A QuikSCAT climatology of tropical cyclone size

D. R. Chavas[1] and K. A. Emanuel[1]

Received 1 July 2010; revised 30 July 2010; accepted 6 August 2010; published 29 September 2010.

[1] QuikSCAT data of near-surface wind vectors for the years 1999–2008 are used to create a climatology of tropical cyclone (TC) size, defined as the radius of vanishing winds. The azimuthally-averaged radius of 12 ms^{-1} wind (r_{12}) is calculated for a subset of TCs (N = 2154) whose centers of circulation were clearly identifiable via subjective analysis of the QuikSCAT-analyzed wind field. The outer radius, r_0, is determined from r_{12} using an outer wind structure model that assumes no deep convection beyond r_{12}. The global median values of r_{12} and r_0 are 197 km and 423 km, respectively, with statistically significant variation across ocean basins. The global distribution of r_{12} is found to be approximately log-normal, the distribution of r_0 is quantitatively much closer to log-normal, and the improvement in fit between r_{12} and r_0 is attributed to the combined effect of the nature of the model employed and the paired distributions of r_{12} and f. Moreover, the normalization employed by Dean et al. (2009) is found to weaken rather than improve the log-normal fit. Finally, within a given storm, both r_{12} and r_0 tend to expand very slowly with time early in the storm lifecycle and then becomes quasi-constant, though significant variance exists across storms. **Citation:** Chavas, D. R., and K. A. Emanuel (2010), A QuikSCAT climatology of tropical cyclone size, *Geophys. Res. Lett.*, *37*, L18816, doi:10.1029/2010GL044558.

1. Introduction

[2] In the absence of land interaction, the horizontal extent of the outer circulation of a tropical cyclone (TC) is observed in nature to vary only marginally during the lifetime of a given TC prior to recurvature into the extra-tropics [*Merrill*, 1984; *Frank*, 1977], but significant variation exists from storm to storm, regardless of basin, location, intensity, and time of year. *Kimball and Mulekar* [2004] determined from Atlantic Extended Best Track data that as a storm intensifies the radius of outermost closed isobar (ROCI) remains approximately constant despite changes in the radial structure of the intermediate wind field. More recently, modeling work by *Hill and Lackmann* [2009] and *Wang* [2009] showed that TCs tend to be larger when embedded in moister mid-tropospheric environments due to the increase in spiral band activity and subsequent generation of diabatic potential vorticity which acts to expand the wind field laterally.

[3] From a broader perspective, *Merrill* [1984] found frequency distributions of ROCI in the Atlantic and Western North Pacific that qualitatively resemble log-normal distributions. *Dean et al.* [2009, hereafter D09] found that the distribution of normalized storm size, defined as the radius of vanishing winds divided by the ratio of the potential intensity to the Coriolis parameter, is close to log-normal in the Atlantic basin. However, the result of D09 is based on the radius of gale force winds (R34) taken from two datasets that employ very different methodologies and whose R34 values for overlapping cases disagree markedly.

[4] Ideally, one would prefer to characterize the size distribution based upon direct surface wind measurements taken from a single, consistent source. Thus, this work examines the global distribution of TC size, defined here as the radius of vanishing winds, using an independent, high-resolution dataset generated by the QuikSCAT satellite microwave scatterometer. The following sections outline the data and methodology used to generate a climatology of TC size, discuss its characteristic values and distribution, and explore the intra-storm evolution of size.

2. Data

[5] Ocean near-surface (10 m) wind vector data are taken from the QuikSCAT Level 2B dataset on a 12.5 km × 12.5 km grid for the period beginning July 19, 1999 (the start of the satellite's operational life) through December 31, 2008; this dataset is available at http://podaac.jpl.nasa.gov/DATA_CATALOG/quikscatinfo.html. Owing to rain contamination of the signal, QuikSCAT data quality is highest away from strong precipitation, and the instrument is considered very accurate in the range 3–20 ms^{-1} (Jet Propulsion Laboratory (NASA), Sea Winds on QuikSCAT, accessed 9 April 2010, http://winds.jpl.nasa.gov/missions/quikscat/index.cfm); *Chou et al.* [2010] found RMS differences between QuikSCAT wind speeds and dropwindsonde data of 2.6 ms^{-1}. For a complete discussion of potential errors, see *Hoffman and Leidner* [2005].

[6] Tropical cyclone 6-hourly location and intensity data are taken from the National Hurricane Center HURDAT Best Track database (National Hurricane Center HURDAT Hurricane Best Track files, accessed 9 April 2010, http://www.nhc.noaa.gov/pastall.shtml/#hurdat). For calculation of the normalization factor, $\frac{PI}{f}$, which is the natural tropical cyclone length scale [*Emanuel*, 1986], potential intensity values are taken from monthly mean re-analysis data [*Bister and Emanuel*, 2002] bi-linearly interpolated to the storm location.

3. Methodology

3.1. Locating TCs

[7] To create a climatology of tropical cyclones as seen by QuikSCAT, Best Track location and intensity data are spline interpolated iteratively forward until reaching the minimum distance, d, to any valid (i.e., non-rain-flagged) QuikSCAT

[1]Department of Earth, Atmosphere, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

0094-8276/10/2010GL044558

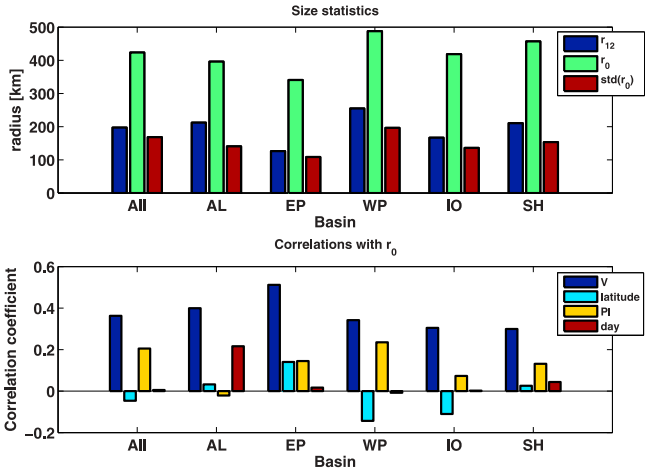

Figure 1. (top) Median values of r_{12} (blue) and r_0 (green) and the standard deviation of r_0 (red) globally and by basin. All units in [km]. (bottom) Correlation coefficients between r_0 and various parameters globally and across basins; "day" represents day of the hurricane season. Basins listed are Atlantic (AL), East Pacific (EP), West Pacific (WP), Indian Ocean (IO), and Southern Hemisphere (SH).

datapoint of a given pass. Cases for which $d > 100$ km or the interpolated intensity $V_{BT} \leq 17.451$ ms^{-1} are skipped.

[8] Next, to identify the TC center of circulation we take as a first guess the interpolated Best Track location, about which we extract all data (including rain-flagged) within a 4° × 4° box. All TC centers are then subjectively identified based on the full QuikSCAT wind vector field in this box. Only those cases for which there exists a single, clearly-defined center of cyclonic circulation are included, based upon the criteria that a) the center is consistent with the wind vectors in the immediate vicinity in all directions, and b) the broad "outer" circulation (i.e., 1–4 degrees from center) is easily discernible and is consistent with the location identified by criterion (a). The authors sought to be conservative in this procedure; when ambiguous, the case was omitted. All data within 2500 km of the center are then used for subsequent analysis.

[9] Only cases over water and for which the potential intensity $PI > 40$ ms^{-1} are included in order to avoid cases in which storms are rapidly transitioning to regions of cold sea surface temperatures where mature tropical cyclones cannot be sustained. The TC translation vector, calculated directly from the full spline interpolation of the Best Track dataset, is then subtracted from all wind vectors. All vectors are projected onto their pure-azimuthal component relative to the TC center and vector magnitudes are signed: positive for cyclonic, negative for anti-cyclonic. Lastly, wind speeds are azimuthally-averaged within 10-km wide rings moving radially outward from center to obtain a radial wind profile for each TC fix.

[10] Finally, we select a single azimuthal-average wind speed, V_{QS}, and for each TC fix determine its radius, r_{QS}, and extrapolate outward to r_0 using a theoretical model of outer wind structure that assumes minimal deep convection in the outer region. This model is described in detail by *Emanuel* [2004] and is reviewed below.

3.2. Selecting V_{QS}

[11] Selection of an optimal QuikSCAT wind speed, V_{QS}, necessitates balancing three key constraints. First, the assumption of constant background flow, represented by the single translation vector subtracted from all points, loses validity far from center; this constraint renders any effort to extract r_0 directly from the QuikSCAT data invalid. Second, the number of data points increases dramatically as one moves outward from the TC center. Finally, *Brennan et al.* [2009] found that QuikSCAT observed winds have a near-zero bias due to rain in the range of 10–15 ms^{-1}. The validity of a given azimuthal-average wind speed depends on the trade-offs between the above three factors. Based on these criteria we set $V_{QS} = 12$ ms^{-1}.

[12] The final result is a dataset of 2154 TC fixes spread across five basins: Atlantic (482), East Pacific (367), West Pacific (640), Indian Ocean (78), and Southern Hemisphere (587).

3.3. Estimating Outer Radius r_0

[13] To estimate the outer radius, r_0, we employ the outer wind structure model derived by *Emanuel* [2004] (for an abridged form, see D09) to extrapolate radially outwards from the QuikSCAT-defined azimuthal-average radius, r_{QS}, of the wind speed V_{QS} described above. Here, we briefly review the model's characteristics. The flow is assumed to be steady and axisymmetric. The model assumes that there is no deep convection beyond r_{QS}, resulting in a local balance between subsidence warming and radiative cooling. Furthermore, given that both the lapse rate and the rate of clear-sky radiative cooling are nearly constant in the tropics, the equilibrium subsidence velocity, w_{rad}, can be taken to be approximately constant. In equilibrium, this subsidence rate must match the rate of Ekman suction-induced entrainment of free tropospheric air into the boundary layer in order to prevent the creation of large vertical temperature gradients across the top of the boundary layer. The radial profile of azimuthal velocity is therefore determined as that which provides the required Ekman suction, and is given by

$$\frac{\partial(rV)}{\partial r} = \frac{2r^2 C_D V^2}{w_{rad}\left(r_0^2 - r^2\right)} - fr \qquad (1)$$

where r is the radius, V is the azimuthal wind speed, f is the Coriolis parameter, C_D is the bulk aerodynamic drag coefficient. We set $C_D = 10^{-3}$ and $w_{rad} = 1.6$ cms^{-1}.

[14] To our knowledge, this nonlinear first order differential equation has no analytical solution. D09 neglected the partial derivative term to derive a simple analytical solution for r_0. However, (1) can also be solved numerically for r_0, and the solution to the full equation is 30–150 km larger than the approximated solution over the typical range of tropical latitudes and r_{QS} values (not shown). Thus, for our purposes we elect to use the full numerical solution.

4. Results

4.1. Basic Statistics

[15] Figure 1 (top) displays the median r_{12} and r_0 values and the standard deviation of r_0 both globally and by basin.

[16] The global median outer radius is 423 km and ranges from a minimum of 341 km in the East Pacific to a maxi-

Table 1. Kolmogorov-Smirnoff p-Values for Statistical Fits to Various Parent Distributions for r_{12}, r_0, r_{12}^*, and r_0^*[a]

Probability Distribution	r_{12}	r_0	r_{12}^*	r_0^*
Log-normal	.028	**.626**	.248	.226
Normal	0	0	0	0
Weibull	.001	0	0	0
Rayleigh	0	0	0	0
Gamma	.05	.11	0	0

[a]Log-normal refers to the normal fit of log(r). Largest p-value is bold.

mum of 488 km in the West Pacific. The standard deviation of r_0 is 168 km and scales across basins in a similar fashion to the median value. The median distance between r_{12} and r_0 is 226 km. These values compare reasonably well with those of previous studies [e.g., *Merrill*, 1984]. Moreover, r_0 is relatively insensitive to variations in w_{rad} and C_D (assumed constant), with changes of approximately 25 km for the rather extreme cases of a halving or doubling of the ratio $\frac{C_D}{w_{rad}}$ for $\phi = 20°$ and $r_{12} = 200$ km.

[17] Figure 1 (bottom) displays correlation coefficients between r_0 and various parameters of interest. The lone correlation of note exists between r_0 and intensity V ($r = .36$) and is relatively consistent across basins; this matches the weak correlation ($r = 0.28$) found by *Merrill* [1984]. Meanwhile, r_0 is effectively independent of latitude, which contradicts the typical finding that TCs tend to expand as they recurve poleward [e.g., *Merrill*, 1984].

4.2. Size Distribution

[18] Table 1 lists the p-values for the statistical fit to various distributions of $\log(r_{12})$, $\log(r_0)$, $\log(r_{12}^*)$, and $\log(r_0^*)$, where the asterisk denotes normalization by $\frac{PI}{f}$ following D09. All p-values are calculated using the Kolmogorov-Smirnoff test statistic. In the case of the normal and log-normal test distributions, the observed data were rescaled to have zero mean and unit variance for comparison to the standard normal parent distribution N (0,1). P-values approaching unity indicate that the observed distribution is close to the parent distribution.

[19] The goodness of fit between the distribution of r_0 and a log-normal parent distribution is the most significant from among the variables and distributions tested here. The null hypothesis that r_0 is gamma distributed ($p = .11$) also cannot be rejected at the 95% confidence level, though based on a χ^2 metric (p =.043) it can be rejected.

[20] For a direct comparison of r_{12} and r_0, their global frequency distributions, along with the Gaussian fit to the mean and variance of the logarithm of each dataset, are displayed in Figure 2. Globally, $p = .028$ for r_{12}, which indicates that the null hypothesis of a log-normal distribution can be rejected at the 95% confidence interval. On the other hand, $p = .626$ for r_0, which indicates that the distribution is reasonably close to log-normal. Moreover, D09 determined that normalization of r_0 by $\frac{PI}{f}$ results in a distribution that is much closer to log-normal. Our results indicate that the distribution of r_0 is significantly closer to log-normal than that of r_{12}, but that the subsequent normalization of r_0 in fact makes the log-normal fit worse. Though normalization does improve the fit for r_{12}, this may be understood in a crude mathematical sense given that $\log(\frac{r}{\frac{PI}{f}}) = \log(r) + \log(\frac{f}{PI})$. The distribution of $\frac{f}{PI}$ itself has a p-value of $p = .165$, which is greater than that of r_{12} but less than r_0, and thus normalization would be expected to improve the fit for r_{12} but to reduce it for r_0. In either case, the important result here is that normalization is not necessary to observe a size distribution that is relatively close to log-normal. The findings are qualitatively similar within individual basins, though it is difficult to draw firm conclusions due to the smaller sample sizes. These results are found to be largely insensitive to the choice of V_{QS} over the range 8–15 ms^{-1} (not shown).

4.3. Control Experiments

[21] To what extent is this log-normal distribution an artifact of the outer wind structure model employed here? Given that our version of r_0 is only a function of r_{12} and f, we perform three test experiments. First, we recalculate r_0 using the observed distribution of f but set all values of r_{12} to be constant and equal to the median value, $r_{12} = 197.15$ km, which results in a p-value of p =.002. Second, we recalculate r_0 using the observed distribution of r_{12} but set all values of f to be constant and equal to the median value, $f = 5 * 10^{-5}$ s^{-1}, which results in a p-value of p =.222.

[22] Finally, we recalculate r_0 using the observed distribution of both r_{12} and f but randomly reshuffle their pairings, the purpose of which is to address the question of whether nature "matches" r_{12} and f in some optimal way as to generate a log-normal distribution. For 100 runs, the p-value for the observed pairings of r_{12} and f is larger than approximately 80% of cases with randomized pairings, which suggests that, though not optimized, how r_{12} and f are paired in nature may play a role in bringing the distribution of r_0 closer to log-normal.

[23] Taken together, these experiments indicate that, though a component of the observed distribution is simply due to the nature of the outer structure model chosen in this work, the actual distributions of r_{12} and f are also central to generating the log-normal distribution.

4.4. Intra-storm Evolution

[24] For the 241 distinct TCs with 4 or more QuikSCAT observations in the dataset used here, the mean intra-storm rate of change of r_{12} and r_0, taken as the slope of the linear

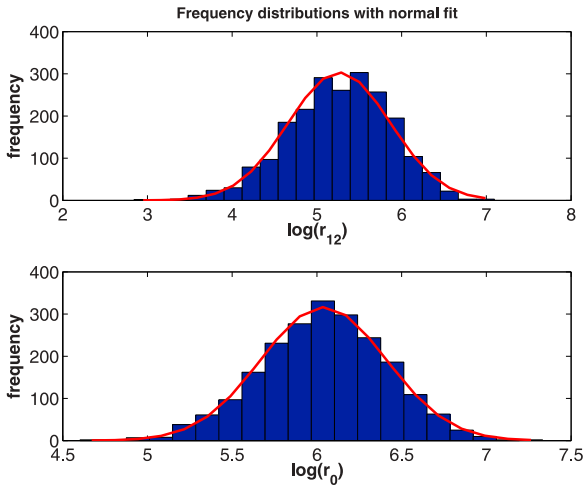

Figure 2. Global frequency distribution with Gaussian fit (red line). (top) $\log(r_{12})$; (bottom) $\log(r_0)$.

least-squares fit to the data, is 18.1 and 10.9 km day^{-1}, or approximately 9 and 2.5 % day^{-1} of the median value, respectively. The respective standard deviations are 43.1 and 53.2 km day^{-1}, indicating significant variance across individual storms; the distribution of rates of change is approximately Gaussian about the mean. Though relatively small, these mean expansion rates are statistically significantly different from zero at the 95% confidence interval (p = 0 and .002, respectively). A slow broadening of the wind field with time has also been noted in previous studies [e.g., *Cocks and Gray*, 2002; *Merrill*, 1984].

[25] Closer inspection reveals that much of this expansion appears to occur early in the storm's evolution. For the 215 distinct TCs whose first 4 observations occur within a 100 hour period, the expansion rate of r_{12} and r_0 over these first 100 hours is 24.0 and 18.7 km day^{-1}, respectively. Meanwhile, for the 35 distinct TCs with 4 or more observations at least 100 hours after the initial observation, the expansion rate beyond 100 hours declines substantially to 8.3 and −0.8 km day^{-1}, respectively, neither of which are statistically significantly different from zero (p = .28 and .92). Significant variance exists, though, as standard deviations are 43.1 and 53.2 km day^{-1}, respectively. If the outer radius of a mature TC truly remains approximately constant with time, then this result may be an indication that our threshold minimum intensity of 17.5 ms^{-1} is capturing TCs at the tail end of the genesis process during which the outer radius has yet to reach its quasi-steady state, but further investigation is needed to validate such a claim.

5. Discussion and Conclusion

[26] Given the high resolution and high precision of QuikSCAT data, the results presented here provide credible evidence that the global distribution of tropical cyclone size, defined as the radius of vanishing winds calculated using an outer wind structure model that assumes vanishing deep convection beyond the azimuthally-averaged radius of 12 ms^{-1} winds, is approximately log-normal. While the distribution of r_{12} is qualitatively log-normal, the distribution of r_0 is quantitatively much closer to log-normal. Moreover, in contrast to the work of D09, we find here that the normalization by the natural length scale of tropical cyclones, defined as the ratio of the potential intensity to the Coriolis parameter, reduces rather than improves the goodness of fit of the observed distribution to log-normal. Control experiments indicate that the choice of model alone is insufficient to explain the observed p-values for the distribution of outer radius; the distributions observed in nature of r_{12} and f, from which the distribution of r_0 is derived, appear to play an important role as well. Finally, analysis of the intra-storm evolution of size indicates that both r_{12} and r_0 tend to expand very slowly with time early in the storm lifecycle, after which size appears to remain nearly constant, although significant variance exists across storms.

[27] What is the implication of the log-normal distribution in the context of tropical cyclones? As noted earlier, in the absence of significant external environmental forcing, there is evidence that the spatial extent of a given tropical cyclone remains relatively constant throughout its lifetime, suggesting that the existence of this distribution may be a manifestation of the processes that generate tropical cyclones in the first place and/or of the distribution of their precursor disturbances. However, with respect to size, there is no obvious single multiplicative process during genesis that is amenable to isolation. This will be the subject of future work.

[28] **Acknowledgments.** This material is based in part upon work supported by the National Science Foundation under grant ATM-0850639. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Bister, M., and K. A. Emanuel (2002), Low frequency variability of tropical cyclone potential intensity: 1. Interannual to interdecadal variability, *J. Geophys. Res.*, *107*(D24), 4801, doi:10.1029/2001JD000776.

Brennan, M. J., C. C. Hennon, and R. D. Knabb (2009), The operational use of QuikSCAT ocean surface vector winds at the National Hurricane Center, *Weather Forecast.*, *24*, 621–645, doi:10.1175/2008WAF2222188.1.

Chou, K.-H., C.-C. Wu, P.-H. Lin, and S. Majumdar (2010), Validation of QuikSCAT wind vectors by dropwindsonde data from Dropwindsonde Observations for Typhoon Surveillance Near the Taiwan Region (DOTSTAR), *J. Geophys. Res.*, *115*, D02109, doi:10.1029/2009JD012131.

Cocks, S. B., and W. M. Gray (2002), Variability of the outer wind profiles of western North Pacific typhoons: Classifications and techniques for analysis and forecasting, *Mon. Weather Rev.*, *130*, 1989–2005.

Dean, L., K. A. Emanuel, and D. R. Chavas (2009), On the size distribution of Atlantic tropical cyclones, *Geophys. Res. Lett.*, *36*, L14803, doi:10.1029/2009GL039051.

Emanuel, K. A. (1986), An air-sea interaction theory for tropical cyclones. Part I: Steady state maintenance, *J. Atmos. Sci.*, *42*, 1062–1071.

Emanuel, K. A. (2004), Tropical cyclone energetics and structure, in *Atmospheric Turbulence and Mesoscale Meteorology*, edited by E. E. Fedorovich et al., p. 240, Cambridge Univ. Press, Cambridge, U. K.

Frank, W. M. (1977), Structure and energetics of the tropical cyclone, part I: Storm structure, *Mon. Weather Rev.*, *105*, 1119–1135.

Hill, K., and G. M. Lackmann (2009), Influence of environmental humidity on tropical cyclone size, *Mon. Weather Rev.*, *137*, 3294–3315.

Hoffman, R. N., and S. M. Leidner (2005), An introduction to the near real time QuikSCAT data, *Weather Forecast.*, *20*, 476–493.

Kimball, S. K., and M. S. Mulekar (2004), A 15-year climatology of North Atlantic tropical cyclones. Part I: Size parameters, *J. Clim.*, *17*, 3555–3575.

Merrill, R. T. (1984), A comparison of large and small tropical cyclones, *Mon. Weather Rev.*, *112*, 1408–1418.

Wang, Y. (2009), How do outer spiral rainbands affect tropical cyclone structure and intensity?, *J. Atmos. Sci.*, *66*, 1250–1273.

D. R. Chavas, Department of Earth, Atmosphere, and Planetary Sciences, Massachusetts Institute of Technology, Bldg. 54, Room 1715, Cambridge, MA 02139, USA. (drchavas@gmail.com)

K. A. Emanuel, Department of Earth, Atmosphere, and Planetary Sciences, Massachusetts Institute of Technology, Bldg. 54, Room 1620, Cambridge, MA 02139, USA. (emanuel@mit.edu)