PDFs of Tropical Tropospheric Humidity: Measurements and Theory

JU-MEE RYOO

Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland

TAKERU IGUSA

Department of Civil Engineering, Johns Hopkins University, Baltimore, Maryland

DARRYN W. WAUGH

Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland

ABSTRACT

The spatial variations in the probability density functions (PDFs) of relative humidity (RH) in the tropical and subtropical troposphere are examined using observations from the Atmospheric Infrared Sounder (AIRS) and the Microwave Limb Sounder (MLS) instruments together with a simple statistical model. The model, a generalization of that proposed by Sherwood et al., assumes the RH is determined by a combination of drying by uniform subsidence and random moistening events and has two parameters: r, the ratio of the drying time by subsidence to the time between moistening events, and k, a measure of the variability of the moistening events. The observations show that the characteristics of the PDFs vary between the tropics and subtropics, within the tropics or subtropics, and with altitude. The model fits the observed PDFs well, and the model parameters concisely characterize variations in the PDFs and provide information on the processes controlling the RH distributions. In tropical convective regions, the model PDFs that match the observations have large r and small k, indicating rapid random remoistening, which is consistent with direct remoistening in convection. In contrast, in the nonconvective regions there are small r and large k, indicating slower, less random remoistening, consistent with remoistening by slower, quasi-horizontal transport. The statistical model derived will be useful for quantifying differences between, or temporal changes in, RH distributions from different datasets or models, and for examining how changes in physical processes could alter the RH distribution.

1. Introduction

Water vapor plays a crucial role in the earth's climate system, and the potential for water vapor feedbacks is a major challenge for understanding and predicting climate change. It is therefore important to know the distribution of atmospheric water vapor and the processes controlling this distribution. As the radiative effect of water vapor is roughly logarithmic in the concentration (e.g., Spencer and Braswell 1997; Held and Soden 2000; Pierrehumbert et al. 2006), it is important to know the full distribution of atmospheric water vapor and not just the mean and variance.

In recent years there have been several studies that have aimed at addressing these issues by examining the probability density functions (PDFs) of observed tropospheric humidity (Soden and Bretherton 1993; Gierens et al. 1999; Spichtinger et al. 2002; Zhang et al. 2003; Sherwood et al. 2006, hereafter S06; Luo et al. 2007; Ekström et al. 2007; Read et al. 2007). These studies all showed that the water vapor PDFs are very broad and non-Gaussian, but the characteristics of the PDFs have varied between studies. For example, Soden and Bretherton (1993) noted a lognormal distribution for 200–500-hPa upper-tropospheric humidity, whereas Zhang et al. (2003) and Luo et al. (2007) have reported bimodal PDFs. It is unclear whether the differences in reported PDFs are caused by differences in the

Corresponding author address: Ju-Mee Ryoo, Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD 21218.
E-mail: jryoo1@jhu.edu

instruments making the measurements (including remote versus in situ measurements), differences in the space–time resolution of the data used, or whether the differences are due to the different regions and time periods considered in the studies.

We investigate some of these issues by examining PDFs for different subregions of the tropics and subtropics (and the whole tropics) using daily measurements from three satellite instruments: the Atmospheric Infrared Sounder (AIRS) instrument on the *Aqua* satellite (Aumann et al. 2003) and the Microwave Limb Sounder (MLS) instruments on the *Upper Atmospheric Research Satellite* (*UARS*) (Read et al. 2001) and *Aura* satellite (Read et al. 2007). We examine how the PDFs vary between regions and between measurements.

We also examine whether the observed PDFs can be reproduced by simple theoretical models. One is the model recently derived by S06. In this model the relative humidity (RH) is assumed to be determined by uniform subsidence and random remoistening process. These simple assumptions are supported by studies that show tropical humidity can be reproduced using the large-scale to advect a water tracer with no microphysics other than condensation when RH exceeds 100% (e.g., Sherwood 1996; Salathe and Hartmann 1997; Pierrehumbert 1998; Dessler and Sherwood 2000; Galewsky et al. 2005). In the S06 model, the PDF of RH has a simple algebraic form (with exponent related to the ratio of drying to remoistening time). We also consider a generalization of this model that includes an additional parameter, which can be interpreted as a measure of the variability of the remoistening events. It will be shown that this generalized model fits the observations better than the model that S06 proposed and can capture the spatial variations in the PDFs.

The data and theoretical models used in this study are described in the next section. In section 3, the spatial and vertical variations of AIRS PDFs of RH are examined and compared with the theoretical model. Measurements from other instruments are considered in section 4, to check the robustness of the results based on AIRS measurements. Finally, conclusions and future work are discussed in section 5.

2. Data and methods

a. Data

The AIRS data examined are level 2 data retrievals (version 5) that have been binned into a $1° \times 1°$ latitude–longitude grid as in Gettelman et al. (2006). The level 2 data include temperature and water vapor at vertical resolution of around 1–2 km and horizontal

resolution of around 50 km. RH is computed from AIRS water vapor and temperature retrievals as in Gettelman et al. (2006): RH over water is calculated for temperatures >273 K, RH over ice is calculated for temperatures <253 K, and a linear combination is calculated between these temperatures. The AIRS water vapor is an average for a layer between two pressure levels and is archived on 28 levels from the surface to the mesosphere. Following the AIRS convention, each layer is referenced by the pressure at the bottom of the layer; for example, RH at 250 hPa corresponds to the RH averaged from 250 to 200 hPa. We examine AIRS RH from 2002 to 2007 on layers with bottoms between 850 and 200 hPa.

The *UARS* MLS instrument made upper-tropospheric water measurements from September 1991 to July 1997 (with limited coverage after 1994). Measurements were made every 4.1° along an orbit track, with 15 orbits per day, on four pressure surfaces between 147 and 464 hPa (Read et al. 2001). We examine here version 4.9 RH measurements on the 215-hPa surface for northern winters [December–February (DJF)] 1991/92 to 1993/94. These measurements have approximately 3-km vertical resolution, with accuracy and precision of 22% and 10%, respectively (Read et al. 2001).

Aura MLS provides water vapor measurements since July 2004. Water vapor mixing ratios are retrieved from calibrated *Aura* MLS observations (Livesey et al. 2006), and RH with respect to ice is computed from water vapor and temperature retrievals as for AIRS (Read et al. 2007). We examine here *Aura* MLS version 2.2 RH measurements at 215 hPa for the northern winters 2005/06 and 2006/07.

A detailed comparison of AIRS and *Aura* MLS water vapor measurements has recently been reported by Fetzer et al. (2008). They showed that there was a high correlation between the measurements from the two instruments at 250 hPa, although *Aura* MLS had a greater dynamical range, with drier values in dry regions and wetter values in moist regions.

b. Probability density functions

Our primary method of analysis in this paper is examination of PDFs. We form and examine PDFs of RH from each of the above datasets. An important issue when calculating and examining PDFs is the space–time scales included in the PDFs. The characteristics of the PDFs can be sensitive to these scales. This is illustrated in Fig. 1, which shows PDFs of AIRS RH measurements at 400 hPa, for different spatial regions and temporal resolution.

Figure 1a compares PDFs for all AIRS data at 400 hPa within the "whole tropics" (30°S–30°N, 0°–360°E) for
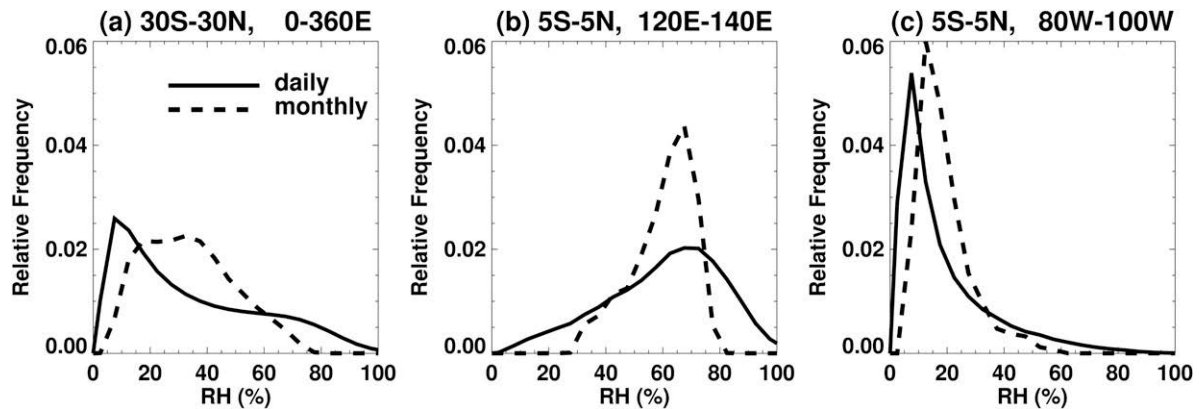
FIG. 1. PDFs of AIRS 400-hPa RH data for (a) the whole tropics (30°S–30°N, 0°–360°E), (b) a tropical convective subregion (5°S–5°N, 120°–140°E), and (c) a tropical nonconvective subregion (5°S–5°N, 80°–100°W). Solid curves show PDFs using daily data, while dashed curves show PDFs from monthly-mean data. All PDFs are formed using data at resolution 1° longitude by 1° latitude.

the 2002/03 to 2006/07 northern winters (December to February), using either daily data (solid curve) or monthly-mean data (dashed curve). In both cases the data are on a 1° longitude by 1° latitude grid. There are significant differences between the two PDFs, even though the same measurements where used. The PDF using daily data is much broader with a large peak at low RH and a second, much broader, peak at high RH. In contrast, the PDF of monthly-mean data is much narrower (no very low or very high values) and has a broader peak at moderate RH. In other words, the averaging process in computing monthly-mean data tends to remove extreme values and produce a PDF with a peak close to the long-term average.

The characteristics of the PDFs also depend on the regions considered. This can be seen by comparing the three panels in Fig. 1. Figures 1b,c show PDFs using same data source as Fig. 1a except for two different 10° latitude by 20° longitude subregions. Again there are significant differences between PDFs using daily and monthly-mean data. There are also, as noted by Ryoo et al. (2008), significant variations between regions. Whereas the PDF for the whole tropics is bimodal, the PDFs for the subregions are unimodal with peaks at high (Fig. 1b) or low (Fig. 1c) RH. These differences are examined in more detail below.

As mentioned in the introduction, the logarithmic dependence of water vapor absorption to the water vapor concentration means that it is important to quantify the full variation in RH. Because of this we focus on PDFs of daily data rather than monthly-mean data, which averages out extreme values. Furthermore, as we would also like to link the characteristics of the PDFs to the processes controlling the humidity distribution and the key processes varying between regions (e.g., Ryoo et al. 2008), we examine the PDFs of

10° latitude by 20° longitude subregions as well as PDFs of the whole tropics.

c. Statistical model

In addition to examining the PDFs from the various measurements, we also compare these observed PDFs with a statistical model for distributions of RH. This model is a generalization of the model derived in S06. As in the S06 model, the generalized model is based on the "time of last saturation" paradigm for tropospheric humidity, in which a parcel's humidity is equal to the lowest saturation value it has experienced since it has left the boundary layer.

In deriving their model, S06 assumed there is uniform subsidence, and the relative humidity R can then be approximated as

$$R = \exp\left(-\frac{t}{\tau_{\text{dry}}}\right), \qquad (1)$$

where t is the time since the parcel was last saturated and τ_{dry} is the uniform drying time by subsidence. S06 further assumed that "remoistening" of parcels occurs by random moistening events, which are independent of the parcel history (i.e., a Poisson process). The PDF of the time t of last saturation is then

$$P(t) = \exp(-t/\tau_{\text{moist}})/\tau_{\text{moist}}, \qquad (2)$$

where τ_{moist} is the mean time of last saturation. The standard deviation of t is also τ_{moist}, which implies that the coefficient of variation (CV) of the time of last saturation is 1. Combining Eqs. (1) and (2) yields the PDF of the relative humidity R:

$$P(R) = rR^{r-1}, \qquad (3)$$

where $r = \tau_{\text{dry}}/\tau_{\text{moist}}$ is the ratio of drying to moistening time. This distribution is a special case of the beta distribution

$$P_{\text{Beta}}(R) = \frac{R^{\alpha-1}(1-R)^{\beta-1}}{B(\alpha,\beta)}, \qquad (4)$$

where $B(\alpha, \beta)$ is the beta function (Wilks 1995). The PDF in (3) is the beta distribution with $\beta = 1$ and $\alpha = r$. The corresponding cumulative distribution is $C(R) = R^r$.

In this paper we consider a two-parameter generalization of the PDF (3). A purely statistical approach would naturally lead to the two-parameter form of the beta PDF where we do not restrict β to be equal to 1 as in (3). This approach, however, cannot be explained in terms of the physics of the underlying phenomena. A more physically based approach that keeps the concepts of uniform subsidence and random remoistening events is to retain Eq. (1) to model the RH in terms of the time t of last saturation and to generalize the PDF in (2) for this time t. The natural generalization in this context is to use the gamma PDF, given by

$$P(t) = \frac{\exp(-kt/\tau_{\text{moist}})t^{k-1}k^k}{\tau_{\text{moist}}^k \Gamma(k)}, \qquad (5)$$

where $\Gamma(k)$ is a gamma function, rather than the exponential PDF in (2). This PDF is still a member of the family of PDFs associated with Poisson processes with the same mean time of last saturation, τ_{moist}, but includes an additional parameter k, which is a measure of the variability of this time. The standard deviation of the time t of last saturation is $\tau_{\text{moist}}/\sqrt{k}$, so the CV of t is $1/\sqrt{k}$. For $k = 1$, the CV is 1 and the PDF reduces to the exponential form (2), while larger k corresponds to less variable events. For the data considered herein, we find that k is less than 10, which corresponds to moderate to large relative variability in the time of last saturation.

The PDF of the relative humidity R is now given by

$$P(R) = \frac{k^k r^k R^{kr-1}}{\Gamma(k)}(-\log R)^{k-1}, \qquad (6)$$

and the cumulative distribution function (CDF) is

$$C(R) = 1 - \gamma\left(-\frac{\log R}{\tau_{\text{dry}}}\right), \qquad (7)$$

where γ is the incomplete gamma function. Both equations reduce to the original S06 distributions in the limit $k = 1$. In the following we refer to distributions of RH given by (3) as the S06 model and distributions given by (6) as the "generalized" model. Below we compare PDFs of the form (6) with the PDFs of the

observations discussed in the previous section. However, before this we examine the characteristics of PDFs given by (6).

Figure 2 illustrates the relationship between the relative subsidence model, the gamma PDF, and the PDF of the RH of the generalized model. In the top plots, the inverse of the relative subsidence model (1) is shown, where the normalized time of last saturation t/τ_{dry} is plotted with respect to the relative humidity R. To represent the range of variability that is implicit in the gamma PDF, we indicate with horizontal lines the mean and the mean plus or minus one standard deviation of the time of last saturation. On the left, these three values are shown for $r = 2$ and $k = 3$, which, as shown later in the paper, roughly corresponds to a convective region of the tropics, while on the right the three values are shown for $r = 0.6$ and $k = 10$, which corresponds to a nonconvective region. The bottom pair of plots shows the PDFs of the RH (6) associated with the values for r and k in the top plots. The purpose of this figure is to illustrate the relationships between the mean and standard deviation of the time of last saturation and the PDF of the RH. It can be seen that short mean times correspond to high RH, while long mean times correspond to low RH, which is physically explained by the drying process during subsidence. The figure also shows that the standard deviation of the time of last saturation causes larger variability in the RH at long mean times as compared with short mean times. This is because the RH curves shown in the top plots have a more highly negative slope at long mean times.

The $P(R)$ for several values of r and k are illustrated in Fig. 3. This shows that a wide range of PDFs can be formed by varying the two parameters. For example, the location of the peak RH varies with r: the peak occurs at RH = 0 for $r < 1$ regardless of the value of k, whereas the peak is at high RH for $r > 1$. It can be shown that the peak of the PDF occurs at

$$R_{\text{peak}} = \exp\left(-\frac{k-1}{kr-1}\right) \qquad (8)$$

for $k > 1$. This shows that as r increases (for fixed k) the peak occurs at larger values, with $P \to 0$ as $r \to 0$ and $P \to 1$ as $r \to \infty$. This differs from the S06 model where the peaks only occur at RH = 0 (for $r < 1$) or 1 (for $r > 1$). Figure 3 also shows that the widths of the PDFs vary with k, with narrower distribution for larger k and smaller r.

It is more common to examine the mean and standard deviation of the RH distribution than the full PDFs, so it is of interest to consider the mean and standard deviation of the above distributions. For distributions given by (6) the mean is
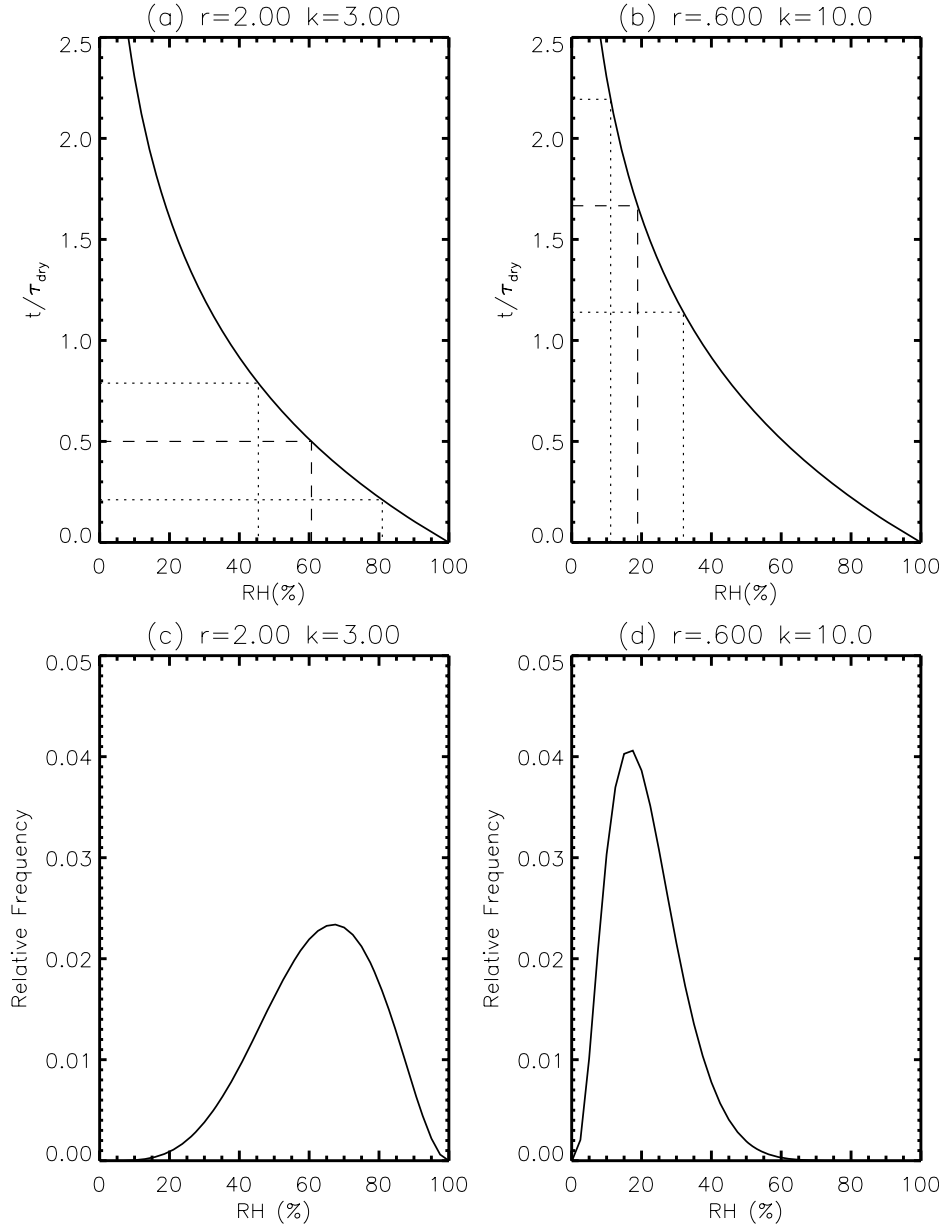
FIG. 2. Relationship between time to last saturation t and RH for (a) $r = 2$, $k = 3$ and (b) $r = 0.6$, $k = 10$. Horizontal dashed lines show mean time to last saturation, and horizontal dotted lines show mean plus and minus standard deviation. (c),(d) Corresponding PDFs of RH.

$$\mu_R = \left(\frac{r}{r + 1/k}\right)^k, \tag{9}$$

and standard deviation is

$$\sigma_R = k^{k/2} r^{k/2} \frac{\sqrt{(kr+1)^{2k} \quad (kr+2)^k k^k r^k}}{(kr+2)^{k/2}(kr+1)^k} \Bigg(\tag{10}$$

The variation of μ_R and σ_R with r and k is shown in Fig. 4. Both μ_R and σ_R depend on r and k, but μ_R de-

pends primarily on r and is only weakly dependent on k (Fig. 4a), and σ_R is primarily dependent on k and only weakly dependent on r (Fig. 4d). As a result, r can be estimated from μ_R using (9) with, say, $k = 2$, and k can be estimated from σ_R using (10) with, say, $r = 2$.

We note that while the generalization in the proposed model can be described mathematically by the gamma PDF (5) with the additional parameter k, there is also an important physical difference from S06. In S06, the air parcels in the tropics are assumed to be remoistened
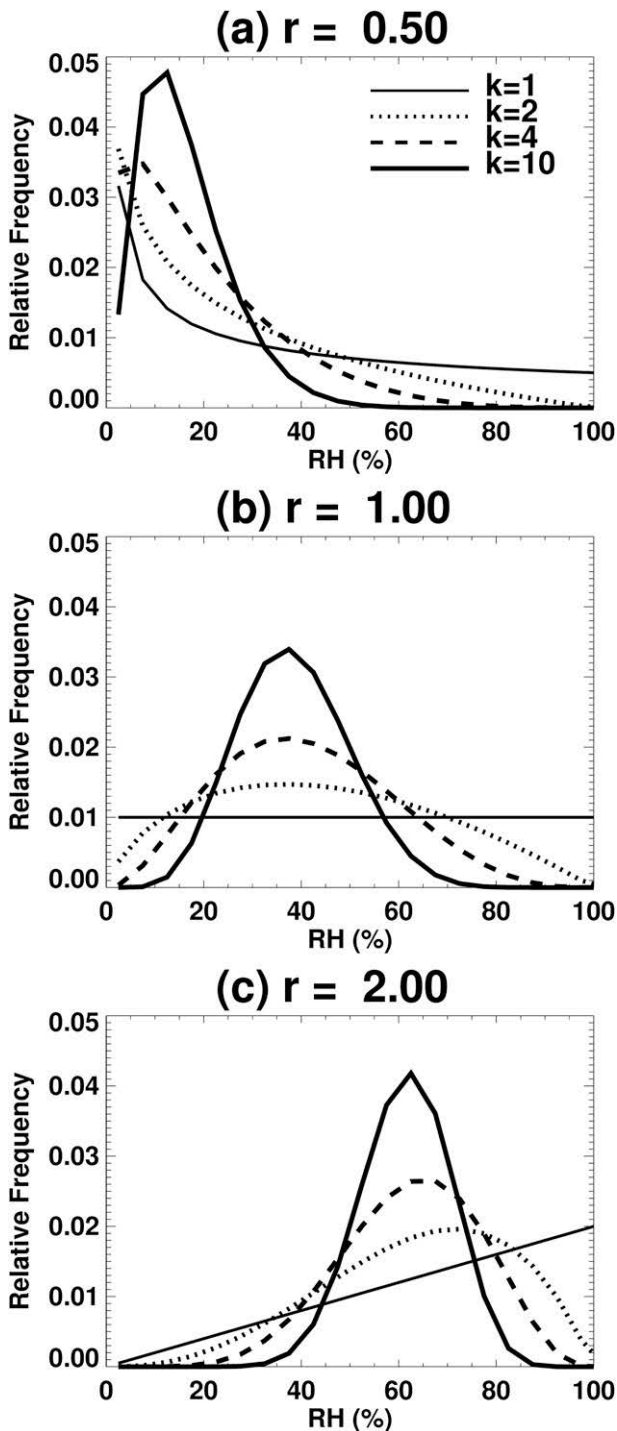
FIG. 3. PDFs of RH for various r and k generated by generalized models: (a) $r = 0.5$, (b) $r = 1$, and (c) $r = 2$ for $k = 1, 2, 4$, and 10, respectively.

according to a Poisson process, which, along with the subsidence model for the RH (1), implies that a single PDF would be used to approximate the RH in the entire tropical region at a given pressure level. In the gener-

alized model, we focus only on the remoistening time at last saturation to allow for any dependence of the mean and standard deviation of this last saturation time on location in the tropics. For instance, in the subtropics, we expect that the last saturation would occur closer to the tropics and that some time would elapse as the air parcels subside into the subtropical region. This would correspond to the right plots of Fig. 2, where the mean value of the time of last saturation is relatively large and the corresponding PDF of the RH has a peak at low RH. The point here is that the remoistening events are location dependent and can no longer be considered as a Poisson process. However, it is not necessary to model the entire process of remoistening; only the time of last saturation is relevant to the RH and we are using the gamma PDF (5) for this event. In the next sections we show how we can model location dependence in the remoistening process using the generalization of S06.

3. AIRS PDFs

We now examine the PDFs from AIRS RH measurements and compare with the above theoretical distributions. We first investigate AIRS measurement for the 250 hPa layer during Northern Hemisphere winter (DJF), and then consider other seasons and altitudes.

a. 250 hPa, DJF

The symbols in Fig. 5 show the PDF (Fig. 5a), and corresponding CDF (Fig. 5b), of AIRS RH for all northern winter data (2002–07) within the tropics and subtropics (30°S–30°N, 0–360°E). The observed PDF is broad and asymmetric with a peak around 20% and long tail of moist air. Such broad distributions of upper-tropospheric RH have also been observed in data from other satellite instruments, for example, MLS, global positioning system (GPS) (S06), and *Odin* (Ekström et al. 2007).

Also shown in Fig. 5 are fits to the AIRS data for the S06 and generalized models. The values of r and, in the case of the generalized model, k, are found by minimizing the mean square error between model and observed PDFs. We choose this method over other standard statistical techniques, such as the maximum likelihood method, because it is not overly sensitive to the low-probability regions of the PDF. These comparisons show that the generalized model is a better fit to the observed PDF and CDF than the original S06 model. In particular, the generalized model can reproduce the peak of the PDF at RH close to 20%, whereas the peak of the S06 model occurs at RH = 0% (as noted in the discussion of Fig. 1, the peaks of the PDFs from the S06 model can only appear at RH equal to 0% or 100%).
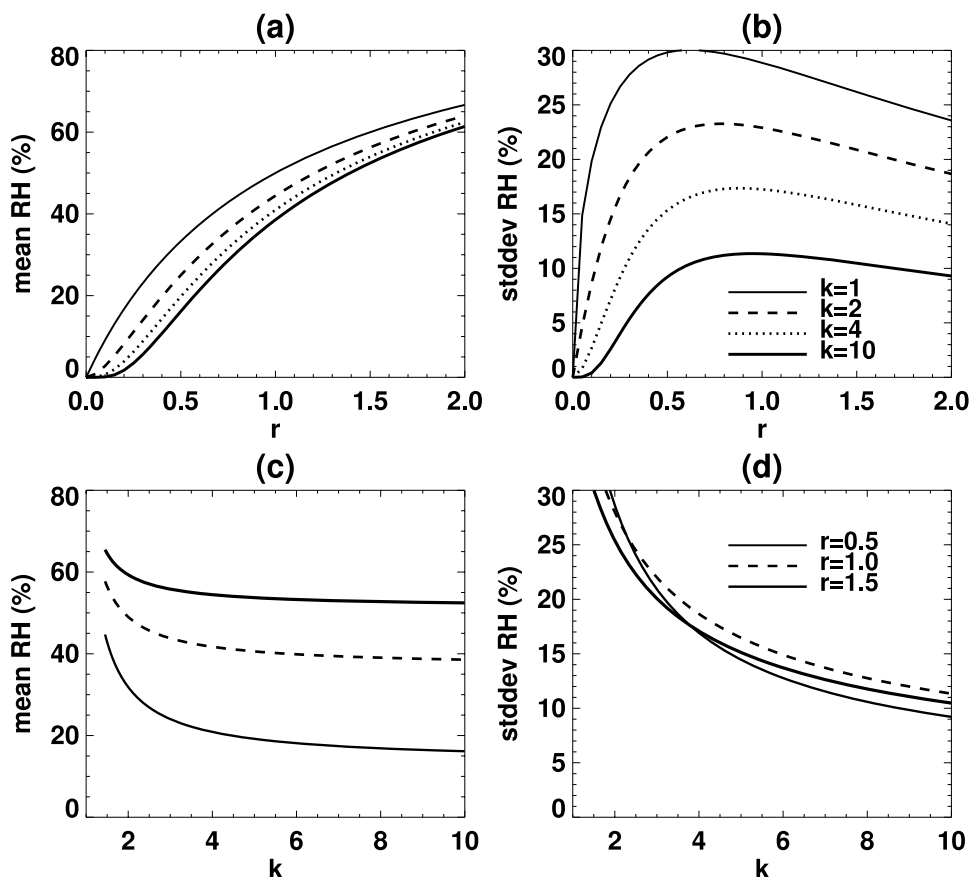
FIG. 4. Plots of (a) mean (μ_R) and (b) standard deviation (σ_R) vs r, and (c) mean (μ_R) and (d) standard deviation (σ_R) vs k, for generalized model.

This comparison indicates that inclusion of an additional parameter in the statistical model greatly improves the agreement with the AIRS data. Further evidence of this improved agreement is presented below. Note that S06 examined CDFs from GPS and the two MLS instruments but not AIRS data. The agreement between the S06 model and GPS PDFs is better than that in Fig. 5b, but the disagreement with MLS PDFs is similar to that for AIRS. The differences between different datasets are examined further in section 5.

We now consider the PDFs for smaller subregions than the whole tropics. As the differences between distributions are more visible if we are considering the PDFs rather than the CDFs, in the remainder of the paper we focus on the PDFs of RH, rather than the CDFs, as considered by S06, but similar results are obtained if CDFs were used (i.e., the best-fit values of r and k are very similar for fits to PDFs or CDFs). Figure 5 shows the PDFs of AIRS RH for six 10° latitude by 20° longitude regions in the tropics (5°S–5°N) and the subtropics (15°–25°N). As noted by Ryoo et al. (2008), the PDFs vary between regions, both with longitude and

between the tropics and subtropics. The location of the peaks of the PDFs varies from around 20% to around 60%, and the width and skewness of the distributions also vary. The fits to the AIRS PDFs for the various subregions for the S06 ($k = 1$) and generalized (variable k) models are also shown in Fig. 6. The generalized model can fit the data for all subregions. This includes not only the peak values but also the range and skewness of the PDFs. Some differences between generalized model and observed PDFs can be seen for high RH, especially in the tropical eastern Pacific (5°S–5°N, 80°–100°W). However, these are relatively small differences.

The S06 model cannot match the variations in the peak of the observed PDFs for subregions. As noted in section 2c, the S06 model was originally developed to model the PDF of the RH for the entire tropics. Hence the S06 model is not expected to produce a close fit to the PDFs of the RH in subregions. However, it will be shown later that the S06 model does yield good results for r in most subregions.

Figure 6 shows that the PDFs of AIRS RH for different tropical or subtropical regions can be represented
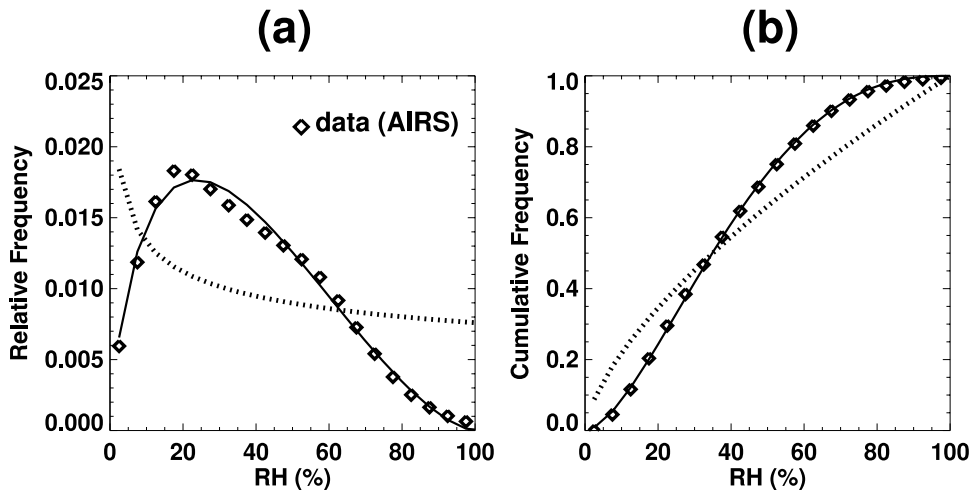
(a)　　　　　　　　　　(b)

Fig. 5. (a) PDFs of 250-hPa RH and the whole tropics (30°S–30°N, 0°–360°E) from AIRS data (symbol) and fits by S06 ($k = 1$; dotted) and generalized (variable k; solid) models. (b) As in (a) except for CDFs.

by the theoretical generalized model. The variations in the PDFs can hence be summarized by variations in r and k. Figure 7 shows the longitudinal variation of r (Figs. 7a,d), k (Figs. 7b,e), and error ϵ (Figs. 7c,f) (see below) for the S06 and generalized model fits to PDFs for 10° by 20° regions in the (upper panels) subtropics or (lower panels) tropics. As could be expected from Fig. 6, both r and k vary with longitude and latitude.

The value of r for the tropics is generally larger than in the subtropics, and the longitudinal variation of r is much larger in the tropics than the subtropics (Figs. 7a,d). The r for the S06 and generalized models have very similar spatial variations and even quantitative agreement, except in tropical Indian (~40°E), western Pacific (~120°E), and Atlantic (~50°W) Oceans. This is somewhat surprising given the different shapes of the PDFs for the S06 and generalized models (e.g., Fig. 6). The similarity in r between the fits using the S06 and generalized model can be understood in terms of relationships between r and the mean value μ_R. As discussed in section 2, r is closely related to μ_R, with only weak sensitivity to k. Hence, for both $k = 1$ (S06 model) and $k > 1$ (generalized model), r will depend primarily on the mean, and not other characteristics, of the RH distributions.

The parameter k also varies with both longitude and latitude; see Figs. 7b,e. In the subtropics k varies between 2 and 6, while in the tropics k varies from 3 to 10. In both the tropics and subtropics the longitudes with maximum in k are generally the longitudes where r is a minimum; for example, in the tropics large k and small r occur around 60°E and 90°W.

The r and k shown above were determined by minimizing the error between the observed and modeled

PDF. To estimate the uncertainty in these estimates, a moving-blocks bootstrap analysis (Künsch 1989) has been performed (where moving time blocks of data are used to account for correlation in time). The vertical bars in Fig. 7 show the uncertainty ($\pm 1\ \sigma$) in the calculated r and k. The uncertainty for r is very small in the both subtropics and tropics and much smaller than the spatial variations in r. The uncertainty in k is larger but still less than spatial variations in k.

As discussed above, the generalized model fits the data better than the S06 model. To quantify this, and the spatial variations in how well the models fit the data, we calculate the root-mean-square error ϵ between the PDFs of the model and data. Figures 7c,f show that the error for the generalized model is about 10 times less than that of S06 model. Also, longitudinal variations in ϵ for the S06 model are very similar to the variations of k of generalized model. This is because the S06 model is identical to the generalized model when $k = 1$, and as k of the generalized model increases the S06 model deviates from the data, resulting in larger ϵ when k is larger.

To examine the spatial variations of r and k further we compare maps of these fields. Figures 8a,b show that r and k have similar spatial variations, with small (large) k in regions of large (small) r. Specifically, there is large r and small k in the tropical western Pacific (5°S–5°N, 120°E) and tropical America (5°S–5°N, 60°W), and small r and large k in the tropical eastern Pacific (5°S–5°N, 120°W) and the northern subtropical mid-Pacific (15°–25°N, 150°E). Maps of μ_R and σ_R calculated from the AIRS data are shown in Figs. 8c,d. As expected from Fig. 4 and the related discussion, there is a strong resemblance between maps of μ_R and σ_R to those of
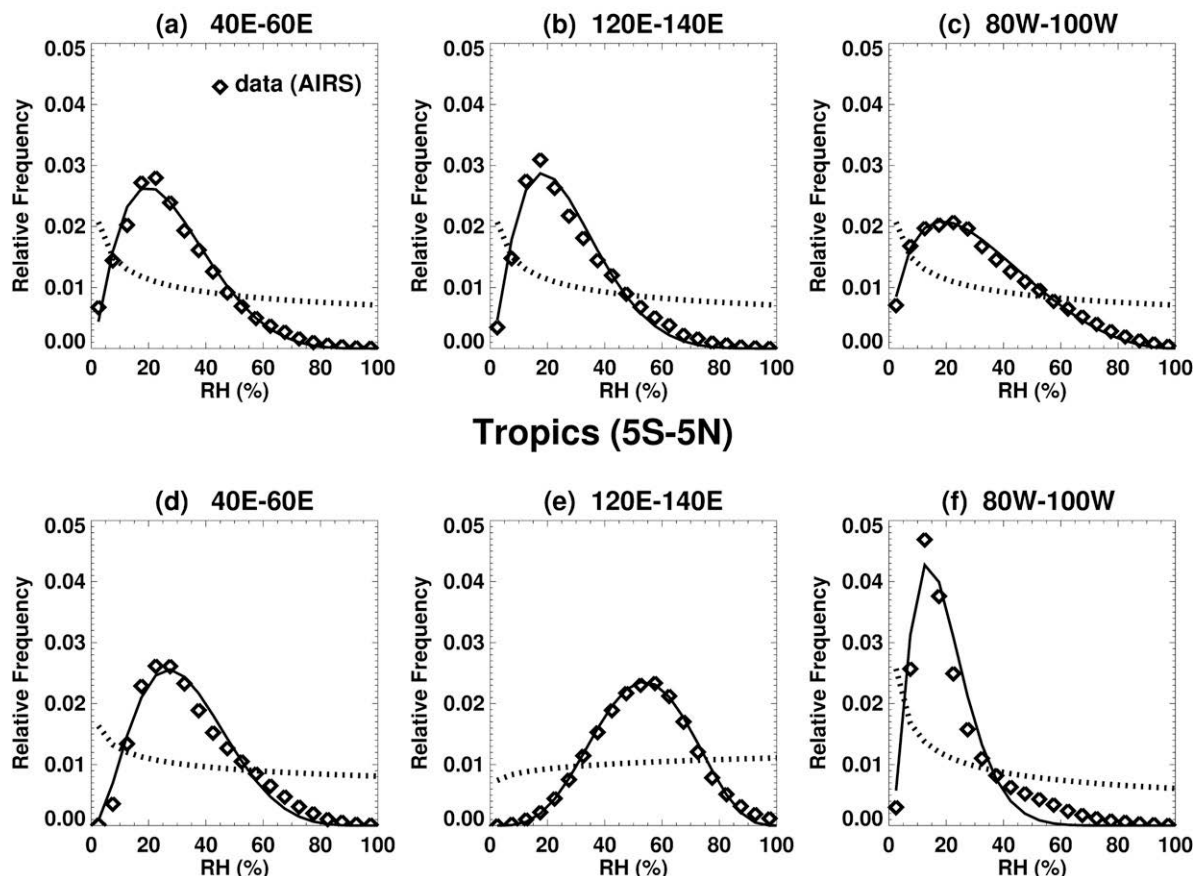
Subtropics (15N–25N)

Tropics (5S–5N)

FIG. 6. As in Fig. 3a except for subregions in the subtropics (15°–25°N): (a) 40°–60°E, (b) 120°–140°E, and (c) 80°–100°W; and the tropics (5°S–5°N): (d) 40°–60°E, (e) 120°–140°E, and (f) 80°–100°W.

r and k, respectively; that is, there is large μ_R where r is large and large σ_R where k is small.

The variations in r and k could provide insight into the variations in the characteristics of the moistening processes. If the drying time is assumed constant, large r and small k indicates rapid, random remoistening, whereas small r and large k implies slower, more regular moistening processes. There are large r and small k, and hence by the above arguments rapid, random remoistening, in the tropical convective regions. In contrast, there are small r and larger k in the dry regions, indicating slower more regular remoistening.

We conjecture that these variations in the remoistening are consistent with our understanding of the physical processes. In the tropical convective regions the moistening is thought to occur by direct rapid moistening by vertical transport in convective systems, whereas lateral mixing by "large scale" advection plays a larger role in the remoistening the drier tropical and

subtropical regions (e.g., Sherwood 1996; Salathe and Hartmann 1997; Pierrehumbert 1998; Waugh 2005; Ryoo et al. 2008). This lateral mixing is produced by transient wave activity (Pierrehumbert and Roca 1998), including Rossby wave breaking along the tropopause (Waugh 2005; Ryoo et al. 2008), and is slower and more regular than convection. The process is not so regular as to be periodic (which would correspond to k of the order of 100) but is considerably less random than processes where k is of order 1.

b. Seasonal and altitudinal variations

The above analysis considered only northern winter data at 250 hPa. We now examine the seasonal and altitudinal variations of AIRS PDFs and whether these PDFs are still well fit by the generalized model.

Figure 9 shows PDFs of AIRS data for the whole tropics (30°S–30°N, 0°–360°E) for different seasons and at several different altitudes. At all levels there are only
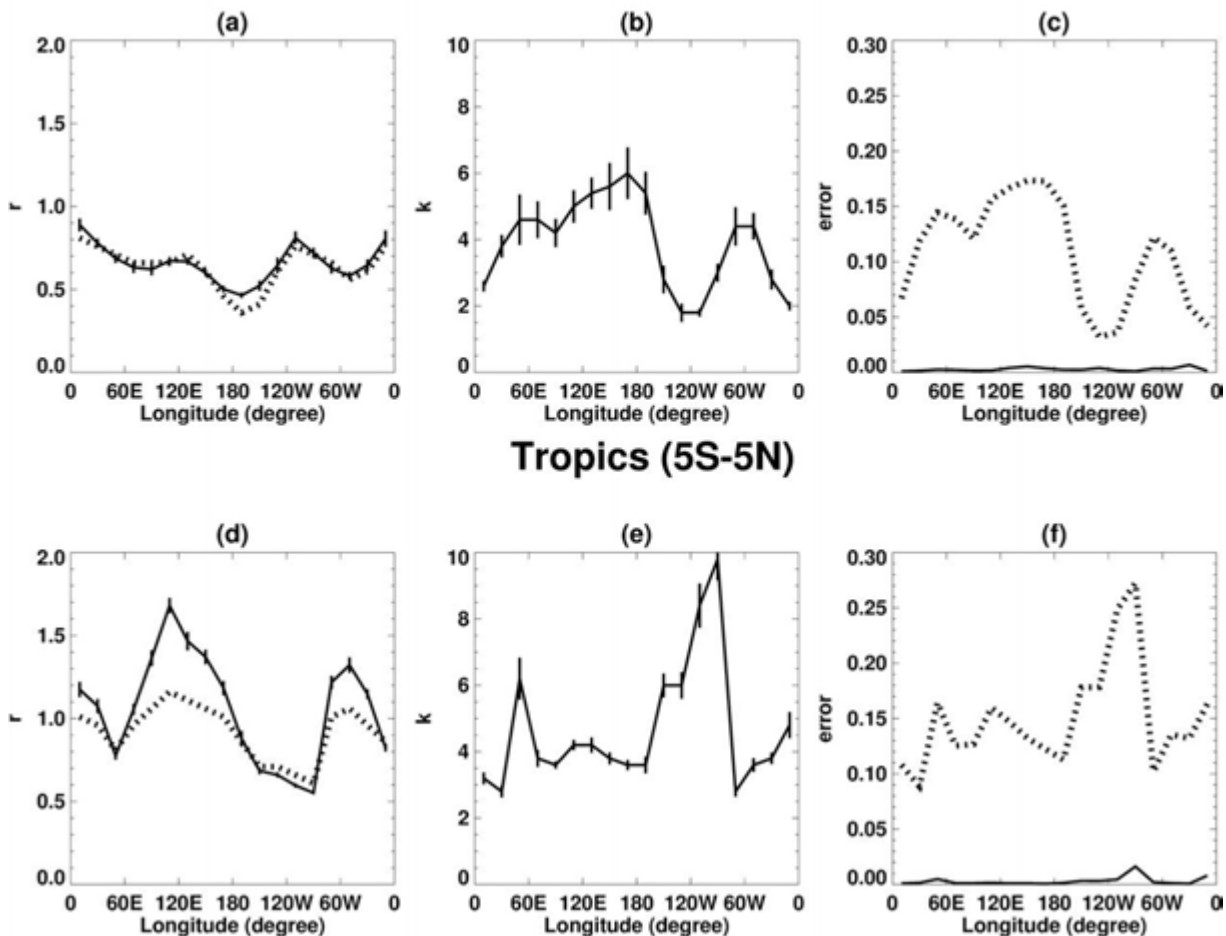
Subtropics (15N–25N)

Tropics (5S–5N)

FIG. 7. Longitudinal variation (a),(d) r; (b),(e) k; and (c),(f) error ϵ for S06 (dotted curves) and generalized (solid) models for (top) subtropics and (bottom) tropics. The vertical bars indicate the one-sigma bounds computed by the moving-blocks bootstrap distribution.

weak seasonal variations, which are reasonable for the tropics. There are, however, large variations in the shape of PDFs with altitude. At 400 and 600 hPa the peak occurs at or less than RH = 10%, which is much drier than the peak at 250 hPa, whereas at 850 hPa there is limited dry air and the dry peak occurs around RH = 40%. In contrast to 250 hPa, the PDFs are bimodal at 400, 600, and 850 hPa, with a second moist peak at RH $\simeq$ 70%–80%. Also shown in Fig. 9 are the fits to the data using the generalized model. Because the seasonal variations are small, we only show the fit for DJF data. The model can capture the general characteristics of the vertical variations, in particular the variation in the peak values. As at 250 hPa, the generalized model is a much better fit than the S06 model at the lower levels (not shown). However, the generalized model cannot reproduce the observed bimodal PDFs, and the dis-

agreements between the observed and generalized model is largest when the observed PDFs are most bimodal.

The PDFs shown in Fig. 9 come from the collection of dry and moist RH over the whole tropical region, which includes moist air from the convective region and dry air from the nonconvective region. When we look at the PDFs for smaller 10° latitude by 20° longitude subregions, most of them are unimodal. For example, in tropical convective regions [e.g., the western Pacific (5°S–5°N, 100°–140°E)], the PDFs have a peak in high RH, while in nonconvective regions like the eastern Pacific (5°S–5°N, 120°–160°W) the peak of PDF is at low RH. These observed PDFs for the subregions can be well fit by the generalized model. When these different regions are combined, the resulting PDF is simply the average of the PDFs of all subregions. Hence, given that
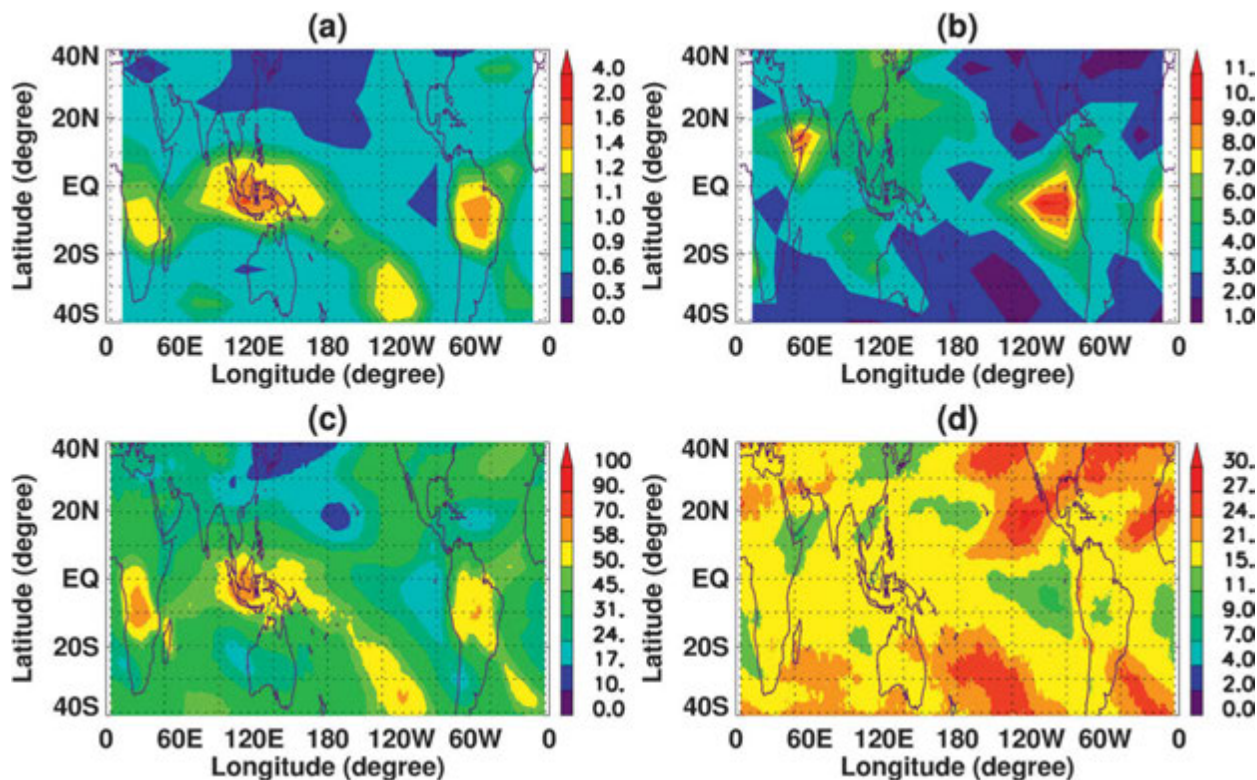
FIG. 8. Maps of (a) best-fit r, (b) best-fit k, (c) mean (μ_R), and (d) standard deviation (σ_R) of 250-hPa AIRS RH.

the peaks of the PDFs of convective and nonconvective subregions are at different RH values, bimodality in the PDF of the combined region is expected. There are some subregions at the edge of the tropical convective regions where bimodal PDFs occur (e.g., 5°S–5°N, 40°–60°W), but the distinction between low values and high values is small (not shown). In addition, this bimodal behavior is tightly related to temporal variations because of the movement of convection into or out of a region, where in this case the PDF is a temporal rather than a spatial average of PDFs with variable locations of peaks.

The bimodal distributions in the midtroposphere are consistent with Zhang et al. (2003), who observed bimodal features in the PDFs of precipitable water using monthly-mean data. Furthermore, if monthly-mean AIRS data rather than daily values are used, the PDFs for midtropospheric RH look similar to the PDFs of precipitable water averaged over 500–300 hPa shown in Zhang et al. (2003); for example, compare Fig. 1a with Fig. 8 of Zhang et al. (2003). [The difference is due to differences in measurements, regions of interest, and time periods. The similarity is much clearer when we compare 3-monthly-mean data with them (not shown).]

The vertical variations of r, k, and ϵ for PDFs, for the whole tropics and whole year, are shown in Fig. 10. There is a minimum in r in the midtroposphere, for both

the S06 and generalized models. A midlevel minimum in the mean RH from AIRS has already been reported (Gettelman et al. 2006; Ryoo et al. 2008), and the minimum in r could be expected given the close relationship between r and μ_R. Also, S06 found a minimum at the same altitude in their calculations of r from GPS data.

The parameter k also varies in the vertical, with an increase with altitude above 500 hPa. As larger k reflects less variable remoistening processes, so this increase of k suggests the moistening processes in the upper troposphere are more affected by more regular and relatively slow large-scale process such as subsidence, rather than by rapid moistening by convective updraft from the surface.

The error ϵ between model and data is shown in Fig. 10c. As at 250 hPa, ϵ for the S06 model is much larger than the generalized model, and the variation of ϵ for the S06 model is similar to the variation in k. The vertical variation of ϵ for the generalized model differs from that of r and k, with local maximum between 300 and 400 hPa and between 600 and 700 hPa. These are altitudes where the PDFs are most bimodal.

Similar vertical variations of r and k occur for subregions, although there are variations with longitude (see Fig. 11). As expected from the discussion in section 2, the
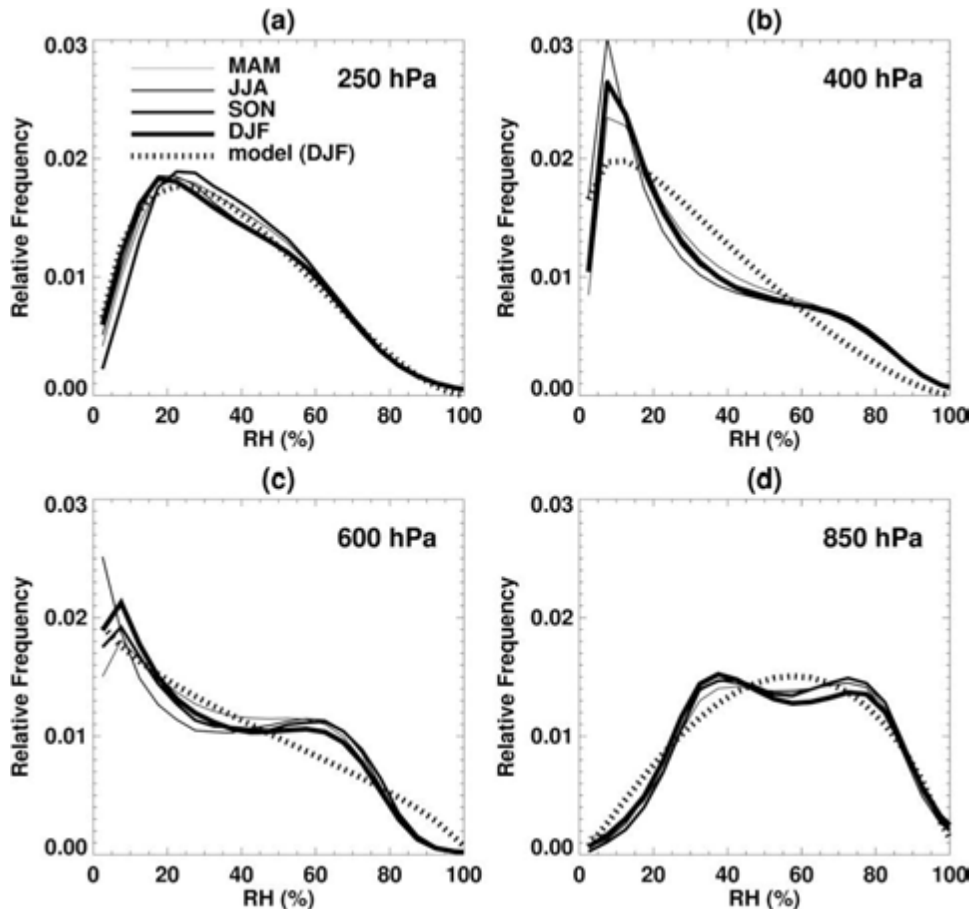
FIG. 9. PDFs for the whole tropics (30°S–30°N, 0°–360°E) for AIRS data at 250, 400, 600, and 850 hPa. Different colors are for different seasons, and dotted curve is fit to DJF data for generalized model.

vertical and longitudinal variations of r and k are similar to those of μ_R and σ_R, respectively (see Fig. 1 of Ryoo et al. 2008). It is interesting to note that in the tropical convective regions (100°–140°E, 40°–60°W) there are local minima in both r and k at midlevels (~400 hPa), whereas in nonconvective regions there is a maximum in k at midlevels.

The vertical variations of r and k in convective regions are consistent with analysis of radiative processes and the energy balance, which show a minimum in convective detrainment at midlevels (Hartmann and Larson 2002; Folkins et al. 2002, 2008). This analysis indicates that in convective regions the air at and above 200 hPa is composed mainly of very moist air parcels that have just detrained from convection, whereas around 400 hPa there is a combination of moist air from recent detrainment and very dry air that has subsided from 200 hPa. As a result the remoistening time at midlevels is longer than aloft, resulting in smaller r (and μ_R). Also, there is larger variability in the moisture and more regularity in remoistening at midlevels, resulting in a larger k.

4. Other data

Having examined PDFs from AIRS we now consider the PDFs of RH measurements from other instruments to test the robustness of the above results. We first compare with measurements made by *UARS* MLS (1992–94) and *Aura* MLS (2005–07) instruments. The latter overlaps with the AIRS data record enabling a comparison of PDFs for the same time periods. We also compare our results with those shown in S06 for GPS data.

For our analysis of MLS measurements we focus on northern winter (DJF) measurements at 215 hPa (which can be compared with the AIRS 200–250-hPa layer). Figure 12 shows the PDFs of AIRS, *UARS* MLS (1992–94), and *Aura* MLS (2005–07) RH for subregions in the tropics (5°S–5°N) and subtropics (15°–25°N). Here two different AIRS PDFs are shown. One was formed using all available data and the other using only data sampled at the same locations as *Aura* MLS. PDFs from different datasets show good agreement except the
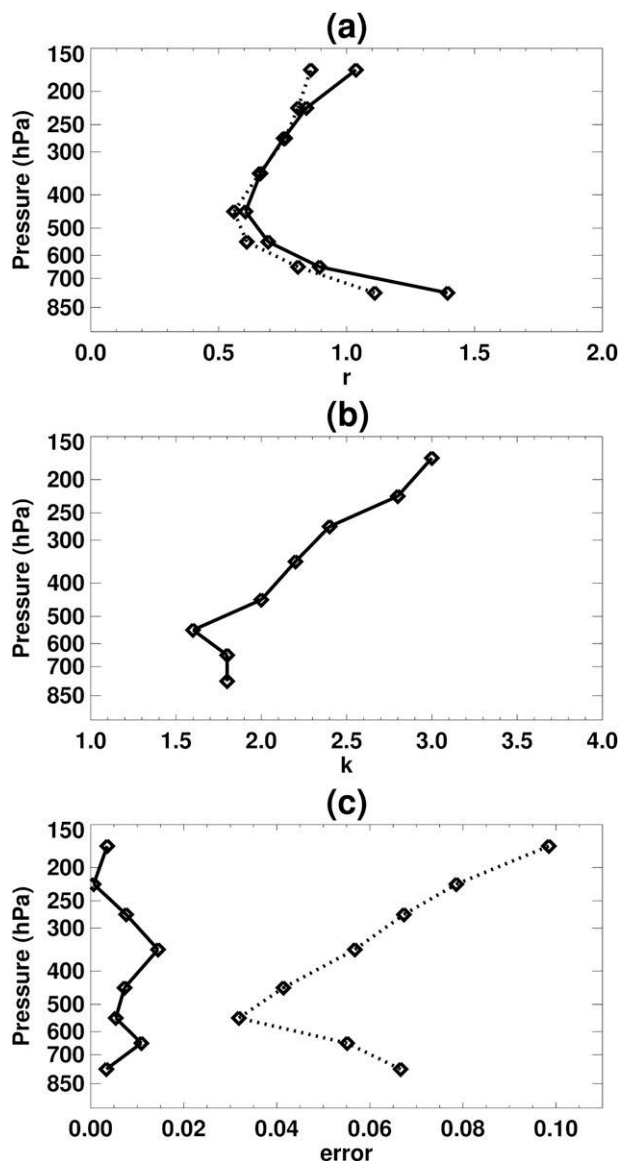
FIG. 10. Vertical variation of (a) r, (b) k, and (c) error ϵ for fits to PDFs of whole tropics and whole year for S06 (dotted) and generalized (solid) models.

tropical convective region (5°S–5°N, 120°–140°E). In this region the AIRS PDFs are narrower, with peak around 60%, whereas the *UARS* MLS and *Aura* MLS PDFs are broader, with low values less than 20% and high values larger than 100% (see also Read et al. 2007). According to Fig. 4d, this would imply that k from AIRS should be considerably larger than k from both MLS measurements.

Figure 13 compares the longitudinal variation of r and k for the tropics (5°S–5°N) and subtropics (15°–25°N) for PDFS of AIRS, *Aura* MLS, and *UARS* MLS measurements shown in Fig. 12. Consider first the parameter r. There is good agreement in r from all three datasets

even though they cover different years. All three datasets show generally higher values of r in the tropics than the subtropics, with largest values in the tropical convective regions (around 120°E and 60°W), and larger longitudinal variations in the tropics than the subtropics. The largest disagreements between the values of r from the different measurements are in the tropical locations with local maximum in r, where r from AIRS is generally larger than from both MLS measurements. This is true even if the same measurement locations are used for the AIRS and *Aura* MLS PDFs, indicating that this is a difference in the measurements and not due to differences in the sampling or different years.

The agreement between k from the different datasets is not as good as for r. There is qualitative agreement in the longitudinal variations of k, but there are quantitative differences. In the tropics, k from AIRS are consistently larger than those from both MLS measurements, even when sampling the same air as *Aura* MLS. The largest difference between AIRS and MLS occur in tropical convective regions (see above). There is better agreement between the two MLS datasets, although differences occur when k is larger (with larger k from *Aura* MLS). The differences in subtropical k among the datasets are not as consistent as the tropics, but the general tendency is the same; for example, k is generally larger from AIRS than MLS.

The above comparison is focused only on upper-tropospheric measurements. As a check on the robustness of the vertical variations, we briefly compare our results with the GPS data shown in S06. As discussed above, r from the GPS data show a minimum at the same height as that from AIRS (see Fig. 6 of S06). The values from GPS are smaller than those of AIRS; for example, for the S06 model the r at 400 hPa is 0.42 from GPS (for January 2002 measurements) compared with 0.53 from data (for DJF 2002/03–2006/07 measurements). S06 fit their model to CDFs rather than PDFs and used a different criterion to determine the best-fit r, but tests show that neither of these causes significant differences in the estimates of r. The difference in our estimate r from AIRS and S06 calculations of r from GPS data is thus due to actual differences in the PDFs from the two datasets.

The exact cause of the above differences between the different datasets is not known, but the broader PDFs from *Aura* MLS than from AIRS is consistent with the analysis of Fetzer et al. (2008). Even though there are some quantitative differences in the values of r and k for PDFs from different datasets, there is overall good agreement in the spatial variations of r and k, both horizontally and vertically. This gives us some confidence in general conclusions primarily based on analysis of the AIRS data.
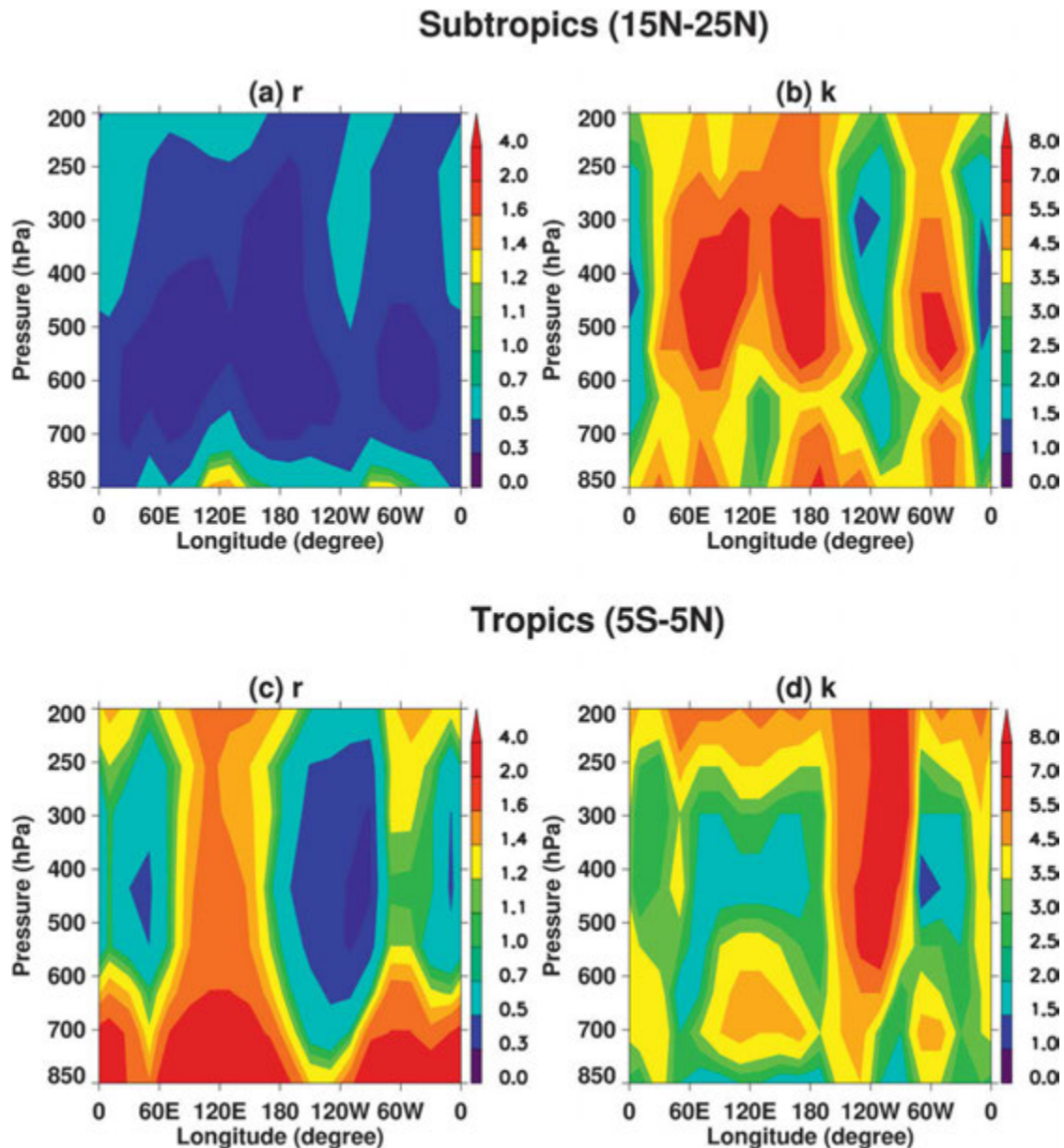
Subtropics (15N–25N)

Tropics (5S–5N)

FIG. 11. The cross section of longitudinal vs altitudinal variation of (a) *r* and (b) *k* for the subtropics (15°–25°N) and tropics (5°S–5°N).

5. Conclusions

Measurements of tropospheric relative humidity (RH) from three different satellite instruments indicate that the probability density functions (PDFs) of daily RH are broad and non-Gaussian. This applies not only for PDFs of the whole extended tropical region (30°S–30°N) but also for PDFs of smaller 10° latitude by 20° longitude subregions. Although the ''local'' PDFs are all broad, the location of the peak, the skewness, and the width vary between the tropics and subtropics, within the tropics or subtropics, and with altitude.

The observed PDFs for all subregions can be well fit using a simple statistical model that is a generalization of that proposed by S06. This model assumes the RH is determined by a combination of drying by uniform subsidence and random remoistening events and has two parameters: *r*, the ratio of drying time (via subsidence) and remoistening time, and *k*, a measure of variability of the remoistening time.
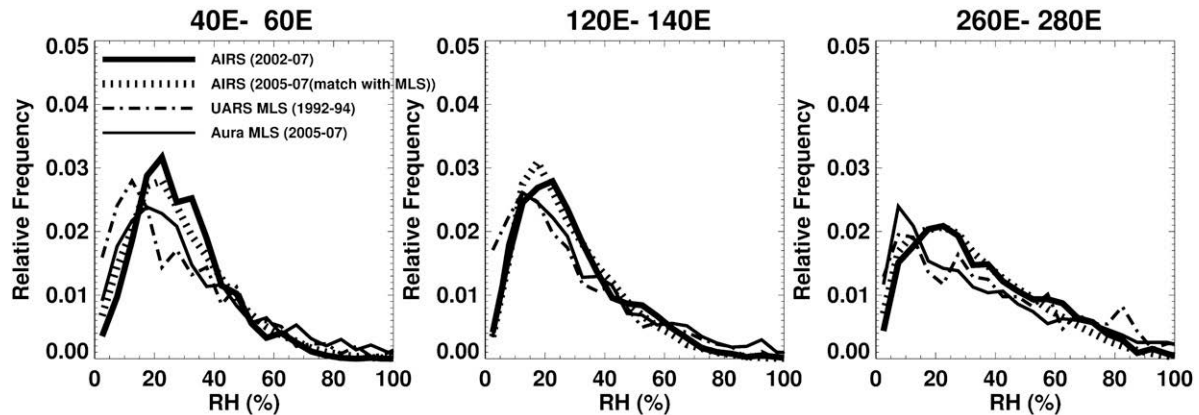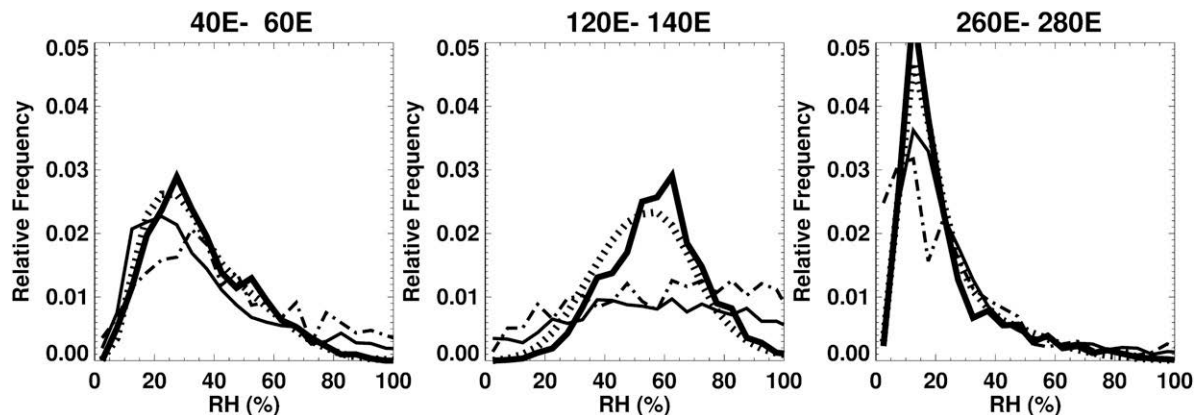
(a) Subtropics (15N-25N)

(b) Tropics (5S-5N)

FIG. 12. PDFs for three subregions in the (a) subtropics (15°–25°N) and (b) tropics (5°S–5°N) for 250-hPa AIRS, *Aura* MLS, and *UARS* MLS measurements, respectively. Dashed curves are for AIRS data sampled at *Aura* MLS locations.

The parameters r and k not only provide a concise way to characterize the RH distributions, but also may provide insight into the processes controlling the RH distributions. In the tropical convective regions there is large r and small k in the upper troposphere, indicating rapid, more variable remoistening in these regions. In contrast, in dry regions in the subtropics and tropical eastern Pacific there is small r and large k, indicating slower, more regular remoistening there.

We conjecture that these variations in the remoistening process are consistent with our understanding of the physical processes in different regions. Previous studies have shown that convection and vertical mixing play the key role in regulating humidity near tropical convective regions, but remoistening in the subtropics comes from lateral advection of moist air from convective regions (e.g., Sherwood 1996; Salathe and Hartmann

1997; Pierrehumbert 1998; Dessler and Sherwood 2000; Galewsky et al. 2005; Dessler and Minschwaner 2007). Thus, in tropical convective regions we expect direct remoistening by rapid, random vertical motions, whereas in dry, nonconvective regions the remoistening occurs by slower, more regular lateral mixing by large-scale advection. The r and k that fit the observed PDFs also vary in the vertical. In the tropics r and k both have a midlevel (300–500 hPa) minimum, indicating slower and more regular remoistening in the midtroposphere. This is consistent with a midlevel minimum in convective de-trainment and midlevel air being a mixture of recently detrained moist air and very dry air that has subsided from below (Folkins et al. 2002).

Although the satellite datasets considered here show a consistent spatial variation in the PDFs, there are some quantitative differences. For example, the MLS
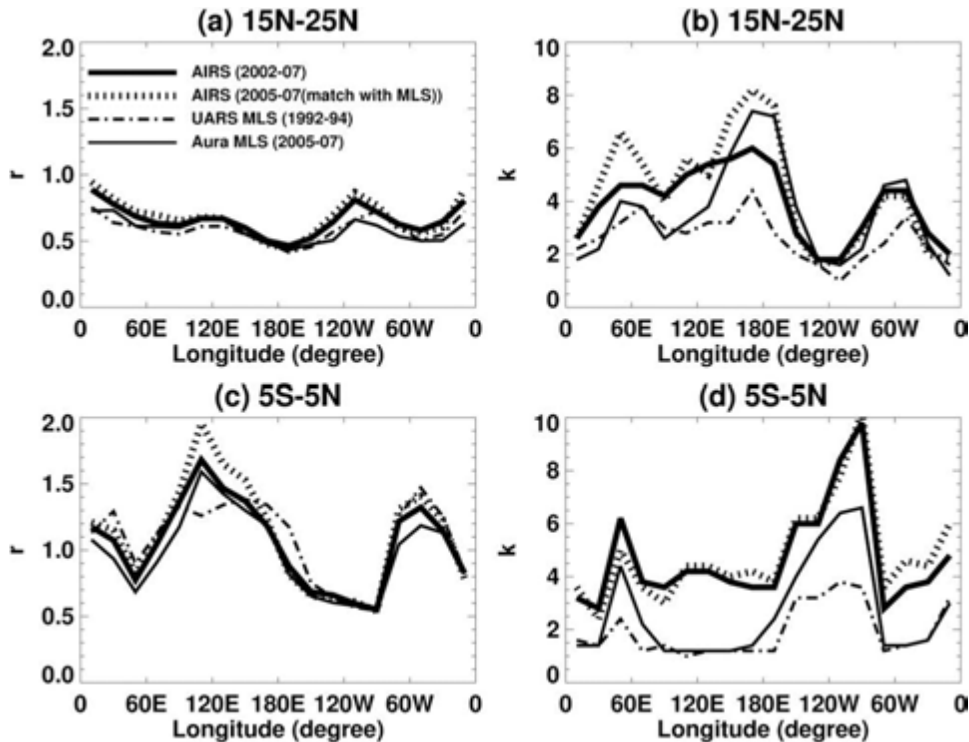
FIG. 13. Longitudinal variation (a),(c) *r* and (b),(d) *k* for generalized model fit to AIRS, *Aura* MLS, and *UARS* MLS measurements. Dashed curves are for AIRS data sampled at *Aura* MLS locations.

PDFs are generally broader than the AIRS PDFs, with a higher probability of low RH and high RH in the MLS data (and as a result k is smaller from MLS data). The magnitude of these differences varies with location, and in some regions the differences are very small (see Figs. 10 and 11). The cause of these differences needs to be examined further. It will also be important to consider other water vapor datasets, in particular those from in situ measurements. Luo et al. (2007) recently presented PDFs of upper-tropospheric (UT) RH from the Measurement of Ozone and Water Vapor by Airbus In-Service (MOZAIC) aircraft program. These PDFs are often bimodal, and appear to differ from the AIRS and MLS PDFs for similar regions and seasons. More analysis is needed to quantify and understand the differences between different datasets.

As discussed above, the spatial variations in r and k appear consistent with our understanding of the physical processes controlling RH distribution. However, this is primarily a qualitative comparison and a more quantitative link between the different physical processes and the parameters r and k is needed. One approach to do this might involve using trajectory-based water vapor simulations. Previous studies have shown that trajectory-based simulations can reproduce upper-tropospheric RH observations (Pierrehumbert and Roca

1998; Dessler and Sherwood 2000; Waugh 2005; Dessler and Minschwaner 2007). Analysis of these calculations would enable some of the assumptions used to derive the statistical model to be tested and would provide an opportunity to examine the origin of moisture and the control mechanisms.

There are several potential uses of the statistical model derived here. Given that the model parameters r and k concisely characterize the RH distributions, fitting this model to climate model output may be useful for quantifying differences in RH distributions between climate models and observations. Although, care will be needed to make sure similar spatial and temporal scales are used for both the data and models. The statistical model may also provide a concise way to characterize any temporal changes in simulated RH distributions (e.g., in simulations with increasing greenhouse gases). Finally, the statistical model may also be useful for exploring how changes in physical processes could alter the RH distribution.

REFERENCES

Aumann, H. H., and Coauthors, 2003: AIRS/AMSU/HSB on the Aqua Mission: Design, science objectives and data products. *IEEE Trans. Geosci. Remote Sens.,* **41,** 253–264.

Dessler, A. E., and S. C. Sherwood, 2000: Simulations of tropical upper tropospheric humidity. *J. Geophys. Res.,* **105,** 20 155–20 163.

——, and K. Minschwaner, 2007: An analysis of the regulation of tropical tropospheric water vapor. *J. Geophys. Res.,* **112,** D10120, doi:10.1029/2006JD007683.

Ekström, M., P. Eriksson, B. Rydberg, and D. P. Murtagh, 2007: First Odin sub-mm retrievals in the tropical upper troposphere: Humidity and cloud ice signals. *Atmos. Chem. Phys.,* **7,** 459–469.

Fetzer, E., and Coauthors, 2008: Comparison of upper tropospheric water vapor observations from the Microwave Limb Sounder and Atmospheric Infrared Sounder. *J. Geophys. Res.,* **113,** D22110, doi:10.1029/2008JD010000.

Folkins, I., K. K. Kelly, and E. M. Weinstock, 2002: A simple explanation for the increase in relative humidity between 11 and 14 km in the tropics. *J. Geophys. Res.,* **107,** 4736, doi:10.1029/2002JD002185.

——, S. Fueglistaler, G. Lesins, and T. Mitovski, 2008: A low-level circulation in the tropics. *J. Atmos. Sci.,* **65,** 1019–1034.

Galewsky, J., A. Sobel, and I. Held, 2005: Diagnosis of subtropical humidity using tracers of last saturation. *J. Atmos. Sci.,* **62,** 3353–3367.

Gettelman, A., W. D. Collins, E. J. Fetzer, A. Eldering, F. W. Irion, P. B. Duffy, and G. Bala, 2006: Climatology of upper-tropospheric relative humidity from the Atmospheric Infrared Sounder and implications for climate. *J. Climate,* **19,** 6104–6121.

Gierens, K., U. Schumann, M. Helten, H. Smit, and A. Marence, 1999: A distribution law for relative humidity in the upper troposphere and lower stratosphere derived from three years of MOZAIC measurements. *Ann. Geophys.,* **17,** 1218–1226.

Hartmann, D. L., and K. Larson, 2002: An important constraint on tropical cloud–climate feedback. *Geophys. Res. Lett.,* **29,** 1951, doi:10.1029/2002GL015835.

Held, I. M., and B. J. Soden, 2000: Water vapor feedback and global warming. *Annu. Rev. Energy Environ.,* **25,** 441–475.

Künsch, H. R., 1989: The jackknife and the bootstrap for general stationary observations. *Ann. Stat.,* **17,** 1217–1241.

Livesey, N. J., L. J. Kovalenko, R. J. Salawitch, I. A. MacKenzie, M. P. Chipperfield, W. G. Read, R. F. Jarnot, and J. W. Waters, 2006: EOS Microwave Limb Sounder observations of upper stratospheric BrO: Implications for total bromine. *Geophys. Res. Lett.,* **33,** L20817, doi:10.1029/2006GL026930.

Luo, Z., D. Kley, R. H. Johnson, and H. Smit, 2007: Ten years of measurements of tropical upper-tropospheric water vapor by MOZAIC. Part I: Climatology, variability, transport, and relation to deep convection. *J. Climate,* **20,** 418–435.

Pierrehumbert, R. T., 1998: Lateral mixing as a source of subtropical water vapor. *Geophys. Res. Lett.,* **25,** 1784–1806.

——, and R. Roca, 1998: Evidence for control of Atlantic subtropical humidity by large scale advection. *Geophys. Res. Lett.,* **25,** 4537–4540.

——, H. Brogniez, and R. Roca, 2006: On the relative humidity of the atmosphere. *The Global Circulation of the Atmosphere,* T. Schneider and A. H. Sobel, Eds., Princeton University Press.

Read, W. G., and Coauthors, 2001: UARS Microwave Limb Sounder upper tropospheric humidity measurement: Method and validation. *J. Geophys. Res.,* **106,** 32 207–32 258.

——, and Coauthors, 2007: Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H_2O and relative humidity with respect to ice validation. *J. Geophys. Res.,* **112,** D24S35, doi:10.1029/2007JD008752.

Ryoo, J.-M., D. W. Waugh, and A. Gettelman, 2008: Variability of subtropical upper tropospheric humidity. *Atmos. Chem. Phys.,* **8,** 1041–1067.

Salathe, E. P., and D. L. Hartmann, 1997: A trajectory analysis of tropical upper-tropospheric moisture and convection. *J. Climate,* **10,** 2533–2547.

Sherwood, S. C., 1996: Maintenance of the free-troposphere tropical water vapor distribution. Part II: Simulation by large-scale advection. *J. Climate,* **9,** 2919–2934.

——, E. R. Kursinski, and W. G. Read, 2006: A distribution law for free-tropospheric relative humidity. *J. Climate,* **19,** 6267–6277.

Soden, B. J., and F. P. Bretherton, 1993: Upper tropospheric relative humidity from the GOES 6.7 μm channel: Method and climatology for July 1987. *J. Geophys. Res.,* **98** (D9), 16 669–16 688.

Spencer, R. W., and W. D. Braswell, 1997: How dry is the tropical free troposphere? Implications for global warming theory. *Bull. Amer. Meteor. Soc.,* **78,** 1097–1106.

Spichtinger, P., K. Gierens, and W. Read, 2002: The statistical distribution law of relative humidity in the global tropopause region. *Meteor. Z.,* **11,** 83–88.

Waugh, D. W., 2005: Impact of potential vorticity intrusions on subtropical upper tropospheric humidity. *J. Geophys. Res.,* **110,** D11305, doi:10.1029/2004JD005664.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.

Zhang, C., B. E. Mapes, and B. J. Soden, 2003: Bimodality in tropical water vapor. *Quart. J. Roy. Meteor. Soc.,* **129,** 2847–2866.