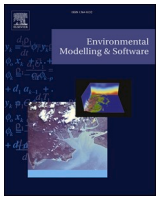
Environmental Modelling and Software 167 (2023) 105773

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



Chasing rainfall: estimating event precipitation along tracks of tropical cyclones via reanalysis data and in-situ gauges

Jasmine B.D. Jaffrés [a,b,*], Jessie L. Gray [c]

[a] C&R Consulting, 188 Ross River Rd, Townsville, Qld 4814, Australia
[b] College of Science and Engineering, James Cook University, Townsville, Australia
[c] School of Health, Medical and Applied Sciences, Central Queensland University, Townsville, Australia

ARTICLE INFO

Handling Editor: Daniel P Ames

Keywords:
Tropical cyclones
Rainfall accumulation
Event tracking
Reanalysis
Station data
Toolbox

ABSTRACT

Tropical cyclones (TCs) are an important water source in many regions around the world, replenishing local dams, waterways and groundwater systems. Three diverse precipitation datasets were tested for dissimilarities in their rainfall characteristics via a new, freely available rainfall tracking toolbox for MATLAB and GNU Octave users: 1) the ERA5 global reanalysis, 2) the Global Historical Climatology Network (GHCN)-Daily station dataset and 3) the regional SILO (Scientific Information for Land Owners) database. Although SILO only covers Australia, its relatively high resolution (0.05°) provides advantages for studies in that region. To test the differences in precipitation datasets, six episodes (eight individual TC events) in all major basins affected by TCs have been selected. These include two instances in which consecutive TCs severely impacted the same region (TCs Idai and Kenneth in south-eastern Africa during March/April 2019 – and hurricanes Eta and Iota in Central America in November 2020). Precipitation for TC episodes was explored through event totals and the proportional contribution to water years within each dataset. Each precipitation dataset demonstrated its inherent advantages and drawbacks, highlighting the benefits of using more than one source to thoroughly evaluate an individual event. These attributes – coupled with the associated impacts of cyclonic events – reinforce the importance of developing tools that can aid in managing TC-related rainfall and flooding potential.

Software availability

Name of software: *rainfall_tracker*.
Developers: Jasmine B. D. Jaffrés, Jessie L. Gray.
Maintainer: Jasmine B. D. Jaffrés.
Contact email: Jasmine@candrconsulting.com.au.
Year of first release: 2023.
Software required: MATLAB or GNU Octave.
Program language: MATLAB (compatible with GNU Octave).
Availability: Freely accessible via GitHub (https://github.com/jjaffres/rainfall_tracker).

1. Introduction

1.1. TC-related rainfall studies

The contribution of tropical cyclones (TCs) to cumulative precipitation is spatially and temporally highly variable. Several studies have explored the rainfall impact of TCs on a global scale. Using remote-sensing data, Jiang and Zipser (2010) determined that TC-related rainfall ranges from approximately 3% to 11% in the various ocean basins (highest in the north-west Pacific) during the TC season, with the contribution locally increasing to about 55% (south-west of the Baja California coast and north-west of the Australia coast), with the north-west Pacific near Taiwan also exceeding 35%. Similar results were also obtained by Khouakhi et al. (2017) who investigated both annual and seasonal TC contribution to rainfall via daily station data derived from the Global Historical Climatology Network (GHCN-Daily). Further, using the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) data, Prat and Nelson (2013) have highlighted the sharp decrease in TC rain contribution further inland – and the high variability in TC rain in East Africa, especially compared to East Asia. If the observed upward trend in TC rain rate between 1998 and 2016 (Guzman and Jiang, 2021) persists into the future, then the TC contribution to rainfall totals may further rise. Thus, our ability to accurately represent and predict TC-related rainfall will need to improve

* Corresponding author. C&R Consulting, 188 Ross River Rd, Townsville, Qld 4814, Australia.
E-mail address: Jasmine@candrconsulting.com.au (J.B.D. Jaffrés).

https://doi.org/10.1016/j.envsoft.2023.105773
Received 5 January 2023; Received in revised form 10 June 2023; Accepted 3 July 2023
Available online 4 July 2023

to combat the impacts.

In Australia, TCs are estimated to provide between 10% and 40% on average (mean) to seasonal (November to April) rainfall totals based on high-resolution, gridded data (Australian Water Availability Project [AWAP]), with the highest mean input observed in the western half of Australia (west of 120°; Dare et al., 2012). Ng et al. (2015) obtained similar percentages – for annual TC contribution to rainfall – in north-western Australia. However, for seasonal rainfall (October to December – and January to March), Ng et al. (2015) determined values over 70% for individual, near-coastal weather stations. Over the entire country, Dare (2013) has calculated a mean TC contribution of 7.6% to the total, seasonal (November to April) rainfall volume, with most of this precipitation attributable to land-falling events. Rainfall is notoriously unreliable in many regions (e.g. Australia), exhibiting significant variability on interannual timescales (Nicholls et al., 1997). This unpredictability is also evident in TC-sourced rainfall totals (e.g. Dare et al., 2012).

Deo et al. (2021) have demonstrated that the contribution of TCs to seasonal peak daily precipitation in the Pacific regions can exceed 60%, with the rainfall impact substantially modulated by El Niño – southern oscillation and Madden–Julian oscillation patterns. A similar percentage (up to 54.2%) was also found for annual rainfall in the north-western Philippines, with the proportional rainfall contribution decreasing to 6% towards the equator (Bagtasa, 2017). Bagtasa (2017) attributed the high rainfall fraction to the enhancement of the Asian south-west monsoon when TCs were situated to the north-east of the country. Because Bagtasa (2017) applied a relatively large radius (10°) from the TC centres to ascribe rainfall to storm events, precipitation only indirectly related to the TCs was also captured. In the north-eastern Pacific, Corbosiero et al. (2009) found that individual storms can account for as much as 95% of precipitation in the warm season. In the south-eastern United States of America (USA), a TC rainfall contribution of up to 15% was determined during the storm season, with the highest proportion associated with the coastal Carolinas (Knight and Davis, 2007). In Mexico, the largest proportion of TC-related precipitation (up to 80%) is associated with the arid western region (Baja California Peninsula), whereas the largest amount (in mm) was observed in the southern regions (Breña-Naranjo et al., 2015). On the African continent, Howard et al. (2019) have attributed up to 70% of austral summer (November to March) rainfall to tropical lows, with the highest totals in south-eastern Angola. Over all of India, Hunt and Fletcher (2019) have estimated that tropical low-pressure systems contribute 37% of June–September rainfall when a 800 km radius is applied, although clustering techniques suggest a slightly larger proportion.

1.2. Flood impact of rainfall from TCs

Flooding is one of the more destructive impacts from TCs. Flood severity is intrinsically linked to rainfall intensity, duration and event total. Additional, localised factors include differences of land use, basin characteristics (shape, size), geographic location and antecedent soil moisture (e.g. Villarini et al., 2014). Extensive economic, environmental and social damage is among the on-flow effects of flooding. Severe flooding can cause major disruptions to transportation networks, isolation and evacuation of townships, inundation of property, and death (Rappaport, 2000). Globally, TC-related floods and related mortalities are most commonly associated with eastern seaboards (e.g. eastern India and the east coast of the USA; Hu et al., 2018).

Although common, flooding from TC events is not restricted to coastal regions, with many TCs responsible for major inundation to inland areas (Villarini et al., 2014). Inland flooding was the predominate cause of death from TCs in the USA from 1979 to 1999 (Rappaport, 2000). However, storm surges were associated with more fatalities (but fewer TCs) when a 50-year period (1963–2012, i.e. including TC Katrina in 2005) was considered (Rappaport, 2014). Inland flooding on relatively flat terrain tends to persist longer, with water levels only slowly subsiding compared to near-coastal areas, thus causing additional damage in the process. Avoidance and mitigation of flooding from TCs is a key challenge for the preservation of economic revenue and a key concern for susceptible global communities. Thus, tools that contribute to the estimation of the extent of TC-based rainfall are important to manage future flooding potential.

1.3. Lack of in-situ datasets and the reliance on reanalysis/satellite data for rainfall

Globally, the availability of in-situ rainfall data is highly deficient, especially in sparsely populated regions such as mountainous terrains, with weather stations preferentially deployed in more urbanised areas. Consequently, scientific studies frequently rely on satellite and reanalysis data as an alternative. Reanalysis and satellite sources provide the benefit of spatially more homogenous and detailed datasets. However, this advantage is offset by their relatively short temporal coverage (Hassler and Lauer, 2021; Sun et al., 2018; Thorne and Vose, 2010) as a consequence of their reliance on high-quality satellite and in-situ data. Specifically, although some weather station time series are very extensive (i.e. covering more than a century), most sites have only been operational for a limited period (Jaffrés, 2019). Another issue with global reanalysis and satellite datasets is their limited ability to resolve fine-scale rainfall features at regional spatial resolutions. Although newer versions of reanalysis data have decreased grid spacing, rainfall studies at a catchment-scale level are still a challenge with respect to accurately representing real-world conditions, especially in regions with higher spatial rainfall variability (e.g. orographic precipitation; Sun et al., 2018). Ground-truthing via gauging data is not always feasible because of limited station network density in some areas, especially in sparsely population regions (Hassler and Lauer, 2021).

1.4. Aims

Numerous precipitation datasets are available at various temporal and spatial scales (in terms of both temporal/spatial resolution and coverage) to investigate rainfall properties of TCs. For the doctoral thesis by Gray (2023), the target period was 1950–2019. Consequently, Gray (2023) selected rainfall databases that encompass that multi-decadal time span to evaluate precipitation patterns. Among these were 1) the SILO (Scientific Information for Land Owners) database, a regional, high-resolution (0.05° grid) dataset for Australia (Jeffrey et al., 2001) and 2) ERA5 (5th-generation ECMWF [European Centre for Medium-Range Weather Forecasts] reanalysis), a global reanalysis dataset (Hersbach et al., 2020). To facilitate these analyses, a rainfall tracking toolbox (*rainfall_tracker*) was developed to enable the extraction of precipitation within a chosen distance from each track position. Specifically, a MATLAB toolbox with GNU Octave compatibility was created to obtain spatial rainfall patterns (event rainfall, peak intensities, totals per time step) associated with individual cyclone tracks.

The present study focuses on the comparison of three diverse precipitation datasets (regional vs global, gridded vs station data, daily vs sub-daily) with long-term coverage. These datasets include 1) the ERA5 global reanalysis, 2) the GHCN-Daily station dataset (Menne et al., 2012a) and 3) the regional SILO database. To determine biases, advantages and disadvantage of each precipitation data source, six episodes (eight individual TC events) were analysed. These case studies were selected by two main attributes: 1) their documented rainfall impact and 2) the affected region, ensuring that all major ocean basins with TC genesis were represented. Differences in estimated rainfall accumulations are discussed in terms of precipitation datasets, track properties and applied radius.

2. Data

2.1. TC track data

TC track data were sourced from the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al., 2010, Table 1). IBTrACS, version 4, is a global TC database that spans from 1841 to present (Knapp et al., 2018), at three-hourly time steps and with a spatial resolution of 0.1°. The database consists of seven basins: 1) northern Indian Ocean, 2) southern Indian Ocean, 3) north-western Pacific, 4) north-eastern Pacific, 5) South Pacific, and the 6) North and 7) South Atlantic (Knapp et al., 2010). A second, regional TC database – by the Australian Bureau of Meteorology (BoM; Table 1) – was also accessed for track confirmation purposes.

2.2. Rainfall datasets

A variety of rainfall datasets – from regional to global, and rectilinearly gridded vs point data (Table 1) – was accessed to highlight 1) the differences in observed or estimated precipitation in individual data sources and 2) the capacity of the rainfall tracking toolbox (*rainfall_tracker*). Several gridded datasets (in netCDF format) have been incorporated and tested, with the code also able to access rainfall data extracted from GHCN-Daily.

2.2.1. ERA5 – global

ERA5 is a global reanalysis dataset that has replaced ERA-Interim (Hersbach et al., 2020). At the time of the study, ERA5 was split into two time-periods: 1) 1950–1978 (preliminary) and 2) 1979–current. Because of difficulties in accurately representing TCs, the preliminary version of ERA5 is not recommended for usage in TC studies (Bell et al., 2021). Consequently, only ERA5 data from 1979 were accessed. The horizontal data resolution of ERA5 is 31 km, reprojected on a 0.25° grid (Hersbach et al., 2020). The sub-daily (hourly) total precipitation variable was extracted.

2.2.2. GHCN-Daily – global

GHCN-Daily is a freely available dataset encompassing over 118,000 stations from 218 countries and territories – for core variables such as precipitation, snowfall, snow depth, minimum and maximum temperature (Jaffrés, 2019; Menne et al., 2012b). The database (in the .dly file format) can be accessed via https://www.ncei.noaa.gov/pub/data/ghcn/daily/. For this study, the entire GHCN-Daily archive (3.28-upd-2021091218) was accessed, with the daily precipitation data extracted via the freely available *ghcnd_access* toolbox (Jaffrés, 2019) before applying the *rainfall_tracker* code.

2.2.3. SILO – regional

Daily gridded rainfall data for Australia were sourced from the SILO database (Jeffrey et al., 2001). Rainfall data within the SILO database cover the period from 1889 to current. The daily gridded data are interpolated on a 0.05° grid using point data sourced from the BoM (Jeffrey et al., 2001). The spatial extent of the grid extends from 10°S to 44°S and 112°E to 154°E.

3. Methods

Although numerous studies have previously focussed on TC-related precipitation, suitable toolboxes for MATLAB or GNU Octave users to conduct such analyses are not readily available or well-known. Further, relatable TC analysis tools built on other coding platforms such as Python (Brackins and Kalyanapu, 2020) have been useful for TC analysis but have generally focused on forecasting and modelling. Brackins and Kalyanapu (2020) have specifically highlighted that they were unable to locate existing code for any of the six existing parametric models (for TC rainfall) that they identified. Among these models is R-CLIPER (rainfall climatology and persistence; Tuleya et al., 2007), predominantly used as a statistical baseline for TC rainfall predictions (e.g. Ko et al., 2020).

An exception – in terms of accessibility and application – is the TempestExtremes software package, which permits the extraction of TC tracks and associated precipitation from gridded model, reanalyses and observational data products (Ullrich et al., 2021). The TempestExtremes software package intricately contructs TC tracks by evaluating and tracing sea level pressure minima within regional or global datasets (Stansfield et al., 2020; Ullrich et al., 2021). Using wind speed and the trajectory, Stansfield et al. (2020) estimated TC size by determining where the azimuthally averaged wind speed exceeds 8 m/s, then applying this area to extract event-based precipitation in the USA.

The toolbox presented here is designed for MATLAB and GNU Octave. External databases (separate TC track and rainfall datasets) are applied as input variables for the *rainfall_tracker* toolbox. The toolbox has been tested on rainfall data in netCDF format (global ERA5 and regional SILO) and the GHCN-Daily database (preprocessed via the freely available *ghcnd_access* toolbox; Jaffrés, 2019). Consequently, a diverse range of station and reanalysis rainfall datasets – rather than output from climate models – are available to extract TC-related precipitation via the *rainfall_tracker* toolbox.

3.1. Data organisation and requirements of the rainfall_tracker toolbox

The *rainfall_tracker* toolbox requires two input files: 1) a .csv containing the TC track information and 2) the rainfall dataset that can either be station data (GHCN-Daily) or a gridded dataset with regional (SILO) or global (ERA5) coverage. For the TC input, a unique TC name is required, along with the latitude, longitude, date and time. If GHCN-Daily data are chosen, the downloaded rainfall files (.dly format) will first need to be converted into a .mat file using the freely available *ghcnd_access* toolbox (Jaffrés, 2019) before applying the *rainfall_tracker* code. Detailed directions are provided in the user's guide for the *rainfall_tracker* toolbox and a sample TC input file is also supplied.

Along with selecting TC tracks and a precipitation dataset, the applicable time zones also have to be stipulated before the toolbox can be run correctly. TC track data are usually associated with Coordinated Universal Time (UTC+0). The time zone also has to be specified for the rainfall data, although UTC+0 generally applies for global, gridded datasets. Conversely, the SILO rainfall is relative to the local time zone of each region (UTC+8 to UTC+11), shifting biannually by one hour in regions where daylight saving is implemented. Similarly, time zones of

Table 1
TC track and rainfall datasets accessed for this study.

Data type	Name	Period	Website
TC tracks	IBTrACS (International Best Track Archive for Climate Stewardship)	1841– present	https://www.ncei.noaa.gov/products/international-best-track-archive?name=ibtracs-data-access
	BoM (Australian Bureau of Meteorology) tropical cyclone database	1907– present	http://www.bom.gov.au/cyclone/tropical-cyclone-knowledge-centre/databases/
Precipitation	ERA5 (5th-generation ECMWF reanalysis)	1979– present	https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form
	GHCN-Daily (Global Historical Climatology Network for daily station data)	1781–present	https://www.ncei.noaa.gov/pub/data/ghcn/daily/
	SILO (Scientific Information for Land Owners)	1889– present	https://www.longpaddock.qld.gov.au/silo/gridded-data/

station data in GHCN-Daily are source-dependent. Further, the BoM reports 24-hour rainfall with respect to local time at 9 a.m. Consequently, if SILO or Australian GHCN-Daily data are accessed, the *rainfall_tracker* toolbox automatically employs a time shift to the TC tracks (+14.99 hours) to align them with the rain period (Fig. 1). An equivalent shift was also applied to other GHCN-Daily subsets when a cut-off period was identified.

Further, TC tracks may transect and affect multiple time zones (Table 2). Thus, for rainfall datasets with non-uniform dates (i.e. GHCN-Daily and SILO), the most relevant time zone has to be determined – in terms of either 1) track proximity to land (if a land-based rainfall dataset or study), 2) time spent in a region, or 3) expected maximum impact in the area of interest. In some instances, applying different time zones to a track may be appropriate (e.g. for trans-Atlantic TCs like hurricane Charley in 1986, moving from the Americas to Europe).

3.2. Rainfall extraction

For this study, six case studies (eight individual TCs during six episodes of storm activity; Table 2; Fig. 2) were selected to illustrate the

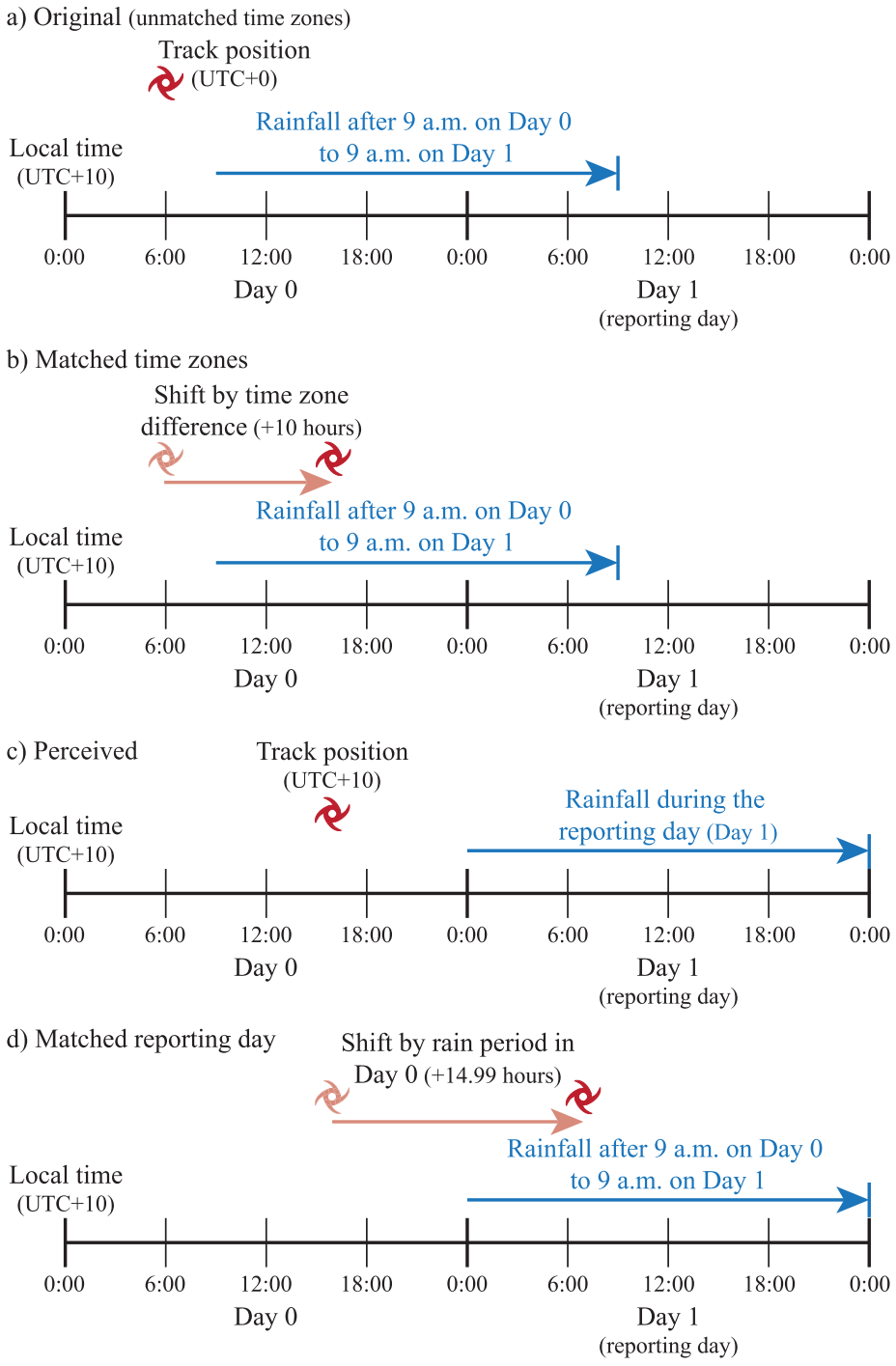


Fig. 1. A schematic illustrating the time zone compatibility of rainfall and track data: a) unmatched time zones, b) track time shifted to match rainfall, c) example scenario if rainfall does not cover midnight to midnight and d) consequent shift of track to ensure extraction of correct rain day.

Table 2
List of TCs and the range of time zones (UTC) covered by the cyclone tracks. Relevant time zones (for SILO and GHCN-Daily in this study) are highlighted in bold. The lowest pressure, peak wind speed and applied water year for each event are also provided.

Name	Year	Lowest pressure (mb) / Peak wind speed (kts[a])	Location	Time zones	Water year
TC Idai	2019	940 / 105^{10min}	Africa (Mozambique, Madagascar)	**2**, 3	August–July
TC Kenneth	2019	934 / 115^{10min}	Africa (Mozambique, Madagascar)	**2**, 3	August–July
Cyclone Nisha	2008	996 / 45^{3min}	India	**5.5**	February–January
TC Seroja	2021	971 / 65^{10min}	South-east Asia (Indonesia, Timor-Leste), Australia	7, **8**, 9	August–July[b]
TC Oswald	2013	991 / 35^{10min}	Australia	**10**	September–August
Typhoon Morakot	2009	945 / 75^{10min}	Asia (Taiwan, mainland China)	**8**, 9	January–December
Hurricane Eta	2020	922 / 130^{1min}	Central America (Nicaragua, Honduras)	-5, **-6**	March–February
Hurricane Iota	2020	917 / 135^{1min}	Central America (Nicaragua, Honduras)	-5, **-6**	March–February

[a] The averaging period for maximum sustained wind speed differs depending on the monitoring agency. Whereas most regions apply a 10-minute averaging period, 3-minute and 1-minute intervals are used in the northern Indian Ocean (cyclone Nisha) and North Atlantic (hurricanes Eta and Iota), respectively.
[b] The water year of the most impacted region (Indonesia, Timor-Leste) was employed. For south-west Australia, a water year of February–January would be more appropriate.

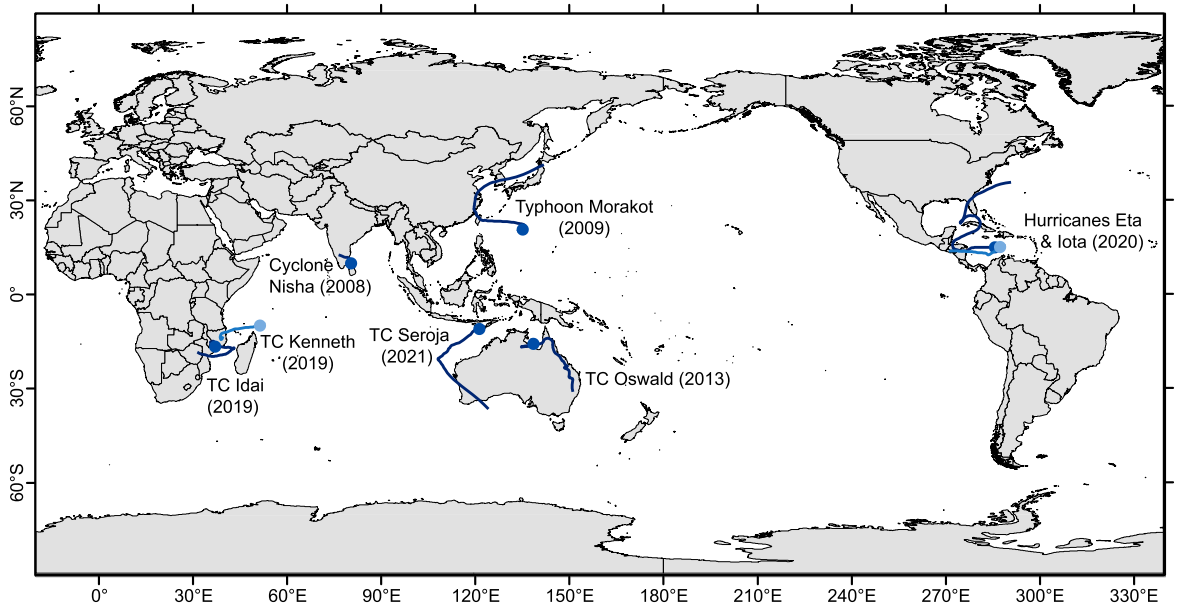


Fig. 2. World map showing the tracks of the eight TCs selected for the six case studies.

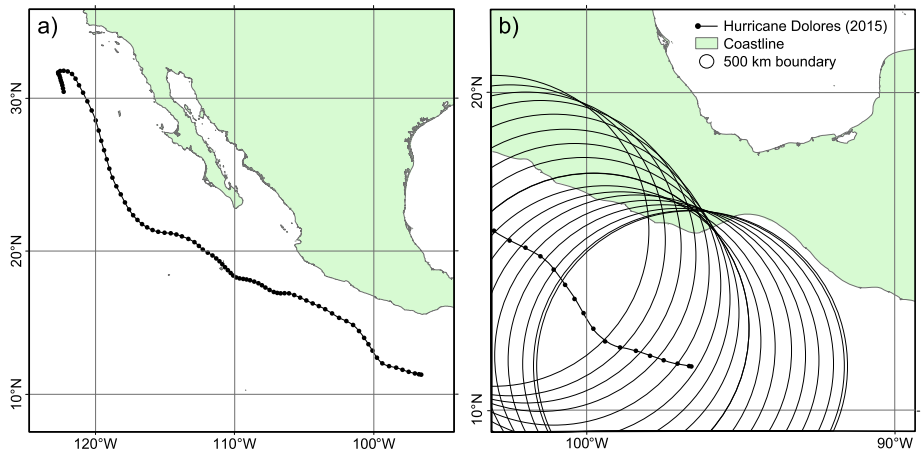


Fig. 3. a) A sample track and b) illustration of radial tracing via circles of fixed radius along the TC track. Rainfall data encompassed by the area within the circles are extracted by the *rainfall_tracker* toolbox.

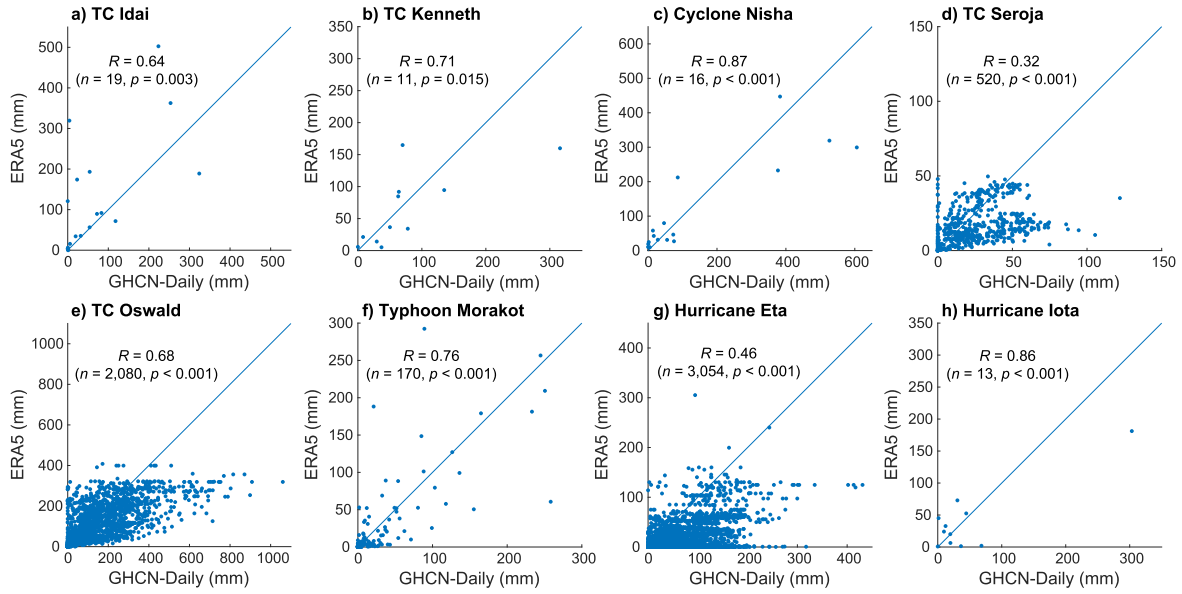


Fig. 4. Correlation of event rainfall estimates (mm) derived from GHCN-Daily and ERA5 for a) TC Idai, b) TC Kenneth, c) cyclone Nisha, d) TC Seroja, e) TC Oswald, f) typhoon Morakot, g) hurricane Eta and h) hurricane Iota.

toolbox capabilities and differences in the tested rainfall datasets. For all events, a 500 km radius was initially defined and aligned with each track position of all individual TCs (Fig. 3), along with a temporal window of one hour. The circles were defined based on the haversine formula:

$$A = \sin^2\left(\frac{lat_{rad2} - lat_{rad1}}{2}\right) + \cos(lat_{rad1}) \times \cos(lat_{rad2}) \times \sin^2\left(\frac{lon_{rad2} - lon_{rad1}}{2}\right), \text{ and}$$

$$\text{distance (km)} = 2 \times r \times \arcsin\left(\sqrt{A}\right),$$

where lat_{rad1} and lat_{rad2} (lon_{rad1} and lon_{rad2}) are the radial latitude (longitude) of two points (track point vs rainfall location), and $r = 6371$ km is the radius of Earth. The *rainfall_tracker* toolbox then extracted the relevant rainfall data for each track position. If multiple track points (time steps) were present during one rainfall interval, the circle areas for these track points were merged to create one output field per rainfall period. This process was repeated for all tracks, providing total rainfall

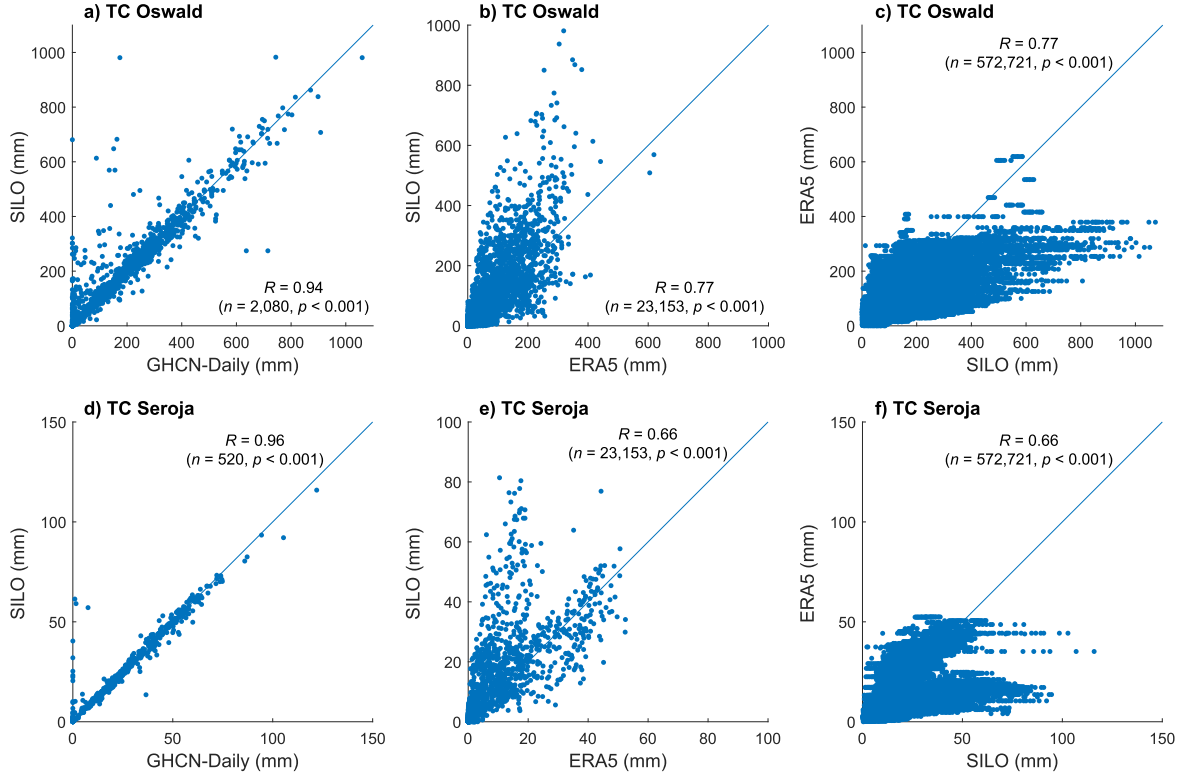


Fig. 5. Correlation of event rainfall estimates (mm) derived from SILO vs a,d) GHCN-Daily and b-c,e-f) ERA5 for a-c) TC Oswald and d-f) TC Seroja.

Table 3
Significance level of paired Wilcoxon signed-rank tests comparing event rainfall estimates by different precipitation datasets. The listed grid point numbers for ERA5 and SILO (in brackets) refer to cells with rainfall data (i.e. non-empty fields) within the SILO region used for the paired Wilcoxon signed-rank test.

TC event	Station number	ERA5 (SILO) grid points	*p*-value			
			GHCN-Daily vs ERA5	GHCN-Daily vs SILO	ERA5 vs SILO	SILO vs ERA5
TC Idai	19	–	0.009	–	–	–
TC Kenneth	11	–	0.465	–	–	–
Cyclone Nisha	16	–	0.918	–	–	–
TC Seroja	520	1354 (33,656)	<0.001	0.629	<0.001	<0.001
TC Oswald	2080	3382 (84,273)	<0.001	<0.001	<0.001	<0.001
Typhoon Morakot	170	–	0.353	–	–	–
Hurricane Eta	3054	–	<0.001	–	–	–
Hurricane Iota	13	–	0.966	–	–	–

Table 4
Maximum contribution (%) to the annual total rainfall (of the relevant water year) of TC events based on ERA5, GHCN-Daily and SILO and the relevant IBTrACS pathways. The tandem events (TCs Idai and Kenneth – and hurricanes Eta and Iota) were assessed jointly. For TC Seroja, a second ERA5 estimate is provided for the BoM track.

TC events	ERA5 (BoM) %	GHCN-Daily %	SILO %
TCs Idai and Kenneth	50.0	–	–
Cyclone Nisha	30.6	–	–
TC Seroja	45.4 (46.0)	16.7	21.3
TC Oswald	49.4	63.0	57.1
Typhoon Morakot	42.7	17.8	–
Hurricanes Eta and Iota	34.7	16.4	–

depth per event and grid point (or weather station).

For two events (cyclone Nisha and TC Oswald), four different radii (500 km, 600 km, 750 km and 1000 km) were applied to investigate the extent of their rainfall impact and, thus, to determine the most suitable radius for these specific storms.

3.3. Statistical analysis

A non-parametric, paired Wilcoxon signed-rank test was applied to evaluate whether event rainfall estimates were statistically different between individual precipitation datasets. For each database combination, the collocated rainfall estimates were extracted (e.g. the nearest ERA5 grid point for each relevant GHCN-Daily station). For SILO vs ERA5, rainfall accumulations were paired and tested in both directions (i.e. each ERA5 grid point within the SILO region was matched to the nearest SILO position – and each SILO value was matched with the nearest ERA5 estimate). For each combination, the correlation coefficient (R) was also obtained. Each rainfall event was further assessed in terms of its contribution to total rainfall during the assigned water year (Table 2), with the percentage of TC-attributed precipitation extracted for the 1-year period. The water year for each event was selected based on the most impacted region.

4. Results

Three datasets (ERA5, GHCN-Daily and SILO) were accessed to obtain rainfall totals for individual TC events. The regional SILO database only covers Australia and was thus only applied for TCs Oswald and Seroja. For each storm, differences in rainfall estimates were compared. All correlation coefficients between pairs of rainfall databases are significant (p-value < 0.05; Figs. 4 and 5). Conversely, the paired Wilcoxon signed-rank tests revealed biases for several storm events, with hurricane Eta, and TCs Idai, Seroja and Oswald all returning p-values below 0.05 when comparing GHCN-Daily with its corresponding ERA5 estimate (Table 3). In terms of contribution to annual precipitation during the respective water year, the largest percentage was obtained for TC Oswald, with GHCN-Daily (SILO) estimating up to 63.0% (57.1%) of the annual rainfall accumulation (Table 4). Conversely, the ERA5 approximation is lower at 49.4% – and surpassed by its assessment for the tandem event of TCs Idai and Kenneth (50.0%).

4.1. TC idai and TC Kenneth (March/April 2019) – south-western Indian Ocean (Madagascar, Mozambique, Malawi, Zimbabwe, Comoros)

Two intense TCs severely impacted the south-western African region about a month apart, with TC Idai making landfall during the night of 14 to 15 March 2019 – and TC Kenneth on 25 April 2019. Especially Mozambique was greatly affected by both events, with catastrophic flooding affecting vast areas. For TC Idai, the greatest rainfall totals were associated with the central west-coast of Madagascar, central Mozambique, southern Malawi and eastern Zimbabwe. The track of TC Kenneth was located slightly more northward, with a concomitant shift of the associated precipitation that was more moderate – albeit still deadly – in comparison with the precursory storm event. Besides Mozambique, TC Kenneth also severely impacted on the Comoros, an archipelagic country situated north-west from Madagascar.

The maximum event rainfall approximated by ERA5 amounts to 713.1 mm for TC Idai (Fig. 6a). In comparison, GHCN-Daily stations only measured precipitation up to 324.1 mm (Fig. 6c), less than half (45.4%) of the reanalysis estimate. Similarly, the largest event total for TC Kenneth differed from 315.9 mm (GHCN-Daily; Fig. 6d) to 439.3 mm (ERA5; Fig. 6b). However, only the rainfall estimates for TC Idai are statistically different (p-value = 0.009) between the two datasets (Table 3), with ERA5 approximations generally exceeding those from GHCN-Daily (Fig. 4a). For the combined precipitation of the two events, ERA5 estimates that up to 50.0% of the annual total for the water year (August 2018 to July 2019) is associated with TCs Idai and Kenneth (Fig. 6e; Table 4).

The in-situ coverage via relevant (i.e. south-east African) stations contained in GHCN-Daily is patchy. Accordingly, a very limited number of stations is available to provide in-situ rainfall estimates, with few sites located within the main rain zone highlighted by elevated ERA5 precipitation. Although ERA5 suggests that precipitation associated with TC Idai mostly fell offshore, broad bands of significant rain are also evident inland. Further, ERA5 presents the largest rainfall estimate over land for TC Idai, near the decay location of TC Kenneth – a region not

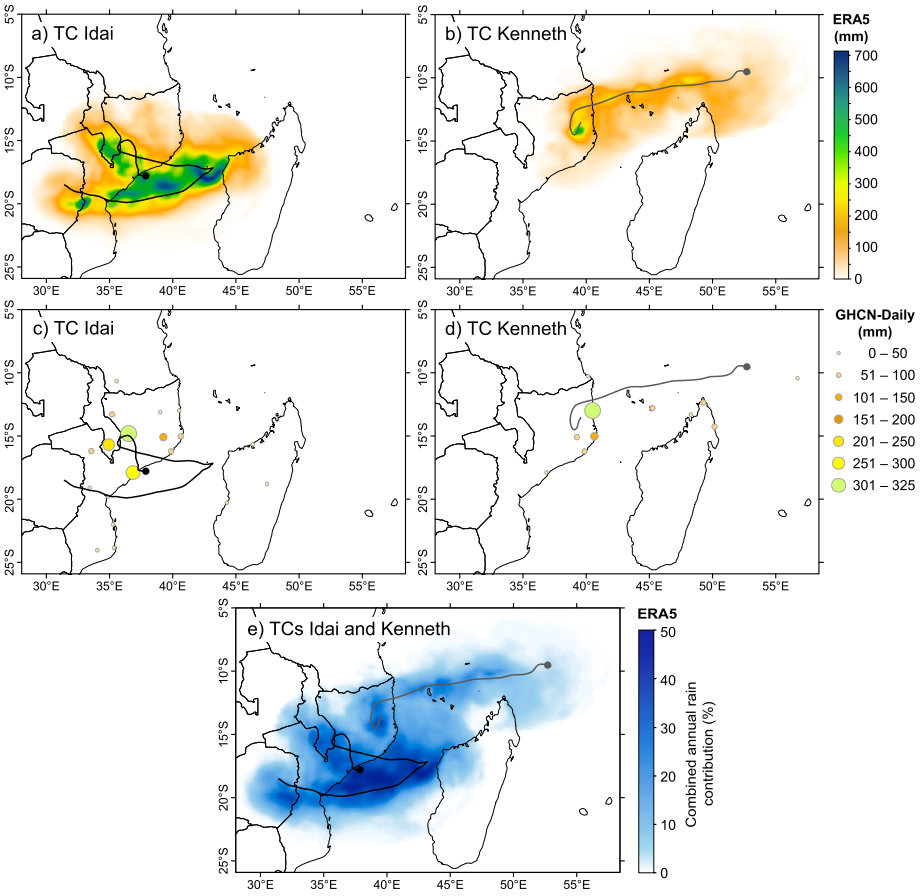


Fig. 6. a-d) Event rainfall accumulation (mm) and e) combined percentage of total rainfall during the water year (August 2018 – July 2019) associated with the passage of a,c,e) TC Idai (black) and b,d,e) TC Kenneth (dark grey) based on a,b,e) ERA5 and c,d) GHCN-Daily.

covered by any station. Instead, the peak gauge-derived rainfall depth is from a coastal GHCN-Daily site, for which ERA5 indicates relatively moderate event rainfall. Because of data gaps both during the event (Fig. S1) and the remainder of the water year, GHCN-Daily gauges could not offer an estimate for the proportion of annual precipitation attributable to the tandem TC event.

4.2. Cyclone Nisha (November 2008) – northern Indian Ocean (Sri Lanka, India)

In terms of intensity, cyclone Nisha was a relatively weak storm event that nevertheless had a devastating impact on both Sri Lanka and India. The agreement between the two rainfall datasets (GHCN-Daily and ERA5) is relatively strong for the entire region (Fig. 4c), except for the coastal area to the north of the track where the estimated event rainfall depth by GHCN-Daily is notably higher than the ERA5 approximation (Fig. 7a,b). Peak rainfall by GHCN-Daily reaches 701 mm in this area, whereas ERA5 provides a maximum rainfall accumulation of 583 mm, with the heaviest rainfall focussed south of the track – and mostly offshore. For ERA5, the peak estimate for annual precipitation contribution is 30.6% during the water year (February 2008 to January 2009; Fig. 7c; Table 4). As with the tandem TC event in south-eastern Africa, the temporal coverage of all relevant weather stations was incomplete during the water year and – on many instances – also during the passage of cyclone Nisha (Fig. S2).

Both datasets notably highlight the focus of precipitation to the north-east of the track (i.e. front-right quadrant), with a much larger area receiving substantial rainfall totals in that region. This spatial feature is attributable to the interaction between the landmass and converging onshore winds, enhancing rainfall in that quadrant.

4.3. TC Seroja (April 2021) – south-eastern Indian Ocean (Indonesia, Timor-Leste, Australia)

Severe TC Seroja formed in the Timor Sea in April 2021. The storm brought extensive rainfall and flooding to Indonesia and Timor-Leste – nations where TCs rarely make landfall (Kurniawan et al., 2021). TC Seroja was initially very slow-moving, aggravating the amount of rainfall delivered to the nearby regions. When the storm finally moved south-westward parallel to the Western Australian coastline, TC Seroja closely interacted with nearby TC Odette, with the two cyclonic vortices approaching and moving around each other in a rare phenomenon called the Fujiwhara effect, named after Japanese meteorologist Sakuhei Fujiwhara (e.g. Fujiwhara, 1922). TC Seroja ultimately absorbed TC Odette, making landfall in south-western Australia before dissipating in the Southern Ocean.

Based on the track from the IBTrACS dataset, a rainfall accumulation of 406 mm is estimated for ERA5, with the heaviest precipitation focussed on southern Indonesia between the islands of Sumba and Timor (Fig. 8a). However, event rainfall derived from GHCN-Daily is much lower, peaking at 122.0 mm. This discrepancy is attributable to the near-complete absence of relevant GHCN-Daily measurements in Indonesia and Timor-Leste (Fig. 8d), despite Kurniawan et al. (2021) demonstrating that several stations were operational in that region during the passage of the storm. Precipitation data for the event are only available for a single Indonesian station (IDM00097260 – Sumbawa Besar) within the vicinity of TC Seroja, although only 50% of the relevant days contained rainfall information (Fig. S3). The GHCN-Daily estimates thus nearly exclusively reflect the Australian rain conditions. Because the creation of SILO was based on station data, its output (Fig. 8b) closely corresponds to the GHCN-Daily precipitation patterns.

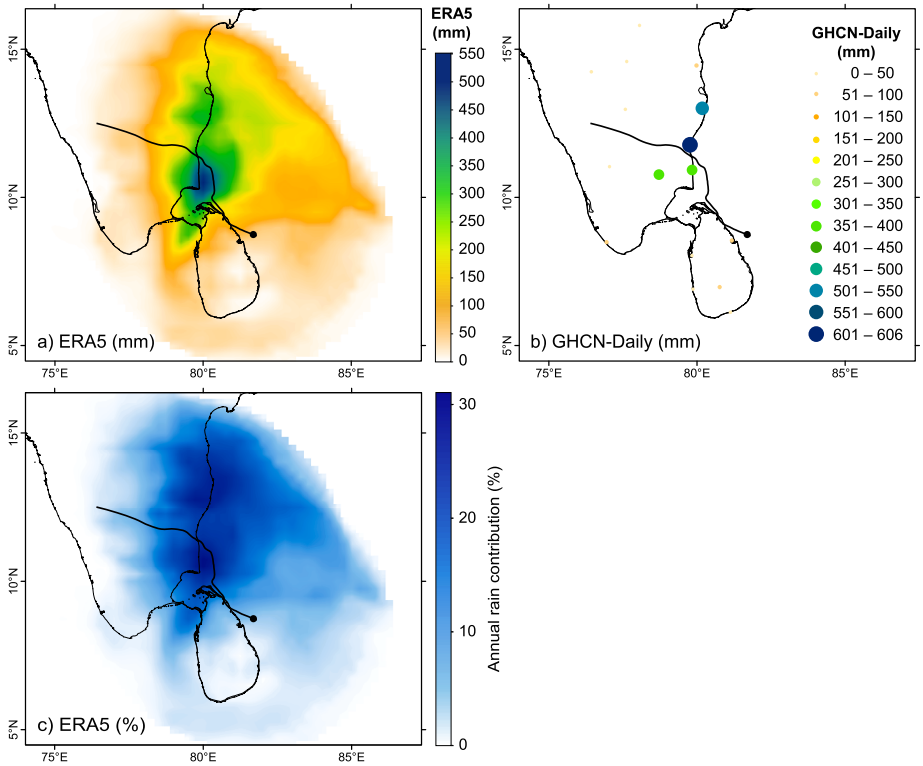


Fig. 7. a,b) Event rainfall accumulation (mm) and c) percentage of total rainfall during the water year (February 2008 – January 2009) associated with the passage of cyclone Nisha (2008) for a,c) ERA5 and b) GHCN-Daily. The results are based on a 500 km radius along the cyclone track.

A notable aspect of the IBTrACS track for TC Seroja is that the trajectory excludes the initial two days, when the low was still strengthening in the Timor Sea while quasi-stationary. Consequently, relevant rainfall was omitted when utilising the IBTrACS data, demonstrated when re-running the *rain_tracker* toolbox with the BoM-derived track that begins 48 hours earlier (Fig. 8c). ERA5 produces a maximum event rainfall of 511 mm, an increase of 25.9% compared to the IBTrACS result. Precipitation estimates specifically increased for islands of both Indonesia and Timor-Leste, regions that experienced extensive damage from the deluge. Rainfall depths south-east from the initial track position are also substantially higher. An area relatively distant from the track exhibits high precipitation amounts, reflective of the still weak low that has not yet contracted the moisture closer to its core. Because the two tracks (IBTrACS vs BoM) were comparable during the overlapping period, neither the SILO nor the GHCN-Daily outputs differed significantly for the two trajectories.

In terms of contribution to annual precipitation, ERA5 estimates that up to 45.4% (46.0%) is associated with TC Seroja based on the IBTrACS (BoM) path (Fig. 9a,c; Table 4) – and a water year (August 2020 to July 2021) tailored to the most storm-affected, northern region (Indonesia and Timor-Leste). In comparison, SILO (21.3%) and GHCN-Daily (16.7%) ascribe lower proportions of the annual precipitation to TC Seroja (Fig. 9b,d), attributable to the absence of relevant data outside Australia. Although the water year could therefore have been adjusted for the south-western Australian region (February 2021 to January 2022), the northern period was maintained for consistency.

4.4. TC Oswald (January 2013) – south-western Pacific (Australia)

TC Oswald affected a broad swath of the Australian continent in January 2013, producing record-breaking, severe precipitation and associated flooding (Leroux et al., 2020). Rainfall was heaviest in coastal regions of north-eastern and eastern Australia (Fig. 10). Inland rainfall was very limited, demonstrating the influence of converging onshore winds that encourage elevated precipitation in near-coastal settings – especially in the front-left TC quadrant, as previously determined by Deng and Ritchie (2020). Deng and Ritchie (2020) have also detailed the mechanisms involved in producing several coastal precipitation hotspots during the event.

Statistically, all three precipitation datasets significantly differ in terms of their rainfall estimates (Table 3). The close relationship between SILO and GHCN-Daily estimates (Fig. 5a) is attributable to the SILO-derivation method that heavily relies on in-situ information, with rainfall accumulations in excess of a metre during the passage of TC Oswald (Fig. 10b,c). In comparison, the maximum estimated event rainfall based on the ERA5 dataset is limited to about 600 mm (Fig. 10a), nearly half of what was observed by weather stations. This discrepancy is probably ascribable to the lower spatial resolution of ERA5, disallowing the identification of the most intense rainfall that is frequently evident on very small, localised scales. However, ERA5 has the benefit to also provide rainfall estimates for offshore regions. In this instance, the largest rainfall total was registered for the north-eastern Gulf of Carpentaria, in northern Australia.

GHCN-Daily estimates that up to 63.0% of the annual precipitation is linked to TC Oswald during the water year (September 2012 to August 2013; Fig. 11c; Table 4). At the station with the largest proportional contribution (ASN00040334 – Biggenden Alert), 324.6 mm (92.6% of the event rainfall) was received within 24 hours. In comparison, the next-highest 1-day precipitation total at that site was limited to 42.0 mm during that water year, with an annual total of 556.4 mm. A comparable peak contribution to annual rainfall accumulation was also derived from SILO (57.1%; Fig. 11b), whereas the ERA5 approximation was lower at 49.4% (Fig. 11a).

4.5. Typhoon Morakot (August 2009) – north-western Pacific (Taiwan, Philippines, mainland China)

Originating in the western Pacific, typhoon Morakot was the

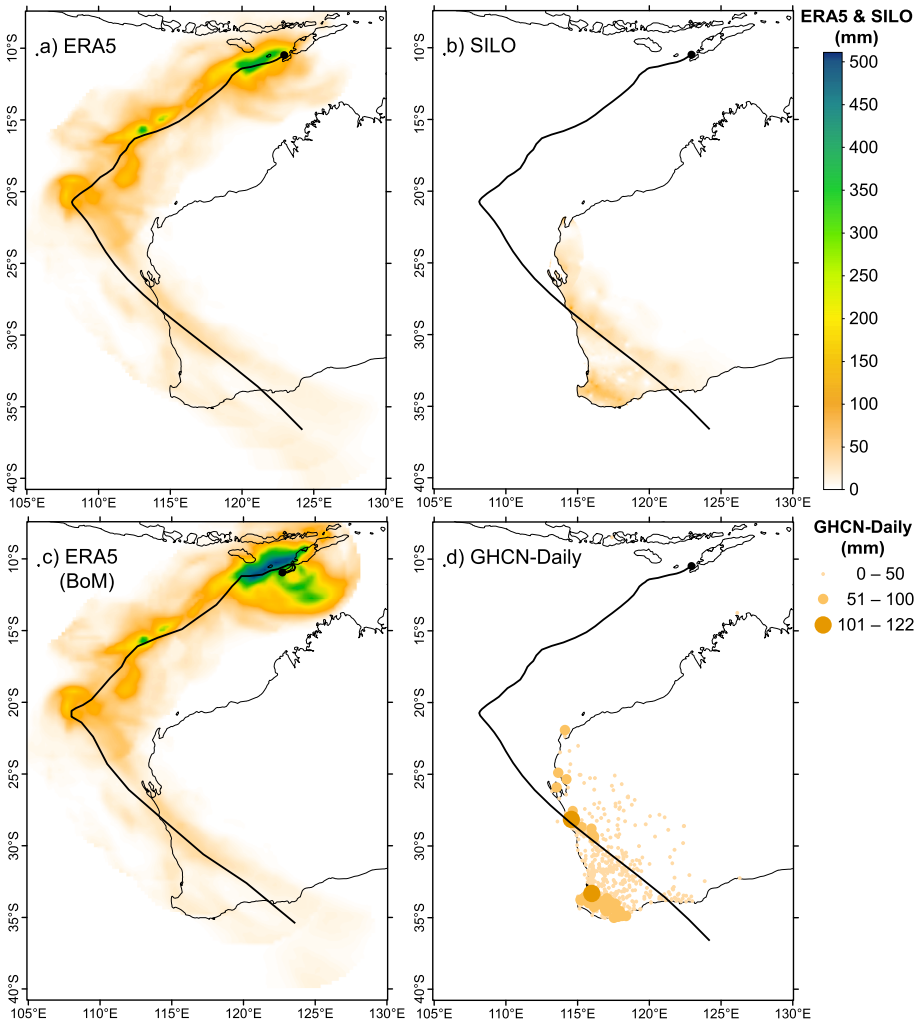


Fig. 8. Total rainfall accumulation (mm) during the passage of TC Seroja (2021) based on a) ERA5, b) SILO and d) GHCN-Daily along the IBTrACS path – and c) the ERA5 rainfall associated with the BoM track.

deadliest cyclone to affect Taiwan, delivering extreme, record-breaking rainfall to the island in 2009 (Lin et al., 2011). Typhoon Morakot subsequently also passed through mainland China and South Korea. ERA5 estimates a maximum event rainfall of 1197 mm for the event, with the most severe precipitation focussed on the southern parts of Taiwan (Fig. 12a). Although large, this rainfall amount is less than half the reported accumulation over a 9-day period (3031.5 mm) for the Weiliaoshan station in Taiwan (Wu et al., 2011), and the total over 100 hours (2855.0 mm) at the A-Li station (Lin et al., 2011). Further, GHCN-Daily only reached a maximum of 260 mm (a coastal station in mainland China; Fig. 12b), attributable to the complete absence of Taiwanese stations within the database.

On an annual scale, ERA5 credits up to 42.7% of the total precipitation during the water year (January to December 2009) to typhoon Morakot (Table 4), with the highest percentages observed in southern Taiwan (Fig. 12c). In comparison, the GHCN-Daily estimate is relatively modest at 17.8% because of the absence of rain gauges in the most affected areas (Fig. 12d).

4.6. Hurricanes Eta and Iota (November 2020) – Caribbean Sea / western Atlantic (Central America)

Somewhat similar to the dual cyclone events in south-eastern Africa (TCs Idai and Kenneth; section 4.1), two tandem, major hurricanes impacted the Caribbean region in November 2020. The initial storm, hurricane Eta, first traversed Nicaragua's east coast before erratically moving northwards through several other nations during its two-week long path, bringing extensive rainfall to many regions. Before its decay, the low associated with the follow-up event – TC Iota – formed, later replicating the initial path of TC Eta by moving westward over Nicaragua, dissipating near the Pacific coast of the country.

For hurricane Eta, the *rain_tracker* toolbox estimates event rainfall of up to 455.6 mm for ERA5 (Fig. 13a), the largest totals mainly concentrated along the Central American coastline and its offshore regions. Although the maximum GHCN-Daily estimates are similar, it places the peak rainfall (up to 422.3 mm) in the USA (Florida; Fig. 13b), with Central America only reaching 243.6 mm – associated with a Honduras station. Conversely, in-situ records from Nicaragua are notably absent.

The successor (hurricane Iota) delivered comparable rainfall totals to Central America, with the peak ERA5 rain accumulation (up of 486.7 mm) observed just offshore from Nicaragua (Fig. 13c). In comparison, GHCN-Daily stations recorded a maximum of 303.3 mm – in a Honduras site (Fig. 13d). Nicaragua was again devoid of in-situ measurements.

The combined precipitation of hurricanes Eta and Iota accounts for up to 34.7% of the total accumulation during the water year (March 2020 to February 2021) based on ERA5 (Fig. 13e). In contrast, the peak contribution by GHCN-Daily rain gauges (16.4%; Fig. 13f) is less than half the estimated ERA5 value (Table 4). Similar to TC Seroja and typhoon Morakot, this comparatively low percentage is ascribable to the absence of weather stations in the most affected regions. Assessments of

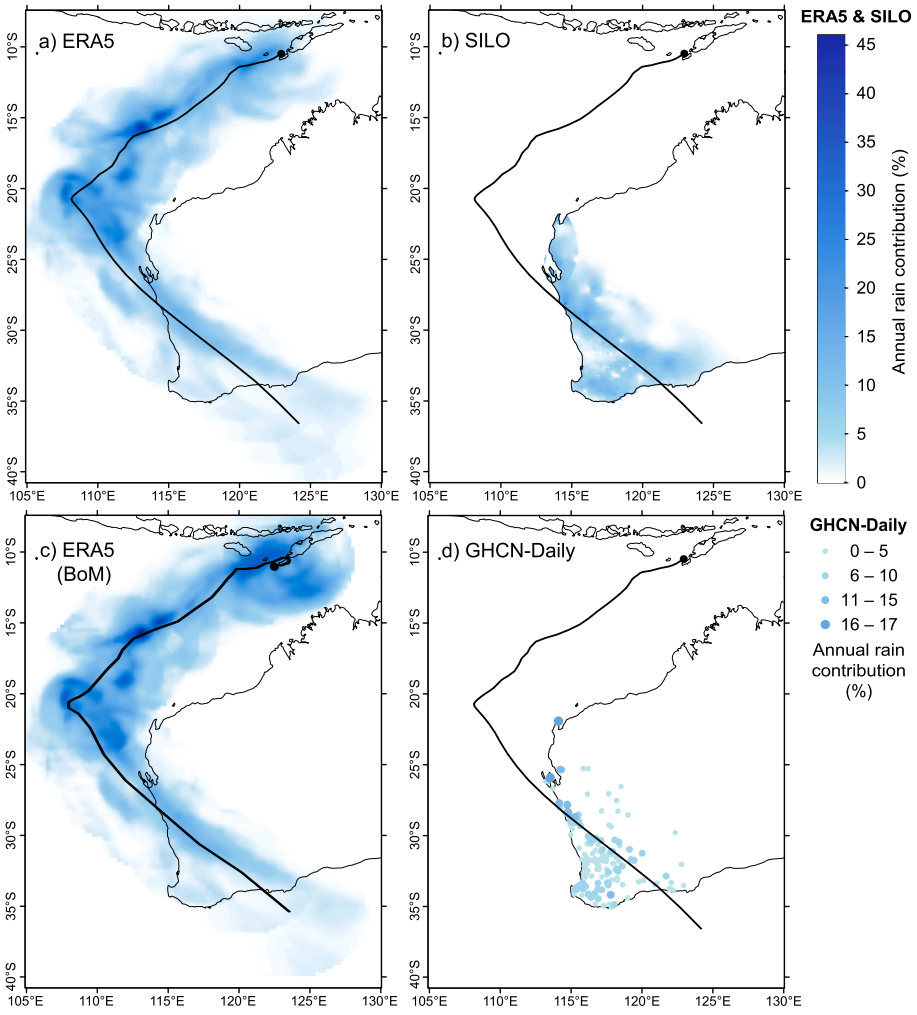


Fig. 9. Percentage of total rainfall during the water year (August 2020 – July 2021) associated with the passage of TC Seroja (2021) based on a) ERA5, b) SILO and d) GHCN-Daily along the IBTrACS path – and c) the ERA5 rainfall associated with the BoM track.

annual rainfall accumulation are further hampered by incomplete data records at all Central American rain gauges.

4.7. Applicability of 500 km radius

For two TC events, the extracted rainfall patterns – when a radius of 500 km was applied – suggest that the range was too small to capture the full impact of the cyclones. Affected storms include cyclone Nisha (section 4.2) and TC Oswald (section 4.4). Consequently, several larger radii (600 km, 750 km and 1000 km) were also tested to determine whether the larger ranges are potentially more optimal for these two cyclones.

For cyclone Nisha, a rapid drop-off in extracted ERA5 rainfall is observed to the north-east of the track with the 500 km radius (Fig. 14a). This spatial decline in event rainfall becomes more gradual when larger ranges are applied (Fig. 14c,e,g). Further, at longer radii, rainfall produced by feeder bands becomes evident south of the track (Fig. 14g). For GHCN-Daily data, variations in radii have a very limited effect on extracted rainfall totals (Fig. 14b,d,f,h). The incorporation of additional weather stations constituted the main difference (Fig. S7b,d,f). One of these additional stations measured 46 mm during the passage of cyclone Nisha.

Rainfall accumulations vary more substantially with different radii for TC Oswald, with one GHCN-Daily station (ASN00032119 – Mena Creek Post Office) increasing its event tally by 550 mm when a 1000 km range is applied (Fig. S8i). In comparison, SILO (446.5 mm) and ERA5 (364.7 mm) are associated with slightly more modest rises in event rainfall (Fig. S8g,h). Although the overall rainfall patterns do not change substantially with modified radii, the most notable variations are in coastal regions of northern and south-eastern Australia (Fig. 15).

5. Discussion

On a global scale, many underlying factors (e.g. TC duration, strength, size and location – and topography) influence the relationship of rainfall and TC occurrence. Rainfall estimates for a specific event and location can also greatly vary among different data sources – both within a dataset (e.g. two nearby weather stations) and separate sources (e.g. GHCN-Daily vs ERA5).

Although gauge data have the advantage to represent direct measurements of in-situ precipitation, weather stations are frequently affected by outages caused by extreme weather events (e.g. wind or flood impact). Further, strong winds may reduce the accuracy of precipitation measurements (Jarraud, 2008). In addition, the station network tends to be biased towards more populated areas, with less urbanised regions often underrepresented – including elevated, relatively remote areas most severely affected by rainfall enhancement because of orographic uplift, resulting in an underestimation of actual precipitation. In terms of GHCN-Daily, even though the applied version contains over 118,000 individual station sources, several countries and

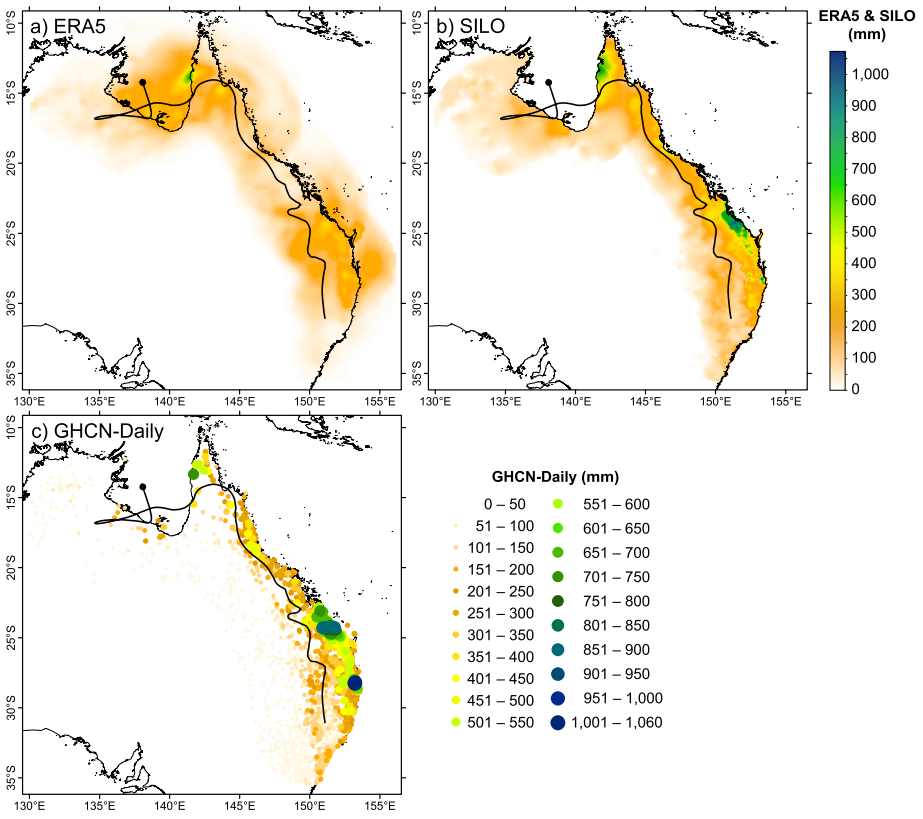


Fig. 10. Total rainfall accumulation (mm) during the passage of TC Oswald (2013) for a) ERA5, b) SILO and c) GHCN-Daily. The rainfall totals are based on a 500 km radius along the TC track.

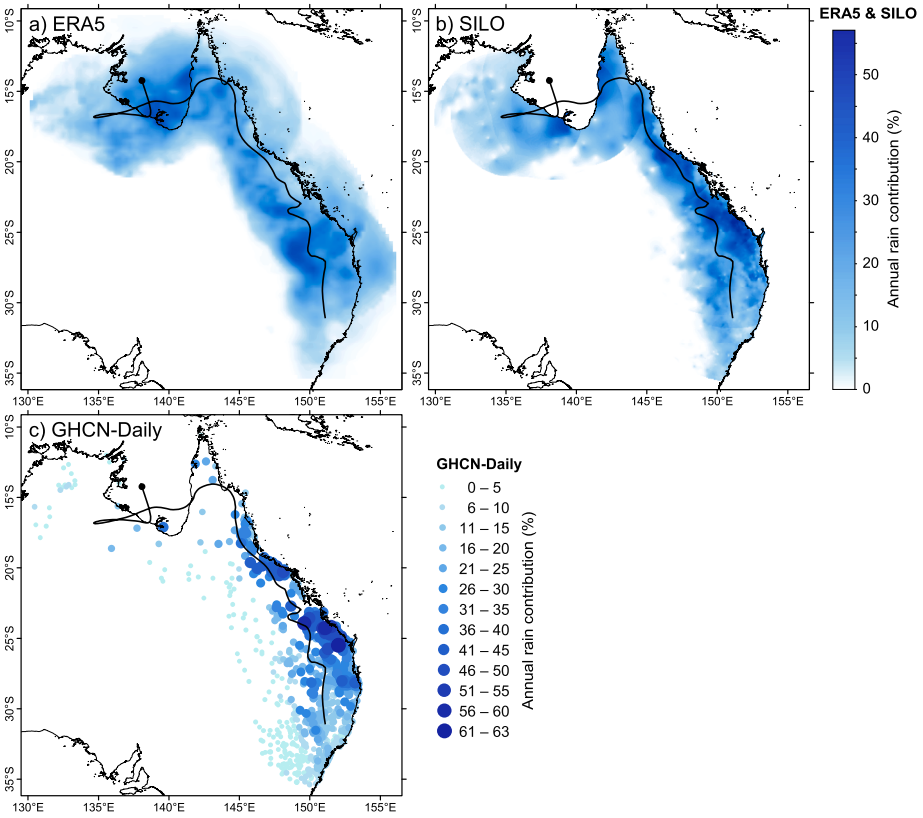


Fig. 11. Percentage of total rainfall during the water year (September 2012 – August 2013) associated with the passage of TC Oswald (2013) for a) ERA5, b) SILO and c) GHCN-Daily. The results are based on a 500 km radius along the TC track.

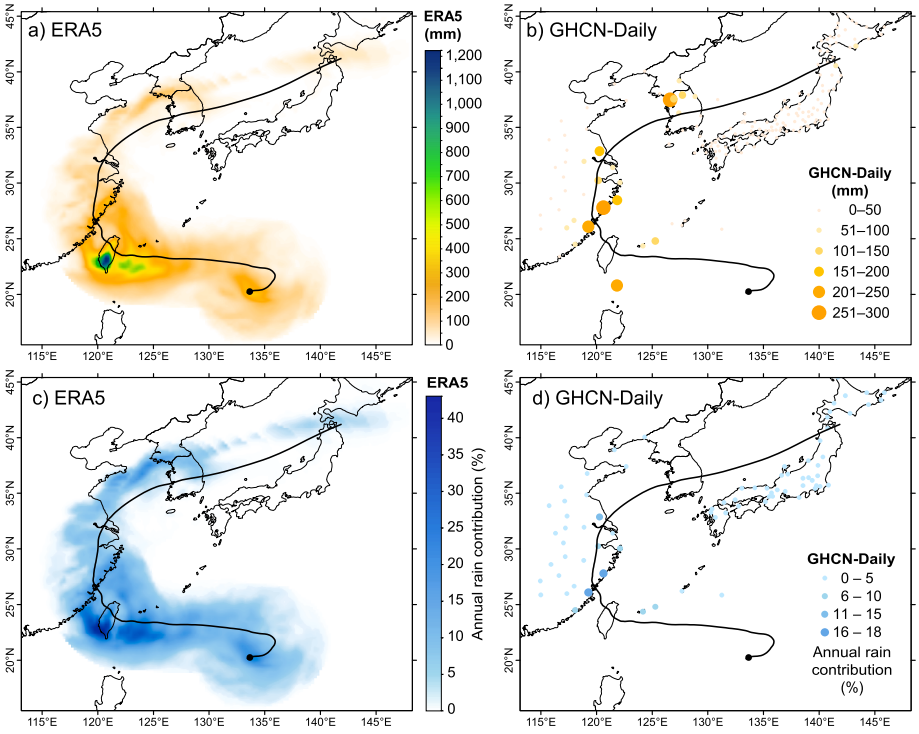


Fig. 12. a,b) Event rainfall accumulation (mm) and c,d) percentage of total rainfall during the water year (February 2008 – January 2009) associated with the passage of typhoon Morakot (2009) based on a,c) ERA5 and b,d) GHCN-Daily.

regions were completely absent during either the relevant event (e.g. Nicaragua for hurricanes Eta and Iota, and Indonesian areas severely affected by TC Seroja) or throughout the dataset (e.g. Taiwan and Timor-Leste).

5.1. Rainfall dataset intercomparison

5.1.1. ERA5

The relatively coarse (0.25°) representation of the precipitation field in ERA5 is not conducive to capturing finer-scale rainfall heterogeneity. However, high (hourly) temporal resolution is beneficial for studies investigating rainfall intensity (rain bursts) at sub-daily timescales.

For TC Idai, Emerton et al. (2020) estimated rainfall accumulations close to 800 mm based on the Integrated Multi-SatellitE Retrievals for Global Precipitation Measurement (IMERG) algorithm by NASA (National Aeronautics and Space Administration) – and a radius of 5° around individual IBTrACS track positions. Although their estimate for the rainfall event surpasses our ERA5 peak (713.1 mm; section 4.1), Emerton et al. (2020) also highlighted that IMERG tends to overestimate high-intensity precipitation.

5.1.2. GHCN-Daily

Direct observations of rainfall via weather stations often play a crucial role for validation of reanalysis datasets. However, the quality of these in-situ data is imperative – in terms of both accuracy and reliability (i.e. near-continuous coverage, with few data gaps over time). For TC Idai, ERA5 measured significantly higher rainfall totals than the corresponding GHCN-Daily station (Fig. 4a), at least partially attributable to frequent data gaps in relevant gauges (Fig. 16a and Fig. S1a). Similarly, of the stations that were operational during TC Kenneth or cyclone Nisha, fewer than 50% (Fig. 16b,c; Fig. S1b and Fig. S2) were active throughout the relevant period (i.e. when within 500 km from the storm). In terms of temporal coverage, GHCN-Daily data were much more reliable during TC Seroja (Fig. S3), TC Oswald (Fig. S4), typhoon Morakot (Fig. S5), and hurricanes Eta and Iota (Fig. S6), with most sites operational throughout the passage of the storm (Fig. 16d–h).

Further, some regions are characterised by paucity in GHCN-Daily station availability throughout the period covered by the database. Taiwanese gauges are completely absent from GHCN-Daily, whereas the number of incorporated stations for Nicaragua (six sites) is very limited. In comparison, Australia and the USA provide relatively dense station networks. GHCN-Daily also comprises 104 Indonesian stations. However, only one gauge within the vicinity of TC Seroja contained some (albeit limited) precipitation data for the event. Consequently, in-situ measurements for TC Seroja are mostly derived from Australian GHCN-Daily sites, thus not capturing the most intense and prolonged rainfall in the region of Indonesia and Timor-Leste. Similarly, because of the absence of Taiwanese stations within GHCN-Daily, the maximum event rainfall estimated by that dataset for typhoon Morakot is much lower than by ERA5 (Fig. 12a,b).

5.1.3. SILO

SILO exhibits a very strong relationship with in-situ data (i.e. GHCN-Daily; Fig. 5a,d), attributable to the derivation method of the database. Nevertheless, although the paired Wilcoxon signed-rank test confirmed that the two databases are statistically comparable for TC Seroja (p-value = 0.629; Table 3), the test returned a significant difference for TC Oswald (p-value <0.001), with SILO displaying a tendency for higher estimates. Because of its higher spatial resolution (0.05°), SILO is able to capture finer, more nuanced details of the heterogeneous rainfall patterns compared to ERA5, as especially evident for TC Oswald (Fig. 10b). However, SILO is simultaneously also limited by its regional scale (solely covering Australia) and relatively low temporal (daily) data frequency. Further, the dataset is also highly reliant on dense station networks for its derivation and may thus have reduced accuracy in underrepresented areas (i.e. especially in the sparsely populated interior – and also in mountainous regions). Nevertheless, TCs are predisposed to more severely impact coastal regions, zones that tend to have the highest network density of rain gauges and are thus more reliably characterised by SILO.

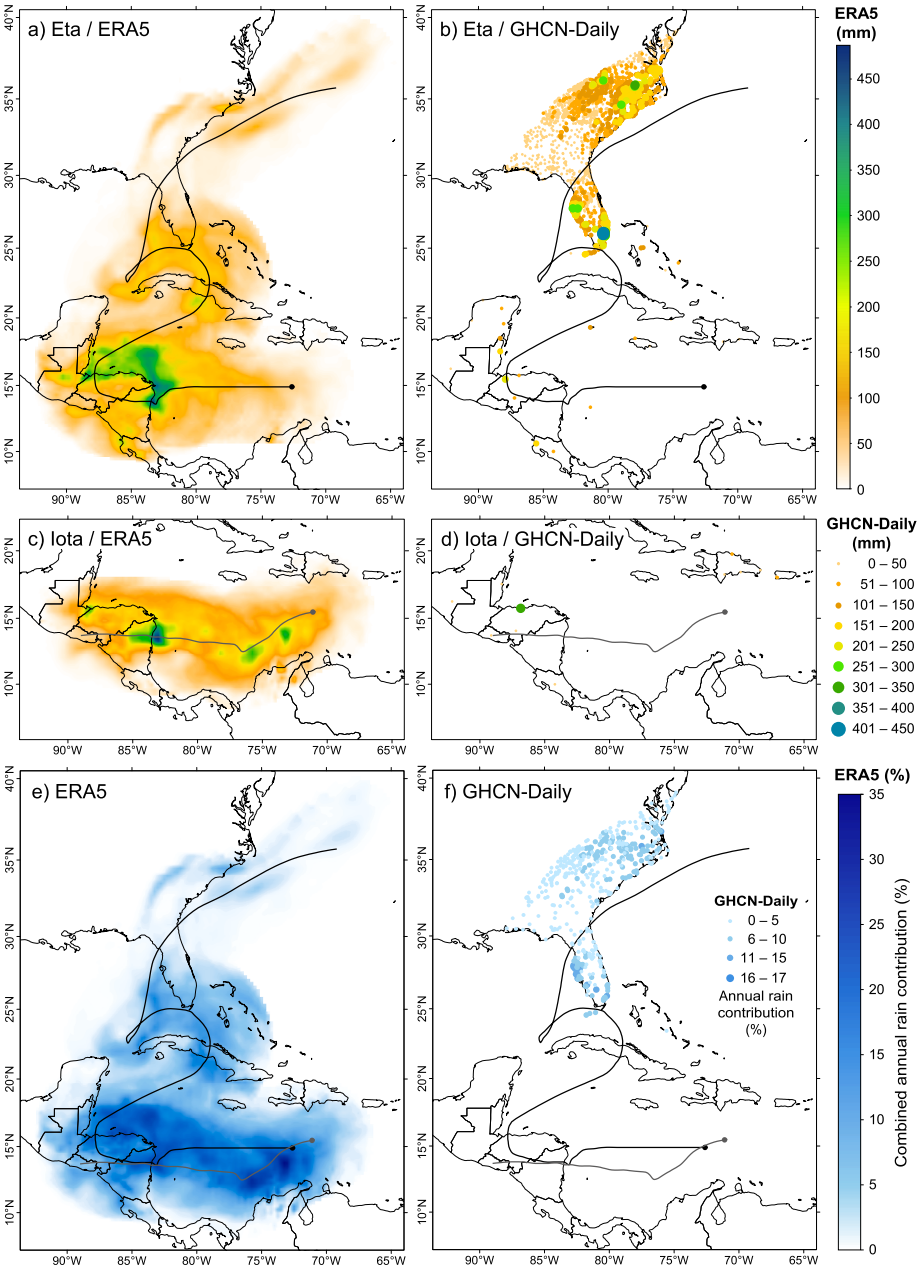


Fig. 13. a-d) Event rainfall accumulation (mm) and e,f) combined percentage of total rainfall during the water year (March 2020 – February 2021) associated with the passage of a,b,e,f) hurricane Eta (black) and c,d,e,f) hurricane Iota (dark grey) based on a,c,e) ERA5 and b,d,f) GHCN-Daily.

5.2. TC-related rainfall impact region

For TC-related rainfall analyses, a 500 km buffer is frequently applied along the track for the entire duration of the TC (Stansfield et al., 2020). However, the rainfall impact of TCs is dependent on both size and intensity, with Deo et al. (2021) suggesting that a 500 km radius may be slightly conservative to attribute rainfall to passing storms. Depending on the size of the TC and the study focus, a different radius may be more suitable. For example, Jiang et al. (2008) applied a 10° (~1100 km) radius to capture extreme storm sizes when investigating rain potential. Conversely, Ng et al. (2015) employed a range of radii (200–1000 km) to determine TC-related rainfall based on station data in Western Australia, settling on 650 km as their preferred choice. However, a static radius of 500 km (or higher) could overestimate TC-related precipitation over land because of the size fluctuation of the storms (Stansfield et al., 2020), with TCs generally rapidly declining in strength and size after landfall. Zhou and Matyas (2018) first created polygon features for rainfall fields above their selected threshold rates (2.5 mm/h and 5.0 mm/h). They then applied a combination of a 500 km radius (for the polygon centroids) and the radius of the outmost, closed isobar (for the polygons) to extract features that satisfy both criteria.

Storm size and strength both vary during the lifetime of a cyclone, significantly influencing the areal extent of precipitation. Because stronger events have a greater capacity for rainfall contraction, rainbands in weaker storms tend to affect a larger area (Touma et al., 2019). Further, TC-associated precipitation is frequently asymmetric around the storm circulation, with the unevenness inversely related to storm intensity (Chen et al., 2006). For example, offshore storms near the east coast of Australia deliver more rainfall to their southern flanks (i.e. front-left quadrant) because of moist onshore airstreams converging in the clockwise rotation. Conversely, drier offshore circulation in the northern segment (front-right quadrant) limits precipitation to areas

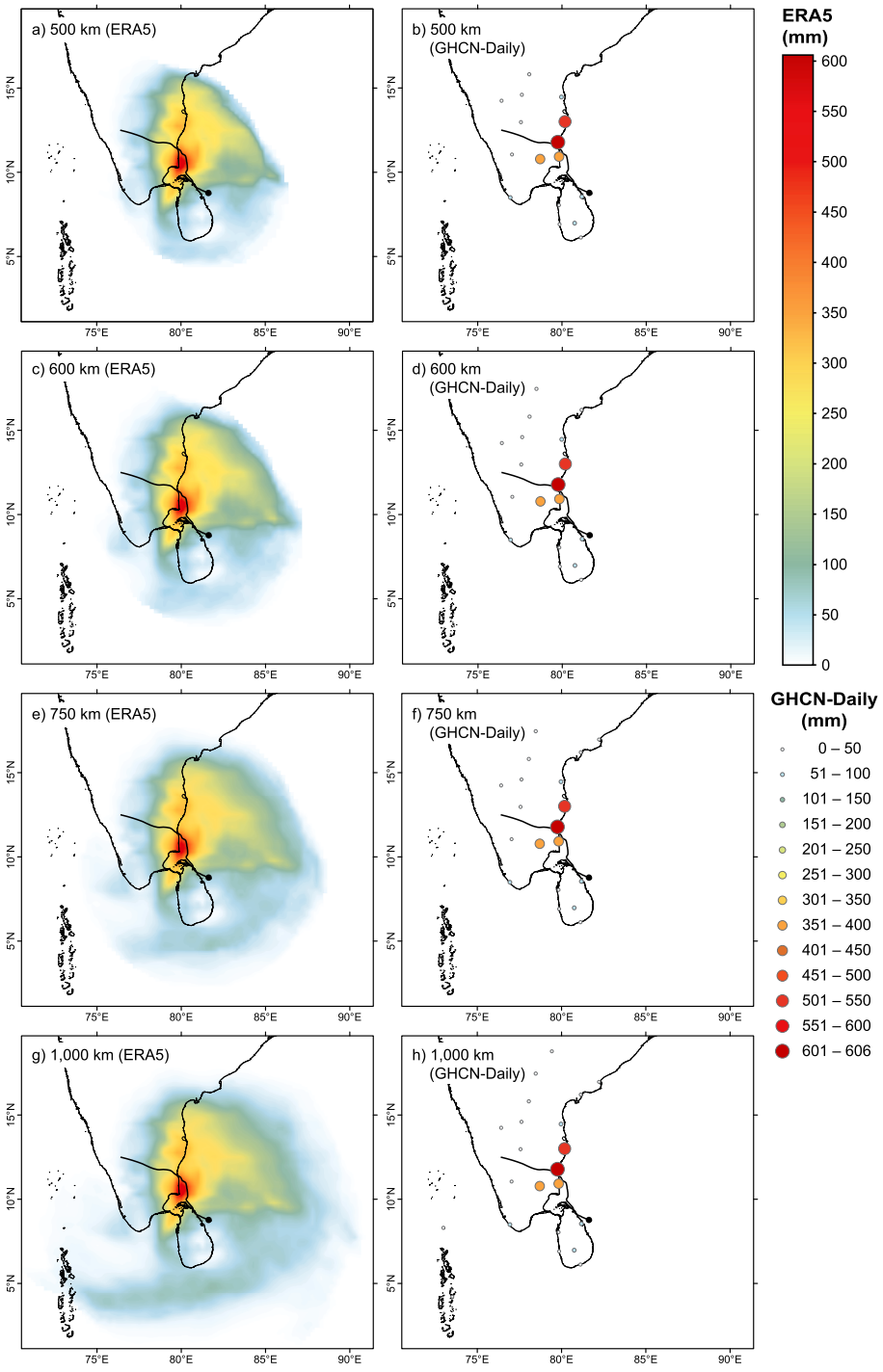


Fig. 14. Estimated event rainfall total for cyclone Nisha based on different radii. The applied radii include a,b) 500 km, c,d) 600 km, e,f) 750 km and g,h) 1000 km. Panels a,c,e,g) represent ERA5 and b,d,f,h) GHCN-Daily.

closer to the core in eastern Australia. For coastlines in the northern hemisphere, the opposite is applicable because of the reverse (anti-clockwise) rotation of these storms, delivering more rainfall to the front-right quadrant of TCs passing nearby landmasses in the Caribbean Sea, north-western Pacific and Bay of Bengal. Nevertheless, the application of a static radius for rainfall extraction does not necessitate symmetric precipitation. Instead, the assumption is made that any rainfall within the enclosed area is attributable to the TC.

The asymmetric rainfall pattern is also clearly evident during TC Oswald. The heaviest rainfall at individual sites along the east coast was generally observed while the storm was still situated to the north of the gauge, with negligible rainfall measured after the TC moved south of the station. For example, the Tully Sugar Mill weather station (ASN00032042) recorded 246.0 mm and 280.4 mm on 22 and 23 January (AEST), respectively (corresponding to the 48-hour period from 23:00 UTC on 20 January; cf. Fig. 1). TC Oswald surpassed the corresponding latitude (17.94°S) after 6 p.m. on January 23, 2013 (UTC), corresponding to the rain reporting date of 24 January local time. After the southward passage, minimal rainfall was recorded throughout the remainder of the storm's existence.

Furthermore, the application of a 500-km radius (or 1-hour window) revealed itself to be too conservative to encapsulate the full extent of the

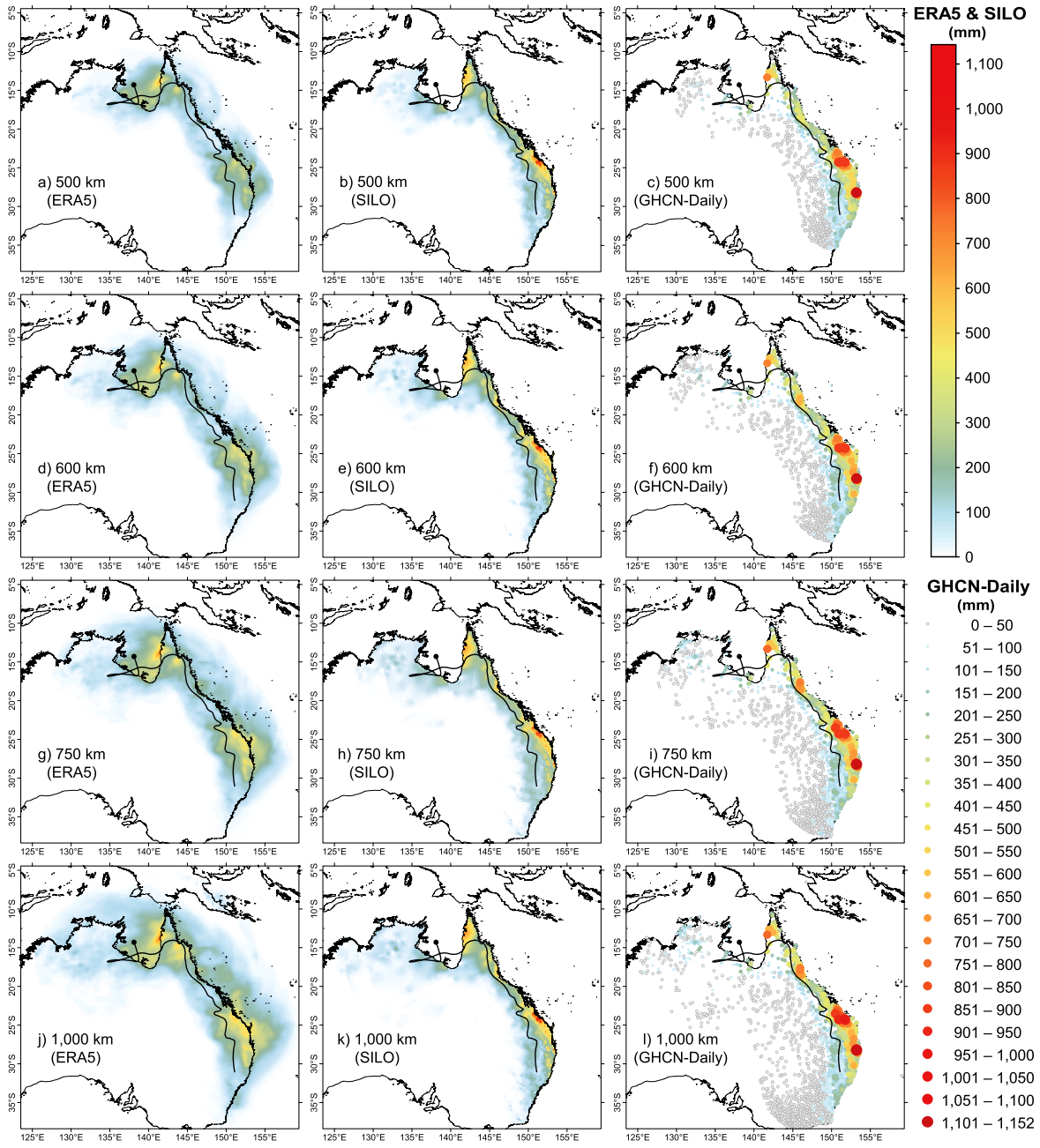


Fig. 15. Estimated event rainfall total for TC Oswald based on different radii. The applied radii include a-c) 500 km, d-f) 600 km, g-i) 750 km and j-l) 1000 km. Panels a,d,g,j) represent ERA5, b,e,h,k) SILO and c,f,i,l) GHCN-Daily.

rainfall impact by TC Oswald. For the Tully Sugar Mill gauge, three rain days were extracted by the toolbox (a total of 367.8 mm from 23 to 25 January local time), including one day immediately prior and after the passage of the storm. Consequently, with a further 78.0 mm measured on 21 January, the estimated rain accumulation constitutes a significant underestimation (by more than 324.0 mm) of the actual precipitation produced by TC Oswald. The actual rainfall accrual was thus close to 700 mm for that site. Similarly, the spatial distribution of estimated event rainfall for cyclone Nisha suggests that not all precipitation associated with the storm was aggregated when the 500 km radius was applied on ERA5 data – specifically to the north-east of the track – with values abruptly changing from over 50 mm to zero at the edge of the enclosed area.

Here, a range of larger radii (600 km, 750 km and 1000 km) was used – to compare with the results from the default (500 km) radius (Figs. S7 and S8) and evaluate which distance (from the cyclone centre) would be sufficient to capture most of the storm-relevant precipitation. In the case of TC Oswald, a radius of 750 km (Fig. 15g–i) appears most suited to aggregate pertinent rainfall, with a total of 698.2 mm extracted for the Tully Sugar Mill gauge (GHCN-Daily). At 600 km, the total for that site was significantly raised (614.8 mm) compared to the default 500 km radius (367.8 mm), although the result is still considered an underestimation. Conversely, at 1000 km, the site total increased to 732.2 mm but the larger radius notably caused the incorporation of unrelated precipitation signatures in other locations, most evident in the north-eastern section of the ERA5 output (Fig. 15j). However, an even larger radius (i.e. in excess of 1000 km) could also be argued for, noting that – even when the 1000 km radius was applied – the initial westward shift in the TC track resulted in the omission of precipitation (40 mm) that fell on the second day of the event registered at ASN00032119 (near-coastal

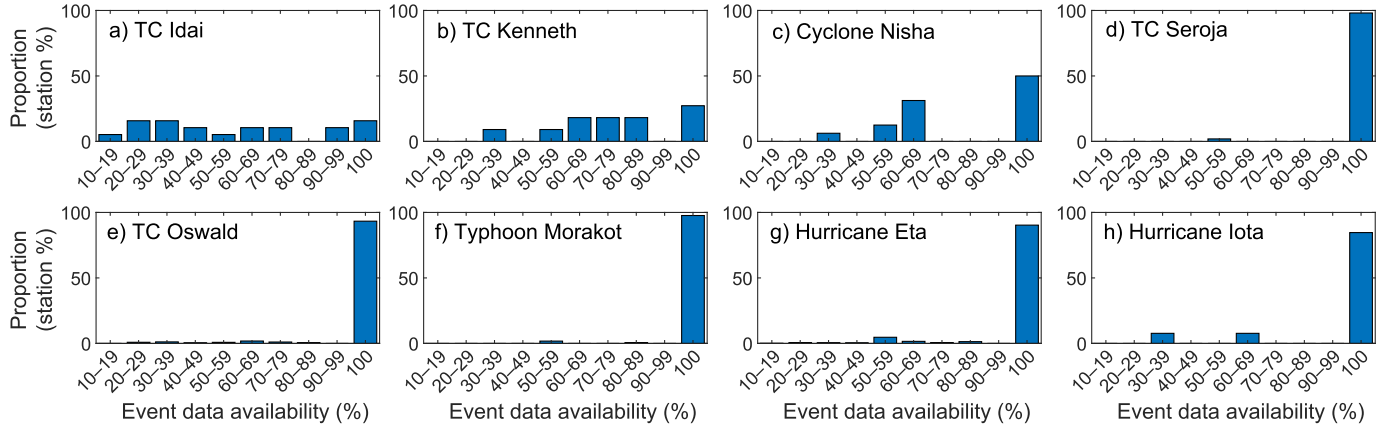


Fig. 16. Event coverage (%) for GHCN-Daily stations during relevant days (i.e. within 500 km from the storm centre) for a) TC Idai, b) TC Kenneth, c) cyclone Nisha, d) TC Seroja, e) TC Oswald, f) typhoon Morakot, g) hurricane Eta and h) hurricane Iota.

site in north-eastern Australia).

For cyclone Nisha, the abrupt drop in event rainfall in the north-eastern region indicates that an insufficiently large area was considered for the rainfall extraction (Fig. 7a). Thus, the three additional, larger radii were also evaluated for that cyclonic event (Fig. 14c–h). However, in terms of GHCN-Daily data, increasing the radius up to 1000 km did not result in greater event rainfall accumulations being detected (Fig. S7b,d,f), with the only difference the greater number of stations (and their associated rainfall) contained within the larger radii. Conversely, ERA5 – with its much more expansive spatial coverage – reveals the effects of feeder bands to the south of the cyclone track (south of Sri Lanka) when larger radii are applied (Fig. 14e,g). Unlike for TC Oswald, a radius of 1000 km does appear to be appropriate to capture the rainfall characteristics of this specific storm.

The catchment response (i.e. streamflow behaviour and flood severity) to intense precipitation events can widely vary dependent on antecedent conditions. In mid-January 2019, tropical storm Desmond pre-wetted the Mozambiquan region later struck by TC Idai (Emerton et al., 2020). Precipitation by TC Idai, in turn, partially overlapped areas affected by TC Kenneth a few weeks later. Similarly, hurricanes Eta and Iota impacted Central America just two weeks apart in November 2020, with hurricane Iota closely matching the initial path of hurricane Eta. These tandem events would therefore have greatly exacerbated the severity and spatio-temporal extent of flash and riverine floods because of elevated antecedent soil moisture. Additionally, TCs are not only an important source for replenishing water resources, they can also result in rainfall deficits in distant areas affected by moisture airflow redirection.

5.3. Limitation and inconsistencies

Within a database, the time zone may not necessarily be constant. Although UTC+0 is commonly applied – especially for global reanalysis and TC track datasets – several sources have variable reference times (i.e. based on local time). This aspect can be problematic when conducting studies in areas that cover multiple time zones. Further, some regions undergo daylight saving – or may have in the past – further modulating the reference time. GHCN-Daily and some gridded rainfall datasets (e.g. SILO and AWAP) are among the affected. In addition, the 24-hour precipitation cut-off is not necessarily at midnight local time (Fig. 1) – and may historically have changed (e.g. DeGaetano, 2000). For example, the daily rainfall station data of the BoM cover 24-hour periods ending at 9 a.m. local time (Fig. 1), with the end-time designating the assigned date. With BoM data incorporated into GHCN-Daily and SILO, this has implications for matching the track position with the correct rain period. Although the *rainfall_tracker* toolbox automatically adjusts the time for the two regional data grids and Australia-based (AS code) GHCN-Daily rainfall to offset the 9 a.m. cut-off, the code expects only one time zone per run. Thus, if a TC travels over multiple time zones, the toolbox may need to be run separately for each time zone (but with the complete track in each instance) and the extracted rainfall data aggregated for individual regions.

Table 5
The 24-hour rainfall cut-off times for GHCN-Daily applied by the *rainfall_tracker* toolbox – based on current time zones and daylight-saving practises.

Local (UTC) 24-hour end-time	Daylight saving	Country	Source
8:30 (3:00)	No	India	Yatagai et al. (2012)
9:00 (23:00 to 2:00)	Some states/ territories	Australia	BoM
20:00 (12:00)[a]	No	China	Yatagai et al. (2012); Ying et al. (2014)
00:00 (15:00)	No	Japan	Yatagai et al. (2012)

[a] Applicable from 1985. For 1949–1984, Ying et al. (2014) referred to a 23:59 UTC end-time (7:59 local time).

Available information in terms of gauging reporting periods are generally very limited and not well-published (Beck et al., 2019). For example, limited information exists in terms of whether the precipitation data in GHCN-Daily cover the period midnight to midnight or – similar to the BoM stations – terminate during the day. Conversely, for some countries, more than one reporting period may be applicable for its sites (e.g. in the Netherlands, UTC+0 for automatic weather stations vs UTC+8 for manual stations; Eden et al., 2018). Because the cut-off time is unknown, the toolbox generally assumes a midnight break unless a nationwide reporting time was determined (Table 5), potentially misaligning the track and rainfall data by more than half a day.

Not all regions are equally represented in GHCN-Daily – and data provision may not include the full operational period of a weather gauge. In addition, station data do not always provide accurate measurements and monitoring of extreme weather events. Often, monitoring stations may become inundated, damaged or relocated due to the extreme conditions, resulting in minimal to no data being recorded. This may have been the case for TC Iota (2020) because another extreme event (TC Eta) impacted the same area two weeks earlier, producing widespread and severe infrastructure damage. An underestimation of the severity, damage and insurance costs for such events (immediately following a major disaster) is also possible.

Among planned future enhancements for the *rainfall_tracker* toolbox are the expansion of tested data formats for precipitation (i.e. file structures in addition to rectilinear netCDF and pre-processed GHCN-Daily .dly files). Further, providing an option to permit variable, user-defined radii during the course of a cyclone track is also considered.

Although the toolbox was specifically developed for TCs, the code potentially has a wide range of applications. The script could also be employed to select other types of point or gridded data based on a chosen distance from a mobile object. For example, Villarini et al. (2014) selected stream gauges within 500 km from a storm centre for their TC-related flood analyses.

6. Conclusions

Three rainfall datasets were evaluated for differences in estimated rainfall accumulation: 1) the ERA5 global reanalysis, 2) the global GHCN-Daily station dataset and 3) the regional SILO database. Six storm episodes (including four single and two tandem TC events) – encompassing all major, tropical basins – illustrated distinct strengths and flaws in individual datasets. ERA5 benefits from its global data coverage and high temporal (hourly) resolution. However, its relatively coarse spatial scale is not conducive for capturing the highly heterogeneous nature of precipitation in some areas. This is evident in the estimated event rainfall for TC Oswald, with localised much higher totals approximated by SILO. Because of its regional status, SILO is nevertheless limited to Australian settings. Further, its non-uniform reference time (cut-off at 9 a.m. local time) complicates the determination of the correct rain day with respect to datasets that are most commonly in UTC, including TC tracks. GHCN-Daily is also affected by this feature. Moreover, information about the applicable 24-hour precipitation period for individual countries or data sources are challenging to obtain, potentially misaligning rainfall and track data by more than half a day. Individual gauges can also experience extensive data gaps, potentially hampering some types of analyses. Despite all these caveats, GHCN-Daily still offers a valuable source for ground-truthing weather and climate features in gridded reanalysis datasets. These drawbacks highlight the importance of developing tools – like the one (*rainfall_tracker*) presented in this article – for enhancing future TC research and mitigating the increasingly devastating impacts of cyclonic storms.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The *rainfall_tracker* toolbox is accessible through the GitHub (https://github.com/jjaffres/rainfall_tracker) file exchange.

Acknowledgements

The complete GHCN-Daily database was accessed via https://www.ncei.noaa.gov/pub/data/ghcn/daily/. Gridded ERA5 total precipitation data were obtained from https://cds.climate.copernicus.eu. The regional, gridded SILO precipitation data were downloaded via https://www.longpaddock.qld.gov.au. IBTrACS track data were derived from https://www.ncei.noaa.gov. The BoM best track TC data were retrieved from http://www.bom.gov.au. This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envsoft.2023.105773.

References

Bagtasa, G., 2017. Contribution of tropical cyclones to rainfall in the Philippines. J. Clim. 30 (10), 3621–3633.
Beck, H.E., Wood, E.F., Pan, M., Fisher, C.K., Miralles, D.G., van Dijk, A.I.J.M., McVicar, T.R., Adler, R.F., 2019. MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment. Bull. Am. Meteorol. Soc. 100 (3), 473–500.
Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Soci, C., Villaume, S., Bidlot, J.-R., Haimberger, L., Woollen, J., Buontempo, C., Thépaut, J.-N., 2021. The ERA5 global reanalysis: preliminary extension to 1950. Q. J. R. Meteorol. Soc. 147 (741), 4186–4227.
Brackins, J.T., Kalyanapu, A.J., 2020. Evaluation of parametric precipitation models in reproducing tropical cyclone rainfall patterns. J. Hydrol. 580, 124255.
Breña-Naranjo, J.A., Pedrozo-Acuña, A., Pozos-Estrada, O., Jiménez-López, S.A., López-López, M.R., 2015. The contribution of tropical cyclones to rainfall in Mexico. Phys. Chem. Earth, Parts A/B/C 83–84, 111–122.
Chen, S.S., Knaff, J.A., Marks, F.D., 2006. Effects of vertical wind shear and storm motion on tropical cyclone rainfall asymmetries deduced from TRMM. Mon. Weather Rev. 134 (11), 3190–3208.
Corbosiero, K.L., Dickinson, M.J., Bosart, L.F., 2009. The contribution of eastern North Pacific tropical cyclones to the rainfall climatology of the southwest United States. Mon. Weather Rev. 137 (8), 2415–2435.
Dare, R.A., 2013. Seasonal tropical cyclone rain volumes over Australia. J. Clim. 26 (16), 5958–5964.
Dare, R.A., Davidson, N.E., McBride, J.L., 2012. Tropical cyclone contribution to rainfall over Australia. Mon. Weather Rev. 140 (11), 3606–3619.
DeGaetano, A.T., 2000. A serially complete simulated observation time metadata file for U.S. daily Historical Climatology Network stations. Bull. Am. Meteorol. Soc. 81 (1), 49–68.
Deng, D., Ritchie, E.A., 2020. Rainfall mechanisms for one of the wettest tropical cyclones on record in Australia—Oswald (2013). Mon. Weather Rev. 148 (6), 2503–2525.
Deo, A., Chand, S.S., Ramsay, H., Holbrook, N.J., McGree, S., Magee, A., Bell, S., Titimaea, M., Haruhiru, A., Malsale, P., Mulitalo, S., Daphne, A., Prakash, B., Vainikolo, V., Koshiba, S., 2021. Tropical cyclone contribution to extreme rainfall over southwest Pacific Island nations. Clim. Dynam. 56 (11), 3967–3993.
Eden, J.M., Kew, S.F., Bellprat, O., Lenderink, G., Manola, I., Omrani, H., van Oldenborgh, G.J., 2018. Extreme precipitation in The Netherlands: an event attribution case study. Weather Clim. Extrem. 21, 90–101.
Emerton, R., Cloke, H., Ficchi, A., Hawker, L., de Wit, S., Speight, L., Prudhomme, C., Rundell, P., West, R., Neal, J., Cuna, J., Harrigan, S., Titley, H., Magnusson, L., Pappenberger, F., Klingaman, N., Stephens, E., 2020. Emergency flood bulletins for Cyclones Idai and Kenneth: a critical evaluation of the use of global flood forecasts for international humanitarian preparedness and response. Int. J. Disaster Risk Reduc. 50, 101811.
Fujiwhara, S., 1922. On the growth and decay of vortical systems and the mechanism of extratropical cyclones. Jpn. J. Astron. Geophys. 1, 125–182.
Gray, J.L., 2023. Assessing the spatial and temporal characteristics of cyclone tracks along the east coast of Australia. Central Queensland University, Townsville, Australia. PhD thesis.
Guzman, O., Jiang, H., 2021. Global increase in tropical cyclone rain rate. Nat. Commun. 12 (1), 5344.
Hassler, B., Lauer, A., 2021. Comparison of reanalysis and observational precipitation datasets including ERA5 and WFDE5. Atmosphere 12 (11), 1462.
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., Thépaut, J.-N., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049.
Howard, E., Washington, R., Hodges, K.I., 2019. Tropical lows in southern Africa: tracks, rainfall contributions, and the role of ENSO. J. Geophys. Res. Atmos. 124 (21), 11009–11032.
Hu, P., Zhang, Q., Shi, P., Chen, B., Fang, J., 2018. Flood-induced mortality across the globe: spatiotemporal pattern and influencing factors. Sci. Total Environ. 643, 171–182.
Hunt, K.M.R., Fletcher, J.K., 2019. The relationship between Indian monsoon rainfall and low-pressure systems. Clim. Dynam. 53 (3), 1859–1871.
Jaffrés, J.B.D., 2019. GHCN-Daily: a treasure trove of climate data awaiting discovery. Comput. Geosci. 122, 35–44.
Jarraud, M., 2008. Guide to Meteorological Instruments and Methods of Observation (WMO-No. 8). World Meteorological Organisation, Geneva, Switzerland, p. 681.
Jeffrey, S.J., Carter, J.O., Moodie, K.B., Beswick, A.R., 2001. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environ. Model. Software 16 (4), 309–330.
Jiang, H., Halverson, J.B., Simpson, J., Zipser, E.J., 2008. Hurricane "rainfall potential" derived from satellite observations aids overland rainfall prediction. J. Appl. Meteorol. Climatol. 47 (4), 944–959.
Jiang, H., Zipser, E.J., 2010. Contribution of tropical cyclones to the global precipitation from eight seasons of TRMM data: regional, seasonal, and interannual variations. J. Clim. 23 (6), 1526–1543.

Khouakhi, A., Villarini, G., Vecchi, G.A., 2017. Contribution of tropical cyclones to rainfall at the global scale. J. Clim. 30 (1), 359–372.
Knapp, K.R., Diamond, H.J., Kossin, J.P., Kruk, M.C., Schreck, C.J.I., 2018. International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4.
Knapp, K.R., Kruk, M.C., Levinson, D.H., Diamond, H.J., Neumann, C.J., 2010. The International Best Track Archive for Climate Stewardship (IBTrACS): unifying tropical cyclone best track data. Bull. Am. Meteorol. Soc. 91 (3), 363–376.
Knight, D.B., Davis, R.E., 2007. Climatology of tropical cyclone rainfall in the southeastern United States. Phys. Geogr. 28 (2), 126–147.
Ko, M.-C., Marks, F.D., Alaka, G.J., Gopalakrishnan, S.G., 2020. Evaluation of hurricane Harvey (2017) rainfall in deterministic and probabilistic HWRF forecasts. Atmosphere 11 (6), 666.
Kurniawan, R., Harsa, H., Nurrahmat, M.H., Sasmito, A., Florida, N., Makmur, E.E.S., Swarinoto, Y.S., Habibie, M.N., Hutapea, T.F., Sudewi, R.S., Fitria, W., Praja, A.S., Adrianita, F., 2021. The Impact of Tropical Cyclone Seroja to the Rainfall and Sea Wave Height in East Nusa Tenggara, the 3rd International Conference on Maritime Sciences and Advanced Technology. IOP Conference Series: Earth and Environmental Science, Pangandaran, West Java, Indonesia, 012049.
Leroux, M.-D., Nguyen-Hankinson, M.C., Davidson, N.E., Callaghan, J., Tory, K., Wain, A., Huang, X., 2020. Environmental interactions during the extreme rain event associated with ex-tropical cyclone Oswald (2013). Journal of Southern Hemisphere Earth Systems Science 69 (1), 216–238.
Lin, C.Y., Hsu, H.m., Sheng, Y.F., Kuo, C.H., Liou, Y.A., 2011. Mesoscale processes for super heavy rainfall of typhoon Morakot (2009) over southern taiwan. Atmos. Chem. Phys. 11 (1), 345–361.
Menne, M.J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R.S., Gleason, B.E., Houston, T.G., 2012a. Global Historical Climatology Network - Daily (GHCN-Daily) (v3.28-upd-2021091218). NOAA National Climatic Data Center. https://doi.org/10.7289/V5D21VHZ.
Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G., 2012b. An overview of the global historical climatology network-daily database. J. Atmos. Ocean. Technol. 29 (7), 897–910.
Ng, B., Walsh, K., Lavender, S., 2015. The contribution of tropical cyclones to rainfall in northwest Australia. Int. J. Climatol. 35 (10), 2689–2697.
Nicholls, N., Drosdowsky, W., Lavery, B., 1997. Australian rainfall variability and change. Weather 52 (3), 66–72.
Prat, O.P., Nelson, B.R., 2013. Mapping the world's tropical cyclone rainfall contribution over land using the TRMM Multi-satellite Precipitation Analysis. Water Resour. Res. 49 (11), 7236–7254.
Rappaport, E.N., 2000. Loss of life in the United States associated with recent Atlantic tropical cyclones. Bull. Am. Meteorol. Soc. 81 (9), 2065–2074.
Rappaport, E.N., 2014. Fatalities in the United States from Atlantic tropical cyclones: new data and interpretation. Bull. Am. Meteorol. Soc. 95 (3), 341–346.
Stansfield, A.M., Reed, K.A., Zarzycki, C.M., Ullrich, P.A., Chavas, D.R., 2020. Assessing tropical cyclones' contribution to precipitation over the eastern United States and sensitivity to the variable-resolution domain extent. J. Hydrometeorol. 21 (7), 1425–1445.
Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.-L., 2018. A review of global precipitation data sets: data sources, estimation, and intercomparisons. Rev. Geophys. 56 (1), 79–107.
Thorne, P.W., Vose, R.S., 2010. Reanalyses suitable for characterizing long-term trends. Bull. Am. Meteorol. Soc. 91 (3), 353–362.
Touma, D., Stevenson, S., Camargo, S.J., Horton, D.E., Diffenbaugh, N.S., 2019. Variations in the intensity and spatial extent of tropical cyclone precipitation. Geophys. Res. Lett. 46 (23), 13992–14002.
Tuleya, R.E., DeMaria, M., Kuligowski, R.J., 2007. Evaluation of GFDL and simple statistical model rainfall forecasts for US landfalling tropical storms. Weather Forecast. 22 (1), 56–70.
Ullrich, P.A., Zarzycki, C.M., McClenny, E.E., Pinheiro, M.C., Stansfield, A.M., Reed, K.A., 2021. TempestExtremes v2.1: a community framework for feature detection, tracking, and analysis in large datasets. Geosci. Model Dev. (GMD) 14 (8), 5023–5048.
Villarini, G., Goska, R., Smith, J.A., Vecchi, G.A., 2014. North Atlantic tropical cyclones and U.S. flooding. Bull. Am. Meteorol. Soc. 95 (9), 1381–1388.
Wu, L., Liang, J., Wu, C.-C., 2011. Monsoonal influence on typhoon Morakot (2009). Part I: observational analysis. J. Atmos. Sci. 68 (10), 2208–2221.
Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., Kitoh, A., 2012. APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bull. Am. Meteorol. Soc. 93 (9), 1401–1415.
Ying, M., Zhang, W., Yu, H., Lu, X., Feng, J., Fan, Y., Zhu, Y., Chen, D., 2014. An overview of the China Meteorological Administration tropical cyclone database. J. Atmos. Ocean. Technol. 31 (2), 287–301.
Zhou, Y., Matyas, C.J., 2018. Spatial characteristics of rain fields associated with tropical cyclones landfalling over the western Gulf of Mexico and Caribbean Sea. J. Appl. Meteorol. Climatol. 57 (8), 1711–1727.