

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

3

4 Guiling Ye^{1,2,3}, Jeremy Cheuk-Hin Leung^{2*}, Wenjie Dong^{1,3*}, Jianjun Xu⁵, Weijing Li⁶, Weihong
5 Qian⁷, Hoiio Kong⁴, Banglin Zhang^{2,8}

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

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 College of Atmospheric Science, Lanzhou University, Lanzhou, China

17

18

19

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

22 **Abstract**

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

36 **1. Introduction**

37 Tropical cyclones (TCs), also known as hurricanes or typhoons, are intense weather systems
38 that form over tropical and subtropical oceans and can cause severe disasters over the coastal regions
39 and even inland areas (Qin et al., 2024; Zhu and Quiring, 2022). Globally, approximately 80 TCs
40 are generated each year (Emanuel, 2018). As one of the most destructive weather systems
41 (Bloemendaal et al., 2022; Dinan, 2017; Emanuel, 2017), TCs significantly impact society and the
42 economy (Kunze, 2021; Lenzen et al., 2019; Noy, 2016). These impacts are expected to be
43 exacerbated by climate change in the future (Chan, 2023; Hassanzadeh et al., 2020; Knutson et al.,
44 2020; Moon et al., 2023; Murakami and Wang, 2022; Yamaguchi et al., 2020). Therefore, research
45 on TCs has become increasingly vital in climate change and prediction (Bhatia et al., 2019; Chan,
46 2019; Lanzante, 2019; Moon et al., 2019; Sharmila and Walsh, 2018; Zhang et al., 2019). However,
47 past variability of TC activity and underlying mechanisms remains challenging due to incomplete
48 early historical TC observation records, which may lead to controversies (Chan et al., 2022a, b;
49 Knutson et al., 2019; Lee et al., 2020).

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

66 Given the limitations of historical TC records, a promising approach is to utilize reanalysis for
67 TC identification (Li et al., 2024; Truchelut and Hart, 2011). Reanalysis combines historical
68 observational data with modern numerical weather models to produce comprehensive, continuous
69 datasets of global atmospheric conditions that adhere to physical principles (Compo et al., 2011;
70 Kalnay et al., 1996; Parker, 2016; Slivinski, 2018). The Twentieth Century Reanalysis (20CR)
71 (Compo et al., 2011), provided by the National Oceanic and Atmospheric Administration (NOAA),
72 is a global reanalysis dataset that covers the longest period among all other reanalyses. The 20CR
73 was designed for long-term analyses from individual extreme weather events to climate variability,

74 and has been applied to a wide range of studies, including those on wave height, storm surge,
75 Madden-Julian Oscillations, and TCs (Chand et al., 2022; Cid et al., 2017; Gergis et al., 2020; Lee
76 et al., 2023; Leung et al., 2022; Moore and Babij, 2017; Slivinski et al., 2019; Truchelut et al., 2013;
77 Wang et al., 2012). The fact that the 20CR only assimilates surface pressure and sea level pressure
78 fields, instead of other observations such as satellites and aircraft, makes it less sensitive to the
79 temporal inhomogeneity of observations (Slivinski et al., 2019, 2021).

80 Several independent studies have documented the feasibility of reproducing the characteristics
81 of some historical TC events based on the 20CR (Emanuel, 2010; Lee et al., 2023; Slivinski et al.,
82 2019; Truchelut et al., 2013; Truchelut and Hart, 2011). For example, following Emanuel (Emanuel,
83 2010), who first expanded and revised TC climatology based on the 20CR, (Truchelut and Hart,
84 2011) employed the 20CR to identify previously unknown TCs in the Atlantic and demonstrated
85 that the 20CR can accurately describe large-scale TC thermodynamic structure. Recently, Truchelut
86 et al. (2013) noted that the 20CR has the ability to investigate TC events that were previously
87 undetected in the pre-satellite era. Compared to other reanalyses, the 20CR well captures the
88 intensity of the 1915 Galveston hurricane (Slivinski et al., 2019) and also offers a more accurate
89 representation of landfalling TCs in East Asia (Lee et al., 2023). These previous studies have
90 demonstrated the effectiveness of the 20CR as a tool for characterizing historical TCs (Emanuel,
91 2010; Truchelut et al., 2013; Truchelut and Hart, 2011). Taking advantage of the 20CR, some
92 researchers have extracted the century-long TC information from the reanalysis product (Chand et
93 al., 2022; Lee et al., 2023; Yeasmin et al., 2023), suggesting its potential as a tool for studying
94 historical changes in TCs under anthropogenic climate change.

95 While the 20CR has been applied to studying the relationship between historical climate change
96 and TC variability, the primary focus was mostly on the TC occurrence frequency, and little attention
97 was given to other TC metrics such as intensity, duration, and location. More importantly, to date,
98 there is no publicly available reanalysis-based global TC dataset covering a century-long period.
99 Therefore, the main objective of this study is to extract TC information (including location, intensity,
100 and lifetime) from the 20CR and reconstruct a historical global TC track dataset spanning 1850–
101 2014. The produced dataset is named the Reanalysis-Based Global Tropical Cyclone Tracks Dataset
102 for the Twentieth Century (RGTracks-20C) and is open to the public for research use. This paper
103 first introduces the production details of the RGTracks-20C and then discusses the validity, key
104 strengths, and usage notes of the datasets. We anticipate that the RGTracks-20C will provide
105 valuable insights into the changing patterns of historical TC activity, improving our understanding
106 of the response of TCs to climate change.

107 **2. Data and methods**

108 **2.1 Data**

109 The primary objective of this study was to reconstruct a 20th century global TC dataset from
110 the 20th Century Reanalysis version 3 (20CRv3) (Slivinski et al., 2019, 2021), the latest version of

111 the 20CR produced by NOAA. Then, the validity of the reconstructed 20th century global TC data
112 was evaluated based on the observed TC records, i.e., the International Best Track Archive for
113 Climate Stewardship (IBTrACS) dataset (Knapp et al., 2010).

114 **2.1.1 20th Century Reanalysis**

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

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

135 **2.1.2 IBTrACS**

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

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

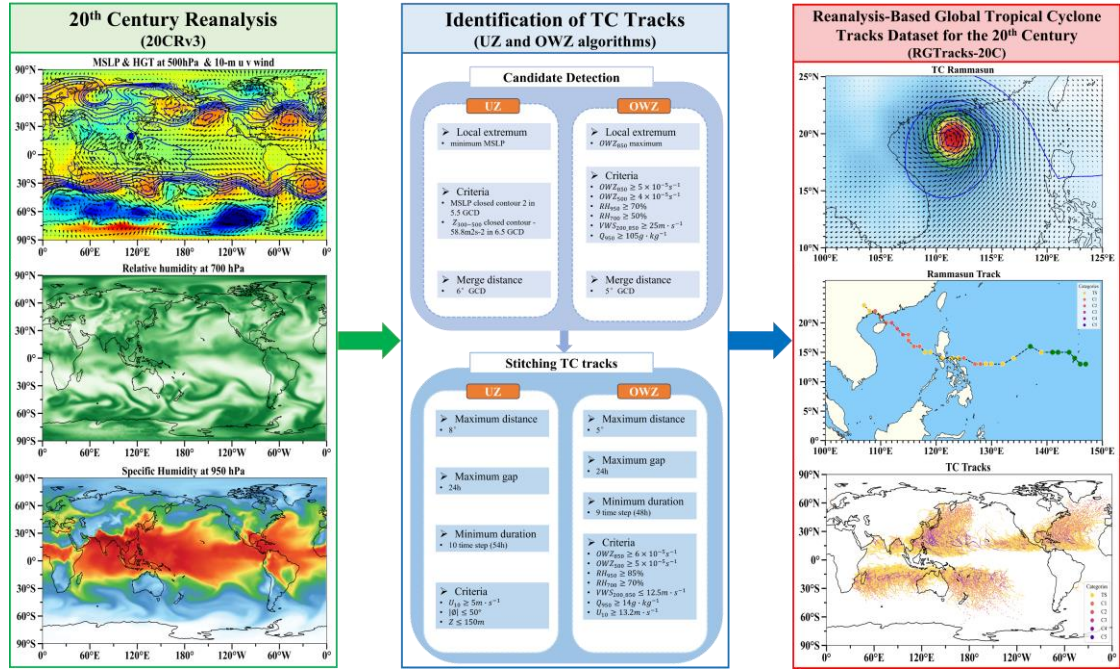
153 Although the IBTrACS has time coverage dating back to the early 20th century, we utilize the
154 data only for the post-satellite period (1979–2014) due to the early data incompleteness issues
155 (Chang and Guo, 2007; Lee et al., 2020; Truchelut et al., 2013). Given that the IBTrACS is the most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

159 **2.2 Production of the RGTracks-20C**

160 **2.2.1 Procedure**

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

176



177
178 **Figure 1:** Schematic diagram showing the production of the RGTracks-20C from the 20CRv3 based on the
179 UZ and OWZ tracking algorithms. Variables shown include U10: 10-m wind speed, ϕ : latitude, z : altitude,
180 GCD: great circle distance.

181 2.2.2 TC tracker

182 i. OWZ Tracker

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

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

$$191 \quad OWZ = \text{sgn}(f) \times (\zeta + f) \times \max \left[\frac{\zeta^2 - (E^2 + F^2)}{\zeta^2}, 0 \right] \quad (1)$$

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

$$195 \quad E = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \quad (2)$$

$$196 \quad F = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \quad (3)$$

197 **First step: Candidate detection.**

198 The OWZ tracker begins by identifying local maxima of OWZ at 850 hPa. Any candidate with
199 a stronger OWZ maximum within 5° of great circle distance (GCD) is excluded. Next, only

200 candidates that meet the six initial threshold conditions shown in Table 1 within a 2° GCD of the
 201 identified maximum are retained. Based on the information provided in Table 1, besides the required
 202 minimum threshold values for the OWZ parameter at 850 *hPa* and 500 *hPa*, additional dynamical
 203 and thermodynamic parameters related to TC formation are taken into account. These parameters
 204 include the maximum threshold for the wind vector difference (vertical wind shear) between 850
 205 *hPa* and 200 *hPa*, as well as the relative humidity at 950 *hPa* and 700 *hPa*, and the minimum
 206 threshold for the specific humidity at 950 *hPa*. This step primarily aims to identify grid points that
 207 contain essential components of a storm. Subsequently, neighboring grid points are grouped together
 208 to define potential TCs.

209 **Second step: Stitching TC tracks.**

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

215

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

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

219

220

221 **ii. UZ tracker**

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

224 **First step: Candidate detection.**

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

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

229 2. The geopotential thickness between 300 and 500 *hPa* (denoted as $Z_{300-500}$) must decrease by
 230 $58.8 \text{ m}^2 \text{ s}^{-2}$ over a distance of 6.5° GCD from the maximum center of $Z_{300-500}$ within 1° GCD of the
 231 center of minimum SLP.

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

233 **Second step: Stitching TC tracks.**

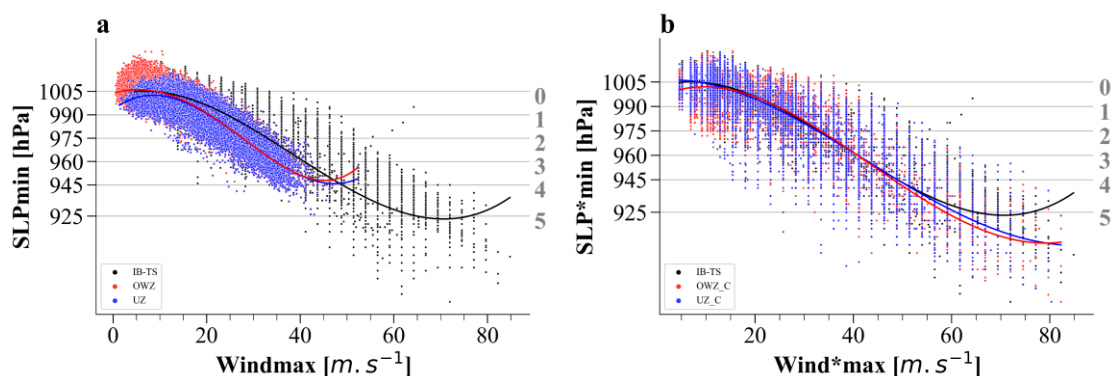
234 The candidates are subsequently linked in time to create paths, ensuring a maximum distance
 235 of 8° GCD between candidates. Each path must last for at least 54 hours without gaps longer than
 236 24 hours. Additionally, ten 6-hourly time steps (equivalent to 54 hours) must satisfy the following
 237 thresholds: wind speed at 10m ≥ 10 m/s and $z \leq 150$ m (where z represents the altitude), and the
 238 storm must form between 0° and 50° .

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

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

247 2.2.3 Bias Correction of TC intensity

248 Given the low horizontal resolution in the 20CRv3, TC intensities derived directly from the
 249 reanalysis generally underestimated compared to observations (Fig. 2a) (Bourdin et al., 2022;
 250 Roberts et al., n.d.). To address this issue, a quantile mapping bias correction, similar to the method
 251 used by Zhao and Held (2010), was applied to adjust for the TC intensity bias within the dataset.
 252 The main idea is to fit the 20CRv-derived TC intensity distributions, either probability distribution
 253 functions (PDFs) or cumulative distribution functions (CDFs), to the observed distributions. This
 254 method has demonstrated significant efficacy in enhancing the accuracy of TC intensity within
 255 models or reanalyses (Faranda et al., 2023; Yoshida et al., 2017). This adjustment resulted in a
 256 wind-pressure relationship in RGTracks-20C that aligns more closely with observational data (Fig.
 257 2b).



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

264

265 2.3 Verification of RGTracks-20C

266 2.3.1 Tracks matching

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

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

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

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

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

283

284 2.3.2 Track verification

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

$$289 \quad POD = \frac{H}{H + M} \quad (4)$$

290

$$291 \quad FAR = \frac{FAs}{H + FAs} \quad (5)$$

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

295 2.4 Definitions

296 2.4.1 TC intensity

297 In assessing the TC intensity, SLP_{min} and $WIND_{max}$ are two commonly used metrics in TC
298 research. However, because $WIND_{max}$ in both observations and reanalysis exhibits relatively higher
299 uncertainties (Bourdin et al., 2022; Chavas et al., 2017; Knapp et al., 2010; Knutson et al., 2015;

300 Schreck et al., 2014), this study opted to use SLP_{min} as the only indicator of TC intensity when
 301 verifying the validity of RGTracks-20C. Nevertheless, $WIND_{max}$ of detected TCs is also provided in
 302 the RGTracks-20C (Table 2) as a reference for researchers who wish to use and improve the dataset,
 303 though it is not discussed in the paper.

304

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

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

308

309 **Storm categories:** the Saffir-Simpson Hurricane Scale (SSHS) from 1 to 5 based on their peak
 310 1-minute wind speed at 10 meters above the surface. In this study, given the significant uncertainties
 311 in WINDmax due to differences between institutions and the limitations of model simulation
 312 capabilities (Chavas et al., 2017; Klotzbach et al., 2020; Knutson et al., 2015), we have chosen to
 313 classify based on SLP_{min} , following the definition of (Klotzbach et al., 2020).

314 2.4.2 Basins

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

318 2.4.3 TC days

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

$$322 \quad TC \text{ days} = \frac{1}{4} \sum_{i=0}^n L_i \quad (6)$$

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

324 3. Results and discussion

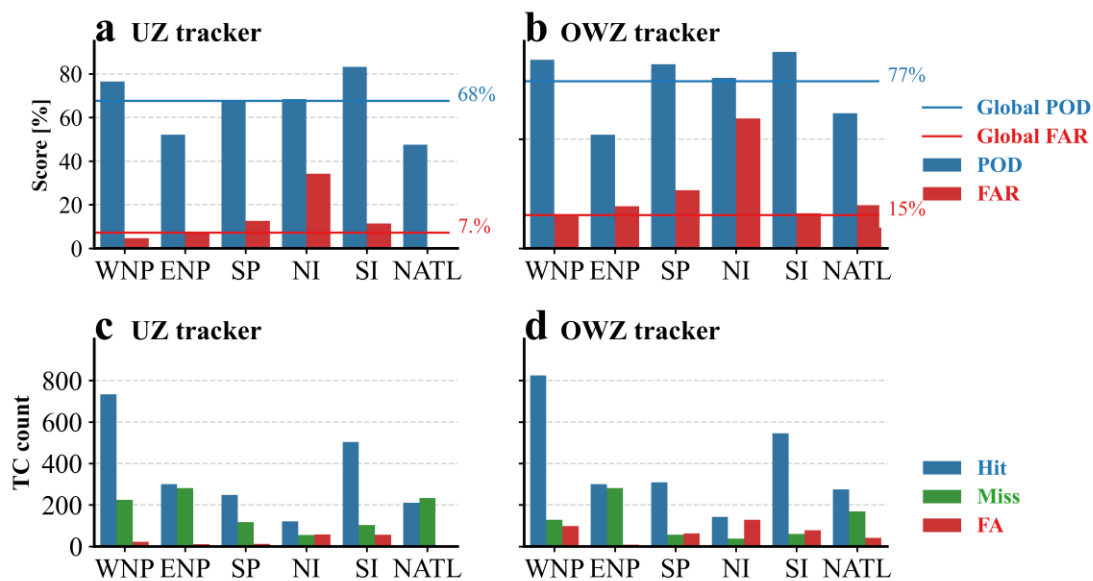
325 3.1 Data Records

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

336 3.2. Validity of trackers

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

342 The POD and FAR of TCs identified by the UZ and OWZ trackers are calculated to assess the
343 ability of the trackers to detect TCs from the 20CRv3 globally and across six basins (see Track
344 verification). Globally, the overall POD and FAR of TCs detected by the UZ tracker are 68% and
345 7% (Fig. 3a), while those by the OWZ tracker are 77% and 15%, respectively (Fig. 3b). Detailed
346 comparisons of each component of POD and FAR, including the number of hits, false alarms, and
347 misses, are provided in Supplementary Sect. S1 and Fig. S1.



348

349 **Figure 3: Accuracy of TC number detection of the RGTracks-20C. a–b, POD (blue bars and line, unit: %)**
350 **and FAR (red bars and line, unit: %) for TC number detected by the UZ (a) and OWZ (b) trackers in each**
351 **basin (bars), compared to the global mean (lines). Blue and red horizontal lines denote the POD and FAR over**
352 **the globe. c–d, same as a–b, except for the number of hits (blue bars), misses (green bars), and false alarms**
353 **(red bars) detected by the UZ (c) and OWZ (d) trackers.**
354

355 For each basin, the distributions of the POD of TCs (Figs. 3a–b) and the number of hits (Figs.
356 3c–d) between the two trackers show high similarities. Specifically, both trackers report higher POD
357 values in the SI (90% for OWZ tracker, 83% for UZ tracker), WNP (86% for OWZ tracker, 77% for
358 UZ tracker), and SP (84% for OWZ tracker, 68% for UZ tracker), followed by the NI (78% for OWZ
359 tracker, 68% for UZ tracker). Lower POD values are observed in the NATL (62% for OWZ tracker,
360 48% for UZ tracker) and the ENP (52% for both OWZ and UZ trackers). Similarly, the largest
361 number of TC hits is observed in the WNP (824 for OWZ tracker, 733 for UZ tracker) and SI (543
362 for OWZ tracker, 503 for UZ tracker), followed by the ENP, SP and NATL, each with approximately
363 200–300 TCs, and the NI with fewer than 200 TCs.

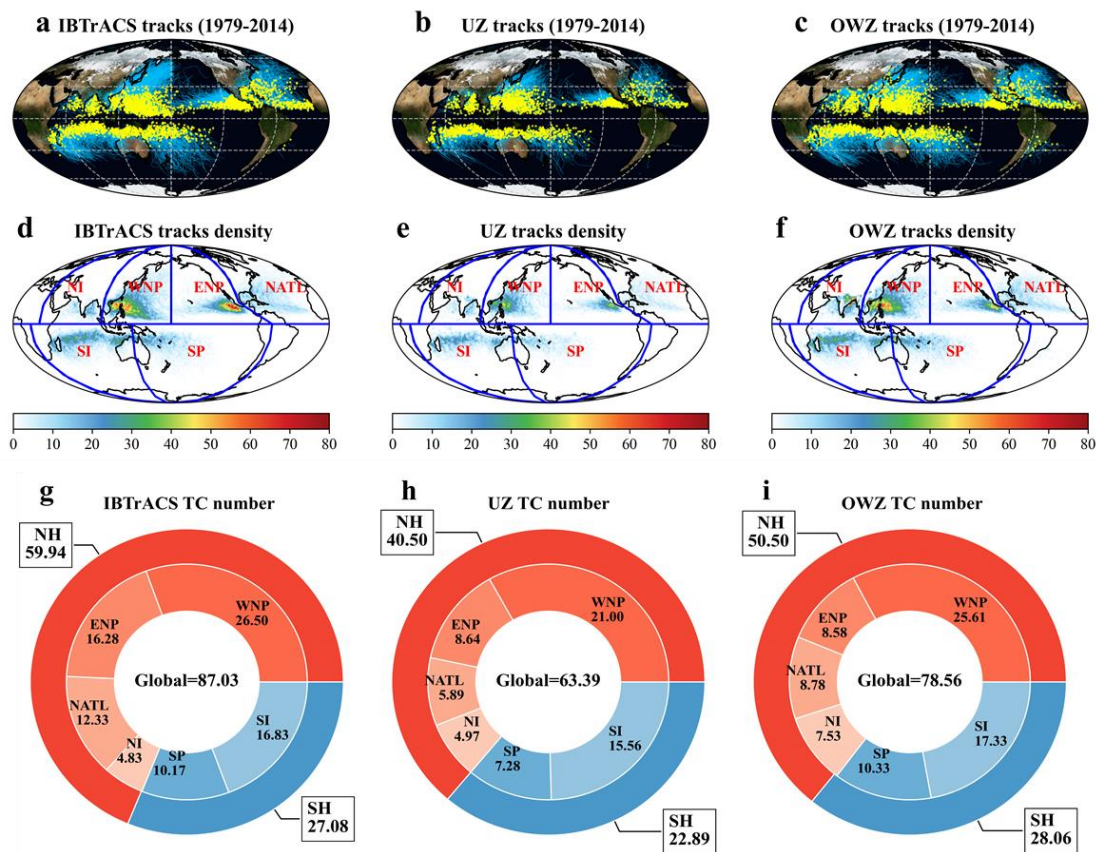
364 The FAR of TCs (Figs. 3a–b), and the number of false alarms (FAs) and misses (Figs. 3c–d)
365 vary between the two trackers. The UZ tracker exhibits FARs below 15% across all basins except
366 the NI. Notably, in the ENP and NATL, the FAR of TCs is below the global average of 7%, with
367 the number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

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

382 **3.3 Climatology of TC activity**

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

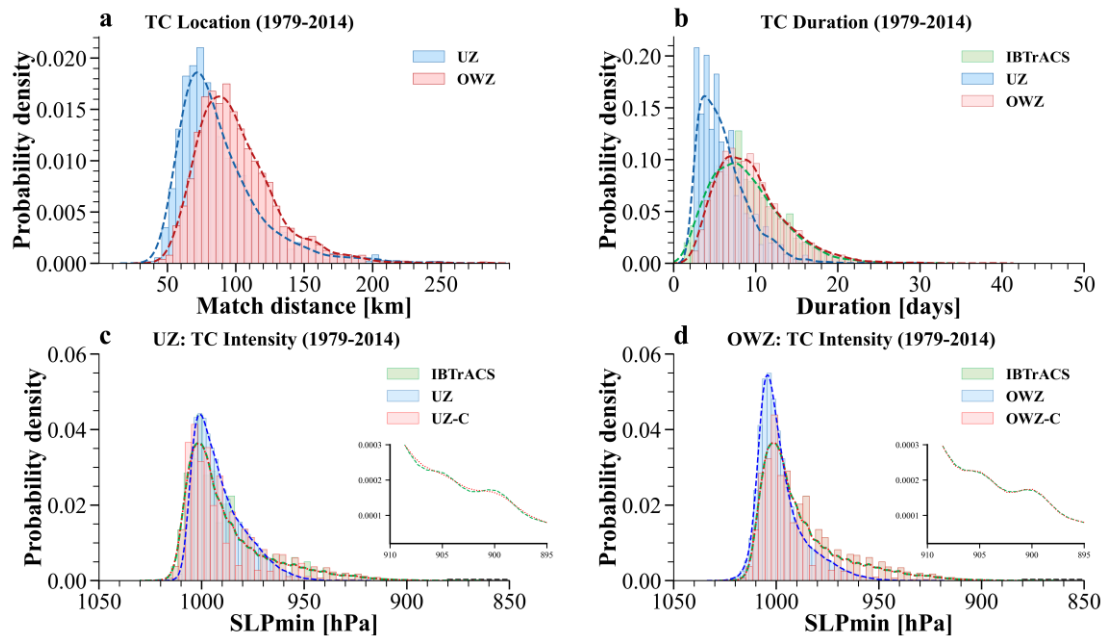
386 In terms of climatology, the RGTracks-20C is able to capture the major spatial patterns of TC
 387 genesis locations and track density over most ocean basins (Figs. 4a–f), indicating its effectiveness
 388 in reproducing the spatial distribution of historically observed TCs. The annual mean TC numbers
 389 in most ocean basins detected by the UZ and OWZ trackers are consistent with observations (Figs.
 390 4g–i). The OWZ tracker especially captures the observed annual mean TC number in the NWP, SI,
 391 and SP well, with discrepancies ranging from -0.48 to 0.89 . Notably, the UZ tracker also accurately
 392 estimates observed annual mean TC number in the NI, demonstrating a relatively small error (4.83
 393 versus 4.97) between the two. However, the UZ and OWZ trackers estimate the annual mean number
 394 of TCs to be 63.39 and 78.56 , respectively (Figs. 4h–i), which are relatively lower than the observed
 395 values (87.03 , Fig. 4g). The main reason for the global underestimation compared to IBTrACS is
 396 the discrepancies in the ENP and NATL, of which the reasons are discussed in Supplementary Sects.
 397 S2.1–2.2. Despite the underestimations in individual basins, the overall TC detection rates resemble
 398 previous publications that aimed to extract TCs from higher-quality reanalyses (Bourdin et al., 2022;
 399 Murakami, 2014). This result verifies the RGTracks-20C’s ability to reproduce the climatology of
 400 the TC number globally and in most basins.



401
 402 **Figure 4: TC genesis locations, tracks, and annual average number from IBTrACS and RGTracks-20C.** a–c,
 403 TC genesis locations (yellow dots) and tracks (blue lines) from IBTrACS (a), and RGTracks-20C using the UZ
 404 (b) and OWZ (c) trackers. d–f, TC tracks density (shading, number of TC occurrence per $1^\circ \times 1^\circ$ latitude-
 405 longitude grid box, 1979–2014) from IBTrACS (c), and RGTracks-20C using the UZ (e) and OWZ (f) trackers.
 406 g–i, mean number of TCs per year globally and for the six basins from IBTrACS (g), and RGTracks-20C using
 407 the UZ (h) and OWZ (i) trackers.

408

409 We further evaluate the accuracies of detected TC tracks in the RGTracks-20C by comparing
410 the arc length of TC tracks between RGTracks-20C and IBTrACS. Results indicate that the global
411 TC location errors range from 10 to 300 km, with the majority between 50–100 km for the UZ
412 tracker and 75–125 km for the OWZ tracker (Fig. 5a). Additionally, the peak errors for both trackers
413 are below 100 km, with the UZ and OWZ trackers showing peak values of approximately 75 km
414 and 95 km, respectively. These findings are consistent across all basins (Fig. 6a). Given that the
415 lower limit of the average TC location error expected from the coarse horizontal resolution of the
416 20CRv3 (1 degree×1 degree) is approximately 100 km, the above-mentioned small mean values of
417 TC location biases confirm that the RGTracks-20C is capable of reproducing most observed TC
418 tracks and locations.
419



420

421 **Figure 5: Distribution of TC characteristics on the IBTrACS and RGTracks-20C. a, Distribution of the mean**
422 **TC location error from 1979–2014 (unit: km) between IBTrACS and the RGTracks-20C by the UZ (blue) and**
423 **OWZ (red) algorithms. b, TC duration (unit: days) from 1979 to 2014 in IBTrACS (green) and the RGTracks-**
424 **20C by the UZ (blue) and OWZ (red) algorithms. c, same as (b), but for TC intensity (SLP_{min} , unit: hPa),**
425 **based on the UZ tracker, before (blue) and after (red) bias correction. d, same as (c), but for the OWZ tracker.**
426 **(UZ: UZ tracker, OWZ: OWZ tracker. UZ-C and OWZ-C represent bias-corrected results for the UZ and**
427 **OWZ trackers, respectively.)**

428

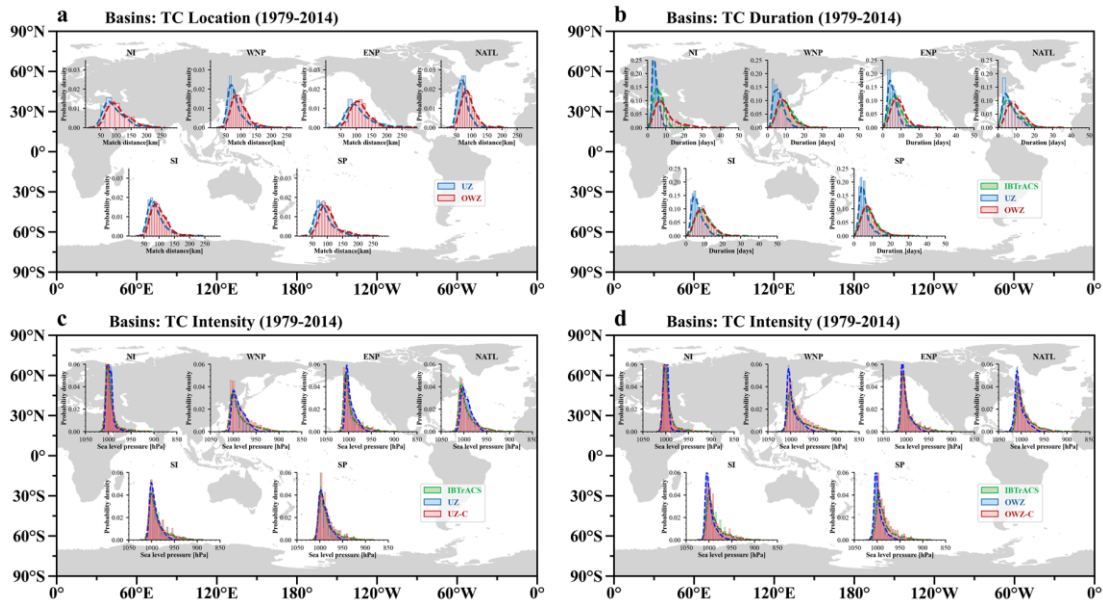


Figure 6: As in Fig. 5, but for six individual basins.

The duration and intensity of TCs are crucial in climate change research, as global warming may lead to stronger and longer-lasting TCs (Knutson et al., 2010). However, observational limitations make these findings more controversial compared to those on TC frequency (Knutson et al., 2010). The RGTracks-20C provides additional support in resolving this controversy. Based on the IBTrACS, the majority of observed TCs globally last fewer than 20 days, with a peak around 8 days (Fig. 5b). Evaluation results (Fig. 5b and Supplementary Fig. S3) show that TCs detected by the OWZ tracker exhibit durations that are close to the observations, and accurately reproduce the TC duration distribution with a peak of 8 days. However, bias is found in the durations of those detected by the UZ tracker, which exhibits a duration peak of approximately 5 days. This is mainly due to the dynamics-based OWZ tracker having the ability to detect storms early in their development (Bell et al., 2018; Bourdin et al., 2022) (Supplementary Fig. S3), while the UZ tracker easily misses weak and short storms (Supplementary Figs. S1a, c) from the 20CRv3 (Bourdin et al., 2022; Tory et al., 2013; Zarzycki and Ullrich, 2017) (Supplementary Sect. S2.3). Similar results are obtained in different basins (Fig. 6b), thus, it is recommended to use the OWZ output when analyzing the durations of TCs and to study the genesis locations of TCs.

For TC intensity, given the relatively considerable uncertainty in $WIND_{max}$ compared to SLP_{min} in both reanalyses and IBTrACS (see Methods) (Bourdin et al., 2022; Chavas et al., 2017; Knapp et al., 2010; Knutson et al., 2015; Schreck et al., 2014), this study exclusively utilizes SLP_{min} to evaluate the capability of RGTracks-20C in representing the intensity of TCs. According to IBTrACS (Figs. 5c–d), the intensity of TCs is mainly distributed between 900 and 1020 hPa, peaking around 1000 hPa, with a long tail on the lower SLP_{min} side. In contrast, the SLP_{min} in RGTracks-20C is mainly distributed in the range of 950 – 1020 hPa, with peaks at 1000 hPa and 1005 hPa for the UZ (Fig. 5c) and OWZ (Fig. 5d) trackers, respectively. This suggests that the 20CRv3 generally underestimates the TC intensity (Fig. 2a), which, as expected, is primarily

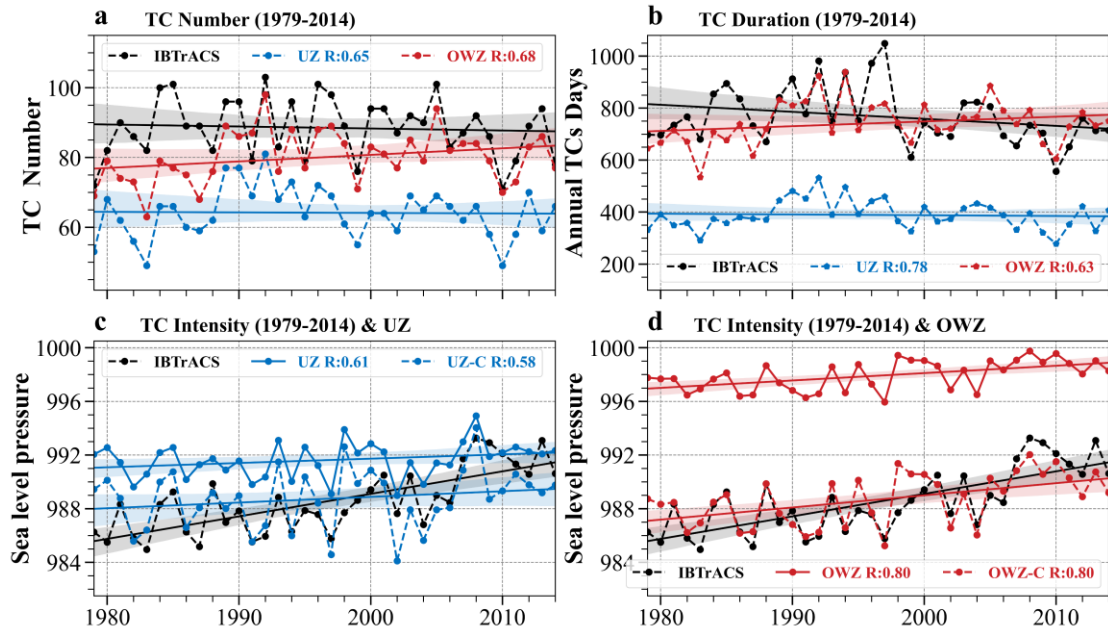
456 because the relatively low spatial resolution of the reanalysis may cause smoothing effects on the
457 sea level pressure field. Apart from spatial resolution, the model's dependence on parameterization
458 processes, along with other factors, may also influence its ability to reproduce TC intensity in the
459 reanalysis (Aarons et al., 2021; Hodges et al., 2017; Malakar et al., 2020).

460 To address this issue, an intensity bias correction was implemented using quantile mapping
461 bias correction (see Methods) (Zhao and Held, 2010). After intensity correction, the TC intensity
462 distribution in RGTracks-20C is more consistent with IBTrACS (Fig. 2b and Figs. 5c–d), especially
463 in terms of peak positions, and accurately reproduces the skewed distribution of TC intensity. In
464 particular, the RGTracks-20C reproduces TC intensity values with SLP_{min} below 940 hPa, which
465 were not found before the intensity bias correction. Notably, while the fitted curves show consistent
466 patterns following correction, they do not perfectly overlap, suggesting that certain discrepancies
467 persist (Figs. 5c–d subplot). This consistency is observed not only on a global scale but also across
468 various basins (Figs. 6c–d).

469 **3.4 Long-term variability of TC activity**

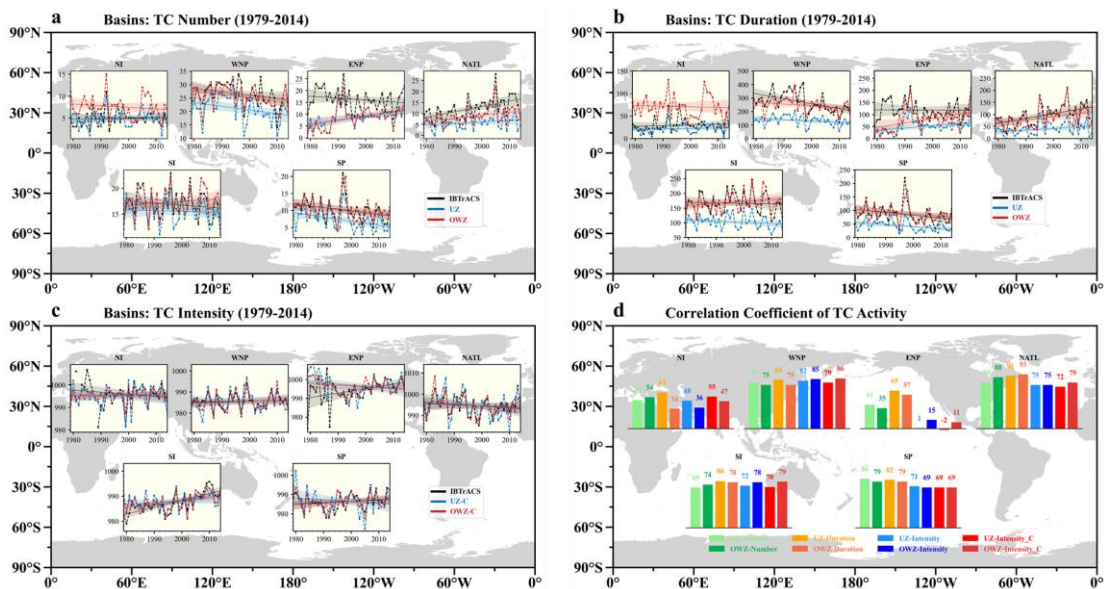
470 This section evaluates the long-term variability of TC activity in the RGTracks-20C by
471 comparing it with the IBTrACS from 1979 to 2014.

472 Firstly, the RGTracks-20C is able to capture the observed interannual variability of global TC
473 number (Fig. 7a), as indicated by the significant correlations between the TC counts derived from
474 the UZ and OWZ trackers and observations, with correlation coefficients of 0.65 and 0.68 (in the
475 following context, all correlations are significant at the 99% confidence level unless otherwise
476 specified), respectively. This is also true for individual basins (Figs. 8a, d), with the correlation
477 coefficients exceeding 0.70 in most basins. Among the six basins, the highest correlation is observed
478 in the NATL, where the correlation coefficient for the OWZ tracker reaches 0.88 (0.79 for the UZ
479 tracker). Subsequent regions with notable correlations include the WNP (0.75 for OWZ tracker,
480 0.79 for UZ tracker), SP (0.79 for OWZ tracker, 0.84 for UZ tracker), and SI (0.74 for OWZ tracker,
481 0.69 for UZ tracker). However, the correlation coefficients are relatively lower in the ENP and NI
482 (Supplementary Table S2 and Fig. S4a, e), of which the reasons are discussed in Supplementary
483 Sect. S2.2 and Table S3. Notably, the long-term trends in the number of TCs recorded by the two
484 datasets are consistent globally and across most of the ocean basins (Supplementary Table S4).



485
486
487
488
489
490
491
492
493
494
495

Figure 7: Time series of globally TC activities from IBTrACS and RGTracks-20C during the periods 1979-2014. TC activities are from the IBTrACS and RGTracks-20C using UZ (blue), and OWZ (red) trackers. a, TC number. b, TC days (unit: days). c, TC intensity in 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.



496
497
498
499
500

Figure 8: As in Fig. 7, but for six basins. a, TC number. b, TC days (unit: days). c, TC intensity in SLP_{min} (unit: hPa) in IBTrACS (black) and RGTracks-20C (after bias correction) using UZ (blue) and OWZ (red) trackers. d, the correlation coefficients (R) between the from IBTrACS and RGTracks-20C. Note*: The R values for TC number and TC intensity are not statistically significant at the 99% confidence level in the NI

501 and ENP. For TC days, the R value is not statistically significant only in the NI. The R values need to be
502 divided by 100.

503

504 TC days, an important metric, encompasses both TC frequency and lifespan (Bell et al., 2018).
505 The RGTracks-20C is able to reproduce the interannual variability of TC days, which is consistent
506 with that in IBTrACS (Fig. 7b), with high correlation coefficients of 0.78 and 0.63 for the UZ and
507 OWZ trackers, respectively. Moreover, these results are further confirmed across basins (Fig. 8b),
508 with correlation coefficients generally exceeding 0.75. In particular, in the NATL, the correlation
509 coefficient exceeds 0.90 (UZ tracker: 0.93, OWZ tracker: 0.91), followed by the SP (UZ tracker:
510 0.82, OWZ tracker: 0.79), the SI (UZ tracker: 0.80, OWZ tracker: 0.78) and the WNP (UZ tracker:
511 0.84, OWZ tracker: 0.75). However, being influenced by the observation biases, the correlation
512 coefficients for TC days are also relatively low in the ENP and NI (Supplementary Table S2 and Fig.
513 S4b, f). Nevertheless, the above results indicate that the RGTracks-20C provides a satisfactory
514 representation of the interannual and long-term variability (Supplementary Sect. S2.4, Table S4) of
515 the TC days globally and across most of the ocean basins.

516 In addition, the global TC intensity series based on RGTracks-20C significantly correlates with
517 that based on IBTrACS, with correlation coefficients of 0.61 and 0.80 for the UZ (Fig. 7c) and OWZ
518 (Fig. 7d) trackers, respectively. This indicates that the TC intensity (SLP_{min}) in RGTracks-20C
519 effectively captures the observed interannual variability. Most basins further validate these results
520 (Fig. 8d). The highest correlation coefficients are observed in the WNP, exceeding 0.80 (UZ tracker:
521 0.82, OWZ tracker: 0.85). Following closely are NATL (UZ tracker: 0.75, OWZ tracker: 0.75) and
522 SI (UZ tracker: 0.72, OWZ tracker: 0.78), while SP (UZ tracker: 0.71, OWZ tracker: 0.69) also
523 demonstrates correlation coefficients of around 0.70.

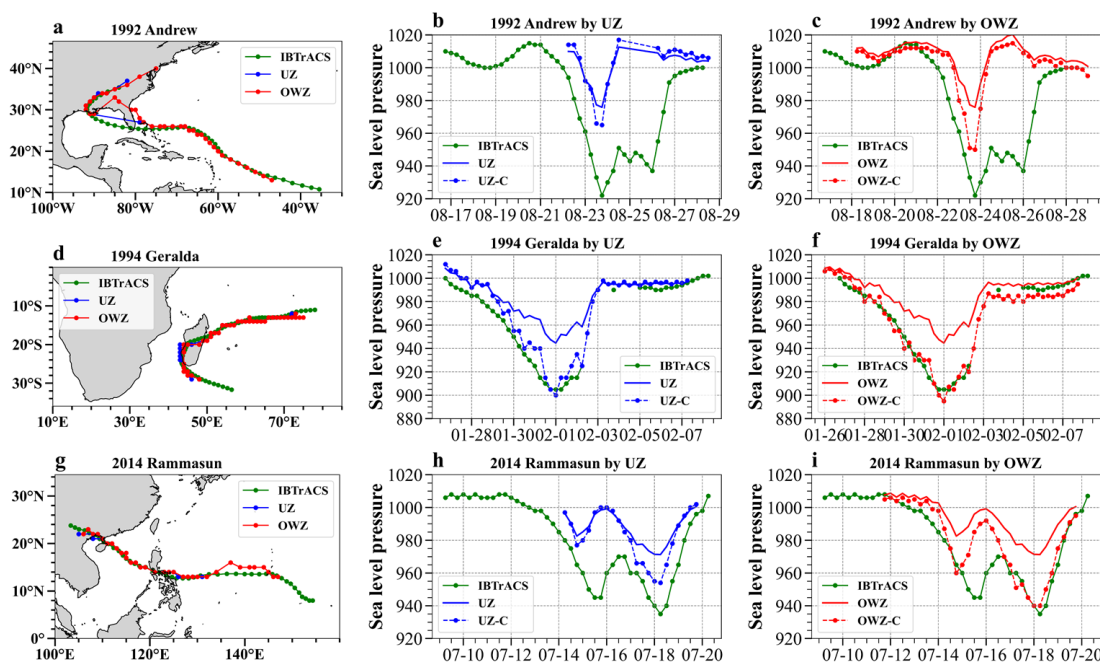
524 The 20CRv3 tends to underestimate the TC intensities, due to its coarse resolution, which
525 suggests the need of a bias correction process during the production of the RGTracks-20C (see
526 Methods). By performing intensity bias corrections to the detected TCs, the TC intensity (SLP_{min})
527 in RGTracks-20C exhibits interannual and long-term variations that are more consistent with the
528 observations (Figs. 2, 7c–d, and Supplementary Tables S2, S4), especially in the WNP, NATL, and
529 SI basins (Figs. 8c–d). These results indicate that the RGTracks-20C can reasonably capture the
530 interannual variability and trends (Supplementary Sect. S2.4 and Table S4) of TC intensity globally
531 and across most basins. Discrepancies in the interannual variability of TC intensity between the
532 RGTracks-20C and IBTrACS are also noted over ENP and NI (Supplementary Table S5 and Fig.
533 S4c–d, f–h), similar to the above findings on TC number and days (Supplementary Sect. S2.2 and
534 Tables S6–S7).

535 **3.5 Key strengths of the RGTracks-20C**

536 The above evaluation analyses confirm that the RGTracks-20C effectively captures both the
537 climatology and long-term variability of TC activity across global and major oceanic basins. In this
538 section, we discuss the key strengths of the RGTracks-20C, specifically its capacity to reconstruct

539 track and intensity information of early-year TCs that may not be included in the observed data
 540 records. Such an advantage of the RGTracks-20C could benefit research about how climate change
 541 has affected TCs over the past century.

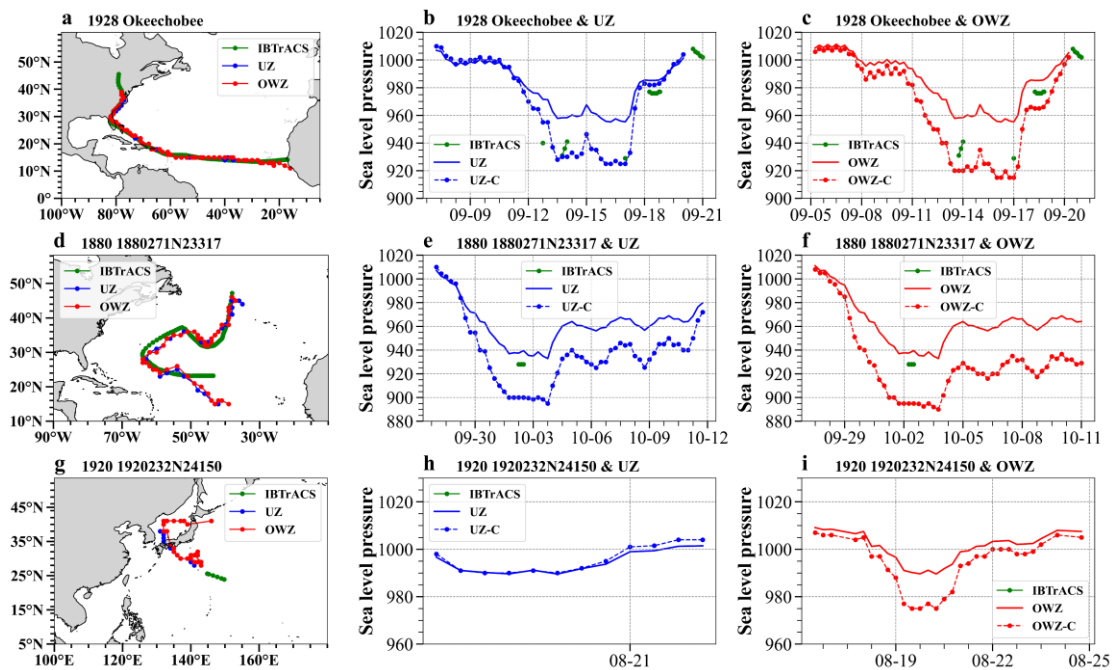
542 Before digging into early-year TCs, we first demonstrate the RGTracks-20C's accuracy in
 543 reproducing specific TCs by making comparisons with observations. Three representative TCs that
 544 caused significant human casualties and economic losses in the NATL, SI, and WNP are analyzed
 545 here: Hurricane 'Andrew' in 1992 (Pimm et al., 1994) (Figs. 9a–c), TC 'Geralda' in 1994 (Hoarau
 546 et al., 2012) (Figs. 9d–f), and Super Typhoon 'Rammasun' in 2014 (Zhang et al., 2017) (Figs. 9h–
 547 i). Compared with IBTrACS, the RGTracks-20C performs exceptionally well in representing the
 548 track and duration of these TCs. However, some discrepancies were observed during landfall (Fig.
 549 9a), possibly due to small TC size, which were not captured by the low-resolution 20CRv3
 550 (Supplementary Sect. S3.1 and Fig. S5). While the 20CRv3 tends to underestimate the intensity of
 551 TCs, the corrected intensity in the RGTracks-20C is highly consistent with observations and
 552 accurately captures the temporal evolutions of TC intensities. This evidence confirms RGTracks-
 553 20C's ability to capture not only the climatology and variability of TC activity, but also the detailed
 554 information on specific TC events.



555
 556 **Figure 9: The historical tracks and intensity records of individual tropical cyclones in the IBTrACS and**
 557 **RGTracks-20C. a–c, Track (a) and intensity (SLP_{min} , unit: hPa. b: UZ tracker, c: OWZ tracker) of Hurricane**
 558 **“Andrew”. d–f, same as a–c, but for track (d) and intensity (SLP_{min} , unit: hPa. e: UZ tracker, f: OWZ tracker)**
 559 **of tropical cyclone “Geralda”. g–i, same as a–c, track (g) and intensity (SLP_{min} , unit: hPa. h: UZ tracker, i:**
 560 **OWZ tracker) of Super typhoon “Rammasun”. Green, blue, and red lines denote results based on the**
 561 **IBTrACS, UZ tracker, and OWZ tracker, respectively. The UZ-C (blue dotted dashed line) and OWZ-C (red**
 562 **dotted dashed line) indicate after intensity bias correction.**

563

564 Prior to the satellite era, limitations in observation systems often led to incomplete records of
 565 early TCs, particularly for TC intensity. An example is hurricane Okeechobee in 1928, which was
 566 one of the deadliest to hit the United States in the early 20th century. Hurricane Okeechobee was
 567 recorded in the IBTrACS (Blake et al., 2011; Mitchell, 1928) (Supplementary Sect. S3.2). However,
 568 during Okeechobee’s lifetime, there were only 16 time points of the TC intensity that were recorded
 569 when it passed the Lesser Antilles and Puerto Rico, and made landfall in the United States (Figs.
 570 10a–c, Supplementary Fig. S6 and Table S8). Similar missing data are common in the IBTrACS
 571 records of early TCs, especially when the TCs were located over the ocean (Figs. 10d–f). Moreover,
 572 the problem of missing TC intensity records is especially evident in other basins (Supplementary
 573 Table S3). For instance, Typhoon No. 8, which made landfall and caused serious damage in Japan
 574 (Supplementary Sect. S3.3), has only track records in the IBTrACS, but with intensity information
 575 missing (Figs. 10g–i). In such cases, taking advantage of the 20CRv3, the RGTracks-20C addresses
 576 these deficiencies by filling in these gaps, substantially enhancing the completeness of early TC
 577 intensity records.



578
 579 **Figure 10:** As in Fig. 9, but for Hurricane “Okeechobee” (a–c), Hurricane ‘1880271N23317’ (d–f), typhoon
 580 ‘192023N24150’(g–i).

581
 582 In addition, not only is the TC intensity missing, but the track records in the IBTrACS may also
 583 be incomplete, such as the above-mentioned Typhoon No.8 in 1920 (Fig. 10g), despite the existence
 584 of historical observation records (Supplementary Sect. S3.3). In this case, the RGTracks-20C not
 585 only provides the missing TC intensity but also fills gaps in IBTrACS during the latter stages of the
 586 typhoon’s development, especially during the landfall phase (Fig. 10g and Supplementary Figs. S7–
 587 9). Moreover, prior to the satellite era, the RGTracks-20C often reports a higher number of TCs than
 588 the IBTrACS, particularly from the early to mid-20th century (Supplementary Sect. 4), which

589 suggests that the RGTracks-20C is also able to detect historical TCs not being recorded in the
590 IBTrACS. These findings demonstrate that the RGTracks-20C can compensate for the incomplete
591 TC track records in the IBTrACS, especially for those in the pre-satellite era.

592 To evaluate the accuracy of early TC records provided by RGTracks-20C, we take the 1928
593 Okeechobee hurricane as a case study. The RGTracks-20C nearly fully reproduces the hurricane's
594 lifespans as recorded in IBTrACS, with the OWZ tracker performing exceptionally well, differing
595 by only one day from the IBTrACS record. Okeechobee's latitude and longitude variations in the
596 RGTracks-20C are highly consistent with those in the IBTrACS, with a positional bias within ± 1
597 degree (Fig. 10a and Supplementary Fig. S6). By comparing Okeechobee's intensity in RGTracks-
598 20C with observational data, we find that the RGTracks-20C reliably reproduces Okeechobee's
599 intensity and its variations (Figs. 10b–c and Supplementary Table S8). For instance, as the hurricane
600 passed over Guadeloupe, IBTrACS recorded a SLP_{min} of 940 hPa, which is closely matched by
601 RGTracks-20C (UZ tracker: 955 hPa; OWZ tracker: 940 hPa). Moreover, the RGTracks-20C
602 captures the weakening and re-intensification of the hurricane between Puerto Rico and its landfall
603 in Florida, where the IBTrACS lacks intensity records, demonstrating the RGTracks-20C's
604 reliability in representing intensity changes (Supplementary Sect. S3.2).

605 **4. Usage notes**

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

611 As documented in prior studies, biases are unavoidable when extracting TCs from reanalyses
612 due to the data quality of reanalyses and the limitations of TC trackers. Some usage notes and
613 cautionary remarks are listed in this section to assist readers in understanding or using the RGTracks-
614 20C.

615 (1) Due to model resolution and parameterization, TC intensity detected directly from the
616 20CRv3 is underrepresented compared to observations. To address this issue in the RGTracks-20C,
617 we corrected the biases using a simple quantile mapping method, assuming that systematic biases
618 primarily cause the TC intensity errors from 20CRv3. While this is generally true, the quantile
619 mapping correction did not account for other factors that may also affect TC intensity biases. The
620 inherent challenges in modeling weaker TCs in 20CRv3, which are largely attributed to the
621 limitations of resolution and parameterization of subgrid-scale processes in numerical models, often
622 result in lower detection rates for tropical depressions and weaker tropical storms (e.g., Category 1)
623 (Hodges et al., 2017). Given the uncertainties in the wind fields of reanalysis data, the current version
624 of the RGTracks-20C does not provide information of TC stages. This will be improved with more
625 advanced correction approaches of TC intensity in future versions of RGTracks-20C (Han and

626 Ullrich, 2025).

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

641 (3) Currently, there are no perfect algorithms for tracking TCs from reanalyses. Although the
642 TC trackers employed in the RGTracks-20C (UZ and OWZ) are two widely recognized algorithms,
643 they were built with different properties and have different limitations. The above evaluation
644 analyses show that the OWZ tracker is closer to the observations in terms of TC number and TC
645 days (Bourdin et al., 2022), while the UZ tracker produces tracks with a shorter duration than the
646 observations, which is mainly related to its physically based tracker intensity threshold (Horn et al.,
647 2014). However, the UZ has a lower FAR, suggesting that it has an advantage in recognizing real
648 TCs and is less likely to misclassify other weather systems as TCs. Generally, since the OWZ tracker
649 demonstrates overall higher stability in detecting TCs, it is recommended to primarily utilize the
650 OWZ tracker results in most cases, with the UZ tracker as a supplementary reference for analyses.
651 In addition, in the production of the RGTracks-20C, globally identical thresholds were used in the
652 TC tracking procedure. However, given the differences in structure and behavior of TCs in different
653 basins and the influence of different meteorological systems and topography, the use of a globally
654 identical tracker may affect the accuracies of TC detection in specific regions (Fu et al., 2021; Raavi
655 and Walsh, 2020a, b). This suggests the need for further improvements in the TC tracking
656 approaches.

657 (4) The assimilation of SLP_{min} from IBTrACS into the 20CRv3 may lead to another limitation.
658 As discussed in Supplementary Sect. S4, the annual number of available observations and
659 assimilated observations (including some IBTrACS) increases over time, with both showing
660 accelerated growth, especially after 1950 (Supplementary Fig. S10). This increasing number of
661 available observation data could improve the quality of the reanalysis. Consequently, the RGTracks-
662 20C exhibits consistent trends and variations with IBTrACS from 1850 to 2014 (Supplementary Fig.
663 S11). In particular, the growth trends in TC numbers from both datasets during the mid-20th century
664 are highly correlated, primarily resulting from the artificial increase in TC detection associated with

665 advancements in observational technologies. In addition, RGTracks-20C currently uses the
666 ensemble mean field of 20CRv3 as input data, which further affects this similarity by inherently
667 weakening the intensity and character of extreme events and introducing smoothing effects
668 (Emanuel, 2024). On the other hand, the assimilation of IBTrACS data has, to some extent, also
669 improved 20CRv3's representation of TC intensity and structure, enabling TC tracker to more
670 effectively detect and identify TCs that actually occurred (Slivinski et al., 2019, 2021), such as the
671 typhoon that made landfall in Japan in 1920 (Fig. 10g) and hurricane Andrew (1992) (Fig. 9a). Due
672 to the changing number of assimilated observation data, RGTracks-20C may not capture the realistic
673 number of TCs in early years. Therefore, caution should be exercised when analyzing the long-term
674 trend of TC activities. Future versions should improve the reliability of RGTracks-20C by tracking
675 TCs using individual ensemble members (Emanuel, 2024).

676 The above factors will be thoroughly considered and addressed in the future versions of
677 RGTracks-20C to enhance its accuracy and applicability. In the next version of RGTracks-20C, a
678 few improvements will be included: (1) We detect TCs separately from all 80 ensemble members of
679 the 20CRv3, in order to avoid the smoothing effects caused by the ensemble mean of reanalyses and
680 to provide reliable estimates of uncertainty and confidence (Emanuel, 2024); (2) we will calibrate
681 algorithm thresholds according to TC characteristics in different ocean basins; (3) more TC tracking
682 algorithms will be included to address the uncertainty of the TC track data (Flaounas et al., 2023;
683 Han and Ullrich, 2025). The algorithm developed by Han and Ullrich (2025) enables the
684 classification of tropical cyclones across different stages of their lifespans.

685 **5. Data Availability**

686 The RGTracks-20C is publicly available at <https://doi.org/10.5281/zenodo.14411917> (Ye et al.,
687 2024). The Other datasets utilized in this study are available: the IBTrACS at
688 <https://www.ncdc.noaa.gov/ibtracs/>; and the 20CRv3 at
689 <https://portal.nersc.gov/archive/home/projects/incite11/www/> (Slivinski et al., 2019). Historical
690 weather chart of the 1920 typhoon that made landfall in Japan from [http://agora.ex.nii.ac.jp/cgi-
691 bin/weather-chart/calendar.pl?year=1920&month=8&lang=en&type=as](http://agora.ex.nii.ac.jp/cgi-bin/weather-chart/calendar.pl?year=1920&month=8&lang=en&type=as).

692 **6. Code Availability**

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

696 **7. Conclusion**

697 In this study, we introduce the RGTracks-20C, a century-long reanalysis-based historical global
698 TC dataset. Statistical evaluations and case studies confirm its reliability in capturing the

699 climatology and interannual variability of observed TC activity on both global and regional scales.
700 A major key strength of the RGTracks-20C is its ability to fill the missing intensity and location
701 records of observed TCs in early years. This dataset provides a reliable alternative for researchers to
702 study the long-term variability of TC characteristics, which will help us to better understand changes
703 and trends in historical TC activity, as well as their relationship with climate change.

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

709 **Competing interests**

710 The authors declare no competing interests.

711 **Author contributions**

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

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

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

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

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

720 **Acknowledgements**

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

728 This study is primarily supported by the National Natural Science Foundation of China (Grant
729 No. U21A6001). G.Y. and W.D. are also supported by the Southern Marine Science and Engineering
730 Guangdong Laboratory (No. SML2023SP208). G.Y., J.C.H.L., J.X., W.L., W.Q., and B.Z. are
731 supported by the Guangdong Province Introduction of Innovative R&D Team Project China
732 (2019ZT08G669). J.X. is supported by the National Natural Science Foundation of China

733 (42130605). J.C.H.L. is supported by the National Natural Science Foundation of China (42405038)
734 and the Guangdong Basic and Applied Basic Research Foundation (2020A1515110275).

735 **Author contributions**

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

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

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

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

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

744

745 **References**

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

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

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

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

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

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

764 Chan, J. C. L.: Frequency and intensity of landfalling tropical cyclones in East Asia: Past variations
765 and future projections, *Meteorology*, 2, 171–190, <https://doi.org/10.3390/meteorology2020012>,
766 2023.

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

769 Chan, K. T. F., Zhang, K., Wu, Y., and Chan, J. C. L.: Publisher Correction: Landfalling hurricane
770 track modes and decay, *Nature*, 608, E14–E14, <https://doi.org/10.1038/s41586-022-05078-1>,
771 2022a.

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

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

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

781 Chavas, D. R., Reed, K. A., and Knaff, J. A.: Physical understanding of the tropical cyclone wind-
782 pressure relationship, *Nat Commun*, 8, 1360, <https://doi.org/10.1038/s41467-017-01546-9>,
783 2017.

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

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

789 Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E.,
790 Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I.,
791 Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J.,
792 Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D.,
793 and Worley, S. J.: The Twentieth Century Reanalysis Project, *Q.J.R. Meteorol. Soc.*, 137, 1–
794 28, <https://doi.org/10.1002/qj.776>, 2011.

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

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

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

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

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

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

814 Emanuel, K.: Atlantic tropical cyclones downscaled from climate reanalyses show increasing
815 activity over past 150 years, *Nat Commun*, 12, 7027, [https://doi.org/10.1038/s41467-021-](https://doi.org/10.1038/s41467-021-27364-8)
816 [27364-8](https://doi.org/10.1038/s41467-021-27364-8), 2021.

817 Emanuel, K.: Limitations of reanalyses for detecting tropical cyclone trends, *Nat. Clim. Chang.*, 14,
818 143–145, <https://doi.org/10.1038/s41558-023-01879-y>, 2024.

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

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

828 Fu, D., Chang, P., Patricola, C. M., Saravanan, R., Liu, X., and Beck, H. E.: Central American
829 mountains inhibit eastern North Pacific seasonal tropical cyclone activity, *Nat Commun*, 12,
830 4422, <https://doi.org/10.1038/s41467-021-24657-w>, 2021.

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

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

836 Hassanzadeh, P., Lee, C.-Y., Nabizadeh, E., Camargo, S. J., Ma, D., and Yeung, L. Y.: Effects of
837 climate change on the movement of future landfalling Texas tropical cyclones, *Nat Commun*,
838 11, 3319, <https://doi.org/10.1038/s41467-020-17130-7>, 2020.

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

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

843 Horn, M., Walsh, K., Zhao, M., Camargo, S. J., Scoccimarro, E., Murakami, H., Wang, H., Ballinger,
844 A., Kumar, A., Shaevitz, D. A., Jonas, J. A., and Oouchi, K.: Tracking scheme dependence of

845 simulated tropical cyclone response to idealized climate simulations, *J. Climate*, 27, 9197–9213,
846 <https://doi.org/10.1175/JCLI-D-14-00200.1>, 2014.

847 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S.,
848 White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo,
849 K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The
850 NCEP/NCAR 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc.*, 77, 437–472,
851 [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.

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

854 Klotzbach, P. J., Bell, M. M., Bowen, S. G., Gibney, E. J., Knapp, K. R., and Schreck, C. J.: Surface
855 pressure a more skillful predictor of normalized hurricane damage than maximum sustained
856 wind, *Bull. Amer. Meteor. Soc.*, 101, E830–E846, <https://doi.org/10.1175/BAMS-D-19-0062.1>,
857 2020.

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

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

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

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

872 Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., Holland, G., Landsea, C., Held, I., Kossin, J.
873 P., Srivastava, A. K., and Sugi, M.: Tropical cyclones and climate change, *Nature Geosci*, 3,
874 157–163, <https://doi.org/10.1038/ngeo779>, 2010.

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

879 Kossin, J. P., Knapp, K. R., Olander, T. L., and Velden, C. S.: Global increase in major tropical
880 cyclone exceedance probability over the past four decades, *Proc. Natl. Acad. Sci. USA*, 117,
881 11975–11980, <https://www.pnas.org/doi/full/10.1073/pnas.1920849117>, 2020.

882 Kunze, S.: Unraveling the effects of tropical cyclones on economic sectors worldwide: Direct and
883 indirect impacts, *Environ Resource Econ* 78, 545–569, [https://doi.org/10.1007/s10640-021-](https://doi.org/10.1007/s10640-021-00541-5)
884 00541-5, 2021.

885 Lai, Y., Li, J., Gu, X., Chen, Y. D., Kong, D., Gan, T. Y., Liu, M., Li, Q., and Wu, G.: Greater flood
886 risks in response to slowdown of tropical cyclones over the coast of China, *Proceedings of the*
887 *National Academy of Sciences*, 117, 14751–14755, <https://doi.org/10.1073/pnas.1918987117>,
888 2020.

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

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

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

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

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

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

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

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

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

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

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

922 Li, Y., Tang, Y., Li, X., Song, X., and Wang, Q.: Recent increase in the potential threat of western
923 North Pacific tropical cyclones, *npj Clim Atmos Sci*, 6, 1–8, [https://doi.org/10.1038/s41612-](https://doi.org/10.1038/s41612-023-00379-2)
924 [023-00379-2](https://doi.org/10.1038/s41612-023-00379-2), 2023.

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

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

931 Mitchell, C. L.: The West Indian hurricane of September 10–20, 1928, *Mon. Wea. Rev.*, 56, 347–
932 350, [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.

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

935 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.:
936 Rainfall strength and area from landfalling tropical cyclones over the North Indian and western
937 North Pacific oceans under increased CO2 conditions, *Weather Clim. Extrem.*, 41, 100581,
938 <https://doi.org/10.1016/j.wace.2023.100581>, 2023.

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

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

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

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

949 Parker, W. S.: Reanalyses and Observations: What’s the difference?, *Bull. Amer. Meteor. Soc.*, 97,
950 1565–1572, <https://doi.org/10.1175/BAMS-D-14-00226.1>, 2016.

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

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

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

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

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

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

971 Sharmila, S. and Walsh, K. J. E.: Recent poleward shift of tropical cyclone formation linked to
972 Hadley cell expansion, *Nature Clim Change*, 8, 730–736, [https://doi.org/10.1038/s41558-018-](https://doi.org/10.1038/s41558-018-0227-5)
973 [0227-5](https://doi.org/10.1038/s41558-018-0227-5), 2018.

974 Slivinski, L. C.: Historical Reanalysis: What, How, and Why?, *J. Adv. Model. Earth Syst.*, 10, 1736–
975 1739, <https://doi.org/10.1029/2018MS001434>, 2018.

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

985 Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan,
986 P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster, D.
987 C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram,
988 T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H.,
989 Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C.,
990 Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński,
991 P.: An evaluation of the performance of the Twentieth Century Reanalysis Version 3, *J. Climate*,
992 34, 1417–1438, <https://doi.org/10.1175/JCLI-D-20-0505.1>, 2021.

993 Torn, R. D. and Snyder, C.: Uncertainty of tropical cyclone best-track information, *Wea. Forecasting*,
994 27, 715–729, <https://doi.org/10.1175/WAF-D-11-00085.1>, 2012.

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

998 Truchelut, R. E. and Hart, R. E.: Quantifying the possible existence of undocumented Atlantic warm-
999 core cyclones in NOAA/CIRES 20th Century Reanalysis data, *Geophys. Res. Lett.*, 38, L08811,
1000 <https://doi.org/10.1029/2011GL046756>, 2011.

1001 Truchelut, R. E., Hart, R. E., and Luthman, B.: Global identification of previously undetected Pre-
1002 satellite-era tropical cyclone candidates in NOAA/CIRES Twentieth-Century Reanalysis Data,
1003 *J. Appl. Meteor. Climatol.*, 52, 2243–2259, <https://doi.org/10.1175/JAMC-D-12-0276.1>, 2013.

1004 Tu, S., Xu, J., Chan, J. C. L., Huang, K., Xu, F., and Chiu, L. S.: Recent global decrease in the inner-
1005 core rain rate of tropical cyclones, *Nat Commun*, 12, 1948, [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-021-22304-y)
1006 [021-22304-y](https://doi.org/10.1038/s41467-021-22304-y), 2021.

1007 Tu, S., Chan, J. C. L., Xu, J., Zhong, Q., Zhou, W., and Zhang, Y.: Increase in tropical cyclone rain
1008 rate with translation speed, *Nat Commun*, 13, 7325, [https://doi.org/10.1038/s41467-022-](https://doi.org/10.1038/s41467-022-35113-8)
1009 [35113-8](https://doi.org/10.1038/s41467-022-35113-8), 2022.

1010 Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., and Reed, K.
1011 A.: TempestExtremes v2.1: a community framework for feature detection, tracking, and
1012 analysis in large datasets, *Geosci. Model Dev.*, 14, 5023–5048, [https://doi.org/10.5194/gmd-](https://doi.org/10.5194/gmd-14-5023-2021)
1013 [14-5023-2021](https://doi.org/10.5194/gmd-14-5023-2021), 2021.

1014 Wang, S. and Toumi, R.: More tropical cyclones are striking coasts with major intensities at landfall,
1015 *Sci Rep*, 12, 5236, <https://doi.org/10.1038/s41598-022-09287-6>, 2022.

1016 Wang, X. L., Feng, Y., and Swail, V.: North Atlantic wave height trends as reconstructed from the
1017 20th century reanalysis, *Geophys. Res. Lett.*, 39, L18705,
1018 <https://doi.org/10.1029/2012GL053381>, 2012.

1019 Yamaguchi, M., Chan, J. C. L., Moon, I.-J., Yoshida, K., and Mizuta, R.: Global warming changes
1020 tropical cyclone translation speed, *Nat Commun*, 11, 47, [https://doi.org/10.1038/s41467-019-](https://doi.org/10.1038/s41467-019-13902-y)
1021 [13902-y](https://doi.org/10.1038/s41467-019-13902-y), 2020.

1022 Ye, G., Jeremy Cheuk-Hin, L., Dong, W., Xu, Ji., LI, W., Qian, W., Kong, H., and Zhang, B.: A
1023 Reanalysis-Based Global Tropical Cyclone Tracks Dataset for the Twentieth Century
1024 (RGTracks-20C) [data set], <https://doi.org/10.5281/zenodo.14411917>, 2024.

1025 Yeasmin, A., Chand, S., and Sultanova, N.: Reconstruction of tropical cyclone and depression
1026 proxies for the South Pacific since the 1850s, *Weather Clim. Extrem.*, 39, 100543,
1027 <https://doi.org/10.1016/j.wace.2022.100543>, 2023.

1028 Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., and Chen, D.: An overview of the
1029 China meteorological administration tropical cyclone database, *J. Atmos. Oceanic Technol.*,
1030 31, 287–301, <https://doi.org/10.1175/JTECH-D-12-00119.1>, 2014.

1031 Yoshida, K., Sugi, M., Mizuta, R., Murakami, H., and Ishii, M.: Future changes in tropical cyclone
1032 activity in high-resolution large-ensemble simulations, *Geophys. Res. Lett.*, 44, 9910–9917,
1033 <https://doi.org/10.1002/2017GL075058>, 2017.

1034 Zarzycki, C. M. and Ullrich, P. A.: Assessing sensitivities in algorithmic detection of tropical
1035 cyclones in climate data, *Geophys. Res. Lett.*, 44, 1141–1149,
1036 <https://doi.org/10.1002/2016GL071606>, 2017.

1037 Zhang, B., Zhang, R., Pinker, R. T., Feng, Y., Nie, C., and Guan, Y.: Changes of tropical cyclone
1038 activity in a warming world are sensitive to sea surface temperature environment, *Environ. Res.*
1039 *Lett.*, 14, 124052, <https://doi.org/10.1088/1748-9326/ab5ada>, 2019.

1040 Zhang, G.: Warming-induced contraction of tropical convection delays and reduces tropical cyclone
1041 formation, *Nat Commun*, 14, 6274, <https://doi.org/10.1038/s41467-023-41911-5>, 2023.

1042 Zhang, X., Duan, Y., Wang, Y., Wei, N., and Hu, H.: A high-resolution simulation of Super typhoon
1043 Rammasun (2014)—Part I: Model verification and surface energetics analysis, *Adv. Atmos.*
1044 *Sci.*, 34, 757–770, <https://doi.org/10.1007/s00376-017-6255-7>, 2017.

1045 Zhao, M. and Held, I. M.: An analysis of the effect of global warming on the intensity of Atlantic
1046 hurricanes using a GCM with statistical refinement, *J. Climate*, 23, 6382–6393,
1047 <https://doi.org/10.1175/2010JCLI3837.1>, 2010.

1048 Zhu, L. and Quiring, S. M.: Exposure to precipitation from tropical cyclones has increased over the
1049 continental United States from 1948 to 2019, *Commun Earth Environ*, 3,
1050 312 <https://doi.org/10.1038/s43247-022-00639-8>, 2022.

1051