https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Open Access Earth System Science Data Discussions

A Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth Century (RGTracks-20C)

Guiling Ye[1,2,3], Jeremy Cheuk-Hin Leung[2,4]*, Wenjie Dong[1,3]*, Jianjun Xu[5], Weijing Li[6], Weihong Qian[7], Hoiio Kong[4], Banglin Zhang[2,8]

Affiliations

[1] School of Atmospheric Sciences, Key Laboratory of Tropical Atmosphere-Ocean System, Ministry of Education, Sun Yat-sen University, Zhuhai, China

[2] Hunan Institute of Advanced Technology, Changsha, China

[3] Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China

[4] Faculty of Data Science, City University of Macau, Macau, China

[5] CMA-GDOU Joint Laboratory for Marine Meteorology, South China Sea Institute of Marine Meteorology, Guangdong Ocean University, Zhanjiang, China

[6] National Climate Center, China Meteorological Administration, Beijing, China

[7] Department of Atmospheric and Oceanic Sciences, Peking University, Beijing, China

[8] College of Atmospheric Science, Lanzhou University, Lanzhou, China

Corresponding author(s): Jeremy Cheuk-Hin Leung (chleung@pku.edu.cn); Wenjie Dong (dongwj3@mail.sysu.edu.cn)

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Abstract

Tropical cyclones (TCs) are among the deadliest disasters affecting human society, and their response to climate change has widely drawn attention from the public. However, assessing how historical TC activity changed with climate change has proven challenging due to incomplete TC records in the early years. Here, we introduce the Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth Century (RGTracks-20C) (Ye et al., 2024), a publicly available century-long global TC track dataset spanning from 1850–2014. The RGTracks-20C is reconstructed from the National Oceanic and Atmospheric Administration Twentieth Century Reanalysis using two independent TC tracking algorithms. Validation based on observations confirms that the RGTracks-20C effectively captures the climatology and long-term variability of TC numbers, tracks, duration, and intensity across various ocean basins. A remarkable key strength of the RGTracks-20C is its ability to fill the missing intensity and location records of TCs observed in early years. This dataset serves as a valuable historical data reference for future research on climate change and TC-related disasters.

1. Introduction

Tropical cyclones (TCs), also known as hurricanes or typhoons, are intense weather systems that form over tropical and subtropical oceans and can cause severe disasters over the coastal regions and even inland areas (Qin et al., 2024; Zhu and Quiring, 2022). Globally, approximately 80 TCs are generated each year (Emanuel, 2018). As one of the most destructive weather systems (Bloemendaal et al., 2022; Dinan, 2017; Emanuel, 2017), TCs significantly impact society and the economy (Kunze, 2021; Lenzen et al., 2019; Noy, 2016). These impacts are expected to be exacerbated by climate change in the future (Chan, 2023; Hassanzadeh et al., 2020; Knutson et al., 2020; Moon et al., 2023; Murakami and Wang, 2022; Yamaguchi et al., 2020). Therefore, research on TCs has become increasingly vital in climate change and prediction (Bhatia et al., 2019; Chan, 2019; Lanzante, 2019; Moon et al., 2019; Sharmila and Walsh, 2018; Zhang et al., 2019). However, past variability of TC activity and underlying mechanisms remains challenging due to incomplete early historical TC observation records, which may lead to controversies (Chan et al., 2022a, b; Knutson et al., 2019; Lee et al., 2020).

Previous research has revealed significant issues related to the completeness of historical TC observational data (Lee et al., 2020), which are highly dependent on the development of the global TC observation system, data analysis techniques, and other factors (Klotzbach and Landsea, 2015; Knapp et al., 2010; Kossin et al., 2020; Landsea et al., 2010; Mann et al., 2007; Ying et al., 2014). Before the introduction of satellite observation, TC information (e.g., intensity and location) primarily relied on conventional coastal weather stations and ship observation reports (Landsea et al., 2006, 2008). Aircraft reconnaissance emerged in the North Atlantic (NATL) and western North Pacific (WNP) after World War II (Emanuel, 2008). However, these observational techniques could not capture all occurred TCs due to their limited observation range. It is possible that an existing TC was unrecorded in the early years. In addition, even if a TC was observed and recorded, its track and intensity information may be discontinuous due to the absence of meteorological satellite observations. For instance, there were no observational records of TC wind speeds in the southern hemisphere before 1956 (Emanuel, 2021). Storm intensity in the Indian Ocean is weaker compared to other basins, partly due to the lack of direct coverage by geostationary satellites in that region before 1998 (Schreck et al., 2014). The incomplete observed data of TCs in the early years, mainly before the 1970s, is a commonly-known unsolved issue in the community.

Given the limitations of historical TC records, a promising approach is to utilize reanalysis for TC identification (Li et al., 2024; Truchelut and Hart, 2011). Reanalysis combines historical observational data with modern numerical weather models to produce comprehensive, continuous datasets of global atmospheric conditions that adhere to physical principles (Compo et al., 2011; Kalnay et al., 1996; Parker, 2016; Slivinski, 2018). The Twentieth Century Reanalysis (20CR) (Compo et al., 2011), provided by the National Oceanic and Atmospheric Administration (NOAA), is a global reanalysis dataset that covers the longest period among all other reanalyses. The 20CR was designed for long-term analyses from individual extreme weather events to climate variability, and has been applied to a wide range of studies, including those on wave height, storm surge, Madden-Julian Oscillations, and TCs (Chand et al., 2022; Cid et al., 2017; Gergis et al., 2020; Lee et al., 2023; Leung et al., 2022; Moore and Babij, 2017; Slivinski et al., 2019; Truchelut et al., 2013; Wang et al., 2012). The fact that the 20CR only assimilates surface

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

pressure and sea level pressure fields, instead of other observations such as satellites and aircraft, makes it less sensitive to the temporal inhomogeneity of observations (Slivinski et al., 2019, 2021).

Several independent studies have documented the feasibility of reproducing the characteristics of some historical TC events based on the 20CR (Emanuel, 2010; Lee et al., 2023; Slivinski et al., 2019; Truchelut et al., 2013; Truchelut and Hart, 2011). For example, following Emanuel (Emanuel, 2010), who first expanded and revised TC climatology based on the 20CR, Truchelut and Hart (2011) employed the 20CR to identify previously unknown TCs in the Atlantic and demonstrated that the 20CR can accurately describe large-scale TC thermodynamic structure. Recently, Truchelut et al. (2013) noted that the 20CR has the ability to investigate TC events that were previously undetected in the pre-satellite era. Compared to other reanalyses, the 20CR well captures the intensity of the 1915 Galveston hurricane (Slivinski et al., 2019) and also offers a more accurate representation of landfalling TCs in East Asia (Lee et al., 2023). These previous studies have demonstrated the effectiveness of the 20CR as a tool for characterizing historical TCs (Emanuel, 2010; Truchelut et al., 2013; Truchelut and Hart, 2011). Taking advantage of the 20CR, some researchers have extracted the century-long TC information from the reanalysis product (Chand et al., 2022; Lee et al., 2023; Yeasmin et al., 2023), suggesting its potential as a tool for studying historical changes in TCs under anthropogenic climate change.

While the 20CR has been applied to studying the relationship between historical climate change and TC variability, the primary focus was mostly on the TC occurrence frequency, and little attention was given to other TC metrics such as intensity, duration, and location. More importantly, to date, there is no publicly available reanalysis-based global TC dataset covering a century-long period. Therefore, the main objective of this study is to extract TC information (including location, intensity, and lifetime) from the 20CR and reconstruct a historical global TC track dataset spanning 1850–2014. The produced dataset is named the Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth Century (RGTracks-20C) and is open to the public for research use. This paper first introduces the production details of the RGTracks-20C and then discusses the validity, key strengths, and usage notes of the datasets. We anticipate that the RGTracks-20C will provide valuable insights into the changing patterns of historical TC activity, improving our understanding of the response of TCs to climate change.

2. Data and methods

2.1 Data

The primary objective of this study was to reconstruct a 20th century global TC dataset from the 20th Century Reanalysis version 3 (20CRv3) (Slivinski et al., 2019, 2021), the latest version of the 20CR produced by NOAA. Then, the validity of the reconstructed 20th century global TC data was evaluated based on the observed TC records, i.e., the International Best Track Archive for Climate Stewardship (IBTrACS) dataset (Knapp et al., 2010).

2.1.1 20th Century Reanalysis

The 20CRv3 is led by NOAA's Physical Sciences Laboratory (PSL) and the Cooperative Institute for Research in Environmental Sciences (CIRES) at the University of Colorado, supported by the U.S. Department of Energy (DOE) (Slivinski et al., 2019, 2021). It, by combining advanced data assimilation

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

and numerical prediction techniques with historical observation data, provides long-term historical weather data with diverse variables, complete spatial and temporal coverage. The 20CRv3 employs sea-surface temperature and sea-ice distributions as its boundary conditions and assimilates only surface pressure reports from the International Surface Pressure Databank (ISPD) version 4.7 (Compo et al., 2019; Cram et al., 2015), which include observations from stations and ships, as well as TC intensity (the minimum central pressure (SLP_{min}) from the IBTrACS (Knapp et al., 2010). As such, it is more consistent and homogeneous with time than other reanalyses (Slivinski et al., 2019).

One should note that the IBTrACS and 20CRv3 are not two independent datasets because the SLP_{min} records in the IBTrACS are partly assimilated in the production of 20CRv3. However, reports show that 20CRv3 shows TCs structure and intensity more accurately and closer to observations than other 20th century reanalyses as a result of the assimilation of IBTrACS (Laloyaux et al., 2018; Slivinski et al., 2019). And, it provides a four-dimensional global gridded atmospheric dataset that spans the whole 20th century and part of the 19th century (1836–2015, with an experimental extension spanning 1806–35), with a 3-hour temporal resolution and 1°×1° horizontal resolution (Slivinski et al., 2021). Thus, the 20CRv3 was applied to the production of the RGTracks-20C in this paper.

2.1.2 IBTrACS

The IBTrACS (Knapp et al., 2010), published by the NOAA, merges recent and historical TC data from meteorological agencies worldwide. These include the Regional Specialized Meteorological Centers (RSMC) and Tropical Cyclone Warning Centers (TCWC) of the World Meteorological Organization (WMO), as well as non-WMO Centers, such as the China Meteorological Administration, the Hong Kong Observatory and the Joint Typhoon Warning Center. The IBTrACS is the most comprehensive and publicly available global TC best-track dataset. It has been widely applied in previous research to investigate the characteristics of TCs (Lai et al., 2020; Li et al., 2023; Tu et al., 2021, 2022; Wang and Toumi, 2022; Zhang, 2023), and has served as a criterion for assessing TC records derived from reanalysis (Bell et al., 2018; Bourdin et al., 2022; Chand et al., 2022; Hodges et al., 2017; Lee et al., 2023). In this study, the most updated version of IBTrACS (v04) (Knapp et al., 2018) serves as an observation reference for evaluating the reliability of the RGTracks-20C. This dataset was cleaned before being used for analyses. Details about the data pre-processing procedures are referred to in Figure B1 in Bourdin et al. (2022). In particular, we standardized maximum sustained wind speeds ($WIND_{max}$) in IBTrACS to 10-minute sustained wind speeds to ensure a consistent global standard(Knapp et al., 2010). We then removed tracks that did not reach the tropical storm stage ($WIND_{max} < 16\ m \cdot s^{-1}$) and those that lasted shorter than two days.

Although the IBTrACS has time coverage dating back to the early 20th century, we utilize the data only for the post-satellite period (1979–2014) due to the early data incompleteness issues (Chang and Guo, 2007; Lee et al., 2020; Truchelut et al., 2013). Given that the IBTrACS is the most reliable record of TCs after the 1970s, the IBTrACS serves as the best benchmark for validating the data quality of RGTracks-20C.However, because the starting years of records vary across basins within the IBTrACS, biases may occur in the assessment results (see Supplementary Sects. S2.2 and S2.4 for more details).

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Open Access Earth System Science Data Discussions

2.2 Production of the RGTracks-20C

2.2.1 Procedure

The RGTracks-20C was constructed from the latest version of 20CR (20CRv3). The relatively short and imperfectly sampled observational record of TCs introduces considerable uncertainty in their data over the past century (Landsea, 2007; Landsea et al., 2010), hindering accurate detection of interannual variability and long-term trends (Knutson et al., 2019; Lee et al., 2020). Reanalysis is an effective way to reduce this uncertainty (Chand et al., 2022; Truchelut et al., 2013). Since TC information is not directly provided in the 20CRv3, objective TC trackers were applied to detect and track TCs in this dataset. Numerous trackers have been developed by operational centers and research institutions to meet various application needs (Hodges et al., 2017; Horn et al., 2014; Tory et al., 2013; Zarzycki and Ullrich, 2017). In this study, as the first version of the RGTracks-20C, we applied two widely used, publicly available, and effective trackers: (1) the physically-based Ullrich & Zarzycki (UZ) tracker (Zarzycki and Ullrich, 2017) and (2) the dynamics-based Okubo-Weiss-Zeta (OWZ) tracker (Tory et al., 2013). Both trackers have been reported to effectively capture TC systems from coarse resolution gridded data uncertainty (Chand et al., 2022; Truchelut et al., 2013), such as the 20CRv3. Figure 1 shows the procedure of producing the RGTracks-20C, and details of the methodology are provided in the following.

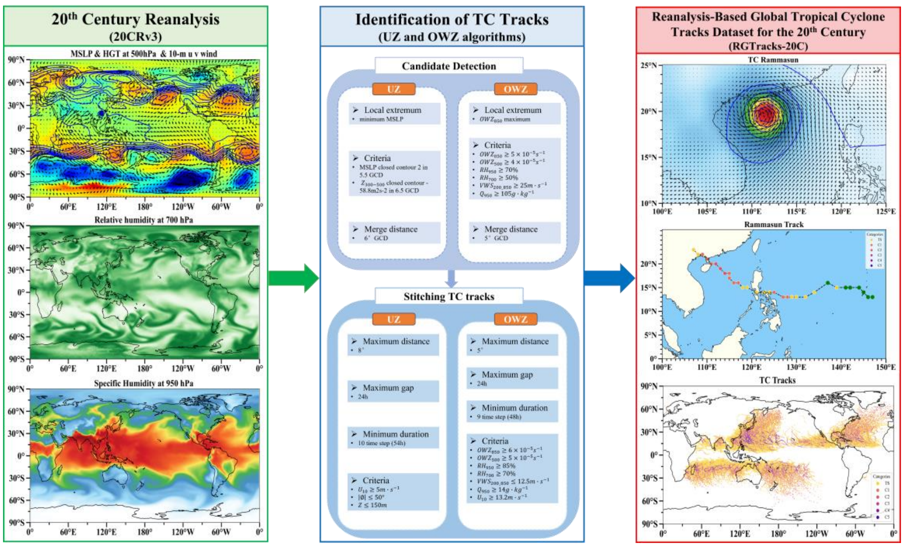

Figure 1: Schematic diagram showing the production of the RGTracks-20C from the 20CRv3 based on the UZ and OWZ tracking algorithms. Variables shown include U10: 10-m wind speed, $\emptyset$: latitude, z:altitude, GCD: great circle distance.

2.2.2 TC tracker

i. OWZ Tracker

The OWZ tracker, initially proposed by Tory et al. (2013), is designed to detect low-deformation vorticity regions within large-scale disturbances, typically situated within the so-called "marsupial pouch", which have the potential for tropical storm formation. Given that the OWZ approach relies solely on large-scale variables, it is particularly effective in detecting TC in coarse-resolution models or reanalysis (Bell et al., 2018; Bourdin et al., 2022).

The OWZ tracker involves a low-deformation vorticity variable parameter, which is the product of absolute vorticity and the Okubo-Weiss parameter normalized by the vertical components of relative vorticity squared (Eq. 1):

$$OWZ = \mathrm{sgn}(f) \times (\zeta + f) \times max\left[\frac{\zeta^2-(E^2+F^2)}{\zeta^2}, 0\right] \tag{1}$$

where f is the Coriolis parameter, $\zeta = \partial v/\partial x - \partial u/\partial y$ is the vertical component of relative vorticity, $(\zeta + f)$ is the absolute vorticity, E is the stretching deformation (Eq. 2), and F is the shearing deformation (Eq. 3):

$$E = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \tag{2}$$

$$F = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \tag{3}$$

First step: Candidate detection.

The OWZ tracker begins by identifying local maxima of OWZ at 850 hPa. Any candidate with a stronger OWZ maximum within 5° of great circle distance (GCD) is excluded. Next, only candidates that meet the six initial threshold conditions shown in Table 1 within a 2° GCD of the identified maximum are retained. Based on the information provided in Table 1, besides the required minimum threshold values for the OWZ parameter at 850 hPa and 500 hPa, additional dynamical and thermodynamic parameters related to TC formation are taken into account. These parameters include the maximum threshold for the wind vector difference (vertical wind shear) between 850 hPa and 200 hPa, as well as the relative humidity at 950 hPa and 700 hPa, and the minimum threshold for the specific humidity at 950 hPa. This step primarily aims to identify grid points that contain essential components of a storm. Subsequently, neighboring grid points are grouped together to define potential TCs.

Second step: Stitching TC tracks.

Consecutive TC points are linked together if their distance does not exceed 5° of GCD and there is a maximum gap of 24 hours between them. To be considered as a valid TC, additional core thresholds (shown in Table 1) must be met for at least 9 time-steps (48 hours). Finally, tracks that do not maintain tropical storm intensity (wind speed at 10 m ≥ 12.3 $m \cdot s^{-1}$) for at least 1 time step are excluded.

Table 1. Parameter threshold values for the OWZ detection criteria. Subscripts stand for isobaric levels in hPa (OWZ: Obuko-Weiss-Zeta s^{-1}, RH: relative humidity %; VWS: vertical wind shear $m \cdot s^{-1}$; Q: specific humidity $g \cdot kg^{-1}$.)

Critertion	OWZ_{850}	OWZ_{500}	RH_{950}	RH_{700}	VWS_{200_850}	Q_{950}
Initial	50×10^{-6}	40×10^{-6}	70	50	25	10
Core	60×10^{-6}	50×10^{-6}	85	70	12.5	14

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

ii. UZ tracker

The UZ tracker , originally proposed by Zarzycki and Ullrich (2017), utilizes sea level pressure on the model grid, incorporating criteria for warm-core structures and storm lifetime.

First step: Candidate detection.

Initially, candidates are identified based on the SLP minimum. And, only those candidates that meet the following two closed-contour criteria are kept:

1. An increase in SLP minimum of at least 2 hPa within a 5.5° GCD from the candidate point to ensure the presence of a sufficiently strong and coherent low-pressure area.

2. The geopotential thickness between 300 and 500 hPa (denoted as $Z_{300\text{-}500}$) must decrease by 58.8 m^2s^{-2} over a distance of 6.5° GCD from the maximum center of $Z_{300\text{-}500}$ within 1° GCD of the center of minimum SLP.

Finally, candidates with a stronger SLP minimum within a 6°=GCD are excluded.

Second step: Stitching TC tracks.

The candidates are subsequently linked in time to create paths, ensuring a maximum distance of 8° GCD between candidates. Each path must last for at least 54 hours without gaps longer than 24 hours. Additionally, ten 6-hourly time steps (equivalent to 54 hours) must satisfy the following thresholds: wind speed at 10m $\geq 10\ m/s$ and $z \leq 150$m (where z represents the altitude), and the storm must form between 0° and 50°.

The UZ tracker, developed specifically for high-resolution models and reanalysis data, is designed to maintain a low false-alarm rate, which may lead to a larger number of misses of weaker storms(Roberts et al., n.d.). In contrast, the OWZ tracker, based on the large-scale environmental conditions favorable for TC formation, addresses this limitation. Thus, combining these two TC trackers can effectively enhance the reliability of RGTracks-20C.

A command-line software, TempestExtremes, developed by Zarzycki and Ullrich (2017), enables fast and versatile and versatile implementation of TC trackers, was used in this study. For further details, please refer to Ullrich et al. (2021).

2.2.3 Bias Correction of TC intensity

Given the low horizontal resolution in the 20CRv3, TC intensities derived directly from the reanalysis generally underestimated compared to observations (Fig. 2a) (Bourdin et al., 2022; Roberts et al., n.d.). To address this issue, a quantile mapping bias correction, similar to the method used by Zhao and Held (2010), was applied to adjust for the TC intensity bias within the dataset. The main idea is to fit the 20CRv-derived TC intensity distributions, either probability distribution functions (PDFs) or cumulative distribution functions (CDFs), to the observed distributions. This method has demonstrated significant efficacy in enhancing the accuracy of TC intensity within models or reanalyses (Faranda et al., 2023; Yoshida et al., 2017). This adjustment resulted in a wind-pressure relationship in RGTracks-20C that aligns more closely with observational data (Fig. 2b).

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

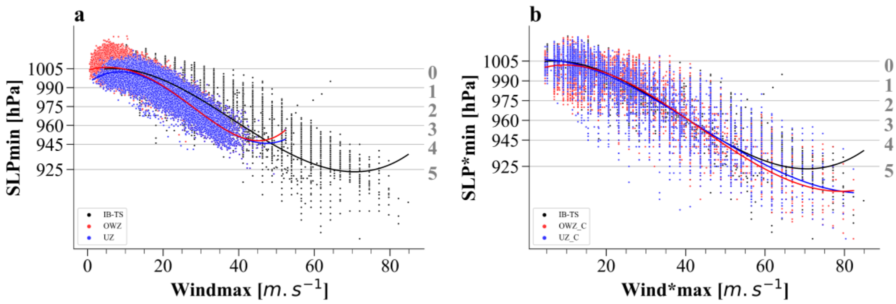

Figure 2: Wind–pressure relationships for IBTrACS and RGTracks-20C. a–b, Scatter plots of SLPmin (unit: hPa) against maximum sustained wind speeds (WINDmax) (unit: $m \cdot s^{-1}$), based on the TCs from IBTrACS (black), OWZ (red), and UZ (blue) trackers, before (a) and after (b) intensity bias correction (see Methods). The curves represent fourth-order polynomial fit results. Storm categories, as defined in the section 'TC intensity', are indicated by horizontal gray lines.

2.3 Verification of RGTracks-20C

2.3.1 Tracks matching

After utilizing the UZ and OWZ trackers to detect TC vortices from the 20CRv3, the resulting tracks are matched one-to-one with those observed in the International Best Track Archive for Climate Stewardship (IBTrACS). The specific procedures are detailed in the "2.4 Tracks Matching" section by Bourdin et al. (2022).

Specifically, a detected track D consists of n points (d_1, d_2, ..., d_n) corresponding to the moments (t_1, t_2, ..., t_n). Similarly, a track O observed in IBTrACS consists of a collection of points at a given time. For every point d_i (t_i) on track D, points from O at the same time t_i located within a 300 km radius of d_i are linked. There may be instances where no such points are found in O. The subset of points in O that are linked to any point in D is labeled as $O_{D-paired}$. It consists of $|O_{D\text{-}paired}|$. There are three possible scenarios:

1. $|O_{D-paired}| = 0$: If none of the points in the RGTracks-20C track D match any points in track O, then track D is classified as a False Alarm (FA).

2. $|O_{D-paired}| > 0$: If all points in $O_{D-paired}$ track correspond to points in the same observed track O, then track O is identified as the closest match for D.

3. $|O_{D-paired}| > 0$: If the points in $O_{D-paired}$ correspond to several observed tracks in O, the observed track with the most points paired with D is regarded as the best match for D.

2.3.2 Track verification

Following the approach suggested by Bourdin et al. (2022), this study compares TC tracks detected from the 20CRv3 with observed tracks from the IBTrACS. The probability of detection (POD) (Eq. 4) and false alarm rate (FAR) (Eq. 5) are used to assess the detection skills of the two trackers.

$$POD = \frac{H}{H + M} \tag{4}$$

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

$$FAR = \frac{FAs}{H + FAs} \tag{5}$$

where hits (H) refer to TC tracks detected from the 20CRv3 that are also present in IBTrACS, misses (M) denote those tracks that are recorded in IBTrACS but were not detected in the 20CRv3, and false alarms (FAs) refer to non-existing TCs that were detected from the 20CRv3.

2.4 Definitions

2.4.1 TC intensity

In assessing the TC intensity, SLP_{min} and $WIND_{max}$ are two commonly used metrics in TC research. However, because $WIND_{max}$ in both observations and reanalysis exhibits relatively higher uncertainties (Bourdin et al., 2022; Chavas et al., 2017; Knapp et al., 2010; Knutson et al., 2015; Schreck et al., 2014), this study opted to use SLP_{min} as the only indicator of TC intensity when verifying the validity of RGTracks-20C. Nevertheless, $WIND_{max}$ of detected TCs is also provided in the RGTracks-20C (Table 2) as a reference for researchers who wish to use and improve the dataset, though it is not discussed in the paper.

Table 2. Data format of the RGTracks-20C. track_id: storm identifier, lat: latitude degrees_north, lon: longitude degrees_east, SLPmin: minimum central pressure ($unit: hPa$), WINDmax: maximum wind speed (units: $m \cdot s^{-1}$), WIND*max and SLP*min denotes TC intensities after bias correction.

track_id,	year	month	day	hour	lon	lat	$WIND_{max}$	SLP_{mi}	hemisphere	basin	season	$WIND^*_{max}$	SLP^*_{mi}
0	1979	1	1	0	142.00	15.00	13.57	996.09	S	SP	1979	13.57	990.00
0	1979	1	1	6	144.00	15.00	14.95	995.27	S	SP	1979	14.95	980.27
…	…	…	…	…	…	…	…	…	…	…	…		
…	…	…	…	…	…	…	…	…	…	…	…		
…	…	…	…	…	…	…	…	…	…	…	…		
2880	2014	12	31	18	120.00	9.00	11.122	1006.20	N	WNP	2014	22.12	998.20

Storm categories: the Saffir-Simpson Hurricane Scale (SSHS) from 1 to 5 based on their peak 1-minute wind speed at 10 meters above the surface. In this study, given the significant uncertainties in WINDmax due to differences between institutions and the limitations of model simulation capabilities (Chavas et al., 2017; Klotzbach et al., 2020; Knutson et al., 2015), we have chosen to classify based on SLPmin, following the definition of Klotzbach et al. (2020).

2.4.2 Basins

We explore the performance of TCs in RGTracks-20C on global and regional scales. The regional division is mainly based on the appendix guide of Knutson et al. (2015), which divides the globe into six basins: the WNP, ENP, South Pacific (SP), NI, South Indian (SI), and NATL.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Open Access Earth System Science Data Discussions

2.4.3 TC days

TC days is defined as the number of 6-hour periods during which an active TC occurs within a basin, divided by 4 (to convert 6-hour blocks into days) and accumulated for the year under consideration such that:

$$TC\ days = \frac{1}{4}\sum_{i=0}^{n} L_i \tag{6}$$

where L_i is the individual lifetime of a TC within the bounds of a basin.

3. Results and discussion

3.1 Data Records

The constructed RGTracks-20C (Ye et al., 2024) provides a century-long collection of global TCs identified from the 20CRv3. The RGTracks-20C is publicly available at the https://github.com/jeremychleung/RGTracks-20C/ and https://zenodo.org/record/8410597. This dataset provides detailed TC information, including location (longitude, latitude, hemisphere, and basin), time (year, month, day, hour, and season), and intensity (SLP_{min} and $WIND_{max}$), with a temporal resolution of 6 hours, spanning from 1850 to 2014 and covering the globe. The dataset is provided as a comma separated values (.csv) file and has a format similar to that of the IBTrACS (Table 2). It is noted that, in the RGTracks-20C, $WIND_{max}$ serves, in addition to SLP_{min}, as a supplementary reference of TC intensity for researchers, but is not discussed here due to accuracy issues and should be used cautiously.

3.2. Validity of trackers

As documented in prior studies, biases are unavoidable when extracting TCs from reanalyses, given the limitations of reanalysis in reproducing the high-resolution TC structure and circulation patterns, as well as the errors caused by the application of different trackers (Bell et al., 2018; Horn et al., 2014; Lee et al., 2023; Slivinski et al., 2019; Truchelut et al., 2013). Before verifying the reliability of RGTracks-20C, it is necessary to evaluate the performance of the two trackers applied.

The POD and FAR of TCs identified by the UZ and OWZ trackers are calculated to assess the ability of the trackers to detect TCs from the 20CRv3 globally and across six basins (see Track verification). Globally, the overall POD and FAR of TCs detected by the UZ tracker are 68% and 7% (Fig. 3a), while those by the OWZ tracker are 77% and 15%, respectively (Fig. 3b). Detailed comparisons of each component of POD and FAR, including the number of hits, false alarms, and misses, are provided in Supplementary Sect. S1 and Fig. S1.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

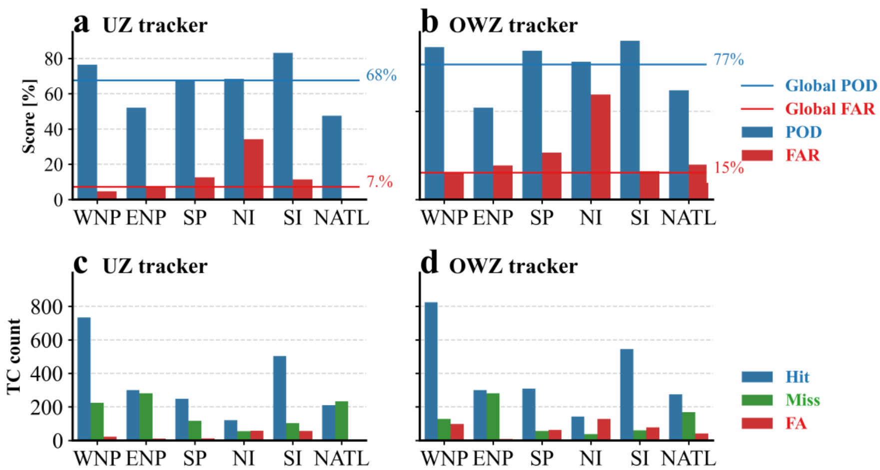

Figure 3: Accuracy of TC number detection of the RGTracks-20C. a–b, POD (blue bars and line, unit: %) and FAR (red bars and line, unit: %) for TC number detected by the UZ (a) and OWZ (b) trackers in each basin (bars), compared to the global mean (lines). Blue and red horizontal lines denote the POD and FAR over the globe. c–d, same as a–b, except for the number of hits (blue bars), misses (green bars), and false alarms (red bars) detected by the UZ (c) and OWZ (d) trackers.

For each basin, the distributions of the POD of TCs (Figs. 3a–b) and the number of hits (Figs. 3c–d) between the two trackers show high similarities. Specifically, both trackers report higher POD values in the SI (90% for OWZ tracker, 83% for UZ tracker), WNP (86% for OWZ tracker, 77% for UZ tracker), and SP (84% for OWZ tracker, 68% for UZ tracker), followed by the NI (78% for OWZ tracker, 68% for UZ tracker). Lower POD values are observed in the NATL (62% for OWZ tracker, 48% for UZ tracker) and the ENP (52% for both OWZ and UZ trackers). Similarly, the largest number of TC hits is observed in the WNP (824 for OWZ tracker, 733 for UZ tracker) and SI (543 for OWZ tracker, 503 for UZ tracker), followed by the ENP, SP and NATL, each with approximately 200–300 TCs, and the NI with fewer than 200 TCs.

The FAR of TCs (Figs. 3a–b), and the number of false alarms (FAs) and misses (Figs. 3c–d) vary between the two trackers. The UZ tracker exhibits FARs below 15% across all basins except the NI. Notably, in the ENP and NATL, the FAR of TCs is below the global average of 7%, with the number of FAs fewer than 20. The OWZ tracker shows a FAR close to the global average (15%) in the WNP and SI, while in the ENP, SP, and NATL, the FAR values range between 15% and 20%. In the NI, however, the two trackers show a relatively higher FAR and more FAs compared to other basins. In terms of missed TC detections, both trackers show relatively few misses, less than 120, in the SP, NI, and SI basins. On the other hand, misses are higher in the ENP and NATL. Overall, the UZ tracker consistently shows a higher number of missed TCs across all basins than the OWZ tracker. This is particularly evident in the WNP and SI, the two basins that account for nearly two-thirds of global TC activity, where the OWZ tracker exhibits fewer missed TC detections (Fig. 3d). Supplementary Sect. S2.1 provides further

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

explanations of the high FAR of TCs observed in the NI, the higher number of missed TCs in the ENP and NATL (Supplementary Fig. S2).

Overall, the accuracies of TC detection by the two tracking algorithms, especially that by the OWZ tracker, have reached the accuracy reported by recent works that extracted TCs from other modern-era reanalyses, such as the fifth generation ECMWF reanalysis (ERA5) (Supplementary Table S1) (Bourdin et al., 2022; Murakami, 2014). This confirms the effectiveness of both trackers in detecting and tracking the majority of TCs from the 20CRv3.

3.3 Climatology of TC activity

Since our target of constructing the RGTracks-20C is to aid the community in studying the response of TCs to climate change, we will focus on the ability of the RGTracks-20C to capture the climatology and long-term variability of TC activity in the following sections.

In terms of climatology, the RGTracks-20C is able to capture the major spatial patterns of TC genesis locations and track density over most ocean basins (Figs. 4a–f), indicating its effectiveness in reproducing the spatial distribution of historically observed TCs. The annual mean TC numbers in most ocean basins detected by the UZ and OWZ trackers are consistent with observations (Figs. 4g–i). The OWZ tracker especially captures the observed annual mean TC number in the NWP, SI, and SP well, with discrepancies ranging from –0.48 to 0.89. Notably, the UZ tracker also accurately estimates observed annual mean TC number in the NI, demonstrating a relatively small error (4.83 versus 4.97) between the two. However, the UZ and OWZ trackers estimate the annual mean number of TCs to be 63.39 and 78.56, respectively (Figs. 4h–i), which are relatively lower than the observed values (87.03, Fig. 4g). The main reason for the global underestimation compared to IBTrACS is the discrepancies in the ENP and NATL, of which the reasons are discussed in Supplementary Sects. S2.1–2.2. Despite the underestimations in individual basins, the overall TC detection rates resemble previous publications that aimed to extract TCs from higher-quality reanalyses (Bourdin et al., 2022; Murakami, 2014). This result verifies the RGTracks-20C's ability to reproduce the climatology of the TC number globally and in most basins.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

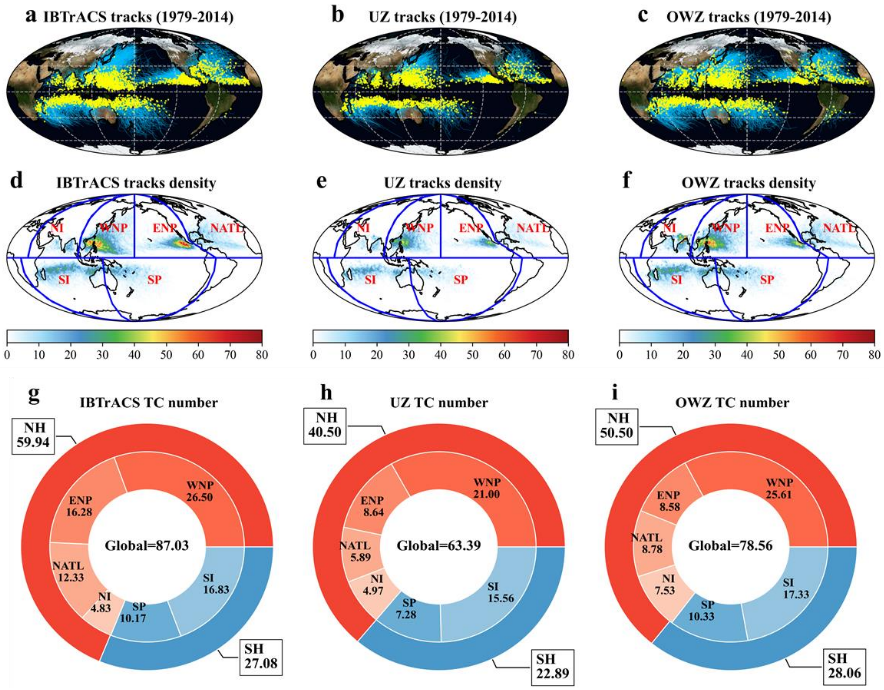

Figure 4: TC genesis locations, tracks, and annual average number from IBTrACS and RGTracks-20C. a–c, TC genesis locations (yellow dots) and tracks (blue lines) from IBTrACS (a), and RGTracks-20C using the UZ (b) and OWZ (c) trackers. d–f, TC tracks density (shading, number of TC occurrence per 1° × 1° latitude-longitude grid box, 1979-2014) from IBTrACS (c), and RGTracks-20C using the UZ (e) and OWZ (f) trackers. g–i, mean number of TCs per year globally and for the six basins from IBTrACS (g), and RGTracks-20C using the UZ (h) and OWZ (i) trackers.

We further evaluate the accuracies of detected TC tracks in the RGTracks-20C by comparing the arc length of TC tracks between RGTracks-20C and IBTrACS. Results indicate that the global TC location errors range from 10 to 300 km, with the majority between 50–100 km for the UZ tracker and 75–125 km for the OWZ tracker (Fig. 5a). Additionally, the peak errors for both trackers are below 100 km, with the UZ and OWZ trackers showing peak values of approximately 75 km and 95 km, respectively. These findings are consistent across all basins (Fig. 6a). Given that the lower limit of the average TC location error expected from the coarse horizontal resolution of the 20CRv3 (1 degree×1 degree) is approximately 100 km, the above-mentioned small mean values of TC location biases confirm that the RGTracks-20C is capable of reproducing most observed TC tracks and locations.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

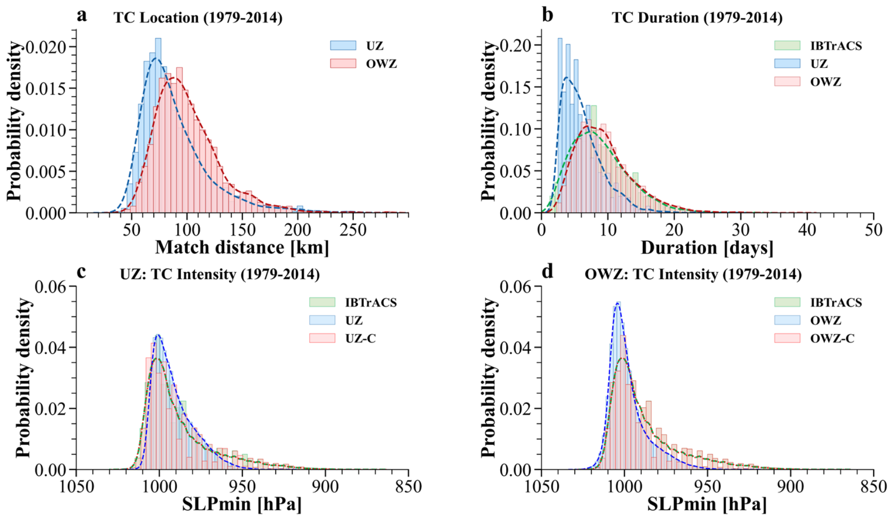

Figure 5: Distribution of TC characteristics om the IBTrACS and RGTracks-20C. a, Distribution of the mean TC location error from 1979–2014 (unit: km) between IBTrACS and the RGTracks-20C by the UZ (blue) and OWZ (red) algorithms. b, TC duration (unit: days) from 1979 to 2014 in IBTrACS (green) and the RGTracks-20C by the UZ (blue) and OWZ (red) algorithms. c, same as (b), but for TC intensity (*SLPmin*, unit: hPa), based on the UZ tracker, before (blue) and after (red) bias correction. d, same as (c), but for the OWZ tracker. (UZ: UZ tracker, OWZ: OWZ tracker. UZ-C and OWZ-C represent bias-corrected results for the UZ and OWZ trackers, respectively.)

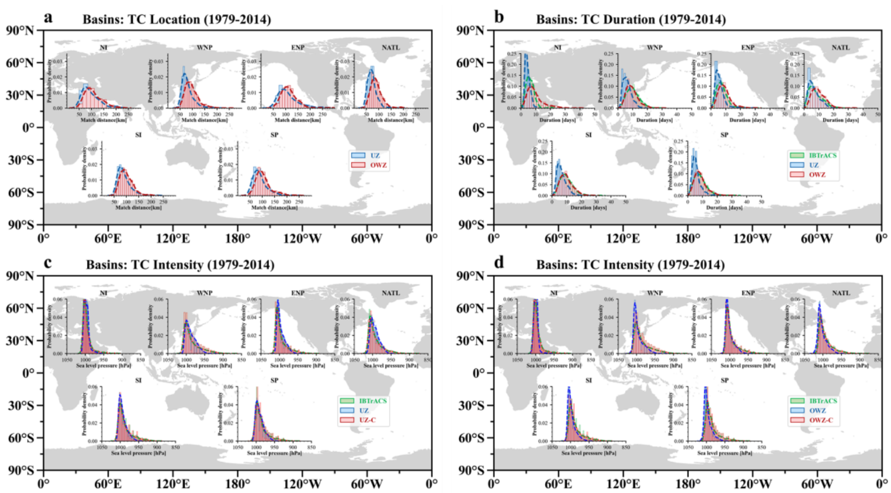

Figure 6: As in Fig. 5, but for six individual basins.

The duration and intensity of TCs are crucial in climate change research, as global warming may lead to stronger and longer-lasting TCs (Knutson et al., 2010). However, observational limitations make these findings more controversial compared to those on TC frequency (Knutson et al., 2010). The

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

RGTracks-20C provides additional support in resolving this controversy. Based on the IBTrACS, the majority of observed TCs globally last fewer than 20 days, with a peak around 8 days (Fig. 5b). Evaluation results (Fig. 5b and Supplementary Fig. S4) show that TCs detected by the OWZ tracker exhibit durations that are close to the observations, and accurately reproduce the TC duration distribution with a peak of 8 days. However, bias is found in the durations of those detected by the UZ tracker, which exhibits a duration peak of approximately 5 days. This is mainly due to the dynamics-based OWZ tracker having the ability to detect storms early in their development (Bell et al., 2018; Bourdin et al., 2022) (Supplementary Fig. S4), while the UZ tracker easily misses weak and short storms (Supplementary Figs. S1a, c) from the 20CRv3 (Bourdin et al., 2022; Tory et al., 2013; Zarzycki and Ullrich, 2017) (Supplementary Sect. S2.3). Similar results are obtained in different basins (Fig. 6b), thus, it is recommended to use the OWZ output when analyzing the durations of TCs.

For TC intensity, given the relatively considerable uncertainty in $WIND_{max}$ compared to SLP_{min} in both reanalyses and IBTrACS (see Methods) (Bourdin et al., 2022; Chavas et al., 2017; Knapp et al., 2010; Knutson et al., 2015; Schreck et al., 2014), this study exclusively utilizes SLP_{min} to evaluate the capability of RGTracks-20C in representing the intensity of TCs. According to IBTrACS (Figs. 5c–d), the intensity of TCs is mainly distributed between 900 and 1020 hPa, peaking around 1000 hPa, with a long tail on the lower SLP_{min} side. In contrast, the SLP_{min} in RGTracks-20C is mainly distributed in the range of 950 – 1020 hPa, with peaks at 1000 hPa and 1005 hPa for the UZ (Fig. 5c) and OWZ (Fig. 5d) trackers, respectively. This suggests that the 20CRv3 generally underestimates the TC intensity (Supplementary Fig. 2a), which, as expected, is primarily because the relatively low spatial resolution of the reanalysis may cause smoothing effects on the sea level pressure field. Apart from spatial resolution, the model's dependence on parameterization processes, along with other factors, may also influence its ability to reproduce TC intensity in the reanalysis (Aarons et al., 2021; Hodges et al., 2017; Malakar et al., 2020).

To address this issue, an intensity bias correction was implemented using quantile mapping bias correction (see Methods) (Zhao and Held, 2010). After intensity correction, the TC intensity distribution in RGTracks-20C is more consistent with IBTrACS (Figs. 5c–d, and Supplementary Fig. 2b), especially in terms of peak positions, and accurately reproduces the skewed distribution of TC intensity. In particular, the RGTracks-20C reproduces TC intensity values with SLP_{min} below 940 hPa, which were not found before the intensity bias correction. This consistency is observed not only on a global scale but also across various basins (Figs. 6c–d).

3.4 Long-term variability of TC activity

This section evaluates the long-term variability of TC activity in the RGTracks-20C by comparing it with the IBTrACS from 1979 to 2014.

Firstly, the RGTracks-20C is able to capture the observed interannual variability of global TC number (Fig. 7a), as indicated by the significant correlations between the TC counts derived from the UZ and OWZ trackers and observations, with correlation coefficients of 0.65 and 0.68 (in the following context, all correlations are significant at the 99% confidence level unless otherwise specified), respectively. This is also true for individual basins (Figs. 8a, d), with the correlation coefficients exceeding 0.70 in most basins. Among the six basins, the highest correlation is observed in the NATL,

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

where the correlation coefficient for the OWZ tracker reaches 0.88 (0.79 for the UZ tracker). Subsequent regions with notable correlations include the WNP (0.75 for OWZ tracker, 0.79 for UZ tracker), SP (0.79 for OWZ tracker, 0.84 for UZ tracker), and SI (0.74 for OWZ tracker, 0.69 for UZ tracker). However, the correlation coefficients are relatively lower in the ENP and NI (Supplementary Table S2), of which the reasons are discussed in Supplementary Sect. S2.2. Notably, the long-term trends in the number of TCs recorded by the two datasets are consistent globally and across most of the ocean basins (Supplementary Table S4).

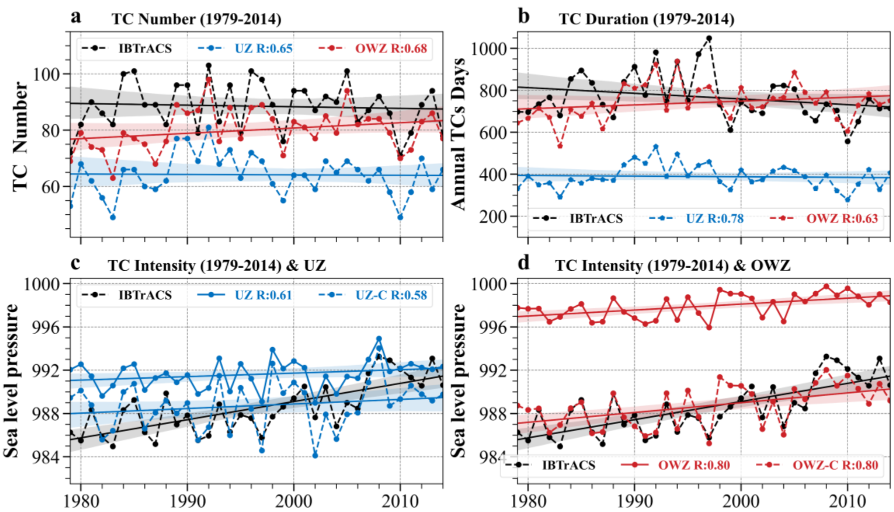

Figure 7: Time series of globally TC activities from IBTrACS and RGTracks-20C during the periods 1979-2014. TC activities are from the IBTrACS and RGTracks-20C using UZ (blue), and OWZ (red) trackers. a, TC number. b, TC days (unit: days). c, TC intensity in *SLPmin* (unit: hPa) in IBTrACS (black) and RGTracks-20C using UZ tracker before (blue solid line) and after (blue dotted line) bias correction. d, same as (c), except for TC intensity in *SLPmin* (unit: hPa) in IBTrACS (black) and RGTracks-20C using OWZ tracker before (red solid line) and after (red dotted line) bias correction. Shaded areas are the two-sided interval of the linear trend at the 95% confidence level. Straight lines are the linear regression. The correlation coefficients (R) between from IBTrACS and RGTracks-20C are marked in the figure legends. All correlation coefficients are statistically significant at the 99% confidence level.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

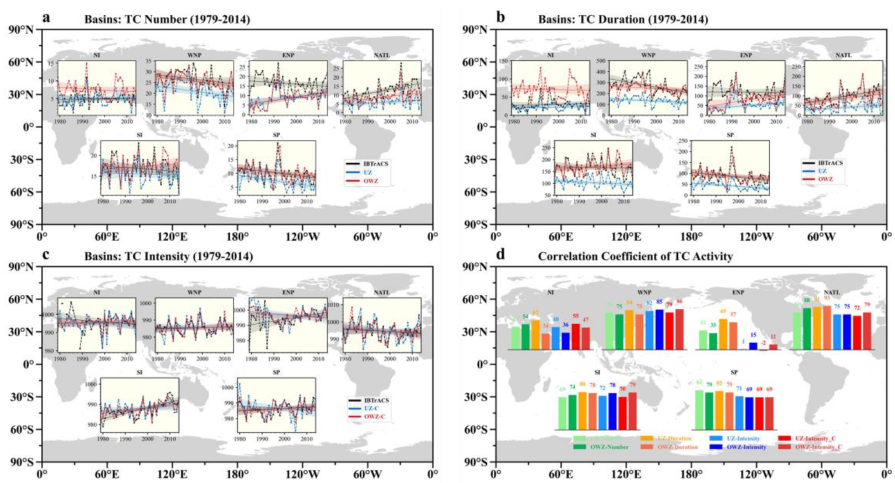

Figure 8: As in Fig. 7, but for six basins. a, TC number. b, TC days (unit: days). c, TC intensity in *SLPmin* (unit: hPa) in IBTrACS (black) and RGTracks-20C (after bias correction) using UZ (blue) and OWZ (red) trackers. d, the correlation coefficients (R) between the from IBTrACS and RGTracks-20C. Note*: The R values for TC number and TC intensity are not statistically significant at the 99% confidence level in the NI and ENP. For TC days, the R value is not statistically significant only in the NI. The R values need to be divided by 100.

TC days, an important metric, encompasses both TC frequency and lifespan (Bell et al., 2018). The RGTracks-20C is able to reproduce the interannual variability of TC days, which is consistent with that in IBTrACS (Fig. 7b), with high correlation coefficients of 0.78 and 0.63 for the UZ and OWZ trackers, respectively. Moreover, these results are further confirmed across basins (Fig. 8b), with correlation coefficients generally exceeding 0.75. In particular, in the NATL, the correlation coefficient exceeds 0.90 (UZ tracker: 0.93, OWZ tracker: 0.91), followed by the SP (UZ tracker: 0.82, OWZ tracker: 0.79), the SI (UZ tracker: 0.80, OWZ tracker: 0.78) and the WNP (UZ tracker: 0.84, OWZ tracker: 0.75). However, being influenced by the observation biases, the correlation coefficients for TC days are also relatively low in the ENP and NI (see Supplementary Table S2). Nevertheless, the above results indicate that the RGTracks-20C provides a satisfactory representation of the interannual and long-term variability (Supplementary Sect. S2.4, Table S4) of the TC days globally and across most of the ocean basins.

In addition, the global TC intensity series based on RGTracks-20C significantly correlates with that based on IBTrACS, with correlation coefficients of 0.61 and 0.80 for the UZ (Fig. 7c) and OWZ (Fig. 7d) trackers, respectively. This indicates that the TC intensity (SLP_{min}) in RGTracks-20C effectively captures the observed interannual variability. Most basins further validate these results (Fig. 8d). The highest correlation coefficients are observed in the WNP, exceeding 0.80 (UZ tracker: 0.82, OWZ tracker: 0.85). Following closely are NATL (UZ tracker: 0.75, OWZ tracker: 0.75) and SI (UZ tracker: 0.72, OWZ tracker: 0.78), while SP (UZ tracker: 0.71, OWZ tracker: 0.69) also demonstrates correlation coefficients of around 0.70.

The 20CRv3 tends to underestimate the TC intensities, due to its coarse resolution, which suggests the need of a bias correction process during the production of the RGTracks-20C (see Methods). By

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

performing intensity bias corrections to the detected TCs, the TC intensity (SLP_{min}) in RGTracks-20C exhibits interannual and long-term variations that are more consistent with the observations (Figs. 7c–d, and Supplementary Fig. 2 and Tables S2, S4), especially in the WNP, NATL, and SI basins (Figs. 8c–d). These results indicate that the RGTracks-20C can reasonably capture the interannual variability and trends (Supplementary Sect. S2.4 and Table S4) of TC intensity globally and across most basins. Discrepancies in the interannual variability of TC intensity between the RGTracks-20C and IBTrACS are also noted over ENP and NI, similar to the above findings on TC number and days (Supplementary Sect. S2.2 and Tables S6–S7).

3.5 Key strengths of the RGTracks-20C

The above evaluation analyses confirm that the RGTracks-20C effectively captures both the climatology and long-term variability of TC activity across global and major oceanic basins. In this section, we discuss the key strengths of the RGTracks-20C, specifically its capacity to reconstruct track and intensity information of early-year TCs that may not be included in the observed data records. Such an advantage of the RGTracks-20C could benefit research about how climate change has affected TCs over the past century.

Before digging into early-year TCs, we first demonstrate the RGTracks-20C's accuracy in reproducing specific TCs by making comparisons with observations. Three representative TCs that caused significant human casualties and economic losses in the NATL, SI, and WNP are analyzed here: Hurricane 'Andrew' in 1992 (Pimm et al., 1994) (Figs. 9a–c), TC 'Geralda' in 1994 (Hoarau et al., 2012) (Figs. 9d–f), and Super Typhoon 'Rammasun' in 2014 (Zhang et al., 2017) (Figs. 9h–i). Compared with IBTrACS, the RGTracks-20C performs exceptionally well in representing the track and duration of these TCs. However, some discrepancies were observed during landfall (Fig. 9a), possibly due to complex topography and TC size, which were not captured by the low-resolution 20CRv3. While the 20CRv3 tends to underestimate the intensity of TCs, the corrected intensity in the RGTracks-20C is highly consistent with observations and accurately captures the temporal evolutions of TC intensities. This evidence confirms RGTracks-20C's ability to capture not only the climatology and variability of TC activity, but also the detailed information on specific TC events.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Open Access Earth System Science Data Discussions

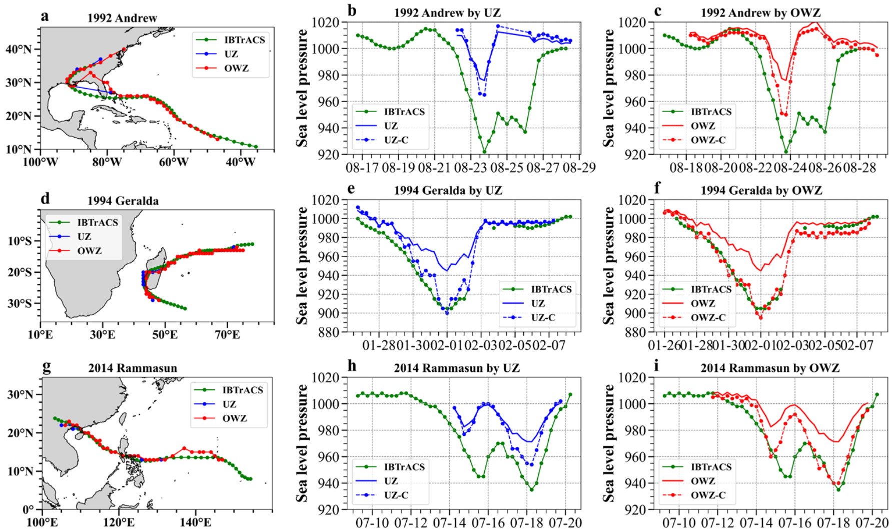

Figure 9: The historical tracks and intensity records of individual tropical cyclones in the IBTrACS and RGTracks-20C. a–c, Track (a) and intensity (*SLPmin*, unit: hPa. b: UZ tracker, c: OWZ tracker) of Hurricane "Andrew". d–f, same as a–c, but for track (d) and intensity (*SLPmin*, unit: hPa. e: UZ tracker, f: OWZ tracker) of tropical cyclone "Geralda". g–i, same as a–c, track (g) and intensity (*SLPmin*, unit: hPa. h: UZ tracker, i: OWZ tracker) of Super typhoon "Rammasun". Green, blue, and red lines denote results based on the IBTrACS, UZ tracker, and OWZ tracker, respectively. The UZ-C (blue dotted dashed line) and OWZ-C (red dotted dashed line) indicate after intensity bias correction.

Prior to the satellite era, limitations in observation systems often led to incomplete records of early TCs, particularly for TC intensity. An example is hurricane Okeechobee in 1928, which was one of the deadliest to hit the United States in the early 20th century. Hurricane Okeechobee was recorded in the IBTrACS (Blake et al., 2011; Mitchell, 1928) (Supplementary Sect. S3.1). However, during Okeechobee's lifetime, there were only 16 time points of the TC intensity that were recorded when it passed the Lesser Antilles and Puerto Rico, and made landfall in the United States (Figs. 10a–c, Supplementary Fig. S5 and Table S8). Similar missing data are common in the IBTrACS records of early TCs, especially when the TCs were located over the ocean (Figs. 10d–f). Moreover, the problem of missing TC intensity records is especially evident in other basins (Supplementary Table S3). For instance, Typhoon No. 8, which made landfall and caused serious damage in Japan (see Supplementary Sect. S3.2), has only track records in the IBTrACS, but with intensity information missing (Figs. 10g–i). In such cases, taking advantage of the 20CRv3, the RGTracks-20C addresses these deficiencies by filling in these gaps, substantially enhancing the completeness of early TC intensity records.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Open Access Earth System Science Data Discussions

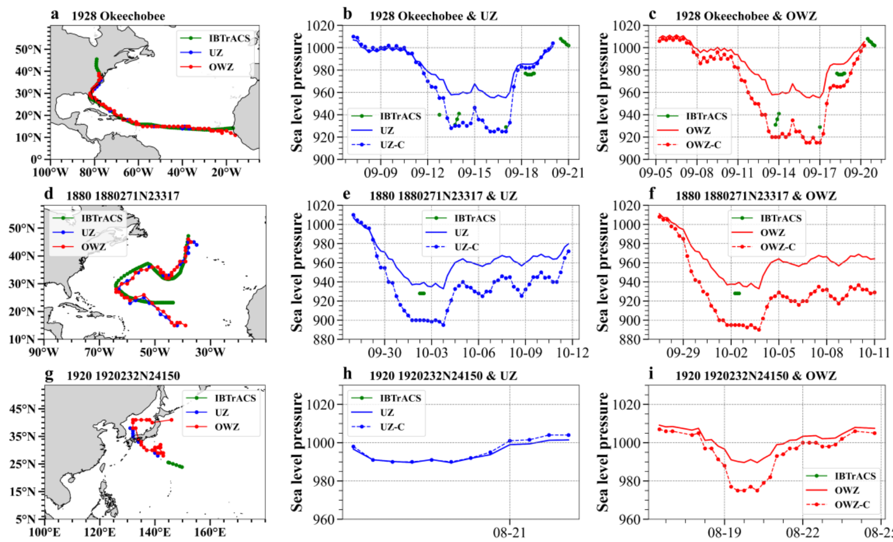

Figure 10: As in Fig. 9, but for Hurricane "Okeechobee" (a–c), Hurricane '1880271N23317' (d–f), typhoon '192023N24150'(g–i).

In addition, not only is the TC intensity missing, but the track records in the IBTrACS may also be incomplete, such as the above-mentioned Typhoon No.8 in 1920 (Fig. 10g), despite the existence of historical observation records (see Supplementary Sect. S3.2). In this case, the RGTracks-20C not only provides the missing TC intensity but also fills gaps in IBTrACS during the latter stages of the typhoon's development, especially during the landfall phase (Fig. 10g and Supplementary Figs. S6–8). Moreover, prior to the satellite era, the RGTracks-20C often reports a higher number of TCs than the IBTrACS, particularly from the early to mid-20th century (Supplementary Fig. S10), which suggests that the RGTracks-20C is also able to detect historical TCs not being recorded in the IBTrACS. These findings demonstrate that the RGTracks-20C can compensate for the incomplete TC track records in the IBTrACS, especially for those in the pre-satellite era.

To evaluate the accuracy of early TC records provided by RGTracks-20C, we take the 1928 Okeechobee hurricane as a case study. The RGTracks-20C nearly fully reproduces the hurricane's lifespans as recorded in IBTrACS, with the OWZ tracker performing exceptionally well, differing by only one day from the IBTrACS record. Okeechobee's latitude and longitude variations in the RGTracks-20C are highly consistent with those in the IBTrACS, with a positional bias within ±1 degree (Fig. 10a and Supplementary Fig. S5). By comparing Okeechobee's intensity in RGTracks-20C with observational data, we find that the RGTracks-20C reliably reproduces Okeechobee's intensity and its variations (Figs. 10b–c and Supplementary Table S8). For instance, as the hurricane passed over Guadeloupe, IBTrACS recorded a SLP_{min} of 940 hPa, which is closely matched by RGTracks-20C (UZ tracker: 955 hPa; OWZ tracker: 940 hPa). Moreover, the RGTracks-20C captures the weakening and re-intensification of the hurricane between Puerto Rico and its landfall in Florida, where the IBTrACS lacks intensity records, demonstrating the RGTracks-20C's reliability in representing intensity changes (Supplementary Sect. S3.1).

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

4. Usage notes

In this study, we introduce the RGTracks-20C, a century-long reanalysis-based historical global TC dataset. Statistical evaluations and case studies confirm RGTracks-20C's reliability in capturing the climatology and interannual variability of observed TC activity on both global and regional scales in the modern satellite era. A major key strength of the RGTracks-20C is its ability to fill the missing intensity or location records of observed TCs in early years.

As documented in prior studies, biases are unavoidable when extracting TCs from reanalyses due to the data quality of reanalyses and the limitations of TC trackers. Some usage notes and cautionary remarks are listed in this section to assist readers in understanding or using the RGTracks-20C.

(1) Due to model resolution and parameterization, TC intensity detected directly from the 20CRv3 is underrepresented compared to observations. To address this issue in the RGTracks-20C, we corrected the biases using a simple quantile mapping method, assuming that systematic biases primarily cause the TC intensity errors from 20CRv3. While this is generally true, the quantile mapping correction did not account for other factors that may also affect TC intensity biases. The inherent challenges in modeling weaker TCs in 20CRv3, which are largely attributed to the limitations of resolution and parameterization of subgrid-scale processes in numerical models, often result in lower detection rates for tropical depressions and weaker tropical storms (e.g., Category 1) (Hodges et al., 2017). This can be improved with more advanced correction approaches of TC intensity in the future.

(2) Discrepancies between the RGTracks-20C and IBTrACS should not be solely attributed to errors in RGTracks-20C, as limitations in IBTrACS may also influence the evaluation results. For example, the classification of TC often relies on forecasters' subjective judgment, which affects whether these systems are included in best track datasets (Torn and Snyder, 2012). Additionally, , differences in observation start times and data sources across basins (Supplementary Table S3) can introduce uncertainties in the IBTrACS data (Chan et al., 2022b). For example, the RGTracks-20C shows relatively large discrepancies with observations in the ENP (Supplementary Sect. S2.2), which may be attributed to the biases of IBTrACS prior to 1988. Similar issues exist for the NI basin. When limiting the study periods to 1988–2014 for the ENP and 1990–2014 for the NI, the RGTracks-20C exhibits good consistency with IBTrACS in TC activity trends, and the correlation significantly improves (Supplementary Fig. S3 and Tables S2, S5). These suggest that the reliability of observational data has been changing over time and may serve as a factor affecting the comparison results between the RGTracks-20C and observational records. Detailed analyses on these two basins can be found in Supplementary Sect. S2.2.

(3) Currently, there are no perfect algorithms for tracking TCs from reanalyses. Although the TC trackers employed in the RGTracks-20C (UZ and OWZ) are two widely recognized algorithms, they were built with different properties and have different limitations. The above evaluation analyses show that the OWZ tracker is closer to the observations in terms of TC number and TC days (Bourdin et al., 2022), while the UZ tracker produces tracks with a shorter duration than the observations, which is mainly related to its physically based tracker intensity threshold (Horn et al., 2014). However, the UZ has a lower FAR, suggesting that it has an advantage in recognizing real TCs and is less likely to misclassify other weather systems as TCs. Generally, since the OWZ tracker demonstrates overall higher stability in detecting TCs, it is recommended to primarily utilize the OWZ tracker results in most cases, with the UZ

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

tracker as a supplementary reference for analyses. In addition, in the production of the RGTracks-20C, globally identical thresholds were used in the TC tracking procedure. However, given the differences in structure and behavior of TCs in different basins and the influence of different meteorological systems and topography, the use of a globally identical tracker may affect the accuracies of TC detection in specific regions (Fu et al., 2021; Raavi and Walsh, 2020a, b). This suggests the need for further improvements in the TC tracking approaches.

(4) The assimilation of SLP_{min} from IBTrACS into the 20CRv3 may lead to into the 20CRv3 may lead to another limitation. As discussed in Supplementary Sect. S4, the RGTracks-20C exhibits consistent trends and variations with IBTrACS from 1850 to 2014 (Supplementary Fig. S10). In particular, the growth trends in TC numbers from both datasets during the mid-20th century are almost identical, primarily resulting from the artificial increase in TC detection associated with advancements in observational technologies. considering that RGTracks-20C currently uses the ensemble mean field of 20CRv3 as input data, which inherently attenuates the intensity and features of extreme events and introduces smoothing effects. In addition, RGTracks-20C currently uses the ensemble mean field of 20CRv3 as input data, which further affects this similarity by inherently weakening the intensity and character of extreme events and introducing smoothing effects (Emanuel, 2024). On the other hand, the assimilation of IBTrACS data has, to some extent, also improved 20CRv3's representation of TC intensity and structure, enabling TC tracker to more effectively detect and identify TCs that actually occurred (Slivinski et al., 2019, 2021). For example, the typhoon that made landfall in Japan in 1920 (Fig. 10g). Nevertheless, this limitation implies that the RGTracks-20C fails to capture the realistic number of TCs in early years, and suggests the need to employ individual members for TC detections (Emanuel, 2024).

The above factors will be thoroughly considered and addressed in the future versions of RGTracks-20C to enhance its accuracy and applicability. In the next version of RGTracks-20C, a few improvements will be included: (1) We detect TCs separately from all 80 ensemble members of the 20CRv3, in order to avoid the smoothing effects caused by the ensemble mean of reanalyses (Emanuel, 2024); (2) we will calibrate algorithm thresholds according to TC characteristics in different ocean basins; (3) more TC tracking algorithms will be included to address the uncertainty of the TC track data (Flaounas et al., 2023).

5. Data Availability

The RGTracks-20C is publicly available at https://doi.org/10.5281/zenodo.14411917 (Ye et al., 2024). The Other datasets utilized in this study are available: the IBTrACS at https://www.ncdc.noaa.gov/ibtracs/; and the 20CRv3 at https://portal.nersc.gov/archive/home/projects/incite11/www/ (Slivinski et al., 2019). Historical weather chart of the 1920 typhoon that made landfall in Japan from http://agora.ex.nii.ac.jp/cgi-bin/weather-chart/calendar.pl?year=1920&month=8&lang=en&type=as.

6. Code Availability

Bourdin (2022a) provided the code for the UZ and OWZ algorithms, which are available at https://doi.org/10.5281/zenodo.6424432. TempestExtremes can be downloaded from https://climate.ucdavis.edu/tempestextremes.php, and version 1.5.2 is used for this study.

7. Conclusion

In this study, we introduce the RGTracks-20C, a century-long reanalysis-based historical global TC dataset. Statistical evaluations and case studies confirm its reliability in capturing the climatology and interannual variability of observed TC activity on both global and regional scales. A major key strength of the RGTracks-20C is its ability to fill the missing intensity and location records of observed TCs in early years. This dataset provides a reliable alternative for researchers to study the long-term variability of TC characteristics, which will help us to better understand changes and trends in historical TC activity, as well as their relationship with climate change.

This knowledge is crucial for protecting vulnerable coastal areas and mitigating TC-related risks in the future climate change. As the first version, the RGTracks-20C has limitations, which may arise from the reanalysis assimilation process and the threshold settings in the TC tracker. Future versions will further address these issues, refining the dataset to improve accuracy and broaden applicability.

Competing interests

The authors declare no competing interests.

Author contributions

G.Y.: methodology, formal analysis, data curation, visualization, writing—original draft, writing—review and editing, software;

J.C.H.L.: conceptualization, methodology, formal analysis, writing—original draft, writing—review and editing, funding acquisition;

W.D.: writing—review and editing, supervision, funding acquisition;

J.X., W.L. and W.Q.: writing—review and editing;

B.Z.: conceptualization, supervision, methodology, formal analysis, writing—review and editing, funding acquisition.

Acknowledgements

The authors sincerely thank Dr. Stella Bourdin from the Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre Simon Laplace (LSCE-IPSL), Gif-sur-Yvette, for her invaluable assistance and guidance on the TC trackers. The authors are very grateful to Dr. Jennifer Gahtan from NOAA's National Center for Environmental Information for providing information about the starting years of the minimum central pressure in the IBTrACS.

This study is primarily supported by the National Natural Science Foundation of China (Grant No. U21A6001). G.Y. and W.D. are also supported by the Southern Marine Science and Engineering Guangdong Laboratory (No. SML2023SP208). G.Y., J.C.H.L., J.X., W.L., W.Q., and B.Z. are supported by the Guangdong Province Introduction of Innovative R&D Team Project China (2019ZT08G669). J.X. is supported by the National Natural Science Foundation of China (42130605). J.C.H.L. is supported by the National Natural Science Foundation of China (42405038) and the Guangdong Basic and Applied Basic Research Foundation (2020A1515110275).

Author contributions

G.Y.: methodology, formal analysis, data curation, visualization, writing—original draft, writing—review and editing, software;

J.C.H.L.: conceptualization, methodology, formal analysis, writing—original draft, writing—review and editing, funding acquisition;

W.D.: writing—review and editing, supervision, funding acquisition;

J.X., W.L. and W.Q.: writing—review and editing;

B.Z.: conceptualization, supervision, methodology, formal analysis, writing—review and editing, funding acquisition.

References

Aarons, Z. S., Camargo, S. J., Strong, J. D. O., and Murakami, H.: Tropical Cyclone Characteristics in the MERRA-2 Reanalysis and AMIP Simulations, Earth and Space Science, 8, e2020EA001415, https://doi.org/10.1029/2020EA001415, 2021.

Bell, S. S., Chand, S. S., Tory, K. J., and Turville, C.: Statistical Assessment of the OWZ Tropical Cyclone Tracking Scheme in ERA-Interim, Journal of Climate, 31, 2217–2232, https://doi.org/10.1175/JCLI-D-17-0548.1, 2018.

Bhatia, K. T., Vecchi, G. A., Knutson, T. R., Murakami, H., Kossin, J., Dixon, K. W., and Whitlock, C. E.: Recent increases in tropical cyclone intensification rates, Nat Commun, 10, 635, https://doi.org/10.1038/s41467-019-08471-z, 2019.

Blake, E. S., Landsea, C., and Gibney, E. J.: The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2010 (and other frequently requested hurricane facts), 2011.

Bloemendaal, N., de Moel, H., Martinez, A. B., Muis, S., Haigh, I. D., van der Wiel, K., Haarsma, R. J., Ward, P. J., Roberts, M. J., Dullaart, J. C. M., and Aerts, J. C. J. H.: A globally consistent local-scale assessment of future tropical cyclone risk, Science Advances, 8, eabm8438, https://doi.org/10.1126/sciadv.abm8438, 2022.

Bourdin, S., Fromang, S., Dulac, W., Cattiaux, J., and Chauvin, F.: Intercomparison of four algorithms for detecting tropical cyclones using ERA5, Geoscientific Model Development, 15, 6759–6786, https://doi.org/10.5194/gmd-15-6759-2022, 2022.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Chan, J. C. L.: Frequency and Intensity of Landfalling Tropical Cyclones in East Asia: Past Variations and Future Projections, Meteorology, 2, 171–190, https://doi.org/10.3390/meteorology2020012, 2023.

Chan, K. T. F.: Are global tropical cyclones moving slower in a warming climate? Environ. Res. Lett., 14, 104015, https://doi.org/10.1088/1748-9326/ab4031, 2019.

Chan, K. T. F., Zhang, K., Wu, Y., and Chan, J. C. L.: Publisher Correction: Landfalling hurricane track modes and decay, Nature, 608, E14–E14, https://doi.org/10.1038/s41586-022-05078-1, 2022a.

Chan, K. T. F., Chan, J. C. L., Zhang, K., and Wu, Y.: Uncertainties in tropical cyclone landfall decay, npj Clim Atmos Sci, 5, 1–8, https://doi.org/10.1038/s41612-022-00320-z, 2022b.

Chand, S. S., Walsh, K. J. E., Camargo, S. J., Kossin, J. P., Tory, K. J., Wehner, M. F., Chan, J. C. L., Klotzbach, P. J., Dowdy, A. J., Bell, S. S., Ramsay, H. A., and Murakami, H.: Declining tropical cyclone frequency under global warming, Nat. Clim. Chang., 12, 655–661, https://doi.org/10.1038/s41558-022-01388-4, 2022.

Chang, E. K. M. and Guo, Y.: Is the number of North Atlantic tropical cyclones significantly underestimated prior to the availability of satellite observations?, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL030169, 2007.

Chavas, D. R., Reed, K. A., and Knaff, J. A.: Physical understanding of the tropical cyclone wind-pressure relationship, Nat Commun, 8, 1360, https://doi.org/10.1038/s41467-017-01546-9, 2017.

Cid, A., Camus, P., Castanedo, S., Méndez, F. J., and Medina, R.: Global reconstructed daily surge levels from the 20th Century Reanalysis (1871–2010), Global and Planetary Change, 148, 9–21, https://doi.org/10.1016/j.gloplacha.2016.11.006, 2017.

Compo, G., Slivinski, L., Whitaker, J., Sardeshmukh, P., McColl, C., Brohan, P., Allan, R., Yin, X., Vose, R., Spencer, L., and others: The international surface pressure databank version 4, 2019.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, Quarterly Journal of the Royal Meteorological Society, 137, 1–28, https://doi.org/10.1002/qj.776, 2011.

Cram, T. A., Compo, G. P., Yin, X., Allan, R. J., McColl, C., Vose, R. S., Whitaker, J. S., Matsui, N., Ashcroft, L., Auchmann, R., and others: The international surface pressure databank

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

version 2, Geoscience Data Journal, 2, 31–46, 2015.

Dinan, T.: Projected Increases in Hurricane Damage in the United States: The Role of Climate Change and Coastal Development, Ecological Economics, 138, 186–198, https://doi.org/10.1016/j.ecolecon.2017.03.034, 2017.

Emanuel, K.: The Hurricane—Climate Connection, Bulletin of the American Meteorological Society, 89, ES10–ES20, https://doi.org/10.1175/BAMS-89-5-Emanuel, 2008.

Emanuel, K.: Tropical Cyclone Activity Downscaled from NOAA-CIRES Reanalysis, 1908–1958, Journal of Advances in Modeling Earth Systems, 2, https://doi.org/10.3894/JAMES.2010.2.1, 2010.

Emanuel, K.: Will Global Warming Make Hurricane Forecasting More Difficult? Bulletin of the American Meteorological Society, 98, 495–501, https://doi.org/10.1175/BAMS-D-16-0134.1, 2017.

Emanuel, K.: 100 Years of Progress in Tropical Cyclone Research, Meteorological Monographs, 59, 15.1-15.68, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0016.1, 2018.

Emanuel, K.: Atlantic tropical cyclones downscaled from climate reanalyses show increasing activity over past 150 years, Nat Commun, 12, 7027, https://doi.org/10.1038/s41467-021-27364-8, 2021.

Emanuel, K.: Limitations of reanalyses for detecting tropical cyclone trends, Nat. Clim. Chang., 14, 143–145, https://doi.org/10.1038/s41558-023-01879-y, 2024.

Faranda, D., Messori, G., Bourdin, S., Vrac, M., Thao, S., Riboldi, J., Fromang, S., and Yiou, P.: Correcting biases in tropical cyclone intensities in low-resolution datasets using dynamical systems metrics, Clim Dyn, https://doi.org/10.1007/s00382-023-06794-8, 2023.

Flaounas, E., Aragão, L., Bernini, L., Dafis, S., Doiteau, B., Flocas, H., Gray, S. L., Karwat, A., Kouroutzoglou, J., Lionello, P., Miglietta, M. M., Pantillon, F., Pasquero, C., Patlakas, P., Picornell, M. Á., Porcù, F., Priestley, M. D. K., Reale, M., Roberts, M. J., Saaroni, H., Sandler, D., Scoccimarro, E., Sprenger, M., and Ziv, B.: A composite approach to produce reference datasets for extratropical cyclone tracks: application to Mediterranean cyclones, Weather and Climate Dynamics, 4, 639–661, https://doi.org/10.5194/wcd-4-639-2023, 2023.

Fu, D., Chang, P., Patricola, C. M., Saravanan, R., Liu, X., and Beck, H. E.: Central American mountains inhibit eastern North Pacific seasonal tropical cyclone activity, Nat Commun, 12, 4422, https://doi.org/10.1038/s41467-021-24657-w, 2021.

Gergis, J., Ashcroft, L., and Whetton, P.: A historical perspective on Australian temperature

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Earth System Science Data Discussions
Open Access

extremes, Clim Dyn, 55, 843–868, https://doi.org/10.1007/s00382-020-05298-z, 2020.

Hassanzadeh, P., Lee, C.-Y., Nabizadeh, E., Camargo, S. J., Ma, D., and Yeung, L. Y.: Effects of climate change on the movement of future landfalling Texas tropical cyclones, Nat Commun, 11, 3319, https://doi.org/10.1038/s41467-020-17130-7, 2020.

Hoarau, K., Bernard, J., and Chalonge, L.: Intense tropical cyclone activities in the northern Indian Ocean, International Journal of Climatology, 32, 1935–1945, https://doi.org/10.1002/joc.2406, 2012.

Hodges, K., Cobb, A., and Vidale, P. L.: How Well Are Tropical Cyclones Represented in Reanalysis Datasets? Journal of Climate, 30, 5243–5264, https://doi.org/10.1175/JCLI-D-16-0557.1, 2017.

Horn, M., Walsh, K., Zhao, M., Camargo, S. J., Scoccimarro, E., Murakami, H., Wang, H., Ballinger, A., Kumar, A., Shaevitz, D. A., Jonas, J. A., and Oouchi, K.: Tracking Scheme Dependence of Simulated Tropical Cyclone Response to Idealized Climate Simulations, https://doi.org/10.1175/JCLI-D-14-00200.1, 2014.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin of the American Meteorological Society, 77, 437–472, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Klotzbach, P. J. and Landsea, C. W.: Extremely Intense Hurricanes: Revisiting Webster et al. (2005) after 10 Years, Journal of Climate, 28, 7621–7629, https://doi.org/10.1175/JCLI-D-15-0188.1, 2015.

Klotzbach, P. J., Bell, M. M., Bowen, S. G., Gibney, E. J., Knapp, K. R., and Schreck, C. J.: Surface Pressure a More Skillful Predictor of Normalized Hurricane Damage than Maximum Sustained Wind, https://doi.org/10.1175/BAMS-D-19-0062.1, 2020.

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., and Neumann, C. J.: The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying Tropical Cyclone Data, Bulletin of the American Meteorological Society, 91, 363–376, https://doi.org/10.1175/2009BAMS2755.1, 2010.

Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., Schreck, C. J., and others: International best track archive for climate stewardship (IBTrACS) project, version 4, NOAA National Centers for Environmental Information, 10, 2018.

Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C.-H., Kossin, J., Mohapatra, M., Satoh, M., Sugi, M., Walsh, K., and Wu, L.: Tropical Cyclones and Climate Change

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Assessment: Part I: Detection and Attribution, Bulletin of the American Meteorological Society, 100, 1987–2007, https://doi.org/10.1175/BAMS-D-18-0189.1, 2019.

Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C.-H., Kossin, J., Mohapatra, M., Satoh, M., Sugi, M., Walsh, K., and Wu, L.: Tropical Cyclones and Climate Change Assessment: Part II: Projected Response to Anthropogenic Warming, Bulletin of the American Meteorological Society, 101, E303–E322, https://doi.org/10.1175/BAMS-D-18-0194.1, 2020.

Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., Holland, G., Landsea, C., Held, I., Kossin, J. P., Srivastava, A. K., and Sugi, M.: Tropical cyclones and climate change, Nature Geosci, 3, 157–163, https://doi.org/10.1038/ngeo779, 2010.

Knutson, T. R., Sirutis, J. J., Zhao, M., Tuleya, R. E., Bender, M., Vecchi, G. A., Villarini, G., and Chavas, D.: Global Projections of Intense Tropical Cyclone Activity for the Late Twenty-First Century from Dynamical Downscaling of CMIP5/RCP4.5 Scenarios, Journal of Climate, 28, 7203–7224, https://doi.org/10.1175/JCLI-D-15-0129.1, 2015.

Kossin, J. P., Knapp, K. R., Olander, T. L., and Velden, C. S.: Global increase in major tropical cyclone exceedance probability over the past four decades, Proceedings of the National Academy of Sciences, 117, 11975–11980, 2020.

Kunze, S.: Unraveling the Effects of Tropical Cyclones on Economic Sectors Worldwide: Direct and Indirect Impacts, Environmental and Resource Economics, https://doi.org/10.1007/s10640-021-00541-5, 2021.

Lai, Y., Li, J., Gu, X., Chen, Y. D., Kong, D., Gan, T. Y., Liu, M., Li, Q., and Wu, G.: Greater flood risks in response to slowdown of tropical cyclones over the coast of China, Proceedings of the National Academy of Sciences, 117, 14751–14755, https://doi.org/10.1073/pnas.1918987117, 2020.

Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J.-R., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H., Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier, E., and Schepers, D.: CERA-20C: A Coupled Reanalysis of the Twentieth Century, Journal of Advances in Modeling Earth Systems, 10, 1172–1195, https://doi.org/10.1029/2018MS001273, 2018.

Landsea, C.: Counting Atlantic tropical cyclones back to 1900, Eos, Transactions American Geophysical Union, 88, 197–202, https://doi.org/10.1029/2007EO180001, 2007.

Landsea, C. W., Harper, B. A., Hoarau, K., and Knaff, J. A.: Can We Detect Trends in Extreme Tropical Cyclones? Science, 313, 452–454, https://doi.org/10.1126/science.1128448, 2006.

Landsea, C. W., Glenn, D. A., Bredemeyer, W., Chenoweth, M., Ellis, R., Gamache, J.,

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Hufstetler, L., Mock, C., Perez, R., Prieto, R., Sánchez-Sesma, J., Thomas, D., and Woolcock, L.: A Reanalysis of the 1911–20 Atlantic Hurricane Database, Journal of Climate, 21, 2138–2168, https://doi.org/10.1175/2007JCLI1119.1, 2008.

Landsea, C. W., Vecchi, G. A., Bengtsson, L., and Knutson, T. R.: Impact of Duration Thresholds on Atlantic Tropical Cyclone Counts, Journal of Climate, 23, 2508–2519, https://doi.org/10.1175/2009JCLI3034.1, 2010.

Lanzante, J. R.: Uncertainties in tropical-cyclone translation speed, Nature, 570, E6–E15, https://doi.org/10.1038/s41586-019-1223-2, 2019.

Lee, R., Chen, L., and Ren, G.: A comparison of East-Asia landfall tropical cyclone in recent reanalysis datasets--before and after satellite era, Frontiers in Earth Science, 10, 2023.

Lee, T.-C., Knutson, T. R., Nakaegawa, T., Ying, M., and Cha, E. J.: Third assessment on impacts of climate change on tropical cyclones in the Typhoon Committee Region – Part I: Observed changes, detection and attribution, Tropical Cyclone Research and Review, 9, 1–22, https://doi.org/10.1016/j.tcrr.2020.03.001, 2020.

Lenzen, M., Malik, A., Kenway, S., Daniels, P., Lam, K. L., and Geschke, A.: Economic damage and spillovers from a tropical cyclone, Natural Hazards and Earth System Sciences, 19, 137–151, https://doi.org/10.5194/nhess-19-137-2019, 2019.

Leung, J. C.H., Qian, W., Zhang, P., and Zhang, B.: Geopotential-based Multivariate MJO Index: extending RMM-like indices to pre-satellite era, Clim Dyn, 59, 609–631, https://doi.org/10.1007/s00382-022-06142-2, 2022.

Li, J., Tian, Q., Shen, Z., Xu, Y., Yan, Z., Li, M., Zhu, C., Xue, J., Lin, Z., Yang, Y., and Zeng, L.: Fidelity of global tropical cyclone activity in a new reanalysis dataset (CRA40), Meteorological Applications, 31, e70009, https://doi.org/10.1002/met.70009, 2024.

Li, Y., Tang, Y., Li, X., Song, X., and Wang, Q.: Recent increase in the potential threat of western North Pacific tropical cyclones, npj Clim Atmos Sci, 6, 1–8, https://doi.org/10.1038/s41612-023-00379-2, 2023.

Malakar, P., Kesarkar, A. p., Bhate, J. n., Singh, V., and Deshamukhya, A.: Comparison of Reanalysis Data Sets to Comprehend the Evolution of Tropical Cyclones Over North Indian Ocean, Earth and Space Science, 7, e2019EA000978, https://doi.org/10.1029/2019EA000978, 2020.

Mann, M. E., Sabbatelli, T. A., and Neu, U.: Evidence for a modest undercount bias in early historical Atlantic tropical cyclone counts, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL031781, 2007.

Mitchell, C. L.: The West Indian hurricane of September 10–20, 1928, Monthly Weather Review, 56, 347–350, 1928.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Moon, I.-J., Kim, S.-H., and Chan, J. C. L.: Climate change and tropical cyclone trend, Nature, 570, E3–E5, https://doi.org/10.1038/s41586-019-1222-3, 2019.

Moon, M., Ha, K.-J., Kim, D., Ho, C.-H., Park, D.-S. R., Chu, J.-E., Lee, S.-S., and Chan, J. C. L.: Rainfall strength and area from landfalling tropical cyclones over the North Indian and western North Pacific oceans under increased CO2 conditions, Weather and Climate Extremes, 41, 100581, https://doi.org/10.1016/j.wace.2023.100581, 2023.

Moore, G. W. K. and Babij, M.: Iceland's Great Frost Winter of 1917/1918 and its representation in reanalyses of the twentieth century, Quarterly Journal of the Royal Meteorological Society, 143, 508–520, https://doi.org/10.1002/qj.2939, 2017.

Murakami, H.: Tropical cyclones in reanalysis data sets, Geophysical Research Letters, 41, 2133–2141, https://doi.org/10.1002/2014GL059519, 2014.

Murakami, H. and Wang, B.: Patterns and frequency of projected future tropical cyclone genesis are governed by dynamic effects, Commun Earth Environ, 3, 1–10, https://doi.org/10.1038/s43247-022-00410-z, 2022.

Noy, I.: The socio-economics of cyclones, Nature Clim Change, 6, 343–345, https://doi.org/10.1038/nclimate2975, 2016.

Parker, W. S.: Reanalyses and Observations: What's the Difference? Bulletin of the American Meteorological Society, 97, 1565–1572, https://doi.org/10.1175/BAMS-D-14-00226.1, 2016.

Pimm, S. L., Davis, G. E., Loope, L., Roman, C. T., Smith, T. J., and Tilmant, J. T.: Hurricane Andrew, BioScience, 44, 224–229, https://doi.org/10.2307/1312226, 1994.

Qin, L., Zhu, L., Liu, B., Li, Z., Tian, Y., Mitchell, G., Shen, S., Xu, W., and Chen, J.: Global expansion of tropical cyclone precipitation footprint, Nat Commun, 15, 4824, https://doi.org/10.1038/s41467-024-49115-1, 2024.

Raavi, P. H. and Walsh, K. J. E.: Basinwise Statistical Analysis of Factors Limiting Tropical Storm Formation from an Initial Tropical Circulation, Journal of Geophysical Research: Atmospheres, 125, e2019JD032006, https://doi.org/10.1029/2019JD032006, 2020a.

Raavi, P. H. and Walsh, K. j. e.: Sensitivity of Tropical Cyclone Formation to Resolution-Dependent and Independent Tracking Schemes in High-Resolution Climate Model Simulations, Earth and Space Science, 7, e2019EA000906, https://doi.org/10.1029/2019EA000906, 2020b.

Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vannière, B., Mecking, J., Haarsma, R., Bellucci, A., Scoccimarro, E., Caron, L.-P., Chauvin, F., Terray, L., Valcke, S., Moine, M.-P., Putrasahan, D., Roberts, C. D., Senan, R., Zarzycki, C., Ullrich, P., Yamada, Y., Mizuta, R., Kodama, C., Fu, D., Zhang, Q., Danabasoglu, G., Rosenbloom, N., Wang, H., and Wu, L.: Projected Future Changes in Tropical Cyclones Using the CMIP6 HighResMIP Multimodel Ensemble, https://doi.org/10.1029/2020GL088662, n.d.

Schreck, C. J., Knapp, K. R., and Kossin, J. P.: The Impact of Best Track Discrepancies on Global Tropical Cyclone Climatologies using IBTrACS, Monthly Weather Review, 142, 3881–3899, https://doi.org/10.1175/MWR-D-14-00021.1, 2014.

Sharmila, S. and Walsh, K. J. E.: Recent poleward shift of tropical cyclone formation linked to Hadley cell expansion, Nature Clim Change, 8, 730–736, https://doi.org/10.1038/s41558-018-0227-5, 2018.

Slivinski, L. C.: Historical Reanalysis: What, How, and Why? Journal of Advances in Modeling Earth Systems, 10, 1736–1739, https://doi.org/10.1029/2018MS001434, 2018.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F. L., Lee, T.-C., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. W. K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system, Quarterly Journal of the Royal Meteorological Society, 145, 2876–2908, https://doi.org/10.1002/qj.3598, 2019.

Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D. C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3, Journal of Climate, 34, 1417–1438, https://doi.org/10.1175/JCLI-D-20-0505.1, 2021.

Torn, R. D. and Snyder, C.: Uncertainty of Tropical Cyclone Best-Track Information, Weather and Forecasting, 27, 715–729, https://doi.org/10.1175/WAF-D-11-00085.1, 2012.

Tory, K. J., Dare, R. A., Davidson, N. E., McBride, J. L., and Chand, S. S.: The importance of low-deformation vorticity in tropical cyclone formation, Atmospheric Chemistry and Physics, 13, 2115–2132, https://doi.org/10.5194/acp-13-2115-2013, 2013.

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Truchelut, R. E. and Hart, R. E.: Quantifying the possible existence of undocumented Atlantic warm-core cyclones in NOAA/CIRES 20th Century Reanalysis data, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL046756, 2011.

Truchelut, R. E., Hart, R. E., and Luthman, B.: Global Identification of Previously Undetected Pre-Satellite-Era Tropical Cyclone Candidates in NOAA/CIRES Twentieth-Century Reanalysis Data, Journal of Applied Meteorology and Climatology, 52, 2243–2259, https://doi.org/10.1175/JAMC-D-12-0276.1, 2013.

Tu, S., Xu, J., Chan, J. C. L., Huang, K., Xu, F., and Chiu, L. S.: Recent global decrease in the inner-core rain rate of tropical cyclones, Nat Commun, 12, 1948, https://doi.org/10.1038/s41467-021-22304-y, 2021.

Tu, S., Chan, J. C. L., Xu, J., Zhong, Q., Zhou, W., and Zhang, Y.: Increase in tropical cyclone rain rate with translation speed, Nat Commun, 13, 7325, https://doi.org/10.1038/s41467-022-35113-8, 2022.

Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., and Reed, K. A.: TempestExtremes v2.1: a community framework for feature detection, tracking, and analysis in large datasets, Geoscientific Model Development, 14, 5023–5048, https://doi.org/10.5194/gmd-14-5023-2021, 2021.

Wang, S. and Toumi, R.: More tropical cyclones are striking coasts with major intensities at landfall, Sci Rep, 12, 5236, https://doi.org/10.1038/s41598-022-09287-6, 2022.

Wang, X. L., Feng, Y., and Swail, V.: North Atlantic wave height trends as reconstructed from the 20th century reanalysis, Geophysical Research Letters, 39, 2012.

Yamaguchi, M., Chan, J. C. L., Moon, I.-J., Yoshida, K., and Mizuta, R.: Global warming changes tropical cyclone translation speed, Nat Commun, 11, 47, https://doi.org/10.1038/s41467-019-13902-y, 2020.

Ye, G., Leung, J.C.H., Dong, W., Xu, Ji., LI, W., Qian, W., Kong, H., and Zhang, B.: A Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth Century (RGTracks-20C) (v1.0.2). Zenodo. https://doi.org/10.5281/zenodo.14411917, 2024.

Yeasmin, A., Chand, S., and Sultanova, N.: Reconstruction of tropical cyclone and depression proxies for the South Pacific since the 1850s, Weather and Climate Extremes, 39, 100543, https://doi.org/10.1016/j.wace.2022.100543, 2023.

Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., and Chen, D.: An Overview of the China Meteorological Administration Tropical Cyclone Database, Journal of Atmospheric and Oceanic Technology, 31, 287–301, https://doi.org/10.1175/JTECH-D-12-00119.1, 2014.

Yoshida, K., Sugi, M., Mizuta, R., Murakami, H., and Ishii, M.: Future Changes in Tropical

https://doi.org/10.5194/essd-2025-126
Preprint. Discussion started: 7 April 2025

Cyclone Activity in High-Resolution Large-Ensemble Simulations, Geophysical Research Letters, 44, 9910–9917, https://doi.org/10.1002/2017GL075058, 2017.

Zarzycki, C. M. and Ullrich, P. A.: Assessing sensitivities in algorithmic detection of tropical cyclones in climate data, Geophysical Research Letters, 44, 1141–1149, https://doi.org/10.1002/2016GL071606, 2017.

Zhang, B., Zhang, R., Pinker, R. T., Feng, Y., Nie, C., and Guan, Y.: Changes of tropical cyclone activity in a warming world are sensitive to sea surface temperature environment, Environ. Res. Lett., 14, 124052, https://doi.org/10.1088/1748-9326/ab5ada, 2019.

Zhang, G.: Warming-induced contraction of tropical convection delays and reduces tropical cyclone formation, Nat Commun, 14, 6274, https://doi.org/10.1038/s41467-023-41911-5, 2023.

Zhang, X., Duan, Y., Wang, Y., Wei, N., and Hu, H.: A high-resolution simulation of Supertyphoon Rammasun (2014)—Part I: Model verification and surface energetics analysis, Adv. Atmos. Sci., 34, 757–770, https://doi.org/10.1007/s00376-017-6255-7, 2017.

Zhao, M. and Held, I. M.: An Analysis of the Effect of Global Warming on the Intensity of Atlantic Hurricanes Using a GCM with Statistical Refinement, Journal of Climate, 23, 6382–6393, https://doi.org/10.1175/2010JCLI3837.1, 2010.

Zhu, L. and Quiring, S. M.: Exposure to precipitation from tropical cyclones has increased over the continental United States from 1948 to 2019, Commun Earth Environ, 3, 1–8, https://doi.org/10.1038/s43247-022-00639-8, 2022.