

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

3

4 Guiling Ye^{1,2,3}, Jeremy Cheuk-Hin Leung^{2*}, Wenjie Dong^{1,3*}, Ralf Toumi⁴, Jianjun
5 Xu⁵, Weijing Li⁶, Weihong Qian⁷, Hoiio Kong⁸, Banglin Zhang^{2,9}

6 **Affiliations**

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

9 ² Hunan Institute of Advanced Technology, Changsha, China

10 ³ Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China

11 ⁴ Department of Physics, Imperial College London, London, UK

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

14 ⁶ National Climate Center, China Meteorological Administration, Beijing, China

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

16 ⁸ Faculty of Data Science, City University of Macau, Macau, China

17 ⁹ College of Atmospheric Science, Lanzhou University, Lanzhou, China

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

20 **Abstract**

21 Tropical cyclones (TCs) are among the deadliest disasters affecting human society, reconstructing
22 their historical activity has become increasingly important. However, a comprehensive
23 understanding of historical TC activity is challenging due to incomplete TC records in the early years.
24 Here, we introduce the Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth
25 Century (RGTracks-20C) (Ye et al., 2024), a publicly available century-long global TC track dataset
26 spanning from 1850–2014. The first version of RGTracks-20C is reconstructed from the National
27 Oceanic and Atmospheric Administration Twentieth Century Reanalysis (20CRv3) primarily using
28 the Okubo–Weiss–Zeta (OWZ) tracker. RGTracks-20C provides a reasonable representation of TC
29 climatology and interannual variability across the global ocean and most major basins during the
30 satellite era. Case studies further highlight its value in providing supplementary track and intensity
31 information for earlier periods when the historical record is sparse or incomplete. RGTracks-20C
32 therefore offers a useful resource for augmenting the historical TC record prior to the satellite era.

33 1. Introduction

34 Tropical cyclones (TCs), also known as hurricanes or typhoons, are intense weather systems that form
35 over tropical and subtropical oceans and can cause severe disasters over the coastal regions and even
36 inland areas (Qin et al., 2024; Zhu and Quiring, 2022). Globally, approximately 80 TCs are generated
37 each year (Emanuel, 2018). As one of the most destructive weather systems (Bloemendaal et al., 2022;
38 Dinan, 2017; Emanuel, 2017), TCs significantly impact society and the economy (Kunze, 2021; Lenzen
39 et al., 2019; Noy, 2016). These impacts are expected to be exacerbated by climate change in the future
40 (Chan, 2023; Hassanzadeh et al., 2020; Knutson et al., 2020; Moon et al., 2023; Murakami and Wang,
41 2022; Yamaguchi et al., 2020). Therefore, research on TCs has become increasingly vital in climate
42 change and prediction (Bhatia et al., 2019; Chan, 2019; Lanzante, 2019; Moon et al., 2019; Sharmila and
43 Walsh, 2018; Zhang et al., 2019). However, past TC activity and underlying mechanisms remain
44 challenging due to incomplete early historical TC observation records, which may lead to controversies
45 (Chan et al., 2022a, b; Knutson et al., 2019; Lee et al., 2020).

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

62 Given the limitations of historical TC records, a promising approach is to utilize reanalysis for TC
63 identification (Li et al., 2024; Truchelut and Hart, 2011). Reanalysis combines historical observational
64 data with modern numerical weather models to produce comprehensive, continuous datasets of global
65 atmospheric conditions that adhere to physical principles (Compo et al., 2011; Kalnay et al., 1996; Parker,
66 2016; Slivinski, 2018). The Twentieth Century Reanalysis (20CR) (Compo et al., 2011), provided by the
67 National Oceanic and Atmospheric Administration (NOAA), is a global reanalysis dataset that covers the
68 longest period among all other reanalyses. The 20CR was designed for long-term analyses from
69 individual extreme weather events to climate variability, and has been applied to a wide range of studies,
70 including those on wave height, storm surge, Madden-Julian Oscillations, and TCs (Chand et al., 2022;
71 Cid et al., 2017; Gergis et al., 2020; Lee et al., 2023; Leung et al., 2022; Moore and Babij, 2017; Slivinski
72 et al., 2019; Truchelut et al., 2013; Wang et al., 2012). The fact that the 20CR only assimilates surface

73 pressure and sea level pressure fields, instead of other observations such as satellites and aircraft, makes
74 it less sensitive to the temporal inhomogeneity of observations (Slivinski et al., 2019, 2021). *It therefore*
75 *provides a long-term historical weather dataset with diverse variables and complete spatial and temporal*
76 *coverage.*

77 *Such complete atmospheric fields provide an important basis for examining the large-scale structure and*
78 *evolution of TCs over long historical periods.* Several independent studies have documented the
79 feasibility of reproducing the characteristics of some historical TC events based on the 20CR (Emanuel,
80 2010; Lee et al., 2023; Slivinski et al., 2019; Truchelut et al., 2013; Truchelut and Hart, 2011). For
81 example, following Emanuel (Emanuel, 2010), who first expanded and revised TC climatology based on
82 the 20CR, Truchelut and Hart (2011) employed the 20CR to identify previously unknown TCs in the
83 Atlantic and demonstrated that the 20CR can accurately describe large-scale TC thermodynamic structure.
84 Recently, Truchelut et al. (2013) noted that the 20CR has the ability to investigate TC events that were
85 previously undetected in the pre-satellite era. Compared to other reanalyses, the 20CR well captures the
86 intensity of the 1915 Galveston hurricane (Slivinski et al., 2019) and also offers a more accurate
87 representation of landfalling TCs in East Asia (Lee et al., 2023). These previous studies have
88 demonstrated the effectiveness of the 20CR as a tool for characterizing historical TCs (Emanuel, 2010;
89 Truchelut et al., 2013; Truchelut and Hart, 2011). Taking advantage of the 20CR, some researchers have
90 extracted the century-long TC information from the reanalysis product (Chand et al., 2022; Lee et al.,
91 2023; Yeasmin et al., 2023), *suggesting, despite its limitations, its potential as a tool for supplementing*
92 *incomplete historical TC records and extending our knowledge of TC activity into the pre-satellite era.*

93 While the 20CR has been applied to studying the relationship between historical climate change and TC
94 variability, the primary focus was mostly on the TC occurrence frequency, and little attention was given
95 to other TC metrics such as intensity, duration, and location. More importantly, to date, there is no
96 publicly available reanalysis-based global TC dataset covering a century-long period. *The lack of such a*
97 *dataset limits our ability to examine how TC track, intensity, and lifetime evolved in earlier periods when*
98 *the observational record was sparse or incomplete. In many early historical cases, even when TC tracks*
99 *are available, intensity and position information is often missing or discontinuous, making it difficult to*
100 *investigate the full evolution of individual storms and to assess TC activity beyond frequency alone.*

101 Therefore, the main objective of this study is to extract TC information (including location, intensity, and
102 lifetime) from the 20CR and reconstruct a historical global TC track dataset spanning 1850–2014. The
103 produced dataset is named the Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the
104 Twentieth Century (RGTracks-20C) and is open to the public for research use. This paper first introduces
105 the production details of the RGTracks-20C and then discusses the validity, key strengths, and usage
106 notes of the datasets. *We anticipate that RGTracks-20C will serve as a valuable new resource for*
107 *reconstructing historical TC tracks, augmenting the limited observational record prior to the satellite era,*
108 *and supporting future studies.*

109 **2. Data and methods**

110 **2.1 Data**

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

116 **2.1.1 20th Century Reanalysis**

117 The 20CRv3 is led by NOAA's Physical Sciences Laboratory (PSL) and the Cooperative Institute for
118 Research in Environmental Sciences (CIRES) at the University of Colorado, supported by the U.S.
119 Department of Energy (DOE) (Slivinski et al., 2019, 2021). It, by combining advanced data assimilation
120 and numerical prediction techniques with historical observation data, provides long-term historical
121 weather data with diverse variables, complete spatial and temporal coverage. The 20CRv3 employs sea-
122 surface temperature and sea-ice distributions as its boundary conditions and assimilates only surface
123 pressure reports from the International Surface Pressure Databank (ISPD) version 4.7 (Compo et al., 2019;
124 Cram et al., 2015), which include observations from stations and ships, as well as TC intensity (the
125 minimum central pressure (SLP_{min}) from the IBTrACS (Knapp et al., 2010). As such, it is more consistent
126 and homogeneous with time than other reanalyses (Slivinski et al., 2019).

127 One should note that the IBTrACS and 20CRv3 are not two independent datasets because the SLP_{min}
128 records in the IBTrACS are partly assimilated in the production of 20CRv3. On the other hand, an
129 advantage is that TCs structure and intensity more accurately and closer to observations than other 20th
130 century reanalyses as a result of the assimilation of IBTrACS (Laloyaux et al., 2018; Slivinski et al.,
131 2019). And it provides a four-dimensional global gridded atmospheric dataset that spans the whole 20th
132 century and part of the 19th century (1836–2015, with an experimental extension spanning 1806–35),
133 with a 3-hour temporal resolution and $1^\circ \times 1^\circ$ horizontal resolution (Slivinski et al., 2021). Thus, the
134 20CRv3 was applied to the production of the RGTracks-20C in this paper.

135 **2.1.2 IBTrACS**

136 The IBTrACS (Knapp et al., 2010), published by the NOAA, merges recent and historical TC data from
137 meteorological agencies worldwide. These include the Regional Specialized Meteorological Centers
138 (RSMC) and Tropical Cyclone Warning Centers (TCWC) of the World Meteorological Organization
139 (WMO), as well as non-WMO Centers, such as the China Meteorological Administration, the Hong Kong
140 Observatory and the Joint Typhoon Warning Center. The IBTrACS is the most comprehensive and
141 publicly available global TC best-track dataset. It has been widely applied in previous research to
142 investigate the characteristics of TCs (Lai et al., 2020; Li et al., 2023; Tu et al., 2021, 2022; Wang and
143 Toumi, 2022; Zhang, 2023), and has served as a criterion for assessing TC records derived from
144 reanalysis (Bell et al., 2018; Bourdin et al., 2022; Chand et al., 2022; Hodges et al., 2017; Lee et al.,
145 2023). In this study, the most updated version of IBTrACS (v04) (Knapp et al., 2018) serves as an
146 observation reference for evaluating the reliability of the RGTracks-20C. This dataset was cleaned before

147 being used for analyses. Details about the data pre-processing procedures are referred to in Figure B1 in
148 (Bourdin et al., 2022). In particular, we standardized maximum sustained wind speeds ($WIND_{max}$) in
149 IBTrACS to 10-minute sustained wind speeds to ensure a consistent global standard (Knapp et al., 2010).
150 We then removed tracks that did not reach the tropical storm stage ($WIND_{max} < 16 \text{ m} \cdot \text{s}^{-1}$) and those that
151 lasted shorter than two days.

152 Although the IBTrACS has time coverage dating back to the early 20th century, we utilize the data only
153 for the post-satellite period (1979–2014) due to the early data incompleteness issues (Chang and Guo,
154 2007; Lee et al., 2020; Truchelut et al., 2013). Given that the IBTrACS is the most reliable record of TCs
155 after the 1970s, the IBTrACS serves as the best benchmark for validating the data quality of RGTracks-
156 20C. However, because the starting years of records vary across basins within the IBTrACS, biases may
157 occur in the assessment results (see Section 3.4 for more details).

158 **2.2 Production of the RGTracks-20C**

159 **2.2.1 Procedure**

160 The RGTracks-20C was constructed from the latest version of 20CR (20CRv3). The relatively short and
161 imperfectly sampled observational record of TCs introduces considerable uncertainty in their data over
162 the past century (Landsea, 2007; Landsea et al., 2010), hindering accurate detection and the
163 reconstruction of complete historical TC track data (Knutson et al., 2019; Lee et al., 2020). In this context,
164 reanalysis provides a physically consistent framework that can complement the limited observational
165 record and support the reconstruction of historical TC tracks (Chand et al., 2022; Truchelut et al., 2013).
166 Since TC information is not directly provided in the 20CRv3, objective TC trackers were applied to detect
167 and track TCs in this dataset. A wide range of TC trackers has been developed for gridded datasets,
168 including both physics-based and dynamics-based approaches (Hodges et al., 2017; Horn et al., 2014;
169 Tory et al., 2013; Zarzycki and Ullrich, 2017). In this study, the current first version of RGTracks-20C
170 is based primarily on the dynamics-based Okubo-Weiss-Zeta (OWZ) tracker (Tory et al., 2013), which
171 has been widely applied to coarse-resolution datasets and is often regarded as relatively less sensitive to
172 horizontal resolution. This version is primarily intended to reconstruct historical TC tracks and augment
173 the limited observational record during early periods with sparse observations. We also tested the
174 suitability of the physically-based Ullrich & Zarzycki (UZ) tracker (Zarzycki and Ullrich, 2017). Figure
175 1 illustrates the methodological workflow adopted in this study, and the details are described below.

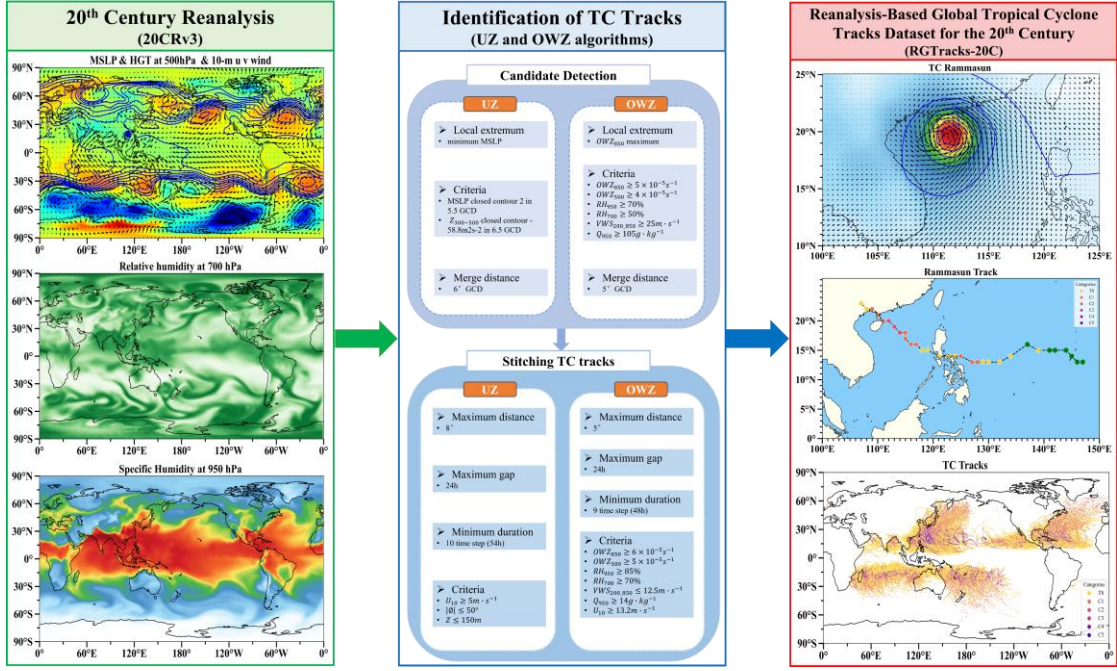


Figure 1: Schematic diagram showing the production of the RGrTracks-20C from the 20CRv3 based on the TC tracking algorithms. Variables shown include U10: 10-m wind speed, ϕ : latitude, z : altitude, GCD: great circle distance.

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):

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

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):

$$E = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \quad (2)$$

$$F = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \quad (3)$$

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

200 retained. Based on the information provided in Table 1, besides the required minimum threshold values
 201 for the OWZ parameter at 850 *hPa* and 500 *hPa*, additional dynamical and thermodynamic parameters
 202 related to TC formation are taken into account. These parameters include the maximum threshold for the
 203 wind vector difference (vertical wind shear) between 850 *hPa* and 200 *hPa*, as well as the relative
 204 humidity at 950 *hPa* and 700 *hPa*, and the minimum threshold for the specific humidity at 950 *hPa*.
 205 This step primarily aims to identify grid points that contain essential components of a storm. Subsequently,
 206 neighboring grid points are grouped together to define potential TCs.

207 Second step: Stitching TC tracks.

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

212
 213 **Table 1.** Parameter threshold values for the OWZ detection criteria. Subscripts stand for isobaric levels in *hPa*
 214 (OWZ: Obuko-Weiss-Zeta s^{-1} , RH: relative humidity %; VWS: vertical wind shear $\text{m} \cdot \text{s}^{-1}$; Q: specific humidity
 215 $\text{g} \cdot \text{kg}^{-1}$.)

Criterion	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

216

217 ii. UZ tracker

218 The UZ tracker, originally proposed by (Zarzycki and Ullrich, 2017), utilizes sea level pressure on the
 219 model grid, incorporating criteria for warm-core structures and storm lifetime. Its design prioritizes a low
 220 false-alarm rate (Roberts et al., n.d.).

221 First step: Candidate detection.

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

- 224 1. An increase in SLP minimum of at least 2 *hPa* within a 5.5° GCD from the candidate point to ensure
 225 the presence of a sufficiently strong and coherent low-pressure area.
- 226 2. The geopotential thickness between 300 and 500 *hPa* (denoted as $Z_{300-500}$) must decrease by
 227 $58.8 \text{ m}^2 \text{ s}^{-2}$ over a distance of 6.5° GCD from the maximum center of $Z_{300-500}$ within 1° GCD of the
 228 center of minimum SLP.

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

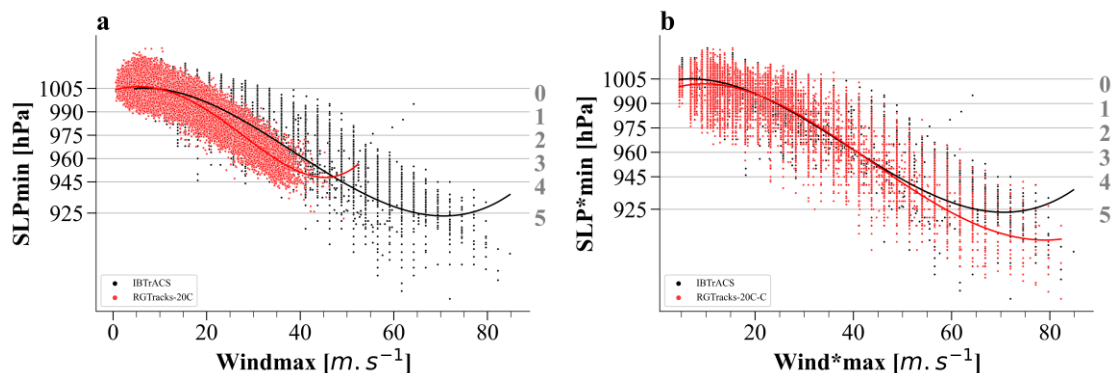
230 Second step: Stitching TC tracks.

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

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

239 2.2.3 Bias Correction of TC intensity

240 Given the low horizontal resolution in the 20CRv3, TC intensities derived directly from the reanalysis
 241 generally underestimated those compared to observations (Fig. 2a) (Bourdin et al., 2022; Roberts et al.,
 242 n.d.). To address this issue, a quantile mapping bias correction, similar to the method used by Zhao and
 243 Held (2010), was applied to adjust for the TC intensity bias within the dataset. The main idea is to fit the
 244 20CRv-derived TC intensity distributions, either probability distribution functions (PDFs) or cumulative
 245 distribution functions (CDFs), to the observed distributions. This method has demonstrated significant
 246 efficacy in enhancing the accuracy of TC intensity within models or reanalyses (Faranda et al., 2023;
 247 Yoshida et al., 2017). This adjustment resulted in a wind-pressure relationship in RGTracks-20C that
 248 aligns more closely with observational data (Fig. 2b).



249 **Figure 2:** Wind–pressure relationships for IBTrACS and RGTracks-20C. a–b, Scatter plots of SLP_{min} (unit: hPa)
 250 against maximum sustained wind speeds ($WIND_{max}$) (unit: $m \cdot s^{-1}$), based on the TCs from IBTrACS (black),
 251 RGTracks-20C (red), before (a) and after (b) intensity bias correction (see Methods). The curves represent fourth-
 252 order polynomial fit results. Storm categories, as defined in the section 'TC intensity', are indicated by horizontal
 253 gray lines.
 254
 255

256 2.3 Evaluation methods

257 2.3.1 Tracks matching

258 After detecting TC vortices from the 20CRv3, the resulting tracks are matched one-to-one with those
 259 observed in the IBTrACS. The specific procedures are detailed in the "2.4 Tracks Matching" section by
 260 (Bourdin et al., 2022).

261 Specifically, a detected track D consists of n points (d_1, d_2, \dots, d_n) corresponding to the moments $(t_1, t_2, \dots,$
 262 $t_n)$. Similarly, a track O observed in IBTrACS consists of a collection of points at a given time. For every
 263 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.
 264 There may be instances where no such points are found in O . The subset of points in O that are linked to
 265 any point in D is labeled as $O_{D-paired}$. It consists of $|O_{D-paired}|$. There are three possible scenarios:

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

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

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

272

273 2.3.2 Track verification

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

$$277 \quad \text{POD} = \frac{H}{H + M} \quad (4)$$

278

$$279 \quad \text{FAR} = \frac{FAs}{H + FAs} \quad (5)$$

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

283 2.4 Definitions

284 2.4.1 TC intensity

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

292 **Table 2.** Data format of the RGTracks-20C. track_id: storm identifier, lat: latitude degrees_north, lon: longitude
 293 degrees_east, SLP_{min} : minimum central pressure (unit hPa), $WIND_{max}$: maximum wind speed (unit: $m \cdot s^{-1}$),
 294 $WIND^*_{max}$ 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

295

296 **Storm categories:** the Saffir-Simpson Hurricane Scale (SSHS) from 1 to 5 based on their peak 1-minute
297 wind speed at 10 meters above the surface. In this study, given the significant uncertainties in $WIND_{max}$
298 due to differences between institutions and the limitations of model simulation capabilities (Chavas et al.,
299 2017; Klotzbach et al., 2020; Knutson et al., 2015), we have chosen to classify based on SLP_{min} , following
300 the definition of (Klotzbach et al., 2020).

301 **2.4.2 Basins**

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

305 **2.4.3 TC days**

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

$$309 \quad TC \text{ days} = \frac{1}{4} \sum_{i=0}^n L_i \quad (6)$$

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

311 **3. Results and discussion**

312 **3.1 Data Records**

313 The constructed RGTracks-20C (Ye et al., 2024) provides a century-long collection of global TCs
314 identified from the 20CRv3. The RGTracks-20C is publicly available at
315 <https://github.com/jeremyhleung/RGTracks-20C/> and <https://zenodo.org/record/8410597>. This dataset
316 provides detailed TC information, including location (longitude, latitude, hemisphere, and basin), time
317 (year, month, day, hour, and season), and intensity (SLP_{min} and $WIND_{max}$), with a temporal resolution of
318 6 hours, spanning from 1850 to 2014 and covering the globe. The dataset is provided as a comma
319 separated values (.csv) file and has a format similar to that of the IBTrACS (Table 2). It is noted that, in
320 the RGTracks-20C, $WIND_{max}$ serves, in addition to SLP_{min} , as a supplementary reference of TC intensity
321 for researchers, but is not discussed here due to accuracy issues and should be used cautiously.

322 **3.2. Tracker performance**

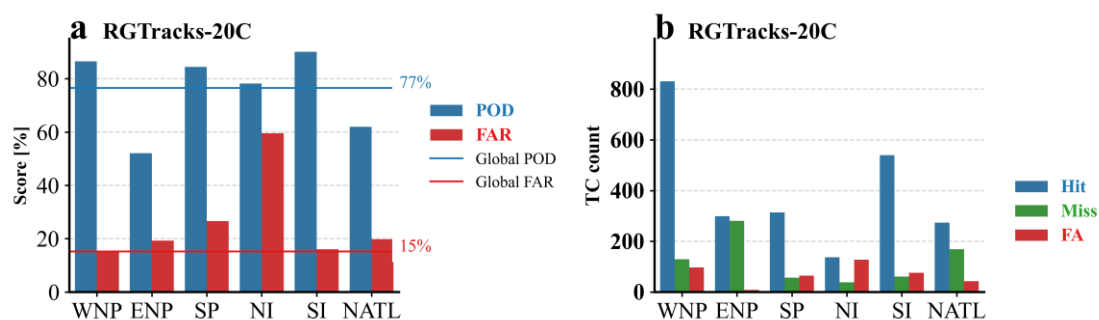
323 As documented in prior studies, biases are unavoidable when extracting TCs from reanalyses, given the
324 limitations of reanalysis in reproducing the high-resolution TC structure and circulation patterns, as well
325 as the errors caused by the application of different trackers (Bell et al., 2018; Horn et al., 2014; Lee et al.,
326 2023; Slivinski et al., 2019; Truchelut et al., 2013). [Before verifying the reliability of RGTracks-20C, it](#)
327 [is necessary to evaluate the performance of the trackers applied. To further support the credibility](#)
328 [assessment of RGTracks-20C, we also evaluate the performance of the UZ tracker as a methodological](#)

329 reference, with a more detailed discussion provided in Sect. 3.5, Supplementary Sects. S1–S2 and Figs.
 330 S1–S7.

331 The POD and FAR of TCs are used to assess OWZ’s performance in detecting TCs in 20CRv3 at both
 332 the global scale and across six ocean basins (see Sect. 2.3.2). Globally, the OWZ tracker yields an overall
 333 POD of 77% and an FAR of 15% (Fig. 3a). Its higher POD and lower number of misses (Fig. 3b) are
 334 advantageous for augmenting the limited observational record during early periods with sparse
 335 observations. Most misses are associated with weak and short-lived TCs that the 20CRv3 fails to simulate
 336 or track (Hodges et al., 2017), as well as its inability to meet the threshold of the tracker. The false alarms,
 337 however, likely represent storms or their developmental stages that were not recorded in IBTrACS,
 338 possibly because they did not reach the tropical storm category (Supplementary Fig. S2).

339 For each basin, the OWZ tracker shows a pattern generally consistent with the global results. Higher POD
 340 values are found in the SI (90%), WNP (86%), and SP (84%), followed by the NI (78%), whereas lower
 341 POD values occur in the NATL (62%) and ENP (52%). The largest numbers of hits are found in the WNP
 342 and the SI Ocean, consistent with the results reported by Bourdin et al. (2022). Overall, the OWZ tracker
 343 identifies most of the observed TCs in the majority of basins, although missed detections remain evident
 344 in the ENP and NATL, where many of the missed TCs are very weak and short-lived TCs (Supplementary
 345 Fig. S3), while the FAR is relatively high in the NI. In addition, some of the false alarms may correspond
 346 to subtropical storms or to weak systems that were not retained in IBTrACS, as suggested in previous
 347 studies (Bell et al., 2018; Bourdin et al., 2022).

348



349

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

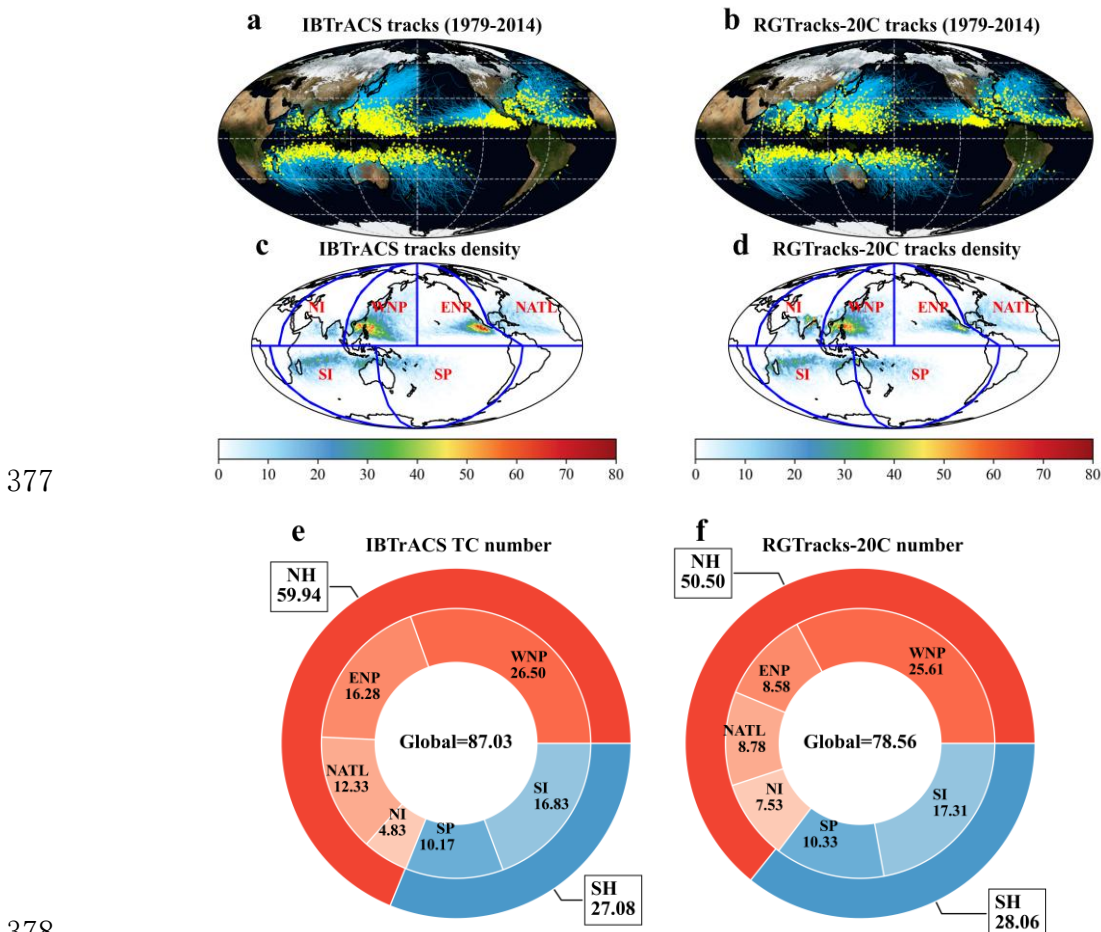
354 Furthermore, comparison with previous studies, as well as the separately evaluated UZ-based results
 355 (Sect. 3.5, and Supplementary Sect. S1 and Fig. S1 and Table S1), provides additional support for the
 356 credibility of the OWZ-based detections in 20CRv3 (Bourdin et al., 2022; Murakami, 2014). Particularly,
 357 the POD and FAR obtained from 20CRv3 exhibit similar magnitudes to the results by applying OWZ to
 358 ERA5, which yields global POD and FAR values of 76% and 25%, respectively (Bourdin et al., 2022;
 359 see Supplementary Table S1). This suggests that the OWZ-based detections provide a reasonable
 360 representation of historical TC activity in 20CRv3 and support their use as the primary dataset for the
 361 first version of RGTracks-20C.

362 **3.3 Climatology**

363 As the first version of RGTracks-20C, the primary objective is to supplement incomplete historical TC records.
 364 In the following sections, we therefore focus on evaluating its ability to capture TC climatology
 365 across various ocean basins.

366 In terms of climatology, the RGTracks-20C is able to capture the major spatial patterns of TC genesis
 367 locations and track density over most ocean basins (Figs. 4a–d), indicating that it provides a reasonable
 368 representation of the spatial distribution of historically observed TCs. The annual mean TC numbers are
 369 also consistent with observations in most ocean basins, especially in the WNP, SI, and SP, where the
 370 discrepancies range from -0.48 to 0.89 (Figs. 4e–f). Globally, the annual average TC number of
 371 RGTracks-20C is 78.56 (Fig. 4f), which is lower than the observed value of 87.03 (Fig. 4e). This
 372 underestimation mainly arises from biases in the ENP and NATL basins. Despite these regional
 373 underestimations, the overall TC detection level is broadly comparable to that reported in previous studies
 374 based on higher-quality reanalyses (Bourdin et al., 2022; Murakami, 2014), suggesting that RGTracks-
 375 20C provides a reasonable representation of global TC frequency climatology.

376



378 **Figure 4:** TC genesis locations, tracks, and annual average number from IBTrACS and RGTracks-20C. a–b, TC
 379 genesis locations (yellow dots) and tracks (blue lines) from IBTrACS (a), and RGTracks-20C (b). c–d, TC tracks
 380 density (shading, number of TC occurrence per $1^\circ \times 1^\circ$ latitude-longitude grid box, 1979–2014) from IBTrACS (c),
 381 and RGTracks-20C (d). e–f, mean number of TCs per year globally and for the six basins from IBTrACS (e), and
 382 RGTracks-20C (f).
 383

384 Based on the spatial distribution shown above, we further compare TC locations between RGTracks-20C
385 and IBTrACS. The results show that global TC location differences range from 10 to 300 km, with most
386 cases falling between 75 and 125 km (Fig. 5a). Peak frequencies occur below 100 km, at approximately
387 95 km. Similar distributions are found across the individual basins (Fig. 5a). Given that the nominal
388 horizontal resolution of 20CRv3 is $1^\circ \times 1^\circ$, corresponding to a characteristic spatial scale of about 100
389 km, this location differences are within a physically reasonable range for TC detection from this
390 reanalysis.

391 TC duration is also an important aspect of TC climatology (Knutson et al., 2010). Based on the IBTrACS,
392 the majority of observed TCs globally last fewer than 20 days, with a peak around 8 days (Fig. 5b). The
393 results show that the TC duration distribution in RGTracks-20C is close to the observations, with a peak
394 near 8 days at both the global and basin scales. (Bourdin et al., 2022; Tory et al., 2013). This is likely
395 related to the ability of the dynamics-based OWZ tracker to identify storms from relatively early stages
396 of development (Supplementary Figs. S4) (Bell et al., 2018; Bourdin et al., 2022).

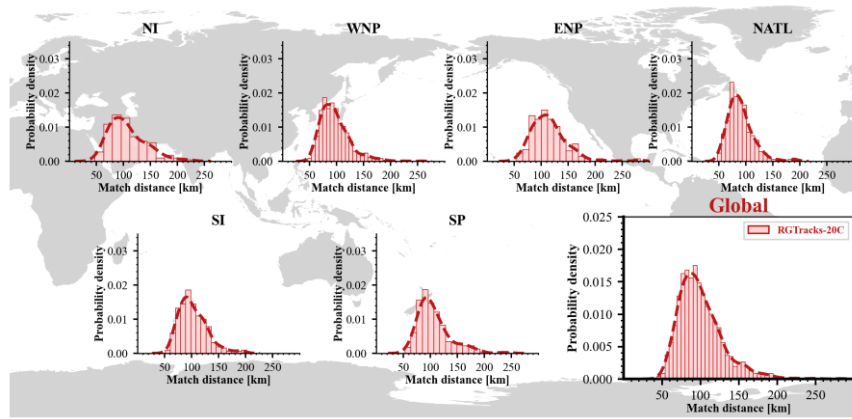
397 For TC intensity, given the relatively considerable uncertainty in $WIND_{max}$ compared to SLP_{min} in both
398 reanalyses and IBTrACS (see Methods) (Bourdin et al., 2022; Chavas et al., 2017; Knapp et al., 2010;
399 Knutson et al., 2015; Schreck et al., 2014), the intensity evaluation of RGTracks-20C is based on SLP_{min}
400 only. According to IBTrACS (Fig. 5c), the intensity of TCs is mainly distributed between 900 and 1020
401 hPa , peaking around 1000 hPa , with a long tail on the lower SLP_{min} side. In RGTracks-20C, the SLP_{min}
402 distribution is mainly concentrated between 950 and 1020 hPa , with a peak near 1005 hPa . This suggests
403 that the 20CRv3 generally underestimates the TC intensity (Fig. 2a), which, as expected, is primarily
404 because the relatively low spatial resolution of the reanalysis may cause smoothing effects on the sea
405 level pressure field. Apart from spatial resolution, the model's dependence on parameterization processes,
406 along with other factors, may also influence its ability to reproduce TC intensity in the reanalysis (Aarons
407 et al., 2021; Hodges et al., 2017; Malakar et al., 2020).

408 To address this issue, an intensity correction based on quantile mapping was applied (see Methods) (Zhao
409 and Held, 2010). After intensity correction, the TC intensity distribution in RGTracks-20C is more
410 consistent with IBTrACS (Fig. 2b and Fig. 5c), especially in terms of peak positions, and accurately
411 reproduces the skewed distribution of TC intensity. In particular, the RGTracks-20C reproduces TC
412 intensity values with SLP_{min} below 940 hPa , which were not found before the intensity bias correction.
413 Notably, while the fitted curves show consistent patterns following correction, they do not perfectly
414 overlap, suggesting that small discrepancies exist (Supplementary Figs. S5d and S6d).

415 In summary, these results suggest that RGTracks-20C provides a reasonable representation of observed
416 TC climatology in terms of spatial distribution, duration, and intensity.

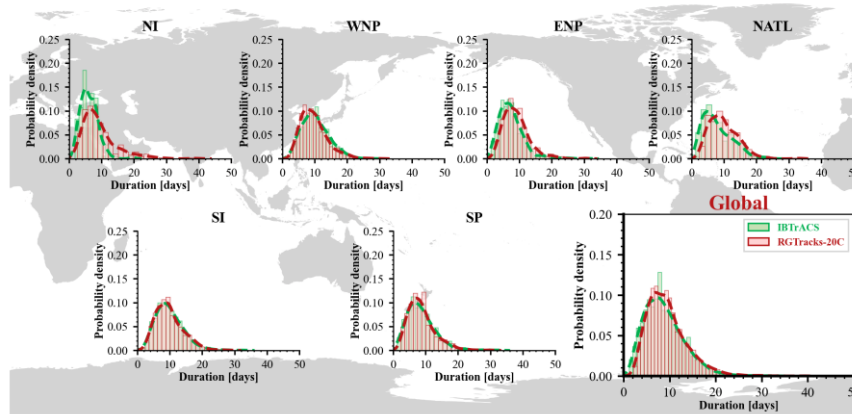
417
418
419

a TC Location (1979-2014)



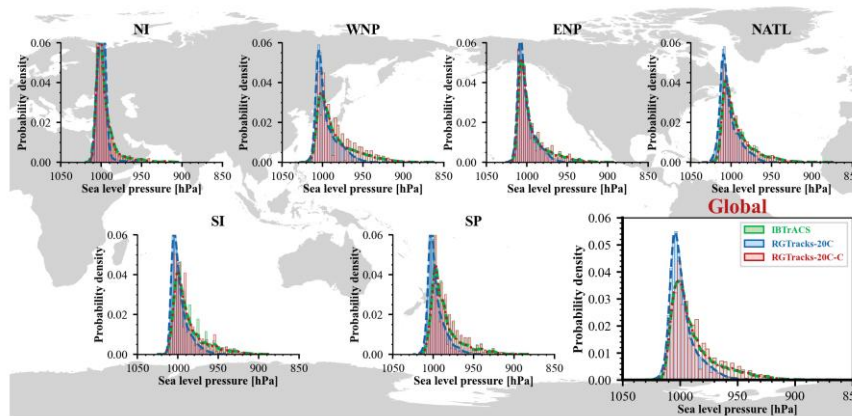
420

b TC Duration (1979-2014)



421

c TC Intensity (1979-2014)



422

423

424

425

426

427

Figure 5: Distribution of TC characteristics on the IBTrACS and RGTracks-20C. a, Distribution of the mean TC location error from 1979–2014 (unit: km) between IBTrACS and the RGTracks-20C. b, TC duration (unit: days) from 1979 to 2014 in IBTrACS (green) and the RGTracks-20C (red) (unit: hPa). c, same as (b), but for TC intensity (SLP_{min} , unit: hPa), before (blue) and after (red) bias correction.

428 **3.4 Interannual variability**

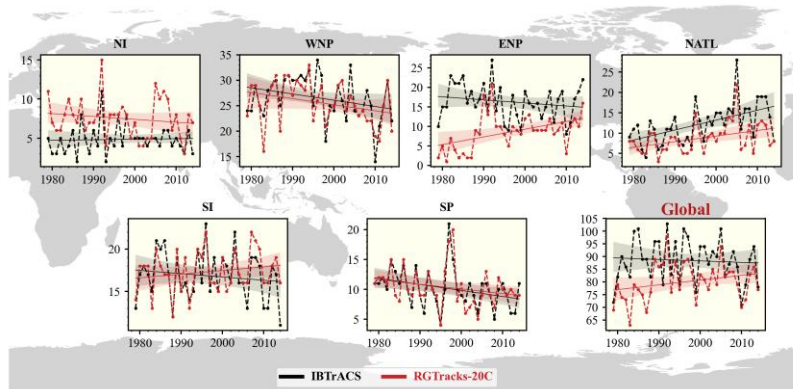
429 To assess whether RGTracks-20C can represent year-to-year variations in TC activity, we compare it
430 with IBTrACS over 1979–2014, when the observational record is relatively complete and more reliable.
431 Firstly, the RGTracks-20C is able to capture the observed interannual variability of global TC number
432 (Fig. 6a), with correlation coefficients of 0.68 (in the following context, all correlations are significant at
433 the 99% confidence level unless otherwise specified). This is also true for individual basins (Fig. 6a and
434 Table 2), with the correlation coefficients exceeding 0.70 in most basins. Among the six basins, the
435 highest correlation is found in the NATL (0.88), followed by the SP (0.79), WNP (0.75), and SI (0.74).
436 Lower correlations are found in the ENP (0.35) and NI (0.54). These results suggest that RGTracks-20C
437 can provide useful information on year-to-year variations in TC number, both globally and across most
438 major ocean basins.

439 TC days, a metric that combines both TC frequency and lifetime (Bell et al., 2018), are also reasonably
440 represented in RGTracks-20C. At the global scale, TC days in RGTracks-20C are significantly correlated
441 with those in IBTrACS, with a correlation coefficient of 0.63 (Fig. 6b). Moreover, these results are further
442 confirmed across basins (Fig. 6b and Table 2), with correlation coefficients generally exceeding 0.75. In
443 particular, the highest correlations occur in the NATL (0.93), followed by the SP (0.79), SI (0.78), and
444 WNP (0.75). However, the correlation coefficients for TC days are also relatively low in the ENP (0.57)
445 and NI (0.34). Overall, these results suggest that RGTracks-20C provides a reasonable representation of
446 the interannual variability of TC days at the global scale and in most ocean basins.

447 In addition, the global TC intensity series based on RGTracks-20C is significantly correlated with that
448 from IBTrACS, with a correlation coefficient of 0.80 after intensity bias correction (Fig. 6c). This
449 suggests that the SLP_{min} TC intensity in RGTracks-20C is broadly consistent with the observed
450 interannual variability. Similar results are found in most basins (Fig. 6c and Table 2), with the highest
451 correlation in the WNP (0.85), followed by the SI (0.78) and NATL (0.72), while the SP also shows a
452 correlation close to 0.69. Because 20CRv3 tends to underestimate TC intensity owing to its coarse spatial
453 resolution, intensity bias correction is necessary during the construction of RGTracks-20C (see Methods).
454 After correction, the interannual variability of TC intensity becomes more consistent with the
455 observations, especially in the WNP, NATL, and SI basins. Discrepancies remain in the ENP (-0.02) and
456 NI (0.47), consistent with the results for TC number and TC days discussed above.

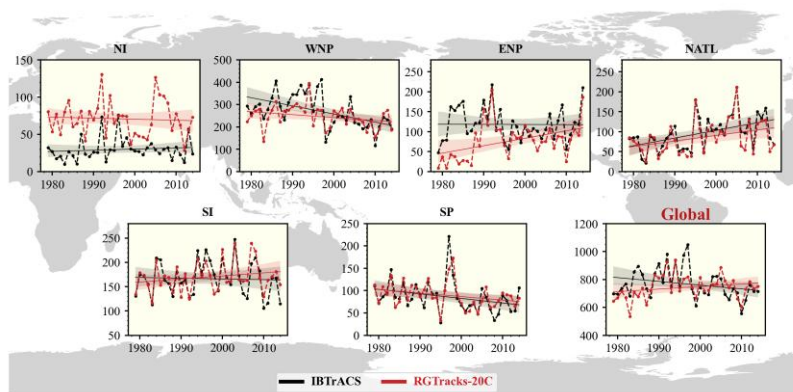
457 Overall, these results suggest that RGTracks-20C provides a reasonable basis for analyzing interannual
458 variations in TC activity, particularly for TC number, TC days, and intensity, in periods when the
459 observational record is more complete.

a TC Number (1979-2014)



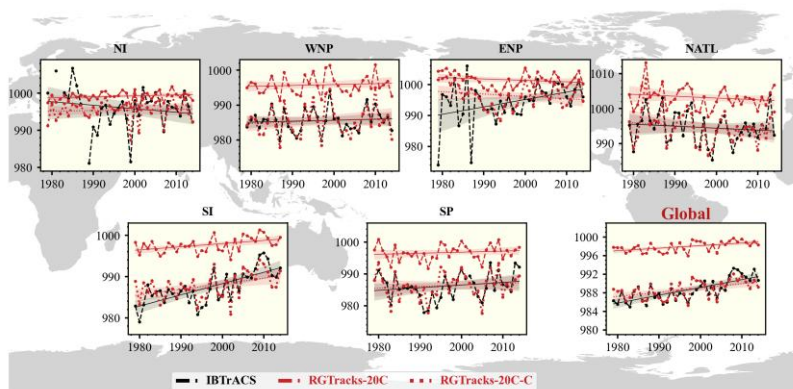
460

b TC Duration (days, 1979-2014)



461

c TC Intensity (1979-2014)



462

463 **Figure 6:** Time series of globally TC activities from IBTrACS and RGTracks-20C during the periods 1979-2014.

464 TC activities are from the IBTrACS and RGTracks-20C (red). a, TC number. b, TC days (unit: days). c, TC intensity

465 in SLP_{min} (unit: hPa) in IBTrACS (black) and RGTracks-20C before (red solid line) and after (red dotted line) bias

466 correction. Shaded areas are the two-sided interval of the linear trend at the 95% confidence level. Straight lines are

467 the linear regression.

468 The basin-dependent differences discussed above are most evident in the ENP, NATL, and NI. Weak and

469 short-lived storms are difficult to identify across all basins (Supplementary Fig. S2), but this problem is

470 particularly pronounced in the ENP and NATL (Supplementary Fig. S3), where such storms account for

471 a larger share of the historical record. The coarse resolution of the reanalysis makes these weak systems

472 more difficult to represent, which in turn reduces the detection skill of the tracking algorithm. In the NI,

473 the lower agreement is more likely related to a different combination of factors, including greater

474 uncertainty in the historical observational record and the tendency of the tracking algorithm to misidentify
475 monsoon depressions as TCs. Consistent with these basin-dependent characteristics, the agreement with
476 IBTrACS is weaker in the ENP and NI: RGTracks-20C tends to underestimate TC activity in the ENP
477 and overestimate it in the NI during the earlier part of the record (Fig. 6a and Supplementary Sect. S3).
478 When the comparison is restricted to the more reliable periods after 1988 in the ENP and after 1990 in
479 the NI, the biases are reduced and the correlations improve substantially (Supplementary Fig. S8 and
480 Tables S2, S5). These results suggest that the discrepancies in these two basins are influenced not only
481 by tracker and reanalysis limitations, but also by temporal inhomogeneities in the observational record
482 (Supplementary Table S3). Because part of the IBTrACS pressure information is assimilated into
483 20CRv3, such inhomogeneities may affect both the representation of TC activity in the reanalysis and
484 the subsequent evaluation against IBTrACS. Consistent with this interpretation, excluding the ENP and
485 NI also increases the global correlation coefficients between RGTracks-20C and IBTrACS (Table S6).
486 Although long-term trend analysis is not the primary purpose of the present dataset, the corresponding
487 trend estimates are provided in Supplementary Sect. S4 and Table S4, S7 for completeness. Over 1979–
488 2014, RGTracks-20C shows generally similar longer-term variations to IBTrACS at the global scale, and
489 several basins exhibit trends of the same sign in both datasets, including significant trends in some cases.
490 These results are provided as supporting information only; the current version of RGTracks-20C is not
491 intended for standalone analyses of long-term trends in TC activity.

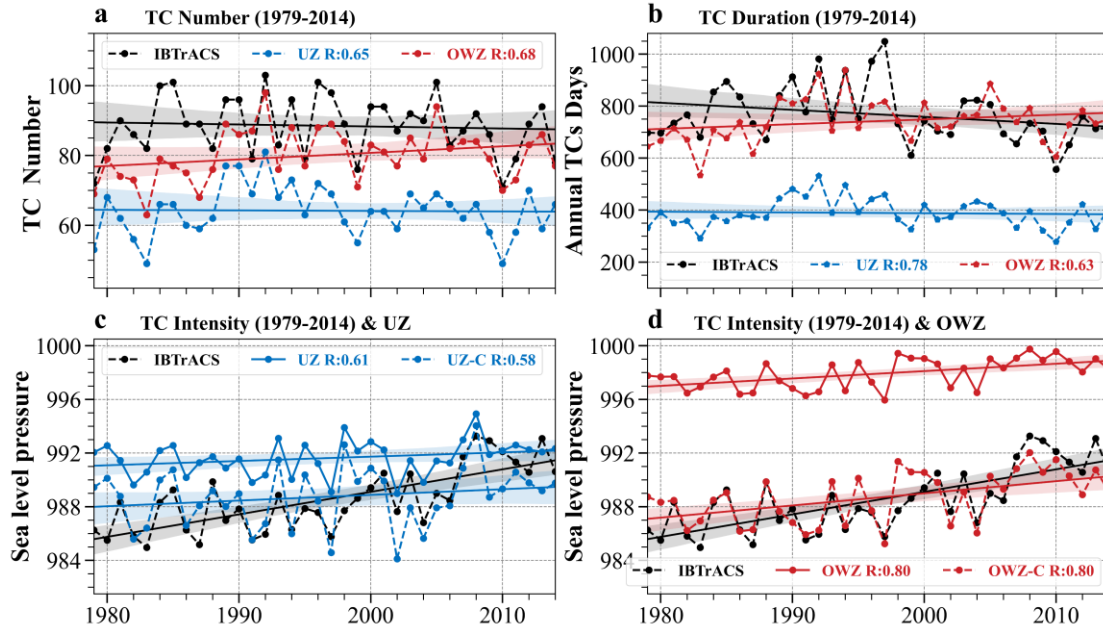
492 **3.5 Comparison of the OWZ and UZ trackers**

493 It is important to recognize that different TC detection and tracking schemes can yield different results,
494 because they rely on different formulations and threshold criteria. In this study, the UZ tracker is included
495 as a methodological reference and sensitivity test (Supplementary Sect. S2).

496 Compared with the OWZ tracker, the UZ tracker shows a lower global POD (68% versus 77%) and a
497 lower FAR (7% versus 15%), but also a larger number of missed detections, especially for weak and
498 short-lived storms (Supplementary Figs. S1–3). The UZ-based annual mean TC number (63.39) is
499 systematically lower than that of RGTracks-20C (78.56) and farther from the observed value (87.03)
500 (Supplementary Fig. S7). In addition, the UZ-based duration distribution is markedly shorter, with a peak
501 near 5 days, whereas both RGTracks-20C and IBTrACS show a peak near 8 days (Supplementary Figs.
502 S5–6). By contrast, differences in spatial location and intensity are relatively smaller.

503 In terms of interannual variability, the UZ-based results are broadly similar to the OWZ-based results
504 (Fig. 7 and Table 2), but they show larger deviations from IBTrACS. In particular, the UZ-based annual
505 TC number is systematically lower than the observations by about 10–20 storms (Fig. 7a), and TC days
506 are also substantially lower than those in IBTrACS (Fig. 7b). For TC intensity, the OWZ-based results
507 are more consistent with the observations, with a correlation coefficient reaching 0.80, whereas the
508 corresponding UZ-based value is only 0.61 (Figs. 7c–d).

509 Therefore, these results indicate that the OWZ-based product provides a more suitable basis for the
510 current first version of RGTracks-20C, while the UZ-based results remain useful as a methodological
511 reference and for comparative analysis.



512
513
514
515
516
517
518
519
520
521
522
523
524

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 and OWZ (red) trackers. a, TC number. b, TC days (unit: days). c, TC intensity in SLP_{min} (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 SLP_{min} (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.

Table 2. The correlation coefficients (R) between the from IBTrACS and RGTracks-20C and UZ tracker. Note*: The R values for TC number and TC intensity are not statistically significant at the 99% confidence level in the NI and ENP.

	Characteristics	WNP	SP	SI	NATL	ENP	NI
RGTracks-20C	Number	0.75	0.79	0.74	0.88	0.35	0.54
	Duration	0.75	0.79	0.78	0.93	0.58	0.34
	Intensity	0.85	0.69	0.78	0.75	0.15	0.36
	Intensity-C	0.86	0.69	0.79	0.79	0.11	0.47
UZ	Number	0.79	0.84	0.69	0.79	0.41	0.49
	Duration	0.84	0.82	0.80	0.91	0.65	0.62
	Intensity	0.82	0.71	0.72	0.75	0.01	0.48
	Intensity-C	0.79	0.69	0.70	0.72	-0.02	0.55

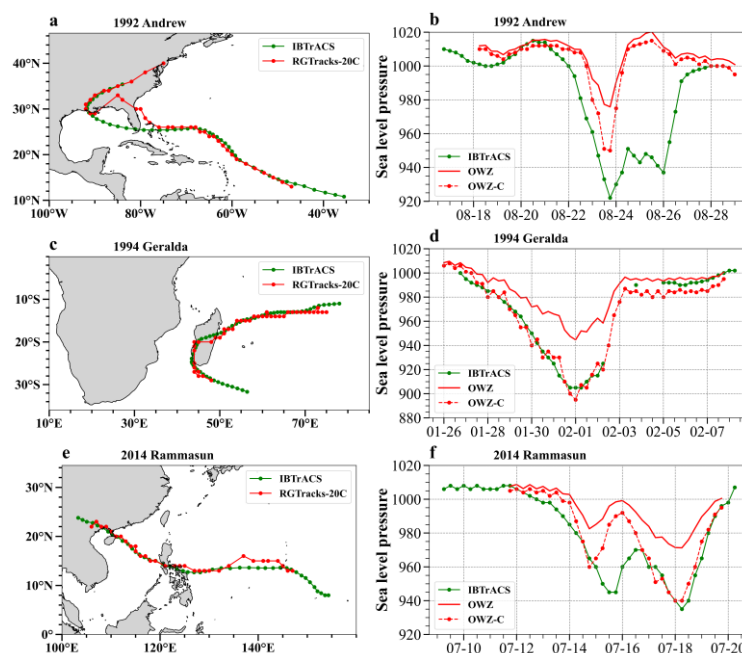
525
526
527
528
529

530 **3.6 Case Studies**

531 Compared with observations, RGTracks-20C provides a reasonable representation of TC climatology and
532 interannual variability at the global scale and across the major ocean basins. In this section, we examine
533 the key strengths of RGTracks-20C, specifically its potential to provide supplementary track and intensity
534 information of early-year TCs that may not be included in the observed data records.

535 Before digging into early-year TCs, we first assess whether RGTracks-20C can reasonably represent
536 several well-documented individual events by comparison with observations. To provide representative
537 examples across different ocean basins, we selected three high-impact TCs with complete observational
538 records in the NATL, SI, and WNP: Hurricane ‘Andrew’ in 1992 (Pimm et al., 1994) (Figs. 8a–d), TC
539 ‘Geralda’ in 1994 (Hoarau et al., 2012) (Figs. 8c–d), and Super Typhoon ‘Rammasun’ in 2014 (Zhang et
540 al., 2017) (Figs. 8e–f). These storms were chosen because they caused substantial human and economic
541 losses and therefore provide useful test cases for assessing the representation of TC track, duration, and
542 intensity evolution in RGTracks-20C.

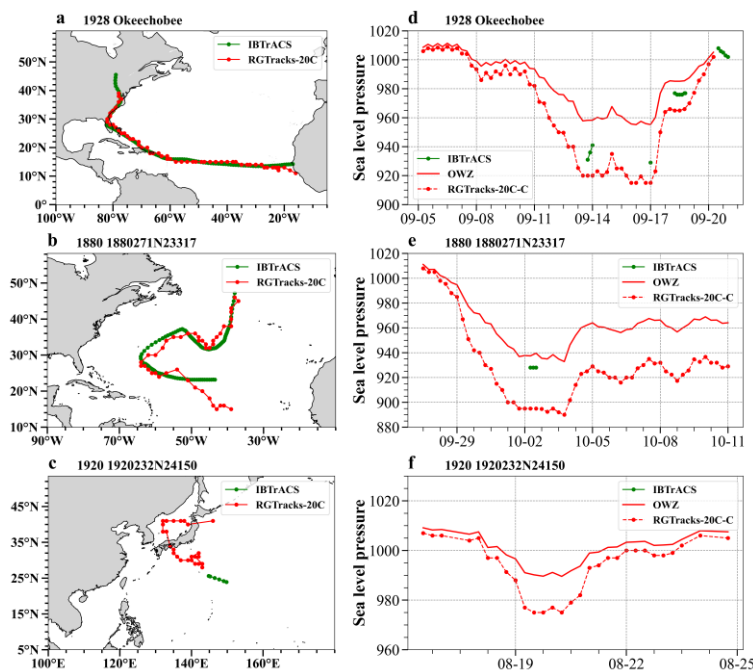
543 Compared with IBTrACS, the RGTracks-20C performs well in representing the track and duration of
544 these TCs. Some discrepancies were observed during landfall (Fig. 8a), possibly due to small TC size,
545 which were not captured by the low-resolution 20CRv3 (Supplementary Sect. S5.1 and Fig. S9). While
546 the 20CRv3 tends to underestimate the intensity of TCs, the corrected intensity in the RGTracks-20C is
547 highly consistent with observations and accurately captures the temporal evolutions of TC intensities.
548 These examples suggest that RGTracks-20C can provide a useful representation not only of TC
549 climatology and variability, but also of the evolution of selected individual TC events.



550 **Figure 8:** The historical tracks and intensity records of individual TCs in the IBTrACS and RGTracks-20C. a–b,
551 Track (a) and intensity (b, SLP_{min} , unit: hPa) of Hurricane “Andrew”. c–d, same as a–b, but for tropical cyclone
552 “Geralda”. e–f, same as a–b, but for Super typhoon “Rammasun”. Green and red lines denote results based on the
553 IBTrACS and RGTracks-20C, respectively.
554

555

556 Prior to the satellite era, limitations in observation systems often led to incomplete records of early TCs,
 557 particularly for TC intensity. An example is Hurricane Okeechobee in 1928, which was one of the
 558 deadliest to hit the United States in the early 20th century. Hurricane Okeechobee was recorded in the
 559 IBTrACS (Blake et al., 2011; Mitchell, 1928) (Supplementary Sect. S5.2). However, during
 560 Okeechobee’s lifetime, there were only 16 time points of the TC intensity that were recorded when it
 561 passed the Lesser Antilles and Puerto Rico, and made landfall in the United States (Figs. 8a–b,
 562 Supplementary Fig. S10 and Table S8). Similar missing data are common in the IBTrACS records of
 563 early TCs, especially when the TCs were located over the open ocean (Figs. 8c–d). In some cases, the
 564 incompleteness is even more pronounced. For example, Typhoon No. 8, which made landfall and caused
 565 serious damage in Japan (Supplementary Sect. S5.3), has only track records in the IBTrACS, but with
 566 intensity information missing (Figs. 8e–f).



567 **Figure 9:** As in Fig. 8, but for Hurricane “Okeechobee” (a–c), Hurricane ‘1880271N23317’ (d–f), typhoon
 568 ‘192023N24150’ (g–i).

569 In such cases, RGTracks-20C can provide supplementary information on both TC track and intensity
 570 where the observational record is incomplete. For Okeechobee, RGTracks-20C reproduces nearly the full
 571 storm lifetime recorded in IBTrACS, and its latitude-longitude evolution is broadly consistent with the
 572 observed track, with positional differences generally within ± 1 (Fig. 8a and Supplementary Fig. S10).
 573 The available observations also suggest that RGTracks-20C provides a reasonable estimate of
 574 Okeechobee’s intensity evolution (Fig. 8d and Supplementary Table S8). In particular, when the
 575 hurricane passed Guadeloupe, IBTrACS records a minimum sea-level pressure of 940 hPa, which is
 576 closely matched in RGTracks-20C. Moreover, RGTracks-20C captures the weakening and re-
 577 intensification of the hurricane between Puerto Rico and its landfall in Florida, a period for which
 578 IBTrACS lacks continuous intensity information. These comparisons indicate that, despite the sparsity
 579 of the historical intensity record, the RGTracks-20C represents the major intensity changes of the storm
 580 reasonably well and gives central pressure estimates that are broadly consistent with the recorded values
 581 at key stages of its evolution (Supplementary Sect. S5.2 and Table S8). Similar value is seen for Typhoon
 582

583 No. 8 in 1920, where RGTracks-20C provides additional information not only on TC intensity but also
584 on the later stage of storm evolution, including the landfall phase (Fig. 9g and Supplementary Sect. S5.3
585 and Figs. S11–S13).

586 Moreover, prior to the satellite era, the RGTracks-20C often reports a higher number of TCs than the
587 IBTrACS, particularly from the early to mid-20th century (Supplementary Sect. 6). This does not by itself
588 confirm that every additional system represents a previously undocumented historical TC, but it does
589 suggest that RGTracks-20C may provide useful supplementary information in periods when the historical
590 observational archive is incomplete. This is in agreement with previous attempts at correcting historical
591 biases (Vecchi and Knutson, 2008). These case studies, therefore, illustrate one of the main strengths of
592 the current version of RGTracks-20C, that is, its potential to extend and complement the historical record
593 of TC track and intensity information during early periods with sparse observations.

594 **4. Usage notes**

595 In this study, we introduce RGTracks-20C, a century-long reanalysis-based global TC dataset for the
596 period 1850–2014. The evaluations above suggest that RGTracks-20C provides a reasonable
597 representation of TC climatology and interannual variability at both global and basin scales during the
598 modern observational era. One important strength of the dataset is its potential to provide supplementary
599 track and intensity information for earlier periods when the historical record is sparse or incomplete.

600 At the same time, several points should be considered when interpreting or using the current version of
601 RGTracks-20C.

602 (1) TC intensity in RGTracks-20C remains relatively uncertain, compared with TC occurrence and track
603 information. Although a bias correction based on quantile mapping was applied to reduce the low-
604 intensity bias in 20CRv3, the corrected intensity should still be interpreted with caution, especially for
605 weaker systems and for analyses requiring finer distinctions in storm intensity (Hodges et al., 2017).
606 Given the relatively large uncertainties in wind fields, the wind information in the current version of
607 RGTracks-20C should be used with caution.

608 (2) Second, discrepancies between RGTracks-20C and IBTrACS in certain basins should not be attributed
609 entirely to limitations of RGTracks-20C. The observational record itself carries uncertainties arising from
610 changes in monitoring technology and data sources over time, which vary considerably across basins
611 (Chan et al., 2022b; Torn and Snyder, 2012) (Supplementary Table S3). These uncertainties are most
612 pronounced in the ENP and NI basins, where the reliability of IBTrACS has changed substantially over
613 time. Users are therefore advised to exercise particular caution when applying RGTracks-20C in these
614 two regions.

615 (3) Two characteristics of the underlying reanalysis should be noted. The use of the ensemble mean field
616 of 20CRv3 introduces a smoothing effect that tends to weaken the representation of intense TCs (Emanuel,
617 2024), contributing to the residual low-intensity bias in RGTracks-20C even after bias correction. In
618 addition, the number of assimilated observations in 20CRv3, including TC-related records from IBTrACS,
619 increases substantially over time, particularly after the mid-twentieth century (Supplementary Figs. S14-
620 15; Slivinski et al., 2019, 2021), meaning that the reliability of RGTracks-20C is higher in more recent
621 decades than in the early record (Supplementary Fig. S15). The growth in TC counts during the mid-

622 twentieth century in both datasets is, moreover, substantially driven by improvements in observational
623 technology rather than actual changes in TC activity. Users are therefore advised to exercise caution when
624 examining long-term trends in TC activity, particularly before 1950.

625 (4) TC tracking algorithms also introduce uncertainties into the dataset. In RGTracks-20C, the OWZ
626 tracker is used as the primary tracking algorithm, as it demonstrates overall stability in TC detection and
627 produces TC counts and TC days that are in closer agreement with IBTrACS at the global scale (Bourdin
628 et al., 2022). It should be noted that globally uniform tracking thresholds were applied across all basins.
629 Given the structural diversity of TCs and the influence of regional meteorological conditions and
630 topography, this may affect detection accuracy in certain regions (Fu et al., 2021; Raavi and Walsh, 2020a,
631 b). More broadly, different tracking algorithms applied to the same reanalysis can yield notably different
632 results in terms of TC counts, track duration, and intensity distributions (Flaounas et al., 2023; Han and
633 Ullrich, 2025). Users should therefore be aware that the TC statistics derived from RGTracks-20C are to
634 some extent dependent on the choice of tracking algorithm and thresholds, and results may differ from
635 those obtained using alternative approaches.

636 **5. Data Availability**

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

642 **6. Code Availability**

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

646 **7. Conclusion**

647 In this paper, we introduce the first version of RGTracks-20C, a century-long reanalysis-based global TC
648 dataset constructed primarily from the OWZ tracker applied to 20CRv3. Comparisons with IBTrACS
649 over 1979–2014 show that RGTracks-20C provides a reasonable representation of TC climatology and
650 interannual variability at both the global scale and across most major ocean basins during the period,
651 when the observational record is relatively more complete. These evaluations suggest that the current
652 version of RGTracks-20C can serve as a useful supplementary dataset for representing TC number, TC
653 days, track location, duration, and intensity (SLP_{min}). The case studies further illustrate the main strength
654 of the current version of RGTracks-20C. For well-observed storms in the modern era, RGTracks-20C
655 shows good agreement with the observed evolution of TC tracks and intensity after bias correction. More
656 importantly, for earlier periods with sparse or incomplete observations, RGTracks-20C can provide

657 supplementary information on TC track and intensity evolution where the historical record is incomplete.
658 In this sense, the present dataset is particularly valuable for augmenting the limited observational record
659 prior to the satellite era. The results presented here indicate that RGTracks-20C provides a useful basis
660 for extending information on historical TC activity into the pre-satellite era and for supporting future
661 studies of TC activity and historical events.

662 **Competing interests**

663 The authors declare no competing interests.

664 **Author contributions**

665 GY: methodology, formal analysis, data curation, visualization, software, and writing: original draft.
666 JCHL and BZ: conceptualization, methodology, formal analysis, funding acquisition, and writing—
667 review and editing. WD: supervision, funding acquisition, and writing: review and editing. RT:
668 manuscript structure refinement, and writing: review and editing. JX, WL, WQ, and HK: writing: review
669 and editing. All authors contributed to reviewing and editing the final version of the manuscript.

670 **Acknowledgements**

671 The authors sincerely thank Dr. Stella Bourdin from the Laboratoire des Sciences du Climat et de
672 l'Environnement, Institut Pierre Simon Laplace (LSCE-IPSL), Gif-sur-Yvette, for her invaluable
673 assistance and guidance on the TC trackers. The authors are very grateful to Dr. Jennifer Gahtan from
674 NOAA's National Center for Environmental Information for providing information about the starting
675 years of the minimum central pressure in the IBTrACS. [The authors are grateful to Yushan Han from the
676 University of California, Davis for providing valuable assistance with the System for Classification of
677 Low-Pressure Systems \(SyCLoPS\) algorithm.](#)

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

685

686

687 **References**

688 Aarons, Z. S., Camargo, S. J., Strong, J. D. O., and Murakami, H.: Tropical cyclone characteristics in the
689 MERRA-2 reanalysis and AMIP simulations, *Earth Space Sci.*, 8, e2020EA001415,
690 <https://doi.org/10.1029/2020EA001415>, 2021.

691 Bell, S. S., Chand, S. S., Tory, K. J., and Turville, C.: Statistical Assessment of the OWZ tropical cyclone
692 tracking scheme in ERA-Interim, *J. Climate*, 31, 2217–2232, <https://doi.org/10.1175/JCLI-D-17->
693 0548.1, 2018.

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

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

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

703 Bourdin, S., Fromang, S., Dulac, W., Cattiaux, J., and Chauvin, F.: Intercomparison of four algorithms
704 for detecting tropical cyclones using ERA5, *Geosci. Model Dev.*, 15, 6759–6786,
705 <https://doi.org/10.5194/gmd-15-6759-2022>, 2022.

706 Chan, J. C. L.: Frequency and intensity of landfalling tropical cyclones in East Asia: Past variations and
707 future projections, *Meteorology*, 2, 171–190, <https://doi.org/10.3390/meteorology2020012>, 2023.

708 Chan, K. T. F.: Are global tropical cyclones moving slower in a warming climate?, *Environ. Res. Lett.*,
709 14, 104015, <https://doi.org/10.1088/1748-9326/ab4031>, 2019.

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

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

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

718 Chang, E. K. M. and Guo, Y.: Is the number of North Atlantic tropical cyclones significantly
719 underestimated prior to the availability of satellite observations?, *Geophys. Res. Lett.*, 34, L14801,
720 <https://doi.org/10.1029/2007GL030169>, 2007.

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

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

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

728 Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E.,
729 Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant,
730 A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok,
731 H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The

732 Twentieth Century Reanalysis Project, Q.J.R. Meteorol. Soc., 137, 1–28,
733 <https://doi.org/10.1002/qj.776>, 2011.

734 Cram, T. A., Compo, G. P., Yin, X., Allan, R. J., McColl, C., Vose, R. S., Whitaker, J. S., Matsui, N.,
735 Ashcroft, L., Auchmann, R., Bessemoulin, P., Brandsma, T., Brohan, P., Brunet, M., Comeaux, J.,
736 Crouthamel, R., Gleason, B.E., Jr, Groisman, P.Y., Hersbach, H., Jones, P.D., Jónsson, T., Jourdain,
737 S., Kelly, G., Knapp, K.R., Kruger, A., Kubota, H., Lentini, G., Lorrey, A., Lott, N., Lubker, S.J.,
738 Luterbacher, J., Marshall, G.J., Maugeri, M., Mock, C.J., Mok, H.Y., Nordli, Ø., Rodwell, M.J.,
739 Ross, T.F., Schuster, D., Srncic, L., Valente, M.A., Vizi, Z., Wang, X.L., Westcott, N., Woollen, J.S.
740 and Worley, S.J. (2015), The International Surface Pressure Databank version 2. *Geosci. Data J.*, 2:
741 31-46. <https://doi.org/10.1002/gdj3.25>, 2015.

742 Dinan, T.: Projected increases in hurricane damage in the United States: The role of climate change and
743 coastal development, *Ecological Economics*, 138, 186–198,
744 <https://doi.org/10.1016/j.ecolecon.2017.03.034>, 2017.

745 Emanuel, K.: The Hurricane—Climate Connection, *Bull. Amer. Meteor. Soc.*, 89, ES10–ES20,
746 <https://doi.org/10.1175/BAMS-89-5-Emanuel>, 2008.

747 Emanuel, K.: Tropical cyclone activity downscaled from NOAA-CIRES Reanalysis, 1908–1958, *J. Adv.*
748 *Model. Earth Syst.*, 2, <https://doi.org/10.3894/JAMES.2010.2.1>, 2010.

749 Emanuel, K.: Will global warming make hurricane forecasting more difficult?, *Bull. Amer. Meteor. Soc.*,
750 98, 495–501, <https://doi.org/10.1175/BAMS-D-16-0134.1>, 2017.

751 Emanuel, K.: 100 Years of progress in tropical cyclone research, *Meteor. Monogr.*, 59, 15.1-15.68,
752 <https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0016.1>, 2018.

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

755 Emanuel, K.: Limitations of reanalyses for detecting tropical cyclone trends, *Nat. Clim. Chang.*, 14, 143–
756 145, <https://doi.org/10.1038/s41558-023-01879-y>, 2024.

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

760 Flaounas, E., Aragão, L., Bernini, L., Dafis, S., Doiteau, B., Flocas, H., Gray, S. L., Karwat, A.,
761 Kouroutzoglou, J., Lionello, P., Miglietta, M. M., Pantillon, F., Pasquero, C., Patlakas, P., Picornell,
762 M. Á., Porcù, F., Priestley, M. D. K., Reale, M., Roberts, M. J., Saaroni, H., Sandler, D.,
763 Scoccimarro, E., Sprenger, M., and Ziv, B.: A composite approach to produce reference datasets for
764 extratropical cyclone tracks: application to Mediterranean cyclones, *Weather Clim. Dynam.*, 4, 639–
765 661, <https://doi.org/10.5194/wcd-4-639-2023>, 2023.

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

769 Gergis, J., Ashcroft, L., and Whetton, P.: A historical perspective on Australian temperature extremes,
770 *Clim Dyn*, 55, 843–868, <https://doi.org/10.1007/s00382-020-05298-z>, 2020.

771 Han, Y. and Ullrich, P. A.: The System for Classification of Low-Pressure Systems (SyCLoPS): An All-
772 In-One objective framework for large-scale data sets, *J. Geophys. Res. Atmos.*, 130,
773 e2024JD041287, <https://doi.org/10.1029/2024JD041287>, 2025.

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

777 Hoarau, K., Bernard, J., and Chalonge, L.: Intense tropical cyclone activities in the northern Indian Ocean,
778 *Int. J. Climatol.*, 32, 1935–1945, <https://doi.org/10.1002/joc.2406>, 2012.

779 Hodges, K., Cobb, A., and Vidale, P. L.: How well are tropical cyclones represented in reanalysis
780 datasets?, *J. Climate*, 30, 5243–5264, <https://doi.org/10.1175/JCLI-D-16-0557.1>, 2017.

781 Horn, M., Walsh, K., Zhao, M., Camargo, S. J., Scoccimarro, E., Murakami, H., Wang, H., Ballinger, A.,
782 Kumar, A., Shaevitz, D. A., Jonas, J. A., and Oouchi, K.: Tracking scheme dependence of simulated
783 tropical cyclone response to idealized climate simulations, *J. Climate*, 27, 9197–9213,
784 <https://doi.org/10.1175/JCLI-D-14-00200.1>, 2014.

785 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White,
786 G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C.,
787 Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR
788 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc.*, 77, 437–472, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2), 1996.

790 Klotzbach, P. J. and Landsea, C. W.: Extremely intense hurricanes: Revisiting Webster et al. (2005) after
791 10 Years, *J. Climate*, 28, 7621–7629, <https://doi.org/10.1175/JCLI-D-15-0188.1>, 2015.

792 Klotzbach, P. J., Bell, M. M., Bowen, S. G., Gibney, E. J., Knapp, K. R., and Schreck, C. J.: Surface
793 pressure a more skillful predictor of normalized hurricane damage than maximum sustained wind,
794 *Bull. Amer. Meteor. Soc.*, 101, E830–E846, <https://doi.org/10.1175/BAMS-D-19-0062.1>, 2020.

795 Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., and Neumann, C. J.: The International Best
796 Track Archive for Climate Stewardship (IBTrACS): Unifying Tropical cyclone data, *Bull. Amer.*
797 *Meteor. Soc.*, 91, 363–376, <https://doi.org/10.1175/2009BAMS2755.1>, 2010.

798 Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., Schreck, C. J., and others: International best
799 track archive for climate stewardship (IBTrACS) project, version 4, NOAA National Centers for
800 Environmental Information, [data set] 10, <https://doi.org/10.25921/82ty-9e16>, 2018.

801 Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C.-H., Kossin, J., Mohapatra, M., Satoh,
802 M., Sugi, M., Walsh, K., and Wu, L.: Tropical cyclones and climate change assessment: part i:
803 Detection and attribution, *Bull. Amer. Meteor. Soc.*, 100, 1987–2007,
804 <https://doi.org/10.1175/BAMS-D-18-0189.1>, 2019.

805 Knutson, T., Camargo, S. J., Chan, J. C. L., Emanuel, K., Ho, C.-H., Kossin, J., Mohapatra, M., Satoh,
806 M., Sugi, M., Walsh, K., and Wu, L.: Tropical cyclones and climate change assessment: Part II:
807 Projected response to anthropogenic warming, *Bull. Amer. Meteor. Soc.*, 101, E303–E322,
808 <https://doi.org/10.1175/BAMS-D-18-0194.1>, 2020.

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

812 Knutson, T. R., Sirutis, J. J., Zhao, M., Tuleya, R. E., Bender, M., Vecchi, G. A., Villarini, G., and Chavas,
813 D.: Global projections of intense tropical cyclone activity for the Late Twenty-First Century from
814 dynamical downscaling of CMIP5/RCP4.5 scenarios, *J. Climate*, 28, 7203–7224,
815 <https://doi.org/10.1175/JCLI-D-15-0129.1>, 2015.

816 Kossin, J. P., Knapp, K. R., Olander, T. L., and Velden, C. S.: Global increase in major tropical cyclone
817 exceedance probability over the past four decades, *Proc. Natl. Acad. Sci. USA*, 117, 11975–11980,
818 <https://www.pnas.org/doi/full/10.1073/pnas.1920849117>, 2020.

819 Kunze, S.: Unraveling the effects of tropical cyclones on economic sectors worldwide: Direct and indirect
820 impacts, *Environ Resource Econ* 78, 545–569, <https://doi.org/10.1007/s10640-021-00541-5>, 2021.

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

824 Lalouaux, P., de Boisseson, E., Balsaseda, M., Bidlot, J.-R., Broennimann, S., Buizza, R., Dalhgren, P.,
825 Dee, D., Haimberger, L., Hersbach, H., Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier,
826 E., and Schepers, D.: CERA-20C: A coupled reanalysis of the Twentieth Century, *J. Adv. Model.*
827 *Earth Syst.*, 10, 1172–1195, <https://doi.org/10.1029/2018MS001273>, 2018.

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

830 Landsea, C. W., Harper, B. A., Hoarau, K., and Knaff, J. A.: Can we detect trends in extreme tropical
831 cyclones?, *Science*, 313, 452–454, <https://doi.org/10.1126/science.1128448>, 2006.

832 Landsea, C. W., Glenn, D. A., Bredemeyer, W., Chenoweth, M., Ellis, R., Gamache, J., Hufstetler, L.,
833 Mock, C., Perez, R., Prieto, R., Sánchez-Sesma, J., Thomas, D., and Woolcock, L.: A Reanalysis of
834 the 1911–20 Atlantic hurricane database, *J. Climate*, 21, 2138–2168,
835 <https://doi.org/10.1175/2007JCLI1119.1>, 2008.

836 Landsea, C. W., Vecchi, G. A., Bengtsson, L., and Knutson, T. R.: Impact of Duration Thresholds on
837 Atlantic tropical cyclone counts, *J. Climate*, 23, 2508–2519,
838 <https://doi.org/10.1175/2009JCLI3034.1>, 2010.

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

841 Lee, R., Chen, L., and Ren, G.: A comparison of East-Asia landfall tropical cyclone in recent reanalysis
842 datasets--before and after satellite era, *Front. Earth Sci.* 10:1026945,
843 <https://doi.org/10.3389/feart.2022.1026945>, 2023.

844 Lee, T.-C., Knutson, T. R., Nakaegawa, T., Ying, M., and Cha, E. J.: Third assessment on impacts of
845 climate change on tropical cyclones in the Typhoon Committee Region – Part I: Observed changes,
846 detection and attribution, *Trop. Cyclone Res. Rev.*, 9, 1–22,
847 <https://doi.org/10.1016/j.tcr.2020.03.001>, 2020.

848 Lenzen, M., Malik, A., Kenway, S., Daniels, P., Lam, K. L., and Geschke, A.: Economic damage and
849 spillovers from a tropical cyclone, *Nat. Hazards Earth Syst. Sci.*, 19, 137–151,
850 <https://doi.org/10.5194/nhess-19-137-2019>, 2019.

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

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

857 Li, Y., Tang, Y., Li, X., Song, X., and Wang, Q.: Recent increase in the potential threat of western North
858 Pacific tropical cyclones, *npj Clim Atmos Sci*, 6, 1–8, <https://doi.org/10.1038/s41612-023-00379-2>,
859 2023.

860 Malakar, P., Kesarkar, A. p., Bhate, J. n., Singh, V., and Deshamukhya, A.: Comparison of reanalysis
861 data sets to comprehend the evolution of tropical cyclones over North Indian Ocean, *Earth Space*
862 *Sci.*, 7, e2019EA000978, <https://doi.org/10.1029/2019EA000978>, 2020.

863 Mann, M. E., Sabbatelli, T. A., and Neu, U.: Evidence for a modest undercount bias in early historical
864 Atlantic tropical cyclone counts, *Geophys. Res. Lett*, 34, L22707,
865 <https://doi.org/10.1029/2007GL031781>, 2007.

866 Mitchell, C. L.: The West Indian hurricane of September 10–20, 1928, *Mon. Wea. Rev.*, 56, 347–350,
867 [https://doi.org/10.1175/1520-0493\(1928\)56<347:TWIHOS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1928)56<347:TWIHOS>2.0.CO;2), 1928.

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

870 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
871 strength and area from landfalling tropical cyclones over the North Indian and western North Pacific
872 oceans under increased CO₂ conditions, *Weather Clim. Extrem.*, 41, 100581,
873 <https://doi.org/10.1016/j.wace.2023.100581>, 2023.

874 Moore, G. W. K. and Babij, M.: Iceland’s Great Frost Winter of 1917/1918 and its representation in
875 reanalyses of the twentieth century, *Q.J.R. Meteorol. Soc.*, 143, 508–520,
876 <https://doi.org/10.1002/qj.2939>, 2017.

877 Murakami, H.: Tropical cyclones in reanalysis data sets, *Geophys. Res. Letts*, 41, 2133–2141,
878 <https://doi.org/10.1002/2014GL059519>, 2014.

879 Murakami, H. and Wang, B.: Patterns and frequency of projected future tropical cyclone genesis are
880 governed by dynamic effects, *Commun Earth Environ*, 3, 1–10, [https://doi.org/10.1038/s43247-022-](https://doi.org/10.1038/s43247-022-00410-z)
881 [00410-z](https://doi.org/10.1038/s43247-022-00410-z), 2022.

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

884 Parker, W. S.: Reanalyses and Observations: What’s the difference?, *Bull. Amer. Meteor. Soc.*, 97, 1565–
885 1572, <https://doi.org/10.1175/BAMS-D-14-00226.1>, 2016.

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

888 Qin, L., Zhu, L., Liu, B., Li, Z., Tian, Y., Mitchell, G., Shen, S., Xu, W., and Chen, J.: Global expansion
889 of tropical cyclone precipitation footprint, *Nat Commun*, 15, 4824, [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-024-49115-1)
890 [024-49115-1](https://doi.org/10.1038/s41467-024-49115-1), 2024.

891 Raavi, P. H. and Walsh, K. J. E.: Basinwise statistical analysis of factors limiting tropical storm formation
892 from an initial tropical circulation, *J. Geophys. Res. Atmos.*, 125, e2019JD032006,
893 <https://doi.org/10.1029/2019JD032006>, 2020a.

894 Raavi, P. H. and Walsh, K. j. e.: Sensitivity of tropical cyclone formation to resolution-dependent and
895 independent tracking schemes in High-Resolution Climate Model simulations, *Earth Space Sci.*, 7,
896 e2019EA000906, <https://doi.org/10.1029/2019EA000906>, 2020b.

897 Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vanni re, B., Mecking, J., Haarsma, R.,
898 Bellucci, A., Scoccimarro, E., Caron, L.-P., Chauvin, F., Terray, L., Valcke, S., Moine, M.-P.,
899 Putrasahan, D., Roberts, C. D., Senan, R., Zarzycki, C., Ullrich, P., Yamada, Y., Mizuta, R., Kodama,
900 C., Fu, D., Zhang, Q., Danabasoglu, G., Rosenbloom, N., Wang, H., and Wu, L.: Projected Future
901 Changes in Tropical Cyclones Using the CMIP6 HighResMIP Multimodel Ensemble. *Geophys. Res.*
902 *Let.*, 47: e2020GL088662. <https://doi.org/10.1029/2020GL088662>

903 Schreck, C. J., Knapp, K. R., and Kossin, J. P.: The impact of best track discrepancies on global tropical
904 cyclone climatologies using IBTrACS, *Mon. Wea. Rev.*, 142, 3881–3899,
905 <https://doi.org/10.1175/MWR-D-14-00021.1>, 2014.

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

908 Slivinski, L. C.: Historical Reanalysis: What, How, and Why?, *J. Adv. Model. Earth Syst.*, 10, 1736–
909 1739, <https://doi.org/10.1029/2018MS001434>, 2018.

910 Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R.,
911 Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Br nnimann, S., Brunet,
912 M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Dom nguez-Castro, F., Freeman, J. E.,
913 Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F. L., Lee, T.-C.,
914 Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. W. K., Przybylak, R., Pudmenzky,
915 C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L.,
916 Wilkinson, C., Wood, K., and Wyszyński, P.: Towards a more reliable historical reanalysis:
917 Improvements for version 3 of the Twentieth Century Reanalysis system, *Q.J.R. Meteorol. Soc.*,
918 145, 2876–2908, <https://doi.org/10.1002/qj.3598>, 2019.

919 Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P.,
920 Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D. C.,
921 Kennedy, J. J., Ashcroft, L., Br nnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A.,
922 Dom nguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C.,
923 Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz,
924 B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: An evaluation of the
925 performance of the Twentieth Century Reanalysis Version 3, *J. Climate*, 34, 1417–1438,
926 <https://doi.org/10.1175/JCLI-D-20-0505.1>, 2021.

927 Torn, R. D. and Snyder, C.: Uncertainty of tropical cyclone best-track information, *Wea. Forecasting*, 27,
928 715–729, <https://doi.org/10.1175/WAF-D-11-00085.1>, 2012.

929 Tory, K. J., Dare, R. A., Davidson, N. E., McBride, J. L., and Chand, S. S.: The importance of low-
930 deformation vorticity in tropical cyclone formation, *Atmos. Chem. Phys.*, 13, 2115–2132,
931 <https://doi.org/10.5194/acp-13-2115-2013>, 2013.

932 Truchelut, R. E. and Hart, R. E.: Quantifying the possible existence of undocumented Atlantic warm-
933 core cyclones in NOAA/CIRES 20th Century Reanalysis data, *Geophys. Res. Lett.*, 38, L08811,
934 <https://doi.org/10.1029/2011GL046756>, 2011.

935 Truchelut, R. E., Hart, R. E., and Luthman, B.: Global identification of previously undetected Pre-
936 satellite-era tropical cyclone candidates in NOAA/CIRES Twentieth-Century Reanalysis Data, *J.*
937 *Appl. Meteor. Climatol.*, 52, 2243–2259, <https://doi.org/10.1175/JAMC-D-12-0276.1>, 2013.

938 Tu, S., Xu, J., Chan, J. C. L., Huang, K., Xu, F., and Chiu, L. S.: Recent global decrease in the inner-core
939 rain rate of tropical cyclones, *Nat Commun*, 12, 1948, <https://doi.org/10.1038/s41467-021-22304-y>,
940 2021.

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

943 Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., and Reed, K. A.:
944 TempestExtremes v2.1: a community framework for feature detection, tracking, and analysis in
945 large datasets, *Geosci. Model Dev.*, 14, 5023–5048, <https://doi.org/10.5194/gmd-14-5023-2021>,
946 2021.

947 Vecchi, G. A., and T. R. Knutson.: On Estimates of Historical North Atlantic Tropical Cyclone Activity.
948 *J. Climate*, 21, 3580 – 3600, <https://doi.org/10.1175/2008JCLI2178.1>, 2008.

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

951 Wang, X. L., Feng, Y., and Swail, V.: North Atlantic wave height trends as reconstructed from the 20th
952 century reanalysis, *Geophys. Res. Lett.*, 39, L18705, <https://doi.org/10.1029/2012GL053381>, 2012.

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

956 Ye, G., Jeremy Cheuk-Hin, L., Dong, W., Xu, Ji., LI, W., Qian, W., Kong, H., and Zhang, B.: A
957 Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth Century (RGTracks-
958 20C) [dataset], <https://doi.org/10.5281/zenodo.14411917>, 2024.

959 Yeasmin, A., Chand, S., and Sultanova, N.: Reconstruction of tropical cyclone and depression proxies
960 for the South Pacific since the 1850s, *Weather Clim. Extrem.*, 39, 100543,
961 <https://doi.org/10.1016/j.wace.2022.100543>, 2023.

962 Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., and Chen, D.: An overview of the China
963 meteorological administration tropical cyclone database, *J. Atmos. Oceanic Technol.*, 31, 287–301,
964 <https://doi.org/10.1175/JTECH-D-12-00119.1>, 2014.

965 Yoshida, K., Sugi, M., Mizuta, R., Murakami, H., and Ishii, M.: Future changes in tropical cyclone
966 activity in high-resolution large-ensemble simulations, *Geophys. Res. Lett.*, 44, 9910–9917,
967 <https://doi.org/10.1002/2017GL075058>, 2017.

968 Zarzycki, C. M. and Ullrich, P. A.: Assessing sensitivities in algorithmic detection of tropical cyclones
969 in climate data, *Geophys. Res. Lett.*, 44, 1141–1149, <https://doi.org/10.1002/2016GL071606>, 2017.

970 Zhang, B., Zhang, R., Pinker, R. T., Feng, Y., Nie, C., and Guan, Y.: Changes of tropical cyclone activity
971 in a warming world are sensitive to sea surface temperature environment, *Environ. Res. Lett.*, 14,
972 124052, <https://doi.org/10.1088/1748-9326/ab5ada>, 2019.

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

975 Zhang, X., Duan, Y., Wang, Y., Wei, N., and Hu, H.: A high-resolution simulation of Supertyphoon
976 Rammasun (2014)—Part I: Model verification and surface energetics analysis, *Adv. Atmos. Sci.*,
977 34, 757–770, <https://doi.org/10.1007/s00376-017-6255-7>, 2017.

978 Zhao, M. and Held, I. M.: An analysis of the effect of global warming on the intensity of Atlantic
979 hurricanes using a GCM with statistical refinement, *J. Climate*, 23, 6382–6393,
980 <https://doi.org/10.1175/2010JCLI3837.1>, 2010.

981 Zhu, L. and Quiring, S. M.: Exposure to precipitation from tropical cyclones has increased over the
982 continental United States from 1948 to 2019, *Commun Earth Environ*, 3,
983 312 <https://doi.org/10.1038/s43247-022-00639-8>, 2022.

984