Data Article

Dataset of outer tropical cyclone size from a radial wind profile

Albenis Pérez-Alarcón\textsuperscript{a,b,*}, Rogert Sorí\textsuperscript{a,c}, José C. Fernández-Alvarez\textsuperscript{a,b}, Raquel Nieto\textsuperscript{a}, Luis Gimeno\textsuperscript{a}

\textsuperscript{a}Centro de Investigación Mariña, Environmental Physics Laboratory (EPhysLab), Universidade de Vigo, Campus As Lagoas s/n, Ourense 32004, Spain
\textsuperscript{b}Departamento de Meteorología, Instituto Superior de Tecnologías y Ciencias Aplicadas, Universidad de La Habana, La Habana 10400, Cuba.
\textsuperscript{c}Instituto Dom Luiz, Faculdade de Ciências da Universidade de Lisboa, Campo Grande 1749-016, Portugal

A R T I C L E   I N F O

Article history:
Received 25 August 2021
Accepted 10 January 2022
Available online 13 January 2022

Keywords:
Tropical cyclone
Tropical cyclones database
Tropical cyclone climatology
Tropical cyclones size

A B S T R A C T

Pérez-Alarcón et al. [1] developed a comparative climatology of the outer radius of tropical cyclones (TCs) from several radial wind profiles. They showed that the Willoughby et al. (2006) (W06) profile can be used to reproduce the TC tangential wind speed; thus, this profile is skilful for estimating the TC outer radius. Here, we present a database of TC sizes estimated from the W06 radial wind profile in each cyclogenetic basin worldwide. The database incorporates the critical wind radii, where the tangential wind speed is approximately 17.5 m s\(^{-1}\) (R\(_{34}\)), 26 m s\(^{-1}\) (R\(_{50}\)), 33 m s\(^{-1}\) (R\(_{64}\)), and 51 m s\(^{-1}\) (R\(_{106}\)), estimated by the W06 profile. The database has a comma-delimited text format with six-hour information on the location, maximum wind speed, central pressure, and the different TC metrics mentioned above. This database has a similar structure to that of the Atlantic Hurricane Database (HURDAT2) of the National Hurricane Center. The database presented here is applicable to studies on TC storm surge risks as well as to the determination of the sources and sinks of atmospheric moisture related to tropical cyclogenesis processes.

DOI of original article: 10.1016/j.wace.2021.100366
* Corresponding author at: Centro de Investigación Mariña, Environmental Physics Laboratory (EPhysLab), Universidade de Vigo, Campus As Lagoas s/n, Ourense 32004, Spain.
E-mail address: albenis.perez.alarcon@uvigo.es (A. Pérez-Alarcón).
Social media: \textsuperscript{a} (@A. Pérez-Alarcón), \textsuperscript{b} (@J.C. Fernández-Alvarez), \textsuperscript{c} (@R. Nieto)

https://doi.org/10.1016/j.dib.2022.107825

2352-3409/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject | Climatology |
|---------|-------------|
| Specific subject area | Tropical cyclones size climatology |
| Type of data | Text |
| How data were acquired | The data were acquired by processing the best-track databases of the National Hurricane Center and the Joint Typhoon Warning Center, which are freely available. |
| Data format | Secondary Data (Analyzed) |
| Parameters for data collection | For database preparation, the entries with missing maximum wind speed values in the best-track database records were excluded. |
| Description of data collection | The database was created from the Willoughby et al. [2] radial wind profile with known tropical cyclone position and intensity in each best-track record. |
| Data source location | Institution: Environmental Physics Laboratory, Faculty of Science, University of Vigo, City/Town/Region: Ourense, Country: Spain |
| Primary data sources: The National Hurricane Center tropical cyclones best track archive (HURDAT2, https://www.nhc.noaa.gov/data/#hurdat) and the Joint Typhoon Warning Center best track tropical cyclones records (https://www.metoc.navy.mil/jtwc/jtwc.html?best-tracks) | |
| Data accessibility | With the article and at http://doi.org/10.17632/8997r89fbf.1 |
| Related research article | A. Pérez-Alarcón, R. Sorí, J. C. Fernández-Alvarez, R. Nieto, L. Gimeno, Comparative climatology of outer tropical cyclone size using radial wind profiles. Weather and Climate Extremes. 33 (2021), 100366. 10.1016/j.wace.2021.100366 |

Value of the Data

- This dataset provides climatology information on the tropical cyclone sizes over each cyclogenetnic basin worldwide since the beginning of cyclone records, and this information is useful for different applications.
- Researchers could use this dataset for many applications, including different risk analyses.
- Researchers could use this dataset to analyse the rain rate radial distribution around the TC centre to map potential storm surge risk impacts on the global population and to determine sources and sinks of atmospheric moisture related to tropical cyclone genesis as well as to intensification and weakening mechanisms.
- Tropical cyclone forecasters could use the methodology used to generate this database to quickly determine the TC size using the position and intensity predicted by numerical weather forecast models and thus make their tropical cyclone warnings applicable to a larger population exposed to the possible impact of the storm.
- Machine learning researchers can use the dataset for benchmarking the performance of different methodologies to obtain tropical cyclone sizes.

1. Data Description

The size of tropical cyclones (TCs), as well as the TC intensity, is a very important structural parameter. Determining the TC size climatology is essential for adopting strategies that minimize the damage caused by these storms.
The TC size database presented here has a comma-delimited text format with six-hourly information on the date, location, maximum wind speed (km h\(^{-1}\)), minimum central pressure (hPa), estimated radius of maximum wind speed (km), estimated \(R_{34}, R_{50}, R_{64},\) and \(R_{100}\) critical wind radii, and outer radius (km) calculated by the Willoughby et al. \cite{2} radial wind profile of all known TCs and subtropical cyclones. The TCSize database format is similar to that described by Landsea and Franklin \cite{3} for the HURDAT2 database. The supplementary files contain the full TCSize database for each TC basin, which also can be download from the Mendeley Data repository \cite{4}. Below is an example of the TCSize database entries, with each part defined.

AL112017, IRMA, 66,
20170902, 1800, 18.7N, 44.1W, 175.7, 973, 29.99, 237.0, 145.5, 93.00, -9999, 786.00
20170903, 0000, 18.5N, 45.5W, 175.7, 973, 29.89, 236.5, 145.0, 93.00, -9999, 784.00
20170903, 0600, 18.2N, 46.7W, 175.7, 973, 29.73, 235.0, 144.0, 92.50, -9999, 781.50
20170903, 1200, 17.9N, 47.9W, 185.0, 969, 28.43, 237.0, 148.5, 97.50, -9999, 771.00
20170903, 1800, 17.6N, 49.2W, 185.0, 965, 28.29, 236.0, 147.5, 97.00, -9999, 768.50
20170904, 0000, 17.3N, 50.4W, 185.0, 959, 28.14, 235.0, 146.5, 96.00, -9999, 766.00
20170904, 0600, 17.0N, 51.5W, 194.2, 952, 26.91, 236.0, 150.0, 100.5, 32.50, 754.00
20170904, 1200, 16.8N, 52.6W, 203.5, 945, 25.77, 236.5, 152.5, 104.0, 37.00, 742.00
20170904, 1800, 16.7N, 53.9W, 212.7, 944, 24.72, 236.5, 154.5, 107.0, 41.00, 730.50
20170905, 0000, 16.6N, 55.1W, 231.2, 943, 22.79, 235.0, 157.5, 112.0, 48.00, 705.50
20170905, 0600, 16.6N, 56.4W, 249.7, 933, 21.04, 232.0, 158.5, 115.5, 53.50, 678.50
20170905, 1200, 16.7N, 57.8W, 277.5, 929, 18.70, 223.5, 156.5, 117.5, 59.50, 635.50
20170905, 1800, 16.9N, 59.2W, 286.7, 926, 18.04, 221.0, 156.0, 118.0, 61.00, 622.00
20170906, 0000, 17.3N, 60.6W, 286.7, 915, 18.16, 222.5, 157.0, 118.5, 61.50, 625.50
20170906, 0545, 17.7N, 61.8W, 286.7, 914, 18.28, 224.0, 158.0, 119.5, 62.00, 629.00

The TCSize database contains two line types. The first type is the heading with information about the basin, the number, and the year of TC genesis, and the second type contains information on the different parameters of the TC. The first line has the following format:

AL112017, IRMA, 66,

The first two spaces correspond to the basin identifier (AL: North Atlantic, EP: East Pacific, CP: Central Pacific, WP: Western North Pacific, IO: North Indian Ocean, SI: South Indian Ocean, and SP: South Pacific Ocean). The third and fourth characters refer to the Automated Tropical Cyclone Forecast (ATCF, Sampson and Schrader \cite{5}) TC number. Characters 5-8 before the first comma represent the year. The characters between the first and the second commas refer to the TC name. The TC entry number in the database appears after the second comma. The rest of the lines contain the data information for each TC. These have the following format:

20170905, 0000, 16.6N, 55.1W, 231.2, 943, 22.79, 235.0, 157.5, 112.0, 48.00, 705.50

The element in the first column represents the date in the yyyydddmm (20170905) format. In the second column, the UTC time of the report appears in the hhmm (0000) format, generally coinciding with the synoptic time, although non-synoptic times are found that indicate either TC landfall or the peak maximum intensity. The third column contains the latitude (16.6N), specifying the hemisphere (N: Northern Hemisphere and S: Southern Hemisphere), and similarly, the fourth column represents the longitude (55.1W) (W: Western Hemisphere and E: Eastern Hemisphere). The fifth and sixth columns contain the maximum wind speed in km h\(^{-1}\) (231.2 km h\(^{-1}\)) and the minimum central pressure in hPa (943 hPa), respectively, while the seventh column records the estimated radius of maximum wind speed in km (26.91 km). The eighth to the twelfth columns contain the estimated \(R_{34}, R_{50}, R_{64},\) and \(R_{100}\) critical wind radii (in km), respectively. The last column contains the TC outer radius in km (754.00 km) estimated from the radial wind profile of Willoughby et al. \cite{2}. Missing values are given by -9999.

The global mean size of tropical cyclones from the database presented in this article is 755.2 km with a standard deviation of 109.5 km and a 95% confidence interval of [754.4, 755.7] km. The interpercentile range (25th–95th percentiles) is from 689.0 to 933.0 km. The median storm size is largest in the North Atlantic and smallest in the North Indian Ocean basins. The
Fig. 1. Spatial distribution map of the outer tropical cyclone sizes estimated using the Willoughby et al. [2] radial wind profile.

Fig. 2. Mean outer TC size (km) in each basin: North Atlantic (NATL), Central and East Pacific (NEPAC) Western North Pacific (WNP), North Indian Ocean (NIO), South Indian Ocean (SIO), and South Pacific Ocean (SPO).

Western North Pacific exhibits the largest variance in size of any basin, while the largest coefficients of variation are found in the basins of the Southern Hemisphere. Fig. 1 shows the global distribution of the TC sizes, whereas Fig. 2 shows a comparison of the mean TC size in each basin.
2. Experimental Design, Materials and Methods

The database of the TC sizes in each cyclogenetic basin worldwide was obtained from the Willoughby et al. [2] radial wind profile. This wind profile is given as follows:

\[
V = \begin{cases} 
V_{\text{max}} \left( \frac{r}{r_m} \right)^n r \leq R_1 \\
V_t (1 - w_1) + V_0 w_1 R_1 \leq r \leq R_2 \\
V_{\text{max}} \left[ (1 - A) e^{-\frac{(r-r_m)}{X_1}} + Ae^{-\frac{(r-r_m)}{X_2}} \right] r > R_2 
\end{cases}
\]

where \( V \) is the tangential wind speed at a distance \( r \) from the centre, \( V_{\text{max}} \) is the maximum wind speed, \( r_m \) is the radius of maximum wind speed, \( A, X_1, \) and \( X_2 \) are parameters associated with the TC intensity, \( V_t \) and \( V_0 \) are the tangential wind speeds at radii of \( R_1 \) and \( R_2 \), respectively, and \( w \) is a weight function (see Willoughby et al. [2]). The radius of maximum wind speed was estimated following Willoughby et al. [2] for the North Atlantic (NATL) and Central and East Pacific (NEPAC) basins, whereas for the Western North Pacific (WNP), North Indian Ocean (NIO), South Indian Ocean (SIO), and South Pacific Ocean (SPO) basins, it was estimated following Tan and Fang [6].

Determination of the radial wind profile requires the position and intensity of the storm. This information, for the NATL and NEPAC basins, was extracted from the National Hurricane Center (NHC) HURDAT2 database [2], which is freely available at https://www.nhc.noaa.gov/data/hurdat, whereas for the rest of the cyclogenetic basins, this information was taken from the Joint Typhoon Warning Center (JTWC) best-track databases available at https://www.metoc.navy.mil/jtwc/jtwc.html?best-tracks. The NATL best-track contains information from 1851 to 2020; NEPAC, from 1949 to 2020; and the rest of the basins, from 1945 to 2019.

Similar to the procedure described by Knaff et al. [7], a tangential wind speed of 2 m\( \text{s}^{-1} \) was assumed as the threshold to define the TC size. Thus, from the Willoughby et al. [2] radial wind profile, the radial distance from the centre where the tangential wind speed was equal to or less than 2 m\( \text{s}^{-1} \) was considered as the tropical cyclone outer radius. A similar procedure was applied to determine the \( R_{34} \), \( R_{50} \), \( R_{64} \), and \( R_{100} \) critical wind radii.

![Fig. 3. Global tropical cyclone tracks by ocean basin based on historical TC best track datasets. Tropical cyclone intensities are displayed with seven scales according to the Saffir-Simpson scale: TD - Tropical Depression, TS - Tropical Storm, H1 - Hurricane category 1, H2 - Hurricane category 2, H3 - Hurricane category 3, H4 - Hurricane category 4, and H5 - Hurricane category 5.](image-url)
During processing of the best-track database records, missing values of the maximum wind speed were disregarded; however, the records encoded with -9999 were incorporated into the TC size database presented in this article, with the intention of keeping all records of each tropical cyclone in each basin. Fig. 3 shows the trajectories of all tropical and subtropical cyclones in each basin since the beginning of the records.

Ethics Statements

Not applicable

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

CRediT Author Statement

Albenis Pérez-Alarcón: Conceptualization, Data curation, Methodology, Formal analysis, Software, Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing; Rogert Sorí: Conceptualization, Methodology, Formal analysis, Investigation, Validation, Writing – review & editing; José C. Fernández-Alvarez: Conceptualization, Data curation, Methodology, Investigation, Validation, Visualization, Writing – review & editing; Raquel Nieto: Methodology, Validation, Writing – review & editing, Supervision; Luis Gimeno: Methodology, Validation, Writing – review & editing, Supervision.

Acknowledgments

A.P-A acknowledges a PhD grant from the Universidade de Vigo. J.C.F-A and R.S acknowledge the support from the Xunta de Galicia (Galician Regional Government) under the Grants no. ED481A-2020/193 and ED481B 2019/070, respectively. The authors acknowledge NHC and JTWC for freely provide the best track archives. We acknowledge the funding for open access from the Universidade de Vigo/Consorcio Interuniversitario do Sistema Universitario de Galicia (CISUG) and the partial support from the Xunta de Galicia under the Project ED431C 2021/44. The authors also thank to anonymous reviewers for their comments and suggestions to improve the manuscript.

Funding

This work is supported by the LAGRIMA project (Grant no. RTI2018-095772-B-I00) funded by the Ministerio de Ciencia, Innovación y Universidades, Spain. Partial support was also obtained from the Xunta de Galicia under the Project ED431C 2021/44 (Programa de Consolidación e Estructuración de Unidades de Investigación Competitivas (Grupos de Referencia Competitiva) and Consellería de Cultura, Educación e Universidade).

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi: 10.1016/j.dib.2022.107825.
A. Pérez-Alarcón, R. Sorí and J.C. Fernández-Alvarez et al. / Data in Brief 40 (2022) 107825

References

[1] A. Pérez-Alarcón, R. Sorí, J.C. Fernández-Alvarez, R. Nieto, L. Gimeno, Comparative climatology of outer tropical cyclone size using radial wind profiles, Weather Clim. Extrem. 33 (2021) 100366, doi:10.1016/j.wace.2021.100366.
[2] H.E. Willoughby, R.W.R. Darling, M. Rahn, Parametric representation of the primary hurricane vortex. Part II: A new family of sectionally continuous profiles, Mon. Weather Rev. 134 (2006) 1102–1120, doi:10.1175/MWR3106.1.
[3] C.W. Landsea, J.L. Franklin, Atlantic hurricane database uncertainty and presentation of a new database format, Mon. Weather Rev. 141 (2013) 3576–3592, doi:10.1175/MWR-D-12-00254.1.
[4] A. Pérez-Alarcón, R. Sorí, J.C. Fernández-Alvarez, R. Nieto, L. Gimeno, Dataset of outer radius of tropical cyclone from a radial wind profile, Mendeley Data V1 (2021), doi:10.17632/8997r89fbf.1.
[5] C.R. Sampson, A.J. Schrader, The automated tropical cyclone forecasting system (version 3.2), Bull. Am. Meteorol. Soc. 81 (2000) 1231–1240, doi:10.1175/1520-0477(2000)081⟨1231:TATCFS⟩2.3.CO;2.
[6] C. Tan, W. Fang, Mapping the wind hazard of global tropical cyclones with parametric wind field models by considering the effects of local factors, Int. J. Disaster Risk Sci. 9 (2018) 86–99, doi:10.1007/s13753-018-0161-1.
[7] J. Knaff, S.P. Longmore, D.A. Molenar, An objective satellite-based tropical cyclone size climatology, J. Clim. 27 (2014) 455–476, doi:10.1175/JCLI-D-15-0610.1.