Aircraft observations of gravity wave activity and turbulence in the tropical tropopause layer: prevalence, influence on cirrus and comparison with global storm-resolving models

Rachel Atlas[1] and Christopher S. Bretherton[1,2]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA
[2]Allen Institute for Artificial Intelligence, Seattle, WA, USA

Correspondence: Rachel Atlas (ratlas@uw.edu)

Abstract. The tropical tropopause layer (TTL) is a sea of vertical motions. Convectively-generated gravity waves create vertical winds on scales of a few to 1000s of kilometers as they propagate in a stable atmosphere. Turbulence from gravity wave breaking, radiatively-driven convection and Kelvin-Helmholtz instabilities stirs up the TTL on the kilometer scale. TTL cirrus, which moderate the water vapor concentration in the TTL and stratosphere, form in the cold phases of large-scale (> 100 km)
5 wave activity. It has been proposed in several modelling studies that small-scale (< 100 km) vertical motions control the ice crystal number concentration and the dehydration efficiency of TTL cirrus. Here, we present the first observational evidence for this.

High-rate vertical winds measured by aircraft are a valuable and underutilized tool for constraining small-scale TTL vertical wind variability, examining its impacts on TTL cirrus, and evaluating atmospheric models. We use 20 Hz data from five
10 National Aeronautics and Space Administration (NASA) campaigns to quantify small-scale vertical wind variability in the TTL, and to see how it varies with ice water content, distance from deep convective cores, and height in the TTL.

We find that 1 Hz vertical winds are well represented by a normal distribution with a standard deviation of 0.2–0.4 m s^{-1}. Consistent with a previous observational study that analyzed two out of the five aircraft campaigns that we analyze here, we find that turbulence is enhanced over the tropical West Pacific and within 100 km of convection, and is most common in the
15 lower TTL (14-15.5 km) closer to deep convection, and in the upper TTL (15.5-17 km) further from deep convection.

An algorithm to classify turbulence, and long wavelength (5 km $< \lambda <$ 100 km) and short wavelength ($\lambda <$ 5 km) gravity wave activity, during level flight legs, is applied to data from the Airborne Tropical TRopopause Experiment (ATTREX). The most commonly sampled conditions are 1) a quiescent atmosphere with negligible small-scale vertical wind variability, 2) long wavelength gravity wave activity (LW GWA), and 3) LW GWA with turbulence. Turbulence rarely occurs in the absence of
20 gravity wave activity.

Cirrus with ice crystal number concentrations exceeding 20 L^{-1} and ice water content exceeding 1 mg m^{-3} are rare in a quiescent atmosphere, but about 20 times more likely when there is gravity wave activity and 50 times more likely when there is also turbulence, confirming the results of the aforementioned modeling studies.

Our observational analysis shows that small-scale gravity waves strongly influence ice crystal number concentration and ice
25 water content within TTL cirrus. Global storm-resolving models have recently been run with horizontal grid spacing between

1 and 10 km, sufficient to resolve some small-scale gravity wave activity. We evaluate simulated vertical wind spectra (10-100 km) from four global-storm resolving simulations that have horizontal grid spacing of 3–5 km, with aircraft observations from ATTREX. We find that all four models have too little resolved vertical wind at horizontal wavelengths between 10 and 100 km, and thus too little small-scale gravity wave activity, although the bias is much less pronounced in global SAM than in the other models. We expect that deficient small-scale gravity wave activity significantly limits the realism of simulated ice microphysics in these models, and that improved representation requires moving to finer horizontal and vertical grid spacing.

1 Introduction

Time-mean vertical motions in the Tropical Tropopause Layer (TTL) are less than 1 cm s^{-1} (Ortland and Alexander, 2014), and synoptic-scale vertical motions on scales exceeding 100 km are typically less than 10 cm s^{-1} (Section 2.1). However, even well away from deep convective updrafts, gravity waves and turbulence can locally produce vertical winds often exceeding 1 m s^{-1}, dwarfing the magnitudes of the synoptic-scale winds.

Vertical motions on all scales influence TTL cirrus clouds, which dehydrate the TTL (Jensen et al., 2013). The dehydrated air is then lofted into the stratosphere (Holton et al., 1995). Decreased water vapor in the stratosphere cools Earth's surface and increases stratospheric ozone (Shindell, 2001). It has been estimated that a 1 ppmv increase in stratospheric water vapor has a radiative forcing of 0.24 W m^{-2} (Solomon et al., 2010). TTL cirrus with higher ice crystal number concentrations dehydrate the TTL and stratosphere, and cool Earth's surface, more effectively (Jensen et al., 2013). Thus, TTL cirrus occurrence and microphysical properties together determine the impact of TTL cirrus on climate.

Recently, studies have used observed temperature fluctuations from aircraft (Kim et al., 2016), satellite (Chang and L'Ecuyer, 2020), and balloon measurements (Bramberger et al., 2022) to show that TTL cirrus occurrence is tightly controlled by predominantly large-scale (> 100 km) wave activity.

Numerous modelling studies have suggested that small-scale (< 100 km) vertical motions strongly influence TTL cirrus microphysics, by initiating new instances of homogeneous freezing (Dinh et al., 2010; Spichtinger and Krämer, 2013; Schoeberl et al., 2015; Jensen et al., 2016; Dinh et al., 2016). However, no existing observational studies have investigated this.

Large-scale wave activity creates large vertical displacements on long timescales and small-scale vertical motions create small vertical displacements on short timescales. On short timescales, cirrus clouds are less able to adjust to rising supersaturations by growing existing ice crystals and are more likely to experience supersaturations exceeding the homogeneous nucleation threshold, forcing them to nucleate new ice particles.

Much of the air in the TTL is highly supersaturated; in the temperature range investigated here (185 - 210 K), the threshold for homogeneous nucleation is between 1.5 and 2.3 times ice saturation (Schneider et al., 2021). If updrafts force atmospheric supersaturation beyond these thresholds, new ice particles will form through homogeneous nucleation. Subsequent downdrafts following homogeneous nucleation will reduce atmospheric saturation. If the timescales are short enough, the cirrus cloud may not have quenched the atmosphere down to supersaturation following homogeneous nucleation, and the ice particles may continue growing in the downdrafts. If the cirrus cloud did have time to quench the atmosphere down to supersaturation, the

downdraft will cause sub-saturation. As most TTL cirrus ice particles are of similar size, this would lead to a reduction of particle size, but is unlikely to significantly decrease the number of ice crystals. Thus, particles homogeneously nucleated in updrafts can persist through downdrafts and irreversibly increase ice crystal number concentrations in cirrus.

Small-scale motions include small-scale gravity wave activity and turbulence. Most of the aforementioned modelling studies have connected small-scale gravity wave activity in particular to homogeneous nucleation in TTL cirrus (Spichtinger and Krämer, 2013; Schoeberl et al., 2015; Jensen et al., 2016; Dinh et al., 2016; Podglajen et al., 2018). Fewer studies have considered the effects of turbulence. Dinh et al. (2010) found that radiatively-driven convective turbulence in combination with radiatively-driven mesoscale circulations helped maintain a simulated TTL cirrus cloud for several days in the absence of strong gravity wave activity. The cirrus cloud in that study, which was 0.5 km thick with ice water contents of up to 1 mg m^{-3}, achieved relatively small turbulent-driven updrafts up to 2.5 cm s^{-1}, which is below the noise floor of aircraft vertical wind measurements. However, TTL cirrus can be several km thick, creating larger radiative destabilization over a deeper layer that could induce stronger turbulent updrafts. No existing studies have examined the relative roles of turbulence and gravity wave activity on TTL cirrus microphysics.

Many studies have used vertical wind measurements from aircraft to characterize small-scale motions over the past two decades. Most have focused on the midlatitude troposphere (Gultepe and Starr, 1995; Koch et al., 2005; Muhlbauer et al., 2014, e.g.), but a series of National Aeronautics and Space Administration (NASA) flight campaigns gathered high-rate vertical wind measurements in the TTL that are also well-suited for such analysis.

Using data from two of those campaigns, the Airborne Tropical TRopopause Experiment (ATTREX) in 2013 and 2014, Podglajen et al. (2017) investigated the frequency and characteristics of turbulence in the TTL, and estimated its effect on transport. They found that turbulence more commonly occurs closer to deep convection and is most common in the lower TTL (< 15.5 km) within 500 km of convection, and in the upper TTL (> 15.5 km) further away from deep convection. Podglajen et al. (2017) also found evidence suggesting clear-sky sources of turbulence are dominant in the TTL.

Our study shows that these findings also hold for other NASA field campaigns that sampled the TTL in different geographical areas, years and seasons (Sections 3.1 and 3.2). It also makes two major new contributions. The first is to distinguish between turbulence and gravity wave activity in the TTL, and examine their separate influences on TTL cirrus microphysics (Section 3.3). The second is to compare the spatial power spectrum of TTL vertical wind in global-storm resolving models with aircraft measurements over the tropical West Pacific (Section 3.4). Section 4 presents our conclusions.

2 Preparing the dataset

2.1 Aircraft measurements of vertical wind

We analyze data from aircraft campaigns that simultaneously measured ice water content (IWC) and high-rate (sampled faster than 1 Hz) vertical wind in and above the TTL, which we define for the purpose of this study as the atmospheric layer between 14 and 18 km altitude and between 30°N and 30°S. Five NASA campaigns meet these criteria: the Airborne Tropical TRopopause EXperiment (Jensen et al., 2017, ATTREX) 2013 and 2014 (we treat these two different years of ATTREX as
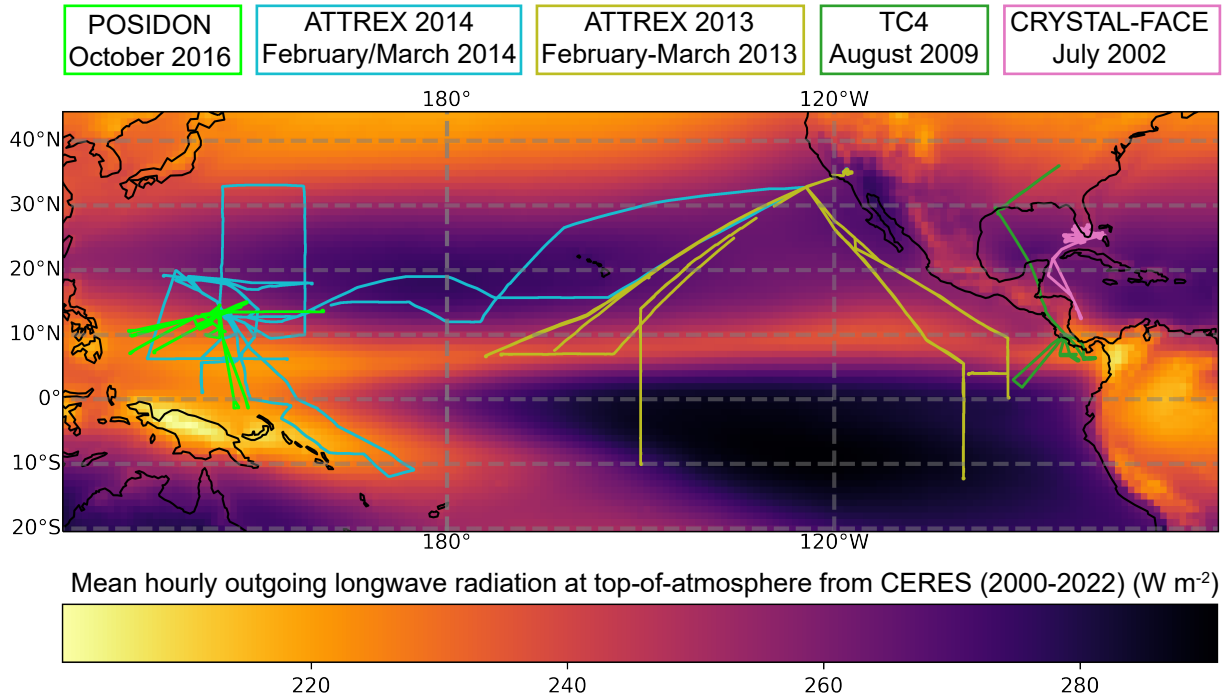
Figure 1. Map of mean top-of-atmosphere outgoing longwave radiation from CERES with flight tracks from the five campaigns used in this study overlaid.

two different campaigns), Pacific Oxidants, Sulfur, Ice, Dehydration, and cONvection (POSIDON), the Tropical Composition, Cloud and Climate Coupling Experiment (Toon et al., 2010, TC4), and the Cirrus Regional Study of Tropical Anvils and Cirrus Layers - Florida Area Cirrus Experiment (Jensen et al., 2004, CRYSTAL-FACE). TC4 and CRYSTAL-FACE used multiple

95   aircraft but only data from the WB-57 meet our criteria. We only include data within the TTL or above the TTL in our analysis.

Figure 1 shows flight tracks from these five campaigns, overlaid on time-mean outgoing longwave radiation (OLR) from the entire Clouds and the Earth's Radiant Energy System (CERES) level 3 satellite-based dataset (Doelling et al., 2013; NASA/LARC/SD/ASDC, 2017), which spans nearly 22 years. Smaller OLRs indicate more frequent deep convection. The

100   frequency of deep convection during the flight campaigns may be different than for the entire CERES dataset, due to seasonal and inter-annual variability. ATTREX 2013 is an outlier in that the majority of its sampling was over the tropical East and Central Pacific, which has infrequent deep convection. TC4, CRYSTAL-FACE and POSIDON sampled primarily close to deep convection near Costa Rica, near Florida, and over the tropical West Pacific, respectively. ATTREX 2014 consisted of two transit flights that sampled far from convection over the East and central Pacific, and six science flights that sampled close to

105   convection over the tropical West Pacific.

All campaigns measured vertical wind at 20 Hz using NASA's Meteorological Measurements Systems (MMS) instrument (Scott et al., 1990). The MMS estimates the vertical wind as the difference between the vertical speed of air relative to the aircraft, and the vertical aircraft speed. These are estimated from pressure sensors and aircraft parameters including pitch, roll and heading. MMS vertical wind data sometimes exhibits discontinuities when the aircraft switches from one maneuver to another (such as from an ascent to a descent) and must be corrected to minimize artifacts caused by this behavior. Following recommendations from Drs. Jonathan Dean-day and Rei Ueyama from NASA, aircraft data is separated into flight legs or maneuvers, which are stretches of time when the airplane has a near constant attack angle, and the vertical wind along each flight leg is demeaned and detrended.

Demeaning and detrending does not introduce significant biases for sufficiently long flight legs because, in the TTL, vertical winds averaged over regions of similar size to a flight leg are typically on the order of a cm s^{-1}. This estimate is based on analyzing vertical winds in the TTL from ERA5 reanalysis averaged over 1° x 1°, or about 100 km x 100 km boxes (not shown). While ERA5 should probably not be trusted in detail for such a purpose, here it is only used to make an order-of-magnitude estimate. We found that 99.9% of these boxes have mean vertical wind with magnitudes less than 10 cm s^{-1}, and 62.5% have mean vertical wind with magnitudes less than 1 cm s^{-1}. The small-scale vertical winds that are the focus of this study have magnitudes of at least 25 cm s^{-1}. Additionally, in the rare event that the mean wind exceeds 10 cm s^{-1} over a 100 km segment of the atmosphere, that mean wind is unlikely to be measured well by the aircraft. Flight legs that are less than 100 km long are removed as demeaning and detrending those legs could remove some small-scale variability and bias our results.

Figure A1 in Appendix A shows corrected and uncorrected vertical wind data for an example flight, demonstrating both the biases in the uncorrected data and the effectiveness of our data correction procedure in removing them. Correcting the MMS data is necessary for constraining the magnitude of the vertical wind, but high frequency (> 1 Hz) vertical wind variance is well constrained in the uncorrected data because the biases in vertical wind have frequencies lower than 1 Hz.

Throughout this study, we use the vertical wind variance as a proxy for turbulence (Atlas et al., 2020). Many studies (Gultepe and Starr, 1995; Muhlbauer et al., 2014; Podglajen et al., 2017, e.g.) have used an estimate of the turbulent eddy dissipation rate (ϵ) instead of vertical wind variance to identify and quantify turbulence. ϵ must be computed from a power spectrum of the aircraft vertical wind. The estimate of ϵ is sensitive to the algorithm used to make the power spectrum (e.g. Fourier decomposition, wavelet analysis), how ϵ is estimated from the power spectrum (e.g. fitting a line, integrating the power spectrum), and the assumed Kolmogorov constant. Vertical wind variance is straightforward to calculate and conveys the same information. Within the inertial range of turbulence, it is proportional to $\epsilon^{2/3}$, with a constant of proportionality that is dependent on the aircraft speed and the sampling frequency (Gultepe and Starr, 1995). Figure A1d-e shows the vertical wind variance and the reported values of ϵ for an example flight, to show that the same turbulent patches are clearly evident in both metrics.

2.2 Aircraft measurements of TTL cirrus microphysics

For all campaigns analyzed, ice water content (IWC) is computed as a difference between total water and water vapor. The Cloud Laser Hygrometer (Davis et al., 2007, CLH) measured total water, and a combination of the Harvard Lyman-α (Weinstock et al., 1994) and the JPL laser hygrometer (May, 1998) measured water vapor during CRYSTAL-FACE and TC4. The
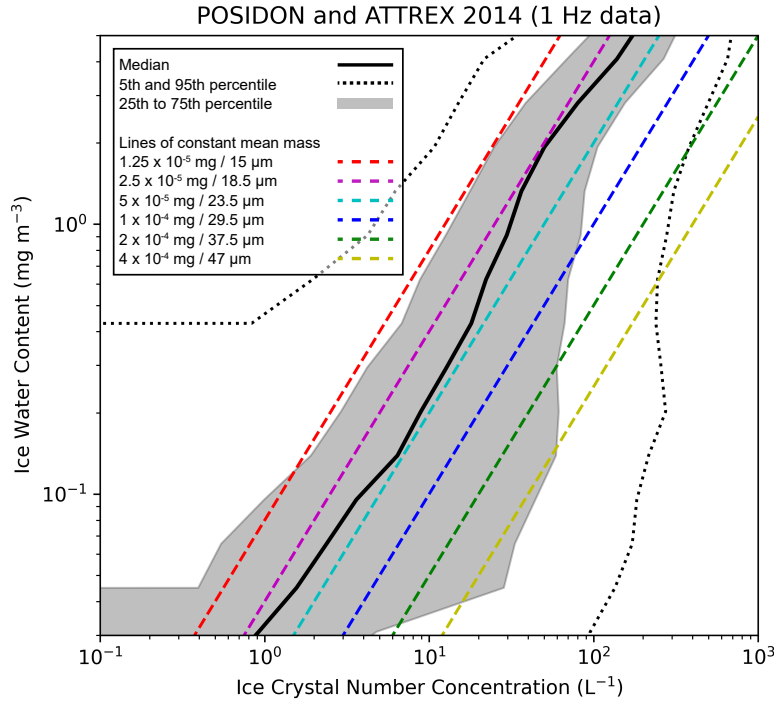
Figure 2. Statistics of ice water content (IWC) are shown as a function of crystal number concentration (NI) for 1 Hz data from POSIDON and ATTREX 2014. The solid line is the median, the dotted lines are the 5th and 95 percentiles, and the shaded area spans the range of the 25th to 75th percentiles. Colored dashed lines indicate constant mean particle mass. The legend lists the mean masses and the corresponding mean equivalent sphere radii for each colored line.

140   reported IWC has a detection limit of about 10^{-1} mg m^{-3}. The NOAA water instrument (Thornberry et al., 2015) measured both total water and water vapor during POSIDON, and ATTREX 2013 and 2014. The reported IWC has a detection limit close to 2×10^{-3} mg m^{-3}. Nonzero IWCs that are below the detection limit are uncertain.

    IWC and NI are tightly linked in TTL cirrus, which tend to have ice size distributions that are more similar than their ice crystal number concentrations. Figure 2a shows percentiles of IWC binned by NI from POSIDON and ATTREX 2014. NI

145   was measured by the Fast Cloud Droplet Probe (Lance et al., 2010, FCDP, for particles 3-24 μm) and the Two Dimensional Stereo Probe (Lawson et al., 2006, 2D-S, for particles 25-3005 μm) for both campaigns. The detection limits for the FCDP and the 2D-S, for an aircraft speed of 100 m s^{-1} and assuming sample areas of 0.03 mm^2 and 80 mm^2, respectively, are about 33 L^{-1} s^{-1} and 0.1 L^{-1} s^{-1}. Median IWC varies nearly linearly with median NI, and the median ice crystal size, assuming single-sized spherical ice particles, varies between about 19 μm (magenta dashed line) and 24 μm (cyan dashed line).

150   IWC and NI have been quality checked and compiled into a single dataset (Krämer et al., 2020a) as described in Krämer et al. (2020b). Since ATTREX 2013 sampled mainly clear sky, it was not included in Krämer et al. (2020b)'s analysis. We still use IWC from this campaign but we note that it has not been subjected to the same quality control as the other campaigns.

2.2.1 Distance from convection

Deep convection influences TTL dynamics by generating gravity waves, so it is useful to look at vertical wind variability as a function of the distance to deep convective cores. We estimate the minimum distance to a deep convective core for each 1 Hz sample of aircraft data using brightness temperature from NCEP/CPC's Merged IR product (Janowiak et al., 2017, MERGIR). This product has 5 km spatial resolution and 30-minute temporal resolution (with output on the hour and half-hour). We only use the data on the half-hour because there is frequently missing data on the hour. We define convective cores as having brightness temperature below 210 K (Gasparini et al., 2021). For every second of aircraft data, we take the snapshot from MERGIR that is closest in time to the aircraft data, and we compute the distance from the aircraft location to the nearest deep convective core in that snapshot, as illustrated in Figure A2. Other studies have used different brightness temperature thresholds to identify deep convection. Podglajen et al. (2017) used 235 K and Wall et al. (2020) used a stricter 200 K. We compare these different thresholds in Appendix A and Figure A2. There we find that a 200 K threshold misses most convective cores entirely, whereas a 235 K threshold includes some anvil cirrus and aging deep convective cores which are less likely to generate gravity waves.

3 Results

3.1 Small-scale vertical wind variability in all campaigns

We analyze the 1 Hz vertical wind ($\overline{w}_1$) in this section and high frequency (> 1 Hz) vertical wind variance ($\sigma^2 w_1$) in the next section, for all five flight campaigns separately. $\overline{w}_1$ is sensitive to both gravity wave activity and turbulence, whereas $\sigma^2 w_1$ is sensitive mainly to turbulence.

Throughout these two sections, we split the data up into categories based on IWC, distance to deep convection cores, and height. Figure B1 shows distributions of these variables for the five flight campaigns and Appendix B discusses our choice of categories. We split the IWC into three categories: clear sky (IWC = 0.0), low-IWC cirrus (between 0.0 and 1 mg m^{-3}), and high-IWC cirrus (IWC > 1 mg m^{-3}). Clear sky and low-IWC cirrus cannot be perfectly discriminated by the measurements, particularly for CRYSTAL-FACE and TC4, which have an IWC detection limit of 0.1 mg m^{-3}.

Figures 3 shows probability distributions of $\overline{w}_1$ for each campaign separately. The first column shows the distribution for all campaign data with the campaign name and the number of 1 Hz samples included in the analysis printed over the plots. The second, third and fourth columns show probability distributions of $\overline{w}_1$ split into categories based on the IWC, distance to deep convective cores, and height in the TTL, respectively, with pie charts showing the distribution of data across the different categories. Normal distributions are fitted to all probability distributions and the fitted standard deviations (σ) are overlaid.

Values of σ vary between 0.17 and 0.4 m s^{-1} for all campaign data (first column of Figure 3). Differences in vertical wind variability across the set of campaigns may arise from sampling closer to or farther from deep convection, the properties of the deep convection (e.g., land vs ocean, shallow vs deep), inter-annual, seasonal and geographic variability in the upper troposphere and TTL including the QBO phase, and sampling different heights within the TTL. CRYSTAL-FACE, which
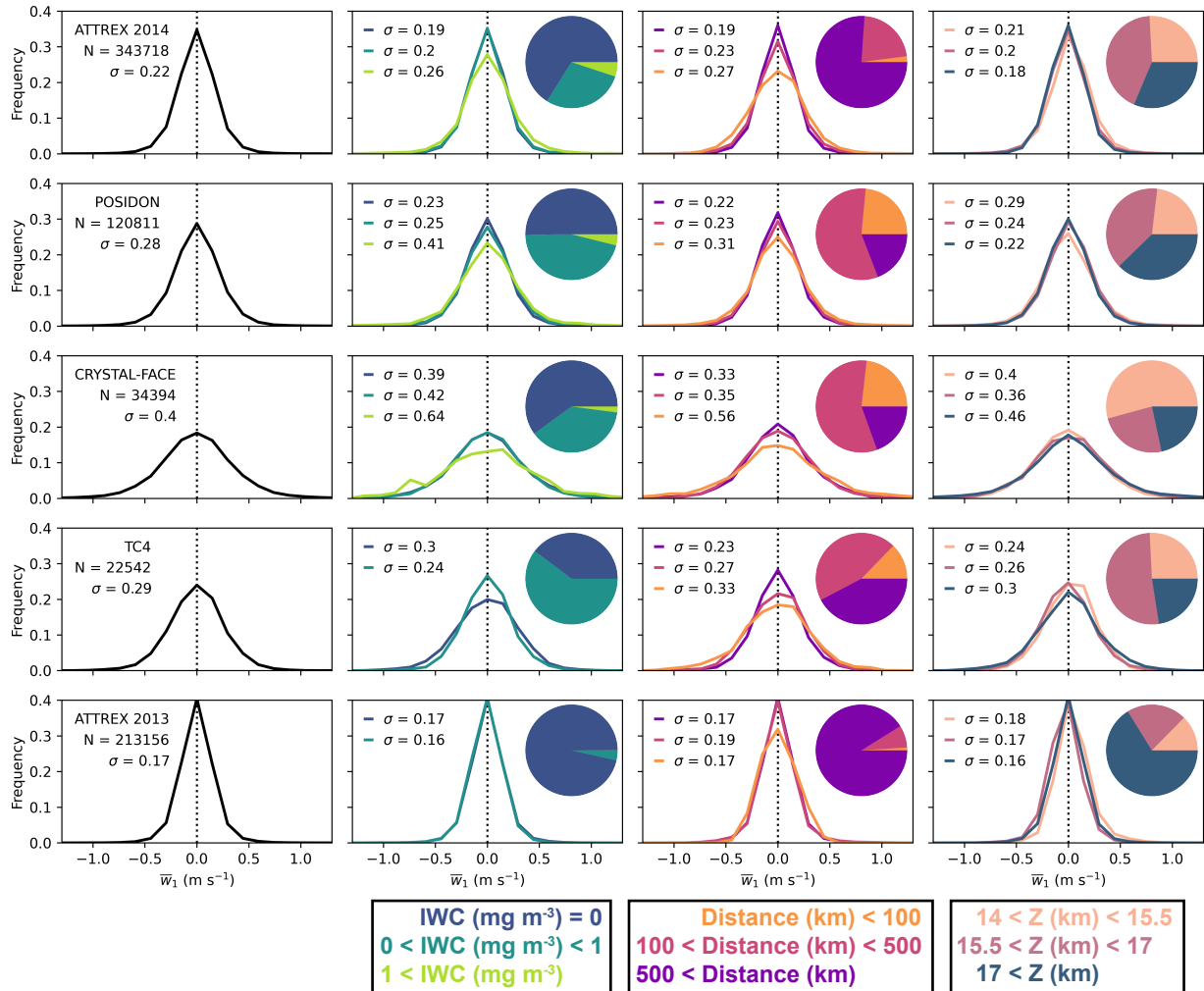
Figure 3. Distributions of 1 Hz mean vertical wind ($\overline{w}_1$) for corrected data from all campaigns, shown in separate rows. Campaign names and the number of 1 Hz samples are given in the leftmost column. Second, third and fourth columns split the data into categories (relative frequencies are shown in pie charts) based on IWC, distance from deep convection, and height in the TTL, respectively. Standard deviations of fitted normal distributions (σ) are printed on all panels.

185 sampled near convection in the Florida region, has the widest distribution of $\overline{w}_1$ (the most vertical wind variability). ATTREX 2013, which sampled the tropical East and Central Pacific, usually far from deep convection, has the least variability. Because the five campaigns sampled such different conditions, it is plausible that they approximately span the vertical wind variability to be expected anywhere in the TTL outside of the immediate vicinity of deep convection.

The remaining columns of Figure 3 partition the data from each experiment into categories of IWC, distance from deep
190 convection, and height, and compare the distributions of $\overline{w}_1$ across these categories. Within ATTREX 2014, POSIDON and CRYSTAL-FACE data, high-IWC cirrus have wider $\overline{w}_1$ distributions (top three rows of second column of Figure 3) and this result is insensitive to the threshold used to define high-IWC cirrus (not shown). In general, vertical wind variability increases closer to deep convection (third column of Figure 3). An exception is ATTREX 2013, where vertical wind variability is low regardless of the distance from deep convection. The first column of Figure B2 shows the same analysis for clear sky data
195 only; the results are the same. Thus, the increase in vertical wind variability close to deep convection is not caused by a higher occurrence of cirrus cloud close to deep convection. There is no consistent relationship between vertical wind variability and height across the five campaigns, or between the analysis of all data (Figure 3) and clear-sky data (Figure B2).

3.2 Turbulence in all campaigns

Figure 4 shows distributions of high frequency (> 1 Hz) vertical wind variance ($\sigma^2 w_1$). $\sigma^2 w_1$ is the variance of the 20 sub-
200 samples within each second of data, and a proxy for turbulence. The distributions are split into the same categories based on IWC, distance from deep convection, and height, as in Figure 3. We define $\sigma^2 w_1$ greater than an empirical detectability threshold of 0.01 m^2 s^{-2} as turbulent. This turbulence threshold is typically lower than the one used by Podglajen et al. (2017), who analyzed strong turbulence rather than just detectable turbulence. This can be seen be comparing the yellow shaded regions of the flight tracks shown in Figure A1d and e. Our methodology therefore flags turbulence with a much
205 higher occurrence frequency. To facilitate comparison with Podglajen et al. (2017) and test the robustness of our results to the turbulence threshold used, Figure B3 shows a version of Figure 4 but with distributions of ϵ instead of $\sigma^2 w_1$, and with a turbulence threshold for ϵ of 10^{-3} m^2 s^{-3}. The sensitivity of turbulence to environmental categories is qualitatively similar to Figure 4, but the implied turbulence frequency is an order of magnitude smaller. We would need to increase our threshold high-frequency vertical velocity variance to over 0.01 m^2 s^{-2} (not shown) to obtain comparable results. We choose our comparably
210 weak detectability threshold because we aim to associate all vertical wind anomalies exceeding 25 cm s^{-1} with a particular type of small-scale vertical wind variability. Using a strong turbulence threshold would result in many of these periods going undetected. Additionally, the infrequency of strong turbulence would result in poorer sampling statistics.

Using our variance-based algorithm, the frequency of turbulence varies between about 6% and 12% for all campaign data (first column of Figure 4). ATTREX 2014 and POSIDON, which both sampled the TTL mainly above the tropical West
215 Pacific, have about twice as frequent turbulence as CRYSTAL-FACE and TC4, even though the latter campaigns had more vertical wind variability. Thus, turbulence accounts for a larger proportion of the vertical variability sampled during ATTREX 2014 and POSIDON than during TC4 and CRYSTAL-FACE.
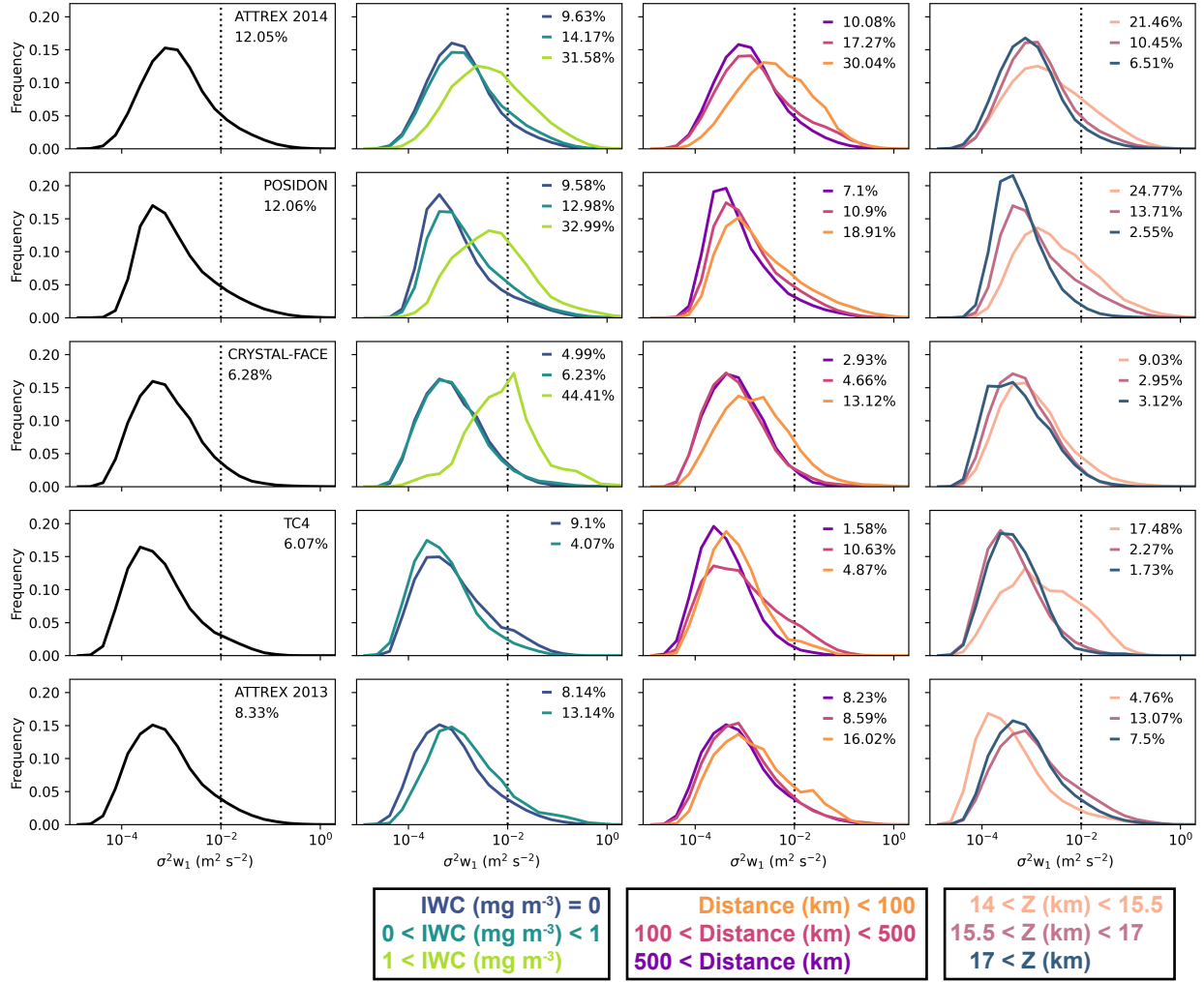
Figure 4. Same as Figure 3 except showing distributions of high frequency (> 1 Hz) vertical wind variance ($\sigma^2 w_1$). The percentage of data exceeding the turbulent threshold (0.01 $m^2 s^{-2}$, vertical dotted line) for each category is printed on each plot.

The frequency of turbulence for clear sky data (dark blue line, second column of Figure 4) varies between about 5% and 10% and is typically just slightly lower than the all-sky frequency, implying that clear-air turbulence accounts for much of the turbulence in the TTL. The frequency of turbulence for high-IWC cirrus is at least three times that for clear-sky data (second column of Figure 4). High-IWC cirrus have more turbulence than clear-sky data independently of the threshold used for turbulence and the threshold used to define high-IWC cirrus (not shown).

Increased turbulence can account for much of the widening of the vertical wind distribution within high-IWC cirrus seen in Figure 3 for ATTREX 2014, POSIDON and CRYSTAL-FACE.

In all campaigns, turbulence is least frequent further than 500 km away from convection (third column of Figure 4), consistent with the findings of Podglajen et al. (2017). In all campaigns except TC4, turbulence is most frequent within 100 km of convection. Column 3 in Figure B2 shows that the same is true for clear-sky data. However, the differences in the frequency of turbulence between the two other categories (100-500 km and > 500 km) are largely diminished in the clear-sky data. If a stricter threshold is used for turbulence, the results are the same except for ATTREX 2013, where the differences in the amount of turbulence within 100 km of convection, and between 100 and 500 km, are diminished (not shown).

Turbulence is enhanced below 15.5 km in all campaigns except ATTREX 2013, which has enhanced turbulence between 15.5 and 17 km. Podglajen et al. (2017) found that turbulence was enhanced below 15.5 km mainly within 500 km of deep convection (Figure 9 in their paper). Consistent with these results, ATTREX 2013 has the most sampling further than 500 km away from deep convection (Figures 3 and B1). These results are insensitive to the turbulence threshold used.

In this section and in the preceding section, we analyzed aircraft data from the TTL from five NASA aircraft campaigns that took place in different years, seasons, and geographical areas, and sampled a different range of distances from deep convection and heights within and above the TTL. Across all campaigns, the probability distribution of $\overline{w}_1$ is well approximated as a normal distribution with a standard deviation between 0.17 and 0.56 m s^{-1}, depending on the distance from deep convection, the height in the TTL, and the presence of cloud. Vertical wind variability, which is influenced by both turbulence and gravity wave activity, is largest during CRYSTAL-FACE, but the frequency of turbulence is largest during ATTREX 2014 and POSIDON. That means that the increased vertical wind variability during CRYSTAL-FACE is due to increased gravity wave activity. It is unlikely that these differences are purely due to different sampling strategies across the campaigns, and we encourage future studies to investigate the causes of geographical differences in the frequency of turbulence and small-scale gravity wave activity.

We verified that several findings reported in Podglajen et al. (2017) about the frequency of turbulence in ATTREX 2013 and 2014 are true across our entire set of campaigns, including that turbulence is enhanced closer to deep convection, below 15.5 km in the TTL when close to deep convection, and above 15.5 km when far away from deep convection. Furthermore, we analyzed turbulence in clear-sky and cloudy data separately and found that turbulence is strongly enhanced within high-IWC TTL cirrus.

3.3 Sources of small-scale vertical wind variability during ATTREX

Level (constant-altitude) flight legs are useful for separately detecting turbulence and gravity wave activity and for performing spectral analyses, because they don't conflate horizontal and vertical scales of variability. Gravity waves often have smaller vertical wavelengths than horizontal wavelengths (Bramberger et al., 2022) so the scale separation between gravity wave activity and turbulence is more pronounced in the horizontal.

ATTREX 2014's flight strategy involved repeatedly flying parallel to the ground at a height of about 14.2 km, typically through cloud, performing a slow ascent, and a quick descent (Figure A1). ATTREX 2013 sometimes employed this flight strategy and sometimes performed ascents and descents with no level legs in between. There are 52 level legs that are at least 100 km long in the ATTREX 2014 dataset, and 13 in the ATTREX 2013 dataset, which are the focus of the rest of this study. In this section, we present an algorithm to distinguish between turbulence and gravity wave activity, and we investigate how the presence of turbulence and gravity wave activity varies with distance to deep convection, IWC and NI. In the following section, we perform spectral analyses of vertical wind on level leg data from ATTREX 2014 and simulated vertical winds from four global storm-resolving models from the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains summer experiment (DYAMOND-1) (Stevens et al., 2019), to evaluate simulated small-scale vertical wind variability over the tropical West Pacific.

Gravity wave activity occurs on a wide range of horizontal scales from 1 km to 1000s of km. We define long wavelength gravity wave activity (LW GWA) as having dominant wavelengths between 5 and 100 km and short wavelength gravity wave activity (SW GWA) as having smaller wavelengths. The purpose of making this distinction is that SW GWA has horizontal length scales that overlap with turbulence, whereas LW GWA does not.

We separate the level legs from ATTREX into non-overlapping 25-second or 5 km segments (the aircraft travels at about 200 m s^{-1} during level legs). LW GWA occurs on spatial scales larger than 5 km, so we classify each entire level leg into one of two categories: 1. negligible LW GWA and 2. LW GWA. All 5 km segments within a level leg receive the same classification.

Turbulence and SW GWA occur on scales smaller than 5 km so we classify each 5 km segment into one of three different categories: 1. negligible sub-5 km variability, 2. turbulence, and 3. SW GWA. Because SW GWA and turbulence occur on similar spatial scales, we can only detect turbulence in the absence of SW GWA. We detail the classification of level legs and 5 km segments in Appendix C.

Figure 5 shows four example time series of vertical wind for four different level legs from the ATTREX 2014 dataset. In the upper time series plot for each example, the 20 Hz vertical wind is shown in grey and the mean vertical wind for the 5 km segments ($\overline{w}_{25}$) is shown in black. In the second row under each example, the 20 Hz vertical wind is color coded according to the sub 5 km variability classifications, and $\overline{w}_{25}$ is color coded according to the LW GWA classification (for the entire level leg). Examples 2 and 3 have LW GWA, whereas examples 1 and 4 have negligible LW GWA. All of the 5 km segments in examples 1 and 2 have negligible sub-5 km variability, whereas in example 3 they are all turbulent. Example 4 has one intermittent patch of turbulence and one patch of SW GWA. These small-scale vertical motions are typically associated with temperature
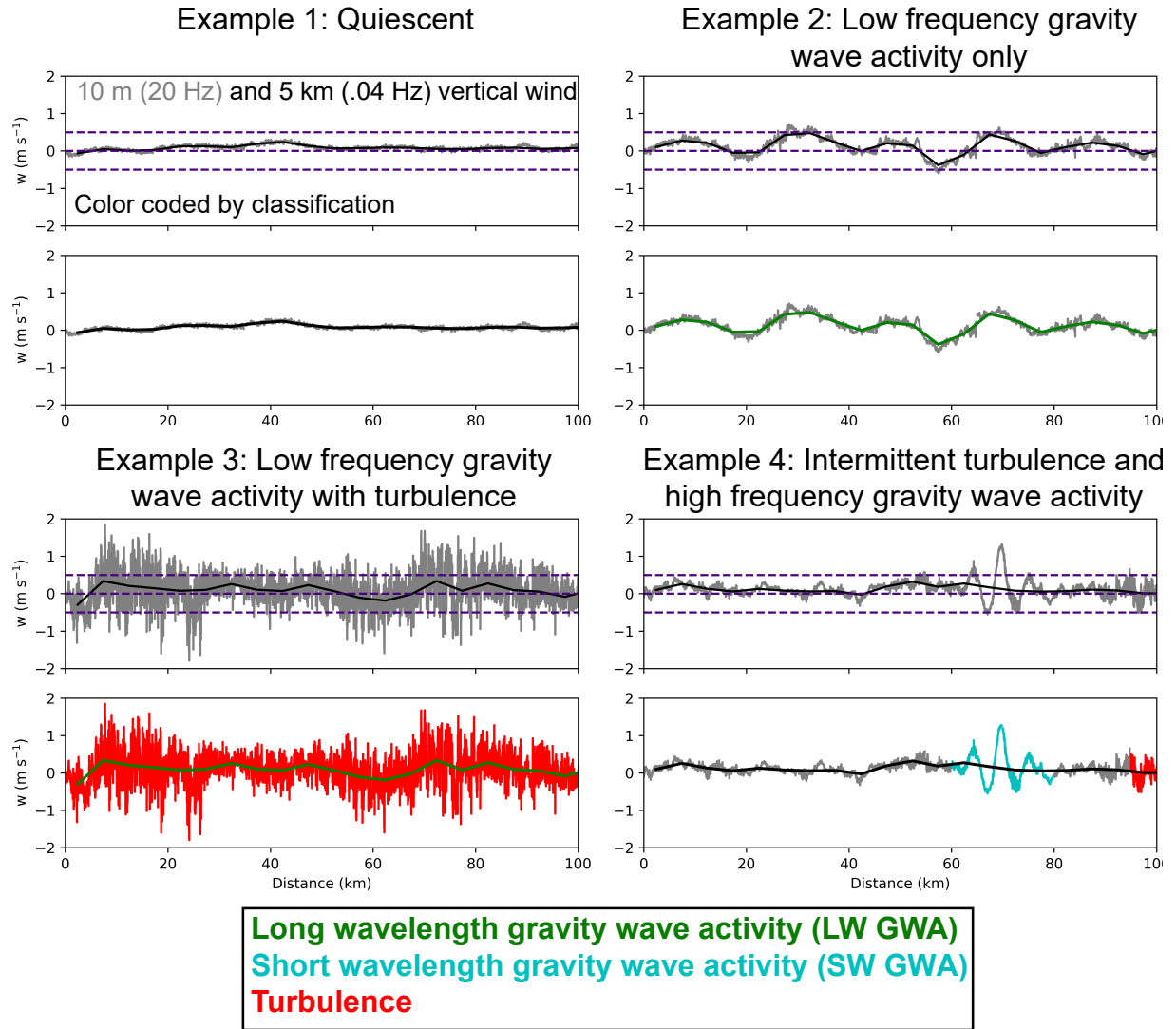
Figure 5. Four example vertical wind time series for 100 km long segments from ATTREX 2014 level legs. All plots show the time series of high-rate (20 Hz/10m) vertical wind and 5 km mean vertical wind ($\overline{w}_{25}$). In the top time series for each example, high-rate vertical wind is in grey and $\overline{w}_{25}$ is in black. Horizontal dashed purple lines indicate 0.0, 0.5 and -0.5 m s^{-1}. In the bottom time series for each example, $\overline{w}_{25}$ is colored green if there is long wavelength gravity wave activity, and the high-rate vertical wind is colored cyan and red, where there is short wavelength gravity wave activity and turbulence, respectively.

anomalies of up to 0.5 K for SW GWA and 0.1-0.2 K for turbulence (not shown; for an example, see Figure 5 of Podglajen et al. (2017)).

285   Each 5 km segment can have three possible sub 5 km classifications, with or without LW GWA. We show the frequency of each combination of classifications in Figure 6a. Hatching indicates LW GWA, and the color represents the sub 5 km variability classification. Three situations are most common: 1. negligible LW GWA with negligible sub 5 km variability (quiescent), 2. LW GWA with negligible sub 5 km variability (LW GWA only), and 3. LW GWA with turbulence. These situations occur 24%, 50% and 22% of the time, respectively. LW GWA is present 75% of the time. Turbulence occurs preferentially with LW GWA,

290 such that only 5% of turbulent segments (11% for clear-sky data) do not have LW GWA (Figure 6a). SW GWA also occurs preferentially with LW GWA; only 8% of segments with SW GWA do not also have LW GWA (these segments are represented by the thin cyan sliver between the areas of the pie chart labelled as quiescent and as LW GWA with turbulence).

    Figure 6b shows histograms of 5 km mean (from top to bottom) NI, IWC, and distance to deep convection. The extra column on the left side of the histograms for NI and IWC corresponds to NI=0 and IWC=0, respectively, and the black lines

295 show thresholds defining high NI ($> 20\,\mathrm{L}^{-1}$) and high IWC ($> 1\,\mathrm{mg\,m}^{-3}$), also used in the category definitions of Figure 6c. Segments classified as LW GWA with turbulence are most common and segments classified as quiescent are rarely seen within high-NI and high-IWC cirrus, or within 500 km of convection (to the left of the black line on the histogram of the distance to deep convection).

    Figure 6c shows pie charts of 5 km mean NI, IWC, and distance from deep convection for the three most common conditions

300 from Figure 6a, with categories used to simplify the analysis. There are fewer total samples in the pie charts for the distance to deep convection than there are in the pie charts for IWC and NI because brightness temperature data is occasionally missing over the tropical West Pacific.

    The likelihoods of occurrence of high-NI cirrus given quiescent conditions, LW GWA only and LW GWA with turbulence are 0.9%, 16% and 40%, respectively. Thus, high-NI cirrus are about 20 times more likely when there is gravity wave activity and

305 50 times more likely when there is also turbulence, compared to in quiescent conditions. The likelihoods of occurrence of high-IWC cirrus given quiescent conditions, LW GWA only and LW GWA with turbulence are 0.7%, 14% and 33%, respectively. Thus, High-IWC cirrus are also about 20 times more likely when there is gravity wave activity and about 50 times more likely when there is also turbulence, compared to in quiescent conditions. High-NI and high-IWC cirrus co-occur with low frequency gravity wave activity 99% of the time.

310   The fractions of 5 km segments within 500 km of convection with quiescent conditions, LW GWA only and LW GWA with turbulence are 5%, 15% and 47%, respectively. Thus, turbulence is enhanced closer to deep convection to a much greater extent than LW GWA.

    Turbulence co-occurs with high-NI cirrus 41% of the time, and with high-IWC cirrus 32% of the time (Figure 6c). However, turbulence is not a necessary condition for high-NI and high-IWC cirrus, as they occur in the absence of turbulence about half

315 of the time. In general, the presence of turbulence is much more highly correlated with the presence of LW GWA than with the presence of high-NI and high-IWC cirrus.
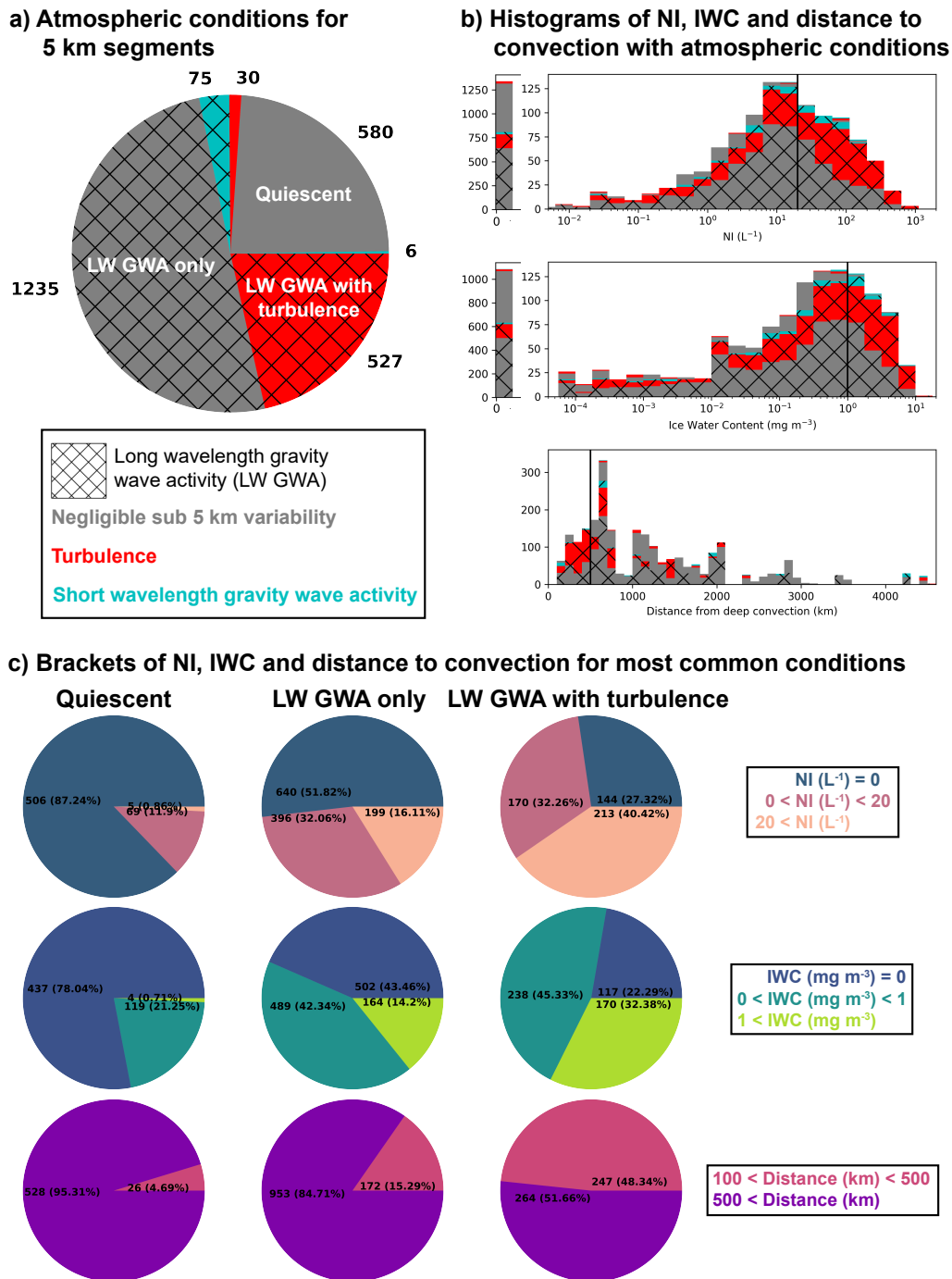
Figure 6. Analysis of 5 km segments from ATTREX 2013 and 2014 level legs: a) Frequency of atmospheric conditions. The most common conditions are labeled. b) Histograms of (top to bottom) NI, IWC, and distance from deep convection, showing the contributions from different atmospheric conditions. Black lines indicate thresholds also used below in c. c) Pie charts showing the relative frequency of categories of (top to bottom) NI, IWC and distance to deep convection for the three most frequent atmospheric conditions.

3.4 Evaluating vertical wind variability in global storm-resolving models

Recently, advances in computing power have made it possible to run global atmospheric models with horizontal grid spacing below 5 km. These models are referred to as global storm-resolving models (GSRMs) because they explicitly resolve deep convection rather than using a deep convective parameterization. Since deep convection is a major source of both gravity waves and the water vapor and ice that form TTL cirrus, this makes GSRMs attractive for studying TTL cloud formation processes, including lifting within gravity waves. However, GSRMs do not resolve turbulence, which they typically parameterize in some form. In this section we use ATTREX observations to address whether TTL vertical winds simulated by GSRMs are sufficiently realistic to form a credible dynamical environment for TTL cirrus formation and evolution.

Stephan et al. (2019) found that explicit convection simulated with a horizontal grid spacing of 5 km produces more gravity wave momentum flux at 30 km in the tropics and subtropics, where convection is the predominant source of gravity wave activity (Fritts and Alexander, 2003, and references therein), and a wider vertical wind distribution throughout the troposphere, compared to parameterized convection. Müller et al. (2018) found that "convective parameterization inhibits gravity wave generation by convective clouds." While these studies suggest gravity wave generation and propagation are more realistic in GSRMs than in coarse resolution models, substantial discrepancies may still exist between GSRMs and the real atmosphere.

In the DYAMOND-1 intercomparison, nine GSRMs were identically initialized from reanalysis and run freely (without nudging) for the 40-day period of 1 August–10 September 2016. Here, we focus on four of those models: Nonhydrostatic ICosahedral Atmospheric Model (NICAM), Global System for Atmospheric Modeling (gSAM), Finite-Volume Cubed-Sphere Dynamical Core (FV3) and ICON (Icosahedral Nonhydrostatic Weather and Climate Model). NICAM, gSAM, FV3 ans ICON have horizontal grid spacing of 3.25 km, 5 km, 3.25 km and 2.5 km, respectively. The vertical grid spacing in the TTL is 400 m for NICAM and close to 500 m for the other three models.

Figure 7 shows snapshots of vertical wind from the 141st hour of simulation at the vertical levels closest to 14.2 km, which is the same level as the ATTREX horizontal aircraft legs. Figure S1 is an animation of vertical wind snapshots for hours 48-957 of the simulations (allowing two days for model spinup), showing that the snapshot in Figure 7 is representative of the simulations. The models differ substantially in their magnitudes of vertical winds and the dominant scales of vertical wind variability. gSAM is an outlier in having more vertical wind variability than the other models, particularly at the grid scale. ICON has more grid scale variability than FV3 and NICAM. NICAM has larger dominant scales of horizontal variability than the other models, and FV3 has weak vertical winds in most grid cells but strong gravity wave activity in the vicinity of deep convection.

Figure 8b shows statistics of power spectra of 20 Hz vertical wind from ATTREX 2014 level legs sampling the tropical west Pacific and from the four GSRMs, for the same regions and vertical levels shown in Figure 7. We perform 1-D fast Fourier transforms (FFTs) for both simulations and observations. FFTs are performed separately along the east-west and north-south direction in the GSRMs and then averaged together. The spectra are similar for both directions (not shown). ICON has an unstructured grid. To enable 1-D FFT analysis, it is interpolated to an 0.01° lat-lon grid. For the observations only, we use
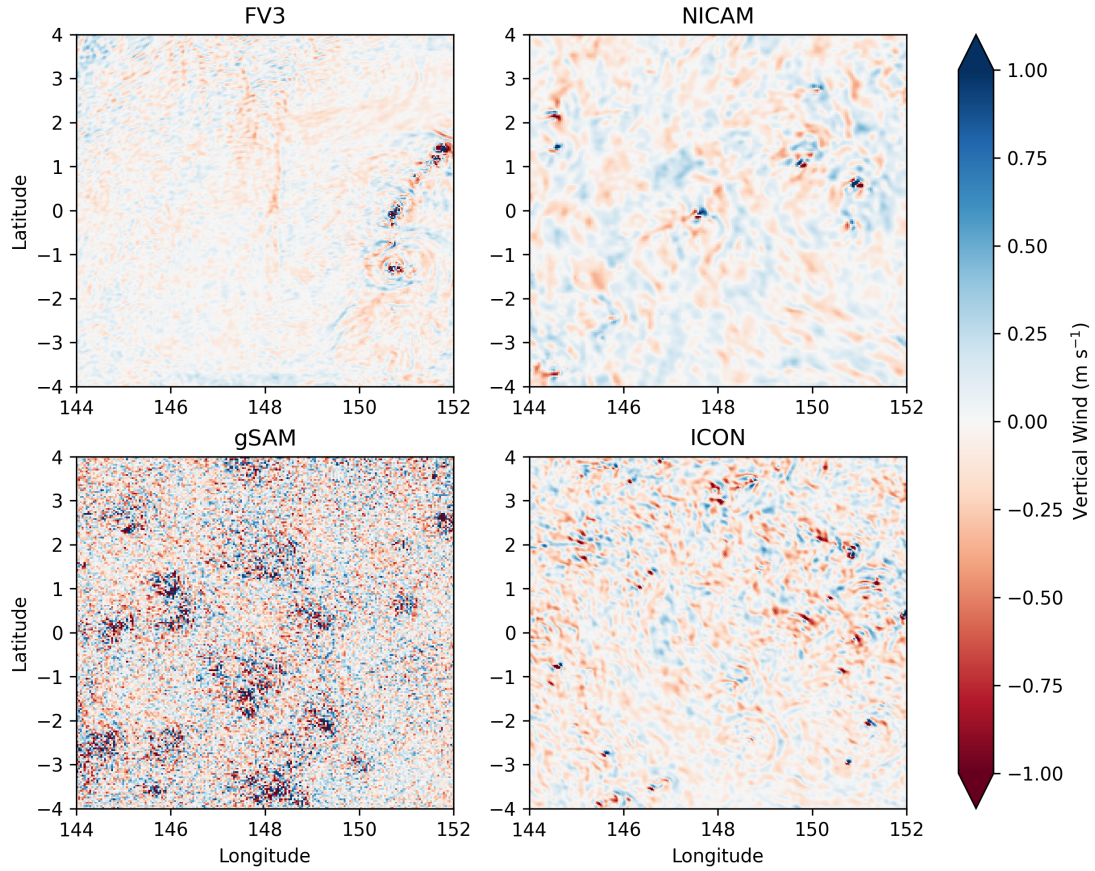
Figure 7. Snapshots of vertical wind from hour 141 in the DYAMOND-1 simulations. Winds are taken from the model level closest to 14 km and the tropical West Pacific region is shown.
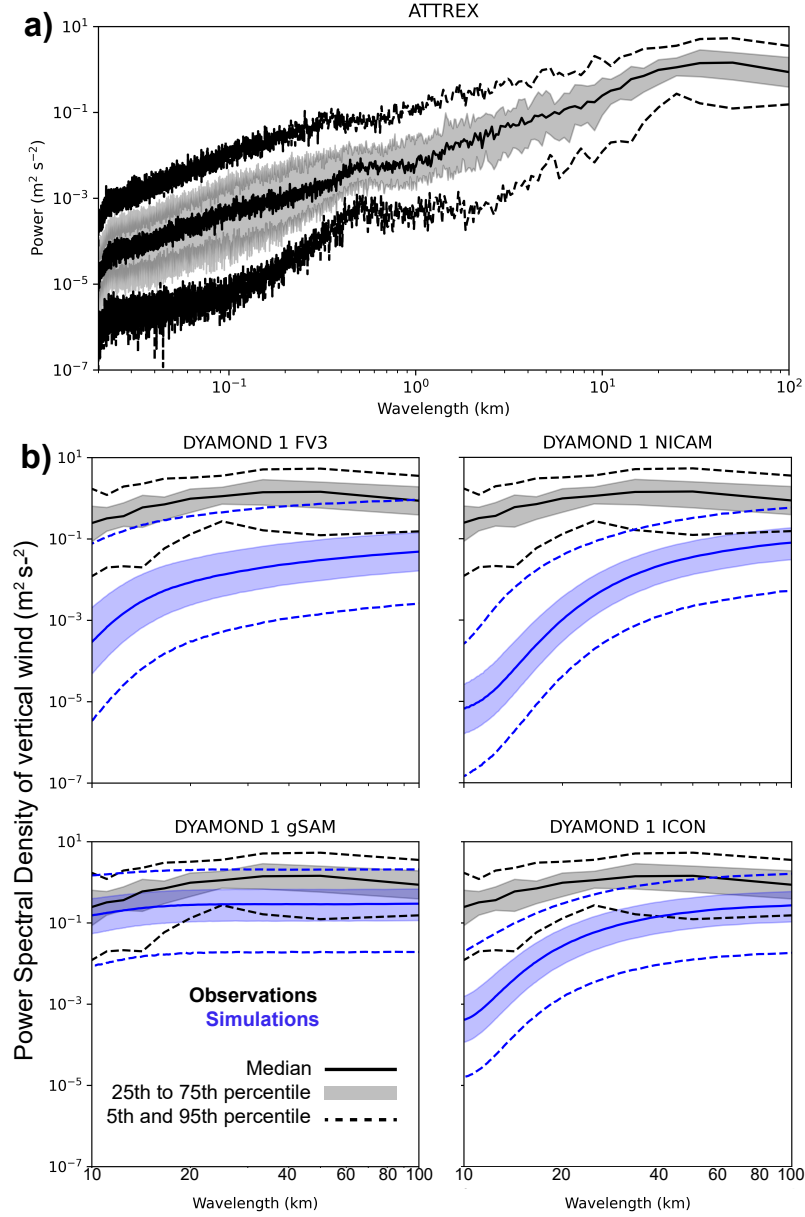
Figure 8. a) Statistics of 1-D power spectral density for the ATTREX 2014 observations over the tropical West Pacific for all measured wavelengths. The increase in power at ~ 500 m is caused by the oscillation of the aircraft. b) Statistics of 1-D power spectral density for the ATTREX 2014 observations over the tropical West Pacific (black) and the DYAMOND-1 simulations (blue) for wavelengths between 10 and 100 km. Solid lines are medians, dotted lines are 5th and 95 percentiles, and shaded areas span the ranges of the 25th to 75th percentiles.

Welch's method for windowed Fourier analysis with a window length of 100 km and a Hann taper to reduce noise in the spectra.

All simulations have too little power in vertical wind at all wavelengths examined. The power spectra for ICON, FV3 and NICAM decrease strongly towards the grid scale, which is where they are furthest from the observed spectra. gSAM has a flatter spectrum which more closely resembles the observations. However, the grid scale variability in gSAM has more resemblance to white noise than to gravity wave activity, so it may not be physically meaningful for this analysis.

We do not look at relationships between cirrus properties and vertical wind here, because none of the models predict NI, and while all models predict IWC, FV3 outputs it on a different grid than the vertical wind, making comparisons difficult. Nonetheless, our observational analysis showed that small-scale gravity wave activity is a strong control on TTL cirrus microphysics, and our comparison here shows that small-scale gravity wave activity is underrepresented in the evaluated GSRMs. If the GSRMs have enough or too much ice within simulated TTL cirrus compared to the real atmosphere, that suggests that over-production of ice within the microphysics schemes is compensating for deficient dynamical contributions to ice production. In this case, the physical mechanisms controlling ice production in simulated TTL cirrus would differ substantially from the real atmosphere.

4   Conclusions

Tropical tropopause layer (TTL) cirrus can be influenced by small-scale vertical motions in the TTL from gravity wave activity and turbulence. The relationships between these phenomena are analyzed using high-rate vertical wind data collected by NASA flight campaigns.

Out of the five campaigns we analyzed, vertical wind variability was largest during CRYSTAL-FACE and TC4 although those campaigns had the lowest frequencies of turbulence, indicating that gravity wave activity was an important source of variability.

Turbulence during ATTREX 2013 and 2014 was analyzed in detail in Podglajen et al. (2017), and we find that some key results from that study hold true across the five campaigns we analyzed: (1) that turbulence is enhanced over the tropical West Pacific and nearer to deep convection, and (2) that turbulence is most frequent in the lower TTL (14-15.5 km) close to deep convection and in the upper TTL (15.5-17 km) further from deep convection.

For the first time, we used aircraft measurements to correlate gravity wave activity and turbulence with TTL cirrus microphysical properties. During ATTREX 2014, 99% of 5 km segments with high ice water content (IWC > 1 mg m^{-3}) and high ice crystal number concentrations (NI > 20 L^{-1}) co-occurred with long wavelength gravity wave activity, and half of those segments co-occurred with turbulence as well. Thus, small-scale vertical motions driven by turbulence and gravity wave activity are key to producing thicker cirrus that contain more ice crystals over the tropical West Pacific.

A strong relationship between small-scale vertical wind variability and TTL cirrus microphysics had been proposed in several modelling studies, but here we present the first observational evidence for it. Each of the modelling studies focused on either

gravity wave activity or turbulence. Here, we show that both sources of small-scale vertical wind variability are important, and that they frequently occur together.

The common co-occurrence of thicker cirrus and turbulence can be explained in two ways: (1) thicker clouds initiate cloud-driven turbulence and (2) clear-air turbulence forms cirrus and/or thickens pre-existing cirrus. These explanations are not mutually exclusive. Our analysis cannot rule out the possibility that cloud-driven turbulence occurs in the TTL, but there are several clues from our study and from Podglajen et al. (2017) that clear-sky sources of turbulence are dominant.

Podglajen et al. (2017) found that turbulent patches were correlated with low Richardson number, indicating the presence of shear. Shear can cause two types of clear-air turbulence: gravity-wave breaking, by creating a critical level, and Kelvin-Helmholtz instabilities. Additionally, gravity wave breaking can create or strengthen shear layers (Dörnbrack, 1998). Here, we found that clear-air turbulence was common in all five flight campaigns analyzed, and was enhanced closer to deep convection, which also is a source of gravity wave activity and layers of locally enhanced vertical wind shear. During ATTREX, turbulence co-occurred with gravity wave activity 95% of the time, and thicker cirrus only 30-40% of the time. We encourage future studies to more closely examine turbulent sources in the TTL.

Another potential source of vertical wind variability that we were not able to examine within this study is cloud-driven mesoscale circulations. Cloud-driven mesoscale circulations and cloud-driven turbulence are both induced by thermal instabilities in cirrus, but they produce vertical motions on different scales. Several modelling studies have suggested that cloud-driven mesoscale circulations can maintain cirrus clouds (Dobbie and Jonas, 2001; Dinh et al., 2010; Jensen et al., 2011), whereas at least one other has found that they cannot (Boehm et al., 1999). However, none of these modelling studies included gravity wave activity. We encourage future modelling studies to analyze the development and influence of cloud-driven mesoscale circulations and turbulence in the presence of realistic gravity wave activity.

We also compared TTL vertical wind variability simulated by global storm-resolving models (GSRMs) in the lower TTL over the tropical West Pacific with ATTREX 2014 data. The four models we evaluated (gSAM, ICON, FV3 and NICAM) had drastically different magnitudes of vertical wind and scales of vertical wind variability. Only gSAM had variability at wavelengths shorter than 100 km comparable to the observations. Thus, GSRMs underestimate the vertical winds that affect TTL cirrus clouds, with potential impacts on their simulated microphysics.

Many aspects of the model dynamics and the experimental setup may affect gravity wave formation and propagation, but horizontal and vertical resolution are likely limiting factors. The effective resolution (the minimum length scale that can be resolved) may be 6 times the horizontal grid spacing for GSRMs (Caldwell et al., 2021), meaning that only gravity waves with wavelengths larger than 19.5 to 25 km can be supported in the GSRMs in this study. Additionally, studies have found that a vertical grid spacing of 200 m or finer in the upper troposphere is necessary to adequately handle gravity wave propagation and achieve convergence (Kuang and Bretherton, 2004; Skamarock et al., 2019), but the GSRMs in this study have a vertical resolution at least twice as coarse as that in the TTL. Thus, we encourage future GSRMs or, more practically, regional cloud-resolving model studies to examine the effects of increased vertical and horizontal resolution on small-scale vertical wind variability.

Data availability. High-rate vertical wind measurements for ATTREX 2013-2014 and POSIDON are available on NASA's ESPO archive. Data from CRYSTAL-FACE and TC4 must be requested from T. Paul Bui. Microphysical measurements for all campaigns are available through NASA's ESPO archive and, for all campaigns except ATTREX 2013, at https://doi.org/10.34730/266ca2a41f4946ff97d874bfa458254c. Brightness temperature data is available at https://doi.org/10.5067/P4HZB9N27EKU

420 *Video supplement.* Video S1 cycles through snapshots of vertical wind for every hour in the DYAMOND-1 simulations, except hours 531-549, for which no vertical wind data is available for ICON. Winds are taken from the model level closest to 14.2 km and the tropical West Pacific region is shown.

Appendix A: Preparing the dataset: additional information

A1 Correcting vertical wind data

425 Figure A1 shows flight tracks for an example flight from ATTREX 2014. Figure A1b color codes the flight track with the uncorrected vertical wind. After 55,000 seconds into the flight, a pattern emerges of apparent downdrafts throughout each descent, apparent updrafts in the level leg and the first half of the following ascent, and apparent downdrafts in the second half. This nonphysical behavior indicates that changing aircraft maneuvers (going from an ascent to a descent, for example) are affecting the measured vertical winds. To mitigate this artifact, we demean and detrend each flight leg (each solid color

430 segment in Figure A1a), and we remove legs that cover a horizontal distance smaller than 100 km, to produce the corrected data shown in Figure A1c. The magnitudes of the corrected vertical winds are smaller and less skewed toward negative values.

Figures A1d-e show the high frequency (> 1 Hz) vertical wind variance ($\sigma^2 w_1$) and the 1-second turbulent kinetic energy dissipation rate (ϵ) as reported in the NASA dataset, respectively. ϵ has rare outliers with unrealistically high values outside patches of turbulence, which are mainly seen between 60,000 and 65,000 seconds into the flight (red points). Podglajen et al.

435 (2017) did not use the reported ϵ, so their analysis was not affected by these outliers.

Outliers aside, the same turbulent patches are evident in both ϵ and $\sigma^2 w_1$. Both metrics are similarly useful for identifying turbulence; we chose $\sigma^2 w_1$ based on its ease of calculation and interpretation. In Figure A1d and e, the yellow color indicates data that is identified as turbulent in this study and in Podglajen et al. (2017), respectively. Our turbulence threshold is lower than the one used in Podglajen et al. (2017) so we identify a larger percentage of the data as turbulent.

440 Figure A1 also demonstrates the sampling strategy for ATTREX 2014 and parts of ATTREX 2013, which included many level legs that are the focus of the analyses in Sections 3.3 and 3.4.

A2 Computing distance from convection

Figure A2 demonstrates how we compute the distance to deep convection for the aircraft data, and compares different brightness temperature thresholds for identifying deep convection. We are interested in deep convective cores as sources of gravity wave

445 activity, moisture, and possible detrained ice in the TTL. Only one pixel in all four snapshots has a brightness temperature below
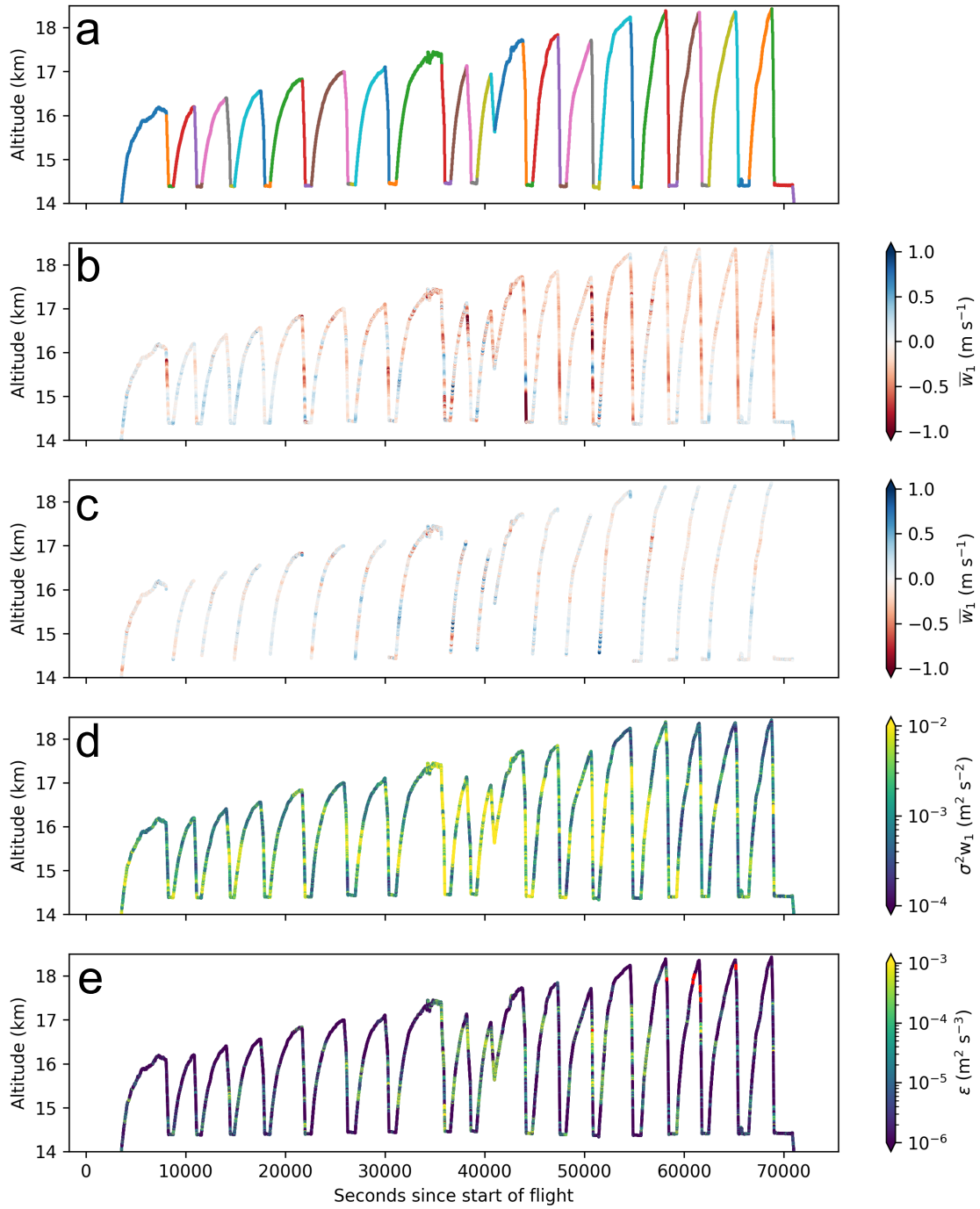
Figure A1. Flight tracks from a flight during ATTREX 2014 shown in time-height space with the color indicating a) different flight legs b) uncorrected mean 1-second vertical wind ($\overline{w}_1$), c) corrected $\overline{w}_1$ d) high frequency (> 1 Hz) vertical wind variance ($\sigma^2 w_1$) and e) turbulent eddy dissipation rate (ϵ). Outliers in ϵ are shown in red.
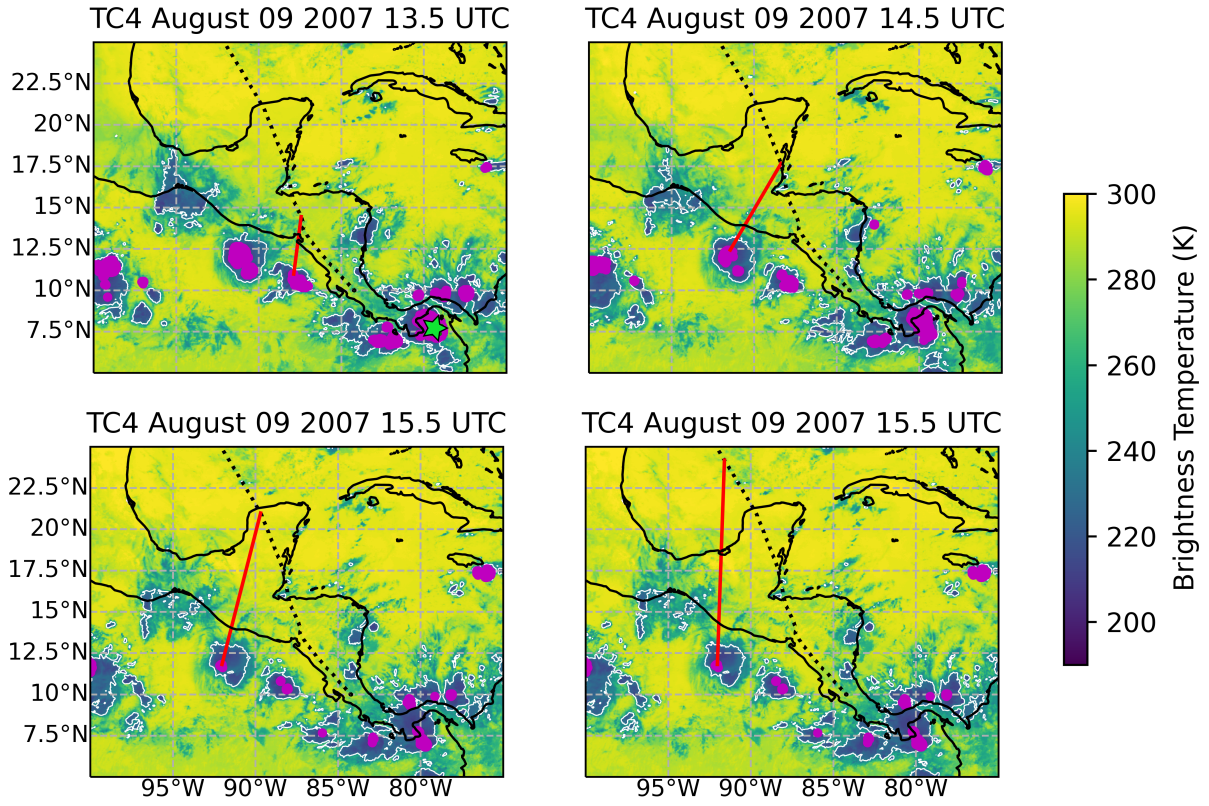
Figure A2. Four snapshots of brightness temperature are shown that overlap an example flight from TC4. The black dotted line shows the flight track. In each snapshot, a red line extends between a point along the flight track that is closest in time to the snapshot shown, and the nearest deep convective core to that point. Pink dots indicate convective cores with brightness temperatures below 210 K. The green star in the upper left plot is the only point that is below 200 K. The white contours surround areas with brightness temperatures below 235 K.

200 K (marked with a green star in the upper right plot), so that threshold is too strict. A threshold of 235 K (white contours) includes some outflow cirrus and remnants of deep convection, which are areas that are influenced by deep convection but that are less likely to generate gravity waves. We select an intermediate value of 210 K (pink points in Figure A2) as our threshold.

Appendix B: Vertical wind variability in all campaigns: additional information

450 **B1 Data categories**

Figure B1 shows histograms of a) IWC, b) NI, c) distance to deep convection and d) altitude, for all flight campaigns separately. Dotted lines shows the boundaries between the categories used in the analyses in Sections 3.1, 3.2, and 3.3. In general, categories are chosen so that each campaign spans at least two categories, making comparisons between categories more
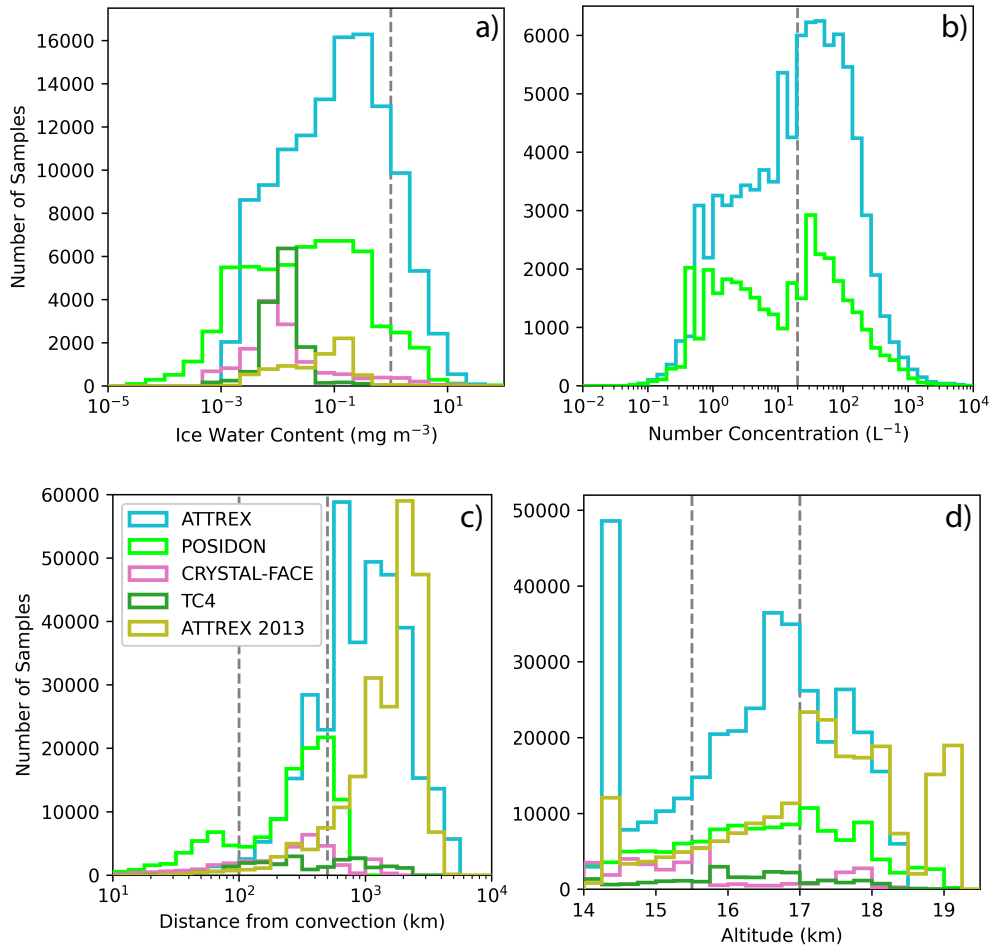
Figure B1. Histogram of a) IWC, b) NI, c) distance from deep convection and d) altitude for all flight campaigns separately. NI is only shown for ATTREX 2014 and POSIDON. Dashed grey lines indicate boundaries used in the study to define categories.

fruitful. NI is only shown for ATTREX 2014 and POSIDON as the other campaigns do not have enough NI data to support a
455 meaningful analysis.

B2 Clear sky analysis

Figure B2 shows distributions of $\overline{w}_1$ and $\sigma^2 w_1$ broken up into categories based on IWC and distance from deep convection, like in Figures 3 and 4, but for clear-sky data only. This analysis shows that the increased vertical wind variability closer to deep convection and lower down in the atmosphere that is shown and discussed in Sections 3.1 and 3.2 is also seen in the
460 clear-sky data.

Clear-sky only
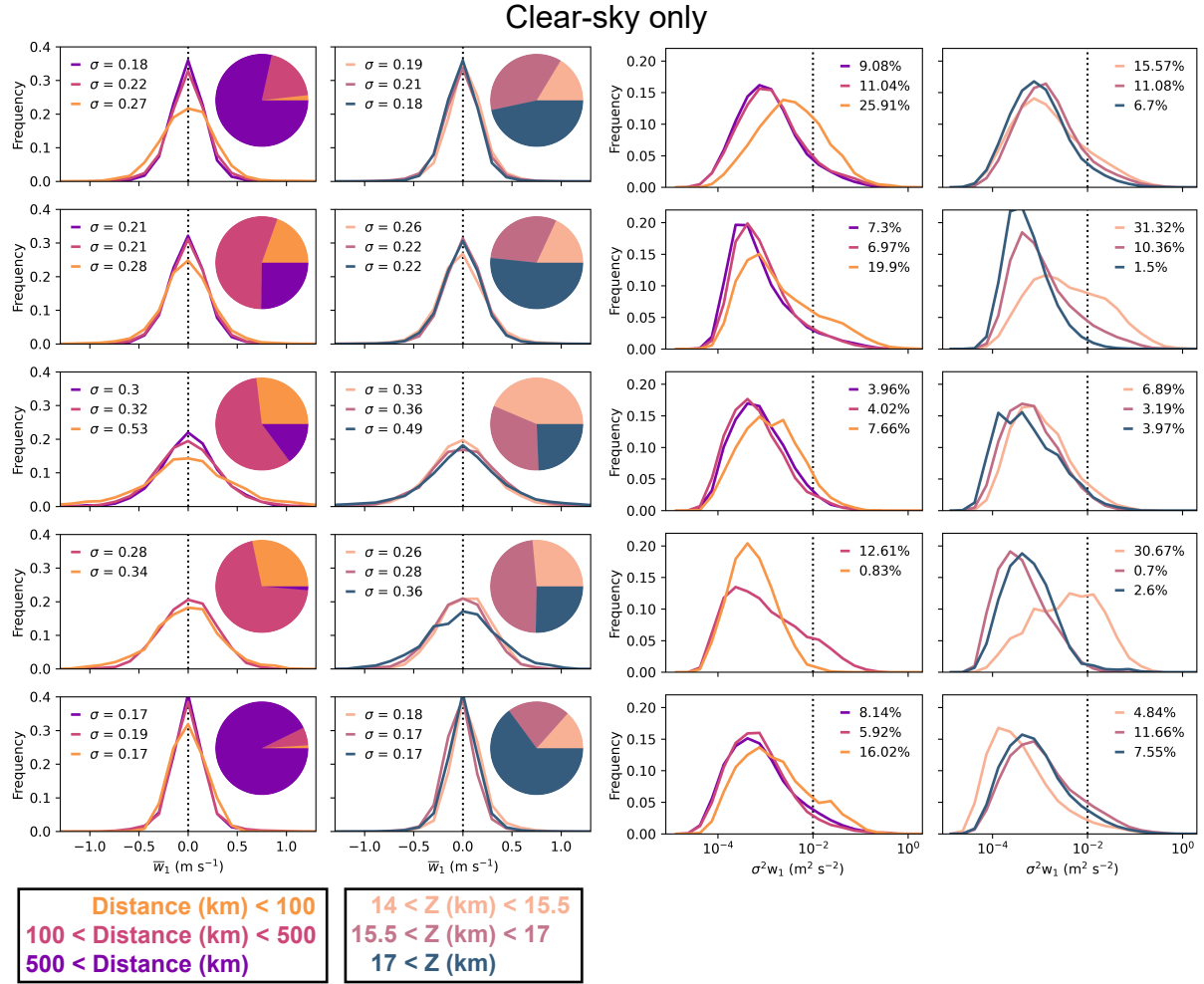
Figure B2. Columns 1-2: Same as columns 3-4 in Figure 3 but for clear-sky data. Columns 3-4: Same as columns 3-4 in Figure 4 but for clear-sky data.
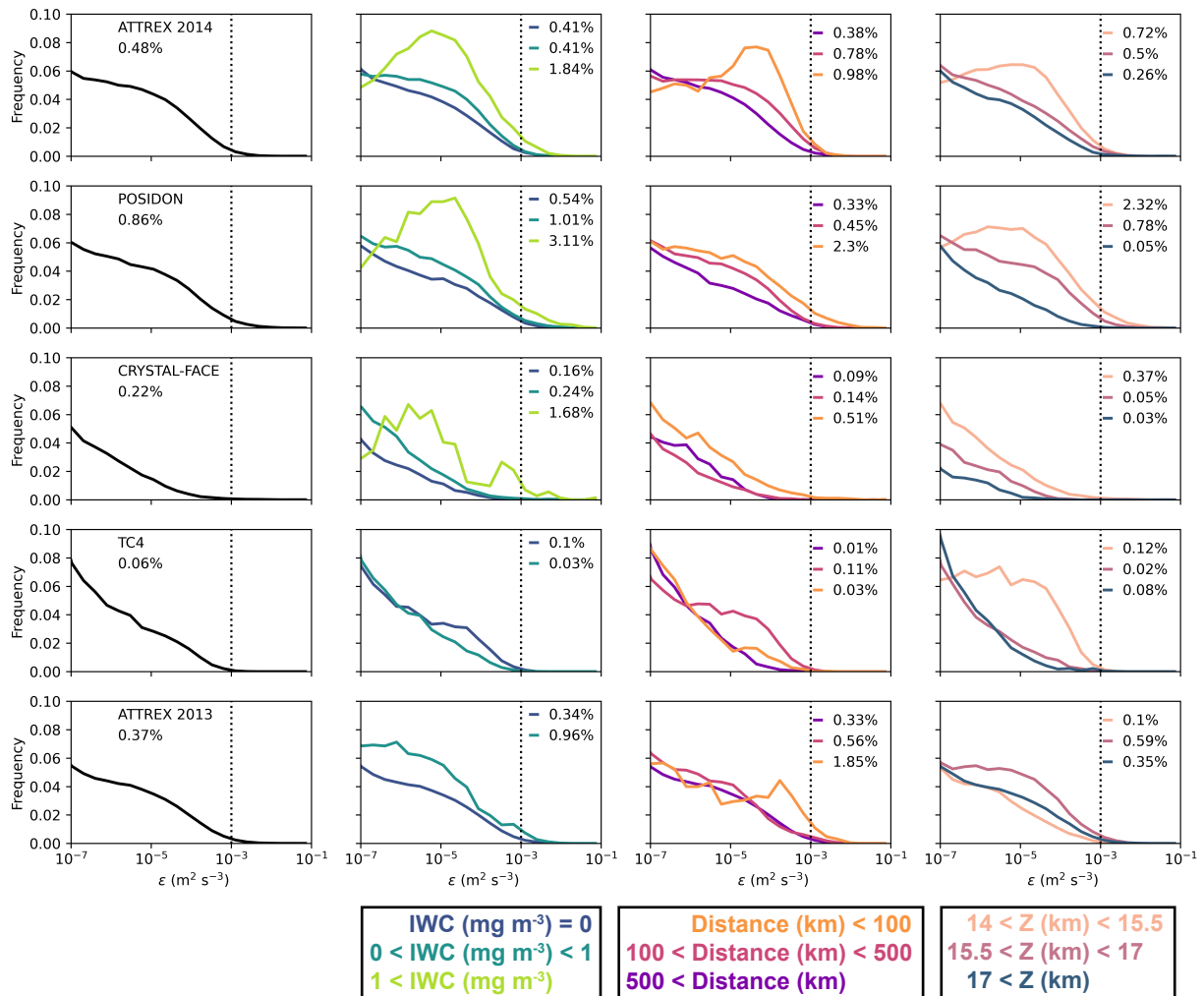
Figure B3. Same as Figure 4 but with distributions of ϵ instead of $\sigma^2 w_1$ as a proxy of turbulence. A threshold of 10^{-3} m^{2}s^{-3} is used as the turbulence threshold.

B3 Turbulent kinetic energy dissipation rate (ϵ)

Figure B3 shows distributions of ϵ broken up into categories based on IWC, distance from deep convection, and altitude, like in Figures 3 and 4. This analysis shows that the relationships discussed in Section 3.1 between the amount of turbulence and IWC, distance from deep convection, and altitude, are robust across different definitions of turbulence, including this much 465 stricter one. Additionally, this figure can be directly compared with Figure 6 from Podglajen et al. (2017).
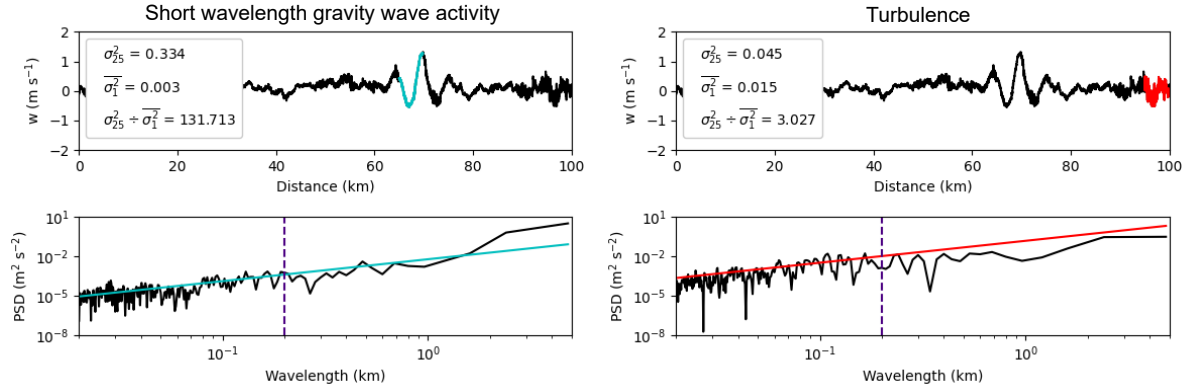
Figure C1. Top row: Vertical wind time series from Example 4 in Figure 5, with specific 5 km sections of interest highlighted in cyan (left) and red (right). $\sigma^2 w_{25}$, $\overline{\sigma^2 w_1}$ and $\sigma^2 w_{25} \div \overline{\sigma^2 w_1}$ for the highlighted sections are overlaid on the plots. Bottom row: Power spectral density for the highlighted 5 km sections (black) with lines fitted to $k^{-5/3}$ between 20 and 100 m (cyan and red).

Appendix C: Gravity wave activity and turbulence detection algorithm

We detect turbulence, LW GWA and SW GWA as follows: For each 25 second/5 km segment within a level leg, we compute the mean vertical wind ($\overline{w}_{25}$), the variance in the high-rate vertical wind ($\sigma^2 w_{25}$), and the mean high frequency (> 1 Hz) vertical wind variance ($\overline{\sigma^2 w_1}$, the average of 25 samples of $\sigma^2 w_1$). Thus, for each level leg we have at least 20 different samples of these variables.

We classify LW GWA based on the difference between the maximum and minimum $\overline{w}_{25}$, which we refer to as $\Delta[\overline{w}_{25}]$, over an entire level leg. If $\Delta[\overline{w}_{25}] > 0.5$ m s^{-1}, we classify the level leg as having LW GWA. Otherwise, we classify the level leg as having negligible LW GWA.

In turbulent conditions, the power spectrum of vertical wind in wavenumbers k is proportional to $k^{-5/3}$ within the inertial sub-range. Figure C1 shows an examples of power spectra for 5 km segments with short wavelength gravity wave activity (bottom left) and turbulence (bottom right). The spectra are assumed to follow $k^{-5/3}$ behavior in the inertial sub-range, and a proportionality constant (related to the turbulent dissipation rate) between the power spectra and $k^{-5/3}$ is fitted to the parts of the spectra between 20 and 100 m, which is approximately the part of the inertial sub-range that can be resolved with 20 Hz (10 m) vertical wind measurement. The cyan and red lines show the predicted power spectra from those fits. In the turbulent case, the power at scales > 1 km is less than what is predicted by $k^{-5/3}$, because those length scales are outside of the inertial sub-range. In the short wavelength gravity wave case, the power at wavelengths > 1 km is greater than what is predicted by $k^{-5/3}$. We interpret that as due to gravity wave activity at these wavelengths.

The vertical wind variance integrated across wavelengths shorter than $l = 2\pi/k$ is proportional to the integral of the power spectrum across wavenumbers greater than k, which is proportional to $k^{-2/3}$ or $l^{2/3}$. Thus, if the ratio of the vertical wind variance across the 5 km (25 s) segment ($\sigma^2 w_{25}$) to the mean vertical wind across 200 m (1 s) sampling windows ($\overline{\sigma^2 w_1}$)

exceeds $25^{\frac{2}{3}}$ or 8.5, we are unlikely to be sampling just turbulence, because the variance in vertical wind is increasing more sharply with wavelength than is plausible for turbulence. In the likely event that one or both of these wavelengths is too long to be in the inertial sub-range of the turbulence, the spectral power will decrease more slowly than $k^{-5/3}$ and the 8.5 ratio threshold is still a sufficient condition that the vertical motions are not just due to turbulence. This can be visualized using Figure C1. The integral of the power spectra up to the blue dashed line is proportional to $\overline{\sigma^2 w_1}$. The integral of the entire power spectrum is proportional to $\sigma^2 w_{25}$. If the power spectra lay along the cyan and red lines, which are the fits to $k^{-5/3}$, then $\sigma^2 w_{25} \div \overline{\sigma^2 w_1}$ would be exactly 8.5. In the short wavelength gravity wave case, the power spectrum is steeper than the cyan line and the ratio is 132. In the turbulence case, the power spectrum is less steep than the red line and the ratio is 3.

Hence we classify 5 km segments as SW GWA if $\sigma^2 w_{25} > 0.04$ m^{2}s^{-1}, and $\sigma^2 w_{25}/\overline{\sigma^2 w_1} > 9.0$. Together, these conditions define a situation where there is a large amount of vertical wind variability on length scales smaller than 5 km that cannot be explained by turbulence. We use an empirically-chosen threshold of 9.0 instead of 8.5, to be slightly conservative in classifying segments as SW GWA.

For 5 km segments that do not have detectable SW GWA, we check for the presence of turbulence. If $\overline{\sigma^2 w_1} > 0.01$ m^{2}s^{-2}, we classify the segment as turbulent.

Competing interests. We declare that no competing interests are present

Acknowledgements. RA acknowledges NSF funding from OISE-1743753 and CB acknowledges funding from AI2. We are extremely grateful to T. Paul Bui for providing high-rate MMS data for CRYSTAL-FACE and TC4, and to Rei Ueyama and Jonathan Dean-Day for their guidance on processing and interpreting the MMS data. We thank Martina Krämer for guidance on the Microphysics Guide to Cirrus, Jacqueline Nugent and Sami Turbeville for providing vertical wind data from DYAMOND-1, and Aurélien Podglajen for helpful commentary on our analysis.

References

Atlas, R. L., Bretherton, C. S., Blossey, P. N., Gettelman, A., Bardeen, C., Lin, P., and Ming, Y.: How Well Do Large-Eddy Simulations and Global Climate Models Represent Observed Boundary Layer Structures and Low Clouds Over the Summertime Southern Ocean?, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 205, https://doi.org/https://doi.org/10.1029/2020MS002205, 2020.

Boehm, M. T., Verlinde, J., and Ackerman, T. P.: On the maintenance of high tropical cirrus, Journal of Geophysical Research: Atmospheres, 104, 24 423–24 433, https://doi.org/https://doi.org/10.1029/1999JD900798, 1999.

Bramberger, M., Alexander, M. J., Davis, S., Podglajen, A., Hertzog, A., Kalnajs, L., Deshler, T., Goetz, J. D., and Khaykin, S.: First Super-Pressure Balloon-Borne Fine-Vertical-Scale Profiles in the Upper TTL: Impacts of Atmospheric Waves on Cirrus Clouds and the QBO, Geophysical Research Letters, 49, e2021GL097 596, https://doi.org/https://doi.org/10.1029/2021GL097596, 2022.

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., Beydoun, H., Taylor, M., Bertagna, L., Bradley, A. M., Clevenger, T. C., Donahue, A. S., Eldred, C., Foucar, J., Golaz, J.-C., Guba, O., Jacob, R., Johnson, J., Krishna, J., Liu, W., Pressel, K., Salinger, A. G., Singh, B., Steyer, A., Ullrich, P., Wu, D., Yuan, X., Shpund, J., Ma, H.-Y., and Zender, C. S.: Convection-Permitting Simulations With the E3SM Global Atmosphere Model, Journal of Advances in Modeling Earth Systems, 13, e2021MS002 544, https://doi.org/https://doi.org/10.1029/2021MS002544, 2021.

Chang, K. W. and L'Ecuyer, T.: Influence of gravity wave temperature anomalies and their vertical gradients on cirrus clouds in the tropical tropopause layer – a satellite-based view, Atmos. Chem. Phys., 20, 12 499–12 514, https://doi.org/10.5194/acp-20-12499-2020, aCP, 2020.

Davis, S. M., Hallar, A. G., Avallone, L. M., and Engblom, W.: Measurement of Total Water with a Tunable Diode Laser Hygrometer: Inlet Analysis, Calibration Procedure, and Ice Water Content Determination, Journal of Atmospheric and Oceanic Technology, 24, 463–475, https://doi.org/10.1175/jtech1975.1, 2007.

Dinh, T., Podglajen, A., Hertzog, A., Legras, B., and Plougonven, R.: Effect of gravity wave temperature fluctuations on homogeneous ice nucleation in the tropical tropopause layer, Atmospheric Chemistry and Physics, 16, 35–46, https://doi.org/10.5194/acp-16-35-2016, 2016.

Dinh, T. P., Durran, D. R., and Ackerman, T. P.: Maintenance of tropical tropopause layer cirrus, Journal of Geophysical Research-Atmospheres, 115, https://doi.org/10.1029/2009jd012735, 2010.

Dobbie, S. and Jonas, P.: Radiative influences on the structure and lifetime of cirrus clouds, Quarterly Journal of the Royal Meteorological Society, 127, 2663–2682, https://doi.org/https://doi.org/10.1002/qj.49712757808, 2001.

Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Nguyen, C., Wielicki, B. A., Young, D. F., and Sun, M.: Geostationary Enhanced Temporal Interpolation for CERES Flux Products, Journal of Atmospheric and Oceanic Technology, 30, 1072–1090, https://doi.org/10.1175/jtech-d-12-00136.1, 2013.

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

Fritts, D. C. and Alexander, M. J.: Gravity wave dynamics and effects in the middle atmosphere, Reviews of Geophysics, 41, https://doi.org/https://doi.org/10.1029/2001RG000106, 2003.

Gasparini, B., Rasch, P. J., Hartmann, D. L., Wall, C. J., and Dütsch, M.: A Lagrangian Perspective on Tropical Anvil Cloud Lifecycle in Present and Future Climate, Journal of Geophysical Research: Atmospheres, 126, e2020JD033 487, https://doi.org/https://doi.org/10.1029/2020JD033487, 2021.

Gultepe, I. and Starr, D. O.: Dynamical structure and turbulence in cirrus clouds: Aircraft observations during FIRE, Journal of the Atmospheric Sciences, 52, 4159–4182, https://doi.org/10.1175/1520-0469(1995)052<4159:dsatic>2.0.co;2, 1995.

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

545 Janowiak, J., Joyce, B., and X., P.: NCEP/CPC L3 Half Hourly 4km Global (60S - 60N) Merged IR V1, https://doi.org/10.5067/P4HZB9N27EKU, 2017.

Jensen, E., Starr, D., and Toon, O. B.: Mission investigates tropical cirrus clouds, Eos, Transactions American Geophysical Union, 85, 45–50, https://doi.org/https://doi.org/10.1029/2004EO050002, 2004.

Jensen, E. J., Pfister, L., and Toon, O. B.: Impact of radiative heating, wind shear, temperature variability, and microphysical processes

550   on the structure and evolution of thin cirrus in the tropical tropopause layer, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/https://doi.org/10.1029/2010JD015417, 2011.

Jensen, E. J., Diskin, G., Lawson, R. P., Lance, S., Bui, T. P., Hlavka, D., McGill, M., Pfister, L., Toon, O. B., and Gao, R.: Ice nucleation and dehydration in the Tropical Tropopause Layer, Proceedings of the National Academy of Sciences, 110, 2041–2046, https://doi.org/10.1073/pnas.1217104110, 2013.

555 Jensen, E. J., Ueyama, R., Pfister, L., Bui, T. V., Alexander, M. J., Podglajen, A., Hertzog, A., Woods, S., Lawson, R. P., Kim, J. E., and Schoeberl, M. R.: High-frequency gravity waves and homogeneous ice nucleation in tropical tropopause layer cirrus, Geophysical Research Letters, 43, 6629–6635, https://doi.org/10.1002/2016gl069426, 2016.

Jensen, E. J., Pfister, L., Jordan, D. E., Bui, T. V., Ueyama, R., Singh, H. B., Thornberry, T. D., Rollins, A. W., Gao, R. S., Fahey, D. W., Rosenlof, K. H., Elkins, J. W., Diskin, G. S., DiGangi, J. P., Lawson, R. P., Woods, S., Atlas, E. L., Rodriguez, M. A. N., Wofsy, S. C.,

560   Pittman, J., Bardeen, C. G., Toon, O. B., Kindel, B. C., Newman, P. A., McGill, M. J., Hlavka, D. L., Lait, L. R., Schoeberl, M. R., Bergman, J. W., Selkirk, H. B., Alexander, M. J., Kim, J. E., Lim, B. H., Stutz, J., and Pfeilsticker, K.: THE NASA AIRBORNE TROP-ICAL TROPOPAUSE EXPERIMENT High-Altitude Aircraft Measurements in the Tropical Western Pacific, Bulletin of the American Meteorological Society, 98, 129–+, https://doi.org/10.1175/bams-d-14-00263.1, 2017.

Kim, J.-E., Alexander, M. J., Bui, T. P., Dean-Day, J. M., Lawson, R. P., Woods, S., Hlavka, D., Pfister, L., and Jensen, E. J.: Ubiquitous

565   influence of waves on tropical high cirrus clouds, Geophysical Research Letters, 43, 5895–5901, https://doi.org/10.1002/2016gl069293, 2016.

Koch, S. E., Jamison, B. D., Lu, C., Smith, T. L., Tollerud, E. I., Girz, C., Wang, N., Lane, T. P., Shapiro, M. A., Parrish, D. D., and Cooper, O. R.: Turbulence and Gravity Waves within an Upper-Level Front, Journal of the Atmospheric Sciences, 62, 3885–3908, https://doi.org/10.1175/jas3574.1, 2005.

570 Krämer, M., Rolf, C., and Spelten, N.: The Cirrus Guide II In-situ Aircraft Data Set, https://doi.org/10.34730/266ca2a41f4946ff97d874bfa458254c, 2020a.

Krämer, M., Rolf, C., Spelten, N., Afchine, A., Fahey, D., Jensen, E., Khaykin, S., Kuhn, T., Lawson, P., Lykov, A., Pan, L. L., Riese, M., Rollins, A., Stroh, F., Thornberry, T., Wolf, V., Woods, S., Spichtinger, P., Quaas, J., and Sourdeval, O.: A microphysics guide to cirrus – Part 2: Climatologies of clouds and humidity from observations, Atmos. Chem. Phys., 20, 12 569–12 608, https://doi.org/10.5194/acp-20-

575   12569-2020, aCP, 2020b.

Kuang, Z. and Bretherton, C. S.: Convective Influence on the Heat Balance of the Tropical Tropopause Layer: A Cloud-Resolving Model Study, Journal of the Atmospheric Sciences, 61, 2919–2927, https://doi.org/10.1175/jas-3306.1, 2004.

Lance, S., Brock, C. A., Rogers, D., and Gordon, J. A.: Water droplet calibration of the Cloud Droplet Probe (CDP) and in-flight performance in liquid, ice and mixed-phase clouds during ARCPAC, Atmos. Meas. Tech., 3, 1683–1706, https://doi.org/10.5194/amt-3-1683-2010,

580   aMT, 2010.

Lawson, R. P., O'Connor, D., Zmarzly, P., Weaver, K., Baker, B., Mo, Q. X., and Jonsson, H.: The 2D-S (Stereo) probe: Design and preliminary tests of a new airborne, high-speed, high-resolution particle Imaging probe, Journal of Atmospheric and Oceanic Technology, 23, 1462–1477, https://doi.org/10.1175/jtech1927.1, 2006.

585 May, R. D.: Open-path, near-infrared tunable diode laser spectrometer for atmospheric measurements of H2O, Journal of Geophysical Research: Atmospheres, 103, 19 161–19 172, https://doi.org/https://doi.org/10.1029/98JD01678, 1998.

Muhlbauer, A., Kalesse, H., and Kollias, P.: Vertical velocities and turbulence in midlatitude anvil cirrus: A comparison between in situ aircraft measurements and ground-based Doppler cloud radar retrievals, Geophysical Research Letters, 41, 7814–7821, https://doi.org/10.1002/2014gl062279, 2014.

Müller, S. K., Manzini, E., Giorgetta, M., Sato, K., and Nasuno, T.: Convectively Generated Gravity Waves in 590 High Resolution Models of Tropical Dynamics, Journal of Advances in Modeling Earth Systems, 10, 2564–2588, https://doi.org/https://doi.org/10.1029/2018MS001390, 2018.

NASA/LARC/SD/ASDC: CERES and GEO-Enhanced TOA, Within-Atmosphere and Surface Fluxes, Clouds and Aerosols 1-Hourly Terra Edition4A, https://doi.org/10.5067/TERRA+AQUA/CERES/SYN1DEG-1HOUR_L3.004A, 2017.

Ortland, D. A. and Alexander, M. J.: The Residual-Mean Circulation in the Tropical Tropopause Layer Driven by Tropical Waves, Journal 595 of the Atmospheric Sciences, 71, 1305–1322, https://doi.org/10.1175/jas-d-13-0100.1, 2014.

Podglajen, A., Bui, T. P., Dean-Day, J. M., Pfister, L., Jensen, E. J., Alexander, M. J., Hertzog, A., Kaercher, B., Plougonven, R., and Randel, W. J.: 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, 3847–3869, https://doi.org/10.1175/jas-d-17-0010.1, 2017.

Podglajen, A., Plougonven, R., Hertzog, A., and Jensen, E.: Impact of gravity waves on the motion and distribution of atmospheric ice 600 particles, Atmospheric Chemistry and Physics, 18, 10 799–10 823, https://doi.org/10.5194/acp-18-10799-2018, 2018.

Schneider, J., Höhler, K., Wagner, R., Saathoff, H., Schnaiter, M., Schorr, T., Steinke, I., Benz, S., Baumgartner, M., Rolf, C., Krämer, M., Leisner, T., and Möhler, O.: High homogeneous freezing onsets of sulfuric acid aerosol at cirrus temperatures, Atmos. Chem. Phys., 21, 14 403–14 425, https://doi.org/10.5194/acp-21-14403-2021, aCP, 2021.

Schoeberl, M. R., Jensen, E. J., and Woods, S.: Gravity waves amplify upper tropospheric dehydration by clouds, Earth and Space Science, 605 2, 485–500, https://doi.org/10.1002/2015ea000127, 2015.

Scott, S. G., Bui, T. P., Chan, K. R., and Bowen, S. W.: The Meteorological Measurement System on the NASA ER-2 Aircraft, Journal of Atmospheric and Oceanic Technology, 7, 525–540, https://doi.org/10.1175/1520-0426(1990)007<0525:tmmsot>2.0.co;2, 1990.

Shindell, D. T.: Climate and ozone response to increased stratospheric water vapor, Geophysical Research Letters, 28, 1551–1554, https://doi.org/https://doi.org/10.1029/1999GL011197, 2001.

610 Skamarock, W. C., Snyder, C., Klemp, J. B., and Park, S.-H.: Vertical Resolution Requirements in Atmospheric Simulation, Monthly Weather Review, 147, 2641–2656, https://doi.org/10.1175/mwr-d-19-0043.1, 2019.

Solomon, S., Rosenlof, K. H., Portmann, R. W., Daniel, J. S., Davis, S. M., Sanford, T. J., and Plattner, G.-K.: Contributions of Stratospheric Water Vapor to Decadal Changes in the Rate of Global Warming, Science, 327, 1219–1223, https://doi.org/doi:10.1126/science.1182488, 2010.

615 Spichtinger, P. and Krämer, M.: Tropical tropopause ice clouds: a dynamic approach to the mystery of low crystal numbers, Atmos. Chem. Phys., 13, 9801–9818, https://doi.org/10.5194/acp-13-9801-2013, aCP, 2013.

Stephan, C. C., Strube, C., Klocke, D., Ern, M., Hoffmann, L., Preusse, P., and Schmidt, H.: Gravity Waves in Global High-Resolution Simulations With Explicit and Parameterized Convection, Journal of Geophysical Research-Atmospheres, 124, 4446–4459, https://doi.org/10.1029/2018jd030073, 2019.

620 Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., Düben, P., Judt, F., Khairoutdinov, M., Klocke, D., Kodama, C., Kornblueh, L., Lin, S.-J., Neumann, P., Putman, W. M., Röber, N., Shibuya, R., Vanniere, B., Vidale, P. L., Wedi, N., and Zhou, L.: DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains, Progress in Earth and Planetary Science, 6, 61, https://doi.org/10.1186/s40645-019-0304-z, 2019.

Thornberry, T. D., Rollins, A. W., Gao, R. S., Watts, L. A., Ciciora, S. J., McLaughlin, R. J., and Fahey, D. W.: A two-channel, tunable diode
625 laser-based hygrometer for measurement of water vapor and cirrus cloud ice water content in the upper troposphere and lower stratosphere, Atmos. Meas. Tech., 8, 211–224, https://doi.org/10.5194/amt-8-211-2015, aMT, 2015.

Toon, O. B., Starr, D. O., Jensen, E. J., Newman, P. A., Platnick, S., Schoeberl, M. R., Wennberg, P. O., Wofsy, S. C., Kurylo, M. J., Maring, H., Jucks, K. W., Craig, M. S., Vasques, M. F., Pfister, L., Rosenlof, K. H., Selkirk, H. B., Colarco, P. R., Kawa, S. R., Mace, G. G., Minnis, P., and Pickering, K. E.: Planning, implementation, and first results of the Tropical Composition, Cloud and Climate Coupling
630 Experiment (TC4), Journal of Geophysical Research: Atmospheres, 115, https://doi.org/https://doi.org/10.1029/2009JD013073, 2010.

Wall, C. J., Norris, J. R., Gasparini, B., Smith, W. L., Thieman, M. M., and Sourdeval, O.: Observational Evidence that Radiative Heating Modifies the Life Cycle of Tropical Anvil Clouds, Journal of Climate, 33, 8621–8640, https://doi.org/10.1175/jcli-d-20-0204.1, 2020.

Weinstock, E. M., Hintsa, E. J., Dessler, A. E., Oliver, J. F., Hazen, N. L., Demusz, J. N., Allen, N. T., Lapson, L. B., and Anderson, J. G.: New fast response photofragment fluorescence hygrometer for use on the NASA ER-2 and the Perseus remotely piloted aircraft, Review
635 of Scientific Instruments, 65, 3544–3554, https://doi.org/10.1063/1.1144536, 1994.