9B.4

A QUIKSCAT CLIMATOLOGY OF TROPICAL CYCLONE SIZE

Daniel R. Chavas* and Kerry A. Emanuel
Program in Atmospheres, Oceans, and Climate, Massachusetts Institute of Technology, Cambridge, MA

1. INTRODUCTION

Despite significant advances in our understanding of the fundamental dynamics and thermodynamics of mature tropical cyclones (TCs), one variable that is not yet predictable, and which has attracted minimal serious theoretical or modeling effort, is storm size. In the absence of land interaction, the horizontal extent of the outer circulation 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) 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.

From a broader perspective, Merrill (1984) found frequency distributions of storm size in the Atlantic and Western North Pacific that qualitatively resemble log-normal distributions, though no formal statistical test was performed. Dean et al. (2009) found that the distribution of 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, Dean et al. (2009) is based on the radius of gale force winds (R34) taken from two datasets that employ very different methodologies and whose R34 values disagree markedly.

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. Relative to previous efforts to analyze TC size, QuikSCAT provides much higher resolution and precision for calculating surface wind vectors and thus is considered significantly more reliable than past datasets. Moreover, QuikSCAT's wider swath provides greater areal coverage and thus contains a significantly larger sample of storms than the lone previous scatterometer-based assessment of TC size (Liu and Chan, 1999). The following sections outline the data and methodology used to generate a climatology of TC size, present the result that the global distribution of outer radius (absent normalization) is closely log-normal, and discuss potential implications of this finding.

2. DATA

Ocean near-surface (10m) wind vector data are taken from the QuikSCAT L2B dataset on a 12.5 km x 12.5 km grid for the period beginning July 19, 1999 (the start of the satellite's operational life) through December 31, 2008. 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 $3-20\ ms^{-1}$ (JPL). Brennan et al. (2009) noted that for wind speeds below $10\ ms^{-1}$, rain resulted in a positive wind speed bias due to anomalous volume backscatter and increased ocean surface roughness; for wind speeds above $15\ ms^{-1}$, rain resulted in a negative wind speed bias due to attenuation of the ocean surface backscatter. For a complete discussion of potential errors, see Hoffman and Leidner (2005).

Tropical cyclone 6-hourly location and intensity data are taken from the National Hurricane Center HURDAT Best Track database (NHC). For calculation of the normalization factor, $\frac{PI}{f}$, potential intensity values are taken from monthly mean re-analysis data (Bister and Emanuel, 2002) bi-linearly interpolated to the place and time of the storm observation.

**Corresponding author address:* Daniel R. Chavas, Massachusetts Institute of Technology 54-1715, Cambridge, MA 02139; email: drchavas@gmail.com

3. METHODOLOGY

3.1 *Locating TCs*

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

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^o x 4^o box. Due to the significant and spatially-asymmetric effects of rain contamination on vector magnitude retrieval, we then calculate at every point (i, j) a modified vorticity variable, ζ^*, based solely on the flow curvature and given by

$$\zeta^*_{i,j} = \sum_{m=1}^{3} (\hat{u}_{i,j+m} - \hat{u}_{i,j-m} + \hat{v}_{i+m,j} - \hat{v}_{i-m,j}) \quad (1)$$

where u and v are the zonal and meridional wind vector components, respectively, and the overhat symbol denotes unit magnitude. Finally, using the location of maxmimum ζ^* as guidance, all TC centers are subjectively identified. 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.

As an additional filter, 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.

In order to separate the flow induced by the tropical cyclone from the background flow, the TC translation vector, calculated directly from the full spline interpolation of the Best Track dataset, is subtracted from all wind vectors. All vectors are then projected onto their pure-azimuthal component relative to the TC center and vector magnitudes are signed: positive for cyclonic, negative for anti-cyclonic. Finally, 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.

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 in Emanuel (2004) and is reviewed below.

3.2 *Selecting* V_{QS}

Selection of an optimal QuikSCAT wind speed, V_{QS}, necessitates evaluations of three key limits on QuikSCAT data quality. First, the assumption of constant background flow, represented by the single translation vector subtracted from all points, becomes increasingly erroneous as one moves away from the center. This constraint renders any effort to extract r_0 directly from the QuikSCAT data invalid. On the other hand, due to simple geometry and a decrease in data coverage caused by rain-contamination, which occurs predominantly near the TC center, the number of data points (and thus the confidence in average values) increases dramatically as one moves outward from center. Finally, as noted by Brennan et al. (2009), QuikSCAT observed winds have a near-zero bias 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 $V_{QS} = 12\ ms^{-1}$, although the overall results presented here are largely insensitive to this selection over the integer range $V_{QS} = [8, 15]$ (not shown).

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

To estimate the outer radius, r_0, we employ the outer wind structure model derived in Emanuel (2004) (for an abridged form, see Dean et al. (2009)) 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

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

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 $w_{rad} = 1.6\ cms^{-1}$.

To our knowledge, this nonlinear first order differential equation has no analytical solution. Dean et al. (2009) argued that the partial derivative term is small except as r approaches r_0, resulting in a simple analytical solution for r_0. However, (2) 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.

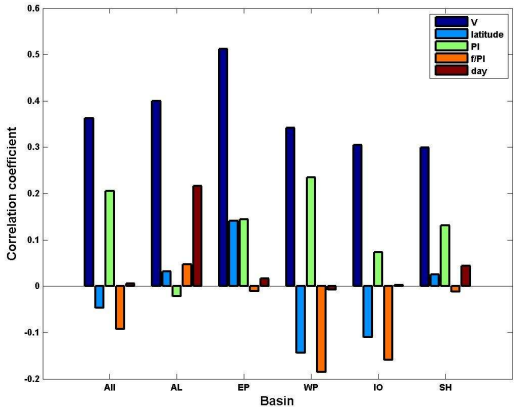

Figure 1: Correlation coefficients between r_0 and various parameters globally and across basins. Basins listed are Atlantic (AL), East Pacific (EP), West Pacific (WP), Indian Ocean (IO), and Southern Hemisphere (SH). For the Southern Hemisphere, "day" is phase shifted by 183 days to coincide with the Northern Hemisphere.

4. RESULTS

4.1 *Correlations and distributions*

Figure 1 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)).

Table 1 lists the p-values for the statistical fit to various distributions of $log(r_{12})$, $log(r_0)$, as well as the logarithm of each radius normalized by the natural tropical cyclone length scale, defined as the ratio of the potential intensity to the Coriolis parameter, denoted $log(r^*_{12})$ and $log(r^*_0)$, respectively. 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 1 indicate that the observed distribution is close to the parent normal distribution.

Table 1: Kolmogorov-Smirnoff p-values for statistical fits to various parent distributions for r_{12}, r_0, r^*_{12}, and r^*_0. Log-normal refers to the normal fit of log(r). Largest p-value is bold.

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

It is clear that the best fit for the distribution of r_0 is the log-normal distribution ($p = .626$) from among those tested here, although the null hypothesis that r_0 is gamma distributed ($p = .11$) also cannot be rejected at the 95% confidence level. Using a χ^2 metric, the p-values for the log-normal and gamma distributions are .494 and .043, respectively, which indicates that the gamma distribution can be rejected at the 95% confidence interval. In either case, although one cannot prove with certainty that an observed distribution has a specified parent distribution, p-values indicate that 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 global frequency distribution of both r_{12} and r_0, along with the Gaussian fit to the mean and variance of the datasets, are displayed in Figure 2. To test the robustness of the result that r_0 is more closely log-normal than r_12, we randomly resample 50% (N=1077) of the global dataset 1000 times. In 98.7% of resamples, $p(r_0) > p(r_{12})$ and the mean $\overline{\Delta p} = .524$, which indicates that the finding

that r_0 is more closely log-normal than r_{12} is robust.

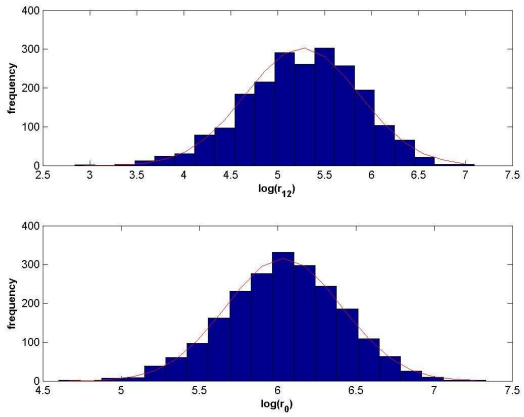

Figure 2: Global frequency distribution with Gaussian fit (red line). Top: $log(r_{12})$; Bottom: $log(r_0)$

Among individual basins (not shown), p-values are greater for r_0 than for r_{12} in all cases, but the largest increase in p-value between r_{12} and r_0 occurs for the full global dataset.

Dean et al. (2009) found that normalizing 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.

4.2 *Control experiments*

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, $r_{12} = 197.15\ km$ (Figure 3a). Although still far from log-normal ($p = .002$), it is interesting that our model, applied using the observed distribution of TC latitudes, manages to bring a purely constant "distribution" a significant ways towards log-normality.

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, $f = 5 * 10^{-5}\ s^{-1}$ (Figure 3b). Interestingly, even with constant f, the distribution is substantially closer to log-normal ($\Delta p = .194$), although the p-value remains only approximately $\frac{1}{3}$ that of the full global dataset.

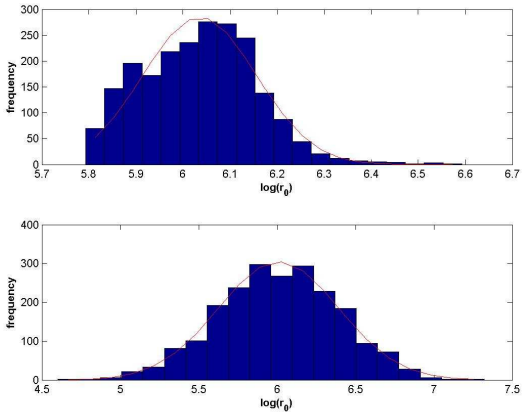

Figure 3: Distribution of $log(r_0)$ for control cases (a) Top: constant $r_{12} = 197.15\ km$ (p=.002), and (b) Bottom: constant $f = 5 * 10^{-5}\ s^{-1}$ (p=.222).

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. The p-value ($N = 100$) 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.

5. DISCUSSION AND CONCLUSION

QuikSCAT measures ocean surface wind vectors directly and at significantly higher resolution and higher precision than other available datasets of tropical cyclone wind structure. Thus, the results presented here provide reliable 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} appears to be broadly log-normal in nature, the distribution of r_0 is quantitatively much closer to log-normal. Moreover, in contrast to the work of Dean et al. (2009), 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.

The control experiments suggest that a component of this result is simply an intrinsic characteristic

of the outer structure model chosen in this work. This is not necessarily an indication that the log-normal distribution is partially an artificial construct; rather, to the extent that the model chosen here represents the outer structure of actual tropical cyclones in nature, it may represent an important piece of the dynamical puzzle that serves to generate a log-normal distribution of r_0. Nonetheless, the choice of model alone is insufficient to explain the observed values of Δp; 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.

What is the implication of the log-normal distribution in the context of tropical cyclones? Log-normal distributions are commonly associated with multiplicative processes and are ubiquitous in science, including the cluster aggregation of particles (Briehl and Urbassek, 1999), molecules and crystals (Espiau de Lamaestre and Bernas, 2006), total rainfall, species abundance, income etc. (see e.g. Limpert et al., 2001; Mitzenmacher, 2004; Koch, 1966). As noted earlier, in the absence of significant external environmental forcing (e.g. interaction with topography, such as Hurricane Ike (Atlantic 2008)), the spatial extent of a given tropical cyclone remains relatively constant throughout its lifetime, suggesting that the existence of this distribution may be fundamental to the processes that generate tropical cyclones in the first place. Thus, this result begs the question of whether tropical cyclogenesis could potentially be represented as a cluster aggregation process in a manner similar to that demonstrated in previous modeling studies of radiative-convective equilibrium (Bretherton et al., 2005; Held et al., 1993). Further investigation is needed to examine how the dynamics of such a process would function for an individual tropical cyclone, and whether these dynamics differ from those acting to generate a log-normal distribution on a global scale.

REFERENCES

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

[2] 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 Forecasting*, 24, 621Ű645. doi:10.1175/2008WAF2222188.1.

[3] Bretherton, C. S., P. N. Blossey, and M. Khairoutdinov, 2005: An energy-balance analysis of deep convective self-aggregation above uniform SST. *J. Atmos. Sci.*, 62, 4273-4292.

[4] Briehl, B., and H. M. Urbassek, 1999: Monte Carlo simulation of growth and decay processes in a cluster aggregation source. *J. Vac. Sci. Technol. A*, 17, 256.

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

[6] Emanuel, K. A., 2004: Tropical cyclone energetics and structure, In *Atmospheric Turbulence and Mesoscale Meteorology*. R. R. a. B. S. E. Federovich. New York, Cambridge University Press, 240.

[7] Espiau de Lamaestre, R., and H. Bernas, 2006: Significance of lognormal nanocrystal size distributions. *Phys. Rev. B*, 73, 125317. doi:10.1103/PhysRevB.73.125317.

[8] Frank, W. M., 1977: Structure and energetics of the tropical cyclone, Part I: Storm structure. *Mon. Wea. Rev.*, 105, 1119-1135.

[9] Held, I. M., R. S. Hemler, and V. Ramaswamy, 1993: Radiative-convective equilibrium with explicit two-dimensional convection. *J. Atmos. Sci.*, 50, 3909Ű3927.

[10] Hill, K., G. M. Lackmann, 2009: Influence of environmental humidity on tropical cyclone size. *Mon. Wea. Rev.*, 137, 3294Ű3315.

[11] Hoffman, R. N., and S. M. Leidner, 2005: An Introduction to the Near Real Time QuikSCAT Data. *Weather and Forecasting*, 20, 4, 476.

[12] Jet Propulsion Laboratory (NASA) Sea Winds on QuikSCAT. http://winds.jpl.nasa.gov/missions/quikscat/index.cfm. Accessed 9 April 2010.

[13] Kimball, S. K., and M. S. Mulekar, 2004: A 15-year climatology of North Atlantic tropical cyclones. Part I: Size parameters. *J. Climate*, 17, 3555Ű3575.

[14] Koch, A. L., 1966: The logarithm in biology. *J. Theoret. Biol.*, 12, 276-290.

[15] Limpert, E., W. A. Stahel, and M. Abbt, 2001: Log-normal distributions across the sciences: keys and clues. *BioScience*, 51, 5, 341Ű352.

[16] Liu, K. S., and J. C. L. Chan, 1999: Size of tropical cyclone as inferred from ERS-1 and ERS-2 data. *Mon. Wea. Rev.*, 127, 2992Ű3001.

[17] Merrill, R. T., 1984: A comparison of large and small tropical cyclones. *Mon. Wea. Rev.*, 112, 1408-1418.

[18] Mitzenmacher M., 2004: A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, 1, 2, 226-251.

[19] National Hurricane Center HURDAT Hurricane Best Track files. http://www.nhc.noaa.gov/pastall.shtml#hurdat. Accessed 9 April 2010.