https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

RABTracks: Reanalysis-Augmented Best Tracks for Tropical Cyclones

Stella Bourdin[1, 2, 3], Kevin Hodges[1, 4], Yushan Han[5], Leo Saffin[1, 4], Alexander J. Baker[1, 4], John Methven[4], and Pier Luigi Vidale[1, 4]

[1]National Centre for Atmospheric Science, UK
[2]Department of Physics, University of Oxford, Oxford, UK
[3]LMD/IPSL, École Polytechnique, Institut Polytechnique de Paris, ENS, PSL Research University, Sorbonne Université, CNRS, Palaiseau, France
[4]Department of Meteorology, University of Reading, Reading, UK
[5]Department of Land, Air and Water Resources, University of California, Davis, CA, USA

Correspondence: Stella Bourdin (stella.bourdin@lmd.ipsl.fr)

Abstract. A new dataset, Reanalysis-Augmented Best Tracks for Tropical Cyclones (RABTracks), is created for the study of Tropical Cyclones (TCs) globally. The goal of RABTracks is to bring together multiple sources of information on historical tropical cyclones in one place. To do so, we augment the International Best Track Archive for Climate Stewardship (IBTrACS) with track data from five analyses (four reanalyses and one operational analysis), using two different tracking methods, in all basins, from 1940 to present. This provides us, in particular, with extended tracks that reconstruct the early and late stages of each recorded cyclone. It is also an opportunity to gather information on cyclones' intensity, size, and to retrieve information about the cyclones' nature at different stages, derived from the Cyclone Phase Space (Hart, 2003), and SyCLoPS classification (Han and Ullrich, 2025). This dataset has a high potential to support robust TC-related climatological studies. It was also designed to be easy to extend with additional data, and all the code for creating the dataset, making the analyses and adding new data is provided in an open-source repository.

1 Introduction

Tropical Cyclones (TCs) are very intense, organised convective systems that cause extreme winds and precipitation. They are responsible for major human and economic losses worldwide, accounting for more than 200,000 deaths and 50% of the natural-disaster-related economic losses worldwide over the 2000-2019 period (CRED and UNDRR, 2020). It is therefore essential to anticipate their impacts at every time horizon. At climate time scales (decadal and more), predicting how cyclonic activity will evolve is necessary to inform preventive and adaptation measures. Making such predictions requires an understanding of the phenomena and their interactions with the large-scale climate environment, which in turn relies on the availability of comprehensive data.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

Best Tracks (BT) are the official, post-seasonally analysed TC tracks produced by different operational TC warning centres and are essential data for studying the evolution of TC activity over long periods of time. Best tracks provide the most accurate estimates of a tropical cyclone's position, intensity, and other characteristics at regular time intervals. Individual agencies, including the World Meteorological Organisation's (WMO) Regional Specialised Meteorological Centres (RSMCs), publish post-season best track data. They are quality-controlled and collected into the International Best Track Archive for Climate Stewardship (IBTrACS, Gahtan et al., 2024), the field's reference dataset.

Because these tracks come from different operational centres, there are inhomogeneities in some aspects of the data, stemming from different operational procedures among the basins, or over time (Schreck et al., 2014). While this is true for the intense tropical stages — where different centres might have different interpretations when it comes to the precise position and intensity of the cyclone — it is even more uncertain for the weak and/or transitional stages: defining the point of genesis or extratropical transition, for example, is subjective; so is distinguishing between a subtropical and a tropical cyclone. Moreover, because the operational centres' mission is to monitor cyclones in the tropics, not in the mid-latitudes, they tend to stop recording the track once they consider it to be no longer tropical, or if they have a longitudinal operational boundary that affects data coverage. This has consequences for anyone trying to make a systematic analysis of these aspects. For example, an analysis of tropical cyclogenesis or a climatology of cyclones of tropical origin in the mid-latitudes. The inhomogeneities among agencies can be, to some extent, reduced by using the Joint Typhoon Warning Center (JTWC) subset, which provides a global coverage using American standards, but is not always the best option for individual basins (Ren et al., 2011; Kumar et al., 2025; Barcikowska et al., 2012). In any case, the JTWC best tracks remain prone to operational changes over time, and to subjective operational judgement, including the treatment of post-TC stages.

In this context, best tracks are not our only source of information about historical TCs, and here we are complementing them with analysis data. An analysis is obtained using data assimilation to combine information from recent observations (usually in a 12-hour window) with a background forecast, which is typically a forecast from the previous cycle of the analysis & forecast system. There are two types of analyses: the operational analyses and the reanalyses. Operational analyses are used to initialise numerical weather prediction model forecasts. They follow the evolution of the operational systems in use at the centre that publishes them. They are created in real time using the data available at that time. Reanalyses are calculated retrospectively and published to serve as climatological references and can be seen as reconstructions of the state of the atmosphere over the past decades. They use a single, consistent model and data assimilation system throughout the entire period they cover. Typically, additional data can be used in a reanalysis because it is available after the fact. For TCs, these include, in particular, TC-related information from the best tracks, which are incorporated into some products (see 2.2). However, due to the computational cost, reanalyses have lower resolution than recent operational analyses.

Since they assimilate observational data and are meant to represent the state of the atmosphere at a given point in the past, analyses, they inherently represent historical tropical cyclones. However, analyses do not provide a perfect reconstruction of historical tropical cyclones because they are constrained by model formulation and resolution, the model used, data assimilation methods, and the availability of observational data. It has been shown that the most recent reanalyses generally include nearly

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

all observed tropical cyclones, but with largely underestimated intensity and a structure that is not fully realistic (Hodges et al., 2017; Schenkel et al., 2017; Zarzycki et al., 2021; Jones et al., 2021; Aarons et al., 2021; Dulac et al., 2023).

Moreover, analyses are distributed as gridded products that represent the atmospheric state as a set of variables, with no specific information regarding meteorological objects such as TCs. TCs, therefore, need to be detected using TC detection algorithms (often called TC trackers). TC trackers have the advantage of objectively applying predefined criteria to identify cyclonic structures. There exists a large number of such TC trackers (Zarzycki and Ullrich, 2017), and they often represent a trade-off between accuracy and the ability to track early and late stages of the cyclones (Bourdin et al., 2022). For the present work, we are interested in methods that allow us to track the cyclones further in space and time than what is recorded in IBTrACS, providing information on the stages not recorded in the best tracks. This ability often comes at the cost of a high number of false alarms, as it requires the tracker to be very sensitive. However, since we will be matching the detected tracks to the observed ones, we will eliminate these false alarms.

The (Re)analysis-Augmented Best Tracks (RABTracks) dataset presented here is an attempt to leverage both product types and create a catalogue of historical tropical cyclones that gathers as much information about past storms as possible in a single place. We consider IBTrACS as the most accurate reference for defining the list of historical tropical storms, and the analyses as useful sources of additional information. The dataset was initially built for the Huracan project, which investigates cyclones of tropical origin reaching Europe, and is designed for students and researchers looking into case studies of historical storms or climatological studies on TC-related topics.

Section 2 provides information about the track data used in this dataset; Section 3 details our methodology for creating the dataset, illustrated with examples; and Section 4 provides validation of the methodology, along with a discussion of the strengths and caveats of the dataset. The conclusion in section 5 discusses the potential for extension of the dataset, and perspectives for analyses based on this data. Finally, section 6 provides information about the data and code distribution.

2 Track data

2.1 IBTrACS

The International Best Track Archive for Climate Stewardship (IBTrACS, Knapp et al., 2010; Gahtan et al., 2024) is the reference collection of cyclone best tracks. In the present dataset, we do not perform any pre-selection of the IBTrACS data: all available attributes are retained, and all tracks are used. Wind variables are converted to m/s, but no conversion is applied for gust time averaging (different centres provide 1-, 3-, or 10-minute-averaged winds), and we prefer to let the user apply the conversion they see fit if needed. Distances in nautical miles and feet have also been converted to km and m, respectively.

We treat each of the basins available from the IBTrACS download page independently: East Pacific (EP), North Atlantic (NA), North Indian (NI), South Atlantic (SA), South Indian (SI), South Pacific (SP), and Western Pacific (WP). Note that some TCs appear in several basin files due to the fact they span several basins. In that case, they will be available in the files corresponding to each of these basins.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Table 1. List of the (re)analyses data sources, and the period for which track data has been extracted.

Name	Resolution	Specific TC assimilation	TRACK	SyCLoPS
ECMWF-OP-AN	$\approx$ 25 to 9 km (T799 to TCo1379)	No	2006 - present	—
JRA3Q	$\approx$ 55 km (TL319)	Bogus vortices based on IBTrACS	1950-present	1948-2024
ERA5	$\approx$ 30 km (TL639)	1940-1978: IBTrACS wind and pressure used / 1979-present: None	1940-present	1980-2024
MERRA2	$\approx 50\,\mathrm{km}\,(0.5° \times 0.625°)$	Vortex relocation	1980-present	1980-2022
NCEP	$\approx$ 22 km (T574)	Vortex relocation & intensity corrections	1979-2024	—

For the present work, we assume IBTrACS is ground truth for what ever was a tropical cyclone (caveats are detailed in section 4.5). We provide data for all tracks in IBTrACS, leaving the choice of whether and how to filter the data to the user. Should the user decide to do any filtering, it is very easy to subset the data based on the information provided: for example, only tracks from a given date, that reached a given intensity, or that are provided by a certain agency.

2.2 (Re)analyses

In RABTracks, we included track data from two TC trackers and five analyses, listed in table 1 and described below. Each tracker-analysis couple is called a "source" hereafter.

We mainly integrate **contemporary reanalyses products**, namely ERA5 (Hersbach et al., 2020; Soci et al., 2024), JRA3Q (Kosaka et al., 2024), MERRA2 (Gelaro et al., 2017; McCarty et al., 2016), and NCEP-CFSR (Saha et al., 2010). The vortices and/or depressions in these products have been tracked using TRACK (Hodges et al., 2017) and/or SyCLoPS (Han and Ullrich, 2025) respectively (more details in 2.2.1). They cover different periods, and the data are provided for the entire period for which we had track data available (see fig. 1). Moreover, the different reanalyses differ in their resolution and assimilation procedures, particularly in whether they assimilated TC-specific information.

We also added data from the **ECMWF operational analysis** (European Centre for Medium-Range Weather Forecasts, 2006; Lang et al., 2023). Cyclones in this analysis were detected using TRACK for the period for which it is available in the TIGGE archive (Swinbank et al., 2016), i.e., since 2006. The main difference between reanalyses and operational analyses is that reanalyses use a single, consistent model and data assimilation system throughout the entire dataset, whereas operational analyses evolve with the operational systems. For the ECMWF operational analysis (EMCWF-OP-AN hereafter), over the 2006-present period, it is important to note that the resolution of the analysis varied from T799 ($\approx$25 km) in 2006 to T1279 ($\approx$16 km) in 2010 and TCo1279 ($\approx$9 km) since 2016 (ECMWF, 2026).

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

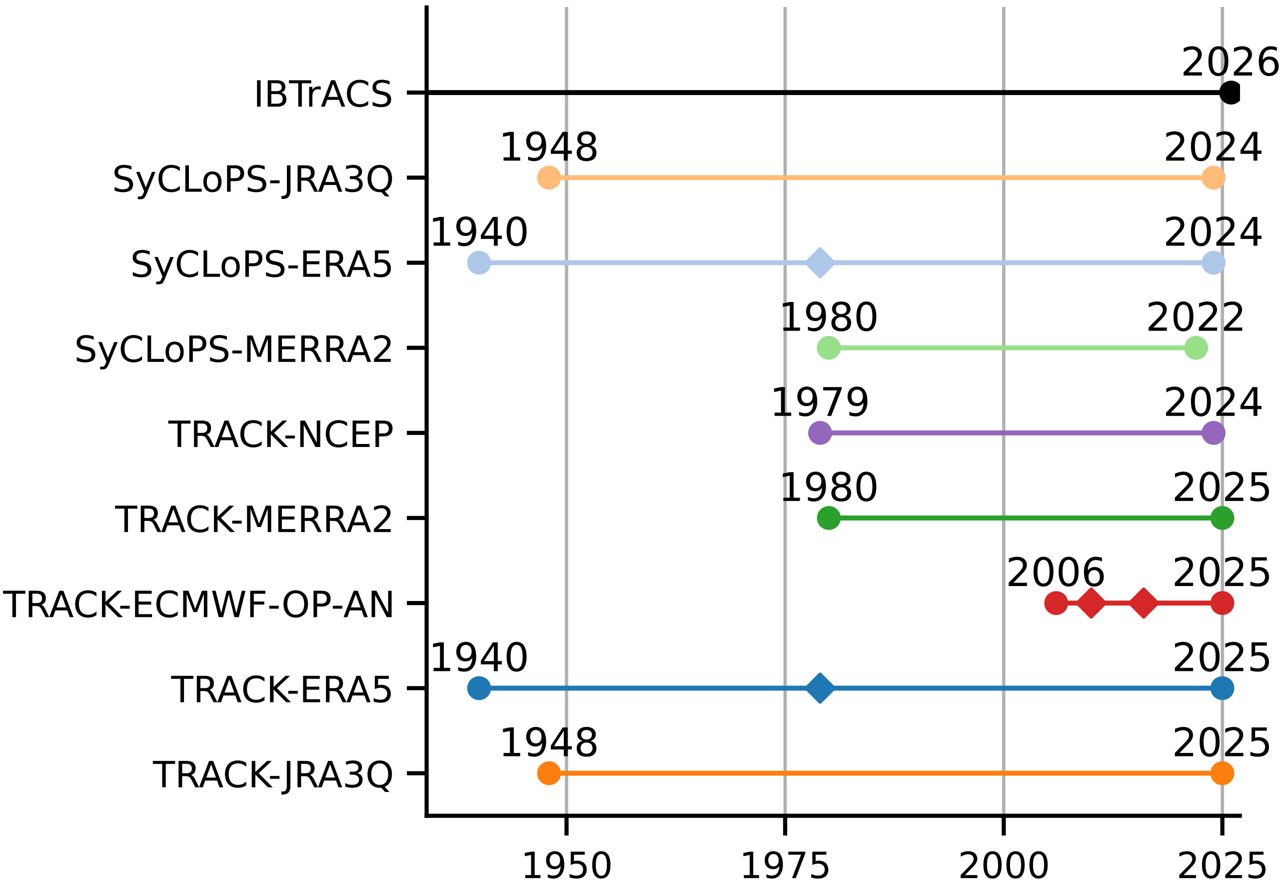

Figure 1. Temporal coverage of the different data sources as used in RABTracks. These correspond to the period for which the track data was available for each tracker, and do not necessarily reflect the temporal coverage of the underlying (re)analyses. Diamonds show breaking points: For ECMWF-OP-AN, they show the changes in resolution; for ERA5, they show the switch from the "back extension" (where TC information is assimilated) and the main product.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

2.2.1 Cyclone tracking

The (re)analyses mentioned above are distributed as gridded products. To obtain cyclone tracks from these sources, we need to use so-called "trackers" that identify cyclones within the gridded data.

The track data used here were obtained using two different tracking algorithms: TRACK (Hodges et al., 2017) and SyCLoPS (Han and Ullrich, 2025). These two methods were selected because they are specifically designed to capture as much of the storm lifecycle as possible. Accuracy in terms of False Alarms is not relevant here, since we will only keep the tracks that match IBTrACS, whereas track length is a desirable feature for our endeavour.

The **TRACK** algorithm uses the vertical average of vorticity between 850 and 600hPa truncated to T63 spectral resolution to reduce noise in the very noisy vorticity field at high resolutions and to provide a common resolution for the identification and tracking of TCs. Initially, all vorticity centres are tracked before TCs are identified. While the algorithm allows for identifying TCs using objective criteria, in the present work, we retain all vortices, and matching to observations serves as TC identification. This approach has been found to be able to identify and track the majority of TCs found in IBTrACS (Hodges et al., 2017) throughout their full lifecycle from weak precursor through to post-tropical cyclone stage (Baker et al., 2021; Sainsbury et al., 2020).

From the TRACK data, we keep the following information: longitude (lon), latitude (lat), minimum sea-level pressure (pres; the minimum SLP within 5 geodesic degrees of the vorticity center), maximum 10m wind speed (wind10; the maximum horizontal 10m wind speed within 6 geodesic degrees of the vorticity center), vorticity at 850hPa (vo850), longitude and latitude of the slp minimum (lon_pres, lat_pres) and of the wind maximum (lon_wind, lat_wind).

For the ERA5 and JRA3Q, Cyclone Phase Space parameters (Hart, 2003, more details in 3.4) have also been computed within the TRACK framework, for cyclones that TRACK identified as having a coherent vorticity structure (see Hodges et al. (2017) for TRACK's TC identification criteria). Based on these, Warm Core Symmetric Intensifying (WCSI) cyclones have been flagged, as described in Saffin et al. (2026) and motivated below.

SyCLoPS is an SLP-based tracking method that detects all types of surface low-pressure systems within a single framework, leveraging the computational efficiency of state-of-the-art feature detection software, TempestExtremes (TE; Ullrich and Zarzycki (2017); Ullrich et al. (2021)). SyCLoPS detects cyclone candidates on the model's native grid when SLP increases by at least 10 Pa in every direction within a 5.5° great-circle distance. TE outputs 16 atmospheric parameters based on geopotential height, temperature, SLP, horizontal winds, and relative humidity fields at specific pressure levels alongside each detected candidate. Tracks are formed if the candidates can be stitched into consecutive time steps of at least 18 hours. These parameters are then used to classify cyclones in a Python classifier, which stores the classification labels for each detected cyclone node and cyclone track. Technical details of the SyCLoPS framework are documented in the SyCLoPS paper (Han and Ullrich, 2025).

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

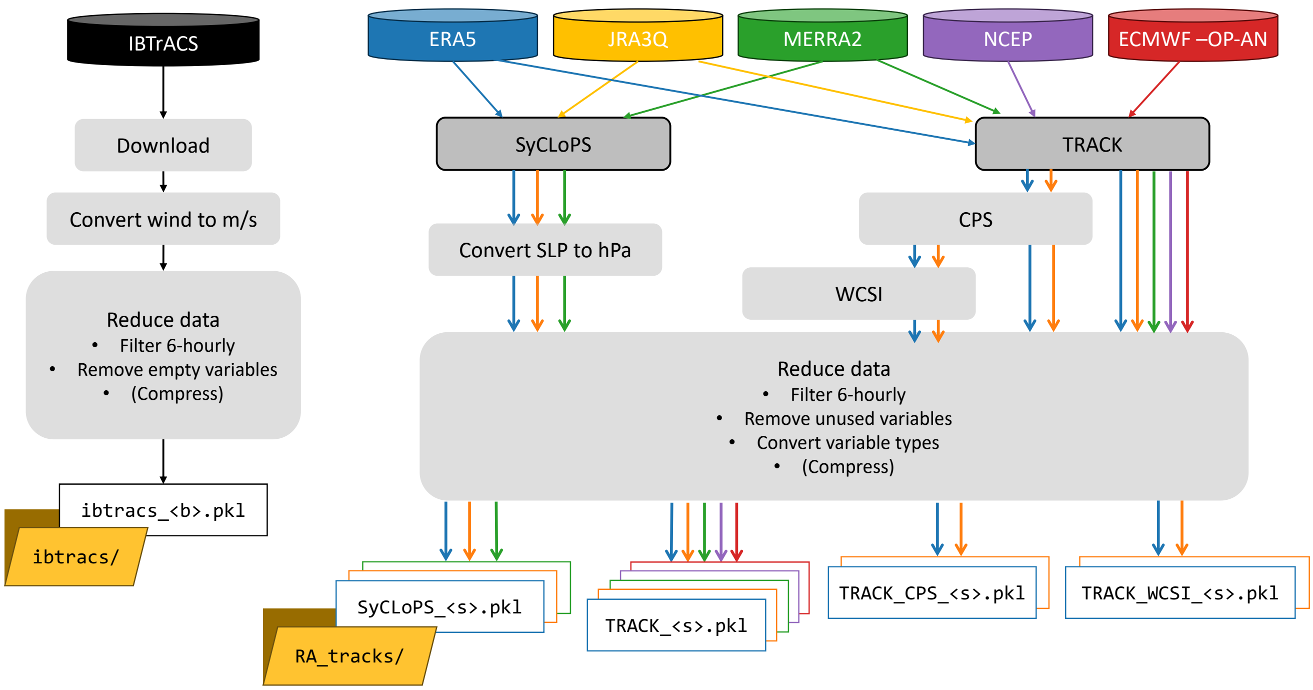

Figure 2. Flowchart of the tracking and data preparation procedure. SyCLoPS and TRACK are two TC trackers (see 2.2.1). The CPS (Cyclone Phase Space) box corresponds to the computation of the CPS parameters within the TRACK framework, and WCSI (Warm Core Symmetric Intensifying) corresponds to the flagging of WCSI storms following Saffin et al. (2026).

From the SyCLoPS data, we keep the following information: longitude (lon), latitude (lat), minimum sea-level pressure (pres), maximum 10m wind speed (wind10; the maximum horizontal 10m wind speed within 2° great-circle distance of the pressure center), and adjusted label (label), low-pressure system area (lpsarea), tropical flag (tropical_flag).

Once the tracks are obtained from the tracker, additional processing steps are performed to make the data homogeneous and lighter: where necessary, pressures are converted to hPa; only 6-hourly time steps are kept across all sources; unused variables are removed.

The whole data preparation process is illustrated in figure 2.

3 Creation of the dataset

The procedure for creating the dataset, described below, is illustrated in figure 3.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

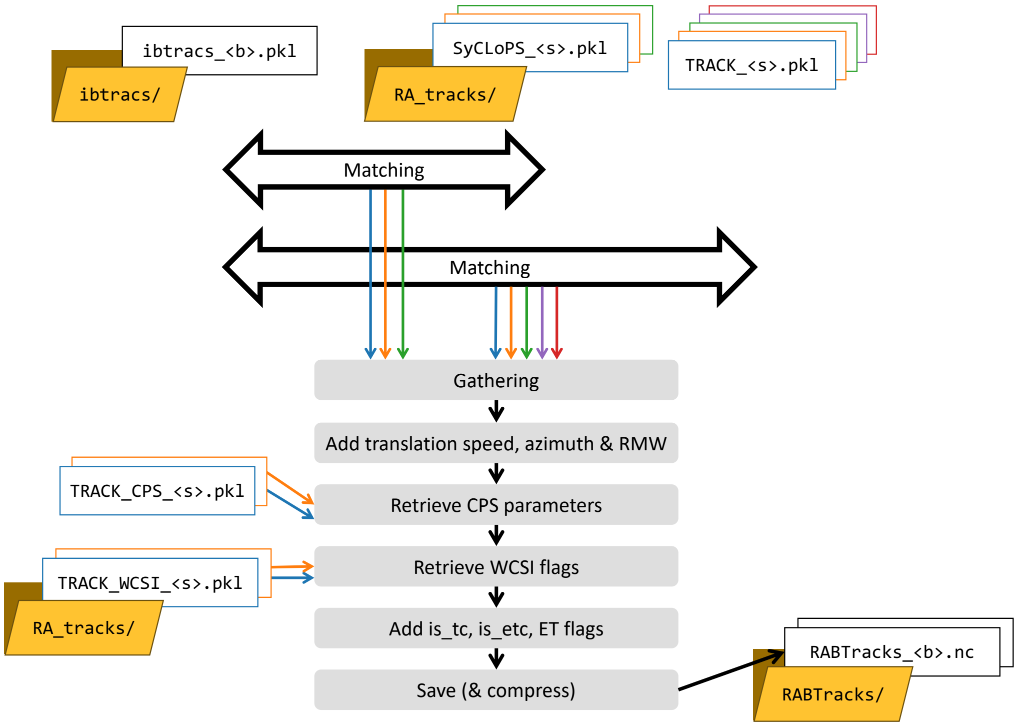

Figure 3. Flowchart of the creation of the dataset. CPS: Cyclone Phase Space; WCSI: Warm Core Symmetric Intensifying; ET: Extra-tropical Transition; TC: Tropical Cyclone; ETC: Extra-Tropical Cyclone.

3.1 Track matching

3.1.1 Procedure

Each cyclone in IBTrACS is matched with the corresponding track in each of the (re)analysis products, if it exists. The matching procedure follows that of Bourdin et al. (2022) (itself similar to that of Hodges et al., 2017):

> “Consider the case of a given detected track D composed of n points $(d_1, d_2, ...d_n)$ defined at times $(t_1, t_2, ..., t_n)$. The observations O consist of a database of tracks and can be seen as a collection of points at given times. For each point $d_i(t_i)$ of the track D, we associated those points of O at time t_i that are located closer than $[d_{max}]$ km from the point d_i. Of course, it is possible that such points do not exist in O. The subset of points of O that have been associated with any point in D is noted $O_{\text{D–paired}}$. It is composed of $|O_{\text{D–paired}}|$ elements. There are three possibilities:
>
> 1. $|O_{\text{D–paired}}| < N_{min}$: [Less than N_{min}] points of D have been paired to a point in O. D is considered to be a false alarm.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

2. $|O_{\text{D–paired}}| \geq N_{min}$ and all the points in $O_{\text{D–paired}}$ belong to the same track D_O in O: D_O is considered to be the match of D.
3. $|O_{\text{D–paired}}| \geq N_{min}$ and the points in $O_{\text{D–paired}}$ belong to more than one track in O: the observed track having the largest number of points paired with D is considered the match of D.

Here, d_{min} was set to 400 km, and $N_{min} = 3$. After this matching is completed for all detected tracks, a final treatment is performed: if an observed track is paired with two or more detected tracks, these detected tracks are merged into a single track. Such cases arise when the detected track corresponds to different parts of the same observed tracks and occur when, for example, the TC temporarily weakened while going over an island before strengthening again."

All distances are computed using the haversine formula. Note that following this procedure, only one point is needed to match two tracks. False alarms are discarded. Observed tracks without corresponding reanalysis tracks are retained, and the corresponding reanalysis data are set to NaN. No matching is done among the (re)analyses themselves.

With this procedure, there are time steps for which data may exist for at least one but not all data sources. In this case, the time steps are present but, NaNs are stored for any source that does not have data. For example, if the reanalysis' track starts earlier than the IBTrACS one, the earlier time step will be present in the dataset, but the IBTrACS data for those time steps are NaNs.

3.1.2 Example

Figure 4 illustrates the track matching procedure. For three tropical cyclones in three different basins, we can see that the algorithm identifies analysis cyclone tracks that match well with the IBTrACS tracks when they exist. We also see that the matched analysis tracks are often longer than the IBTrACS one, providing information about the storm's lifecycle before and after the IBTrACS record.

In the case of Gonzalo (2014), three of the matched reanalysis tracks start days before IBTrACS, and they allow tracing the cyclone back to what was likely an African Easterly Wave. All of the matched reanalysis tracks end after the IBTrACS record, and they all show that the storm continued to move towards Europe. In JRA3Q, both trackers even track it all the way to Italy, Greece, and the Mediterranean Sea.

As a matter of fact, according to the "List of European Tropical Cyclones" Wikipedia page, several media[1,2] at the time reported remnants of Gonzalo causing strong winds and floods from the British Isles to Greece. Our database reveals this extended trajectory of the storm.

Similarly, cyclone Bruce (2013) is tracked in the reanalyses further than in IBTrACS. Even though Bruce began being monitored by the BOM on December 16th, we can identify in the reanalyses a precursor disturbance over the maritime continent since December 9th. The operational monitoring stopped on December 24th as the cyclone made an extra-tropical transition.

[1] BBC (21 October 2014). "Hurricane Gonzalo: Woman hit by tree in London dies", https://www.bbc.co.uk/news/uk-29700594
[2] "Hurricane Gonzalo live: two dead as winds batter Britain". Telegraph.co.uk. 21 October 2014

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

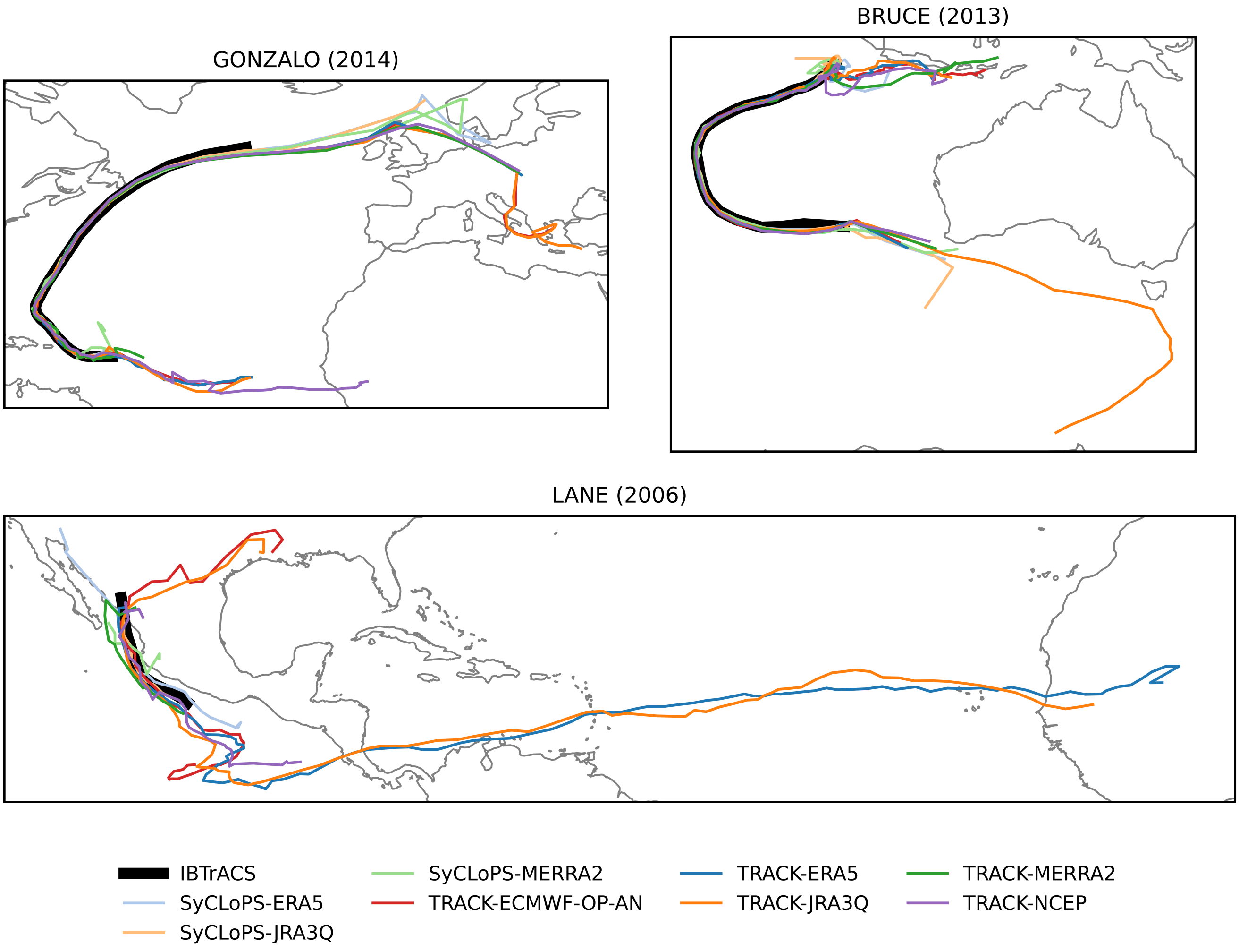

Figure 4. Examples illustrating the track matching: The IBTrACS tracks (thick black line) have been matched to a cyclone track in each source (colored lines)

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

This extra-tropical cyclone can be followed for two more days in most reanalyses. In JRA3Q, TRACK even identifies a track that goes all the way to Tasmania and then recurves polewards and approaches Antarctica.

As for Hurricane Lane (2006), we can see that the analysis tracks start earlier than the IBTrACS record, with two in particular (TRACK-ERA5 and TRACK-JRA3Q) extending back to Africa. This means that Hurricane Lane could probably be traced back to an African Easterly Wave that crossed the Atlantic (Serra et al., 2010).

3.2 Gathering

3.2.1 Procedure

Once each observed track is matched to as many analysis tracks as possible, we gather all the information back into a single xarray object. Doing so, we immediately retrieve all additional attributes about the storm available from IBTrACS and the tracking algorithms, namely: all IBTrACS columns; wind and pressure intensity from all sources; vorticity and position of the SLP minimum and wind maximum from TRACK; labels and LPS area from SyCLoPS. This allows us to have in the same place different sources of information about a storm's intensity, nature and size.

3.2.2 Examples

Figure 5 shows, as an example, all the intensity information available for Hurricane Gordon (2006).

Note that the SyCLoPS lines for ERA5 and JRA3Q are not visible on the plot because they overlap exactly with the TRACK ones from the same reanalysis, since both trackers capture the same cyclone and retrieve the analysis's intensity along the detected track. The SyCLoPS-MERRA2 trajectory jumps away from the IBTrACS trajectory because MERRA2 does not have the storm's surface signature at this point, so the algorithm links with another disturbance to the north.

TC intensity is a known shortcoming of reanalyses (Hodges et al., 2017; Dulac et al., 2022). We can see significant discrepancies across all the sources. While we can see an effect of resolution, with MERRA2 and NCEP simulating lower intensities than ERA5 and ECMWF-OP-AN, the JRA3Q reanalysis stands out with a simulated intensity much closer to the IBTrACS record, despite a lower resolution than ERA5 and ECMWF-OP-AN. This was observed for many cyclones in the database (not shown). We attribute this performance to the assimilation (both the fact that it is done, and the method to do it) of IBTrACS information in JRA3Q.

We also note that the analyses' intensity bias decreases with time. This is because, over time, the cyclone moves poleward, and by the third intensity peak, Gordon had become an extratropical cyclone. We observed this for many cyclones (not shown), and we attribute it mainly to the fact that ETCs are better represented than TCs in analyses, because they are larger, and because the analysis models do not resolve better baroclinic processes at play for the extra-tropical phase than the the deep convective or mesoscale processes that are critical to the intensification mechanism while in the tropical phase.

Finally, while the IBTrACS record stops on Sept. 25th, the cyclone continued to be tracked in MERRA2, NCEP and ERA5. The intensity data for these sources reveal a fourth, weaker intensification cycle that was not captured in IBtrACS.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

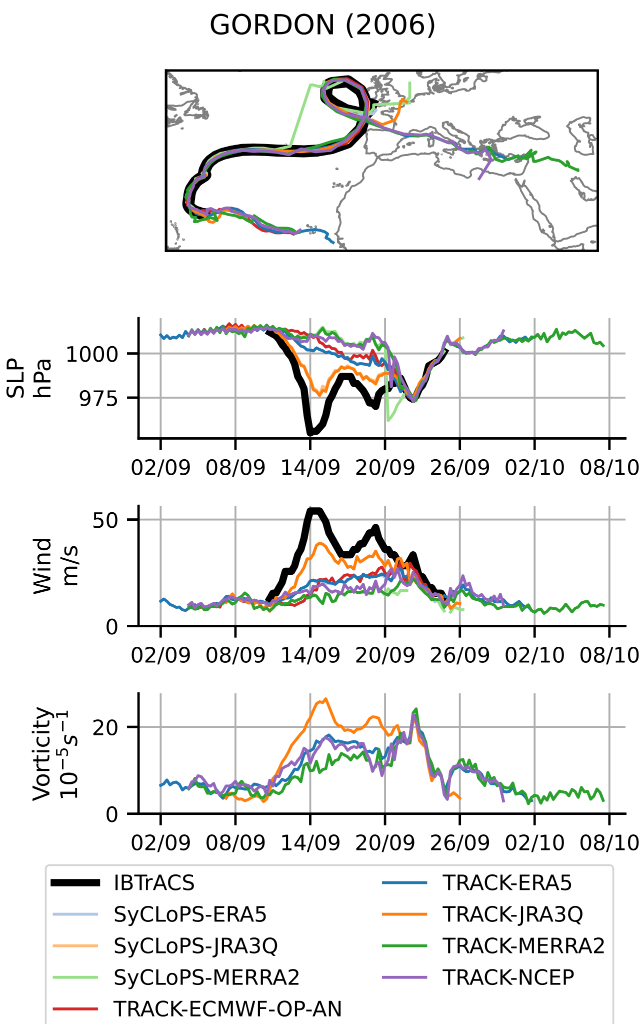

Figure 5. Illustration of the intensity information available in RABTracks for cyclone Gordon (2006): Trajectory is shown at the top, then time series of Sea-Level Pressure (SLP), 10m surface wind speed, and 850-hPa vorticity

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

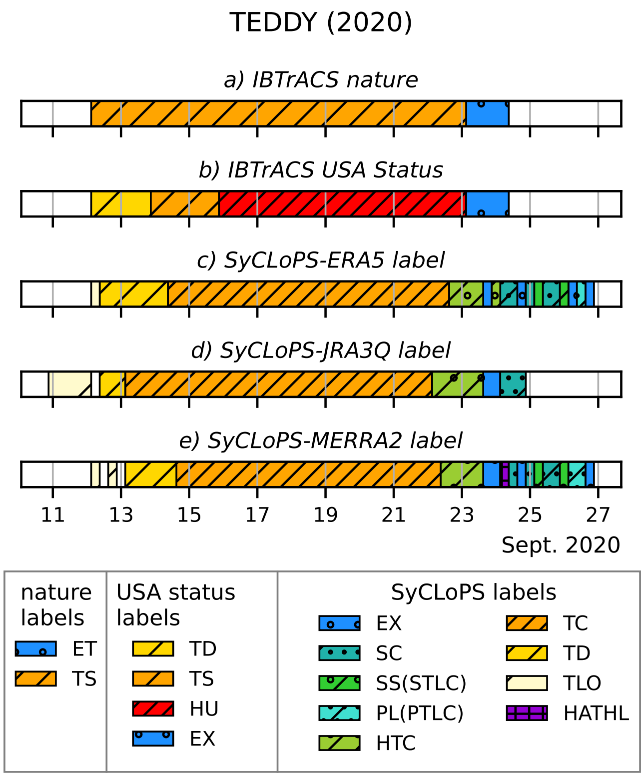

Figure 6. Illustration of the nature information available in RABTracks for one cyclone, through the IBTrACS' nature and USA status labels, and the classification by SyCLoPS

Figure 6 shows, as an example, all the nature information available for Hurricane Teddy (2020).

225 In IBTrACS, we see that the USA's NHC classified the storm successively as a Tropical Depression (TD), and Tropical Storm (TS), and a Hurricane (HU), before they considered the storm was no longer tropical on Sept. 23rd, and classified it as extra-tropical (ET) (panel b). The "nature" label is based on the NHC classification, and gathers the tropical TD, TS and HU stages under a single "TS" label (panel a).

Using the SyCLoPS algorithm (panels c,d,e), the cyclone is identified successively as Tropical Low (TLO), Tropical Dis-
230 turbance (TD) and Tropical Cyclones (TC). SyCLoPS's "TC" label is designed to be analogous to NHC's "TS". In ERA and MERRA2, we see that the storm is identified as TC on Sept. 14, as in IBTrACS, whereas in JRA3Q it is classified as TC about one day earlier. After a long TC stage, it becomes a hybrid TC, then it oscillates between subtropical (SC) and extratropical (EX) statuses.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

3.3 Translation speed, azimuth and radius of maximum wind

3.3.1 Procedure

For all the tracks, translation speed and azimuth (direction of propagation) are computed using the HuracanPy package (Bourdin and Saffin, 2025). Translation speed is computed in m/s as dx/dt, where dx is the geodesic distance between two consecutive points in m (computed using the `pyproj` package), and $dt = 6\,h$ is the time between two consecutive points. Forward Azimuth is provided in degrees using geodesic calculations from the `pyproj` package. 0° corresponds to the North, 90° to the East, 180° to the South and 270° to the West.

Radius of Maximum Wind (RMW) is computed as the distance between the SLP minimum and the 10m wind speed maximum for each TC point. It has been computed only for TRACK sources, since SyCLoPS does not retrieve the wind speed maximum position by default. Note that this is a simplistic way to define the RMW, based on the available attributes, but most of the literature uses more elaborate definitions, often the radius of the maximum azimuthally-averaged tangential wind (e.g. Torgerson et al., 2023; Chavas and Knaff, 2022).

Since both the azimuth and the translation speed can be noisy when computed from 6-hourly data, a smoother version, computed using 1-day rolling-mean coordinates, is also provided. The 1-day rolling mean of the RMW is also provided.

3.3.2 Example

Figure 7 illustrates the computation of the azimuth and translation speed of Hurricane Beryl (2024). In the top panel, we can see Beryl's trajectory in the Atlantic. Beryl had a typical C-shaped trajectory, starting in the tropical Eastern Atlantic, moving north-westwards under the trade winds, recurving northward as it approached continental North America, and then recurving further north-eastward as it progressed over land. At the point where the IBTrACS record stops, the storm is reported to have merged with a frontal system (National Hurricane Center, 2024), and RABTracks extensions allow us to follow the remnants of the storm. In the middle panel, we can see that the azimuth is recorded to shift indeed from West to North, then to East. On the bottom panel, we can see the storm accelerating before it reached the Caribbean, then decelerating in the Caribbean Sea and the Gulf of Mexico, accelerating after its continental landfall, decelerating over land, accelerating again as it became extra-tropical, and finally decelerating again as it dissipated.

Figure 8 shows all the size information available for Cyclone Bruce (2013).

We have information about the RMW provided by three different agencies for this cyclone: Météo France's Regional Specialised Meteorological Centre (La Réunion), the Australian Bureau of Meteorology (BOM), and the USA's Joint Typhoon Warning Center (JTWC); and computed from TRACK data. The RMW from all three agencies match well and show that the cyclone's eye became smaller during the intensification phase (peak intensity was reached on Sept. 21st) and then enlarged during the decay phase. In the analyses, we see that the cyclone was larger prior to being recorded in IBTrACS, followed a similar pattern over the time when the operational agencies monitored it, and enlarged after the extra-tropical transition. The discrepancy between the values in RMW in the TRACK-based data and the observations is likely due to differences in the RMW computation method and to the poor representation of TCs' inner structure in reanalyses (Dulac et al., 2023).

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

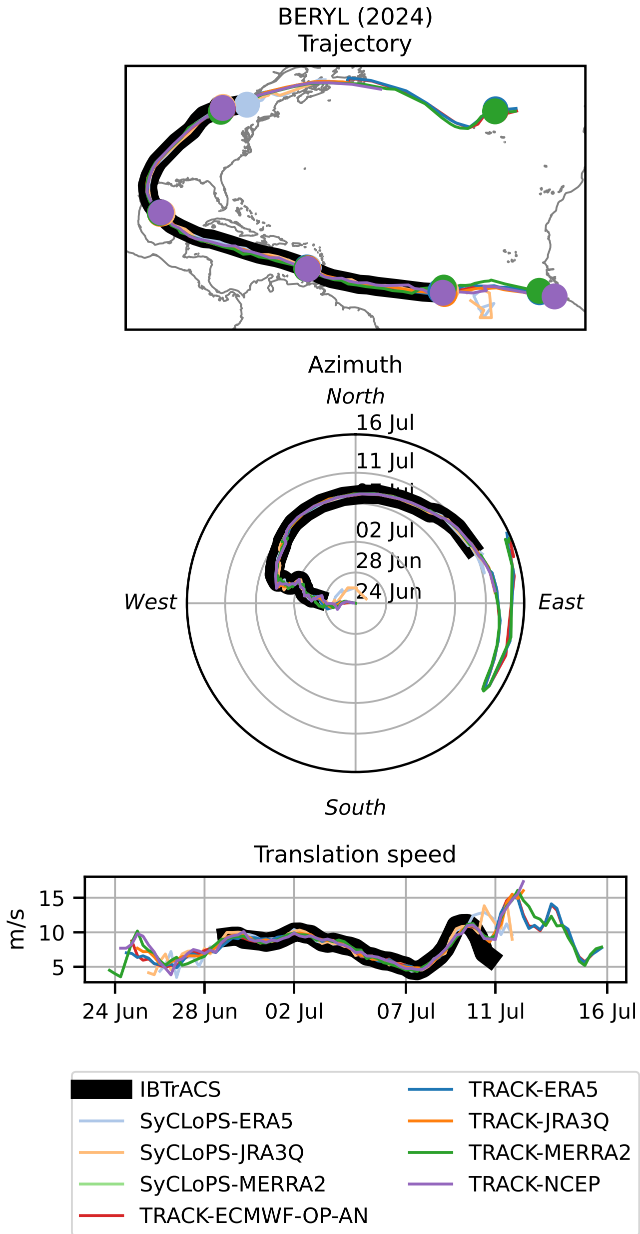

Figure 7. Illustration of the azimuth and translation speed computation: The top left panel shows the trajectory of Beryl (2024) in the North Atlantic, according to all sources; the top right panel shows the azimuth as a function of time; the bottom panel shows the translation speed over time. The dots on the trajectory panel correspond to the same times are the r- and x- ticks on the azimuth and translation speed panels respectively.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

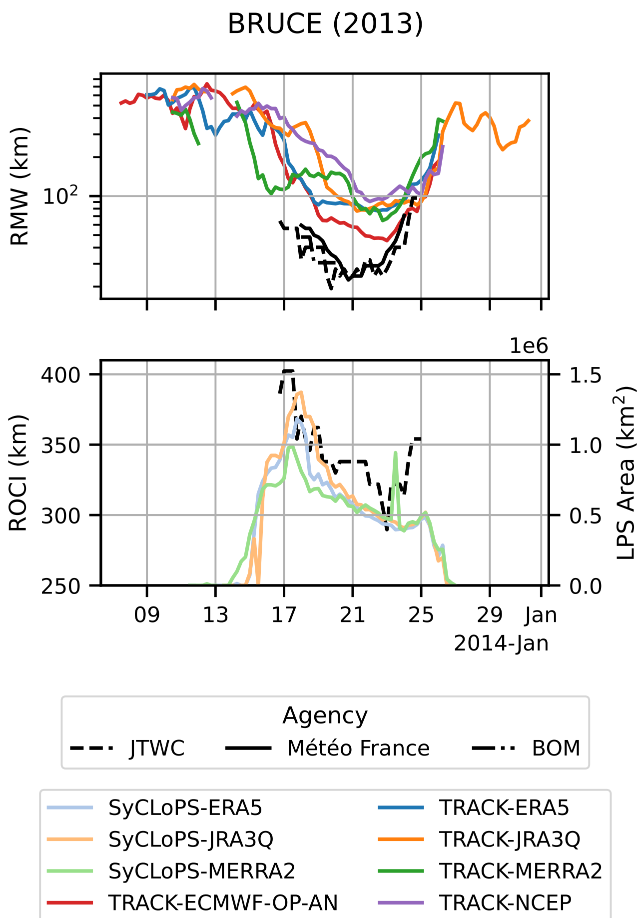

Figure 8. Illustration of the size information available in RABTracks for cyclone Bruce (2013), including Radius of Maximum Winds (RMW) and Radius of the Outermost Closed Isobar (ROCI) provided by the agencies within IBTrACS, Radius of Maximum Winds (RMW) computed from TRACK attributes, and SyCLoPS Low Pressure System (LPS) area.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

For the outer size of the cyclone, we have reports of the Radius of the Outermost Closed Isobar (ROCI) from the JTWC only, complemented by the Low Pressure System (LPS) area computed within the SyCLoPS algorithm. In the case of Bruce, all four show a consistent evolution in the outer size of the cyclone, with an increase in size as the cyclone forms, followed by a gradual decrease over the rest of its lifetime.

3.4 Cyclone Phase Space parameters

The Cyclone Phase Space (CPS) was introduced by Hart (2003) as a tool to classify cyclone stages. It has since become a standard method, in particular to identify extra-tropical transitions (Jones et al., 2003; Evans et al., 2017), and classify hybrid cyclones such as medicanes (Miglietta, 2023).

The CPS is based on three parameters: A symmetry parameter, $B = (Z_{600} - Z_{900})|_{\mathrm{L}} - (Z_{600} - Z_{900})|_{\mathrm{R}}$ which corresponds to the difference in the thickness of the 900-600 hPa layer between the left and right half of the cyclone within 500 km of the cyclone center, relative to the direction of propagation; and two parameters for the lower and upper thermal winds:

$$-|V_{\mathrm{T}}^{\mathrm{L}}| = \frac{\partial \Delta Z}{\partial lnP}\bigg|_{900\mathrm{hPa}}^{600\mathrm{hPa}}, \; -|V_{\mathrm{T}}^{\mathrm{U}}| = \frac{\partial \Delta Z}{\partial lnP}\bigg|_{600\mathrm{hPa}}^{300\mathrm{hPa}}$$

where ΔZ is the difference between the maximum and minimum geopotential height within a radius of 500 km of the cyclone centre, and the gradient is calculated using the difference between the two given levels.

3.4.1 Procedure

In our case, the Cyclone Phase Space parameters have been computed for the TRACK-ERA5 and TRACK-JRA3Q sources using TRACK, using the standard levels and radius (900, 600 and 300 hPa, and 500 km, following Hart, 2003). Note that there is no specific reason why the CPS parameters could not be computed for the other sources, except for the computational cost, which led us to restrict to two sources.

Given that the high-frequency CPS parameters can be quite noisy, a 24h rolling mean of these parameters is also computed and provided. This is a common practice in the field to define cyclone stages (e.g. Bieli et al. (2020)).

Based on these CPS parameters, we define three labels:

- Time steps where $|B| \leq 15\,\mathrm{m}$ and $V_{\mathrm{T}}^{\mathrm{L}} \geq 0\,\mathrm{m}$ are labelled **Warm Core Symmetric (WCS)**.

- If $|B| \leq 15\,\mathrm{m}$, $V_{\mathrm{T}}^{\mathrm{L}} \geq 0\,\mathrm{m}$ and $V_{\mathrm{T}}^{\mathrm{U}} \geq 0\,\mathrm{m}$ (symmetric deep warm core) and the cyclone is intensifying (in terms of vorticity) for at least one day, the cyclone is labelled as **Warm Core Symmetric Intensifying (WCSI)**. It was shown in Saffin et al. (2026) that combining the CPS parameters with an intensification criterion makes for a good discriminant of TC stages.

- Time steps where $|B| > 15\,\mathrm{m}$ and $V_{\mathrm{T}}^{\mathrm{L}} < 0\,\mathrm{m}$ are labelled Cold-Core Asymmetric (CCA).

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

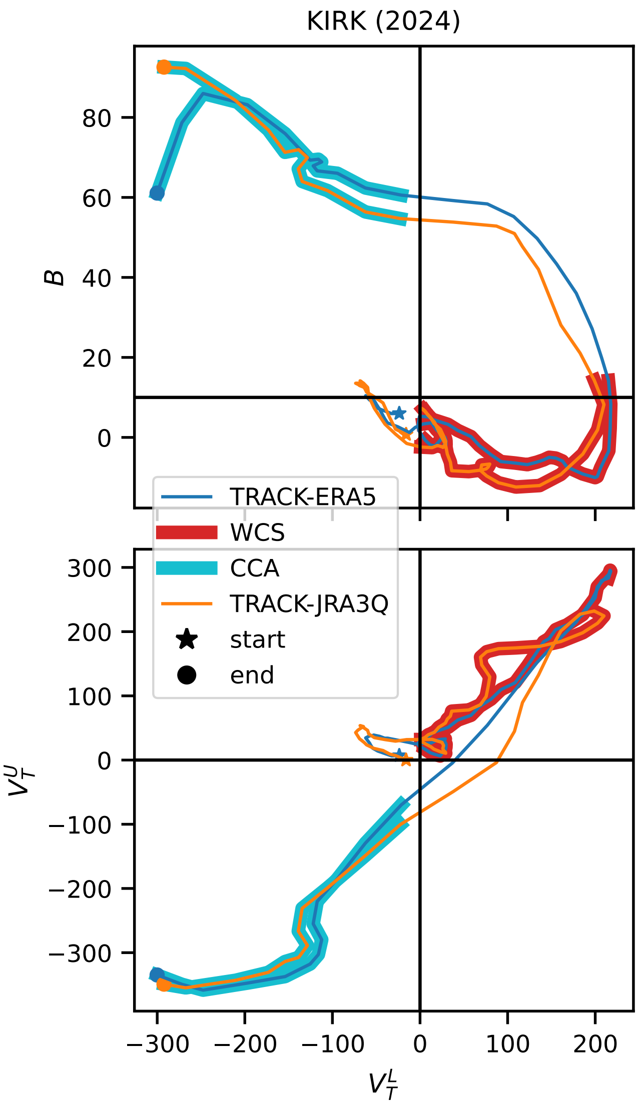

Figure 9. Cyclone Phase Space parameters for Kirk (2024). The red and blue highlights show the WCS and CCA phases, respectively.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

3.4.2 Example

Figure 9 shows the CPS parameters for Hurricane Kirk (2024), along with identified WCS and CCA periods. The ERA5 and JRA3Q data agree well on the phase space trajectory. Kirk started as a symmetric (B < 10 m), neutral ($V_T^U \& V_T^L \approx 0$) structure. It then evolved into a deep warm core ($V_T^U \& V_T^L > 0$) symmetric cyclone, time during which it was identified as Warm Core Symmetric (WCS). The WCS phase ended when the storm became asymmetric (B > 10 m) because it developed a warm sector as it crossed into a baroclinic zone. Long after, the storm lost it warm core and became cold core ($V_T^U \& V_T^L < 0$) asymmetric (CCA).

3.5 Extratropical Transition (ET)

3.5.1 Procedure

For each of the cyclones, we flag the potential onset and offset of extra-tropical transition. To do so, we define the onset as the last point when the cyclone is considered to be a TC, and the offset as the first point after the onset when the cyclone is considered to be an ETC.

"Being a TC" (`is_tc` variable in the dataset) and "being an ETC" (`is_etc` variable in the dataset) is defined differently depending on the source:

- For IBTrACS, the 'nature' "TS" and "ET" labels define the TC and ETC stages, respectively.
- For SyCLoPS sources, the adjusted short labels "TD"+"TC" and "EX" define the TC and ETC stages, respectively.
- For TRACK sources where the CPS parameters are available, we define the TC stage as the time when the cyclone is WCSI plus the preceding and following WCS time steps, and the ETC stage when the cyclone is CCA;
- For TRACK sources where the CPS parameters have not been computed, we cannot define the ET phase.

Note that using CCA for defining ETC stage is unexact: it has been shown that many extra-tropical cyclones exhibit a warm core in the CPS, especially when undergoing warm seclusion (Besson et al., 2025), including the example below (fig. 10). Here, we keep this simplistic definition because our primary goal is to define the time of ET completion, which has traditionnally been defined as the time when the cyclone became CCA (Hart, 2003; Evans et al., 2017).

3.5.2 Example

These flags are shown for hurricane Teddy (2020) in figure 10.

Panel a is the same as figure 6a, and shows the successive classification of the cyclone as TS and ET in IBTrACS. The corresponding `is_tc` and `is_etc` are shown in panel b, and the ET is considered instantaneous since the change in status from TS to ET occurs in two successive time steps.

Panels c,e,g are the same as figure 6's panels c-d-e and show the classification of the cyclone by the SyCLoPS algorithm. The corresponding `is_tc` and `is_etc` flags are shown in panels d, f, h. For all three reanalyses, the end of the TC stage

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

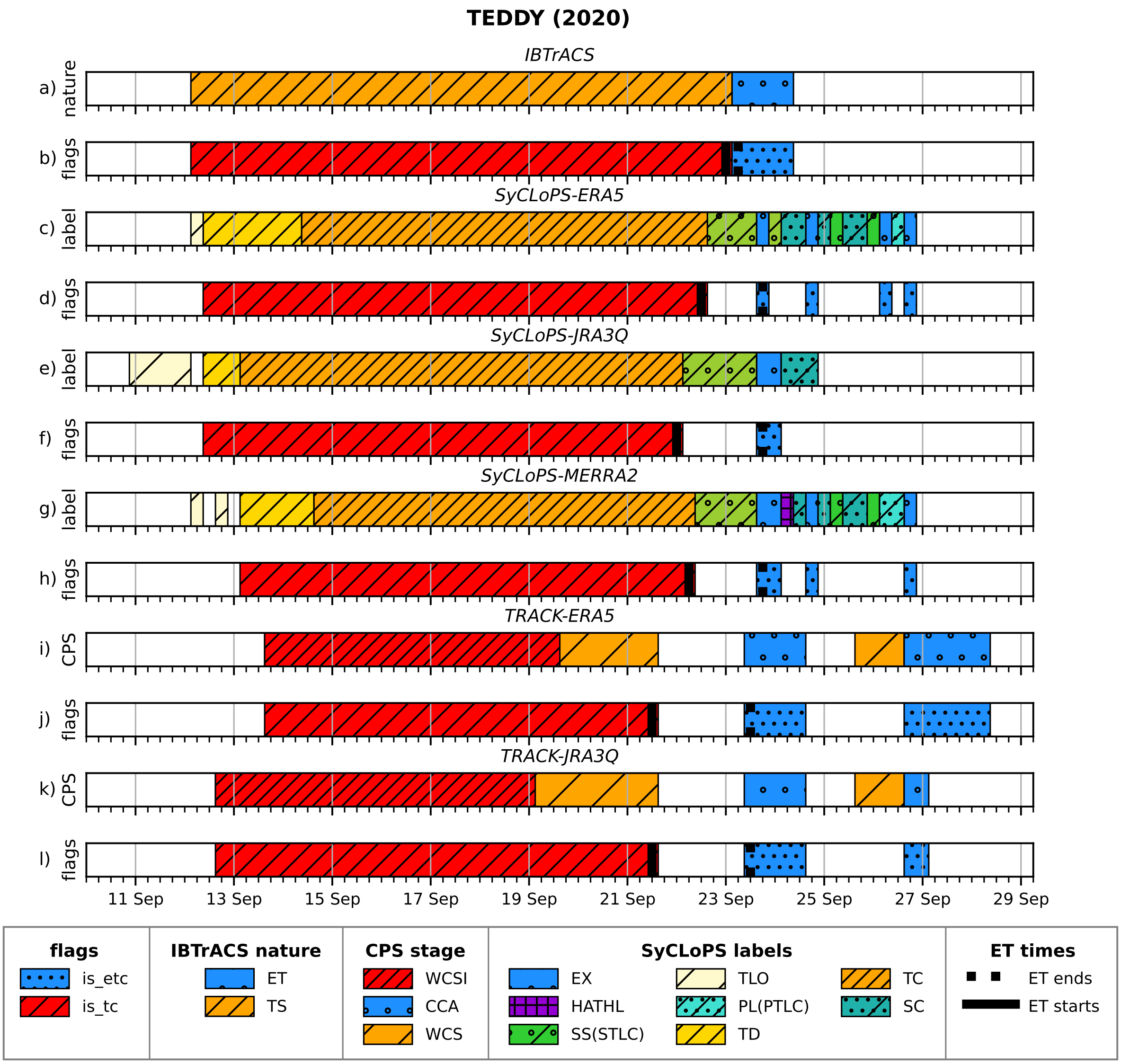

Figure 10. Illustration of the ET identification procedure: for each source, the identification variable is shown (**nature** for IBTrACS, **label** for SyCLoPS, **CPS** parameters for TRACK) over time; the corresponding `is_tc` and `is_etc` **flags** are shown, as well as the identified time of ET onset and offset.

is detected on Sept. 22nd by SyCLoPS, although at slightly different times depending on the reanalysis. The onset of ET is flagged accordingly. In all three reanalyses, the first ETC point is detected at 18:00:00 on Sept 23rd, marking the end of the transition.

Using TRACK and the CPS (panels i, k), the cyclone is identified as Warm Core Symmetric (WCS) from Sept. 12th (resp. Sept. 13th) in ERA5 (resp. JRA3Q), until Sept. 21st. During the first part of this period, the cyclone is also intensifying, hence the WCSI label. In both reanalyses, the cyclone becomes Cold Cold Asymmetric (CCA) on Sept. 23rd for less than 48 h, then briefly returns to WCS before ending as WCS. The `is_tc` stage spans the whole first WCS stage, and ends on Sept 21st, when ET is therefore declared to start. ET is considered to end on Sept. 23rd, when the first `is_etc` flag corresponds to the CCA status.

The timing of the ET onset is consistent among the SyCLoPS sources (Sept. 22nd), and among the TRACK sources (Sept. 21st). However, ET onset occurs earlier in TRACK sources, where the TC stage is identified to be shorter. ET offset is detected on Sept 23rd for all sources. The reasons for these differences are likely linked to differences in methodology and are not investigated here.

4 Validation, added value and caveats

4.1 Hit Rates

In the previous section, we have shown that the matching procedure works for three examples. Here, we provide a more extensive statistical validation.

First, we check the hit rates: The proportion of the observed tracks for which a corresponding (re)analysis track is found (fig. 11). For a fair comparison, these are computed only over the 2006-2022 period, when all sources are available.

Over all basins and all sources, the minimum hit rate is 90.6% (SyCLoPS-MERRA2 in the NI), and the maximum is 100% (TRACK-MERRA2 in in the NA), with a mean and median at 97.4% and 98.6% respectively. These rates are very high and comparable to those reported in previous literature (Bourdin et al., 2022; Han and Ullrich, 2025; Hodges et al., 2017). They mean that for the large majority of the observed storms, we find the corresponding trajectory in every source.

There are small discrepancies among basins, with a minimum average hit rate of 95% in the NI, compared to a maximum of 98.3% in the WP. Across the three reanalyses where both trackers were applied (ERA5, JRA3Q, MERRA2), TRACK yields slightly higher hit rates, averaging 98.7%, compared to 95.7% for SyCLoPS.

We know that observed cyclones that cannot be found in the reanalyses are usually short-lived weak cyclones that the analysis model failed to intensify (Hodges et al., 2017; Bourdin et al., 2022; Saffin et al., 2026; Zarzycki and Ullrich, 2017). Comparing the rates for all cyclones against those computed only for cyclones that reached category 1 or above (solid vs. hollow bars in figure 11), we find higher hit rates: between 92.4% and 100%, with a 99.4% mean and 100% median. It means we are more likely to find corresponding tracks for intense cyclones.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

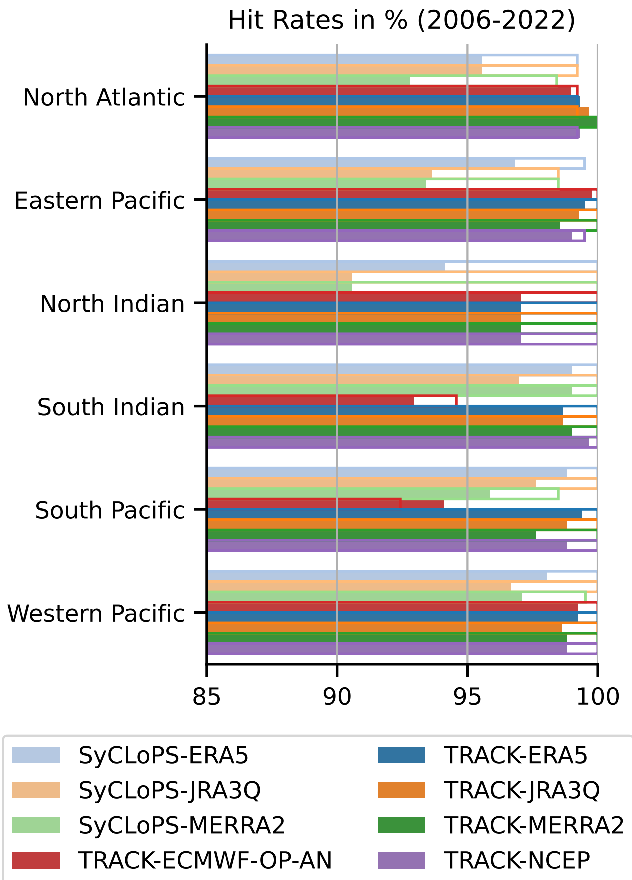

Figure 11. Hit Rates (proportion of observed cyclones matched with a (re)analysed cyclone) for each source and each basin, over the period 2006-2022 common for all sources. Note that the x-axis was cropped. The solid bars show the values when all cyclones in IBTrACS are taken into account, and the hollow bars show the values when only considering cyclones of category 1 and more.

It should be noted that the hit rates can vary in time. In particular, for the reanalyses that cover earlier periods (ERA5 from 1940, JRA3Q from 1948), we find a significant drop in hit rates before 1980, for both trackers (not shown), probably indicative of a reduced ability of the reanalysis to represent tropical cyclones accurately before the satellite era.

4.2 Matching distance

Another metric we examine is the matching distance, i.e., the distance between two matching tracks during the time they match. This is computed as the mean of the geodesic distances between each matched point, and the distributions are shown in figure 12. Note that this distance can exceed 300 km, the distance threshold for the matching, because there might be time

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

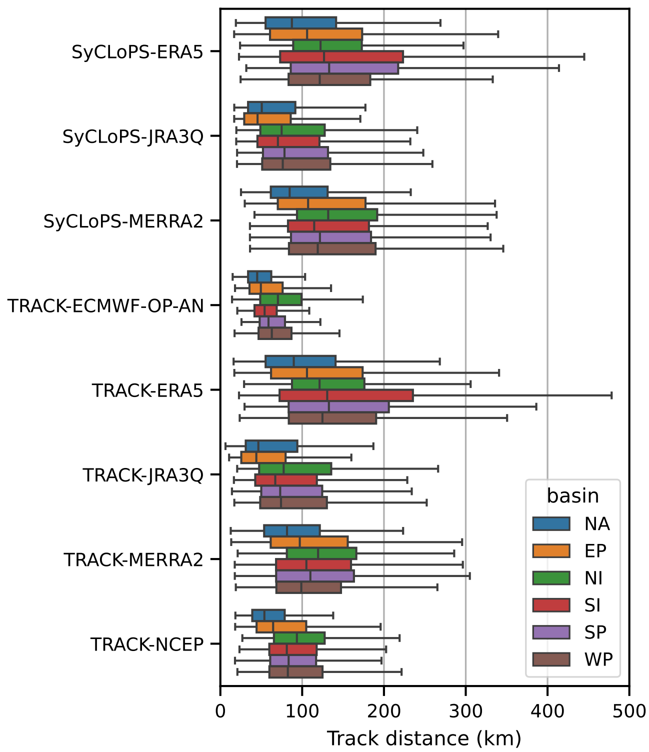

Figure 12. Boxplot of the distributions of the track-wise matching distance in each source and basin. In all cases, the position of the SLP minimum was used

steps over the time when both tracks exist when the two tracks are more than 300km apart. The matching imposes that there is at least one time step when both tracks are within 300km of each other.

We can see that across all sources, most tracks match with a distance of the order of a few grid cells, which is satisfactory. The ECMWF operational reanalysis stands out with very small matching distances, which is likely associated with its finer resolution.

4.3 Track extension

The third and last metric we check is the delay between the start and end of the matched (re)analyses tracks compared to IBTrACS, shown in fig. 13, which illustrates the relevance of our dataset. The large majority (about or more than 75% in all sources and basins) of the matched analysis tracks start earlier and end later than the IBTrACS record. This validates our choice of SyCLoPS and TRACK as tracking methodologies to extend the IBTrACS record and clearly shows the benefit of this approach in reconstructing more of the storms' lifecycles.

More precisely, we observe that the TRACK tracks tend to start even earlier than the SyCLoPS tracks. This is a known difference between vorticity- and SLP-based trackers (Bourdin et al., 2022), stemming from the fact that the vortex typically

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

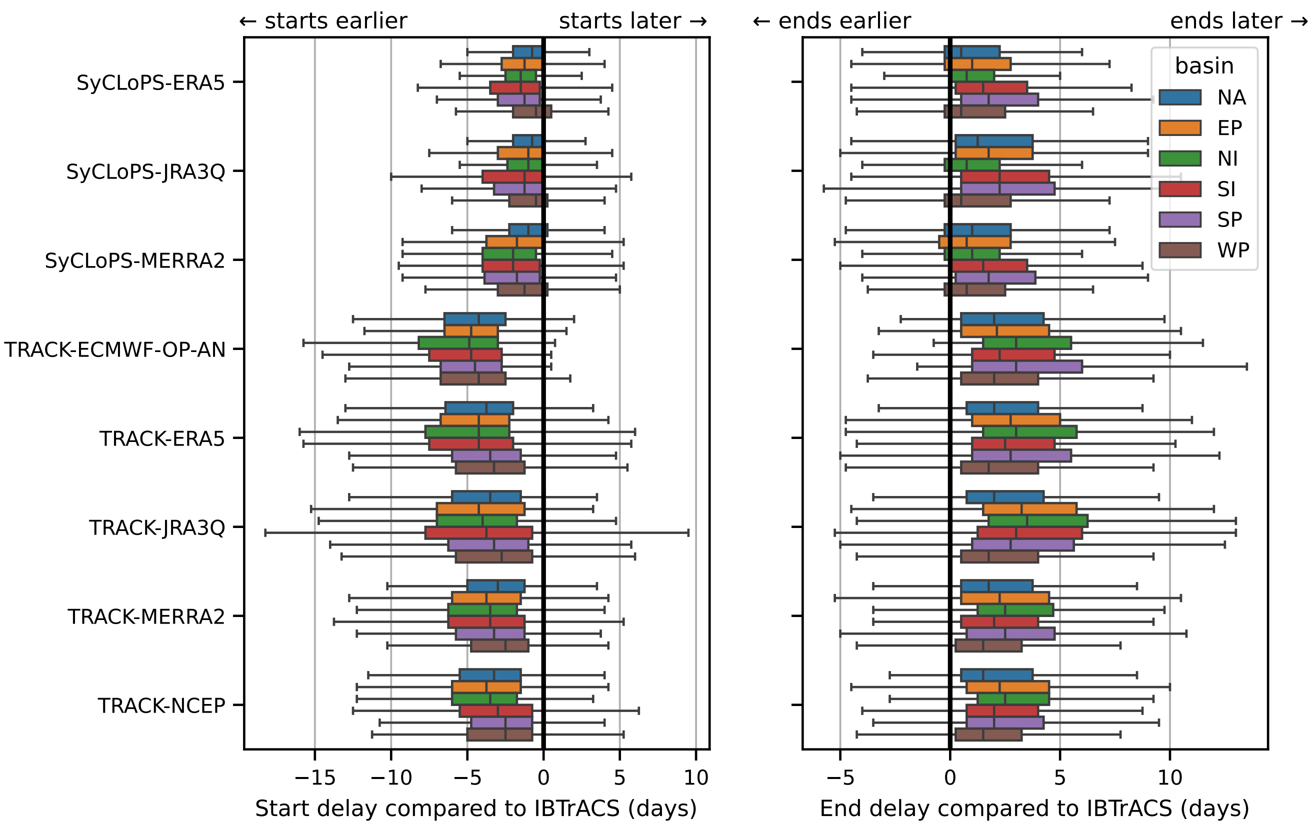

Figure 13. Boxplot of the distribution of the delay between the start and end of the matched (re)analyses tracks compared to IBTrACS, for each source and basin.

forms before the depression. TRACK also tends to follow the disturbance for longer after the IBTrACS record stops, although the difference with SyCLoPS is less than for genesis. This could be because remnants retain a vorticity signature for longer after the SLP minimum disappears. There are only slight variations across basins and across sources.

4.4 Uncertainty and level of confidence

This procedure allows us to define a notion of "level of confidence". Because we merge several sources of information about each track, we can evaluate whether these sources converge or diverge. The more data sources agree, the more confident we are about the information.

The first type of level of confidence is qualitative and case-by-case. For example, if the IBTrACS track for a given storm stops in the middle of the Atlantic, but several reanalysis products have it go all the way to Europe, we gain confidence in the fact that this cyclone indeed went to Europe. On the contrary, if only one reanalysis's track does that, we are more uncertain that this is the truth. In the following we assign equal confidence to all the data source for simplicity, but it must be noted that

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Figure 14. Number of Cyclones of Tropical Origin (CTO) that made landfall in mainland France over the period 1980-2022, depending on the imposed level of confidence.

different weights could be applied depending on our evaluation of the reliability of a given data source. This can be illustrated with the case shown previously: In the case of Gonzalo (fig. 4), we have high confidence that the storm went over the UK and into Europe since all reanalyses track it this way. On the other hand, we can only have moderate confidence that Lane (fig. 4) crossed the entire Atlantic before developping into a hurricane in the EP, since only two of 9 sources show this.

With the present database, we can also define a quantitative level of confidence, somewhat analogous to what was done in Flaounas et al. (2023). It must however be noted that because the different catalogues that were merged in the present database do not cover the same period, interpreting quantitative levels of confidence over time must be done with care, since some period have less data sources than others.

For example, let us find the number of Cyclones of Tropical Origin (CTO) that made landfall in mainland France over the period 1980-2022. For homogeneity, we use all the TRACK sources that cover this period, namely TRACK-ERA5, TRACK-JRA3Q, TRACK-NCEP and TRACK-MERRA2. Results are shown in figure 14. We see that there are 25 cyclones for which at least one source has a track making landfall in France, but this number decreases to 16 if we want at least two sources to agree on the French landfall, 10 if we impose 3 source, and only 7 where all four sources agree (Irene 1981, Floyd 1993, Iris 1995, Karl 1998, Gordon 2006, Henri 2015, and Pablo 2019). For all of these, the IBTrACS record stops before the French landfall, highlighting the benefit of using RABTracks rather than IBTrACS alone to estimate CTO risk in this region.

4.5 Caveats to keep in mind

There are three main caveats to the datasets that the user should keep in mind:

- Inhomogeneities in time: The different sources do not cover the same period. The ECMWF-OP-AN source has a resolution that changes with time. IBTrACS is known to embed strong inhomogeneities in time (Schreck et al., 2014), and is generally not considered robust before the satellite era (1̃980 onwards). Subsetting of sources and specific period might be necessary to have an homogeneous sample.

- Caveats of trackers: Because of the way trackers work, and of the matching procedure, it is possible that some of the apparently continuous tracks actually correspond to two (or more) different systems.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

– Caveats of analyses: We know that in the analyses, model resolution prevents the accurate representation of TC structure and intensity (e.g., Hodges et al., 2017; Dulac et al., 2023). Comparison of intensity values between analyses and IBTrACS, but also among analyses should be done with caution.

5 Conclusions

5.1 Potential for extension

RABTracks was not designed to be a static dataset. The ambition with RABTracks is to make it easy to gather as much information about historical storms as possible in one place. Therefore, it was designed to be easy to extend, using a NetCDF format that allows for increasing the dimensionality, and providing all the code.

Potential extensions include:

1. Including more sources: For example, 20th century reanalyses and regional reanalyses, to cover a longer period, or make use of higher-resolution data. Tracks from any tracking algorithm can also be used.

2. Include more historical storms, such as tropical disturbances that did not develop into TCs (invests) from SuperBT (Fiorino, 2026), for example. We could also consider including the tracks of cyclones not in IBTrACS, but robustly identified by several reanalyses/trackers

3. Computing additional attributes.

4. Attach cyclone-centred snapshots, or static maps to the track, that represent the hazards or the large-scale environment associated with the cyclones. For example, cyclone-centered wind and precipitation snapshots can be used to study the evolution of the hazards footprint of a cyclone throughout it lifecycle; or Sea-Surface Temperature and SLP maps can be attached to study the large scale environment with which the cyclone interacted.

5.2 Perspectives for RABTracks-based analyses

Finally, to give an idea of the investigations that can be performed with RABTracks. Below are descriptions of a few examples.

– Case studies: Thanks to the wealth of complementary information available for each cyclone, RABTracks can be used as a basis for case studies of particular cyclones. One could, in particular, take advantage of the confidence level aspect to discuss critically the evolution of a given case.

– Genesis: Because RABTracks provides in many cases an extension of the track before it was recorded in IBTrACS, it can be useful to trace back TCs to precursor/seeds.

– Extra-tropical Transitions (ET) and post-tropical stages: an update of Bieli et al. (2019)'s ET climatology is underway based on RABTracks, taking advantage of the multiple, recent sources and the two tracking methodologies. Climatology

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Open Access Earth System Science Data Discussions

of Cyclones of Tropical Origin (CTOs) reaching any specific area can be easily done using RABTracks. In the case of the Huracán project, we will use it to obtain the climatology of CTOs reaching Europe and the associated hazards.

- Comparison of reanalyses: RABTracks provides information that makes it very easy to compare the different reanalyses. For example, figure 5 illustrates how the different products represent intensity.
- Comparison of trackers: RABTracks also makes it very easy to compare SyCLoPS and TRACK as cyclone detection algorithms.
- Helping with AI models training: Skillfully handled, the multi-source extended TC tracks in RABTracks could further improve the ability of AI weather models to predict cyclone track and intensity with greater lead time, as well as their post-TC evolution.

6 Code and data availability

Three types of assets are distributed related to RABTracks:

1. The input track data;
2. The code to produce the dataset, and the figures within this paper;
3. The final RABTracks product.

To facilitate distribution, the variables were converted to the smallest suitable type, and all the data was compressed using the Lempel-Ziv-Markov chain algorithm (LZMA).

A static version of those assets, at the time of the submission of this article, is available on Zenodo at https://zenodo.org/records/20426511 (Bourdin, 2026).

If you are interested in obtaining the latest version of the dataset, which will be subject to evolutions, check out the RABTracks GitHub repository at https://github.com/Huracan-project/RABTracks.

Author contributions. SB led the project, designing and implementing the methodology for creating the dataset, leading the discussions, producing the figures and writing the manuscript. KH ran the TRACK detection method and computed the CPS parameters. YH ran the SyCLoPS detection method. LS computed the WCSI flags. KH, LS, AJB, JM and PLV participated in discussions throughout the whole process from conceptualisation to writing. All authors participated in the review of the manuscript. PLV, AJB, and JM obtained funding for the Huracán project.

Acknowledgements. This dataset was created within the scope of the UK NERC Large Grant project Huracán (NE/W009587/1), supporting SB, KH, LS, AJB, JM and PLV. Contributions from YH were performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

This project relied heavily on the storage and computing facilities available through JASMIN, the UK's collaborative data analysis environment (Lawrence et al., 2013, https://www.jasmin.ac.uk).

References

Aarons, Z. S., Camargo, S. J., Strong, J. D. O., and Murakami, H.: Tropical Cyclone Characteristics in the MERRA-2 Reanalysis and AMIP Simulations, Earth and Space Science, 8, e2020EA001 415, https://doi.org/10.1029/2020EA001415, number: 3 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020EA001415, 2021.

Baker, A. J., Hodges, K. I., Schiemann, R. K. H., and Vidale, P. L.: Historical Variability and Lifecycles of North Atlantic Midlatitude Cyclones Originating in the Tropics, Journal of Geophysical Research: Atmospheres, 126, e2020JD033 924, https://doi.org/10.1029/2020JD033924, number: 9 _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD033924, 2021.

Barcikowska, M., Feser, F., and Von Storch, H.: Usability of best track data in climate statistics in the western North Pacific, Monthly weather review, 140, 2818–2830, 2012.

Besson, M., Rivière, G., and Fromang, S.: A cyclone phase space dedicated to extratropical cyclones, EGUsphere, 2025, 1–30, 2025.

Bieli, M., Camargo, S. J., Sobel, A. H., Evans, J. L., and Hall, T.: A Global Climatology of Extratropical Transition. Part II: Statistical Performance of the Cyclone Phase Space, Journal of Climate, 32, 3583–3597, https://doi.org/10.1175/JCLI-D-18-0052.1, number: 12, 2019.

Bieli, M., Sobel, A. H., Camargo, S. J., Murakami, H., and Vecchi, G. A.: Application of the Cyclone Phase Space to Extratropical Transition in a Global Climate Model, Journal of Advances in Modeling Earth Systems, 12, e2019MS001 878, https://doi.org/10.1029/2019MS001878, number: 4 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001878, 2020.

Bourdin, S.: RABTracks V1, https://doi.org/10.5281/zenodo.20426511, data set. In: Earth System Scientific Data, 2026.

Bourdin, S. and Saffin, L.: HuracanPy: A Python package for reading and analysing cyclone tracks, Journal of Open Source Software, 10, 8174, 2025.

Bourdin, S., Fromang, S., Dulac, W., Cattiaux, J., and Chauvin, F.: Intercomparison of four algorithms for detecting tropical cyclones using ERA5, Geoscientific Model Development, 15, 6759–6786, https://doi.org/10.5194/gmd-15-6759-2022, number: 17 Publisher: Copernicus GmbH, 2022.

Chavas, D. R. and Knaff, J. A.: A simple model for predicting the tropical cyclone radius of maximum wind from outer size, Weather and Forecasting, 37, 563–579, 2022.

CRED and UNDRR: Human cost of disasters : An overview of the last 20 years (2000-2019), Tech. rep., http://www.undrr.org/publication/human-cost-disasters-overview-last-20-years-2000-2019, 2020.

Dulac, W., Cattiaux, J., Chauvin, F., Bourdin, S., and Fromang, S.: How Realistic are Tropical Cyclones in the ERA5 Reanalysis?, pp. EGU22–5755, https://doi.org/10.5194/egusphere-egu22-5755, conference Name: EGU General Assembly Conference Abstracts ADS Bibcode: 2022EGUGA..24.5755D, 2022.

Dulac, W., Cattiaux, J., Chauvin, F., Bourdin, S., and Fromang, S.: Assessing the representation of tropical cyclones in ERA5 with the CNRM tracker, Climate Dynamics, https://doi.org/10.1007/s00382-023-06902-8, 2023.

ECMWF: Changes in ECMWF Model, https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model, accessed: 2026-04-17, 2026.

European Centre for Medium-Range Weather Forecasts: ECMWF Operational Analysis: Assimilated Data, http://catalogue.ceda.ac.uk/uuid/c46248046f6ce34fc7660a36d9b10a71, accessed: 2026-04-14, 2006.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Evans, C., Wood, K. M., Aberson, S. D., Archambault, H. M., Milrad, S. M., Bosart, L. F., Corbosiero, K. L., Davis, C. A., Dias Pinto, J. R., Doyle, J., Fogarty, C., Galarneau, T. J., Grams, C. M., Griffin, K. S., Gyakum, J., Hart, R. E., Kitabatake, N., Lentink, H. S., McTaggart-Cowan, R., Perrie, W., Quinting, J. F. D., Reynolds, C. A., Riemer, M., Ritchie, E. A., Sun, Y., and Zhang, F.: The Extratropical Transition of Tropical Cyclones. Part I: Cyclone Evolution and Direct Impacts, Monthly Weather Review, 145, 4317–4344, https://doi.org/10.1175/MWR-D-17-0027.1, number: 11, 2017.

Fiorino, M.: A Global Super Best Track for Tropical Cyclone Forecasting and Research, https://doi.org/10.17605/OSF.IO/AWXUB, 2026.

Flaounas, E., Aragão, L., Bernini, L., Dafis, S., Doiteau, B., Flocas, H., L. Gray, S., Karwat, A., Kouroutzoglou, J., Lionello, P., Pantillon, F., Pasquero, C., Patlakas, P., Picornell, M. A., Porcù, F., D. K. Priestley, M., Reale, M., Roberts, M., Saaroni, H., Sandler, D., Scoccimarro, E., Sprenger, M., and Ziv, B.: A composite approach to produce reference datasets for extratropical cyclone tracks: Application to Mediterranean cyclones, Weather and Climate Dynamics Discussions, pp. 1–32, https://doi.org/10.5194/wcd-2022-63, publisher: Copernicus GmbH, 2023.

Gahtan, J., Knapp, K., Schreck, C., Diamond, H. J., Kossin, J. P., and Kruk, M. C.: International best track archive for climate stewardship (IBTrACS) project, Version 4r01, NOAA National Centers for Environmental Information, 10, 2024.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Silva, A. M. d., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, number: 14 Publisher: American Meteorological Society Section: Journal of Climate, 2017.

Han, Y. and Ullrich, P. A.: The system for classification of low-pressure systems (SyCLoPS): An all-in-one objective framework for large-scale data sets, Journal of Geophysical Research: Atmospheres, 130, e2024JD041 287, 2025.

Hart, R. E.: A Cyclone Phase Space Derived from Thermal Wind and Thermal Asymmetry, Monthly Weather Review, 131, 585–616, https://doi.org/10.1175/1520-0493(2003)131<0585:ACPSDF>2.0.CO;2, number: 4 Publisher: American Meteorological Society Section: Monthly Weather Review, 2003.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, number: 730 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3803, 2020.

Hodges, K., Cobb, A., and Vidale, P. L.: How Well Are Tropical Cyclones Represented in Reanalysis Datasets?, Journal of Climate, 30, 5243–5264, https://doi.org/10.1175/JCLI-D-16-0557.1, number: 14 Publisher: American Meteorological Society, 2017.

Jones, E., Wing, A. A., and Parfitt, R.: A Global Perspective of Tropical Cyclone Precipitation in Reanalyses, Journal of Climate, 34, 8461–8480, https://doi.org/10.1175/JCLI-D-20-0892.1, number: 21 Publisher: American Meteorological Society Section: Journal of Climate, 2021.

Jones, S. C., Harr, P. A., Abraham, J., Bosart, L. F., Bowyer, P. J., Evans, J. L., Hanley, D. E., Hanstrum, B. N., Hart, R. E., Lalaurette, F., Sinclair, M. R., Smith, R. K., and Thorncroft, C.: The Extratropical Transition of Tropical Cyclones: Forecast Challenges, Current Understanding, and Future Directions, Weather and Forecasting, 18, 1052–1092, https://doi.org/10.1175/1520-0434(2003)018<1052:TETOTC>2.0.CO;2, number: 6, 2003.

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., and Neumann, C. J.: The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying Tropical Cyclone Data, Bulletin of the American Meteorological Society, 91, 363–376, https://doi.org/10.1175/2009BAMS2755.1, number: 3 Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2010.

Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., et al.: The JRA-3Q reanalysis, Journal of the Meteorological Society of Japan. Ser. II, 102, 49–109, 2024.

Kumar, S., Chand, S., Ramsay, H., Klotzbach, P. J., Courtney, J., Koschatzky, V., and Kumar, S.: An inter-comparison of tropical cyclone datasets for the Australian region, Environmental Research Letters, 20, 014 039, 2025.

Lang, S., Rodwell, M., and Schepers, D.: IFS upgrade brings many improvements and unifies medium-range resolutions, ECMWF Newsletter, 176, 21–28, 2023.

Lawrence, B. N., Bennett, V. L., Churchill, J., Juckes, M., Kershaw, P., Pascoe, S., Pepler, S., Pritchard, M., and Stephens, A.: Storing and manipulating environmental big data with JASMIN, in: 2013 IEEE international conference on big data, pp. 68–75, IEEE, 2013.

McCarty, W., Coy, L., Gelaro, R., Huang, A., Merkova, D., Smith, E. B., Sienkiewicz, M., and Wargan, K.: MERRA-2 Input Observations: Summary and Assessment, GMAO Office Note No. 9 (Version 1.1), NASA Global Modeling and Assimilation Office (GMAO), https://gmao.gsfc.nasa.gov/pubs/docs/McCarty885.pdf, accessed: 2026-04-17, 2016.

Miglietta, M. M.: Toward the definition of "Medicane", https://scholar.google.com/citations?view_op=view_citation&hl=fr&user=WcI7CCcAAAAJ&sortby=pubdate&citation_for_view=WcI7CCcAAAAJ:-LHtoeeytlUC, 2023.

National Hurricane Center: Tropical Cyclone Report: Hurricane Beryl, Tropical Cyclone Report AL022024, National Hurricane Center, National Oceanic and Atmospheric Administration, https://www.nhc.noaa.gov/data/tcr/AL022024_Beryl.pdf, 2024.

Ren, F., Liang, J., Wu, G., Dong, W., and Yang, X.: Reliability analysis of climate change of tropical cyclone activity over the western North Pacific, Journal of climate, 24, 5887–5898, 2011.

Saffin, L., Hodges, K., Bourdin, S., Fiorino, M., Methven, J., Vidale, P. L., and Baker, A. J.: An improved approach to identifying tropical cyclones in model data using the cyclone phase space: warm core, symmetric, and intensifying, https://essopenarchive.org/doi/full/10.22541/essoar.15003790/v1, 2026.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P. V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Dool, H. v. d., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP Climate Forecast System Reanalysis, Bulletin of the American Meteorological Society, 91, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1, number: 8 Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2010.

Sainsbury, E. M., Schiemann, R. K. H., Hodges, K. I., Shaffrey, L. C., Baker, A. J., and Bhatia, K. T.: How Important Are Post-Tropical Cyclones for European Windstorm Risk?, Geophysical Research Letters, 47, e2020GL089 853, https://doi.org/10.1029/2020GL089853, number: 18 Publisher: John Wiley & Sons, Ltd, 2020.

Schenkel, B. A., Lin, N., Chavas, D., Oppenheimer, M., and Brammer, A.: Evaluating Outer Tropical Cyclone Size in Reanalysis Datasets Using QuikSCAT Data, Journal of Climate, 30, 8745–8762, https://doi.org/10.1175/JCLI-D-17-0122.1, number: 21 Publisher: American Meteorological Society Section: Journal of Climate, 2017.

https://doi.org/10.5194/essd-2026-403
Preprint. Discussion started: 30 July 2026

Schreck, C. J., Knapp, K. R., and Kossin, J. P.: The Impact of Best Track Discrepancies on Global Tropical Cyclone Climatologies using IBTrACS, Monthly Weather Review, 142, 3881–3899, https://doi.org/10.1175/MWR-D-14-00021.1, number: 10 Publisher: American Meteorological Society Section: Monthly Weather Review, 2014.

Serra, Y. L., Kiladis, G. N., and Hodges, K. I.: Tracking and mean structure of easterly waves over the Intra-Americas Sea, Journal of Climate, 23, 4823–4840, 2010.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., et al.: The ERA5 global reanalysis from 1940 to 2022, Quarterly Journal of the Royal Meteorological Society, 150, 4014–4048, 2024.

Swinbank, R., Kyouda, M., Buchanan, P., Froude, L., Hamill, T. M., Hewson, T. D., Keller, J. H., Matsueda, M., Methven, J., Pappenberger, F., et al.: The TIGGE project and its achievements, Bulletin of the American Meteorological Society, 97, 49–67, 2016.

Torgerson, W., Schwendike, J., Ross, A., and Short, C. J.: Intensity fluctuations in Hurricane Irma (2017) during a period of rapid intensification, Weather and Climate Dynamics, 4, 331–359, 2023.

Ullrich, P. A. and Zarzycki, C. M.: TempestExtremes: A framework for scale-insensitive pointwise feature tracking on unstructured grids, Geoscientific Model Development, 10, 1069–1090, https://doi.org/10.5194/gmd-10-1069-2017, number: 3 Publisher: Copernicus Gesellschaft mbH, 2017.

Ullrich, P. A., Zarzycki, C. M., McClenny, E. E., Pinheiro, M. C., Stansfield, A. M., and Reed, K. A.: TempestExtremes v2.1: A Community Framework for Feature Detection, Tracking and Analysis in Large Datasets, Geoscientific Model Development Discussions, pp. 1–37, https://doi.org/https://doi.org/10.5194/gmd-2020-303, publisher: Copernicus GmbH, 2021.

Zarzycki, C. M. and Ullrich, P. A.: Assessing sensitivities in algorithmic detection of tropical cyclones in climate data, Geophysical Research Letters, 44, 1141–1149, https://doi.org/10.1002/2016GL071606, number: 2 _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016GL071606, 2017.

Zarzycki, C. M., Ullrich, P. A., and Reed, K. A.: Metrics for evaluating tropical cyclones in climate data, Journal of Applied Meteorology and Climatology, https://doi.org/10.1175/JAMC-D-20-0149.1, 2021.