JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

Note on the Weekly Cycle of Storm Heights over the SE U.S.

Thomas L. Bell

Laboratory for Atmospheres, NASA/Goddard Space Flight Center, Greenbelt, Maryland, USA

Jung-Moon Yoo

Department of Science Education, Ewha Womans University, Seoul, South Korea

Myong-In Lee

Goddard Earth Sciences and Technology Center, Univ. Maryland Baltimore Cty., Baltimore, Maryland, USA

Thomas L. Bell, NASA/Goddard Space Flight Center, Mail Code 613.2, Greenbelt, Maryland 20771, USA. (Thomas.L.Bell@nasa.gov)

Jung-Moon Yoo, Department of Science Education, Ewha Womans University, 11-1, Daehyun-Dong, Seodaemun-Gu, Seoul, South Korea 120-750. (yjm@ewha.ac.kr)

Myong-In Lee, NASA/Goddard Space Flight Center, Mail Code 610.1, Greenbelt, Maryland 20771, USA. (Myong-In.Lee-1@nasa.gov)

Abstract. An earlier paper by *Bell et al.* [2008] showed satellite evidence that average summertime (1998–2005) rainfall over the non-coastal southeast U.S. varied with the day of the week in a statistically significant way, with the maximum occurring midweek (Tue–Thu). An explanation was proposed in which the recurring midweek increase in air pollution over the area causes a shift in the drop-size distribution in clouds to smaller sizes as the clouds develop. The smaller droplets could be carried to higher altitudes where their freezing releases additional latent heat, invigorating the storms. Evidence for this phenomenon was provided by storm-height distributions obtained from the Tropical Rainfall Measuring Mission radar, but the statistical significance of the midweek increase in storm heights was unclear. An improved statistical analysis of the storm-height distributions is provided here, indicating that the probability that storms climb above altitudes 7–15 km is increased midweek relative to weekends (Sat–Mon) for afternoon storms (1200–2400 local time). The morning storm heights, on the other hand, are found not to exhibit statistically significant shifts, which would be consistent with the above explanation. Morning storm statistics are also found to be much more sensitive than afternoon storm statistics to the exact area over which the averages are taken.

1. Introduction

In an earlier paper by *Bell et al.* [2008] (hereinafter "B08") evidence was presented for a weekly cycle in rain-rate estimates from the Tropical Rainfall Measuring Mission (TRMM) satellite's microwave instruments over the non-coastal southeast U.S. This area was referred to in B08 as "Area B" and is shown in Figure 1 as the red cross-hatched area. Averages of TRMM rain estimates over Area B for the summertime (Jun–Aug) for 1998–2005 showed substantial changes in average rainfall with the day of the week, peaking on Tuesday and remaining large for the next two days. **Figure 1**

The explanation of the dependence of rain rate on the day of the week proposed there invoked the well-known variations in pollution with the day of the week and the theory described by Rosenfeld in the papers by *Williams et al.* [2002]; *Andreae et al.* [2004] and further developed in *Rosenfeld et al.* [2008], that the decrease in droplet sizes in storm clouds forming in "dirty" air enabled more liquid water to reach higher altitudes and to release additional latent heat as it froze, energizing the storms, causing them to grow larger and rain more. The mechanism requires that the storms form in environments such as exist in the SE U.S. during the summertime: highly unstable vertical temperature structures, with ample moisture below and cloud-base temperatures well above freezing. This theory would not apply to the drier western half of the country, and, indeed, no weekly cycle was discernible there. The theory predicts, in fact, that this effect of pollution should be maximum in the afternoons, and this was observed: the statistical significance of the weekly cycle in rainfall over Area B increased considerably when averages were restricted to afternoon (1200–2400 LT) data.

This evidence was reinforced by B08's analysis of surface rain-gauge data and the model reanalysis data [version R-2 of the National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) reanalysis data, *Kanamitsu et al.*, 2002]. As the satellite data suggested, daily rainfall as measured by rain gauges increased in the middle of the week, and lower-level wind convergence, upper-level divergence, and 500-hPa vertical winds over Area B fluctuated with the day of the week in a way that was consistent with the changes in convection implied by the rain activity. Furthermore, although not reported in B08, the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard both NASA's Terra and Aqua satellites [*Remer et al.*, 2005] also shows significant increases in fractional cloud cover during the middle of the week over Area B, accompanied by decreases in cloud-top temperatures, both signs of increased midweek convective activity.

The TRMM satellite precipitation radar (PR) provided additional evidence of storm invigoration, showing that the distribution of storm heights shifts to higher altitudes during the middle of the week compared to weekends. At the time of B08's publication a credible statistical analysis of the changes in the PR storm-height distributions with the day of the week was unavailable. This note is intended, in part, to rectify that. In the following sections we describe the PR storm-height data used in the analysis, present the method of estimating the statistical confidence of the changes we see, provide some discussion of how sensitive the changes in the morning (0000–1200 LT) distributions are to averaging details, and offer our conclusions in the final section.

2. Description of Data

The TRMM precipitation radar (PR) product 2A23 [*TRMM PR Team*, 2005] reports "storm height" for each radar observation where the PR algorithm (version 6) determines that precipitation is detected within the radar beam with a high degree of confidence. "Storm height" here means the height of the highest point in the radar beam with detectable returns ($\sim$ 17–18 dBZ), measured relative to mean sea level. The PR footprint is roughly 4–5 km in diameter. A more complete discussion of the issues involved with the interpretation of this product may be found in B08. We analyze storm-height data here for the same period used in B08, 1998–2005 summers (June–August).

It should be noted that because of the TRMM's low-inclination orbit and the PR's swath width, the PR is unable to see north of about 36.3°. The PR's observations are most frequent in the neighborhood of latitude 33.7N, and our areal statistics consequently weight the higher latitudes more, proportional to the PR's observational frequency.

3. Method of analysis

Instead of averaging over the irregularly shaped Area B in Figure 1, we used a simpler, rectangular box spanning latitudes 32.5N–40N and longitudes 100W–80W. This is the area formed by substituting for the southernmost $2.5^\circ \times 2.5^\circ$ grid box of Area B the bottom right-hand corner of the rectangle, as indicated by the arrow in Figure 1. We shall refer to this area as "Area B′".

Histograms $n(a)$ of storm heights in Area B′ are obtained by counting the number of PR footprints identified by the PR algorithm as containing rain and with storm heights in an altitude bin labeled by altitude a (in km), where the bin extends from $a - 0.5\,\mathrm{km}$ to $a + 0.5\,\mathrm{km}$. These histograms are used to calculate the fraction of footprints in Area

B′ for which storm heights are in bin a or above, for local observation times falling either in Tue–Thu or Sat–Mon. We further subdivide the observations into morning (00–12 LT) and afternoon (12–24 LT) categories. If, for instance, $n_{\mathrm{TWT,m}}(a)$ is the number of footprints in Area B′ for Tue–Thu mornings with storm heights in bin a, then we can write the fraction of footprints with storm heights at a or above as

$$c_{\mathrm{TWT,m}}(a) = \frac{\sum_{a'=a}^{t} n_{\mathrm{TWT,m}}(a')}{\sum_{a'=0}^{t} n_{\mathrm{TWT,m}}(a')} \; , \tag{1}$$

where the denominator in (1) is in effect just the total number of footprints with PR-detected rain and t is the bin with maximum reported storm height. Because no storm heights above 20.5 km are detected, $t = 20$.

As a measure of the change in height distributions with the day of the week, we investigate the ratios

$$r_i(a) = c_{\mathrm{TWT},i}(a)/c_{\mathrm{SSM},i}(a) \; , \tag{2}$$

where the index $i = \{\mathrm{m,a}\}$ indicates whether the data are for mornings or afternoons. The ratio $r_i(a)$ tells us how much more probable it is that storms reach or exceed altitude a Tue–Thu ("midweek") compared to Sat–Mon ("weekends"). If there were no change in behavior with the day of the week, we would expect $r_i(a) = 1$.

3.1. Sampling error estimates

The sampling error in the ratio is represented by δr_i:

$$r_i = \langle r_i \rangle + \delta r_i \; . \tag{3}$$

(We omit specifying the altitudes a here and in the next equation to help simplify the notation.) The angular brackets in (4) indicate the expected values of the quantities that we would calculate if we had infinite amounts of data with the same climatological statistics as the data we actually have. Our best estimate of these expectations will in fact be the averages obtained from the data we actually have. Note that, by definition, $\langle \delta r_i \rangle = 0$.

In B08, Appendix A, Eq. (A8), an estimate of the sampling error variance in these ratios is given,

$$\mathrm{var}(\delta r_i) \approx \langle c_{\mathrm{SSM},i} \rangle^{-2}\, \mathrm{var}(\delta c_{\mathrm{TWT},i}) + \langle c_{\mathrm{TWT},i} \rangle^2 \langle c_{\mathrm{SSM},i} \rangle^{-4}\, \mathrm{var}(\delta c_{\mathrm{SSM},i}) \ . \qquad (4)$$

which we have copied above, but with the dependence on whether it is morning or afternoon, $i =$ m or a, made explicit.

The variances $\mathrm{var}(\delta c_{\mathrm{TWT},i})$ and $\mathrm{var}(\delta c_{\mathrm{SSM},i})$ needed in (4) are estimates of the sampling error variances in the individual distributions $c_{\mathrm{TWT},i}$ and $c_{\mathrm{SSM},i}$. In B08 these error variances were estimated by assuming that storm heights for each PR footprint are statistically independent of each other. The estimates ignored the effects of spatial and temporal correlation in the data, which are surely substantial, with the result that the error estimates in B08 were at best lower bounds for the actual errors. We improve on these error estimates here, so that we get a better sense about which changes in the storm height distributions are "real".

To try to deal with spatial and temporal correlations in the storm heights, we work with the storm-height distributions for *each week*, represented by $n_{p,w}(a)$, where p denotes the period from which the data come: whether the data are from TWT or SSM, and whether they are for mornings or afternoons ($i =$ m or a). Weeks are labeled by integers $w = 1, \ldots, W$. Thus,

$$n_{\mathrm{TWT,m}}(a) = \sum_{w=1}^{W} n_{p,w}(a) \ , \quad p = \{\mathrm{TWT}, \mathrm{m}\} \ , \qquad (5)$$

where the sum is over all W weeks during the summers of the years of interest. There are 13 weeks per summer. Storm behavior is scarcely predictable beyond a few hours, and it is reasonable to assume that the distributions $n_{p,w}(a)$ are not very correlated from week to week. By treating the weekly distributions as a single "measurement", most of the statistical effects of spatial and temporal correlations are captured.

We assume that the sampling error variance of an average over W uncorrelated observations x_w can be approximated when W is large by the normal-statistics result

$$\mathrm{var}(\overline{x_w}) = \frac{1}{W}\mathrm{var}(x_w) \ , \tag{6}$$

where the bar notation indicates the average

$$\overline{x_w} = \frac{1}{W}\sum_{w=1}^{W} x_w \ , \tag{7}$$

and var is the estimated variance of the variable x_w,

$$\mathrm{var}(x_w) = \frac{1}{W-1}\sum_{w=1}^{W}(x_w - \overline{x_w})^2 \ . \tag{8}$$

This gives us estimates of the error variance in the overall storm-height distributions in Eq. (5) for large W:

$$\mathrm{var}[\delta n_p(a)] \approx W\mathrm{var}[n_{p,w}(a)] \ , \tag{9}$$

where $\mathrm{var}[n_{p,w}(a)]$ is estimated as in Eq. (8). [Note that the familiar factor $1/W$ is replaced by W on the right-hand side of Eq. (9) because the factor $1/W$ is absent from the definition of $n_{\mathrm{TWT,m}}(a)$ in Eq. (5).]

Building on this estimate for the error variance of $\delta n_p(a)$, we can make estimates of the error variances of the ratios r_i, since they are all derived from the distributions $n_p(a)$ through Eq. (1). This is the basis for estimating the sampling error of $r_i(a)$ in Eq. (4). Details are given in the Appendix.

4. Storm Height Distributions

PR data for storm heights for the summers of 1998–2005 over Area B′ were analyzed, with error bars for the ratios $r_i(a)$ estimated as described in the Appendix. The results are shown in panel (b) of Figure 2 with one-sigma error bars. This figure should be compared to Figure 10(c) in B08, reproduced here in Figure 2(a). As was found in B08, storms in the afternoon tend to climb to higher altitudes during the midweek period compared with weekends, particularly for those reaching altitudes above 7 to 15 km. There is a 40% higher chance that midweek storm heights will exceed 9 km than weekend storm heights. As expected, the error bars found here are generally much larger than those estimated in B08, almost certainly due to the amount of spatial and temporal correlation in the data.

Figure 2

The behavior of the “morning” data shown in Figure 2 is, however, somewhat different from what was found in B08, where the statistics for $r_{\mathrm{m}}(a)$ appeared to show that the ratios for morning storm heights were significantly above 1 at higher altitudes. Our new results are consistent with there being no change with the day of the week in morning storm-height distributions.

We have been able to discern three possible reasons for this change in the morning results from the results in B08. The primary reason appears to be changing the averaging area from Area B to Area B′. We know that the statistics of storm behavior depend on location. The diurnal cycle of rainfall in the $2.5° \times 2.5°$ grid box in the southeast corner of Area B′ peaks quite strongly in the middle of the afternoon ($\sim$ 1500–1900 LT), whereas the diurnal cycle in the southwest corner of Area B′ is ill-defined [e.g., *Hirose et al.*, 2008]. This suggests that the conditions for storm invigoration may be very different in the mornings for the two grid boxes, perhaps enough to change the overall statistics. The other two reasons for the changes in statistics were 1) that the histograms for each hour of the day and day of the week were smoothed in B08 over a 4–5-h range before computing the ratios, whereas they are not smoothed here; and local times were computed for each $2.5° \times 2.5°$ grid box in B08, whereas the local times at the center of Area B′ used here are calculated to determine whether data are assigned to mornings or afternoons. The changes in statistics due to these last two reasons were relatively minor compared with the change resulting from the shift from Area B to Area B′.

5. Discussion and Conclusions

The increase in afternoon storm heights during the middle of the week compared with weekends seen in Figure 2 is consistent with the physical picture that storm growth is enhanced by the presence of additional particulate pollution in the atmosphere during the middle of the week. The size of the increase appears to be statistically strongest at higher altitudes and it would be hard to attribute the increases to the happenstances of sampling.

The result for morning storm heights obtained here—that there is no clear change in behavior with the day of the week—is easier to understand than the behavior found in B08 for morning storms: the physical mechanism invoked to explain the afternoon changes would suggest that, since convective potential is smaller in the morning hours and there is likely to be less release of latent heat of fusion due to the freezing of water droplets, there should be less invigoration of morning storms and less of a weekly cycle. The sensitivity of the morning results to the exact area over which the statistics are obtained, possibly owing to the changes in the diurnal variations present in Area B, however, suggests that unraveling how pollution affects morning convection will be more difficult.

Appendix

The sampling error variance estimates for the ratios $r_i(a)$ defined in Eq. (2) are made using the weekly values of the storm-height distributions $n_{p,w}(a)$. We present here some details about how these estimates are made. An expression for the error variance of $r_i(a)$ is given in Eq. (4) in terms of the error variances of $\delta c_{\mathrm{TWT},i}$ and $\delta c_{\mathrm{SSM},i}$, and we show here how these two variances can be estimated. As in Eq. (5), we use p to symbolize the period from which the data came, whether TWT or SSM, and whether $i = \mathrm{m}$ or a (morning or afternoon).

We first define the cumulative sum of the storm-height histograms for each week w,

$$N_{p,w}(a) = \sum_{a'=a}^{t} n_{p,w}(a') \ , \tag{A1}$$

and the cumulative sum of the overall storm-height histograms,

$$\begin{aligned} N_p(a) &= \sum_{a'=a}^{t} n_p(a') \ , && \text{(A2)} \\ &= \sum_{w=1}^{W} N_{p,w}(a) \ , && \text{(A3)} \end{aligned}$$

which gives the number of PR-observed storm heights in Area B′ falling in bin a or higher. The fractions $c_p(a)$ defined in Eq. (1) can then be written

$$c_p(a) = \frac{N_p(a)}{N_p(0)} \ . \tag{A4}$$

As was done in Eq. (3), we describe the sampling error in terms of deviations from the climatological mean,

$$N_p(a) = \langle N_p(a) \rangle + \delta N_p(a) \ . \tag{A5}$$

The first term on the right-hand side of (A5) is the expected count, and the second term is the deviation from the expected count due to the particular data sample from which we happened to have calculated $N_p(a)$ from. We can use analogous notation to rewrite Eq. (A4) as

$$
\begin{aligned}
c_p(a) &= \langle c_p(a)\rangle + \delta c_p(a) && \text{(A6)}\\
&= \frac{\langle N_p(a)\rangle + \delta N_p(a)}{\langle N_p(0)\rangle + \delta N_p(0)} \,. && \text{(A7)}
\end{aligned}
$$

Assuming that the fluctuations $\delta N_p(0)$ are not too big relative to $\langle N_p(0)\rangle$, we expand Eq. (A7) to second order in δ and then keep only second-order terms in δ for the expression for $\langle[\delta c_p(a)]^2\rangle$ to obtain

$$
\begin{aligned}
\langle[\delta c_p(a)]^2\rangle &\approx \frac{\langle[\delta N_p(a)]^2\rangle}{\langle N_p(0)\rangle^2} - 2\frac{\langle N_p(a)\rangle}{\langle N_p(0)\rangle}\frac{\langle \delta N_p(a)\delta N_p(0)\rangle}{\langle N_p(0)\rangle^2} \\
&\quad + \frac{\langle N_p(a)\rangle^2}{\langle N_p(0)\rangle^2}\frac{\langle[\delta N_p(0)]^2\rangle}{\langle N_p(0)\rangle^2} && \text{(A8)}\\
&= \frac{1}{\langle N_p(0)\rangle^2}\Big\{\langle[\delta N_p(a)]^2\rangle \\
&\quad -2c_p(a)\langle \delta N_p(a)\delta N_p(0)\rangle \\
&\quad +c_p^2(a)\langle[\delta N_p(0)]^2\rangle\Big\} \,. && \text{(A9)}
\end{aligned}
$$

Following the approach that led to Eq. (9) above, we can estimate the first and last terms inside the curly brackets in Eq. (A9), using the fact that $N_p(a)$ can be written as a sum of weekly contributions (Eq. A3). We find

$$\langle[\delta N_p(a)]^2\rangle \approx W \mathrm{var}[N_{p,w}(a)] \,, \tag{A10}$$

where the variance on the right-hand side is calculated as in Eq. (8).

The same approach allows us to estimate the middle term inside the curly brackets of Eq. (A9). One finds

$$\langle \delta N_p(a)\delta N_p(0)\rangle \approx W \mathrm{cov}[N_{p,w}(a), N_{p,w}(0)] \tag{A11}$$

with the covariance estimated in the usual way: letting $x(w) = N_{p,w}(a)$ and $y(w) = N_{p,w}(0)$,

$$\mathrm{cov}[x(w), y(w)] \approx \frac{1}{W-1} \times \sum_w [x(w) - \overline{x}][y(w) - \overline{y}] \ . \quad \text{(A12)}$$

Given Eq. (4) and the estimates of $\mathrm{var}[\delta c_p(a)]$ provided by Eqs. (A9), (A10), and (A11) here, we can estimate the error variance for the ratios r_i. The square roots of these error variances are used as one-sigma error bars in Figure 2.

Note that because morning and afternoon samples may only be separated by a few hours, sampling errors for the morning and afternoon cases may not be statistically independent of each other. Consequently, the error bars in Figure 2 may not be appropriate for testing whether the ratios r_{m} and r_{a} are statistically different from each other. Sampling errors in $r_i(a)$ at one altitude are also probably correlated with those at other altitudes.

Acknowledgments. Research by TLB was supported by the Science Mission Directorate of the National Aeronautics and Space Administration as part of the Precipitation Measurement Mission program under Dr. Ramesh Kakar.

References

Andreae, M. O., D. Rosenfeld, P. Artaxo, A. A. Costa, G. P. Frank, K. M. Longo, and M. A. F. Silva-Dias (2004), Smoking rain clouds over the Amazon, *Science*, *303*(5662), 1337–1342.

Bell, T. L., D. Rosenfeld, K.-M. Kim, J.-M. Yoo, M.-I. Lee, and M. Hahnenberger (2008), Midweek increase in U.S. summer rain and storm heights suggests air pollution invigorates rainstorms, *J. Geophys. Res.*, *113*(D02209), doi:10.1029/2007JD008623.

Hirose, M., R. Oki, S. Shimizu, M. Kachi, and T. Higashiuwatoko (2008), Finescale diurnal rainfall statistics refined from eight years of TRMM PR data, *J. Appl. Meteor. Climat.*, *47*, 544–561.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter (2002), NCEP–DOE AMIP-II reanalysis (R-2), *Bull. Amer. Meteorol. Soc.*, *83*, 1632–1643, doi:10.1175/BAMS-83-11-1631.

Remer, L. A., et al. (2005), The MODIS aerosol algorithm, products and validation, *J. Atmos. Sci.*, *62*, 947–973.

Rosenfeld, D., U. Lohmann, G. B. Raga, C. D. O'Dowd, M. Kulmala, S. Fuzzi, A. Reissell, and M. O. Andreae (2008), Flood or drought: How do aerosols affect precipitation?, *Science*, *321*(5894), 1309–1313, doi:10.1126/science.1160606.

TRMM PR Team (2005), Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar Algorithm Instruction Manual for Version 6, NASDA/NASA, available online at `http://www.eorc.jaxa.jp/TRMM/document/pr_manual/pr_manual_v6.pdf`.

Williams, E., et al. (2002), Contrasting convective regimes over the Amazon: Implications for cloud electrification, *J. Geophys. Res.*, *107*(D20), 8082, doi:10.1029/2001JD000380.

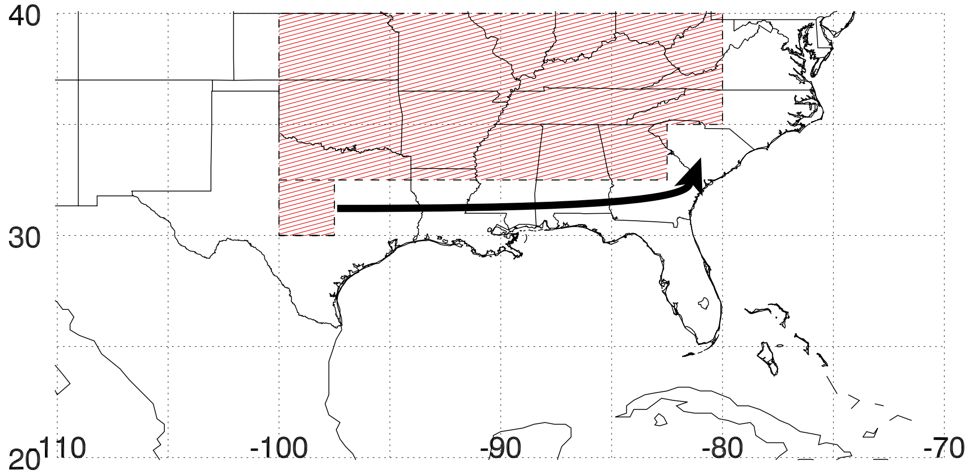

Figure 1. Averaging area (red cross-hatching) used in *Bell et al.* [2008], called "Area B" there. A new rectangular averaging area is used here, by moving the grid box at the bottom to the right-hand side, as indicated by the arrow. This rectangular area is referred to here as "Area B′".

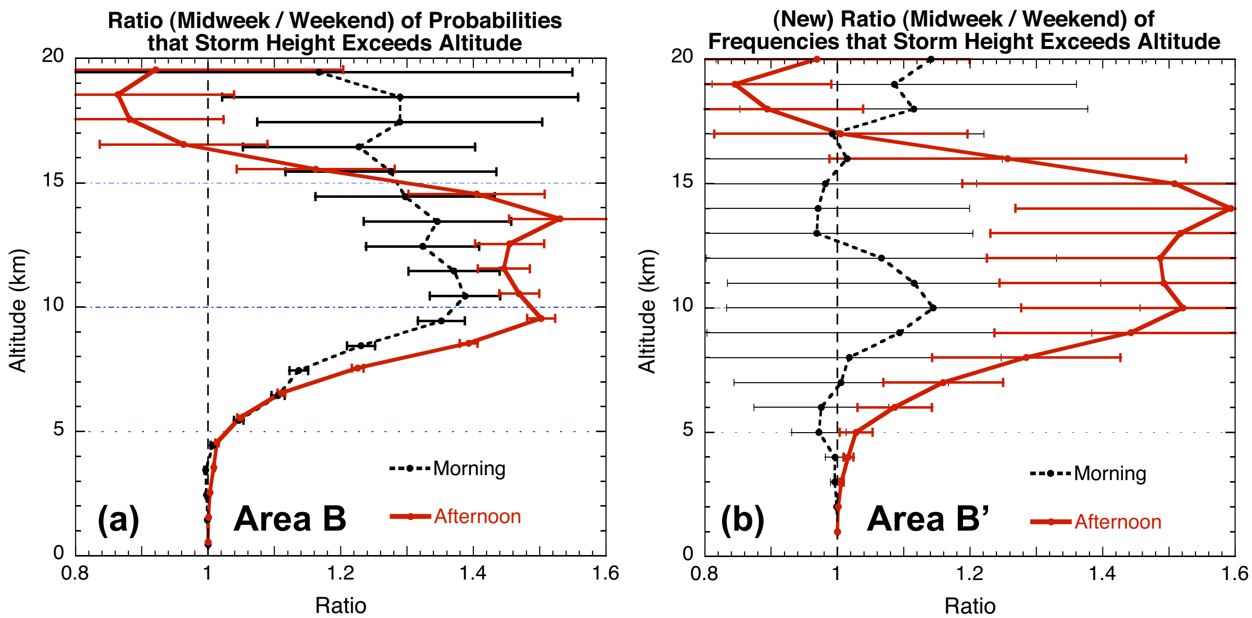

Figure 2. The ratio of the frequency that storm heights exceed a given altitude during the midweek (Tue–Thu) to the frequency that weekend (Sat–Mon) storm heights exceed that altitude. Dashed black lines are for morning (00–12 LT) data; red lines for afternoon (12–24 LT) data. (a) Ratio over Area B, figure reproduced from B08. (b) New results for Area B′ with one-sigma error bars estimated here. A ratio of 1 (dashed line) would suggest that there is no variation in storm-height distributions with the day of the week.