

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

**Please provide feedback**

Please support the ScholarWorks@UMBC repository by emailing [scholarworks-group@umbc.edu](mailto:scholarworks-group@umbc.edu) and telling us what having access to this work means to you and why it's important to you. Thank you.

# A Modern Approach to Stability-Based Definition of the Tropopause

Emily N. Tinney,<sup>a</sup> Cameron R. Homeyer,<sup>a</sup> Lexy Elizalde,<sup>b</sup> Dale F. Hurst,<sup>c,d</sup>  
Anne M. Thompson,<sup>e,f</sup> Ryan M. Stauffer,<sup>e</sup> Holger Vömel,<sup>g</sup> and Henry B. Selkirk<sup>h,i</sup>

<sup>a</sup> *School of Meteorology, University of Oklahoma, Norman, OK, USA*

<sup>b</sup> *South Dakota School of Mines and Technology, Rapid City, SD, USA*

<sup>c</sup> *Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA*

<sup>d</sup> *NOAA Global Monitoring Laboratory, Boulder, CO, USA*

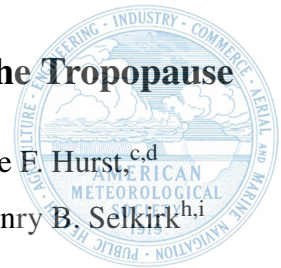
<sup>e</sup> *NASA Goddard Space Flight Center (GSFC), Greenbelt, MD, USA*

<sup>f</sup> *University of Maryland-Baltimore County/Joint Center for Environmental Systems Technology, Baltimore, MD, USA*

<sup>g</sup> *National Center for Atmospheric Research, Boulder, CO, USA*

<sup>h</sup> *Earth Science Division, NASA Headquarters, Washington, DC, USA*

<sup>i</sup> *Agile Decision Sciences, LLC, Beltsville, MD, USA*



*Corresponding author:* Emily N. Tinney, [emily.tinney@ou.edu](mailto:emily.tinney@ou.edu)

**Early Online Release:** This preliminary version has been accepted for publication in *Monthly Weather Review* cited, and has been assigned DOI 10.1175/MWR-D-22-0174.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

**ABSTRACT:** Definition of the tropopause has remained a focus of atmospheric science since its discovery near the beginning of the 20th century. Few universal definitions (those that can be reliably applied globally and to both common observations and numerical model output) exist and many definitions with unique limitations have been developed over the years. The most commonly used universal definition of the tropopause is the temperature lapse-rate definition established by the World Meteorological Organization (WMO) in 1957 (the LRT). Despite its widespread use, there are recurrent situations where the LRT definition fails to reliably identify the tropopause. Motivated by increased availability of coincident observations of stability and composition, this study seeks to re-examine the relationship between stability and composition change in the tropopause transition layer and identify areas for improvement in stability-based definition of the tropopause. In particular, long-term (40+ years) balloon observations of temperature, ozone, and water vapor from six locations across the globe are used to identify co-variability between several metrics of atmospheric stability and composition. We found that the vertical gradient of potential temperature is a superior stability metric to identify the greatest composition change in the tropopause transition layer, which we use to propose a new universally applicable potential temperature gradient tropopause (PTGT) definition. Application of the new definition to both observations and reanalysis output reveals that the PTGT largely agrees with the LRT, but more reliably identifies tropopause-level composition change when the two definitions differ greatly.

**SIGNIFICANCE STATEMENT:** In this study we provide a review of existing tropopause definitions (and their limitations) and investigate potential improvement in the definition of the tropopause using balloon-based observations of stability and atmospheric composition. This work is motivated by the need for correct identification of the tropopause to accurately assess upper troposphere lower stratosphere processes, which in turn has far reaching implications for our understanding of Earth's radiation budget and climate. The result of this research is the creation of a new, universally-applicable stability-based definition of the tropopause: the Potential Temperature Gradient Tropopause (PTGT).

## 1. Introduction

The tropopause is a complex transition layer between the upper troposphere and lower stratosphere (UTLS). Its definition continues to be an important and challenging task due to the assessment of many UTLS-relevant processes such as stratosphere–troposphere exchange being dependent on its location. For example, a tropopause that is identified too high can result in the false assessment of stratospheric air in the troposphere, and vice versa. Ultimately, it is the diverse dynamic, chemical, and radiative coupling between troposphere and stratosphere in the UTLS and the two-way exchange of air via stratosphere–troposphere exchange, which significantly impact Earth's radiation budget and climate, that motivate continued refinement in our approach to tropopause definition (Holton et al. 1995; Stohl et al. 2003; Gettelman et al. 2011; Banerjee et al. 2019).

Since its discovery in the early 20<sup>th</sup> century (see Hoinka (1997) for a review), various approaches to define the tropopause transition layer have been proposed. Most proposed definitions, for convenience or simplicity, define the tropopause as a boundary, despite increasing recognition that it is best characterized as a layer of depth ranging from tens of meters to several kilometers (e.g., Hoinka 1997; Hegglin et al. 2009; Homeyer et al. 2010; Tilmes et al. 2010; Pan et al. 2014). Some notable findings of global tropopause characteristics since its discovery include: i) tropopause altitudes are uniformly high in the tropics and uniformly low in the extratropics, with a discontinuity in altitude occurring near the subtropical jet that is known as the “tropopause break” (Palmén 1948; Danielsen 1959; Randel et al. 2007a), ii) the tropopause can become significantly deformed and difficult to identify in highly dynamic situations (and, often, stratosphere–troposphere exchange events) such as within tropopause folds driven by ageostrophic circulations near upper tropospheric

jets and in regions impacted by Rossby wave breaking that contain wide-spread vertical lamination of tropical upper troposphere air and extratropical lower stratosphere air (Reed 1955; Danielsen 1968; Shapiro 1980; Browell et al. 1987; Newman and Schoeberl 1995; Vaughan and Timmis 1998; Olsen et al. 2008; Pan et al. 2009), iii) the extratropical tropopause is characterized by a strong temperature inversion  $\sim 2\text{--}3$  km in depth, referred to as the “tropopause inversion layer” or TIL, which is driven by radiative processes and dynamic processes such as downwelling and column stretching/shrinking (Birner 2006; Randel et al. 2007b; Son and Polvani 2007; Peevey et al. 2014), and iv) long-term changes in tropopause altitude can be used as an indicator of global climate change (Seidel et al. 2001; Shepherd 2002; Santer et al. 2003; Seidel and Randel 2006; Xian and Homeyer 2019; Meng et al. 2021; Thompson et al. 2021). While dozens of tropopause definitions have been proposed over the years, they can be separated into three main categories: stability-based, dynamics-based, and composition-based definitions.

Stability-based definitions of the tropopause use profiles of temperature to determine the vertical stratification of the atmosphere. The original and most-used stability-based definition is that created by the World Meteorological Organization (WMO) in 1957. The temperature lapse-rate definition of the tropopause (LRT) is defined as “*the lowest level at which the lapse rate decreases to  $2^{\circ}\text{C km}^{-1}$  or less, provided also the average lapse rate between this level and all higher levels within 2 km does not exceed  $2^{\circ}\text{C km}^{-1}$* ” (World Meteorological Organization 1957). This method is one of few globally and universally-applicable definitions (i.e., it can be applied reliably to both balloon observations and model output), although its accuracy can vary seasonally, latitudinally, and under previously-outlined complex synoptic patterns such as tropopause folds and Rossby wave breaking events (Zängl and Hoinka 2001; Homeyer et al. 2010). The most significant and routine errors in tropopause identification using the LRT definition occur in polar regions during hemispheric winter and spring where a stable, near-isothermal middle troposphere and UTLS can prevent the LRT criteria from being met at a level below the lower-to-middle stratosphere, resulting in an erroneously high tropopause altitude (Zängl and Hoinka 2001). Near the subtropical jets and more spatially confined, lamination of tropical tropospheric air and extratropical stratospheric air can result in complex stability profiles that make tropopause definition using any criteria difficult, for which errors in LRT altitude are almost always realized as a high bias and are also unfortunately common (Homeyer et al. 2010). In the tropics, the cold point tropopause (i.e., the UTLS temperature

minimum) is a commonly-used definition, but is susceptible to bias from temperature fluctuations driven by wave activity (Kim and Alexander 2015). Additionally, the cold point tropopause cannot be used poleward of the subtropical jets due to decoupling between the temperature minimum in a profile and the dominant troposphere-stratosphere composition change (e.g., Highwood and Hoskins 1998). In evaluation of modeled convective mass transport, tropopause identification by a potential temperature gradient threshold performed similarly to the LRT, though its potential further use has not been fully explored (Maddox and Mullendore 2018). Other variations of stability-based definitions have also been created in recent years, often leveraging the Brunt–Väisälä frequency  $N$  in their identification (Homeyer et al. 2010; Gettelman and Wang 2015; Duran and Molinari 2019). Note that any stability-based definition applied to radiosonde observations can be sensitive to noise and measurement-related artifacts, such as solar heating of the sensor or adsorption of hydrometeors onto the sensor, but these are typically negligible for tropopause altitude definition.

Dynamic definitions of the tropopause are the most popular alternative to thermal (i.e. stability) definitions in the subtropics and extratropics. Potential vorticity (PV) is commonly used as the basis of a dynamic tropopause definition (Reed 1955; Holton et al. 1995; Kunz et al. 2011). PV is the product of static stability and absolute vorticity and is conserved in frictionless, adiabatic flows. While some studies use the gradient of PV along isentropes to identify the dynamic tropopause, it is most often prescribed as a surface of constant PV expressed as a potential vorticity unit, or PVU, where  $1 \text{ PVU} = 10^{-6} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$ . The PV threshold selected often ranges from  $\pm 1$ – $4$  PVU, with  $\pm 2$  PVU being most common, but the ideal threshold has been found to vary seasonally and by altitude from model analyses (Hoerling et al. 1991; Wernli and Bourqui 2002; Sprenger et al. 2003; Kunz et al. 2011; Škerlak et al. 2014). The PV-based dynamic tropopause cannot be applied in the tropics where the absolute vorticity approaches zero and PV surfaces become nearly vertical. However, the arguably greatest limitation of a PV-based dynamic definition of the tropopause is that it is largely gridded-data-based and therefore cannot be applied in many traditional observational studies without additional support of a gridded dataset. Alternative dynamic definitions exist based on trajectory calculations, called a “Lagrangian tropopause” (e.g., Berthet et al. 2007), but these are more computationally expensive and entirely model-based.

Lastly, composition-based definitions of the tropopause use the often dramatic composition change between troposphere and stratosphere to identify a chemical tropopause. Previous observa-

tional and modeling studies have leveraged the concentrations of various trace gases, such as—but not limited to—ozone ( $\text{O}_3$ ), water vapor ( $\text{H}_2\text{O}$ ), and carbon monoxide ( $\text{CO}$ ), due to their sharp gradients across the tropopause transition layer. These sharp gradients allow the compounds to be considered tracers of stratospheric ( $\text{O}_3$ ) and tropospheric ( $\text{H}_2\text{O}$ ,  $\text{CO}$ ) air. Past studies have either used profiles of one of these trace gases—predominantly  $\text{O}_3$ —or two trace gases to identify the characteristic composition change in the tropopause transition layer. For example, Bethan et al. (1996) used balloon observations of  $\text{O}_3$  at select sites around the world to produce an ‘ozone tropopause’ definition. For two trace gases, plots of coincident tropospheric and stratospheric tracers demonstrating tracer–tracer relationships have become a common technique for a chemical tropopause transition layer definition, with  $\text{O}_3$ – $\text{CO}$  and  $\text{O}_3$ – $\text{H}_2\text{O}$  relationships being most common (Fischer et al. 2000; Hoor et al. 2002; Zahn and Brenninkmeijer 2003; Pan et al. 2004). However, such chemical tropopause definitions can be sensitive to transient processes including stratosphere–troposphere exchange, are not able to be applied to common observations such as radiosondes and models without chemistry, and can require adjustment when applied to locations at a wide range of latitudes. Unique approaches to chemical tropopause definition have been explored using models, such as passive tracer concentrations with a tropospheric source (e.g., Prather et al. 2011), but are ultimately less common and also not applicable to observations.

Despite the wide range of approaches used to identify the tropopause in prior work, the LRT is often characterized as the most reliable definition due to both its universal applicability and superior ability to identify the approximate location of the layer of greatest composition change in the UTLS, both in tropical, high-tropopause regions and extratropical, low-tropopause regions (Gettelman et al. 2011; Pan et al. 2018). The common coincidence of the LRT with the layer of greatest composition change is remarkable, because the LRT definition was developed at a time where coincident profile observations of atmospheric temperature and composition were extremely rare. Recent work has demonstrated that other stability metrics may perform similarly to the LRT, but their potential use and applicability to a wide variety of environments and dynamic scenarios must be further investigated (Maddox and Mullendore 2018). Recognition of this fact and motivated by both the increasing availability of long-term, globally distributed profiles of observed atmospheric temperature and composition, especially  $\text{O}_3$  and  $\text{H}_2\text{O}$ , and the known common failure modes of the LRT definition, this study seeks to evaluate the relationship between

UTLS composition change and common metrics of atmospheric stability. In doing so, the following questions are addressed: 1) Does the temperature lapse rate best correspond to composition change in the UTLS? 2) Is it possible to improve upon the known limitations of the LRT definition by designing an alternative, universally applicable stability-based tropopause definition? Our focus in addressing these questions is on the conventional exercise of identifying a single tropopause level rather than the composition transition layer depth.

To examine relationships between UTLS composition change and stability, 12-40+ years of balloon-based profile observations from six locations ranging in latitude from the high Arctic to the South Pole are used in this study. These observations include coincident measurements of temperature,  $O_3$ , and (in some cases)  $H_2O$ . The observations are first used to statistically examine the relationship between UTLS composition ( $O_3$  only) and three commonly-used metrics of atmospheric stability: the temperature lapse rate, vertical gradient of potential temperature, and  $N$ . Following this analysis, a new stability-based tropopause definition is proposed, applied to profile observations and reanalysis output in comparison to the LRT, and evaluated through tropopause-relative analysis of  $O_3$  and  $O_3$ – $H_2O$  tracer-tracer diagrams.

## 2. Data

### *a. Balloon Observations*

High-resolution balloon observations of temperature,  $O_3$ , and  $H_2O$  used in this study were obtained from NOAA's Earth System Research Laboratories (ESRL) Global Monitoring Laboratory (GML) online archive (NOAA 2021, 5 out of 6 sites) and from NASA's Southern Hemisphere ADDitional OZonesondes (SHADOZ) and Network for the Detection of Atmospheric Composition Change (NDACC) archives (NASA 2022a,b, for the Costa Rica site only). These data include traditional radiosonde observations of air temperature, pressure, and humidity in all cases and one or both of the following instruments, depending on the flight: an electrochemical concentration cell (ECC) ozonesonde for measuring  $O_3$  with an accuracy of  $\pm 5\%$  and a precision of  $\pm 3$ – $4\%$  (Witte et al. 2017; Thompson et al. 2017; Witte et al. 2018; Sterling et al. 2018) and/or a NOAA frost point hygrometer (FPH; Hurst et al. 2011) or cryogenic frost point hygrometer (CFH; Vömel et al. 2007) for measuring tropospheric and stratospheric  $H_2O$  with total uncertainties of  $\pm 10\%$  and  $\pm 6\%$ , respectively (Hall et al. 2016; Vömel et al. 2016). Note that while the radiosonde data include



humidity measurements for all profiles, these data are not suitable for detailed analysis of H<sub>2</sub>O composition in the UTLS given the well-established need for corrections of measurement biases and slow sensor response times (e.g., Miloshevich et al. 2004). Regardless of the composition instrumentation used, every balloon carried a radiosonde and most reached altitudes >30 km, providing measurements of meteorology and composition throughout the troposphere and lower stratosphere. For ECC and radiosonde flights, O<sub>3</sub> and temperature measurements reported at a vertical data spacing of 5–10 m were averaged in 100-m vertical layers, while for flights that also included a FPH/CFH, the 5–10 m-spaced measurements of H<sub>2</sub>O, O<sub>3</sub> and temperature were averaged in 250-m vertical layers. Many of the flights that include both FPH/CFH and ECC instruments are also available in the ECC-only datasets. In such cases, only the multi-instrument 250-m layer data is retained for analysis to prevent double counting. For consistency, all 250-m layer data are linearly interpolated to 100-m vertical grid spacing prior to analysis. For either dataset, only the ascent (i.e., generally higher quality) observations are used for analysis.

To explore relationships between stability and composition across a wide range of latitudes, five NOAA ESRL GML balloon sounding sites and one NASA site were selected for analysis in this study, which are: Summit Station, Greenland (72.58°N); Boulder, Colorado (39.95°N); Hilo, Hawaii (19.72°N); Lauder, New Zealand (45.04°S); South Pole Station, Antarctica (90.00°S); and San José, Costa Rica (~10°N), respectively. Latitude was used as the basis for balloon sounding site selection because it is the leading factor of global variability in tropopause altitude and UTLS composition (e.g., Gettelman et al. 2011). The Boulder, Hilo, San José and Lauder balloon sounding sites include measurements with both ECC and FPH/CFH instruments, while the Summit and South Pole sites are ECC(O<sub>3</sub>)-only. Figure 1 shows a map of these balloon sounding sites and Table 1 summarizes the number of flights from each site at least 18 km in depth from the surface that are used for analysis in this study.

Figures 2 and 3 summarize data availability for all balloon sounding sites in detail as a function of year and season, respectively. The ozonesonde record is the most extensive of the balloon flights, extending back to the late 1960s for the Boulder and South Pole sites. The Boulder, Hilo, and South Pole stations have the longest data records, with sustained observations from the late-1970s or mid-1980s to the Present. Flights from Boulder (2003–Present), Hilo (2005–Present), San José (2005–Present), and Lauder (2011–Present) that include all three instruments (radiosonde, ECC,

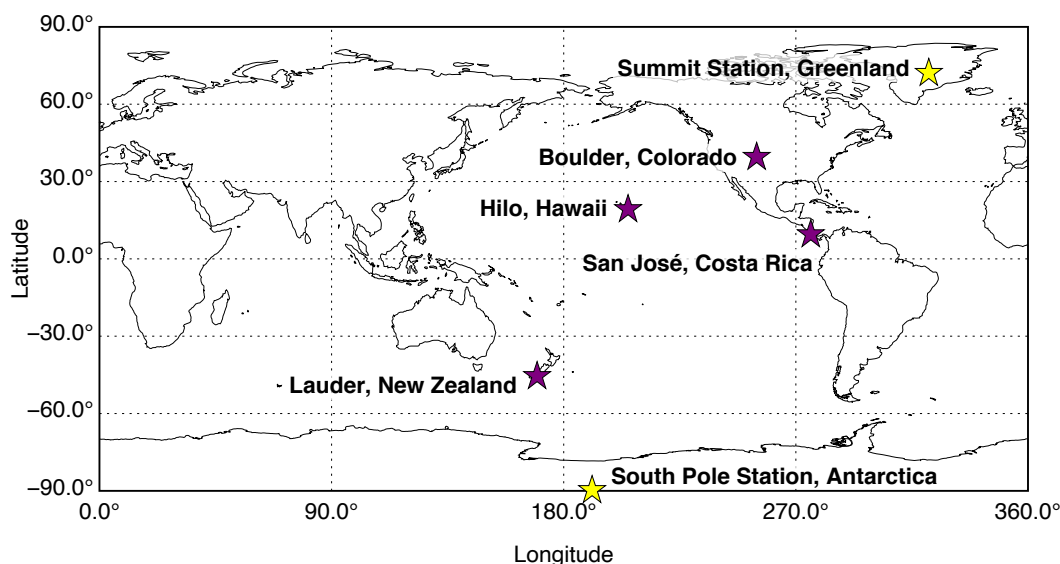


FIG. 1. Sites of the balloon observations used in this study. Yellow stars indicate sites with ozonesonde-only observations and purple stars indicate sites that also have water vapor data.

TABLE 1. The total number of balloon observations at least 18 km in depth by type (ECC or O<sub>3</sub>-only, FPH/CFH or H<sub>2</sub>O-only, and both) and location.

Observation Site	O <sub>3</sub> -only	H <sub>2</sub> O-only	O <sub>3</sub> & H <sub>2</sub> O
Boulder, Colorado	1678	284	241
Hilo, Hawaii	1612	5	126
Lauder, New Zealand	0	5	181
Summit Station, Greenland	527	0	0
San José, Costa Rica	402	1	221
South Pole Station, Antarctica	2134	0	0

and FPH/CFH) amount to ~10 seasonally-distributed flights per year for each site. The Boulder FPH record began in 1980, making it the world's longest record of UTLS H<sub>2</sub>O measurements. There are varying degrees of seasonality in the availability of ozonesonde data for the balloon sounding sites, with the greatest seasonality at the South Pole and Summit Stations, where there are maxima in the number of observations during the seasons when substantial stratospheric O<sub>3</sub> depletion occurs (in September–November and March–May, respectively).

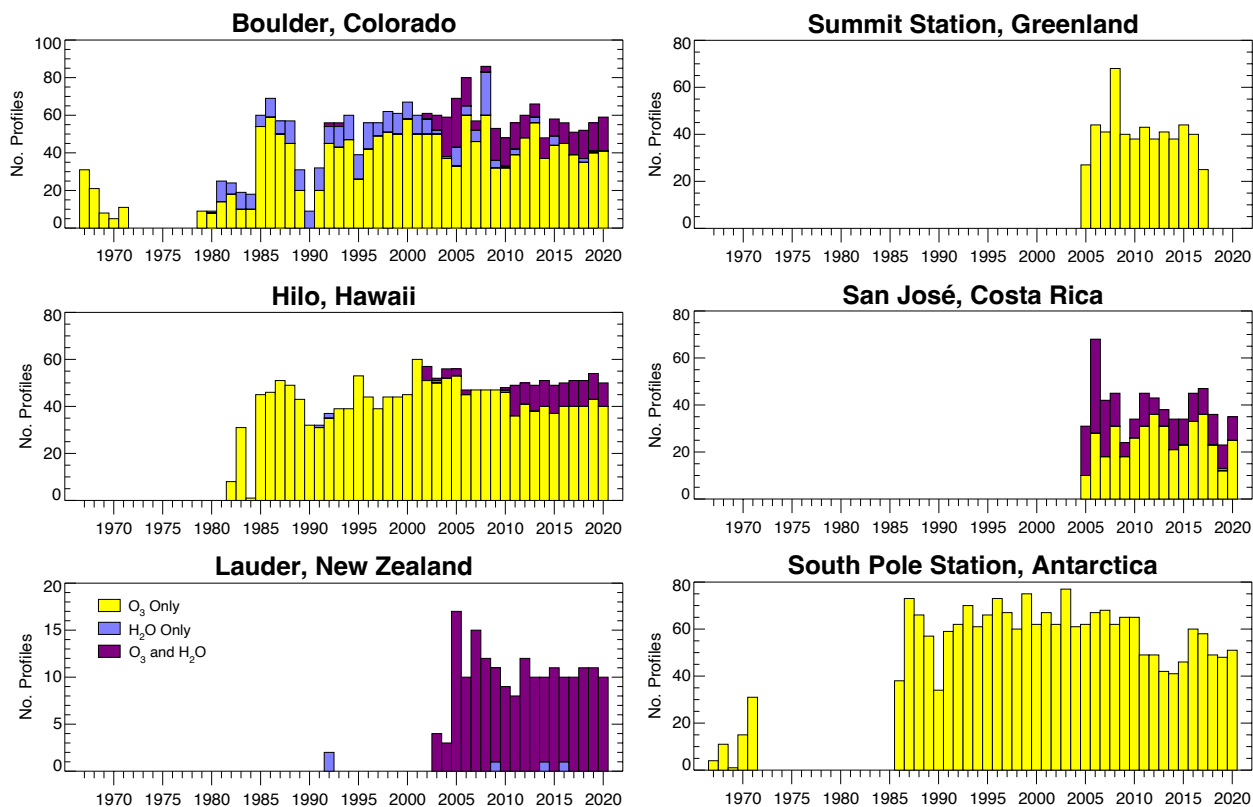


FIG. 2. Stacked bar charts indicating the number of balloon profiles reaching at least 18 km above the surface, as a function of year for the 6 balloon sounding sites used. Yellow bars indicate the total number of ozonesonde-only observations, blue the total number of FPH/CFH-only observations, and purple the total number of balloon observations with both instruments. Note that the ordinate varies by observation site.

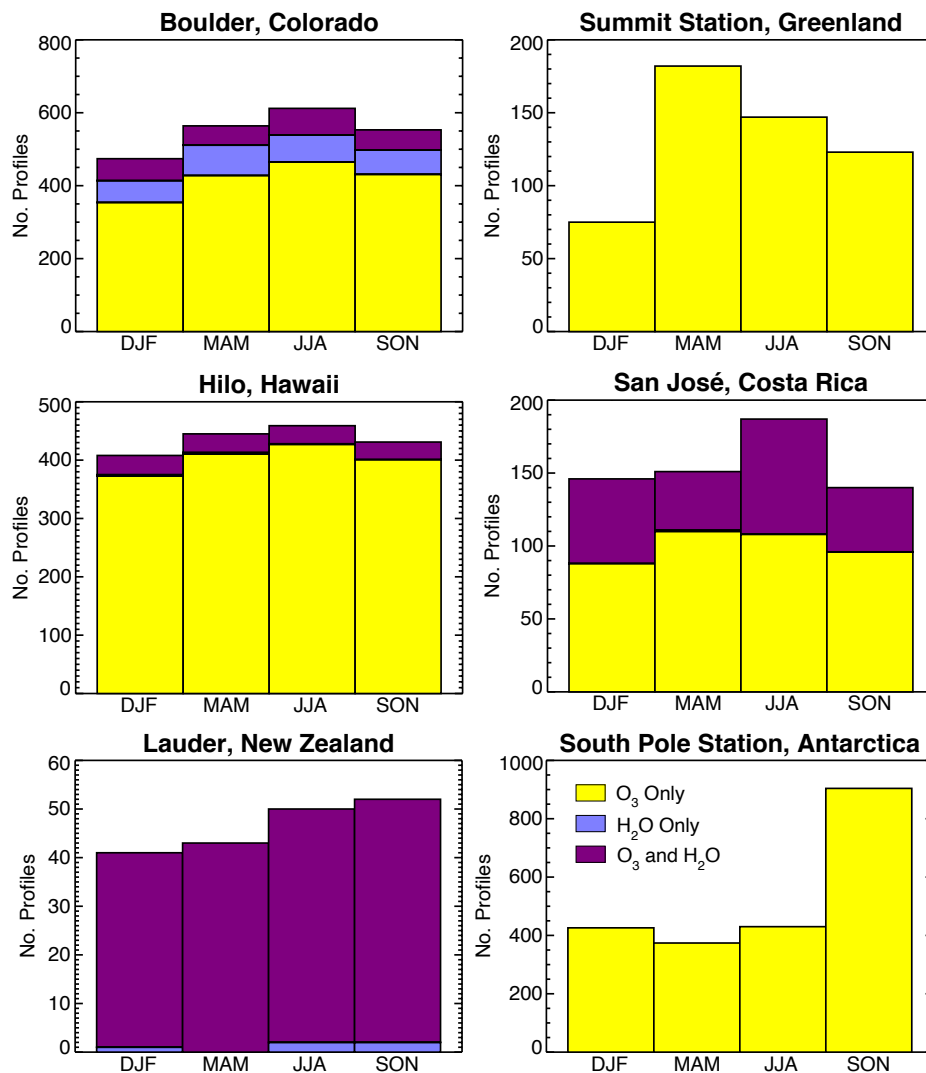


FIG. 3. As in Fig. 2, but by season. Note that the disproportionate number of ECC soundings at South Pole Station during SON is due to an increased frequency of soundings each year during the formation and evolution of the polar ozone holes.

### *b. Reanalysis Output*

To demonstrate global application of tropopause definitions in this study, output from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) is used. MERRA-2 is available every 3 hours from 1979 to the present at a horizontal grid spacing of  $0.625^\circ \times 0.5^\circ$  longitude-latitude and 72 vertical levels, with a vertical grid spacing of  $\sim 1100$  m in the UTLS (Gelaro et al. 2017). Four individual 0000 UTC analysis times during the midpoint of each season of a single year are used here, to demonstrate the seasonality and variability of global tropopause definition and compare performance of the new stability-based tropopause definition introduced here to the lapse-rate and dynamic (i.e. PV-based) tropopauses. Tropopause altitudes are calculated for each reanalysis profile after first interpolating temperature to a regular 100-m vertical grid using cubic splines.

## **3. Results**

### *a. Composition-Stability Relationships*

As outlined in the Introduction, a unique opportunity afforded by modern balloon observations that include measurements of atmospheric composition is that the relationship between composition change in the tropopause transition layer and stability can be extensively evaluated. Here, we explore this using the entire record of balloon observations introduced in Section 2a and multiple metrics of static stability. In particular, we focus attention on three easily computed conventional metrics that can be obtained from any balloon radiosonde profile or model output. Namely, we investigate vertical gradients of temperature ( $\partial T/\partial z$ ), potential temperature  $\theta$  ( $\partial \theta/\partial z$ ), and the Brunt-Väisälä frequency  $N$ , given as:

$$N = \sqrt{\frac{g}{\theta} \frac{\partial \theta}{\partial z}} = \sqrt{\frac{g}{T} \left( \frac{\partial T}{\partial z} + \Gamma_d \right)} \quad (1)$$

where  $g$  is the gravitational constant ( $9.8 \text{ m s}^{-2}$ ) and  $\Gamma_d$  is the dry adiabatic lapse rate ( $9.8 \text{ K km}^{-1}$ ). Note that  $\partial T/\partial z$  is simply the opposite sign of the environmental temperature lapse rate  $\Gamma$ , which the LRT definition is based upon. These three stability metrics have been widely used to both quantitatively and qualitatively depict stability change associated with the tropopause transition layer, but only  $\partial T/\partial z$  (or  $\Gamma$ ) has been routinely used for tropopause definition in the past.

Three example profiles of  $T$ ,  $\theta$ ,  $O_3$ , and the stability metrics are given in Figure 4. The three profiles shown encompass variability in time of year, tropopause altitude, and location, with a high-latitude example from the South Pole Station during austral spring with substantial stratospheric  $O_3$  depletion, a midlatitude example from Boulder, Colorado, and a low-latitude example from San José, Costa Rica. Characteristic profiles of  $T$  and  $\theta$  from each location indicate the variability in UTLS structure that make reliable, globally-applicable identification of the tropopause difficult. Namely, abrupt changes in  $T$  and  $\theta$  and concomitant step-like changes in the stability metrics are found in the Boulder and San José examples, near altitudes of 14 km and 17 km, respectively. Sharp changes in  $O_3$  composition clearly accompany the stability changes in those cases, which is common in environments with such strong stability-based tropopause definition. However, in the South Pole Station example, the tropopause transition in stability is more gradual, despite a pronounced  $O_3$  composition change indicating a tropopause transition layer near 9 km altitude. The three stability metrics show a weak transition that is far less pronounced than in the other example profiles, but there is no pronounced minimum in the temperature profile near the tropopause.

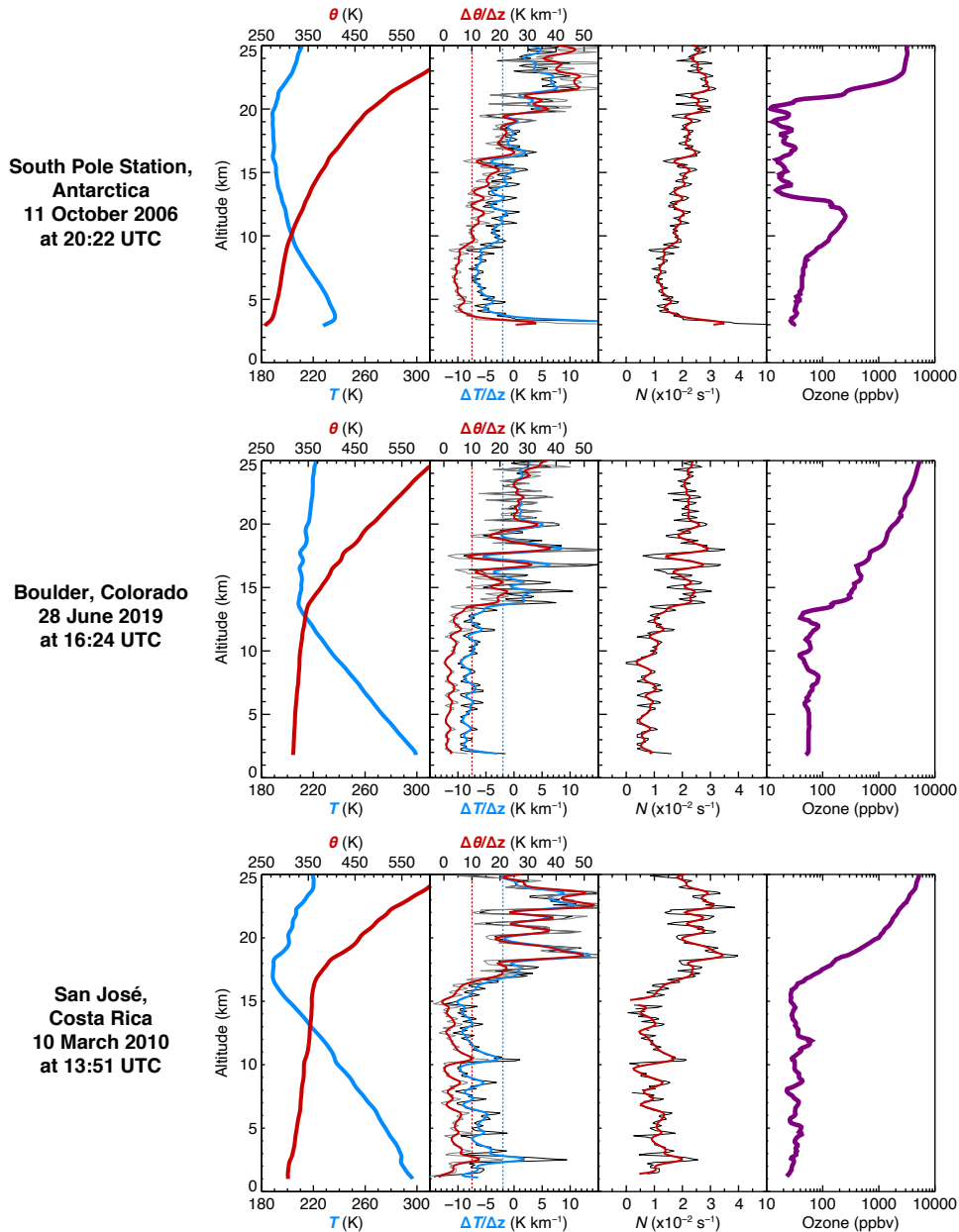


FIG. 4. Example ozonesonde profiles of (from left to right) temperature ( $T$ ; blue) and potential temperature ( $\theta$ ; red), vertical gradients of  $T$  ( $\Delta T/\Delta z$ ) and  $\theta$  ( $\Delta \theta/\Delta z$ ), the Brunt-Väisälä frequency  $N$ , and ozone volume mixing ratio (purple). For the  $T$ ,  $\theta$ , and stability profiles, the raw data are shown by the black and gray profiles in the background, while colored lines show  $2\text{-}\sigma$  Gaussian smoothing of the raw  $T$  and  $\theta$  data. In the second panel, vertical blue and red dashed lines are given at constant values of  $\Delta T/\Delta z = -2 \text{ K km}^{-1}$  and  $\Delta \theta/\Delta z = 10 \text{ K km}^{-1}$ , respectively. The profiles are sourced from (top) South Pole Station, Antarctica on 11 October 2006 at 20:22 UTC, (middle) Boulder, Colorado on 28 June 2019 at 16:24 UTC, and (bottom) San José, Costa Rica on 10 March 2010 at 13:51 UTC.

We consider two main characteristics of stability profiles that aid in establishing a clear relationship between a stability metric and UTLS O<sub>3</sub> composition change: i) increased troposphere-stratosphere contrast of the stability metric (i.e., the relative difference between common low-stability tropospheric and high-stability stratospheric values) co-located with the strong vertical gradient in O<sub>3</sub> and ii) decreased fluctuations of the stability metric that are generally unrelated with the tropopause transition layer. For the example profiles in Figure 4,  $\partial\theta/\partial z$  appears to provide the greatest contrast of the stability metrics, but all metrics are characterized by large variability (especially at higher altitudes - i.e., throughout the stratosphere). This large variability stems from a range of thermodynamic and dynamic processes such as vertically propagating high-frequency (e.g., gravity) waves and complex large-scale lamination of stable layers in latitude from Rossby wave breaking and tropopause folding (e.g., Holton et al. 1995; Stohl et al. 2003; Gettelman et al. 2011). To leverage stability transitions for tropopause definition, measuring changes over substantial depths or smoothing of the underlying  $T$  and  $\theta$  data is sensible. Figure 4 shows the resulting reduction in variability of the stability metrics after application of  $2\text{-}\sigma$  Gaussian smoothing to the 100-m  $T$  and  $\theta$  profiles, which spans a layer  $\sim 1$  km deep in total and has a full width at half maximum of  $\sim 500$  m for the Gaussian weights. Nevertheless, some broad differences among the three stability metrics can be inferred from comparison in these three example profiles: i)  $\partial\theta/\partial z$  provides the greatest troposphere-stratosphere contrast of the three metrics, and ii) of the smoothed metrics,  $N$  provides the least variance in magnitude with changing altitude unrelated with the transition layer. These differences and their relationship to troposphere-stratosphere O<sub>3</sub> composition change are robustly examined in the statistical analysis that follows.

Figure 5 shows two-dimensional frequency distributions of joint O<sub>3</sub> and stability observations for all balloon observations from each site. This analysis is limited to altitudes between 5 and 20 km (5 and 25 km for San José) to focus on observations from the UTLS and prevent contamination from stable planetary boundary layer inversions. Most joint frequency distributions reveal a bimodal pattern, with frequent observations of low O<sub>3</sub> ( $\lesssim 100$  ppbv) at low stability and frequent observations of high O<sub>3</sub> ( $\gtrsim 100$  ppbv) at high stability. The degree of bimodality and 2-D separation of the low and high modes varies by location and stability metric. For low latitude, high tropopause altitude stations (Hilo and San José), there is weaker bimodality and an overall frequency bias toward the low mode due to more of the analyzed layer being in the troposphere. The bimodality increases



with latitude for the remaining stations, reflecting a more equal population of tropospheric and stratospheric air in the contributing observations as a result of the transition to both lower tropopause altitude, and downward transport of high-O<sub>3</sub> air in the midlatitude lower stratosphere, which enhances the near-tropopause gradient in composition and thereby increases O<sub>3</sub> separation of the modes.

Each of the stability parameters offer strong separation in stability magnitudes for the two high-frequency modes, which is indicated by lesser overlap of high- and low-O<sub>3</sub> modes in stability space (Fig. 5), though identification of a tropopause transition layer via stability is challenging for any metric since composition and stability are not always positively related (e.g., at some point there is no relation for air confined to the troposphere or stratosphere and away from the tropopause). However, the stability parameters differ in the consistency of the value at which this separation occurs in different environments. In Figure 5, somewhat arbitrary gray horizontal lines are drawn for each stability metric at a constant value across all stations:  $-2 \text{ K km}^{-1}$  for  $\Delta T/\Delta z$  (i.e., the value consistent with the LRT),  $10 \text{ K km}^{-1}$  for  $\Delta\theta/\Delta z$ , and  $0.017 \text{ s}^{-1}$  for  $N$ . The  $-2 \text{ K km}^{-1}$   $\Delta T/\Delta z$  threshold corresponds relatively well with the location of high-frequency mode separation for Summit Station, Hilo, San José, and Lauder, but overlaps with air with stratospheric O<sub>3</sub> characteristics for Boulder and South Pole Station. Alternatively, both the  $\Delta\theta/\Delta z$  and  $N$  thresholds perform relatively well at all locations. A constant  $\Delta\theta/\Delta z$  threshold, specifically, seems the most appropriate metric for identification of the transition at all stations due to minimized overlap of the high-frequency modes. The consistency of a single  $\Delta\theta/\Delta z$  value ( $10 \text{ K km}^{-1}$  in this case) co-located with the UTLS composition transition in a variety of tropopause environments indicates that this metric provides the most promising opportunity for discriminating between troposphere and stratosphere layers on a global scale. The analysis in Figure 4 is consistent with this conclusion, where a  $10 \text{ K km}^{-1}$   $\partial\theta/\partial z$  threshold is reached at the location of the strong O<sub>3</sub> gradient at each station.

Why is it that the three stability metrics vary in the consistency of their relationship to composition change across a variety of environments? To investigate this, relationships between  $\partial\theta/\partial z$  and the remaining stability metrics as a function of altitude are shown in Figure 6. The relationship between  $\partial\theta/\partial z$  and  $\partial T/\partial z$  is modified by pressure: a  $-2 \text{ K km}^{-1}$   $\partial T/\partial z$  threshold corresponds with smaller  $\partial\theta/\partial z$  values at higher pressures (lower altitudes; e.g.,  $\sim 10 \text{ K km}^{-1}$  at 300 hPa) and

larger values at lower pressures (higher altitudes; e.g.,  $\sim 14 \text{ K km}^{-1}$  at 100 hPa). Similarly, the relationship between  $\partial\theta/\partial z$  and  $N$  is modified by  $\theta$ , where lower  $\theta$  corresponds to lower altitudes and vice versa. This vertical variation of the relationship between stability metrics demonstrates why tropopause identification based on them can differ. Ultimately, the composition-stability analysis shown in Figure 5 indicates that  $\partial\theta/\partial z$  is the most consistent at demarcating the UTLS composition change using a single threshold value. Motivated by the above results, we introduce a new  $\partial\theta/\partial z$ -based tropopause definition in the following subsection.

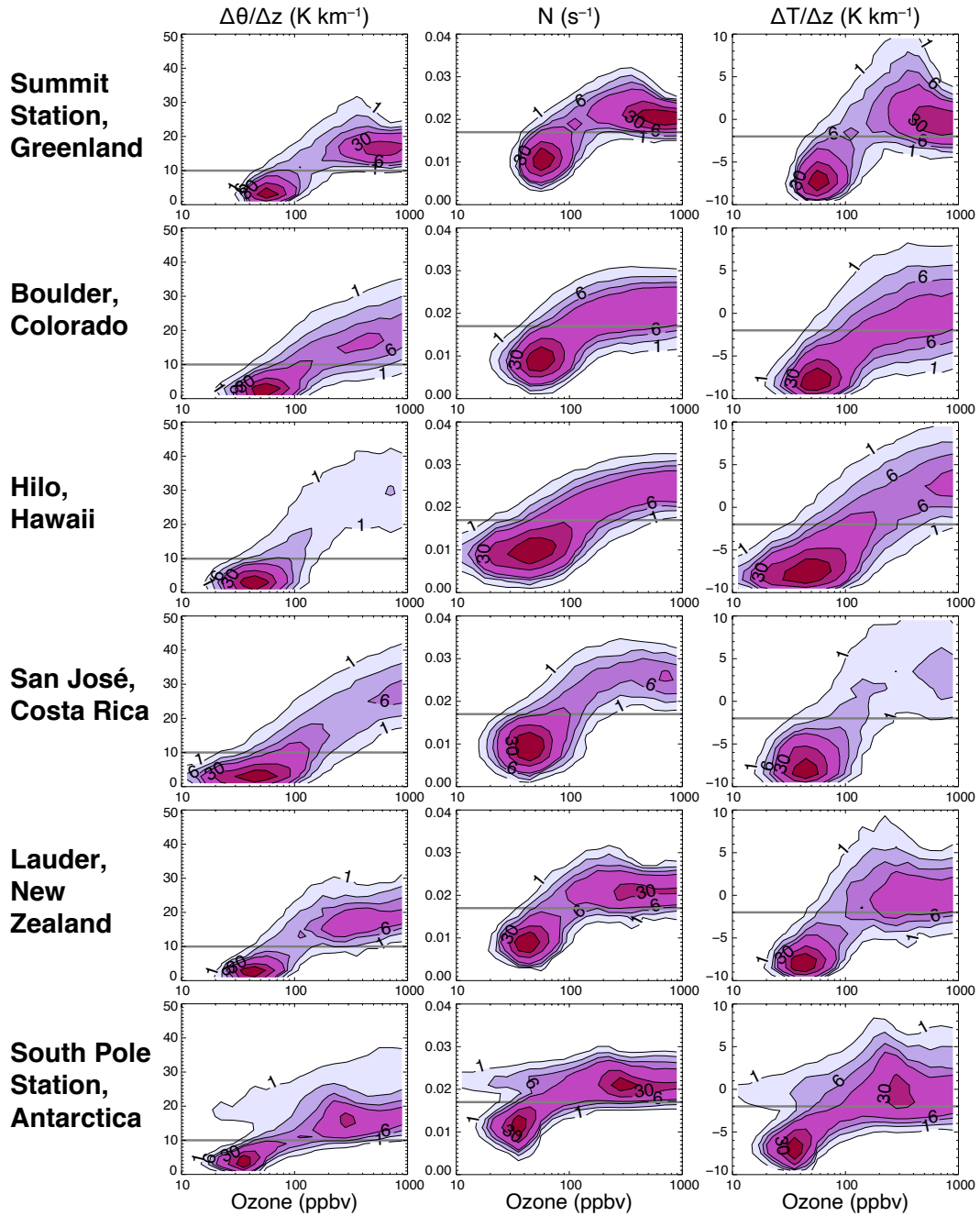


FIG. 5. Two-dimensional frequency distributions of ozone vs. (left) potential temperature gradient  $\Delta\theta/\Delta z$ , (middle) Brunt-Väisälä frequency  $N$ , and (right) temperature gradient  $\Delta T/\Delta z$  for all measurements from each observation site. Ozone is placed into logarithmic bins of size 0.1, while  $\Delta\theta/\Delta z$ ,  $\Delta T/\Delta z$ , and  $N$  are placed into bins of size  $2 \text{ K km}^{-1}$ ,  $1 \text{ K km}^{-1}$ , and  $0.002 \text{ s}^{-1}$ , respectively. Contours represent the frequency of observations relative to the maximum value within any 2-D bin in each panel, with shading at 1, 3, 6, 10, 30, and 60%. Gray horizontal lines are given in each panel at constant values of  $\Delta\theta/\Delta z = 10 \text{ K km}^{-1}$ ,  $N = 0.017 \text{ s}^{-1}$ , and  $\Delta T/\Delta z = -2 \text{ K km}^{-1}$ , for reference.

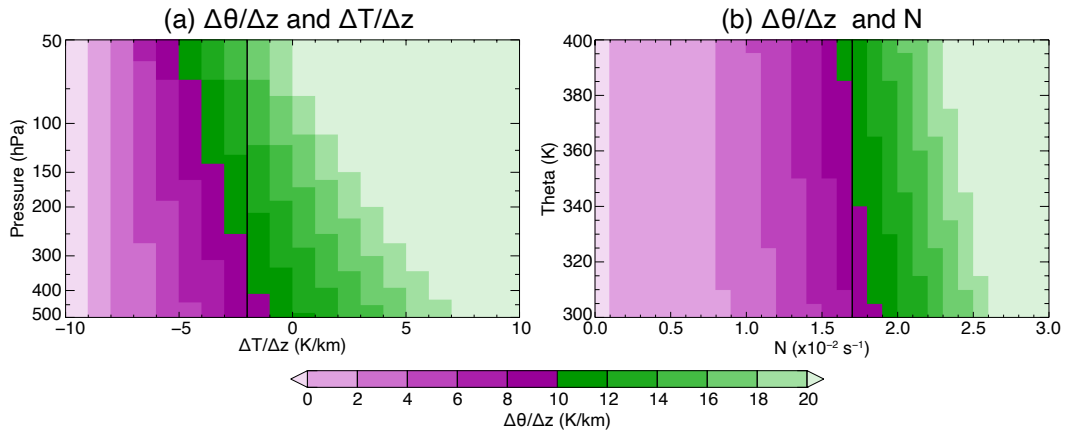


FIG. 6. Values of  $\Delta\theta/\Delta z$  as a function of (a)  $\Delta T/\Delta z$  and pressure, and (b)  $N$  and  $\theta$ . Black vertical lines are given at constant values of (a)  $\Delta T/\Delta z = -2$  K and (b)  $N = 0.017$  s $^{-1}$ .

### b. The Potential Temperature Gradient Tropopause Algorithm

Based on the finding from statistical analysis of balloon observations that the vertical gradient of potential temperature ( $\partial\theta/\partial z$ ) provides the greatest discrimination between air masses with tropospheric or stratospheric characteristics, an algorithm to identify a potential temperature gradient tropopause (hereafter PTGT) was sought. The goals of this new algorithm were to: 1) provide a reliable technique for tropopause altitude identification that captured the typical step-wise change in the magnitude of  $\partial\theta/\partial z$  (and more generally, static stability) from troposphere to stratosphere, 2) to improve stability-based identification of the tropopause in instances where the conventional LRT definition fails, and 3) to better coincide with the sharp composition change commonly observed in the tropopause transition layer. After rigorous tests and evaluation of mathematical approaches to this problem, the most reliable and simplest approach we developed is modeled after the LRT definition.

The PTGT definition is as follows: i) the first tropopause is defined as the lowest level at which the potential temperature gradient increases to  $10 \text{ K km}^{-1}$ , provided also that the potential temperature gradient between this level and all higher levels within 2 km does not fall below  $10 \text{ K km}^{-1}$ , and ii) if above the first tropopause, the potential temperature gradient between any level and all higher levels within 1 km falls below  $10 \text{ K km}^{-1}$ , then a second tropopause may be defined as done for the first tropopause, but using a potential temperature gradient threshold of  $15 \text{ K km}^{-1}$ . The 10 and  $15 \text{ K km}^{-1}$  thresholds were selected based on sensitivity tests, with the ultimate goal being

consistent performance at all latitudes. These  $\partial\theta/\partial z$  thresholds also commonly coincide with composition change observed near the tropopause and in complex transport events such as Rossby wave breaking (e.g., Pan et al. 2009).

Applying this definition reliably and as intended by the authors requires that data are on a regular altitude grid and that forward (i.e., upward) finite differencing is used for computing the potential temperature gradient, as implied by the language in the definition (again, consistent with the LRT). For convenience, the authors' code for the PTGT algorithm and the LRT algorithm is provided in multiple programming languages as Supplemental Material. While most tropopause definitions do not outline the utility and/or necessity of requiring all profile data to conform to a uniform vertical grid, doing so is advantageous to reliable and consistent application of any tropopause definition since nearly all algorithms attempt to identify robust changes in some variable — typically temperature— over a finite depth. To formalize this, we emphasize here that a necessary step to successful implementation of any tropopause algorithm be interpolation of profile data to a uniform grid (100-m data spacing is recommended for most datasets), which we include in the provided code. Note that for temperature, interpolation via cubic splines is often preferred to minimize error in tropopause identification (e.g., Hoffmann and Spang 2022).

### *c. Evaluation of the Tropopause Algorithm*

#### 1) APPLICATION TO OBSERVATIONS

Figure 7 shows the results of applying LRT and PTGT tropopause definitions to the example profiles previously shown in Figure 4. These examples demonstrate near perfect agreement of the two tropopause definitions in the Boulder and San José profiles where the stability transition is sharp and large disagreement in the South Pole Station example where the transition is weaker. Namely, in the South Pole Station profile, the PTGT definition identifies the subtle, lower altitude transition in stability that lies in close proximity to the pronounced O<sub>3</sub> increase near 9 km, indicative of the transition from troposphere to stratosphere air. Alternatively, the LRT definition is biased nearly 5 km high, identifying an arbitrary level that satisfies the lapse-rate criteria and falls within the broad layer of stratospheric O<sub>3</sub> depletion found ~12–22 km in this example. As outlined in Section 1, such high bias in LRT altitude in the polar regions was one motivating factor for revisiting stability-based tropopause definition in this study due to its unfortunately common occurrence

during winter and early spring, which will ultimately introduce significant error to any analysis that relies on accurate identification of the tropopause (e.g., stratosphere–troposphere exchange). As will be shown further below, the improvement in tropopause identification given by the new PTGT definition is not unique to this example.

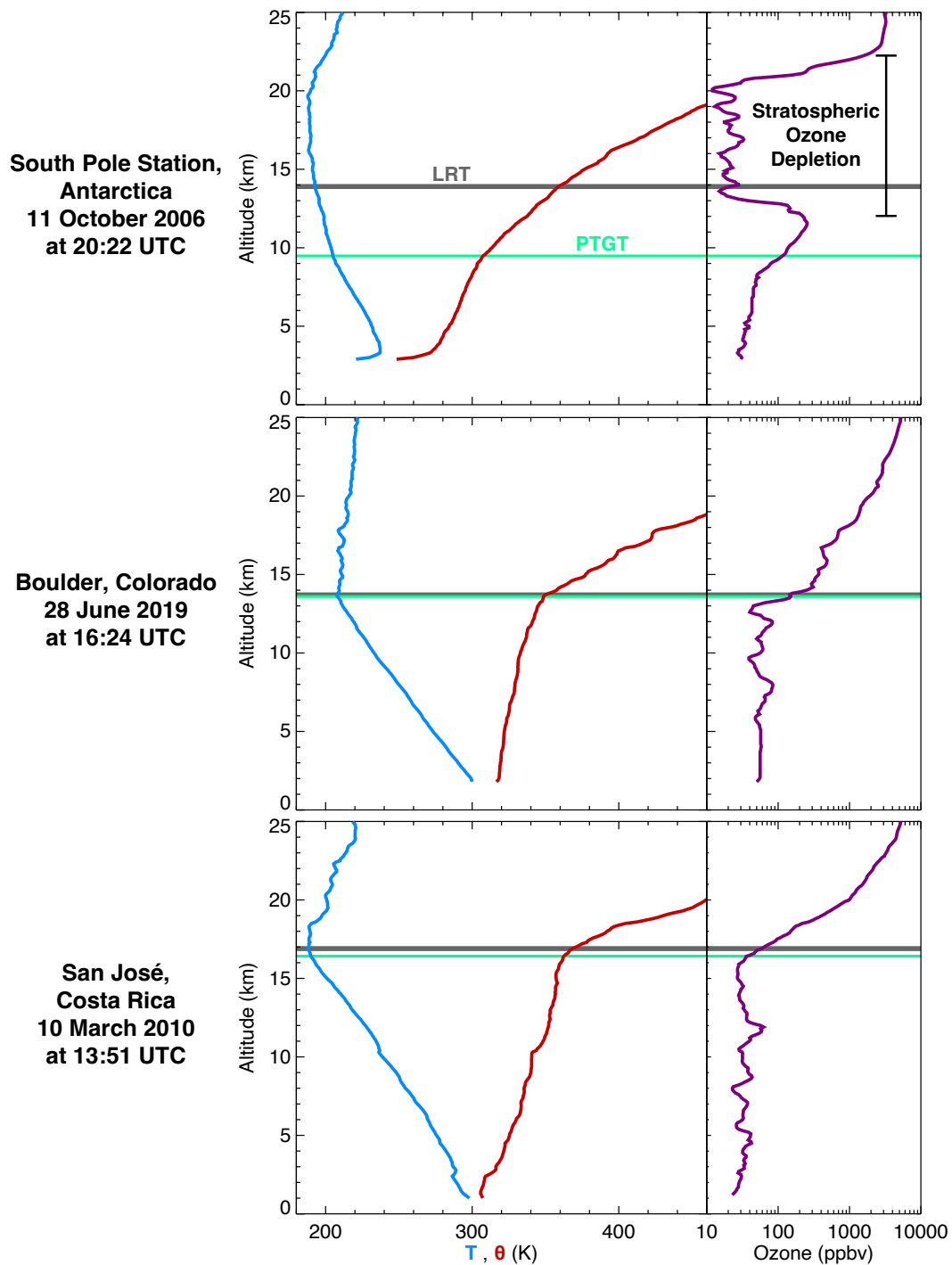


FIG. 7. For the same profiles in Fig. 4: (left) profiles of temperature ( $T$ ; blue) and potential temperature ( $\theta$ ; red) and (right) profiles of ozone volume mixing ratio. Altitudes of objectively diagnosed lapse-rate tropopauses (LRT; gray) and potential temperature gradient tropopauses (PTGT; green) are superimposed.

Figure 8 shows scatterplots comparing PTGT and LRT altitudes for all of the balloon observations. The scatterplots are densely grouped along a 1-to-1 line at all locations, indicating frequent agreement between the two definitions. When the PTGT and LRT disagree, the PTGT is almost always identified at a lower altitude. Lauder and Summit Station show the best agreement among the six stations, with only 4.8% and 1.7% of profiles containing differences larger than 1 km, respectively. For the remaining balloon sounding sites (Boulder, Hilo, San José, and South Pole Station), there are substantial occurrences of  $PTGT < LRT$ . In total, differences between PTGT and LRT larger than 1 km are found for 21.8% of the soundings at Boulder, 38.6% of the soundings at Hilo, 29.0% at San José, and 26.5% at South Pole Station. The frequency of large differences at South Pole Station is likely biased high since 40% of the observations are made in SON when the LRT is frequently biased high (Fig. 3). To evaluate the consequences of this difference in tropopause definition, we focus attention on the observed composition change relative to the diagnosed LRT and PTGT heights in the analysis that follows.



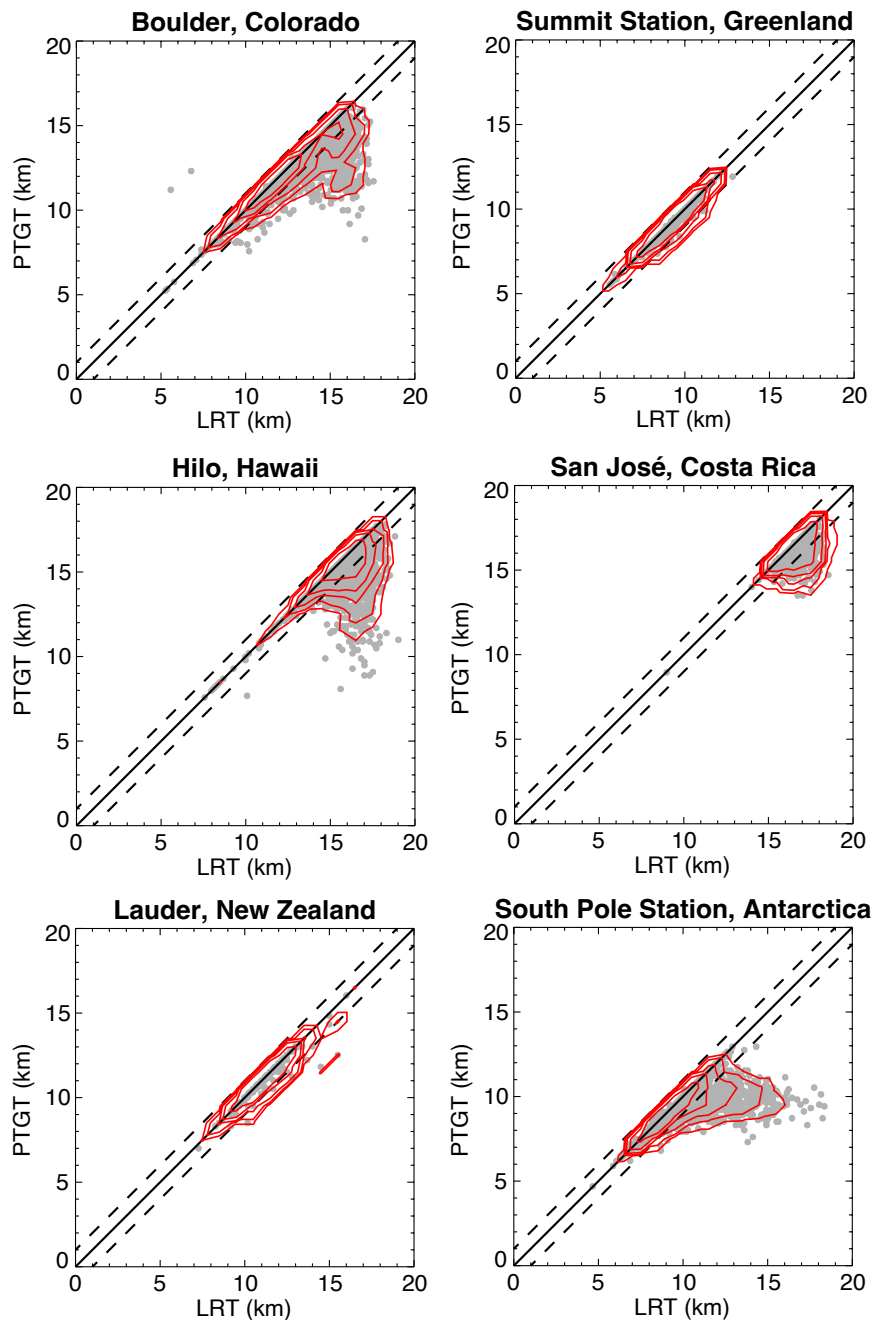


FIG. 8. Scatterplots comparing PTGT and LRT altitudes for all balloon observations used in this study, separated by observation site. The solid black lines are 1-to-1 lines and the dashed black lines show the boundaries of  $\pm 1$  km altitude differences. Red contours indicate the density of observations in 1-km  $\times$  1-km bins as the fraction of the total number of profiles, doubling in magnitude from 0.5 to 8 % (i.e., 0.5, 1, 2, 4, and 8 %).

From a composition perspective, O<sub>3</sub> concentrations relative to an appropriately identified tropopause will be uniformly low in the troposphere in most cases, with a sharp increase in O<sub>3</sub> within the tropopause transition layer and increasingly high values with increasing altitude in the lower stratosphere. Such analysis is featured in Figure 9, which shows tropopause-relative profiles of O<sub>3</sub>,  $\theta$ , and temperature for each station. The vertical structure and median O<sub>3</sub>,  $\theta$ , and temperature change across the LRT and PTGT are nearly indistinguishable for Lauder and Summit Station. The low spread between 10<sup>th</sup> and 90<sup>th</sup> percentile values also suggests consistent tropopause identifications for both definitions relative to O<sub>3</sub> concentrations at these locations, although this is also likely a reflection of their limited sampling (Figs. 3 and 8). Alternatively, tropopause-relative O<sub>3</sub> concentrations for Boulder, Hilo, San José, and South Pole Station exhibit notable differences for the PTGT and LRT definitions.

For Boulder, while PTGT-relative concentrations of O<sub>3</sub> remain low ( $\leq 90$  ppbv) throughout the troposphere (i.e., below-tropopause altitudes), O<sub>3</sub> concentrations begin to increase  $\sim 2$  km below the LRT. This leads to a slightly weaker O<sub>3</sub> gradient across the tropopause, and higher LRT-relative concentrations throughout the stratosphere relative to PTGT-relative concentrations. These differences reveal that the LRT definition is more frequently identifying the tropopause near the top of the transition layer, while the PTGT is identifying the tropopause near the base of the transition layer. A similar pattern is seen in the Hilo and San José tropopause-relative profiles, where LRT-relative O<sub>3</sub> concentrations begin to exceed those based on the PTGT 3–4 km below the tropopause. The Hilo profiles also stand out due to the wider 10<sup>th</sup> to 90<sup>th</sup> percentile spread of PTGT-relative O<sub>3</sub>,  $\theta$ , and temperature, likely reflecting increased complexity of UTLS layers common to the subtropics. For the Hilo and San José profiles, the low PTGT-relative 10<sup>th</sup> percentile stratospheric O<sub>3</sub> concentrations and  $\theta$ , combined with the high 90<sup>th</sup> percentile PTGT-relative temperatures, are indicative of profiles where the PTGT is identified well below the LRT (as indicated in Fig. 8). Lastly, the South Pole Station exhibits the largest differences between PTGT- and LRT-relative observations of any of the balloon sounding sites. The median concentrations for all profiles (Fig. 9) differ slightly, but the 90<sup>th</sup> percentile below-tropopause concentrations of LRT-relative O<sub>3</sub> are an order of magnitude larger than the respective PTGT-relative observations. This is the result of a number of profiles where the lapse-rate criteria are not satisfied near the strong tropopause-level

O<sub>3</sub> gradient, resulting in an LRT altitude that is biased several kilometers high (as in the South Pole Station example in Fig. 7).

Figure 10 displays tropopause-relative observations for only those profiles where the PTGT is identified more than 1 km below the LRT for the 4 balloon sounding sites where there is adequate sampling of such environments. For Hilo and Boulder, 38.6% and 21.8% of profiles contain a PTGT altitude more than 1 km below the LRT and substantial differences between all three PTGT- and LRT- relative parameters are revealed in Figure 10. As in the results for all profiles in Figure 9, O<sub>3</sub> concentrations remain low throughout the PTGT-relative troposphere, with a modest gradient occurring across the tropopause. Alternatively, the LRT-relative O<sub>3</sub> profiles for Hilo and Boulder feature increases throughout the upper troposphere and a stronger gradient across the tropopause than that for PTGT concentrations. The accompanying temperature and potential temperature profiles provide insight into why these differences may be occurring. Specifically, the LRT-relative temperatures exhibit a strong, cold-point minimum at the level of the diagnosed tropopause that is characteristic of tropical environments, while the PTGT-relative temperatures are similar in structure to a double tropopause environment, where a weak transition in stability accompanies the primary tropopause and the colder, higher tropical tropopause remains as the secondary tropopause (and often temperature minimum of the profile). This suggests that Hilo and Boulder profiles where the LRT is located well above the PTGT occur due to the PTGT identifying a lower, weaker stability transition that does not satisfy the criteria of the LRT definition but from a composition perspective (O<sub>3</sub>) is consistent with horizontal layering of tropical troposphere and extratropical stratosphere that is a robust outcome from the formation of a large-scale double tropopause (e.g., Pan et al. 2009).

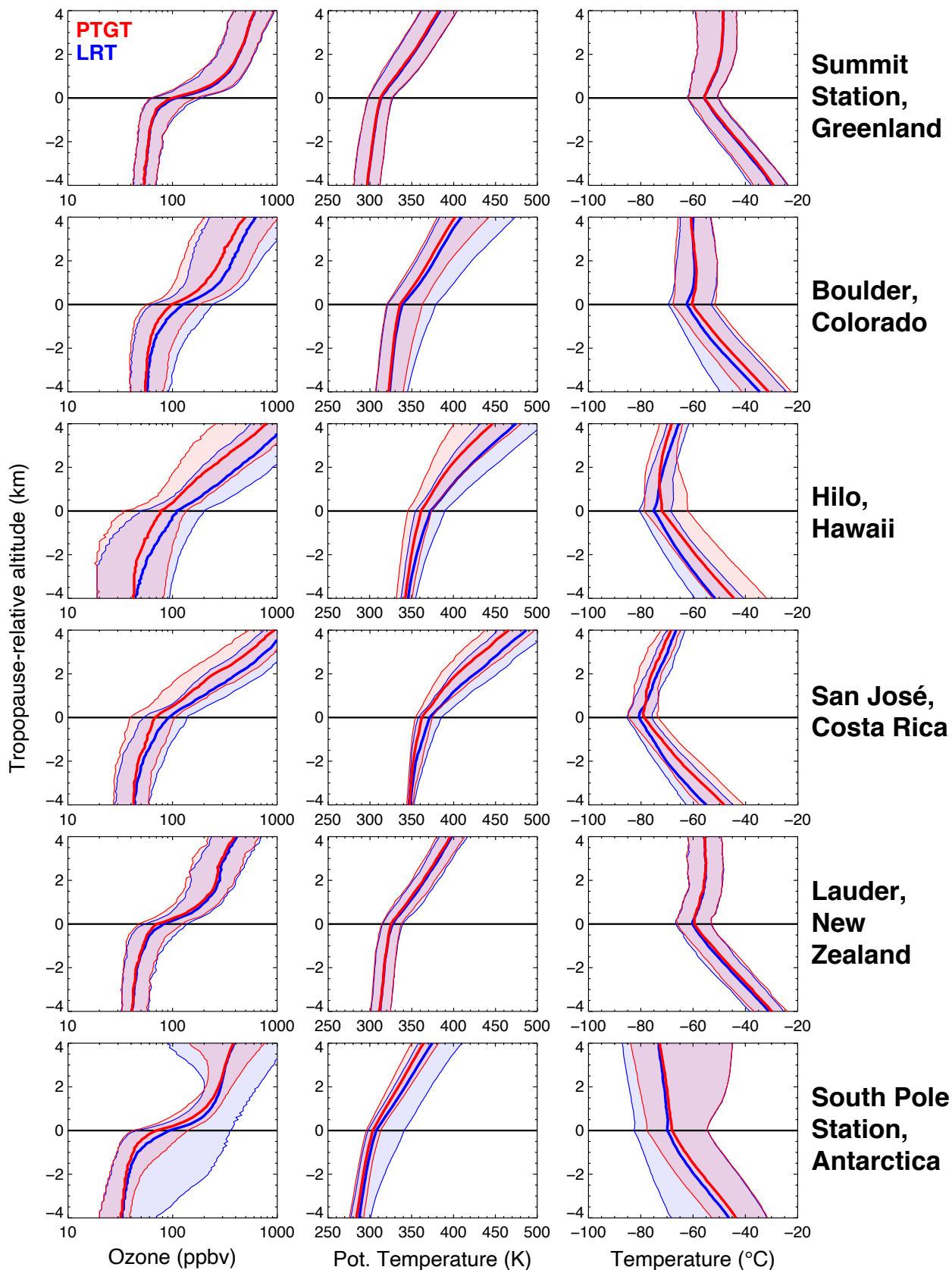


FIG. 9. For all profiles with ozone from each observation site: tropopause-relative profiles of (left) ozone, (middle) potential temperature, and (right) temperature. Thin lines bounding the color-fill in each panel span the 10th- to 90th-percentile observations for each variable, while thick lines superimposed on the color-fill represent the median value. Blue profiles are relative to the LRT, while red profiles are relative to the PTGT.

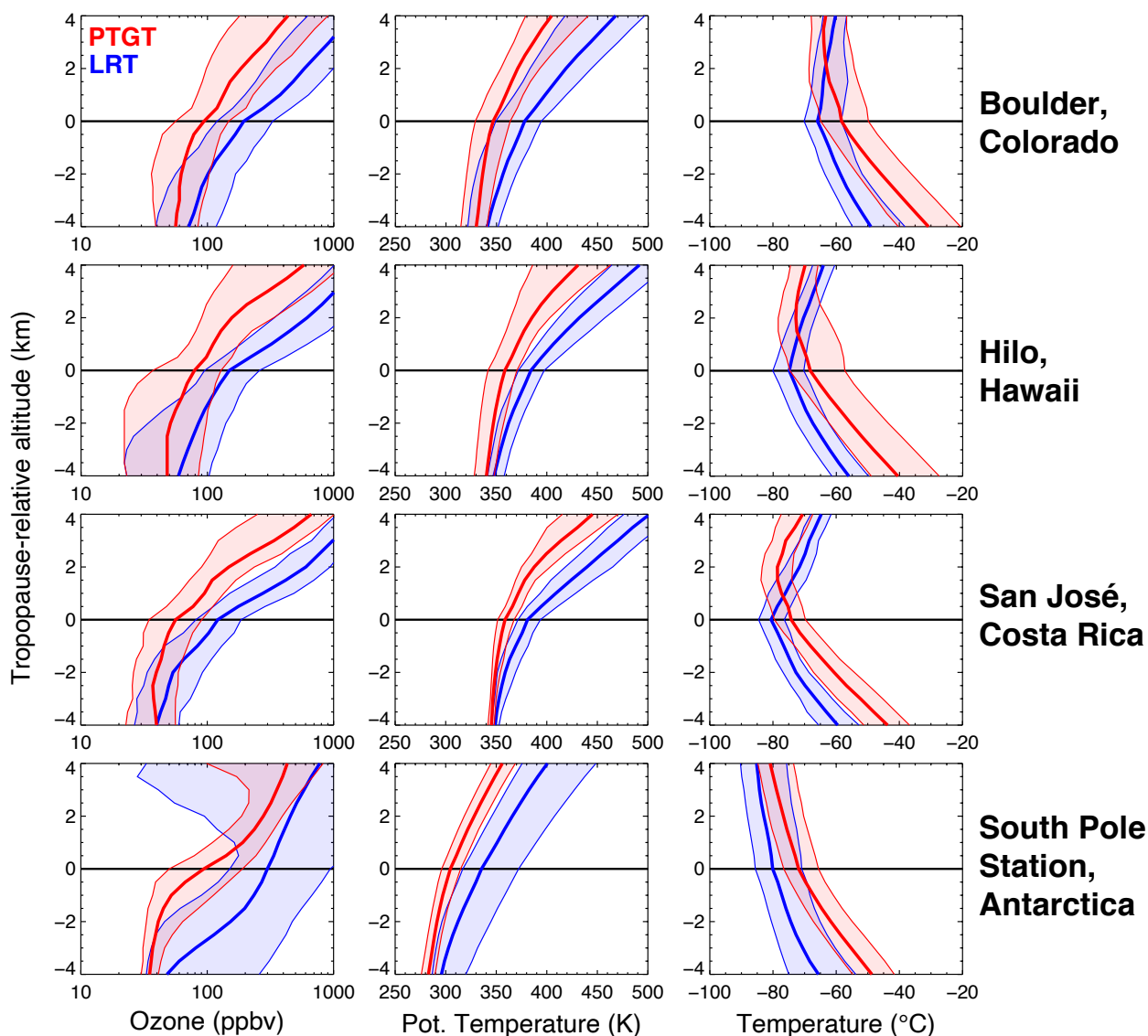


FIG. 10. As in Fig. 9, but only for instances where the PTGT is located more than 1 km below the LRT.

The results for profiles with PTGT altitude more than 1 km below the LRT at San José are similar to those at Hilo and Boulder. While this outcome is somewhat surprising as transport of extratropical lower stratosphere air does not often extend deeply into the tropics, the tropopause-relative  $O_3$  indicates that the PTGT location better captures the tropopause-level composition change at this station. Finally, for the 26.5% of South Pole Station profiles where the PTGT is located more than 1 km below the LRT, the differences in  $O_3$  concentrations are remarkable. As discussed previously, the LRT often fails to capture the weak stability transition that occurs at the poles and is biased several kilometers high. The analysis shown in Figure 10 demonstrates that the PTGT excels at identifying the tropopause transition in these cases, as demonstrated by its approximate co-location with the sharp  $O_3$  increase at the tropopause.

We further examine the appropriateness of each tropopause definition from a composition perspective by analyzing their locations in  $O_3$ – $H_2O$  tracer–tracer space. This is accomplished using profiles from the four balloon sounding sites where coincident  $O_3$  and  $H_2O$  observations are available: Boulder, Hilo, San José, and Lauder. Figure 11 shows joint frequency distributions of  $O_3$  and  $H_2O$  observations with the frequency of PTGT and LRT definitions in  $O_3$ – $H_2O$  space superimposed, where observations from all four sites are combined and separated into high tropopause environments (where both PTGT and LRT exceed 14 km), low tropopause environments (where both PTGT and LRT are below 14 km), and environments where the LRT is identified above 14 km while the PTGT is identified below 14 km (i.e., where they differ substantially).

The  $O_3$ – $H_2O$  joint frequency distributions in Figure 11 reveal features commonly identified in prior work (Pan et al. 2004; Hegglin et al. 2009; Tilmes et al. 2010). Namely, tropical, high-tropopause altitude environments are characterized by two distinct stratospheric and tropospheric branches in tracer–tracer space, which meet near the tropopause at low concentrations of both trace gases and form an “L”-shaped relationship (as found in Fig. 11a). Here, the stratospheric branch has uniformly low  $H_2O$  (<10 ppmv) and a wide range of  $O_3$  concentrations, and the tropospheric branch has uniformly low  $O_3$  (<150 ppbv) and a wide range of  $H_2O$  concentrations. In the subtropics and extratropics, a three-branch structure emerges, with the third branch characterized by a negative slope that broadens the corner of the “L” shape and intersects the tropospheric and stratospheric branches (commonly referred to as the extratropical transition layer or ExTL, or more generally as the “mixing” branch). The degree of mixing between stratospheric and/or tropospheric air in

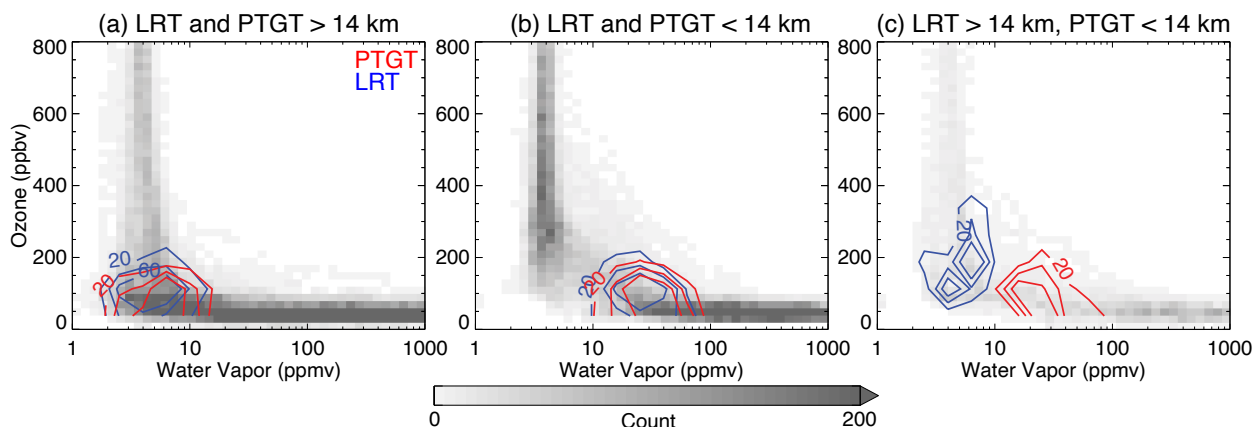


FIG. 11. Locations of identified PTGTs and LRTs in tracer–tracer space for co-located observations of ozone and water vapor from the Boulder, Hilo, San José, and Lauder stations for (a) high-tropopause (tropical) environments, (b) low-tropopause (extratropical) environments, and (c) environments where a tropical and extratropical tropopause are identified by the LRT and PTGT definitions, respectively. Gray shading indicates the number of 250-m layer observations placed into each water vapor–ozone ( $0.1 \log(\text{ppmv}) \times 75 \text{ ppbv}$  sized) bin, and red and blue contours represent the frequency of PTGT and LRT definitions in tracer–tracer space (in %), respectively. The number of profiles contributing to each panel are: (a) 358, (b) 332, and (c) 59.

the UTLS helps to control the extent to which this branch extends toward higher concentrations of the stratospheric and tropospheric tracers (e.g., Pan et al. 2004; Konopka and Pan 2012). For the tropical, high-tropopause environment profiles and extratropical, low-tropopause profiles analyzed here (Figs. 11a and 11b), there is clear agreement between the LRT and PTGT definitions in tracer–tracer space, with the PTGT favoring slightly lower  $\text{O}_3$ . For high-tropopause environments, the tropopause altitudes are consistently identified at the corner of the “L” shape. For low-tropopause environments, the tropopause altitudes from both definitions are consistently identified at the tropospheric end of the ExTL. For the profiles where the PTGT and LRT definitions substantially differ (Fig. 11c), the tracer–tracer distributions appear to be more subtropical in nature with a shallower mixing branch and the PTGT altitudes favor the tropospheric end of the mixing branch (as in low-tropopause, extratropical environments) while the LRT altitudes favor the stratospheric end of the mixing branch. This result demonstrates that the PTGT definition results in more consistent stability-based identification of UTLS composition change than the LRT definition.

## 2) OBSERVATIONS OF MULTIPLE TROPOPAUSES

Environments with multiple stable layers in the UTLS, and therefore multiple tropopause, are not uncommon and have been observed since early radiosondes (e.g., Kochanski 1955). Such environments are of particular interest due to their connection with both large and smaller (i.e. convective) scale stratosphere–troposphere exchange (e.g., Schwartz et al. 2015; Pan et al. 2009; Homeyer et al. 2011, 2014; Solomon et al. 2016; Tinney and Homeyer 2021). Accordingly, double tropopause (DT) environments have been investigated in both observations and model output in recent decades. They occur in both hemispheres and most frequently occur in the midlatitudes during the cool season, often near the subtropical jet (e.g., Randel et al. 2007a; Añel et al. 2008; Schwartz et al. 2015; Manney et al. 2017; Xian and Homeyer 2019). Most frequently, a DT is formed by differential advection of less stable, tropical upper tropospheric and more stable, extratropical lower stratospheric air, often associated with a Rossby wave breaking event (e.g., Randel et al. 2007a; Pan et al. 2009; Castanheira and Gimeno 2011; Peevey et al. 2012; Schwartz et al. 2015; Liu and Barnes 2018). DTs can also form through modification of thermodynamic properties from dynamic stretching or shrinking of layers in the UTLS (e.g., Randel et al. 2007a; Wang and Polvani 2011; Peevey et al. 2014). Additionally, Castanheira et al. (2009) and Xian and Homeyer (2019) have shown an increase in the frequency of DT events since the mid-to-late 20<sup>th</sup> century, for which the processes responsible are not well understood.

With the exception of the proposed PTGT definition, the LRT is the only tropopause definition that provides additional language for DT identification: *“If above the first tropopause the average lapse rate between any level and all higher levels within 1 km exceeds  $3^{\circ}\text{C km}^{-1}$ , then a second tropopause is defined by the same criterion as [the primary definition]. This tropopause may be either within or above the 1 km layer”* (World Meteorological Organization 1957). As described in the algorithm description in Section 3b, the PTGT provides a similar definition based on  $\partial\theta/\partial z$  thresholds. A comparison of LRT- and PTGT-defined DTs is therefore desired, which we include here. The frequency of DT identifications at each station based on the LRT and PTGT definitions are summarized in Table 2. At all stations, the PTGT identifies fewer DTs than the LRT, though the magnitude of this difference varies by station. This is likely the result of the variation in the  $\partial\theta/\partial z$ – $\partial T/\partial z$  relationship with altitude: the LRT and PTGT thresholds are well aligned at lower latitudes where DTs occur at lower pressures, while at higher latitudes it is more difficult for the PTGT DT



requirement of  $15 \text{ K km}^{-1}$  to be met at higher pressures (Fig. 6). The latitudinal variation in DT identifications is similar between the two definitions, with both identifying DTs most frequently at midlatitude stations (Boulder and Lauder). However, the LRT identifies far more DTs in the polar stations (Summit and South Pole) where the PTGT identifies relatively few. To further explore these differences, example profiles are shown in Figure 12 and frequency distributions of tropopause altitude differences are shown in Figure 13.

TABLE 2. The frequency of double tropopause identifications in the observed profiles. For each site, the double tropopause frequency based on the LRT definition, PTGT definition, and both definitions is given.

Observation Site	LRT (%)	PTGT (%)	Both (%)
Summit Station, Greenland	23.7	2.8	2.5
Boulder, Colorado	61.5	39.9	34.1
Hilo, Hawaii	22.2	20.1	8.1
San José, Costa Rica	12.4	11.9	3.9
Lauder, New Zealand	41.1	15.1	13.5
South Pole Station, Antarctica	22.0	6.4	2.7

In cases where both definitions identify a DT (Figs. 12e,f and right column of Fig. 13), the identifications of both the primary and secondary tropopause levels are within 500 m of each other more than 60% of the time at all stations except for the South Pole. Additionally, they are consistent more than 80% of the time at Boulder, the station where such profiles are most common. When a DT is identified by the LRT definition only (Figs. 12a,b and left column of Fig. 13), the PTGT and LRT primary tropopause are often found within 500 m of each other, while the LRT identifies a secondary boundary anywhere from  $\sim 3\text{--}7$  km above, depending on the station. As mentioned above, at least part of this population of DTs identified by the LRT only is likely a result of the vertical variation in the  $\partial\theta/\partial z - \partial T/\partial z$  relationship making the PTGT DT requirement more difficult to attain, especially at higher latitudes. Alternatively, in cases where a DT is identified by the PTGT definition only, the PTGT secondary tropopause often aligns with the LRT primary tropopause (Figs. 12c,d and middle column of Fig. 13). However, the primary PTGT in such cases is most frequently found  $\sim 2\text{--}4$  km below the primary LRT, where it is identifying a stability change that does not satisfy the temperature threshold used in the LRT. This finding provides additional context to the analysis in Figure 10, where PTGT-relative temperature profiles exhibit characteristics of a DT environment.

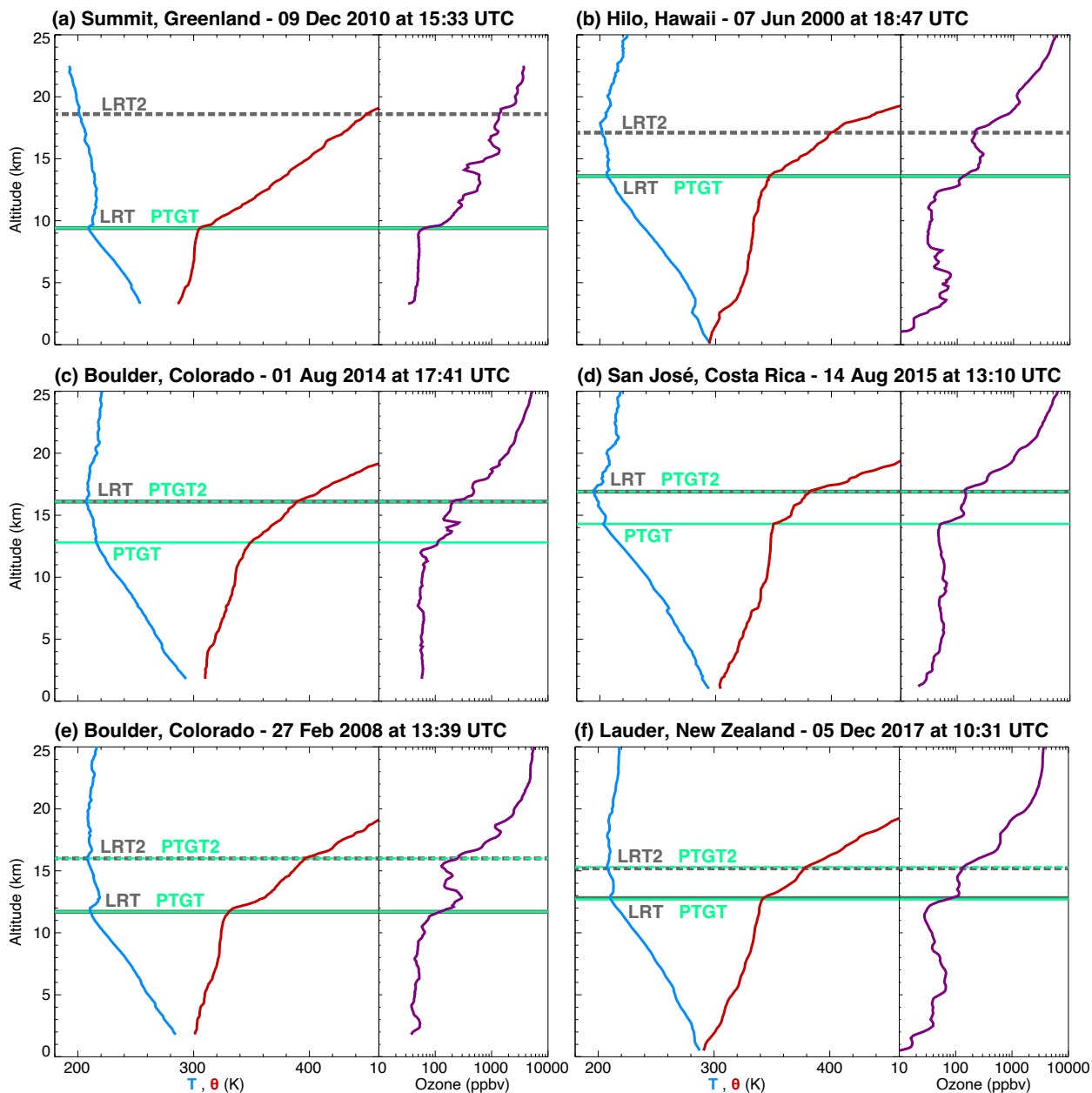


FIG. 12. As in Fig. 7, but for select profiles with a double tropopause identification: (a,b) LRT-only, (c,d) PTGT-only, and (e,f) LRT and PTGT. Secondary tropopauses are given by the dashed horizontal lines.

## Double Tropopause Comparisons

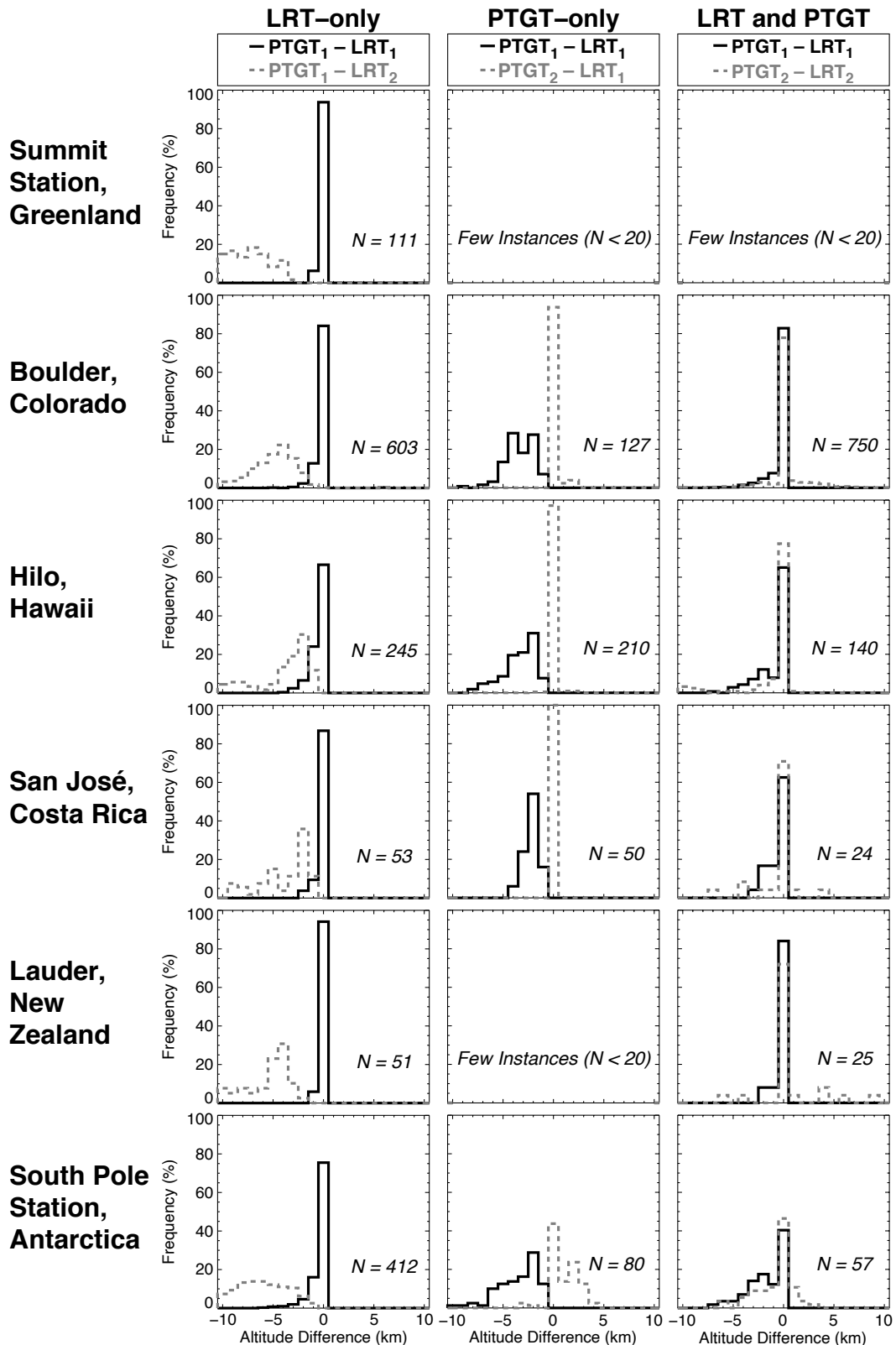


FIG. 13. Frequency distributions of tropopause altitude differences for profiles with a double tropopause (DT), using data from the six balloon sites. From left to right: profiles with a DT identified by only the LRT definition, only the PTGT definition, and both the LRT and PTGT definitions. For cases where only one tropopause definition identifies a DT, black solid lines indicate differences between the primary tropopause altitudes of the two definitions and dashed gray lines indicate differences between the secondary tropopause and the primary tropopause identified by the other definition. For cases where both the LRT and PTGT identify a DT, solid black lines indicate differences between primary tropopauses and dashed gray lines indicate differences between secondary tropopauses. The number of profiles used in each panel is noted.

### 3) APPLICATION TO REANALYSIS OUTPUT

Figures 14, 15 and 16 enable broader evaluation of the performance of the PTGT definition compared to the LRT. Both global maps (Fig. 14) and pole-to-pole cross sections at select longitudes (Fig. 16) demonstrate that, generally, there is broad consistency between the PTGT and LRT in each season. Namely, both definitions represent the tropopause break (indicated by the shift from dark purple to dark green in Fig. 14) similarly, as well as intricate dynamic features occurring in the midlatitudes. In Figure 15, DT features across the midlatitudes are also seen with broadly consistent identification between the definitions, though the DT areas identified by the LRT definition are slightly more expansive, which is consistent with the balloon observation DT frequency differences in Table 2.

There are a few additional key differences between reanalysis-derived PTGT and LRT altitudes worth noting. First, in high-tropopause environments, the PTGT tends to be lower than the LRT. Second, the high bias in LRT altitude in the polar regions is clearly resolved by the PTGT definition in these MERRA-2 examples, specifically evidenced near the South Pole in July and October. Similarly, the LRT identifies a widespread (likely nonphysical) DT feature in the high Arctic during January (Fig. 15) that is not identified by the PTGT definition in this region. Lastly, in Figure 16, we examine the consistency of each definition with the commonly-used dynamic tropopause via several PV surfaces. In general, the PTGT primary tropopause more consistently aligns with the PV contours. This is especially apparent near the subtropical jets, where the PTGT primary tropopause follows the PV contours and identifies a secondary tropopause at higher altitudes where the LRT places the primary —and only— tropopause. Overall, the application of the PTGT algorithm to reanalysis is largely consistent with the observational analysis above for both primary and secondary tropopause identifications.

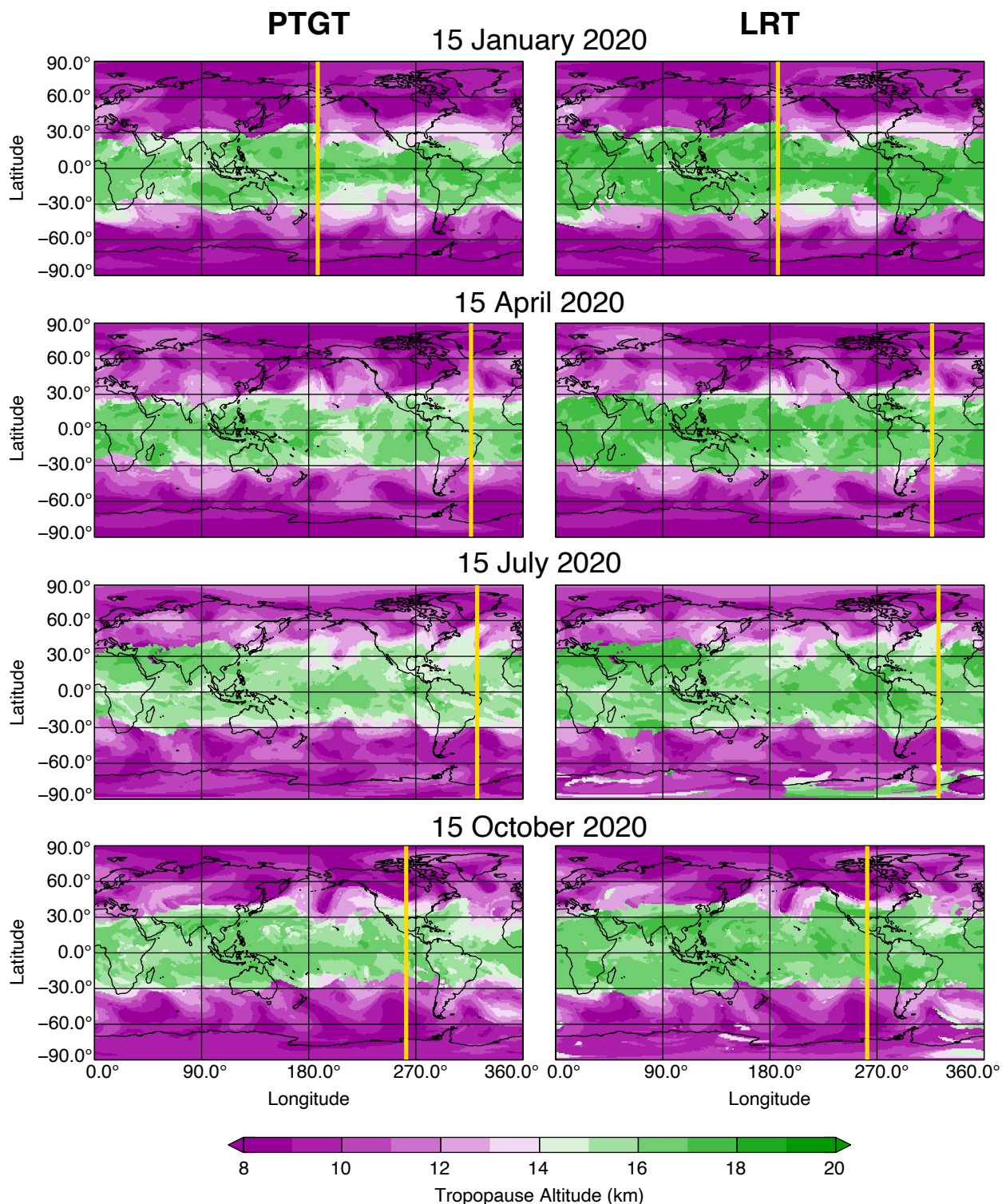


FIG. 14. Maps of primary tropopause altitude diagnosed by the PTGT and LRT definitions for four select MERRA-2 reanalysis times: (top to bottom) 15 January 2020, 15 April 2020, 15 July 2020, and 15 October 2020. Thick yellow lines in each panel correspond to the locations of the vertical cross-sections in Fig. 16.

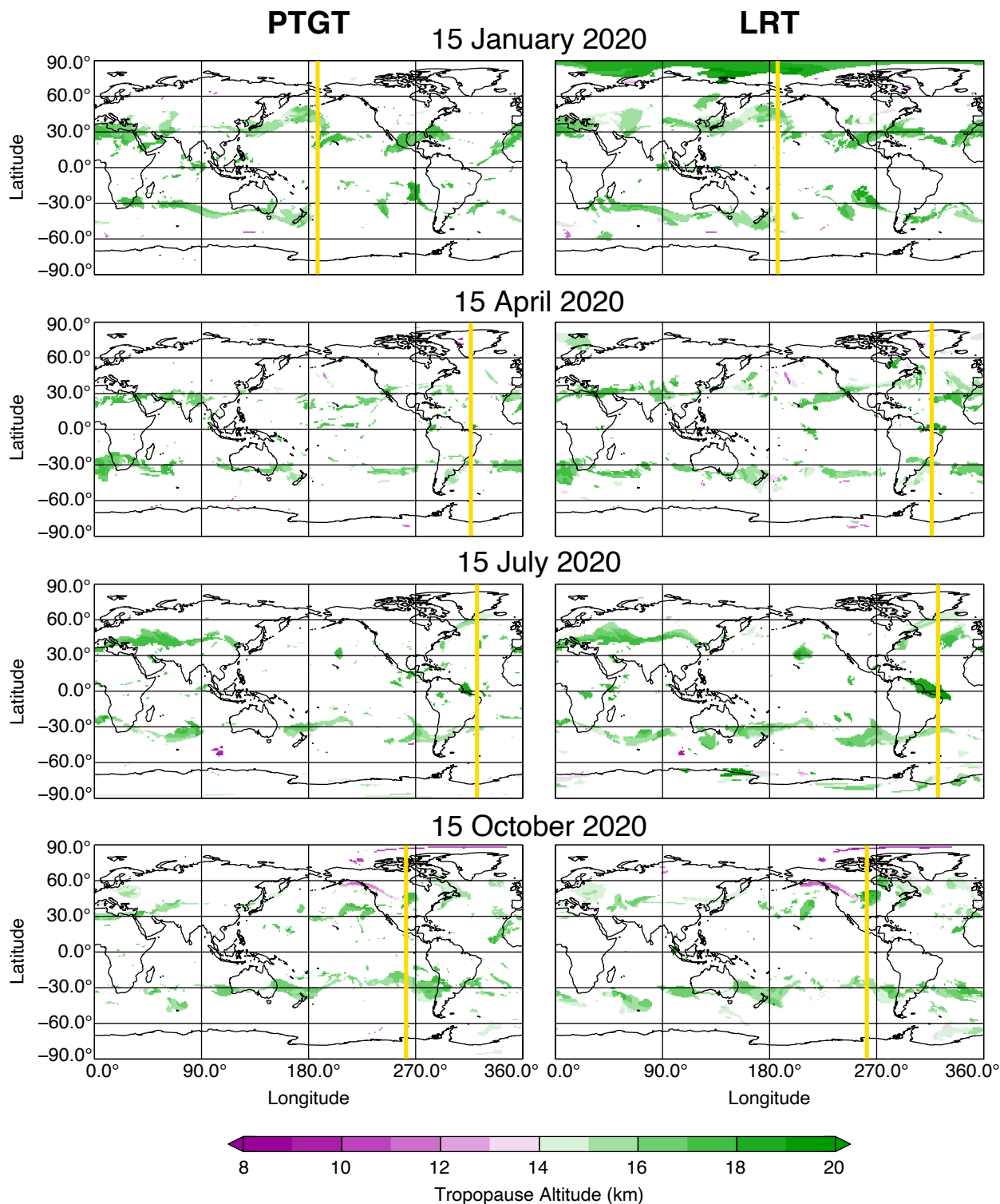


FIG. 15. As in Fig. 14, but for secondary tropopause altitude when present.

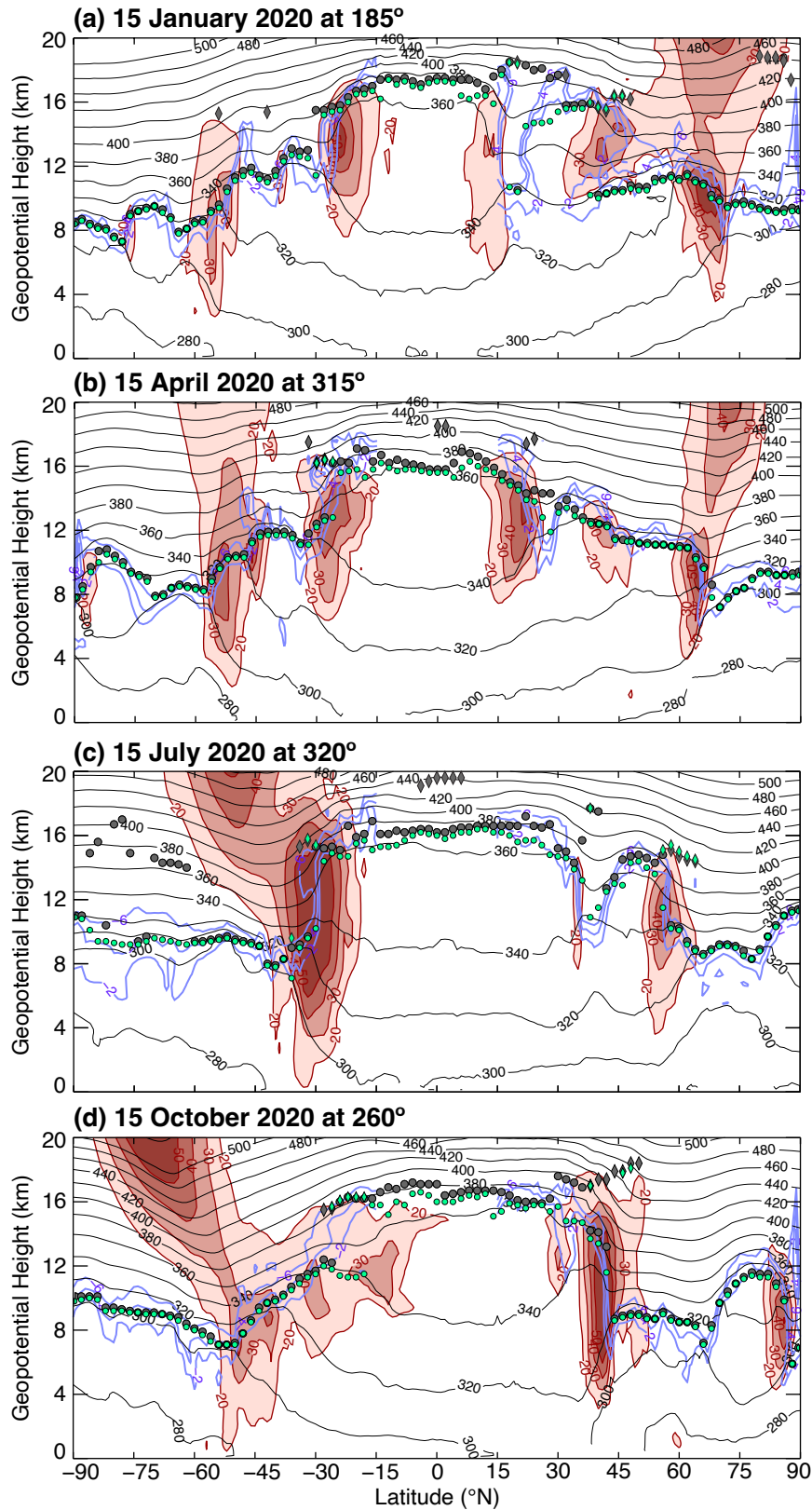




FIG. 16. Pole-to-pole vertical cross-sections of MERRA-2 output for the four times in Fig. 14. Horizontal wind speeds are shown by the red color-filled contours (every  $10 \text{ m s}^{-1}$ , starting at  $20 \text{ m s}^{-1}$ ), potential temperature (in K) by the black contours, potential vorticity above 4 km by the purple contours (every 2 PVU from  $\pm 2$ –6 PVU), PTGT altitudes by the green circles, PTGT secondary tropopause altitudes by the green diamonds, LRT altitudes by the gray circles, and LRT secondary tropopause altitudes by the gray diamonds.

## 4. Summary

This study examines long-term records of balloon-based observations of atmospheric temperature and composition to revisit stability-based definition of the tropopause. More than 7000 O<sub>3</sub> and temperature profiles observed in tropical, extratropical, and polar locations spanning all seasons and multiple decades were used to investigate composition-stability relationships near the tropopause (Figs. 1, 2, and 3). By analyzing coincident observations of O<sub>3</sub> concentrations and various stability metrics in both individual ozonesonde profiles (Fig. 4) as well as a statistical analysis of all observations (Fig. 5), it was shown that the vertical gradient of potential temperature ( $\partial\theta/\partial z$ ) is the most consistent stability-based discriminator for the sharp composition change between troposphere and stratosphere in a wide variety of environments when compared to common alternatives such as the temperature lapse rate and Brunt-Väisälä frequency.

Based on the identified superiority of  $\partial\theta/\partial z$  as a globally-consistent stability-based indicator of the tropopause transition layer, identification of a tropopause level via a PTGT definition is desired. Modeled after the LRT, the PTGT definition simply requires a potential temperature gradient threshold to be met across multiple depths. This PTGT algorithm offers a new globally- and universally-applicable stability-based tropopause definition for future UTLS studies and serves as an alternative to the LRT. Through comparison of PTGT and LRT altitudes in observations (Figs. 7 and 8) and reanalysis model output (Figs. 14 and 16), as well as the examination of PTGT- and LRT-relative O<sub>3</sub> concentrations (Figs. 9 and 10) and their locations in O<sub>3</sub>–H<sub>2</sub>O tracer–tracer space (Fig. 11), it was demonstrated that the PTGT resolves known limitations of the LRT. Moreover, though the PTGT is often found within ~1 km of the LRT (Fig. 8), when it differs greatly from the LRT it is more consistent in identifying the layer of greatest composition change in the UTLS, favoring the tropospheric endpoint of the chemical mixing layer. Instances of the PTGT differing greatly from the LRT are most often negative, with the PTGT well below the LRT altitude. This result from a composition-informed stability-based tropopause definition (i.e., the PTGT) is consistent with findings in past studies using an ozone tropopause (e.g., Bethan et al. 1996).

Criteria for multiple tropopause identification using the potential temperature gradient were also included in the PTGT definition proposed here and modeled after the LRT —the only alternative definition that includes criteria for multiple tropopause identification. Instances of double tropopauses (DTs) were evaluated using the PTGT definition and compared to the LRT definition,

as done for the primary tropopause altitude. It was demonstrated that 1) fewer double tropopauses are identified by the PTGT definition (Table 2 and Fig. 15), 2) when both the LRT and PTGT identify a DT, they are in close agreement (Fig. 13), and 3) cases where only the PTGT identifies a double tropopause are largely those where the primary PTGT falls well below the LRT, such that the LRT altitude closely coincides with the secondary PTGT.

In conclusion, the introduction of a new universal stability-based tropopause definition motivates revisiting past work built upon the LRT definition. In particular, future work with the PTGT should be dedicated toward important UTLS topics such as stratosphere–troposphere exchange, tropopause climatology, and long-term tropopause variability and change.

*Acknowledgments.* We thank Bryan Johnson and five reviewers for helpful feedback on the paper. This work was supported, in part, by the National Aeronautics and Space Administration (NASA) under Award 80NSSC18K0746, and by the National Science Foundation (NSF) under grant no. AGS-1560419.

*Data availability statement.* All of the data used in this study are publicly and freely available from NOAA (most of the ozone and water vapor balloon observations; NOAA (2021)) and NASA (the Costa Rica ozone and water vapor balloon observations and the MERRA-2 output; NASA (2022a), NASA (2022b), and Global Modeling and Assimilation Office (GMAO) (2015)).

## References

- Añel, J. A., J. C. Antuña, L. de la Torre, J. M. Castanheira, and L. Gimeno, 2008: Climatological features of global multiple tropopause events. *J. Geophys. Res.*, **133**, D00B08, <https://doi.org/10.1029/2007JD009697>.
- Banerjee, A., G. Chiodo, M. Previdi, M. Ponater, A. J. Conley, and L. M. Polvani, 2019: Stratospheric water vapor: an important climate feedback. *Climate Dynamics*, **53**, 1697–1710, <https://doi.org/10.1007/s00382-019-04721-4>.
- Berthet, G., J. G. Esler, and P. H. Haynes, 2007: A Lagrangian perspective of the tropopause and the ventilation of the lowermost stratosphere. *J. Geophys. Res.*, **112**, D18 102, <https://doi.org/10.1029/2006JD008295>.
- Bethan, S., G. Vaughan, and S. J. Reid, 1996: A comparison of ozone and thermal tropopause heights and the impact of tropopause definition on quantifying the ozone content of the troposphere. *Q. J. R. Meteorol. Soc.*, **122**, 929–944.
- Birner, T., 2006: Fine-scale structure of the extratropical tropopause. *J. Geophys. Res.*, **111**, D04 104, <https://doi.org/10.1029/2005JD006301>.
- Browell, E. V., E. F. Danielsen, S. Ismail, G. L. Gregory, and S. M. Beck, 1987: Tropopause fold structure determined from airborne lidar in situ measurements. *J. Geophys. Res.*, **92 (D2)**, 2112–2120.

- Castanheira, J. M., J. A. Añel, C. A. F. Marques, J. C. Antuña, M. L. R. Liberato, L. de la Torre, and L. Gimeno, 2009: Increase of upper troposphere/lower stratosphere wave baroclinicity during the second half of the 20th century. *Atmos. Chem. Phys.*, **9** (23), 9143–9153.
- Castanheira, J. M., and L. Gimeno, 2011: Association of double tropopause events with baroclinic waves. *J. Geophys. Res.*, **116**, D19 113, <https://doi.org/10.1029/2011JD016163>.
- Danielsen, E. F., 1959: The laminar structure of the atmosphere and its relation to the concept of a tropopause. *Arch. Meteorol. Geophys. Bioklimatol., Ser. A*, **11**, 293–332.
- Danielsen, E. F., 1968: Stratospheric-tropospheric exchange based on radioactivity, ozone and potential vorticity. *J. Atmos. Sci.*, **25**, 502–518.
- Duran, P., and J. Molinari, 2019: Tropopause evolution in a rapidly intensifying tropical cyclone: A static stability budget analysis in an idealized axisymmetric framework. *Journal of the Atmospheric Sciences*, **76** (1), 209–229.
- Fischer, H., and Coauthors, 2000: Tracer correlations in the northern high latitude lowermost stratosphere: Influence of cross-tropopause mass exchange. *Geophys. Res. Lett.*, **27** (1), 97–100.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Gettelman, A., P. Hoor, L. L. Pan, W. J. Randel, M. I. Hegglin, and T. Birner, 2011: The extratropical upper troposphere and lower stratosphere. *Rev. Geophys.*, **49**, RG3003, <https://doi.org/10.1029/2011RG000355>.
- Gettelman, A., and T. Wang, 2015: Structural diagnostics of the tropopause inversion layer and its evolution. *J. Geophys. Res. Atmos.*, **120**, 46–62, <https://doi.org/10.1002/2014JD021846>.
- Global Modeling and Assimilation Office (GMAO), 2015: MERRA-2 inst3\_3d\_asm\_Nv: 3d, 3-Hourly, Instantaneous, Model-Level, Assimilation, Assimilated Meteorological Fields V5.12.4. Goddard Earth Sciences Data and Information Services Center (GES DISC), URL [https://disc.gsfc.nasa.gov/datasets/M2I3NVASM\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2I3NVASM_5.12.4/summary), <https://doi.org/10.5067/WWQSQXQ8IVFW8>.

- Hall, E. G., A. F. Jordan, D. F. Hurst, S. J. Oltmans, H. Vömel, B. Kühnreich, and V. Ebert, 2016: Advancements, measurement uncertainties and recent comparisons of the NOAA frost point hygrometer. *Atmos. Meas. Tech.*, **9**, 4295–4310, <https://doi.org/10.5194/amt-9-4295-2016>.
- Hegglin, M. I., C. D. Boone, G. L. Manney, and K. A. Walker, 2009: A global view of the extratropical tropopause transition layer from atmospheric chemistry experiment fourier transform spectrometer o-3, h2o, and co. *J. Geophys. Res.*, **114**, D00B11, <https://doi.org/10.1029/2008JD009984>.
- Highwood, E. J., and B. J. Hoskins, 1998: The tropical tropopause. *Quarterly Journal of the Royal Meteorological Society*, **124** (549), 1579–1604, [https://doi.org/https://doi.org/10.1002/qj.49712454911](https://doi.org/10.1002/qj.49712454911).
- Hoerling, M. P., T. K. Schaack, and A. J. Lenzen, 1991: Global objective tropopause analysis. *Mon. Wea. Rev.*, **119**, 1816–1831.
- Hoffmann, L., and R. Spang, 2022: An assessment of tropopause characteristics of the ERA5 and ERA-Interim meteorological reanalyses. *Atmos. Chem. Phys.*, **22**, 4019–4046, <https://doi.org/10.5194/acp-22-4019-2022>.
- Hoinka, K. P., 1997: The tropopause: discovery, definition and demarcation. *Meteorol. Z.*, **6**, 281–303.
- Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglass, and L. Pfister, 1995: Stratosphere-troposphere exchange. *Rev. Geophys.*, **33** (4), 403–439.
- Homeyer, C. R., K. P. Bowman, and L. L. Pan, 2010: Extratropical tropopause transition layer characteristics from high-resolution sounding data. *J. Geophys. Res.*, **115**, D13 108, <https://doi.org/10.1029/2009JD013664>.
- Homeyer, C. R., K. P. Bowman, L. L. Pan, E. L. Atlas, R.-S. Gao, and T. L. Campos, 2011: Dynamical and chemical characteristics of tropospheric intrusions observed during start08. *Journal of Geophysical Research: Atmospheres*, **116** (D6), [https://doi.org/https://doi.org/10.1029/2010JD015098](https://doi.org/10.1029/2010JD015098).

- Homeyer, C. R., and Coauthors, 2014: Convective transport of water vapor into the lower stratosphere observed during double-tropopause events. *Journal of Geophysical Research: Atmospheres*, **119** (18), 10,941–10,958, <https://doi.org/10.1002/2014JD021485>.
- Hoor, P., H. Fischer, L. Lange, J. Lelieveld, and D. Brunner, 2002: Seasonal variations of a mixing layer in the lowermost stratosphere as identified by the CO-O<sub>3</sub> correlation from in situ measurements. *J. Geophys. Res.*, **107** (D5), 4044, <https://doi.org/10.1029/2000JD000289>.
- Hurst, D. F., S. J. Oltmans, H. Vömel, K. H. Rosenlof, S. M. Davis, E. A. Ray, E. G. Hall, and A. F. Jordan, 2011: Stratospheric water vapor trends over boulder, colorado: Analysis of the 30 year boulder record. *J. Geophys. Res.*, **116** (D2), <https://doi.org/10.1029/2010JD015065>.
- Kim, J.-E., and M. J. Alexander, 2015: Direct impacts of waves on tropical cold point tropopause structure. *Geophys. Res. Lett.*, **42**, 1584–1592, <https://doi.org/10.1002/2014GL062737>.
- Kochanski, A., 1955: Cross sections of the mean zonal flow and temperature along 80°w. *J. Meteor.*, **12**, 95–106.
- Konopka, P., and L. L. Pan, 2012: On the mixing-driven formation of the Extratropical Transition Layer (ExTL). *J. Geophys. Res.*, **117**, D18 301, <https://doi.org/10.1029/2012JD017876>.
- Kunz, A., P. Konopka, R. Müller, and L. L. Pan, 2011: Dynamical tropopause based on isentropic potential vorticity gradients. *J. Geophys. Res.*, **116**, D01 110, <https://doi.org/10.1029/2010JD014343>.
- Liu, C., and E. Barnes, 2018: Synoptic formation of double tropopauses. *J. Geophys. Res. Atmos.*, **123**, 693–707, <https://doi.org/10.1002/2017JD027941>.
- Maddox, E. M., and G. L. Mullendore, 2018: Determination of best tropopause definition for convective transport studies. *Journal of the Atmospheric Sciences*, **75** (10), 3433–3446, <https://doi.org/10.1175/JAS-D-18-0032.1>.
- Manney, G. L., and Coauthors, 2017: Reanalysis comparisons of upper tropospheric-lower stratospheric jets and multiple tropopauses. *Atmos. Chem. Phys.*, **17**, 11 541–11 566, <https://doi.org/10.5194/acp-17-11541-2017>.

- Meng, L., J. Liu, D. W. Tarasick, W. J. Randel, A. K. Steiner, H. Wilhelmson, L. Wang, and L. Haimberger, 2021: Continuous rise of the tropopause in the Northern Hemisphere over 1980–2020. *Sci. Adv.*, **7**, eabi8065, <https://doi.org/10.1126/sciadv.abi8065>.
- Miloshevich, L. M., A. Paukkunen, H. Vömel, and S. J. Oltmans, 2004: Development and validation of a time-lag correction for Vaisala radiosonde humidity measurements. *J. Atmos. Oceanic Technol.*, **21**, 1305–1327.
- NASA, 2022a: Costa Rica water vapor and ozonesondes, last accessed January 2022. NASA Network for the Detection of Atmospheric Composition Change, URL <https://www-air.larc.nasa.gov/missions/ndacc/data.html>.
- NASA, 2022b: Southern hemisphere Additional Ozonesondes (SHADOZ), last accessed January 2022. NASA Goddard Space Flight Center, URL <https://tropo.gsfc.nasa.gov/shadoz/>.
- Newman, P. A., and M. R. Schoeberl, 1995: A reinterpretation of the data from the NASA stratosphere-troposphere exchange project. *Geophys. Res. Lett.*, **22** (18), 2501–2504.
- NOAA, 2021: Ozone and water vapor sondes, last accessed November 2021. NOAA Earth System Research Laboratories Global Monitoring Laboratory, URL <https://gml.noaa.gov/aftp/data/ozwv/>.
- Olsen, M. A., A. R. Douglass, P. A. Newman, J. C. Gille, B. Nardi, V. A. Yudin, D. E. Kinnison, and R. Khosravi, 2008: HIRDLS observations and simulation of a lower stratospheric intrusion of tropical air to high latitudes. *Geophys. Res. Lett.*, **35**, L21 813, <https://doi.org/10.1029/2008GL035514>.
- Palmén, E., 1948: On the distribution of temperature and wind in the upper westerlies. *J. Meteorol.*, **5**, 20–27.
- Pan, L. L., S. B. Honomichl, T. V. Bui, T. Thornberry, A. Rollins, E. Hintsä, and E. J. Jensen, 2018: Lapse rate or cold point: The tropical tropopause identified by in situ trace gas measurements. *Geophys. Res. Lett.*, **45**, 10,756–10,763, <https://doi.org/10.1029/2018GL079573>.
- Pan, L. L., L. C. Paulik, S. B. Honomichl, L. A. Munchak, J. Bian, H. B. Selkirk, and H. Vömel, 2014: Identification of the tropical tropopause transition layer using the ozone-water vapor relationship. *J. Geophys. Res. Atmos.*, **119**, 3586–3599, <https://doi.org/10.1002/2013JD020558>.



- Pan, L. L., W. J. Randel, B. L. Gary, M. J. Mahoney, and E. J. Hints, 2004: Definitions and sharpness of the extratropical tropopause: A trace gas perspective. *J. Geophys. Res.*, **109**, D23 103, <https://doi.org/10.1029/2004JD004982>.
- Pan, L. L., and Coauthors, 2009: Tropospheric intrusions associated with the secondary tropopause. *J. Geophys. Res.*, **114**, D10 302, <https://doi.org/10.1029/2008JD011374>.
- Peevey, T. R., J. C. Gille, C. R. Homeyer, and G. L. Manney, 2014: The double tropopause and its dynamical relationship to the tropopause inversion layer in storm track regions. *J. Geophys. Res. Atmos.*, **119**, 10,194–10,212, <https://doi.org/10.1002/2014JD021808>.
- Peevey, T. R., J. C. Gille, C. E. Randall, and A. Kunz, 2012: Investigation of double tropopause spatial and temporal global variability utilizing High Resolution Dynamics Limb Sounder temperature observations. *J. Geophys. Res.*, **117**, D01 105, <https://doi.org/10.1029/2011JD016443>.
- Prather, M. J., X. Zhu, Q. Tang, J. Hsu, and J. L. Neu, 2011: An atmospheric chemist in search of the tropopause. *J. Geophys. Res.*, **116**, D04 306, <https://doi.org/10.1029/2010JD014939>.
- Randel, W. J., D. J. Seidel, and L. L. Pan, 2007a: Observational characteristics of double tropopauses. *J. Geophys. Res.*, **112**, D07 309, <https://doi.org/10.1029/2006JD007904>.
- Randel, W. J., F. Wu, and P. Forster, 2007b: The extratropical tropopause inversion layer: Global observations with GPS data, and a radiative forcing mechanism. *J. Atmos. Sci.*, **64**, 4489–4496.
- Reed, R. J., 1955: A study of a characteristic type of upper-level frontogenesis. *J. Meteorol.*, **12**, 226–237.
- Santer, B. D., and Coauthors, 2003: Behavior of tropopause height and atmospheric temperature in models, reanalyses, and observations: Decadal changes. *J. Geophys. Res.*, **108** (D1), 4002, <https://doi.org/10.1029/2002JD002258>.
- Schwartz, M. J., G. L. Manney, M. I. Hegglin, N. J. Livesey, M. L. Santee, and W. H. Daffer, 2015: Climatology and variability of trace gases in extratropical double-tropopause regions from MLS, HIRDLS, and ACE-FTS measurements. *J. Geophys. Res. Atmos.*, **120**, 843–867, <https://doi.org/10.1002/2014JD021964>.

- Seidel, D. J., and W. J. Randel, 2006: Variability and trends in the global tropopause estimated from radiosonde data. *J. Geophys. Res.*, **111**, D21 101, <https://doi.org/10.1029/2006JD007363>.
- Seidel, D. J., R. J. Ross, J. K. Angell, and G. C. Reid, 2001: Climatological characteristics of the tropical tropopause as revealed by radiosondes. *J. Geophys. Res.*, **106 (D8)**, 7857–7878.
- Shapiro, M. A., 1980: Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the stratosphere and troposphere. *J. Atmos. Sci.*, **37**, 994–1004.
- Shepherd, T. G., 2002: Issues in stratosphere-troposphere coupling. *J. Meteorol. Soc. Japan*, **80 (4B)**, 769–792.
- Škerlak, B., M. Sprenger, and H. Wernli, 2014: A global climatology of stratosphere-troposphere exchange using the ERA-Interim data set from 1979 to 2011. *Atmos. Chem. Phys.*, **14**, 913–937, <https://doi.org/10.5194/acp-14-913-2014>.
- Solomon, D. L., K. P. Bowman, and C. R. Homeyer, 2016: Tropopause-penetrating convection from three-dimensional gridded nexrad data. *Journal of Applied Meteorology and Climatology*, **55 (2)**, 465–478, <https://doi.org/10.1175/JAMC-D-15-0190.1>.
- Son, S.-W., and L. M. Polvani, 2007: Dynamical formation of an extra-tropical tropopause inversion layer in a relatively simple circulation model. *Geophys. Res. Lett.*, **34**, L17 806, <https://doi.org/10.1029/2007GL030564>.
- Sprenger, M., M. C. Maspoli, and H. Wernli, 2003: Tropopause folds and cross-tropopause exchange: A global investigation based upon ECMWF analyses for the time period March 2000 to February 2001. *J. Geophys. Res.*, **108 (D12)**, 8518, <https://doi.org/10.1029/2002JD002587>.
- Sterling, C. W., and Coauthors, 2018: Homogenizing and estimating the uncertainty in NOAA's long-term vertical ozone profile records measured with the electrochemical concentration cell ozonesonde. *Atmos. Meas. Tech.*, **11**, 3661–3687, <https://doi.org/10.5194/amt-11-3661-2018>.
- Stohl, A., H. Wernli, P. James, M. Bourqui, C. Forster, M. A. Liniger, P. Seibert, and M. Sprenger, 2003: A new perspective of stratosphere-troposphere exchange. *Bull. Amer. Meteorol. Soc.*, **84 (11)**, 1565–1573.

- Thompson, A. M., R. M. Stauffer, K. Wargan, J. C. Witte, D. E. Kollonige, and J. R. Ziemke, 2021: Regional and seasonal trends in tropical ozone from SHADOZ profiles: Reference for models and satellite products. *J. Geophys. Res. Atmos.*, **126**, e2021JD034 691, <https://doi.org/10.1029/2021JD034691>.
- Thompson, A. M., and Coauthors, 2017: First reprocessing of southern hemisphere additional ozonesondes (SHADOZ) profile records (1998–2016): 2. comparisons with satellite and ground-based instruments. *J. Geophys. Res. Atmos.*, **122**, 13,000–13,025, <https://doi.org/10.1002/2017JD027406>.
- Tilmes, S., and Coauthors, 2010: An aircraft-based upper troposphere and lower stratosphere O<sub>3</sub>, CO, and H<sub>2</sub>O climatology for the Northern Hemisphere. *J. Geophys. Res.*, **115**, D14 303, <https://doi.org/10.1029/2009JD012731>.
- Tinney, E. N., and C. R. Homeyer, 2021: A 13-year trajectory-based analysis of convection-driven changes in upper troposphere lower stratosphere composition over the united states. *Journal of Geophysical Research: Atmospheres*, **126** (3), e2020JD033 657, <https://doi.org/https://doi.org/10.1029/2020JD033657>.
- Vaughan, G., and C. Timmis, 1998: Transport of near-tropopause air into the lower midlatitude stratosphere. *Q. J. R. Meteorol. Soc.*, **124**, 1559–1578.
- Vömel, H., D. E. David, and K. Smith, 2007: Accuracy of tropospheric and stratospheric water vapor measurements by the cryogenic frost point hygrometer: Instrument details and observations. *J. Geophys. Res.*, **112**, D08 305, <https://doi.org/10.1029/2006JD007224>.
- Vömel, H., T. Naebert, R. Dirksen, and M. Sommer, 2016: An update on the uncertainties of water vapor measurements using cryogenic frost point hygrometers. *Atmos. Meas. Tech.*, **9**, 3755–3768, <https://doi.org/10.5194/amt-9-3755-2016>.
- Wang, S., and L. M. Polvani, 2011: Double tropopause formation in idealized baroclinic life cycles: The key role of an initial tropopause inversion layer. *J. Geophys. Res.*, **116**, D05 108, <https://doi.org/10.1029/2010JD015118>.

- Wernli, H., and M. Bourqui, 2002: A Lagrangian “1-year climatology” of (deep) cross-tropopause exchange in the extratropical Northern Hemisphere. *J. Geophys. Res.*, **107** (D2), 4021, <https://doi.org/10.1029/2001JD000812>.
- Witte, J. C., A. M. Thompson, H. G. J. Smit, H. Vömel, F. Posny, and R. Stübi, 2018: First reprocessing of Southern Hemisphere ADditional OZonesondes profile records: 3. uncertainty in ozone profile and total column. *J. Geophys. Res. Atmos.*, **123**, 3243–3268, <https://doi.org/10.1002/2017JD027791>.
- Witte, J. C., and Coauthors, 2017: First reprocessing of Southern Hemisphere ADditional OZonesondes (SHADOZ) profile records (1998–2015): 1. methodology and evaluation. *J. Geophys. Res. Atmos.*, **122**, 6611–6636, <https://doi.org/10.1002/2016JD026403>.
- World Meteorological Organization, 1957: Meteorology—a three-dimensional science: Second session of the commission for aerology. *World Meteorol. Organ. Bull.*, **4**, 134–138.
- Xian, T., and C. R. Homeyer, 2019: Global tropopause altitudes in radiosondes and reanalyses. *Atmos. Chem. Phys.*, **19**, 5661–5678, <https://doi.org/10.5194/acp-19-5661-2019>.
- Zahn, A., and C. A. M. Brenninkmeijer, 2003: New directions: A chemical tropopause defined. *Atmos. Environ.*, **37**, 439–440.
- Zängl, G., and K. P. Hoinka, 2001: The tropopause in the polar regions. *J. Clim.*, **14**, 3117–3139.