantities and their derivatives to determine a range of different CAT indices (Sharman, Doyle, & Shapiro, 2012; Sharman et al., 2014). It also differs from the German Weather Service method, which uses the subgrid-scale turbulent kinetic energy equation to predict CAT (Goecke & Machulskaya, 2021).

As a result of the validation process outlined in Appendix A, daily runs of a 15-member IFS ensemble at $T_{CO}639$ resolution (about 18 km) were performed for all flight days during the months of September, October, and November 2019. The IFS model version that became operational in October 2021 was used for these runs. These re-forecasts were initialized at 00 UTC using the operational ensemble data analysis, and the output was stored at each hour up to 23 UTC. The EDR predictions based on the CAT2 index were projected onto the HALO observations using the same technical parameters regarding data reduction as described in Appendix A.

4. Results

4.1. Shear and Thermal Stability Along the Flight Tracks

Figure 2 depicts the PDFs of the vertical shear

$$\mathrm{S} = \sqrt{\left(\frac{\partial \mathrm{U}}{\partial \mathrm{z}}\right)^2 + \left(\frac{\partial \mathrm{V}}{\partial \mathrm{z}}\right)^2}, \tag{11}$$

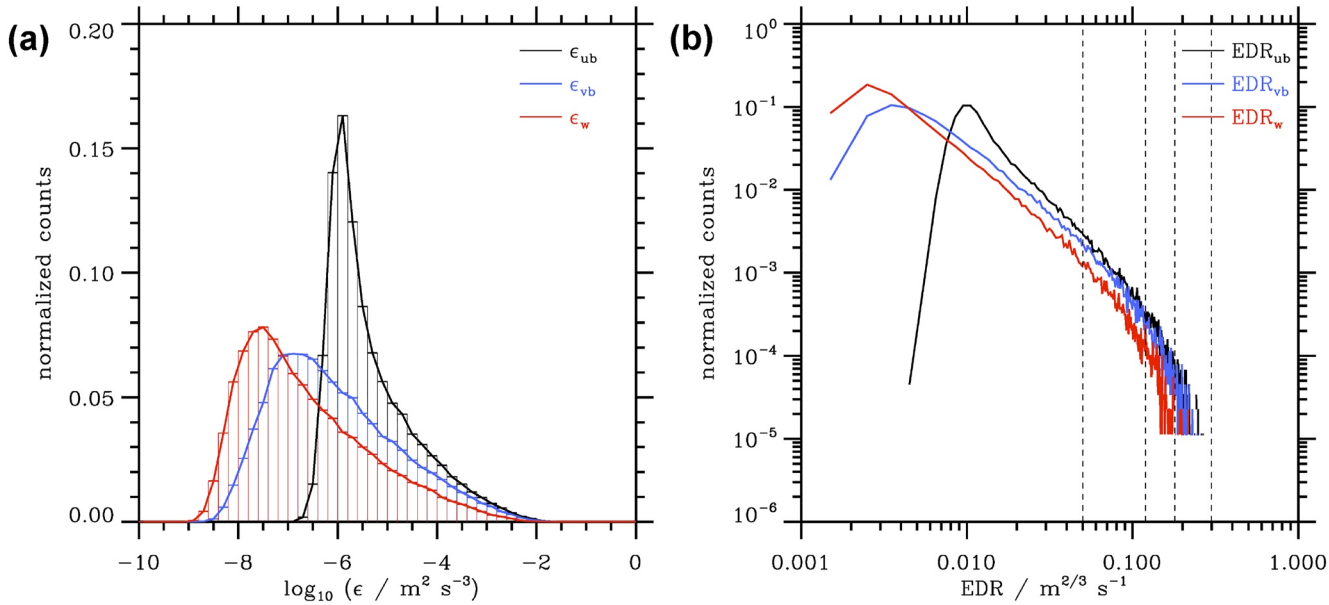

Figure 3. Observed probability density functions of the eddy dissipation rates ϵ_i (a) and the $\mathrm{EDR}_i = \epsilon_i^{1/3}$ (b) derived from the 10 Hz BAHAMAS data of HALO's research flights during SOUTHTRAC. There are 88340 values of ϵ_i and EDR_i evaluated in 10 s intervals every 5 s along all straight and constant altitude legs of the research flights as shown in Figure 1b. This number is used for the normalization of the observed probability density functions. The dashed lines in panel (b) mark the limits of light ($0.05 < \mathrm{EDR}/(\mathrm{m}^{2/3}\,\mathrm{s}^{-1}) < 0.12$), light-to-moderate ($0.12 < \mathrm{EDR}/(\mathrm{m}^{2/3}\,\mathrm{s}^{-1}) < 0.18$), moderate ($0.18 < \mathrm{EDR}/(\mathrm{m}^{2/3}\mathrm{s}^{-1}) < 0.30$), and moderate-to-severe ($\mathrm{EDR} > 0.30\ \mathrm{m}^{2/3}\mathrm{s}^{-1}$) CAT, respectively. The vertical lines in panel (a) represent the data binned according to the width of the bars, and the colored curves connect the values in the middle of the bars.

the buoyancy frequency N (Equation 5), and the Richardson number $\mathrm{Ri} = N^2/S^2$ as calculated from the interpolated one hourly IFS fields along the flights tracks at 5 s intervals, that is, at the same frequency of the observed ϵ_i or EDR_i values. In addition to these variables, the terrain height z_{ter} and the altitude of the thermal tropopause z_{tr} were determined at these intervals. The distribution of N in Figure 2a clearly shows two peaks belonging to the typical tropospheric and stratospheric values of 0.01 s^{-1} and 0.02 s^{-1}, respectively. Since the second peak is obviously dominant, it readily becomes apparent that the majority of the research flight was conducted under strongly stratified thermal conditions: only 6% of all flights took place for $N \leq 0.01\ \mathrm{s}^{-1}$, 42% for $N \geq 0.02\ \mathrm{s}^{-1}$, and the major part in the intermediate range. The shear distribution in Figure 2b has a single peak at around $3 \cdot 10^{-3}\ \mathrm{s}^{-1}$, larger shear values $S > 10^{-2}\ \mathrm{s}^{-1}$ are relatively rare (about 10% of the data) whereas about 77% of the data falls into the range $10^{-3}\ \mathrm{s}^{-1} \leq S \leq 10^{-2}\ \mathrm{s}^{-1}$. Consequently, the local Richardson numbers are predominantly larger than 1, yet there are a few cases where the $\mathrm{Ri} < 1$ (about 8%) or even falls below 0.25 (about 0.6%), Figure 2c. Overall, the local atmospheric conditions during the SOUTHTRAC research flights were dominated by an airflow under thermally stable conditions with low vertical shear and large Richardson numbers.

4.2. Eddy Dissipation Rates Along the Flight Tracks

Figure 3 shows the PDFs of ϵ_i and EDR_i (i stands for the foreward, sideward, and upward velocity components u_b, v_b, and w as outlined above in Section 2.3). For all flights analyzed, ϵ_{i} varies between 10^{-9} and $10^{-2}\ \mathrm{m}^2\ \mathrm{s}^{-3}$, in Figure 3a. The distribution of ϵ_{w} has a maximum at about $10^{-8}\ \mathrm{m}^2\ \mathrm{s}^{-3}$ (expected value of the distribution is $1.8 \cdot 10^{-7}\ \mathrm{m}^2\ \mathrm{s}^{-3}$), larger peak values are found for the foreward and sideward values of ϵ_i (expected values for ϵ_{ub} and ϵ_{vb} are 37.3 and $6.8 \cdot 10^{-7}\ \mathrm{m}^2\ \mathrm{s}^{-3}$, respectively), see Figure 3a. The slope values determined from the w-spectra vary widely around a mean value of about −5/3, due in part to randomness and in part to the lack of an inertial range in the atmospheric flow, see Figure 4. In fact, examination of the ϵ_i distributions shows that the turbulence along the flight tracks is usually strongly anisotropic, so that the assumption of an inertial spectrum is often only a crude approximation in determining ϵ. It is important to note that the results of the 10 and 100 Hz data are similar over the same flight segments, which is an important quality feature of the BAHAMAS data, see the comparison as presented in Appendix B.

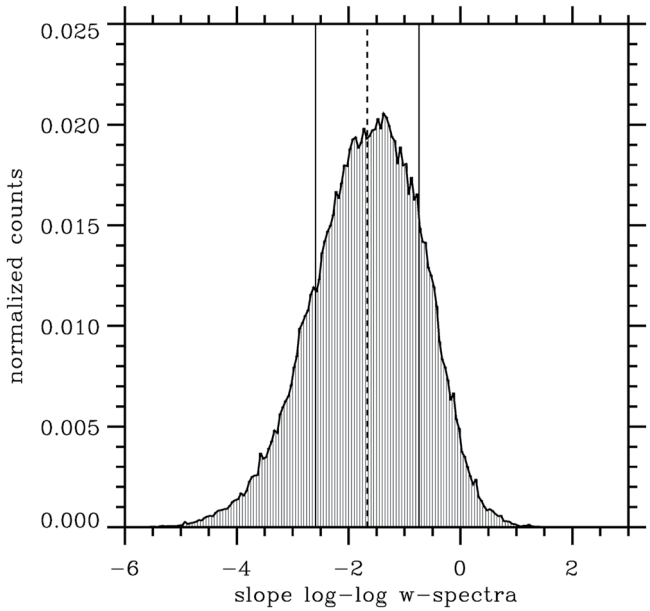

Figure 4. Probability density distribution of all slope values from the log-log w-spectra that were used to determine the eddy dissipation rates from the 10 Hz BAHAMAS spectra. The distribution is nearly normal with respect to mean value of −1.66824 ± 0.923314 (dashed and solid vertical lines). The deviation to the theoretical value of the Kolmogorov spectrum of −1.66667 is 0.095%. Altogether, 65.9% of all determined slope values are in the range −1.66824 ± 0.923314.

The equivalent distributions for EDR_i are shown in Figure 3b where the thresholds of light (0.05 < EDR < 0.12), light-to-moderate (0.12 < EDR < 0.18), moderate (0.18 < EDR < 0.30), and moderate-to-severe (EDR > 0.30) turbulence are indicated by dashed vertical lines. The EDR-values determined from the foreward and sideward velocity components are generally larger than EDR_w, but all three distributions indicate that the airflow along the SOUTHTRAC flights was predominantly calm and no severe turbulence was detected. This finding is reflected in the low expected values of the EDR distributions, which are 0.022, 0.016, and 0.010 for the three velocity components, respectively, well below the threshold for light turbulence but well above the estimated minimum values for SOUTHTRAC as presented in Appendix A. In fact, the probability of light turbulence is only 8.65%, 6.73%, 3.68% based on $EDR_{(ub,vb,w)}$ values and decreases to 0.20%, 0.15%, 0.03% for EDR > 0.18.

It should be mentioned that the experimental determination of eddy dissipation rates by aircraft observations based on ϵ_w seems more reliable than the determination by means of ϵ_{ub} or ϵ_{vb}. The quantity ϵ_w is still measurable at very weak turbulence, while the other quantities are dominated by noise. Thus, ϵ_w covers a wider range of values than ϵ_{ub} and ϵ_{vb}. As mentioned in Section 2.3, the determination of u_b and v_b depends largely on the airspeed, which is much larger than the turbulent velocity fluctuations, while the vertical velocity component is close to zero in the absence of turbulence and waves.

A slightly different view on the turbulent state of the airflow at the SOUTHTRAC flight levels can be obtained by evaluating the Ozmidov and Kolmogorov scales (Equations 3 and 4) as well as their ratio L_O/η as shown in Figure 5. The Ozmidov scale is generally smaller than 10 m, in most cases L_O < 1 m. The Kolmogorov scale is always smaller than 10 cm. As mentioned in Section 2.3, the airflow is actively turbulent when the ratio L_O/η is greater than about 10. The interesting result is that the proportion of ratio $L_O/\eta > 10$ (as determined for the ϵ_w-values) is 22.3%. According to this criterion, only about 1/5 of the airflow can be considered actively turbulent, but obviously—As shown by the small ϵ-values—With a low intensity.

Investigating Figure 5c, there are also values where the ratio is $L_O/\eta < 1$. For example, 19259 (out of 88340) data points with $L_{Ow}/\eta_w < 1$ entered the red curve in Figure 5c. Of these 19259 data points, only 3.2% had buoyancy

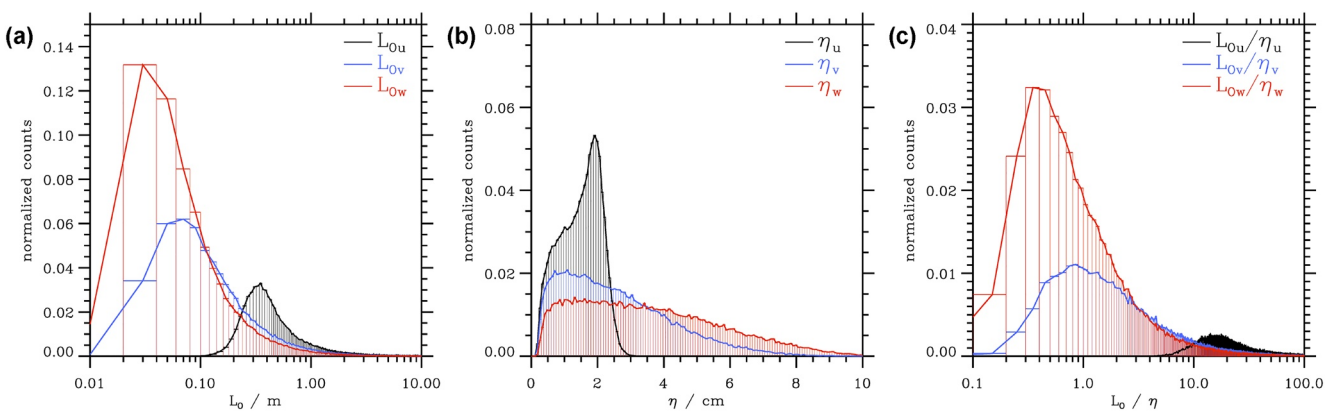

Figure 5. Probability density functions of the Ozmidov scale L_O (a), the Kolmogorov scale η (b), and the ratio L_O/η (c) computed from the 10Hz BAHAMAS ϵ_i data of HALOs research flights during SOUTHTRAC, see Figure 3a. The vertical lines in all panels represent the data binned according to the width of the bars, and the colored curves connect the values in the middle of the bars.

Table 1
Mean and Maximum ϵ Values Sorted for Selected Criteria: Terrain Height Beneath the Flight Tracks z_{ter}, Tropopause Altitude z_{tr}, Buoyancy Frequency N, and Vertical Shear of the Horizontal Wind S

		$\overline{\epsilon_u}$ 10^{-5} m^2 s^{-3}	$\overline{\epsilon_v}$ 10^{-5} m^2 s^{-3}	$\overline{\epsilon_w}$ 10^{-5} m^2 s^{-3}	ϵ_{ub}^{max} 10^{-2} m^2 s^{-3}	ϵ_{vb}^{max} 10^{-2} m^2 s^{-3}	ϵ_{w}^{max} 10^{-2} m^2 s^{-3}
1	all	9.29	7.09	3.04	2.92	3.59	1.19
2	$z_{ter} < 10$ m	9.36	7.26	3.27	2.92	3.59	1.19
3	$z_{ter} > 600$ m	10.62	8.37	3.33	1.91	2.04	0.91
4	$z > z_{tr}$	6.53	4.97	1.77	2.32	3.59	1.19
5	$z < z_{tr}$	16.85	12.89	6.52	2.92	2.38	1.02
6	$N > 0.02\,s^{-1}$	4.88	3.62	0.95	1.52	1.85	0.56
7	$N < 0.01\,s^{-1}$	22.95	20.25	10.06	1.13	2.39	1.19
8	$S < 0.001\,s^{-1}$	3.37	2.74	1.21	1.07	1.09	0.56
9	$0.01 < S/s^{-1} < 0.001$	6.79	5.44	2.09	1.86	2.39	1.01
10	$S > 0.01\,s^{-1}$	33.14	23.12	11.88	2.92	3.59	1.19
11	$L_{Ow}/\eta_w < 1$	1.05 10^{-1}	8.38 10^{-3}	1.11 10^{-3}	1.31 10^{-2}	2.07 10^{-3}	4.34 10^{-6}
12	$L_{Ow}/\eta_w > 10$	16.33	12.53	5.38	2.92	3.59	1.19

Note. In lines 11 and 12, the criteria is the ratio of the Ozmidov scale L_O to the Kolmogorov scale η as determined from the eddy dissipation rate ϵ_w.

frequencies $N < 0.015 s^{-1}$, that is, most of them were from situations with strongly stratified airflow in conditions where the absolute temperature was always below −45°C and 99.4% (70.7%) of the data had temperatures less than −50°C (−55°C), respectively. In a recent DNS study, Okino and Hanazaki (2020) assigned the regime where the Ozmidov scale is smaller than the Kolmogorov scale to decaying turbulence in its final stage under large stable stratification.

Table 1 lists the mean and maximum eddy dissipation rates ϵ_i sorted by selected criteria. The first line gives the total mean $\overline{\epsilon_i}$ and maximum values ϵ_i^{max} for all data: $\overline{\epsilon_i}$ ranges between 3 and $9.3 \cdot 10^{-5}$ m^2 s^{-3} (the smallest value for $\overline{\epsilon_w}$, the largest for $\overline{\epsilon_{ub}}$) and the maximum values are from 1.2 to $3.6 \cdot 10^{-2}$ m^2 s^{-3}. Lines 2 and 3 are intended to illustrate the influence of the terrain beneath the flight tracks. Interestingly, the maximum values do not occur over mountainous terrain ($z_{ter} > 600$ m), but over the lowlands or the sea. Only the mean ϵ values are slightly larger over mountainous terrain, cf. lines 3 and 2.

Distinguishing between tropospheric and stratospheric observations, lines 4 and 5 in Table 1 indicate that the mean ϵ values are reduced by about 30%–40% for the stratospheric observations compared to the all values as listed in line 1, while they are increased by over 80% (more than a factor of 2 for $\overline{\epsilon_w}$) for the tropospheric-only observations, indicating the more turbulent nature of the tropospheric airflow. The same tendency is seen when the buoyancy frequency N is used as a criterion instead of the tropopause height z_{tr}, see lines 6 and 7. A clear tendency for the mean and maximum ϵ values to rise is observed for increasing vertical shear S, see lines 8 to 10 in Table 1. The largest values are obtained for the strongest vertical shear $S > 0.01 s^{-1}$, indicating the controlling influence of the local shear on the turbulence intensity.

The lines 11 and 12 in Table 1 give the mean and maximum eddy dissipation rates ϵ_i sorted by the ratio $L_{Ow}/\eta_w = 1$ and 10, respectively. As expected for decaying turbulence, the ϵ_i-values are very small for $L_{Ow}/\eta_w = 1$. For active turbulence, the mean and maximum values are close to those of the troposphere, as listed in line 5.

4.3. Gravity Wave Momentum and Energy Fluxes

Figure 6 presents the results of the wave momentum and energy flux analysis in two ways. Figure 6a shows the Eliassen-Palm relation for the 357 segments as introduced in Section 2.2. The Eliassen-Palm relation

$$EF_z = -\mathbf{U} \cdot \mathbf{MF}$$

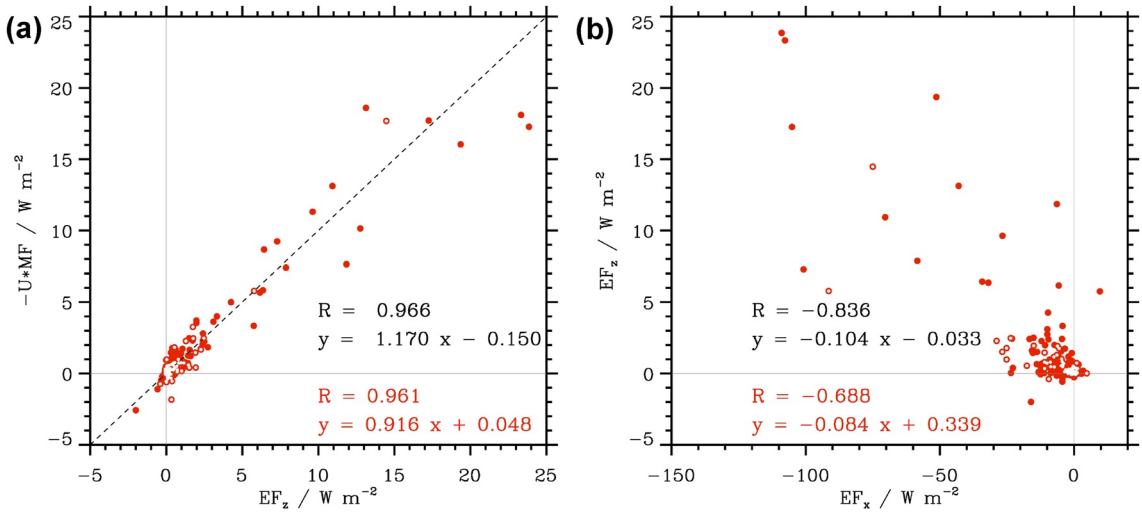

Figure 6. (a) Eliassen-Palm relation and (b) correlation between the vertical and the zonal wave energy fluxes EF_z and EF_x for all straight and constant altitude legs from the SOUTHTRAC research flights (red bullets and red font for the correlation coefficient and the linear fit). In both panels, the red circles filled with a white dot indicate those flight segments where the averaged wind was within a sector of 270° ± 15° (black font for the correlation coefficient and the linear fit).

provides a functional dependence of the vertical flux of wave energy $EF_z = w' p'_c$ and the negative scalar product of the vectors of the horizontal wind **U** and the vertical flux of horizontal momentum $\mathbf{MF} = (MF_x, MF_y)$ as defined in Section 2.2, see Figure S2 in Supporting Information S1. Figure 6a shows that both independently determined quantities EF_z and $-\mathbf{U} \cdot \mathbf{MF}$ are 96% correlated and are, therefore, linearly dependent. Their correlation increases slightly if only segments or flight legs are selected where the flow comes from a 30° sector around west. Since both quantities are in a one-to-one dependence, the majority of the detected waves have characteristics of linearly propagating, stationary, non-dissipative internal gravity waves in accordance with Eliassen and Palm (1960). Altogether, about 83% of all analyzed flight legs have positive wave energy EF_z, that is, the detected waves are propagating and transporting negative momentum upward.

Just as Figure 6a shows the momentum transport by vertically propagating waves, Figure 6b illustrates the horizontal wave propagation by correlating the vertical and zonal fluxes of wave energy. It is obvious that the zonal energy flux EF_x is mostly negative (for about 76% of all analyzed flight legs), that is, the gravity waves propagate against the mean, predominantly westerly flow, see Figure S1d in Supporting Information S1. This finding is particularly evident from the increase in the correlation coefficient from 69% to 84% when only the flight legs where the wind is almost from the west are considered. This means that both panels in Figure 6 indicate that most of the gravity waves detected are generated by stationary sources, that is, they are mountain waves.

In Figure 6, large values of $EF_z > 20$ W m^{-2} belong to research flight ST12, see Tables S1 and S2 in Supporting Information S1. During this particular flight, several short legs were flown in a hexagonal pattern directly above and in the lee of the Andes to capture a large-amplitude mountain wave event by GLORIA. However, for SOUTHTRAC, such large numbers are the exception, and values of $EF_z > 10$ W m^{-2} occurred only in a few segments during research flights ST10, ST12, and ST26, see Figure S2c in Supporting Information S1. The elevated EF_z values of more than 5 W m^{-2} during ST14 and ST18 are due to mountain waves over the Antarctic Peninsula (ST14) and the southern Andes (ST18).

4.4. EDR as Function of Local Flow Parameters

Figure 7 juxtaposes the observed EDR_i as function of the vertical gravity wave energy flux EF_z (panels a and b) and EDR_i as function of the inverse Richardson number Ri^{-1} (panels c and d). The left column in Figure 7 uses leg-averaged quantities for the 357 segments as introduced in Section 2.2. Here, the correlation between EDR_i and EF_z is essentially zero, that is, there is no correspondence of larger vertical wave fluxes implying higher turbulence intensity. The correlation between the leg-averaged EDR_i and Ri^{-1} is about 0.45 for all velocity components, that is, a decreasing leg-averaged Richardson number also indicates a higher turbulence intensity.

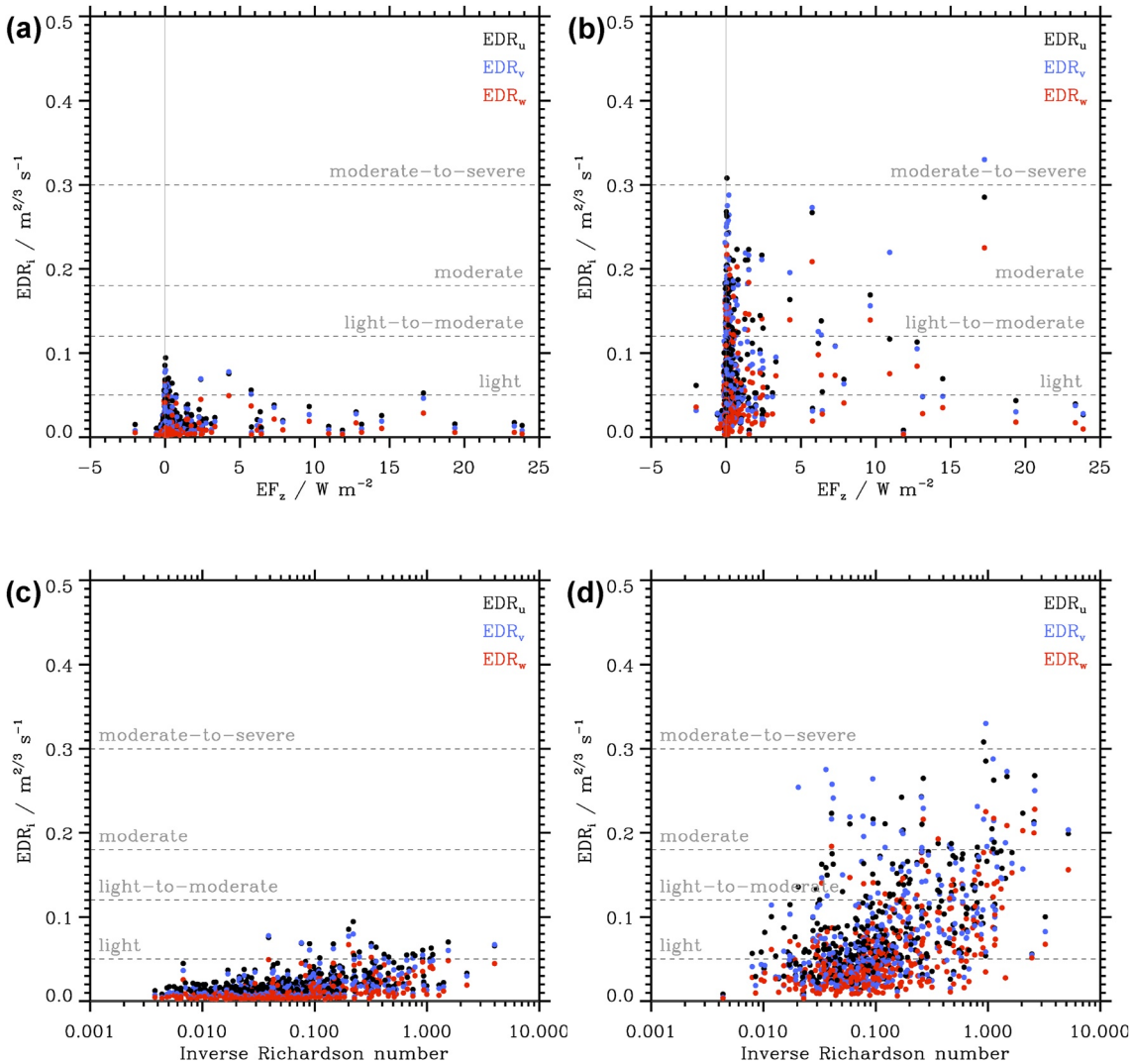

Figure 7. (a) Leg-averaged observed EDR_i-values as determined for all 357 segments as function of the observed vertical energy flux EF_z as determined from the HALO flight level BAHAMAS data as leg averages over these segments. (b) The same as in panel (a) but for the maximum EDR_i-values as observed in the segments. (c and d) the same as for panels (a and b) but as function of the inverse Richardson number. The dashed lines mark the limits of light ($0.05 < EDR/(m^{2/3}s^{-1}) < 0.12$), light-to-moderate ($0.12 < EDR/(m^{2/3}s^{-1}) < 0.18$), moderate ($0.18 < EDR/(m^{2/3}s^{-1}) < 0.30$), and moderate-to-severe ($EDR > 0.30\ m^{2/3}s^{-1}$) CAT, respectively.

This correlation becomes even larger (about 0.5) when the maximum Ri^{-1} values obtained within the 357 segments are correlated with the maximum EDR_i values, as shown in Figure 7d. It is interesting to note that the correlation with the inverse Richardson number decreases to values less than 0.32 if one takes all the 88340 point values independent from the selection of straight legs. These findings suggest that the turbulence observed during the SOUTHTRAC research flights is more dependent on the local shear (correlation coefficient with S ranges from 0.40 to 0.52) and the inverse thermal stratification (correlation coefficient with N^{-1} ranges from 0.3 to 0.45).

5. IFS Prediction Along the SOUTHTRAC Research Flights

Figure 8 shows the PDFs of the observed $\epsilon_{ub,w}$ and $EDR_{ub,w}$ as well as the corresponding numerical predictions of the IFS. For the latter, both the 15 ensemble members and the ensemble mean are plotted. As described in Section 3, the comparison of IFS data with SOUTHTRAC observations was performed for a reduced data set considering only observations, regardless of altitude, that fall within a time window of 30 min length centered

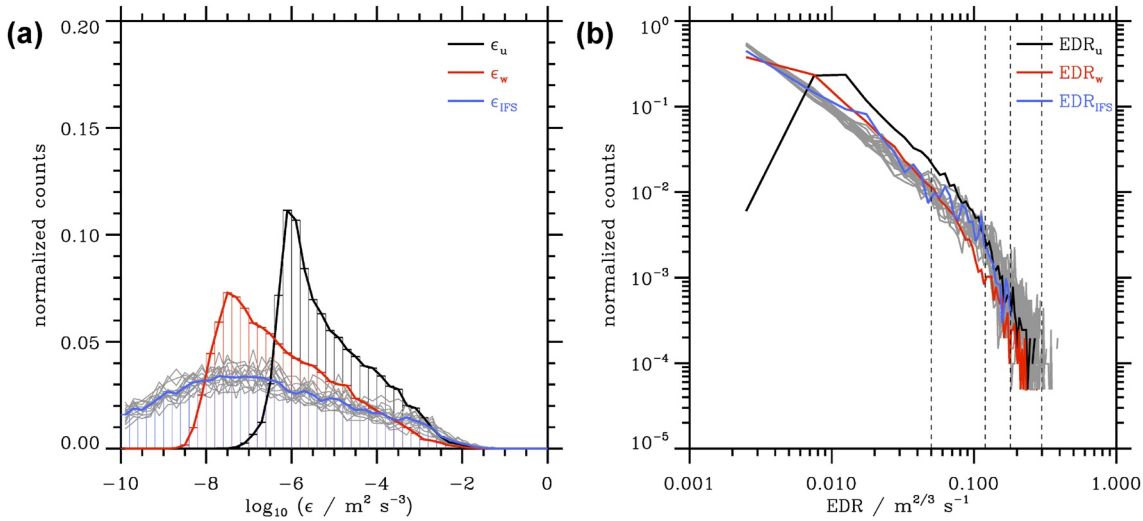

Figure 8. (a) Probability density functions of observed ϵ_{ub} and ϵ_w as well as the computed ϵ as determined from a 15 member IFS ensemble. The gray lines mark the distributions of the individual ensemble members, the blue line is the ensemble mean. Panel (b) the same plot as in (a) but for the EDR values. For the analysis method of the computed ϵ and EDR values, see text. The dashed lines in panel (b) mark the limits of light ($0.05 < \text{EDR}/(\text{m}^{2/3}\text{s}^{-1}) < 0.12$), light-to-moderate ($0.12 < \text{EDR}/(\text{m}^{2/3}\text{s}^{-1}) < 0.18$), moderate ($0.18 < \text{EDR}/(\text{m}^{2/3}\text{s}^{-1}) < 0.30$), and moderate-to-severe ($\text{EDR} > 0.30\,\text{m}^{2/3}\text{s}^{-1}$) CAT, respectively. The vertical lines in panel (a) represent the data binned according to the width of the bars, and the colored curves connect the values in the middle of the bars.

around each even hour in which IFS fields are available and do not deviate more than 160 m from the flight level. Therefore, the shapes of the black and red curves in Figure 8 differ slightly from those in Figure 3.

The PDF curves for the eddy dissipation rates ϵ_{ub} and ϵ_w are quite different from those of the IFS predictions. First of all, the expected values of the distributions differ in such a way that the one for the ensemble mean of the IFS with $1.1 \cdot 10^{-6}$ m^2 s^{-3} lies between the smaller expected value of $0.5 \cdot 10^{-6}$ m^2 s^{-3} for the ϵ_w-PDF and the larger expected value of $5.7 \cdot 10^{-6}$ m^2 s^{-3} for the ϵ_{ub}-PDF. Second, the IFS distributions are much broader than the observed PDFs that do not extend to values as small as those predicted by the numerical model. For the larger ϵ-values — which are, after all, more relevant to aviation turbulence –, the tails of the IFS distributions lie between those of the ϵ_{ub} and ϵ_w PDFs. Table 2a lists the relative fractions of ϵ values for selected thresholds of 10^{-3} m^2 s^{-3}, 10^{-4} m^2 s^{-3}, etc. The gradual decrease in relative fractions with increasing thresholds is well captured by the IFS. Quantitatively, the IFS overestimates the relative fractions of ϵ_w and underestimates the relative fractions of ϵ_{ub} in agreement with the visual impression of Figure 8a. We note that the ϵ_w values derived

Table 2
(a) Relative Fractions of Eddy Dissipation Rates Larger Than a Given Threshold as Derived From the Probability Density Functions of the Observed ϵ_u, ϵ_w and the IFS mean $\overline{\epsilon_{IFS}}$ as Plotted in Figure 8a

(a)	dϵ_u (%)	dϵ_w (%)	d$\overline{\epsilon_{IFS}}$ (%)
$\epsilon > 10^{-3}$ m^2 s^{-3}	3.6	1.4	3.7
$\epsilon > 10^{-4}$ m^2 s^{-3}	16.7	7.5	10.9
$\epsilon > 10^{-5}$ m^2 s^{-3}	38.5	19.8	20.5
$\epsilon > 10^{-6}$ m^2 s^{-3}	76.4	64.9	32.9
(b)	dEDR$_u$ (%)	dEDR$_w$ (%)	d$\overline{\text{EDR}_{IFS}}$ (%)
EDR > 0.05 m$^{2/3}$ s^{-1}	12.9	5.6	8.8
EDR > 0.12 m$^{2/3}$ s^{-1}	2.0	0.9	1.3
EDR > 0.18 m$^{2/3}$ s^{-1}	0.4	0.2	0.0
EDR > 0.30 m$^{2/3}$ s^{-1}	0.0	0.0	0.0

Note. (b) The same as in part (a) but for EDR values from the distributions as plotted in Figure 8b.

from the model reach below 10^{-9} m^2 s^{-3}, that is, the probability density distribution suggests even lower values than the measured ones and the ϵ_{MIN} as estimated in Appendix A.

The decreasing fraction of larger EDR values is represented by the numbers in Table 2b. The ensemble mean of the IFS only predicts light and light-to-moderate turbulence occurring in about 8.8% and 1.3% of all cases. The observed small fraction of moderate turbulence is not reproduced by the ensemble mean, however, investigating the individual ensemble members the following fractions are determined: 27.5%, 9.7%, 3.6%, and 0.2% for the EDR thresholds of 0.05, 0.12, 0.18, and 0.30 m$^{2/3}$ s^{-1}, respectively. These numbers reflect the potential of individual ensemble members to predict stronger turbulence which is also visible in the extent the gray curves have in Figure 8b.

The correlations for the SOUTHTRAC HALO data reveal the following values: Ensemble mean EDR_{IFS} versus $EDR_{ub} = 0.339$ and ensemble mean EDR_{IFS} versus $EDR_w = 0.394$, that is, there is a slightly larger correlation between the IFS ensemble mean and the observed PDFs of the vertical wind component. Also this correlation is slightly larger compared to the values for the commercial aircraft data (0.370) as listed in Table A1 of Appendix A. Interestingly, the correlation coefficients for individual ensemble members are always smaller than the ensemble mean, rarely exceeding 0.3 for EDR_{ub} (two members) and 0.35 for EDR_w (three members). The continuous ranked probability score (CRPS) as introduced in Appendix A gives a value of 0.0187 for ensemble mean EDR_{IFS} versus EDR_{ub} and a value of 0.0113 for EDR_{IFS} versus EDR_w. These values are about a factor 3 smaller than the one found for the comparison with the EDR values of the commercial airliners, see Table A1, Appendix A. Finally, the correlation between the observed EDR_{ub} and EDR_w is 0.899 for data above 8 km altitude. This correlation is really large and relies mostly on the small EDR values that are not relevant for aviation. Taken a threshold $EDR_{ub} > 0.05$, the corresponding correlations decrease to 0.743, and for $EDR_{ub} > 0.10$ to 0.637.

6. Discussion and Conclusion

The observational results of this SOUTHTRAC study are based mainly on the mean wind spectra for the foreward, sideward, and vertical velocity components u_b, v_b, and w obtained from the 10 and 100 Hz BAHAMAS data. The mean wind spectra are consistent with previous results, see Appendix B. The amplitudes of the w-spectra are an order of magnitude larger than those of the NAWDEX campaign (Schäfler et al., 2018). They compare well to the DEEPWAVE (Fritts et al., 2016) campaign results, as shown in Figure B2. Here, the somewhat larger spectral amplitudes of observed vertical winds on horizontal scales near 50 km are most likely associated with vertically propagating mountain waves detected over the Southern Andes during the SOUTHTRAC research flights, see Appendix B.

Addressing research question Q1: The observed mean ϵ_i values derived from these spectra range from 3 to $9 \cdot 10^{-5}$ m^2 s^{-3} and are generally somewhat smaller than previously reported values from radar and radiosonde observations as mentioned in the Introduction. However, other airborne observations also show small eddy dissipation rates, as reported by Schumann et al. (1995), Cho et al. (2003), Podglajen et al. (2017), with values as low as and below 10^{-8} m^2 s^{-3}. Still, however, the majority of the ϵ values at HALO flight levels is close to zero turbulent dissipation ($\epsilon < 10^{-5}$ m^2 s^{-3}, see Table 2). As determined from the PDFs, one important finding of our study is that occurrence of stronger turbulence in the free atmosphere is extremely rare; severe turbulence was not found on all the analyzed research flights. On the other hand, the aircraft measurements also indicate the existence of regions where the Ozmidov scale is smaller than the Kolmogorov scale, indicating decaying turbulence in its final stage at strong thermal stratification.

In agreement with earlier studies, the ϵ_w values are smaller (here about a factor 3) than the ϵ_{vb} and ϵ_{ub} values, for example, Lilly et al. (1974). We attribute this result to the anisotropy of atmospheric turbulence under the prevailing thermally stable conditions. Another influence is the different noise of the sensors responsible for the determination of the corresponding wind components (Giez et al., 2021). This influence also leads to an anisotropy of the derived eddy dissipation rates. As explained in Section 4.2, the determination of eddy dissipation rates based on the spectra of the vertical wind component seems to be the most trustworthy approach.

Maximum eddy dissipation rates reached values up to $3.6 \cdot 10^{-2}$ m^2 s^{-3}, again larger values are found for ϵ_{vb} and ϵ_{ub} compared to ϵ_w. There was no large difference between the eddy dissipation rates observed over lowlands and mountains. As the majority of the flights was in the stratosphere, this finding confirms former findings from radiosonde observations, for example, Ko et al. (2019). However, much larger mean eddy dissipation rates are found in the troposphere whereas the individual maximum ϵ_i-values are not varying much between the troposphere and stratosphere. This findings indicate two facts: the troposphere is generally much more turbulent than the stratosphere, a result as expected. Second, if local instabilities lead to the generation of turbulence, the maximum ϵ_i-values range between and 1 and $3.6 \cdot 10^{-2}$ m^2 s^{-3} for the SOUTHTRAC research flights. These values are an order of magnitude larger than those derived from the results of Whiteway et al. (2003) and could be a consequence of the larger local shear values during the SOUTHTRAC flights.

In search for the reasons of the observed turbulence events and responding to research question Q2, we followed three approaches: First, we explored the hypothesis that enhanced wave energy fluxes are implying also the likelihood of breaking gravity waves and, eventually, the production of turbulence. However, there is no correlation between EDR and EF_z. This result is consistent with the demonstrated linearity of the observed gravity waves under the prevailing stable thermal stratification as shown in Figure 6. On the other hand, and this is the second approach, the background flow can lead to local instabilities if the shear becomes large. Gravity waves can help to modify the airflow in such a way that the shear is locally enhanced. Indeed, it could be demonstrated that for increasing vertical shear of the horizontal wind the eddy dissipation rates increase. Therefore, it is not surprising that the EDR values are also correlated with the inverse Richardson number, indicating the dominant influence of vertical shear in turbulence production.

As a third approach, ensemble predictions of the IFS are used to correlate numerically computed EDR values with the observed EDR from the SOUTHTRAC research flights. This approach allowed us to answer research question Q3: Here, we attain similar, even slightly larger correlation coefficients as obtained by comparing the IFS ensemble with maximum EDR values reported from commercial airliners. The calculated EDR values often reach even smaller values than the measured ones, which underlines the realism of large areas with zero to very low turbulence. Taken all EDR observations into account, the IFS CAT predictions show more skill than a simple correlation with the inverse Richardson number. However, considering only the straight legs above about 8 km altitude, the correlation with Ri^{-1} is slightly superior or comparable to the IFS results. Furthermore, the contribution from the vertical diffusion scheme (Equation 9) dominates the determination of the IFS-CAT index in Equation 10 and supports the previously mentioned result that breaking gravity waves most likely played a minor role during the SOUTHTRAC flights.

Altogether, the study shows that the EDR values observed by a research aircraft are well-suited to be used as basis to validate NWP forecasts of CAT. The EDR estimates from the research aircraft observations are of high quality and large horizontal resolution and could provide a standard for model validation because they are independent of tuning procedures commonly used for EDR estimates of commercial aircraft. In particular, the presence of many straight and level flight tracks and the occurrence of weak to moderate turbulence during SOUTHTRAC flights could be well suited to test and challenge the capabilities of advanced CAT prediction techniques.

Appendix A: Verification Against the NOAA/MADIS Data Set

As part of the preparation of turbulence predictions using the operational IFS, Bechtold, Bramberger, Dörnbrack, Isaksen, and Leutbecher (2021) tested various CAT indices to find an optimal implementation that was both accurate and computationally efficient. In addition to the CAT2 index that was introduced in Section 3 and Equation 10, a weighted combination of the Ellrod1 index (Ellrod & Knapp, 1992) and GWD from Equation 7, referred to as CAT1 and defined as

$$\mathrm{CAT1} = 0.7\,\mathrm{Ellrod1} + \mathrm{GWD} \tag{A1}$$

was implemented on a test basis in the IFS. The Ellrod1 index is defined as the product of the vertical shear S from Equation 11 and the horizontal deformation of the horizontal wind field defined as

$$D = \sqrt{\left(\frac{\partial \mathrm{U}}{\partial y} + \frac{\partial \mathrm{V}}{\partial x}\right)^2 + \left(\frac{\partial \mathrm{U}}{\partial x} - \frac{\partial \mathrm{V}}{\partial y}\right)^2}.$$

A further index, named CAT12, is the arithmetic mean of CAT1 and CAT2. All three indices were verified against EDR reports from commercial aircraft over the continental U.S. for the period January–March 2019. The "maximum EDR" was retrieved from the NOAA/MADIS database, that is, the peak EDR value of 12 individual EDR measurements over a one-minute period. The data set and algorithm onboard commercial aircraft for calculating EDR is described in detail in Sharman et al. (2014).

For the January-March 2019 verification period under consideration, there are more than 4 million airline observations to which the IFS data had to be projected. A large fraction of these observations have an EDR value of zero, while the IFS produces a quasi-continuous field of non-zero CAT indices. Therefore, an EDR threshold of $\mathrm{EDR_{thr}} = 0.005\ \mathrm{m^{2/3}\ s^{-1}}$ was introduced for the observations, leaving only 197000 observations for statistical analysis. The chosen value $\mathrm{EDR_{thr}}$ can be derived approximately if we assume that the ratio $L_O/\eta = 10$ holds for active turbulence. Using Equation 6 for the minimum eddy dissipation rate as derived in Section 4.2, ϵ_{MIN} varies for all SOUTHTRAC measurements between $1.3 \cdot 10^{-8}\ \mathrm{m^2\ s^{-3}}$ and $1.1 \cdot 10^{-6}\ \mathrm{m^2\ s^{-3}}$ with a mean value of $4.1 \cdot 10^{-7}\ \mathrm{m^2\ s^{-3}}$. The respective minimum EDR values vary between $0.0023\ \mathrm{m^{2/3}\ s^{-1}}$ and $0.0104\ \mathrm{m^{2/3}\ s^{-1}}$ with a mean value of $0.0074\ \mathrm{m^{2/3}\ s^{-1}}$, a number corresponding closely to the threshold value $\mathrm{EDR_{thr}}$.

Two different types of IFS runs were conducted. First, daily deterministic 24-h forecasts were produced for the 3-month period January–March 2019. The CAT indices were computed from the hourly IFS output on the full vertical model resolution of 137 levels, but on a reduced horizontal grid of $0.3° \times 0.3°$. The comparison was made for the altitude range of 5–12 km, that is, the cruise levels of commercial aircraft, where the vertical resolution of the IFS is about 300 m. The projection of the IFS forecasts of the different CAT indices onto the observations is done by retaining all EDR observations 15 min before and after the full hour and allowing a maximum height difference between observations and model data of 160 m, that is, about half of the vertical spacing.

In addition, because atmospheric turbulence is intermittent in nature, CAT predictions were computed using an ensemble of IFS integrations. Therefore, the second set of daily IFS forecast runs was conducted using a 15-member ensemble at reduced spectral resolution of $\mathrm{T_{CO}}$ 639 for a two week-period. For this period from January 1-14, 2019, which includes 19600 observations, the high-resolution deterministic forecasts can also be compared to the ensemble predictions. The latter analysis utilizes the fair variant of the ensemble mean correlation and continuous rank probability score defined as

$$\mathrm{CPRS} = \frac{1}{M}\sum_{i=1}^{M}|X_i - Y| - \frac{1}{2\,M(M-1)}\sum_{i,j=1}^{M}|X_i - X_j|, \tag{A2}$$

where $M = 15$ is the number of ensemble members, X_i ($i = 1, 2, \ldots, M$) is an ensemble forecast projected onto the observations and Y the observation vector (Ferro, 2014). For a 1-member (deterministic) forecast, the fair CRPS reduces to the mean absolute error (MAE) function which corresponds to first term of the rhs of Equation A2. Thus, the ensemble CRPS is directly comparable to the MAE of the high-resolution deterministic forecasts, that is, the smaller the CPRS values, the smaller the deviations from the observations.

Figure A1 shows, for the entire 3-month period, the PDFs of the observations and the EDR estimates based on the three CAT indices from the deterministic IFS runs. Both CAT1 and CAT2 underestimate the relative occurrence of weak turbulence intensities with $\mathrm{EDR} < 0.05\ \mathrm{m^{2/3}\ s^{-1}}$ compared to the observations, which deviate from the log-normal law for small values. CAT2 overestimates the observed distribution for moderate to severe turbulence, while CAT1 underestimates the occurrence of severe to heavy turbulence with $\mathrm{EDR} > 0.3\ \mathrm{m^{2/3}\ s^{-1}}$. The best overall agreement with the observed distribution is obtained using a linear combination of CAT1 and CAT2.

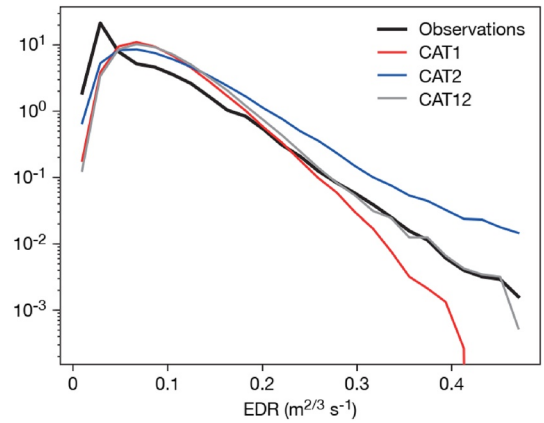

Figure A1. Probability density distributions of the EDR for the period January–March 2019 and for heights between 5 and 12 km, as obtained from the NOAA/MADIS observational data set (black) and from daily 0 to 24-h IFS forecasts projected onto the observation locations, where CAT1 and CAT2 are calculated from Equations A1 and 10 and CAT12 is the arithmetic mean of both. Taken from Bechtold, Bramberger, Dörnbrack, Isaksen, and Leutbecher (2021).

Table A1 (second and third column) shows the point correlations and MAE of the high-resolution deterministic IFS forecasts versus observations for January–March 2019. The average 3-month point correlations of the high-resolution predictions with observations are 0.33 for CAT1, 0.30 for CAT2, and 0.35 for CAT12, while the MAE is about 0.05 $m^{2/3}$ s^{-1}. The errors are slightly higher for CAT2 than for CAT1, which is to be expected since CAT2 is based on parameterized model tendencies and is, therefore, more variable than CAT1, which is based solely on state variables of the IFS.

For the shorter January period (fourth to seventh column in Table A1), the correlations reach values between 0.32 for CAT2 and 0.36 for CAT12, while the MAE remains at about 0.05 $^{2/3}$ s^{-1}. Comparing these values with those of the ensemble, we see that the ensemble performs significantly better. The mean correlations of the ensemble reach 0.37 for CAT2 and up to 0.40 for the combined product CAT12, while the CRPS is around 0.03 $m^{2/3}$ s^{-1}. This means, the IFS ensemble forecasts are of reasonable reliability, given that turbulence intensity is typically sorted in EDR intervals of 0.1 $m^{2/3}$ s^{-1}. The point correlations obtained may still seem relatively low, but their values must be put in relation to point correlations of about 0.53 obtained for 10-m wind speed forecasts over land and point correlations of 0.2–0.4 obtained when checking the forecast results of the same numerical IFS integrations of daily rainfall over tropical land with synoptic observations.

Table A1
Verification of Different EDR Parameters Against Observations for the High-Resolution Forecasts (HRES—Grid Spacing of About 9 km) for January–March 2019 and for HRES and the Ensemble Forecasts (ENS—Grid Spacing of About 18 km) for the Period 1–14 January 2019

EDR parameter	Corr HRES January–March	MAE HRES January–March	Corr HRES 1–14 January	Corr ENS 1–14 January	MAE HRES 1–14 January	CRPS ENS 1–14 January
CAT1	0.33	0.050	0.33	0.38	0.049	0.030
CAT2	0.30	0.057	0.32	0.37	0.054	0.034
CAT12	0.35	0.045	0.36	0.40	0.049	0.029

Note. Verification Statistics are Correlation (Corr), Mean Absolute Error (MAE) and the Fair Variant of the Continuous Ranked Probability Score (CRPS).

Appendix B: Comparison Between 10 and 100 Hz BAHAMAS Data

Figure B1 shows the frequency spectra for the three velocity components u_b, v_b, and w in aircraft-related coordinates as well as the spectra of the vertical body velocity wb of HALO for both the 10 Hz as well as the 100 Hz BAHAMAS data from SOUTHTRAC.Comparing first the wb frequency spectra with those of the velocity components, it becomes clear that the aircraft motion has no influence on the shape of the velocity spectra. Only in the 100 Hz data are the oscillations of the nose boom seen in the v_b and w spectra at 20 Hz. The spectrum of the vertical aircraft velocity wb shows enhanced amplitudes below 1 Hz at low and high ϵ values. This enhancement, at frequencies down to about 0.01 Hz, is a result of velocity- and autopilot-dependent mixed phugoid and pitch-mode body oscillations that occur in both turbulent and calm air (Nelson, 1998). The large peak at 6 Hz is likely due to structural wing or fuselage oscillations. The spectra of the 10 and 100 Hz data are very similar in the 0.4–4 Hz analysis interval used to determine eddy dissipation rates, a result that promotes confidence in the well-calibrated BAHAMAS measurements (Giez et al., 2021).

Figure B2 compares the mean horizontal and vertical kinetic energy spectra during three large airborne field campaigns for which high-resolution, well-calibrated in situ observations are available. During the DEEPWAVE campaign (Fritts et al., 2016), measurements were made with the NOAA/NCAR GV over New Zealand and most flights were mountain-crossing transects to study gravity waves excited by flow over the Southern Alps. The NAWDEX campaign (Schäfler et al., 2018) took place over the North Atlantic and there were few mountain crossings (Bramberger et al., 2020; Wilms et al., 2020). The SOUTHTRAC campaign is the subject of this paper. Additonally, the predictions of the simple model for the vertical energy spectra as introduced by Schumann (2019, Equation 7) are plotted. While the horizontal energy spectra are very similar for all campaigns (black curves in Figure B2), the vertical energy spectra differ (red curves in Figure B2): NAWDEX shows the smallest spectral amplitudes and both DEEPWAVE and SOUTHTRAC show one order of magnitude larger spectral energies. Interestingly, the maximum in SOUTHTRAC is shifted to horizontally longer waves at about 50 km, while DEEPWAVE has a distinct maximum at 10 km. Most likely, the waves observed during DEEPWAVE are trapped in the tropopause inversion layer, while the longer waves during SOUTHTRAC indicate vertically propagating

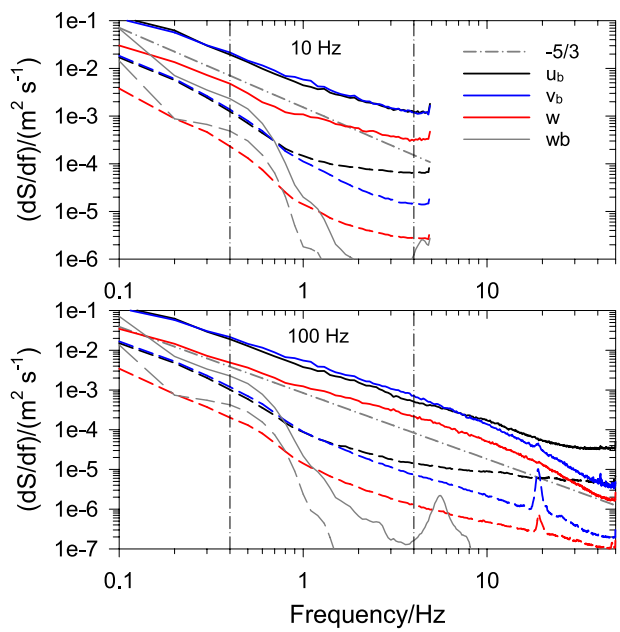

Figure B1. Frequency spectra of the foreward, sideward, and upward velocity components u_b, v_b, and w, respectively. In addition, the frequency spectra of the vertical body velocity wb of HALO are plotted in gray. The upper panel is for the 10 Hz BAHAMAS data, and the lower panel for 100 Hz BAHAMAS data during SOUTHTRAC. The solid and dashed lines distinguish high (full curves, $\epsilon > 2.2 \cdot 10^{-6}$ m^2 s^{-3}) and low (dashed, $\epsilon \leq 2.2 \cdot 10^{-6}$ m^2 s^{-3}) values of the eddy dissipation rates. The dashed-dotted gray line represents the −5/3 slope of the Kolmogorov (1941) spectrum.

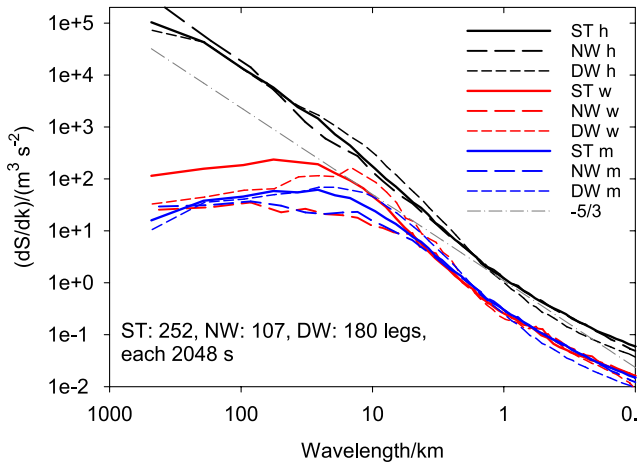

Figure B2. Mean spectra of kinetic energy per wavenumber interval dS/dk versus wavelength λ averaged over 20 log intervals per decade at high wavenumbers, as derived from 252 legs during SOUTHTRAC (ST), 107 legs observed during NAWDEX (NW), 180 legs observed during deep propagating gravity wave experiment of 2048s duration each. The curves denoted by "h" are for the horizontal energy, by "w" for the vertical energy spectra and the letter "m" refers to the model as introduced in Equation 7 by Schumann (2019)). The dashed-dotted gray line represents the −5/3 slope of the Kolmogorov (1941) spectrum.

modes excited by flow over the southern Andes. The spectral slopes in the inertial subrange of the w spectra are similar for all data sets and compare well with the model introduced by Schumann (2019, Equation 7), see blue curves Figure B2.

Data Availability Statement

Both datasets, that is, the 10 Hz as well as 100 Hz BAHAMAS measurements aboard HALO are available via the HALO data base at DLR https://halo-db.pa.op.dlr.de/mission/116. In addition, separate files are available containing eddy dissipation rates and slope values from the HALO observations computed for this work. The respective URLs for the 10 and 100 Hz data are: https://halo-db.pa.op.dlr.de/dataset/8497 and https://halo-db.pa.op.dlr.de/dataset/8496. Interested users must register to the HALO database by sending their request to the corresponding author or to the responsible instrument PI to become a member of the SOUTHTRAC mission. Operational CAT forecasts based on the CAT2 index are available via https://apps.ecmwf.int/codes/grib/param-db/?id=260290. The IFS research experiment "hhhx", which was run for this work, can be found at https://doi.org/10.21957/xbar-5611. The EDR values of the 15 IFS ensemble members used in this paper can be downloaded from the HALO database via https://halo-db.pa.op.dlr.de/dataset/8955. The EDR reports of the commercial airliners are provided by the U.S. National Oceanic and Atmospheric Administration's (NOAA) Meteorological Assimilation Data Ingest System (MADIS) public archive of aircraft data https://madis.ncep.noaa.gov/acars_variable_list.shtml.

Acknowledgments
This work was partly funded by the Federal Ministry for Education and Research under grants 01LG1907 (project WASCLIM) in the frame of the Role of the Middle Atmosphere in Climate (ROMIC)-program. Further support by the German Science foundation under grants GW-TP/DO 1020/9-1 and PACOG/RA 1400/6-1 in the frame of the DFG-research group MS-GWAVES is also acknowledged. Access to the operational ECMWF data used for the interpolation onto the HALO flight tracks was granted through the special project "Gravity Waves and Turbulence over the Andes" by AD. Open Access funding enabled and organized by Projekt DEAL.

References

Andrés Hernández, M. D., Hilboll, A., Ziereis, H., Förster, E., Krüger, O. O., Kaiser, K., et al. (2022). Overview: On the transport and transformation of pollutants in the outflow of major population centres—Observational data from the EMeRGe European intensive operational period in summer 2017. *Atmospheric Chemistry and Physics*, *22*(9), 5877–5924. https://doi.org/10.5194/acp-22-5877-2022

Bechtold, P., Bramberger, M., Dörnbrack, A., Isaksen, L., & Leutbecher, M. (2021). Forecasting clear-air turbulence. *ECMWF Newsletter*, *169*, 32–37.

Bechtold, P., Bramberger, M., Dörnbrack, A., Leutbecher, M., & Isaksen, L. (2021). *With a clear air turbulence (CAT) index from the IFS. In Experimenting* (Vol. 874). ECMWF Technical Memoranda. https://doi.org/10.21957/4l34tqljm

Beljaars, A. C. M., Brown, A. R., & Wood, N. (2004). A new parametrization of turbulent orographic form drag. *Quarterly Journal of the Royal Meteorological Society*, *130*(599), 1327–1347. https://doi.org/10.1256/qj.03.73

Bellenger, H., Wilson, R., Davison, J. L., Duvel, J. P., Xu, W., Lott, F., & Katsumata, M. (2017). Tropospheric turbulence over the tropical open ocean: Role of gravity waves. *Journal of the Atmospheric Sciences*, *74*(4), 1249–1271. https://doi.org/10.1175/JAS-D-16-0135.1

Bramberger, M., Dörnbrack, A., Wilms, H., Ewald, F., & Sharman, R. (2020). Mountain-wave turbulence encounter of the research aircraft HALO above Iceland. *Journal of Applied Meteorology and Climatology*, *59*(3), 567–588. https://doi.org/10.1175/JAMC-D-19-0079.1

Briggs, J., & Roach, W. T. (1963). Aircraft observations near jet streams. *Quarterly Journal of the Royal Meteorological Society*, *89*(380), 225–247. https://doi.org/10.1002/qj.49708938004

Cho, J. Y. N., Newell, R. E., Anderson, B. E., Barrick, J. D. W., & Thornhill, K. L. (2003). Characterizations of tropospheric turbulence and stability layers from aircraft observations. *Journal of Geophysical Research*, *108*(D20), 8784. https://doi.org/10.1029/2002JD002820

Clayson, C. A., & Kantha, L. (2008). On turbulence and mixing in the free atmosphere inferred from high-resolution soundings. *Journal of Atmospheric and Oceanic Technology*, *25*(6), 833–852. https://doi.org/10.1175/2007JTECHA992.1

Dörnbrack, A. (1998). Turbulent mixing by breaking gravity waves. *Journal of Fluid Mechanics*, *375*, 113–141. https://doi.org/10.1017/S0022112098002833

Dörnbrack, A., Kaifler, B., Kaifler, N., Rapp, M., Wildmann, N., Garhammer, M., et al. (2020). Unusual appearance of mother-of-pearl clouds above El Calafate, Argentina (50°21'S, 72°16'W). *Weather*, *75*(12), 378–388. https://doi.org/10.1002/wea.3863

Dutton, J. A. (1967). Belling the CAT in the sky. *Bulletin of the American Meteorological Society*, *48*(11), 813–820. https://doi.org/10.1175/1520-0477-48.11.813

Dutton, J. A., & Panofsky, H. A. (1970). Clear air turbulence: A mystery may Be unfolding. *Science*, *167*(3920), 937–944. https://doi.org/10.1126/science.167.3920.937

Eliassen, A., & Palm, E. (1960). On the transfer of energy in stationary mountain waves. *Geofysiske Publikasjoner*, *22*(3), 1–23.

Ellrod, G. P., & Knapp, D. I. (1992). An objective clear-air turbulence forecasting technique: Verification and operational use. *Weather and Forecasting*, *7*(1), 150–165. https://doi.org/10.1175/1520-0434(1992)007<0150:aocatf>2.0.co;2

Ferro, C. A. T. (2014). Fair scores for ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, *140*(683), 1917–1923. https://doi.org/10.1002/qj.2270

Friedl-Vallon, F., Gulde, T., Hase, F., Kleinert, A., Kulessa, T., Maucher, G., et al. (2014). Instrument concept of the imaging Fourier transform spectrometer GLORIA. *Atmospheric Measurement Techniques*, *7*(10), 3565–3577. https://doi.org/10.5194/amt-7-3565-2014

Fritts, D. C., Smith, R. B., Taylor, M. J., Doyle, J. D., Eckermann, S. D., Dörnbrack, A., et al. (2016). The deep propagating gravity wave experiment (DEEPWAVE): An airborne and ground-based exploration of gravity wave propagation and effects from their sources throughout the lower and middle atmosphere. *Bulletin of the American Meteorological Society*, *97*(3), 425–453. https://doi.org/10.1175/BAMS-D-14-00269.1

Geller, M. A., Love, P. T., & Wang, L. (2021). A climatology of unstable layers in the troposphere and lower stratosphere: Some early results. *Monthly Weather Review*, *149*(5), 1233–1245. https://doi.org/10.1175/MWR-D-20-0276.1

Gettelman, A., Hoor, P., Pan, L. L., Randel, W. J., Hegglin, M. I., & Birner, T. (2011). The extratropical upper troposphere and lower stratosphere. *Reviews of Geophysics*, *49*(3), RG3003. https://doi.org/10.1029/2011RG000355

Giez, A., Mallaun, C., Nenakhov, V., & Zöger, M. (2021). *Calibration of a nose boom mounted airflow sensor on an atmospheric research aircraft by inflight Maneuvers* Tech. Rep. No. 2021-17. Oberpfaffenhofen: DLR. Retrieved from https://elib.dlr.de/145704/

Giez, A., Mallaun, C., Zöger, M., Dörnbrack, A., & Schumann, U. (2017). Static pressure from aircraft trailing-cone measurements and numerical weather-prediction analysis. *Journal of Aircraft*, *54*(5), 1728–1737. https://doi.org/10.2514/1.C034084

Giez, A., Zöger, M., Dreiling, V., & Mallaun, C. (2019). *Static source error calibration of a nose boom mounted air data system on an atmospheric research aircraft using the trailing cone method* (Tech. Rep. No. 2019-07). Oberpfaffenhofen: DLR. Retrieved from https://elib.dlr.de/135789/

Goecke, T., & Machulskaya, E. (2021). Aviation turbulence forecasting at DWD with ICON: Methodology, case studies, and verification. *Monthly Weather Review*, *149*(7), 2115–2130. https://doi.org/10.1175/MWR-D-19-0383.1

Goldburg, A., & Pao, Y.-H. (1969). *Clear air turbulence and its detection*. Plenum Press.

Grubišić, V., Doyle, J. D., Kuettner, J., Mobbs, S., Smith, R. B., Whiteman, C. D., et al. (2008). The terrain-induced rotor experiment: A field campaign overview including observational highlights. *Bulletin of the American Meteorological Society*, *89*(10), 1513–1534. https://doi.org/10.1175/2008BAMS2487.1

Hocking, W. K. (1985). Measurement of turbulent energy dissipation rates in the middle atmosphere by radar techniques: A review. *Radio Science*, *20*(6), 1403–1422. https://doi.org/10.1029/RS020i006p01403

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., & Pfister, L. (1995). Stratosphere-troposphere exchange. *Reviews of Geophysics*, *33*(4), 403–439. https://doi.org/10.1029/95RG02097

Jesswein, M., Bozem, H., Lachnitt, H.-C., Hoor, P., Wagenhäuser, T., Keber, T., et al. (2021). Comparison of inorganic chlorine in the Antarctic and Arctic lowermost stratosphere by separate late winter aircraft measurements. *Atmospheric Chemistry and Physics*, *21*(23), 17225–17241. https://doi.org/10.5194/acp-21-17225-2021

Kantha, L., & Hocking, W. (2011). Dissipation rates of turbulence kinetic energy in the free atmosphere: MST radar and radiosondes. *Journal of Atmospheric and Solar-Terrestrial Physics*, *73*(9), 1043–1051. (Scientific Results from Networked and Multi-instrument studies based on MST Radar). https://doi.org/10.1016/j.jastp.2010.11.024

Kim, J.-H., Sharman, R., Strahan, M., Scheck, J. W., Bartholomew, C., Cheung, J. C. H., et al. (2018). Improvements in nonconvective aviation turbulence prediction for the world area forecast system. *Bulletin of the American Meteorological Society*, *99*(11), 2295–2311. https://doi.org/10.1175/BAMS-D-17-0117.1

Kim, S.-H., Kim, J., Kim, J.-H., & Chun, H.-Y. (2022). Characteristics of the derived energy dissipation rate using the 1 Hz commercial aircraft quick access recorder (QAR) data. *Atmospheric Measurement Techniques*, *15*(7), 2277–2298. https://doi.org/10.5194/amt-15-2277-2022

Knox, J. A. (1997). Possible mechanisms of clear-air turbulence in strongly anticyclonic flows. *Monthly Weather Review*, *125*(6), 1251–1259. https://doi.org/10.1175/1520-0493(1997)125<1251:pmocat>2.0.co;2

Ko, H.-C., Chun, H.-Y., Wilson, R., & Geller, M. A. (2019). Characteristics of atmospheric turbulence retrieved from high vertical-resolution radiosonde data in the United States. *Journal of Geophysical Research: Atmospheres*, *124*(14), 7553–7579. https://doi.org/10.1029/2019JD030287

Kohma, M., Sato, K., Tomikawa, Y., Nishimura, K., & Sato, T. (2019). Estimate of turbulent energy dissipation rate from the VHF radar and radiosonde observations in the Antarctic. *Journal of Geophysical Research: Atmospheres*, *124*(6), 2976–2993. https://doi.org/10.1029/2018JD029521

Kolmogorov, A. (1941). The local structure of turbulence in incompressible viscous fluids for very large Reynolds number. *Doklady Akademii Nauk SSSR*, *30*, 301–305.

Li, Q., Rapp, M., Schrön, A., Schneider, A., & Stober, G. (2016). Derivation of turbulent energy dissipation rate with the middle atmosphere Alomar radar system (MAARSY) and radiosondes at Andøya, Norway. *Annales Geophysicae*, *34*(12), 1209–1229. https://doi.org/10.5194/angeo-34-1209-2016

Lilly, D. K., Waco, D. E., & Adelfang, S. I. (1974). Stratospheric mixing estimated from high-altitude turbulence measurements. *Journal of Applied Meteorology and Climatology*, *13*(4), 488–493. https://doi.org/10.1175/1520-0450(1974)013<0488:smefha>2.0.co;2

Lübken, F.-J. (1993). *Experimental results on the role of turbulence for the heat budget of the ipper atmosphere*. Habilitationsschrift zur Erlangung der venia legendi der Hohen Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität zu Bonn.

Lübken, F.-J. (1997). Seasonal variation of turbulent energy dissipation rates at high latitudes as determined by in situ measurements of neutral density fluctuations. *Journal of Geophysical Research*, *102*(D12), 13441–13456. https://doi.org/10.1029/97JD00853

Luce, H., Wilson, R., Dalaudier, F., Hashiguchi, H., Nishi, N., Shibagaki, Y., & Nakajo, T. (2014). Simultaneous observations of tropospheric turbulence from radiosondes using Thorpe analysis and the VHF MU radar. *Radio Science*, *49*(11), 1106–1123. https://doi.org/10.1002/2013RS005355

Nelson, R. C. (1998). *Flight stability and automatic control* (1st edn). McGraw-Hill.

Okino, S., & Hanazaki, H. (2020). Direct numerical simulation of turbulence in a salt-stratified fluid. *Journal of Fluid Mechanics*, *891*, A19. https://doi.org/10.1017/jfm.2020.146

Oncley, S. P., Friehe, C. A., Larue, J. C., Businger, J. A., Itsweire, E. C., & Chang, S. S. (1996). Surface-layer fluxes, profiles, and turbulence measurements over uniform terrain under near-neutral conditions. *Journal of the Atmospheric Sciences*, *53*(7), 1029–1044. https://doi.org/10.1175/1520-0469(1996)053<1029:slfpat>2.0.co;2

Orr, A., Bechtold, P., Scinocca, J., Ern, M., & Janiskova, M. (2010). Improved middle atmosphere climate and forecasts in the ECMWF model through a nonorographic gravity wave drag parameterization. *Journal of Climate*, *23*(22), 5905–5926. https://doi.org/10.1175/2010JCLI3490.1

Ozmidov, R. V. (1965). On the turbulent exchange in a stably stratified ocean. *Izvestiya Akademii Nauk SSSR, Metally*, *1*, 861–871.

Panofsky, H. A., Dutton, J. A., Hemmerich, K. H., McCreary, G., & Loving, N. V. (1968). Case studies of the distribution of cat in the troposphere and stratosphere. *Journal of Applied Meteorology and Climatology*, *7*(3), 384–389. https://doi.org/10.1175/1520-0450(1968)007<0384:csotdo>2.0.co;2

Pavelin, E., Whiteway, J. A., Busen, R., & Hacker, J. (2002). Airborne observations of turbulence, mixing, and gravity waves in the tropopause region. *Journal of Geophysical Research*, *107*(D10), ACL81–ACL86. https://doi.org/10.1029/2001JD000775

Pepler, S. J., Vaughan, G., & Hooper, D. A. (1998). Detection of turbulence around jet streams using a VHF radar. *Quarterly Journal of the Royal Meteorological Society*, *124*(546), 447–462. https://doi.org/10.1002/qj.49712454605

Piper, M., & Lundquist, J. K. (2004). Surface layer turbulence measurements during a frontal passage. *Journal of the Atmospheric Sciences*, *61*(14), 1768–1780. https://doi.org/10.1175/1520-0469(2004)061<1768:sltmda>2.0.co;2

Podglajen, A., Bui, T. P., Dean-Day, J. M., Pfister, L., Jensen, E. J., Alexander, M. J., et al. (2017). Small-scale wind fluctuations in the tropical tropopause layer from aircraft measurements: Occurrence, nature, and impact on vertical mixing. *Journal of the Atmospheric Sciences*, *74*(11), 3847–3869. https://doi.org/10.1175/JAS-D-17-0010.1

Rapp, M., Kaifler, B., Dörnbrack, A., Gisinger, S., Mixa, T., Reichert, R., et al. (2021). SOUTHTRAC-GW: An airborne field campaign to explore gravity wave dynamics at the world's strongest hotspot. *Bulletin of the American Meteorological Society*, *102*(4), E871–E893. https://doi.org/10.1175/BAMS-D-20-0034.1

Riese, M., Oelhaf, H., Preusse, P., Blank, J., Ern, M., Friedl-Vallon, F., et al. (2014). Gimballed Limb observer for radiance imaging of the atmosphere (GLORIA) scientific objectives. *Atmospheric Measurement Techniques*, *7*(7), 1915–1928. https://doi.org/10.5194/amt-7-1915-2014

Riley, J. J., & Lindborg, E. (2008). Stratified turbulence: A possible interpretation of some geophysical turbulence measurements. *Journal of the Atmospheric Sciences*, *65*(7), 2416–2424. https://doi.org/10.1175/2007JAS2455.1

Rodriguez Imazio, P., Dörnbrack, A., Urzua, R. D., Rivaben, N., & Godoy, A. (2022). Clear air turbulence observed across a tropopause fold over the Drake passage—A case study. *Journal of Geophysical Research: Atmospheres*, *127*(4), e2021JD035908. https://doi.org/10.1029/2021JD035908

Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., et al. (2018). The North atlantic waveguide and downstream impact experiment. *Bulletin of the American Meteorological Society*, *99*(8), 1607–1637. https://doi.org/10.1175/BAMS-D-17-0003.1

Schumann, U. (2019). The horizontal spectrum of vertical velocities near the tropopause from global to gravity wave scales. *Journal of the Atmospheric Sciences*, *76*(12), 3847–3862. https://doi.org/10.1175/JAS-D-19-0160.1

Schumann, U. (2020). *Measurement and model data comparisons for the HALO-FAAM formation flight during EMeRGe on 17 July 2017* (Tech. Rep. No. 2020-48). Oberpfaffenhofen: DLR Insitut für Physik der Atmosphäre. Retrieved from https://zenodo.org/record/4427965

Schumann, U., Konopka, P., Baumann, R., Busen, R., Gerz, T., Schlager, H., et al. (1995). Estimate of diffusion parameters of aircraft exhaust plumes near the tropopause from nitric oxide and turbulence measurements. *Journal of Geophysical Research*, *100*(D7), 14147–14162. https://doi.org/10.1029/95JD01277

Sharman, R. D., Cornman, L. B., Meymaris, G., Pearson, J., & Farrar, T. (2014). Description and derived climatologies of automated in situ eddy-dissipation-rate reports of atmospheric turbulence. *Journal of Applied Meteorology and Climatology*, *53*(6), 1416–1432. https://doi.org/10.1175/JAMC-D-13-0329.1

Sharman, R. D., Doyle, J. D., & Shapiro, M. A. (2012). An investigation of a commercial aircraft encounter with severe clear-air turbulence over western Greenland. *Journal of Applied Meteorology and Climatology*, *51*(1), 42–53. https://doi.org/10.1175/JAMC-D-11-044.1

Sharman, R. D. & Lane, T. (Eds.) (2016). *Aviation turbulence: Processes, detection, prediction*. Springer. https://doi.org/10.1007/978-3-319-23630-8

Sharman, R. D., Trier, S. B., Lane, T. P., & Doyle, J. D. (2012). Sources and dynamics of turbulence in the upper troposphere and lower stratosphere: A review. *Geophysical Research Letters*, *39*(12). https://doi.org/10.1029/2012GL051996

Shutts, G. (2015). A stochastic convective backscatter scheme for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *141*(692), 2602–2616. https://doi.org/10.1002/qj.2547

Smalikho, I. N. (1997). Accuracy of the turbulent energy dissipation rate estimation from the temporal spectrum of wind velocity fluctuations. *Atmospheric and Oceanic Optics*, *10*(8), 559–563.

Smith, R. B., Nugent, A. D., Kruse, C. G., Fritts, D. C., Doyle, J. D., Eckermann, S. D., et al. (2016). Stratospheric gravity wave fluxes and scales during DEEPWAVE. *Journal of the Atmospheric Sciences*, *73*(7), 2851–2869. https://doi.org/10.1175/JAS-D-15-0324.1

Smith, R. B., Woods, B. K., Jensen, J., Cooper, W. A., Doyle, J. D., Jiang, Q., & Grubišić, V. (2008). Mountain waves entering the stratosphere. *Journal of the Atmospheric Sciences*, *65*(8), 2543–2562. https://doi.org/10.1175/2007JAS2598.1

Storer, L. N., Gill, P. G., & Williams, P. D. (2019). Multi-model ensemble predictions of aviation turbulence. *Meteorological Applications*, *26*(3), 416–428. https://doi.org/10.1002/met.1772

Strauss, L., Serafin, S., Haimov, S., & Grubišić, V. (2015). Turbulence in breaking mountain waves and atmospheric rotors estimated from airborne in situ and Doppler radar measurements. *Quarterly Journal of the Royal Meteorological Society*, *141*(693), 3207–3225. https://doi.org/10.1002/qj.2604

Thorpe, S. A. (1977). Turbulence and mixing in a Scottish Loch. *Philosophical Transactions of the Royal Society of London—Series A: Mathematical and Physical Sciences*, *286*(1334), 125–181. https://doi.org/10.1098/rsta.1977.0112

Trier, S. B., Sharman, R. D., Muñoz-Esparza, D., & Lane, T. P. (2020). Environment and mechanisms of severe turbulence in a midlatitude cyclone. *Journal of the Atmospheric Sciences*, *77*(11), 3869–3889. https://doi.org/10.1175/JAS-D-20-0095.1

Turner, H. S. (1955). Clear air turbulence and topography. *Weather*, *10*(9), 294–297. https://doi.org/10.1002/j.1477-8696.1955.tb00218.x

Vinnichenko, N. K., & Dutton, J. A. (1969). Empirical studies of atmospheric structure and spectra in the free atmosphere. *Radio Science*, *4*(12), 1115–1126. https://doi.org/10.1029/RS004i012p01115

White, F. M. (1991). *Viscous fluid flow* (2nd edn), McGraw-Hill.

Whiteway, J. A., Pavelin, E. G., Busen, R., Hacker, J., & Vosper, S. (2003). Airborne measurements of gravity wave breaking at the tropopause. *Geophysical Research Letters*, *30*(20), 2003GL018207. https://doi.org/10.1029/2003GL018207

Wildmann, N., Eckert, R., Dörnbrack, A., Gisinger, S., Rapp, M., Ohlmann, K., & van Niekerk, A. (2021). In situ measurements of wind and turbulence by a motor glider in the Andes. *Journal of Atmospheric and Oceanic Technology*, *38*(4), 921–935. https://doi.org/10.1175/JTECH-D-20-0137.1

Wilms, H., Bramberger, M., & Dörnbrack, A. (2020). Observation and simulation of mountain wave turbulence above Iceland: Turbulence intensification due to wave interference. *Quarterly Journal of the Royal Meteorological Society*, *146*(732), 3326–3346. https://doi.org/10.1002/qj.3848

Zhang, J., Zhang, S. D., Huang, C. M., Huang, K. M., Gong, Y., Gan, Q., & Zhang, Y. H. (2019). Latitudinal and topographical variabilities of free atmospheric turbulence from high-resolution radiosonde data sets. *Journal of Geophysical Research: Atmospheres*, *124*(8), 4283–4298. https://doi.org/10.1029/2018JD029982