



# 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<sup>1</sup> and Christopher S. Bretherton<sup>1,2</sup>

<sup>1</sup>Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA

<sup>2</sup>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) 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, but this has yet to be confirmed with observations.

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 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  $0.2\text{--}0.4$  m s<sup>-1</sup>. We find that turbulence is enhanced within cirrus with high ice water content (IWC) ( $> 10^{-3}$  g m<sup>-3</sup>), which also have high ice crystal number concentrations. Consistent with a previous study, we find that turbulence is enhanced over the tropical West Pacific and within 500 km of convection, and is most common in the 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 waves during level flight legs is applied to data from the Airborne Tropical Tropopause Experiment (ATTREX). The most commonly sampled conditions are a quiescent atmosphere with negligible small-scale vertical wind variability, and long wavelength gravity wave activity with or without turbulence. Turbulence rarely occurs in the absence of gravity wave activity.

High-IWC cirrus are rare in a quiescent atmosphere, but 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.

Spectral analysis of vertical wind observed in level flight legs over the tropical West Pacific TTL is compared with simulations by four global storm-resolving models with horizontal grid spacings of 3-5 km, sampled over the same region. All four models have too little resolved vertical wind at horizontal wavelengths less than 100 km, but this bias is much less pronounced in global SAM than the other models.



## 1 Introduction

Time-mean vertical motions in the Tropical Tropopause Layer (TTL) are less than  $1 \text{ cm s}^{-1}$  (Ortland and Alexander, 2014), and synoptic-scale vertical motions on scales exceeding 100 km are typically less than  $10 \text{ cm s}^{-1}$  (Section 2.1). However, even well away from deep convective updrafts, gravity waves and turbulence can locally produce vertical winds often exceeding  $1 \text{ m s}^{-1}$ , 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  $0.24 \text{ W m}^{-2}$  (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 ( $> 100 \text{ km}$ ) wave activity.

Numerous modelling studies have suggested that small-scale ( $< 100 \text{ km}$ ) 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.

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 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  $0.5 \text{ km}$  thick with ice water contents of up to  $10^{-3} \text{ g m}^{-3}$ , achieved relatively small turbulent-driven updrafts up to  $2.5 \text{ cm s}^{-1}$ , 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 effects 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.

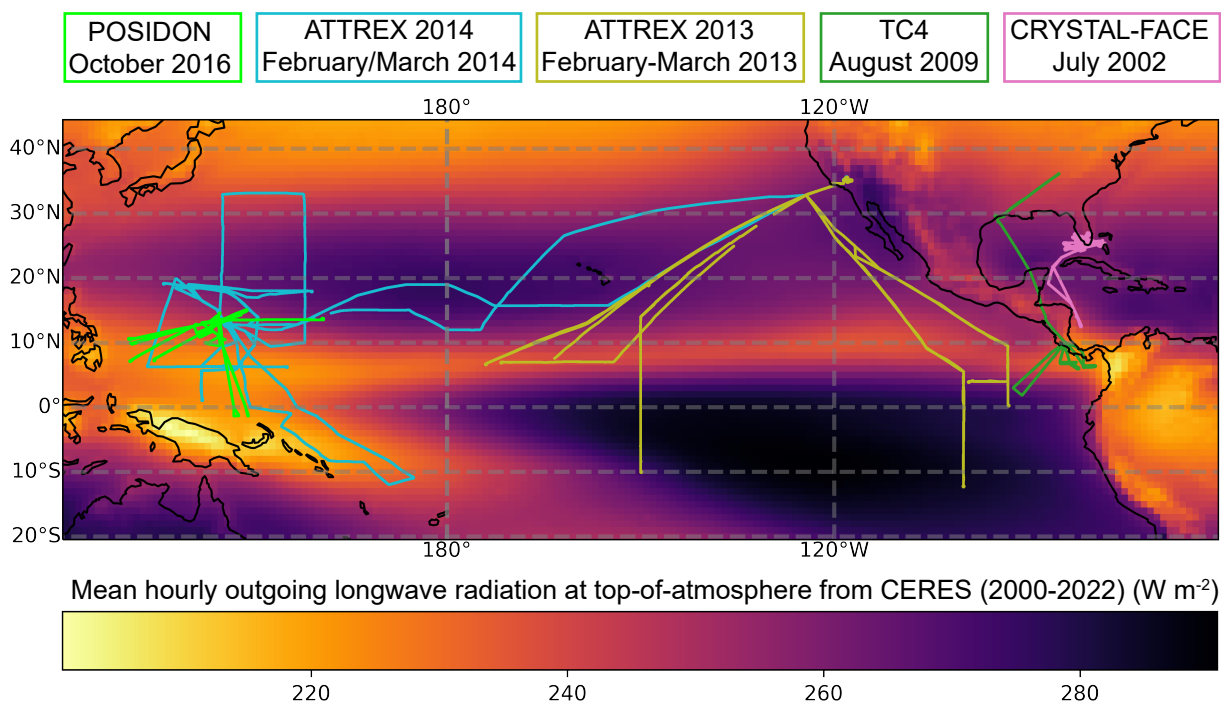
## 2 Preparing the dataset

### 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 two different campaigns), the 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 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 deep convection. The 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 seven science flights that sampled close to 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. Johnathan Dean-day and Rei Ueyama from NASA, aircraft data is separated into flight legs or



**Figure 1.** Map of mean top-of-atmosphere outgoing longwave radiation from CERES with flight tracks from the five campaigns used in this study campaigns overlaid.

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 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 on the order of a  $\text{cm s}^{-1}$ . This estimate is based on analyzing vertical winds in the TTL from ERA5 reanalysis averaged over  $1^\circ \times 1^\circ$ , or about  $100 \text{ km} \times 100 \text{ km}$  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  $10 \text{ cm s}^{-1}$ , and 62.5% have mean vertical wind with magnitudes less than  $1 \text{ cm s}^{-1}$ . The small-scale vertical winds that are the focus of this study have magnitudes of at least  $25 \text{ cm s}^{-1}$ . Flight legs that are less than  $100 \text{ km}$  long and for which the assumption of small mean wind may be less applicable are removed.

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 cleaning procedure in removing them. Correcting the MMS data is necessary for constraining the magnitude of the vertical wind, but sub  $1 \text{ Hz}$  vertical wind variance is well constrained in the uncorrected data because the biases in vertical wind have frequencies lower than  $1 \text{ Hz}$ .



105 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 ( $\epsilon$ ) instead of vertical wind variance to identify and quantify turbulence.  $\epsilon$  must be computed from a power spectrum of  
the aircraft vertical wind. The estimate of  $\epsilon$  is sensitive to the algorithm used to make the power spectrum (e.g. Fourier decom-  
position, wavelet analysis), how  $\epsilon$  is estimated from the power spectrum (e.g. fitting a line, integrating the power spectrum),  
110 and the assumed Kolmogorov constant. For this reason, we use vertical wind variance, which is straightforward to calculate  
and conveys the same information. Vertical wind variance is proportional to  $\epsilon^{2/3}$ , with a constant of proportionality that is  
dependent on the aircraft speed and the sampling frequency. Figure A1d-e shows the vertical wind variance and the reported  
values of  $\epsilon$  for an example flight, to show that the same turbulent patches are clearly evident in both metrics.

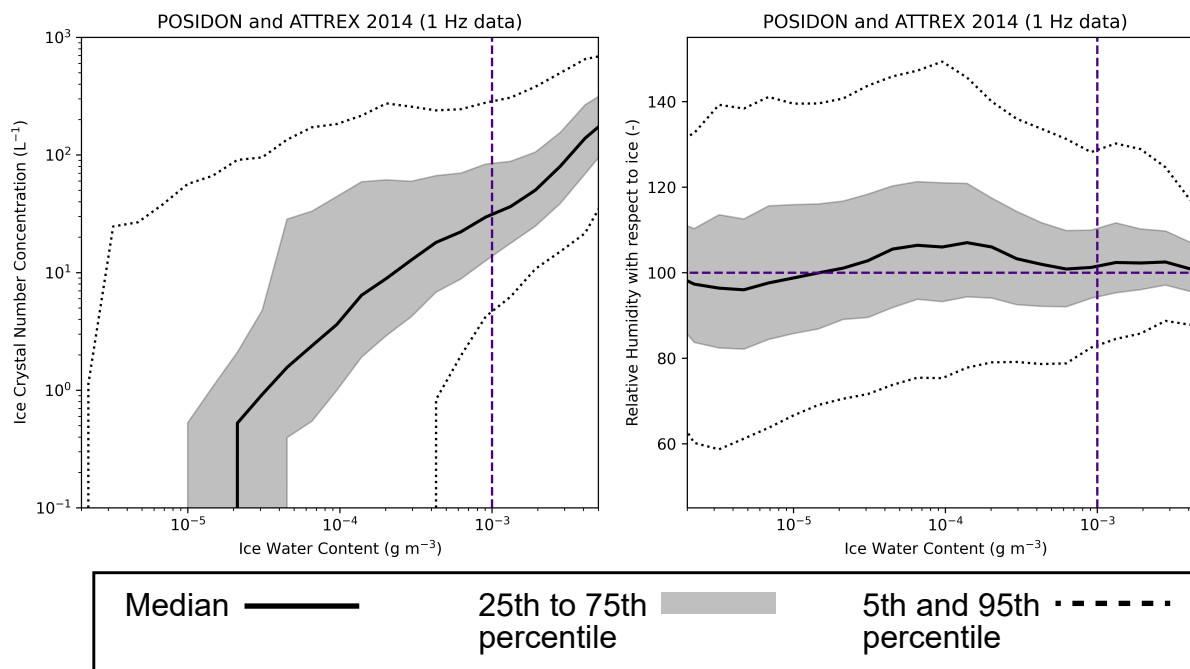
## 2.2 Aircraft measurements of TTL cirrus microphysics

115 For all campaigns analyzed, IWC is computed as a difference between total water and water vapor. A Cloud Laser Hygrometer  
(Davis et al., 2007, CLH) measured total water, and a combination of the Harvard Lyman- $\alpha$  (Weinstock et al., 1994) and the  
JPL laser hygrometer (May, 1998) measured water vapor during CRYSTAL-FACE and TC4. The reported IWC has a detection  
limit of about  $10^{-4}$  g 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  $2 \times 10^{-6}$  g m $^{-3}$ . Nonzero  
120 IWCs that are below the detection limit are uncertain.

Most studies linking TTL cirrus microphysics to dehydration efficiency have focused on ice crystal number concentration  
(NI) as the link. This is plausible because each ice crystal can only grow so large before it falls out of the cirrus cloud. A cirrus  
cloud with more ice crystals thus has more capacity for quenching atmospheric water vapor down to ice saturation through  
depositional growth. Here, we focus on IWC because NI from CRYSTAL-FACE, measured by the Cloud, Aerosol, Precipitation  
125 Spectrometer (Baumgardner et al., 2001, CAPS), is only reported for 30-second intervals and may have contamination from  
shattered artefacts.

IWC and NI are tightly linked in TTL cirrus. Figure 2a shows NI vs IWC from POSIDON and ATTREX 2014. NI was  
measured by the Fast Cloud Droplet Probe (Lance et al., 2010, FCDP, for particles 3-24  $\mu$ m) and the Two Dimensional Stereo  
Probe (Lawson et al., 2006, 2D-S, for particles 25-3005  $\mu$ m) for both campaigns. The detection limits for the FCDP and the  
130 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 $^{-1}$ . We show data from POSIDON and ATTREX 2014 only because they extensively sampled a wide range  
of IWCs above their detection threshold of  $2 \times 10^{-6}$  g m $^{-3}$ .

NI increases with IWC, and the rate of increase is larger for IWC  $> 10^{-3}$  g m $^{-3}$  (to the right of the vertical purple dashed  
line). NI for IWC  $< 2 \times 10^{-5}$  g m $^{-3}$  is frequently below the detection limit of the 2DS and CDP. We expect that TTL cirrus  
135 with higher NI more effectively dehydrate the TTL. Figure 2b shows relative humidity with respect to ice (RH $_i$ ) vs IWC for  
the same two campaigns. Cirrus with IWC between  $2 \times 10^{-5}$  and  $5 \times 10^{-4}$  g m $^{-3}$  are too diffuse to quench atmospheric water  
vapor down to ice saturation. As a result, they have median RH $_i$  near 110% with a large spread around the median. Cirrus with  
higher ice water contents have median RH $_i$  closer to 100% with a narrower spread around the median.



**Figure 2.** Statistics of ice crystal number concentrations (NI, left plot) and relative humidity with respect to ice ( $RH_i$ , right plot), are shown as a function of IWC for 1 Hz data from POSIDON and ATTREX 2014

Throughout this study, we differentiate between low-IWC ( $IWC < 10^{-3} \text{ g m}^{-3}$ ) and high-IWC cirrus ( $IWC > 10^{-3} \text{ g m}^{-3}$ ). The vertical purple dashed lines in Figure 2a-b show the boundary between the two categories. High-IWC cirrus (to the right of the boundary) are associated with high NI and  $RH_i$  near ice saturation, indicating that they are more effective at dehydrating the TTL than lower-IWC cirrus.

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

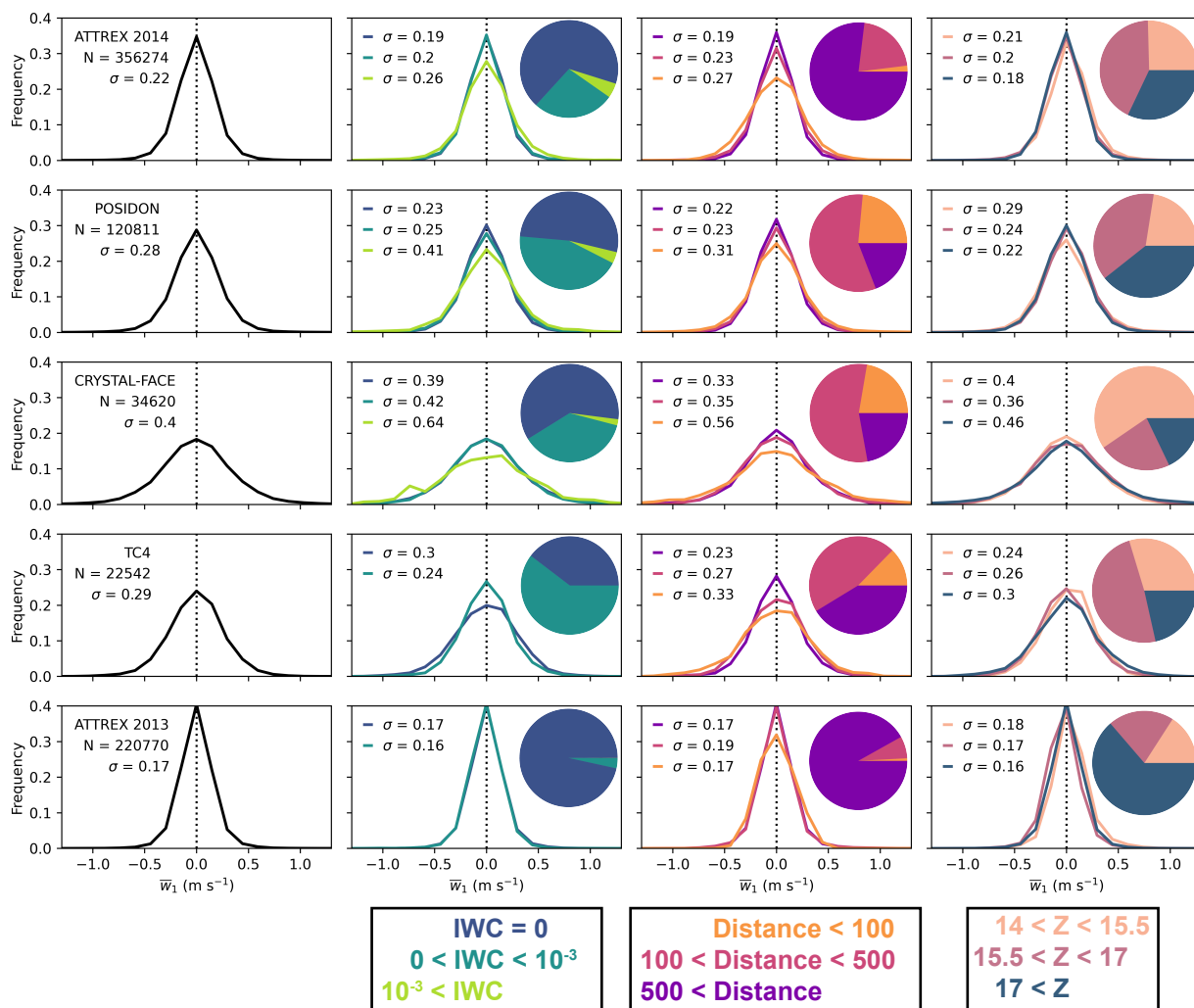
We analyze the 1 Hz vertical wind ( $\bar{w}_1$ ) in this section and sub 1 Hz vertical wind variance ( $\sigma_1^2$ ) in the next section, for all five flight campaigns separately.  $\bar{w}_1$  is sensitive to both gravity wave activity and turbulence, whereas  $\sigma_1^2$  is sensitive mainly to turbulence.

Throughout these two sections, we split the data up into brackets 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 brackets. We split the IWC into three brackets: clear sky (IWC = 0.0), low-IWC cirrus (between 0.0 and  $10^{-3}$  g m $^{-3}$ ), and high-IWC cirrus (IWC >  $10^{-3}$  g 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  $10^{-4}$  g m $^{-3}$ .

Figure 3 shows probability distributions of  $\bar{w}_1$  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 overlaid. The second, third and fourth columns show probability distributions of  $\bar{w}_1$  split into brackets 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 brackets. Normal distributions are fitted to all probability distributions and the fitted standard deviations ( $\sigma$ ) are overlaid.

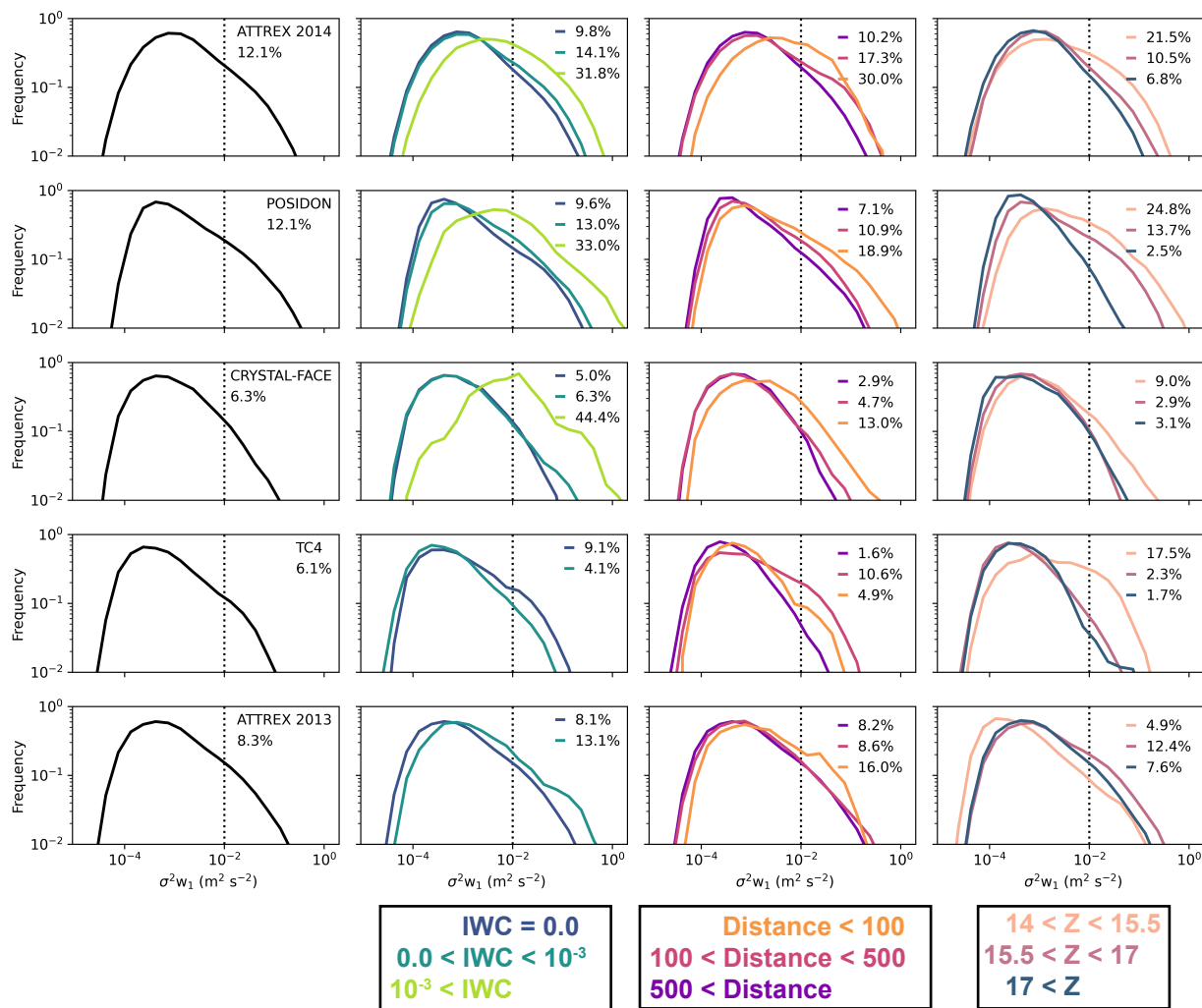
Values of  $\sigma$  vary between 0.17 and 0.4 m s $^{-1}$  for all campaign data (first column of Figure 3). Differences in vertical wind variability across the set of campaigns may arise from sampling closer to or farther from deep convection, the properties of the deep convection (e.g., land vs ocean, 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 sampled near convection in the Florida region, has the widest distribution of  $\bar{w}_1$  (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 convection, and height, and compare the PDF of  $\bar{w}_1$  across these categories. ATTREX 2014, POSIDON and CRYSTAL-FACE sampled high-IWC cirrus, and those data are associated with wider  $\bar{w}_1$  distributions (top three rows of second column of Figure 3). In general, vertical wind variability increases closer to deep convection (third column of Figure 3). An exception



**Figure 3.** First column: Distributions of 1 Hz mean vertical wind ( $\bar{w}_1$ ) for all corrected data. Campaign names and the number of 1 Hz samples are overlaid. Second, third and fourth columns show the same thing split into brackets based on IWC, distance from deep convection, and height in the TTL, respectively. Sigma values for fitted normal distributions are overlaid on all panels. Pie charts showing the distribution of brackets are overlaid on columns 2-4.





**Figure 4.** Same as Figure 2 except showing distributions of sub 1 Hz vertical wind variance ( $\sigma_1^2$ ). The percentage of data exceeding the turbulent threshold ( $0.01 \text{ m}^2 \text{ s}^{-2}$ , vertical dotted line) is overlaid for each category on each plot.

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 but only for clear sky data; the results are the same. Thus, the increase in vertical wind variability closer to deep convection is not caused by more cirrus cloud. The relationship between vertical wind variability and height is inconsistent across the five campaigns.



### 190 3.2 Turbulence in all campaigns

Figure 4 shows distributions of sub 1 Hz vertical wind variance (variance of the 20 sub-samples within each second of data,  $\sigma_1^2$ ), which is a proxy for turbulence. The distributions are split into the same brackets based on IWC, distance from deep convection, and height, as in Figure 3. We define  $\sigma_1^2$  greater than an empirical detectability threshold of  $0.01 \text{ m}^2 \text{ s}^{-2}$  as turbulent, and we mark that threshold with a black dotted line. The percentage of data that exceeds that threshold is printed on the plots. This  
195 turbulence threshold is lower than the one used by Podglajen et al. (2017), as shown in Figure A1d and e.

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 Pacific, have about twice as much turbulence as CRYSTAL-FACE and TC4, even though those 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-  
200 FACE.

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 triple that for clear-sky data (second column of Figure 4). The enhancement is particularly dramatic during CRYSTAL-FACE but there are only 716 1 Hz high-IWC  
205 samples from that campaign, so the statistics are less robust than for POSIDON and ATTREX 2014. Increased turbulence can account for much of the widening of the vertical wind distribution within high-IWC cirrus seen in Figure 3.

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  
210 turbulence between the two other brackets (100-500 km and  $> 500$  km) are largely diminished in the clear-sky data. 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 only within 500 km of deep convection. Consistent with these results, ATTREX 2013 has the most sampling further than 500 km away from deep convection (Figures 3 and B1).

We analyzed aircraft data from the TTL from five NASA aircraft campaigns that took place in different years, seasons, and  
215 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  $\bar{w}_1$  is well approximated as a normal distribution with a standard deviation between 0.17 and  $0.56 \text{ m s}^{-1}$ , 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  
220 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, which are more effective at dehydrating the atmosphere due to having larger NI.

### 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.

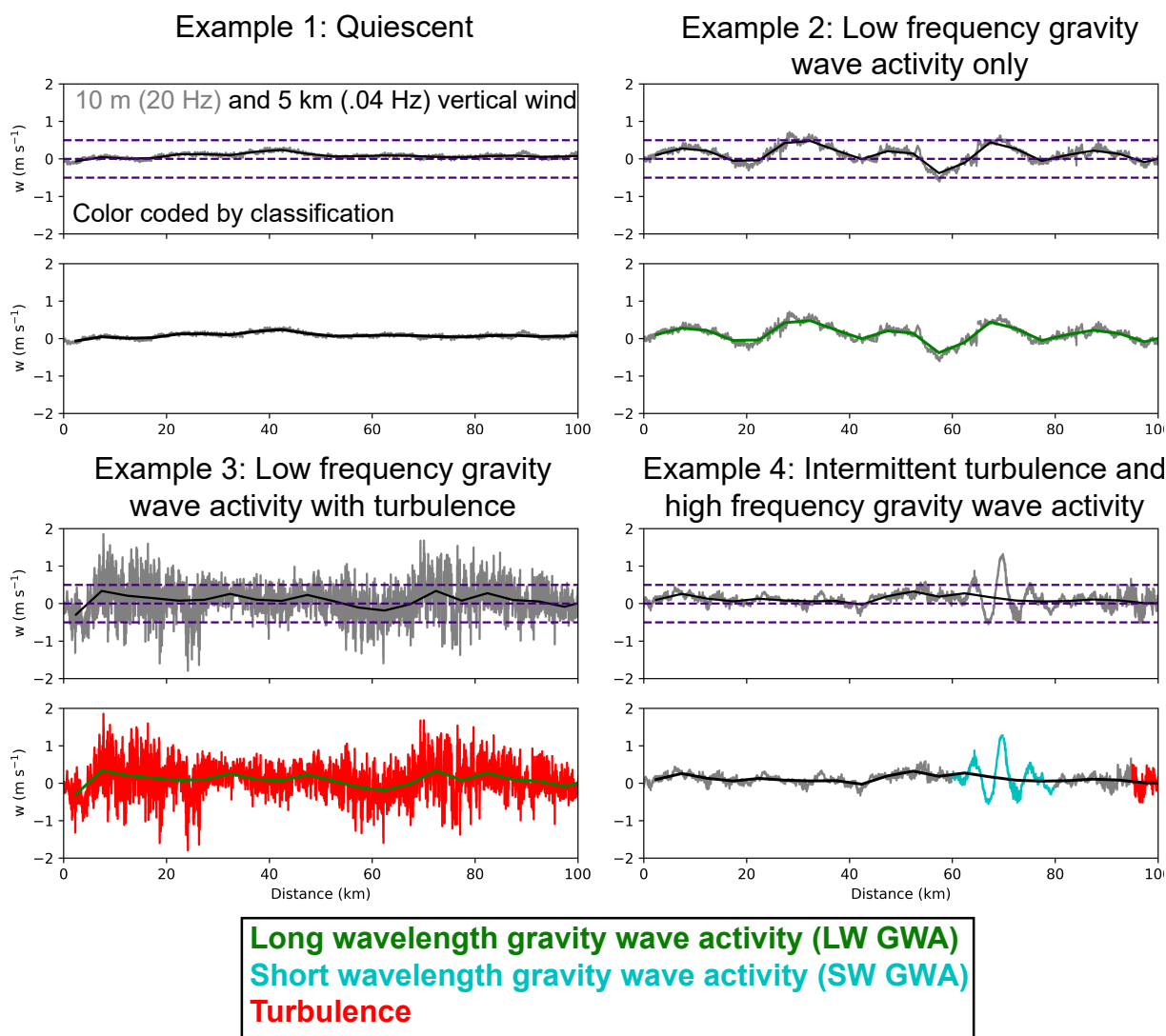
ATTREX 2014's flight strategy involved repeatedly flying parallel to the ground at a height of about 14.2 km, typically through cloud, performing a slow ascent, and a quick descent (Figure A1). ATTREX 2013 sometimes employed this flight strategy and sometimes performed ascents and descents with no level 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 and IWC. 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), to evaluate small-scale simulated 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 larger than 5 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<sup>-1</sup> 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 ( $\bar{w}_{25}$ ) 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  $\bar{w}_{25}$  is color coded according to the LW GWA classification (for the entire level



**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) and 5 km mean vertical wind. In the top time series for each example, high-rate vertical wind is in grey and  $\bar{w}_{25}$  is in black. Horizontal dashed purple lines indicate 0.0, 0.5 and  $-0.5 \text{ m s}^{-1}$ . In the bottom time series for each example,  $\bar{w}_{25}$  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.



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, while 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 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)).

Each 5 km segment can have three possible sub 5 km classifications, with or without LW GWA activity. 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.

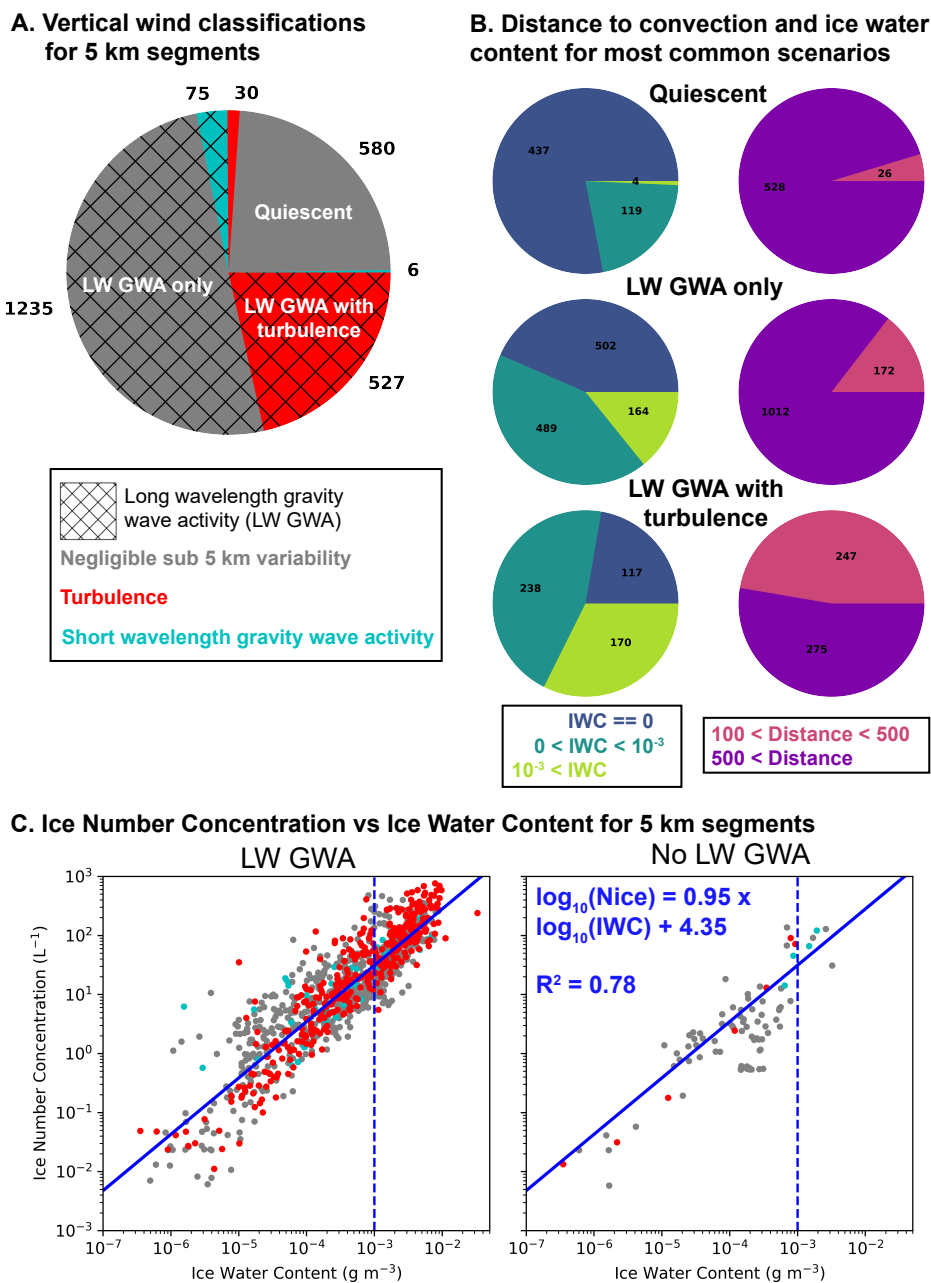
Figure 6b shows pie charts of 5 km mean IWC and distance from deep convection for those three common conditions. The likelihoods of occurrence of high-IWC cirrus given quiescent conditions, LW GWA only and LW GWA with turbulence are 0.6%, 13% and 33%, respectively. Thus, high-IWC cirrus are about 20 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. 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.

Only high-IWC cirrus might be capable of generating cloud-driven turbulence that produces updraft speeds above the noise threshold of the MMS. Turbulence co-occurs with high-IWC cirrus 30% of the time (Figure 6b). Therefore, at most, 30% of the turbulence can be cloud-driven. However, high-IWC cirrus occur in the absence of turbulence about half of the time. In general, the turbulence is much more highly correlated with LW GWA than with high-IWC cirrus. Turbulence co-occurs with LW GWA 95% of the time (Figure 6a). For clear-sky segments only, turbulence co-occurs with LW GWA 89% of the time (not shown).

Figure 6c shows scatter plots of NI vs IWC for 5 km segments with LW GWA (left side) and negligible LW GWA (right side.) The colors of the points shows their sub 5km variability classification. We fit a line to  $\log_{10}(\text{NI})$  vs  $\log_{10}(\text{IWC})$  for all of the data and we overlay the fitted relationship on the left plot, and show the fitted line on both plots. We note that new homogeneous nucleation produces very high NI ( $10^4 L^{-1}$ ) on very small length and time scales that can only be detected in sub 1 Hz data. The increase in NI with IWC is consistent with small-scale vertical motions thickening TTL cirrus by initiating new homogeneous nucleation. The apparent linear relationship between  $\log_{10}(\text{NI})$  and  $\log_{10}(\text{IWC})$  for 5 km segments could be a useful target for ice microphysics parameterization, particularly for global cloud-resolving models, which have horizontal grid spacing close to 5 km.

### 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 spacings 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



**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) Frequency of brackets of IWC and distance from deep convection for the three most common atmospheric conditions, with colors following Figures 2 and 3. c) Ice crystal number concentrations vs IWC for 5 km segments with LW GWA (left) and without LW GWA (right), with colors showing the sub 5 km variability classification. Blue solid lines show the fitted linear relationship, which is overlaid on the left plot. Blue dashed lines show the lower boundary of high-IWC cirrus.



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.

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 initialized in the same way 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 AND ICON have horizontal grid spacing of 3.25 km, 5 km, 3.25 km and 2.5 km, respectively. NICAM has 400 m vertical grid spacing in the TTL, whereas the other three models have close to 500 m.

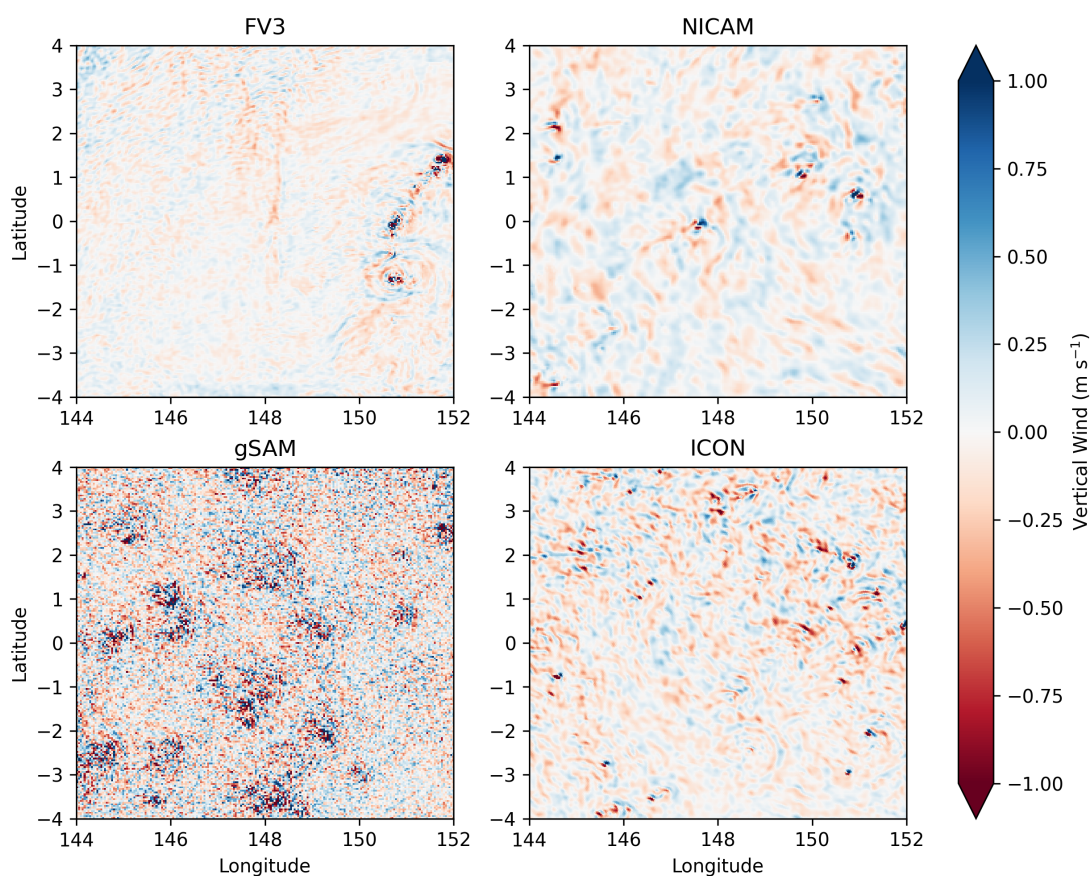
Figure 7 shows snapshots of vertical wind from the 141st hour of simulation at the vertical levels closest to 14 km. Figure S1 is an animation of vertical wind snapshots for hours 48-957 of the simulations (allowing two days for model spinup), showing that the snapshot in Figure 7 is representative of 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 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 8 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  $0.01^\circ$  lat-lon grid. For the observations only, we use 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.

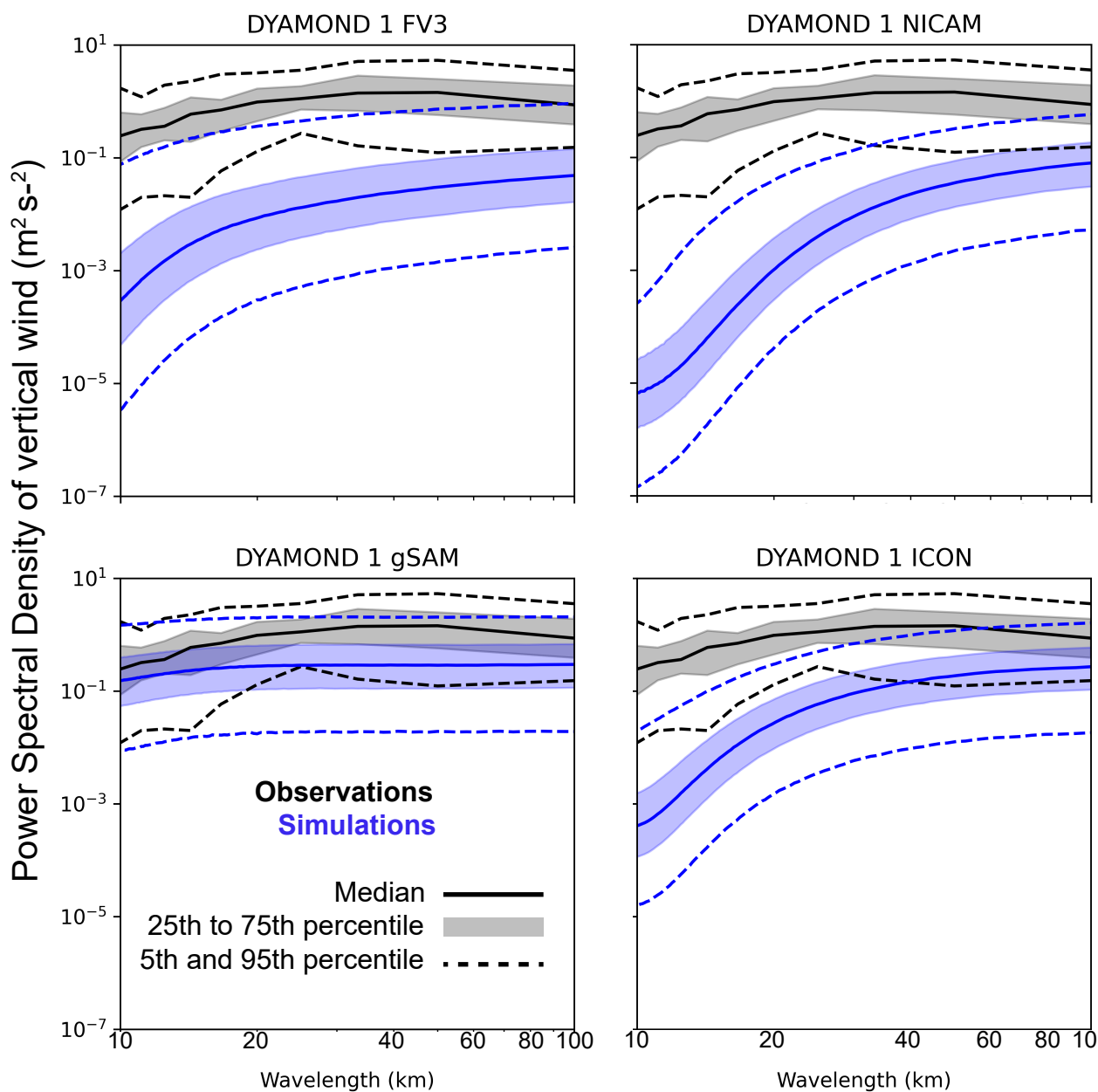


DYAMOND-1 hour = 141



**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.** Statistics of 1-D power spectral density for the ATTREX 2014 observations over the tropical West Pacific (black) and the DYAMOND-1 simulations (blue). Solid lines are medians, dotted lines are 5th and 95 percentiles, and shaded areas span the ranges of the 25th to 75th percentiles.



#### 4 Conclusions

325 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  
330 variability.

Turbulence during ATTREX was analyzed in detail in [Podglajen et al. \(2017\)](#), and we find that several 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.

335 For the first time, we used aircraft measurements to correlate gravity wave activity and turbulence with TTL cirrus microphysical properties. Turbulence is strongly enhanced in cloudy regions with ice water contents (IWCs) larger than  $10^{-3} \text{ gm}^{-3}$ . These high-IWC cirrus have high ice crystal number concentrations that make them more effective at dehydrating the atmosphere than lower-IWC cirrus, so understanding their formation mechanisms is particularly important. During ATTREX 2014, 99% of 5 km segments with mean IWCs larger than  $10^{-3} \text{ gm}^{-3}$  co-occurred with long wavelength gravity wave activity, and  
340 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 high IWC 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. Additionally, each of the modelling studies focused on either gravity wave activity or turbulence. Here, we show that both sources of small-scale vertical wind  
345 variability are important, and that they frequently occur together.

The common co-occurrence of high ice-water content 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.

350 [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  
355 co-occurred with gravity wave activity 95% of the time, and high-IWC cirrus only 30% of the time. We encourage future studies to more closely examine turbulent sources in the TTL.



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. That is, 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 less 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 GSRM or, more practically, regional cloud-resolving model studies to examine the effects of increased vertical and horizontal resolution on 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 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>

*Video supplement.* Video S1 cycles through snapshots of vertical wind from every hour in the DYAMOND-1 simulations, except hours , 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. Hours 531-549 are excluded because there is no ICON vertical wind data during those times.



## 380 Appendix A: Preparing the dataset: additional information

### A1 Cleaning vertical wind data

Figure A1 shows flight tracks for an example flight from ATTREX 2014 and Figure A1b color codes the flight track with the uncorrected vertical wind. After 55,000 seconds into the flight, a pattern emerges where there are apparent downdrafts throughout the entire descent, apparent updrafts in the first half of the ascent, and apparent downdrafts in the second half. This behavior is nonphysical and 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 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.

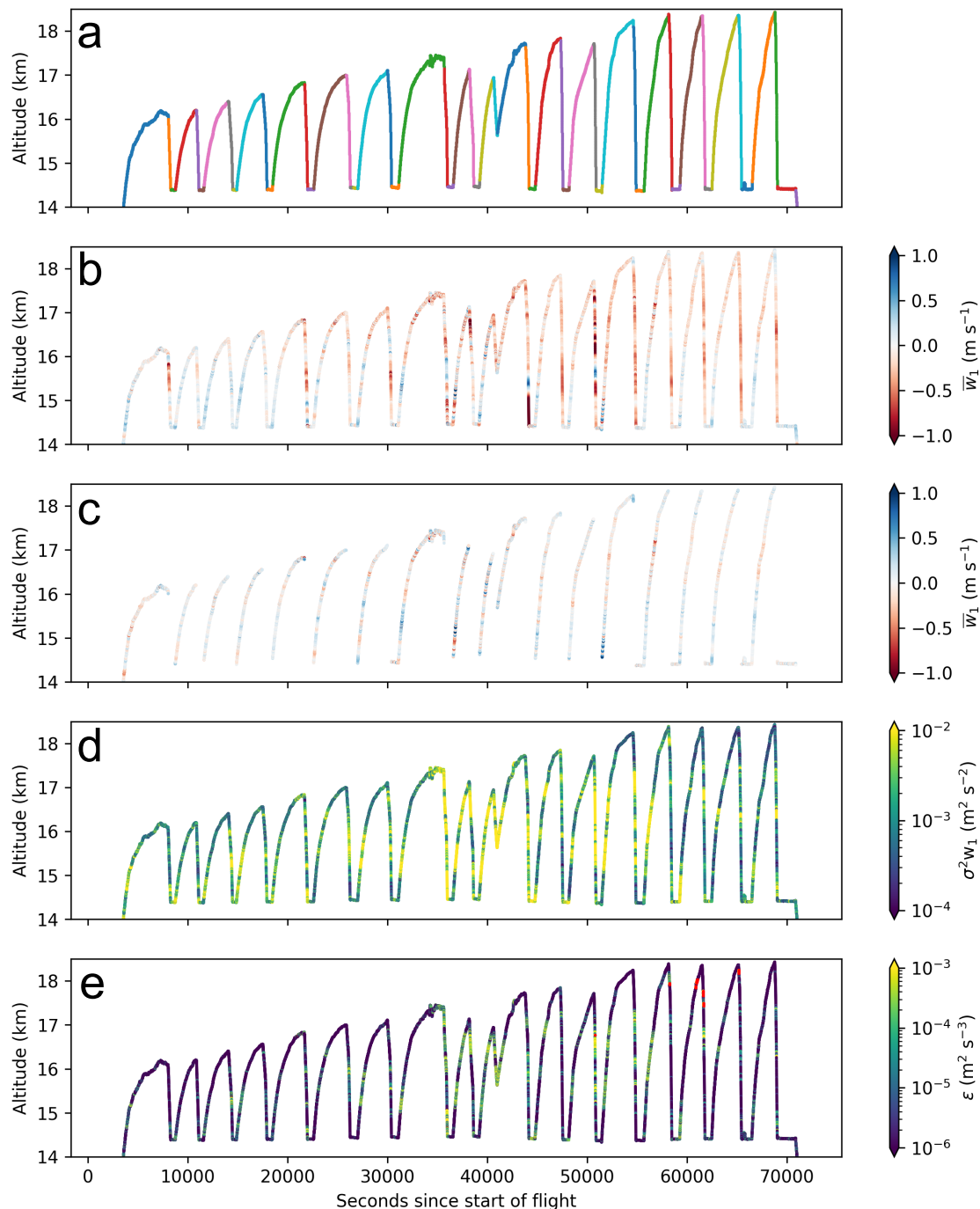
Figures A1d-e show the 1-second vertical wind variance ( $\sigma_1^2$ ) and the 1-second  $\epsilon$  as reported in the NASA dataset, respectively.  $\epsilon$  has rare outliers with unrealistically high values outside patches of turbulence, which are mainly seen between 60,000 and 65,000 seconds into the flight (red points). We note that Podglajen et al. (2017) did not use the reported  $\epsilon$ , so their analysis was not affected by these outliers.

Outliers aside, the same turbulent patches are evident in both  $\epsilon$  and  $\sigma_1^2$ . Both metrics are similarly useful for identifying turbulence so we choose  $\sigma_1^2$  based on its greater intuitiveness and ease of calculation. 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.

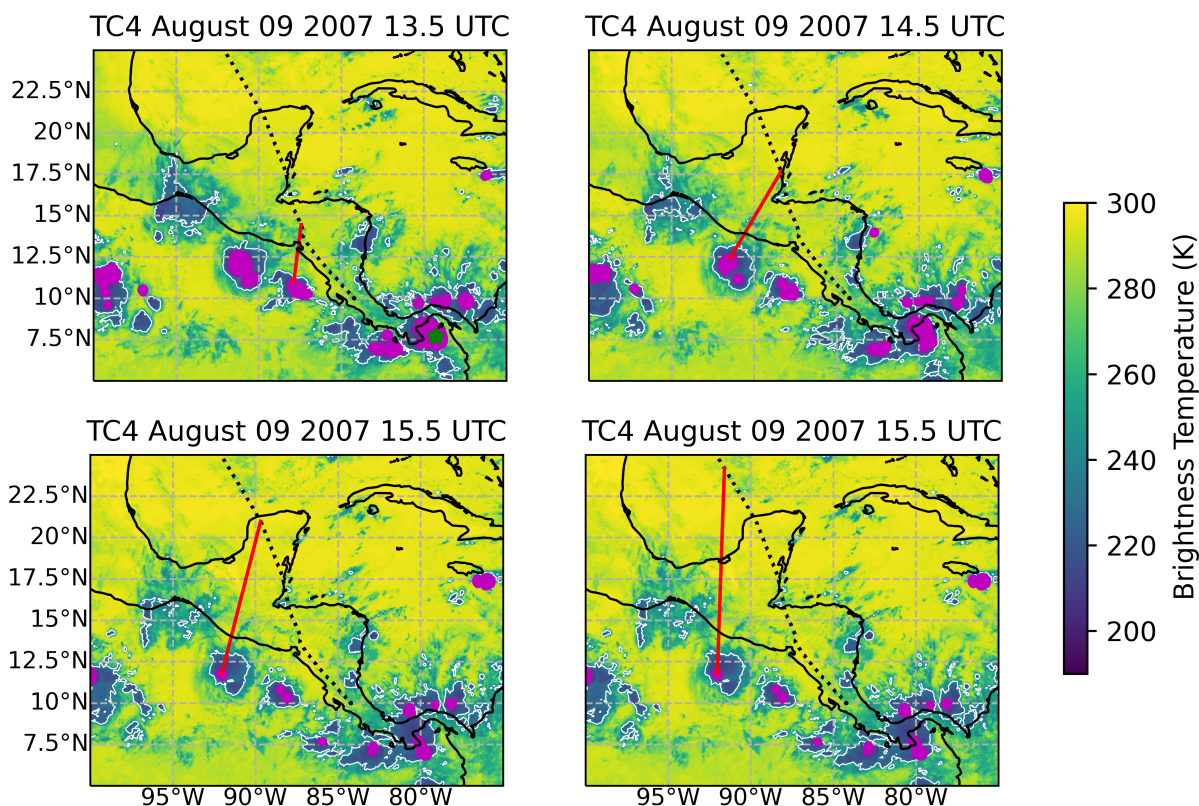
Figure A1 also demonstrates the sampling strategy for ATTREX 2014 and parts of ATTREX 2013, which included many level legs that are the focus of the analyses in Sections 3.2 and 3.3.

### 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 activity, moisture, and possible detrained ice in the TTL. Only one pixel in all four snapshots has a brightness temperature below 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.



**Figure A1.** Flight tracks from a flight during ATTREX 2014 shown in time-height space with the color indicating a) different flight legs b) mean 1-second uncorrected vertical wind ( $\bar{w}_1$ ), and C) mean 1-second corrected vertical wind ( $\bar{w}_1$ ) d) variance in 1-second vertical wind ( $\sigma^2 w_1$ ) and e) turbulent eddy dissipation rate ( $\epsilon$ ). Outliers in  $\epsilon$  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 right plot is the only point that is below 200 K. The white contours surround areas with brightness temperatures below 235 K.

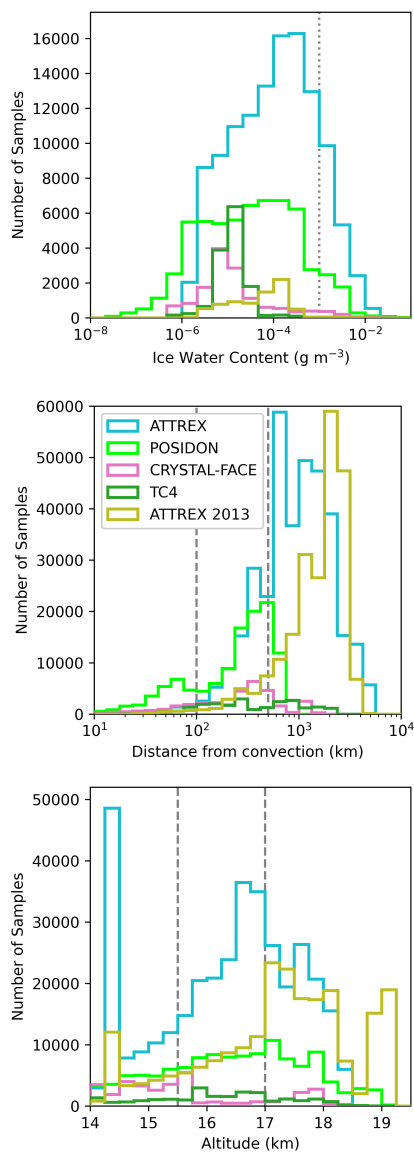
## Appendix B: Vertical wind variability in all campaigns: additional information

### B1 Bracketing the data

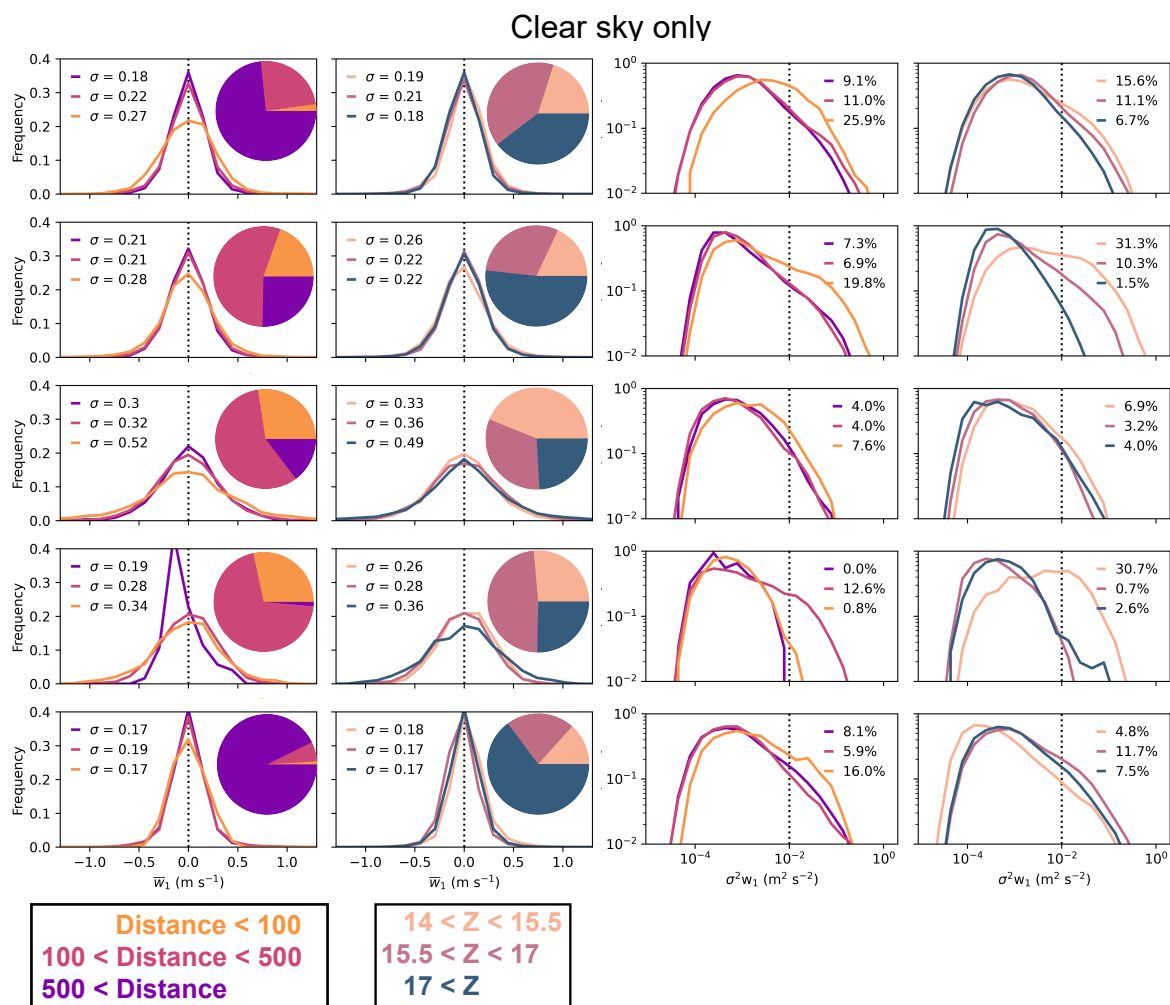
Figure B1 shows histograms of IWC, distance to deep convection and altitude for all flight campaigns separately. Dotted lines show the boundaries between the brackets used in the analyses in Sections 3.1 and 3.2. In general, brackets are chosen so that  
410 each campaign spans at least two brackets, making comparisons between brackets more fruitful.

### B2 Clear sky analysis

Figure B2 shows distributions of  $\bar{w}_1$  and  $\sigma_1^2$  broken up into brackets based on IWC and distance from deep convection, like in Figures 2 and 3, but for clear-sky data only. This analysis shows that the increased vertical wind variability closer to deep



**Figure B1.** Histogram of (top to bottom) IWC, distance from deep convection and altitude for all flight campaigns separately. Dotted lines indicate bracket boundaries.



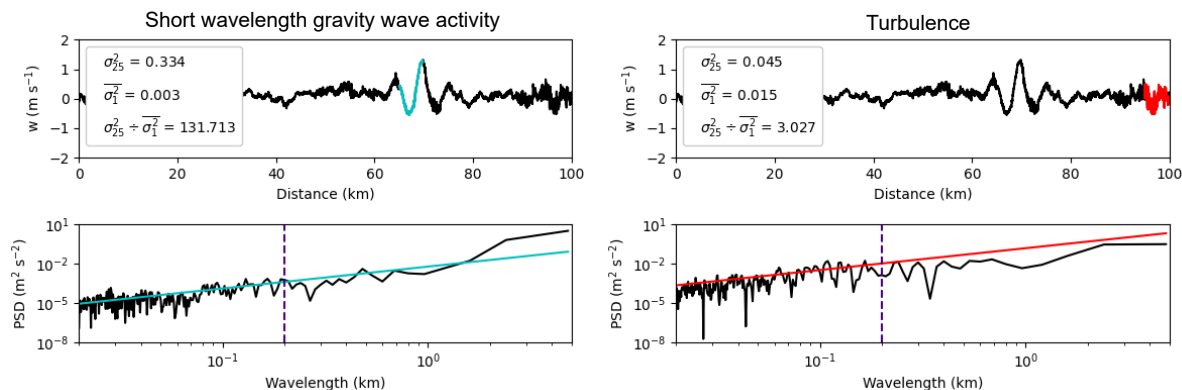
**Figure B2.** Columns 1-2: Same as columns 2-3 in Figure 2 but for clear-sky data. Columns 3-4: Same as columns 3-4 in Figure 3 but for clear-sky data.

415 convection and lower down in the atmosphere that is shown and discussed in Sections 3.1 and 3.2 is also seen in the clear-sky data.

### 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 ( $\sigma_{25}^2$ ), and the mean sub 1 Hz vertical wind variance ( $\overline{\sigma_1^2}$ ). Thus, for each level leg we have at least 20 different samples of these variables.





**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).  $\sigma_{25}^2$ ,  $\overline{\sigma_1^2}$  and  $\sigma_{25}^2 \div \overline{\sigma_1^2}$  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).

420 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  $\Delta[\overline{w}_{25}] > 0.5 \text{ m s}^{-1}$ , 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 right) and turbulence (bottom left). 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  $> 1 \text{ km}$  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  $> 1 \text{ km}$  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 ( $\sigma_{25}^2$ ) to the mean vertical wind across 200 m (1 s) sampling windows ( $\overline{\sigma_1^2}$ ) exceeds  $25^{2/3}$  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  $\overline{\sigma_1^2}$ . The integral of the entire power spectrum is proportional to  $\sigma_{25}^2$ . If the power spectra lay along the cyan and red lines, which are the fits to  $k^{-5/3}$ , then  $\sigma_{25}^2 \div \overline{\sigma_1^2}$  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  $\sigma_{25}^2 > 0.04 \text{ m}^2\text{s}^{-1}$ , and  $\sigma_{25}^2/\overline{\sigma_1^2} > 9.0$ . 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  
445 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  $\overline{\sigma_1^2} > 0.01 \text{ m}^2\text{s}^{-2}$ , we classify the segment as turbulent.

*Author contributions.* RA did all analysis and wrote and edited the manuscript. CB provided guidance on the analysis and edited the  
450 manuscript.

*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 Paul Bui for providing high-rate MMS data for CRYSTAL-FACE and TC4, and to Rei Ueyama and Johnathon 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  
455 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  
460 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.
- Baumgardner, D., Jonsson, H., Dawson, W., O'Connor, D., and Newton, R.: The cloud, aerosol and precipitation spectrometer: a new instru-  
ment for cloud investigations, *Atmospheric Research*, 59-60, 251–264, [https://doi.org/https://doi.org/10.1016/S0169-8095\(01\)00119-3](https://doi.org/https://doi.org/10.1016/S0169-8095(01)00119-3),  
2001.
- Bramberger, M., Alexander, M. J., Davis, S., Podglajen, A., Hertzog, A., Kalnajs, L., Deshler, T., Goetz, J. D., and Khaykin, S.: First Super-  
465 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  
470 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  
475 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-*  
480 *Atmospheres*, 115, <https://doi.org/10.1029/2009jd012735>, 2010.
- 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.: Geosta-  
tionary 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,  
485 <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,  
490 <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 Atmo-  
spheric Sciences*, 52, 4159–4182, [https://doi.org/10.1175/1520-0469\(1995\)052<4159:dsatic>2.0.co;2](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/10.1029/95RG02097>, 1995.
- 495 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/10.1029/2004EO050002>, 2004.
- 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  
500 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.
- 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.
- 505 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., 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 TROPICAL 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.
- 510 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 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  
515 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.
- 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.,  
520 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-20-12569-2020>, 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.
- 525 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>, 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,  
530 1462–1477, <https://doi.org/10.1175/jtech1927.1>, 2006.



- May, R. D.: Open-path, near-infrared tunable diode laser spectrometer for atmospheric measurements of H<sub>2</sub>O, *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 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](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 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 particles, *Atmospheric Chemistry and Physics*, 18, 10 799–10 823, <https://doi.org/10.5194/acp-18-10799-2018>, 2018.
- Schoeberl, M. R., Jensen, E. J., and Woods, S.: Gravity waves amplify upper tropospheric dehydration by clouds, *Earth and Space Science*, 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:tmmstot>2.0.co;2](https://doi.org/10.1175/1520-0426(1990)007<0525:tmmstot>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.
- 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.
- 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.
- 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 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.,



- 570 Minnis, P., and Pickering, K. E.: Planning, implementation, and first results of the Tropical Composition, Cloud and Climate Coupling  
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.
- 575 Weinstock, E. M., Hints, 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  
of Scientific Instruments*, 65, 3544–3554, <https://doi.org/10.1063/1.1144536>, 1994.