https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Hurricanes that haven't happened, yet: Identifying unprecedented tropical cyclone scenarios

Dorothy Heinrich[1,2], Erin Coughlan de Perez[2,3], Leanne Archer[4], Nadia Bloemendaal[5], Kevin I. Hodges[1], Helen Hooker[1], Theodore G. Shepherd[1], Nathan Sparks[6], Ralf Toumi[6], Elisabeth Stephens[1,2]

[1]Department of Meteorology, University of Reading, UK. [2]Red Cross Red Crescent Climate Centre, Netherlands. [3]Feinstein International Center, Friedman School of Nutrition Science and Policy, Tufts University, USA. [4] University of Bristol,UK. [5] Institute for Environmental Studies, Vrije Universiteit Amsterdam and Royal Netherlands Meteorological Institute, Netherlands. [6]Department of Physics, Imperial College London, UK.

Correspondence: Dorothy Heinrich (d.f.heinrich@pgr.reading.ac.uk)

Abstract.
Tropical cyclones (TCs) can be unprecedented in many dimensions and can result in disasters when they are unforeseen. We conduct an intercomparison of four TC databases to identify plausible synthetic events that would exceed observational records, and argue that these provide robust and evidence-based scenarios for disaster management use. We compare datasets produced by two statistical TC track models and a newly published TC track hindcast archive from numerical weather predictions: STORM (n = 712,800), IRIS (n= 472,162), and WATTCH (n= 36,793). For all six TC basins, we explore how each dataset characterises unprecedented extreme events in terms of lifetime maximum intensity, 24h changes in wind speed, monthly frequency of Category 4 and 5 storms, and latitude at first landfall. We assess how each dataset represents the basin-level observational record from IBTrACS by conducting a series of fidelity tests (mean, standard deviation, kurtosis, and skewness) to assess whether their most extreme events could be considered plausible in the current climate. Between 50% and 89% of dataset-basin combinations pass at least 2/4 fidelity tests if we include where models indicate underestimation. From this, we identify several hundreds of plausible simulated TCs that exceed historical records in different ways. Where datasets show good fidelity, we illustrate the potential use of these datasets by extracting unprecedented scenarios such as a Category 5 TC hitting southern Madagascar or a TC making landfall on the city of Xai-Xai in Mozambique south of the country's most southerly landfall. Based on this work, we underscore an opportunity for disaster management practitioners to access unprecedented TC scenarios relevant to their work and that would be both robust and imaginative, going beyond current practice.

1 Introduction

In 1900, the town of Galveston, Texas (USA) was emerging as a booming metropolis – the bay around which it was built thought to be safe from the hurricanes which often brought devastation to the region. At the end of August of that year, residents met news of a hurricane moving towards Galveston Bay first with disbelief, then with panic - no disaster preparedness plans existed for this possibility (Frank, 2003(Frank, 2003). The storm killed between 8,000 and 12,000 people, an estimate with

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

such a wide range because the city was so devastated that most bodies were never found. To this day, the hurricane remains the deadliest natural hazard to have occurred in the USA.

Preparedness for unprecedented tropical cyclones, like the 1900 Galveston hurricane, hinges on an ability to imagine what is possible and an understanding of what is plausible. Between 1970 and 2019, tropical cyclones killed over 779,324 people and caused economic losses of over 1.4 trillion USD globally whilst, in the same period, the population living in tropical risk zones has increased two-fold (WMO, 2021). Some of these events have been called "unprecedented": Hurricane Katrina (2005), super typhoon Haiyan (2013, and tropical cyclone Freddy (2023), to name a few.

To prepare for extreme weather events like tropical cyclones, disaster managers often use hazard scenarios in desk-based or live simulations to strengthen disaster preparedness plans (Mahdi et al., 2023). However, one known limitation within these practices is that the scenarios are often designed based on past events (Jeffries et al., Submitted), thus failing to imagine all the different dimensions in which a future event might be different from the past – unprecedented, for example, in magnitude, duration, or location (Heinrich et al., 2024).

1.1 Literature review

Various methods have been used to assess extreme weather risk at different timescales, utilising observations, forecasts, and climate models. To assess the risk of a frequently occurring hazard like tropical cyclones, the starting point is often the observational record, such as that compiled in NOAA's International Best Track Archive for Climate Stewardship (IBTrACS). However, whilst IBTrACS goes as far back as recording a storm in the North Atlantic in 1851, there are limits to data coverage, differences and inconsistencies in reporting between different ocean basins, and changes in operational practices (Meiler et al., 2022; Schreck et al., 2014). The reliability brought by satellite imagery means that the 1980s are often benchmarked as the decade when IBTrACS become more complete, and windspeeds and pressure recorded more reliably across all basins (Gahtan et al., 2024; Knapp et al., 2010). For instance, Hodges et al. (2017) use the period from 1979 onward for their analysis. All in all, the relatively short and heterogeneous observational record of tropical cyclones poses limits to observation-based risk analyses, particularly for low-probability events or for locations which experience infrequent storms. This can lead to a lack of accurate understanding of risks of higher magnitude and impactful events in different places (Bloemendaal et al., 2020).

Two main types of approaches can help bridge these gaps. The first looks towards the past to lengthen the record. For instance, coastal sediment analysis can provide insights into paleo tropical cyclones (a field of study called "paleo tempestology") and enable a reconstruction of past tropical cyclone activity (Bloemendaal et al., 2020). Oral and written history can also provide evidence of past records held in museums, national archives, or other places where memory of past disasters is conserved, visualised, and transmitted (Ballard et al., 2020; Frew & White, 2020; Sloan, 2008). For example, compelling research has turned to newspaper archives to uncover past tropical cyclones that had previously been overlooked or forgotten (Chenoweth & Landsea, 2004; Msemo et al., 2022; Sobel et al., 2019).

Looking to the past has its limitations, however. It can be both expensive to process this data and difficult to extrapolate beyond local scales (Bloemendaal et al., 2020). Therefore, a second approach consists of modelling synthetic tropical cyclones in different ways. In this approach, tracks and intensities are derived through statistical resampling and simulations based on an underlying observational or climate model dataset and on knowledge of the physical processes which drive or constrain the plausibility of extreme values (Bloemendaal et al., 2020; Vickery et al., 2000).

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

This synthetic modelling creates coherent tropical cyclones with physically consistent characteristics to the original dataset although can retain biases from the training data (Bloemendaal et al., 2020; Vickery et al., 2000). Resampling can be run repeatedly, creating large synthetic records which would generally include lower-probability and higher magnitude events than exist in the observational record. The models used for this can be either coupled statistical-dynamical (e.g. (Carozza et al., 2024) or fully statistical (e.g. (Bloemendaal et al., 2020; Sparks & Toumi, 2024); with some run at global scales whilst others built specifically for regional coverage (e.g. RAFT from Xu et al., 2023).

In their seminal article, Tales of Future Weather, Hazeleger et al.(2015) advocate for a combination of climate models and weather models to analyse risks of impactful extreme weather. In recent years, ensemble forecasts from Numerical Weather Prediction (NWP) models have emerged as state-of-the-art for tropical cyclone prediction at short lead-times (Emerton et al., 2024; Hooker et al., 2023; Titley et al., 2020). From these ensemble forecasts, it is possible to assign probabilities to a range of forecast outcomes (Emerton et al., 2024). While the usual usage of ensemble forecasts is to predict the weather, the range of realisations produced by the ensemble has also been shown to have value in identifying risks of plausible but as yet unprecedented extreme events (Kelder et al., 2022; Thompson et al., 2017). The use of ensemble forecasts to extend observational datasets and estimate extreme value statistics was notably pioneered through the Unprecedented Simulated Extremes using Ensembles (UnSEEn) protocol (Thompson et al., 2017). Since then, the approach has been used in numerous studies (Berghald et al., 2024; Coughlan de Perez et al., 2025; Fischer et al., 2023; Kay et al., 2024; Kelder et al., 2020; Thompson et al., 2023). However, it has not yet directly be adapted for tropical cyclones. Bourdin et al (2025) similarly make use of tracks from ECMWF hindcasts in their framework to select analogue TC events. Other studies have taken counterfactual approaches to modelling plausible tropical cyclones using ensemble reforecasts, such as (Philp et al., 2022) who derive downward counterfactual hurricanes from an initial-condition ensemble reforecast for the North Atlantic, and Rye & Bond (2022) who model counterfactuals for major hurricanes as an exercise for the insurance sector.

All these approaches mainly provide tropical cyclone tracks and intensities in the current climate, although they do allow for future climate simulations. Climate models can generate tropical cyclone-like vortices and be used to understand expected future changes to tropical cyclone characteristics as a result of anthropogenic global warming (Roberts et al., 2020). Research has shown limits in the resolution of traditional Global Climate Models (GCMs) to adequately represent convective processes and, therefore, tropical cyclone intensity and frequency in the current, let alone the future, climate (Davis, 2018; Russotto et al., 2022). Increasingly however, convection-permitting models are being run at high-enough resolution to resolve sea-air interactions and therefore used to estimate both extreme rainfall and the pressure minima of tropical cyclones (Buonomo et al., 2024; M. Lee et al., 2023), although there remain questions about their ability to model the most intense storms (Baker et al., 2024). These types of approaches have been used to estimate both tropical cyclone-driven flooding projections and projections of changing tropical cyclone intensity, including rapid intensification (Baker et al., 2024), and have been argued critical to understanding emerging trends of tropical cyclone risk in a changing climate (Archer et al., 2024; Baker et al., 2024).

Tropical cyclone modelling can be coupled with other physical models to derive fluvial flooding (Brunner & Slater, 2022) and storm surge (Benito et al., 2025) cascading from the storms. Coupled approaches, such as those which combine meteorological and hydrological models, can provide a comprehensive picture of the risk posed by tropical cyclones in a changing climate (Hooker et al., 2026). Due to the computational power required to run these simulations at a global scale, however, (Archer et al., 2024; Bloemendaal et al., 2020; Sparks & Toumi, 2024), simulations are often restricted to repurposing those used for operational forecasting (Benito et al., 2025; Brunner & Slater, 2022) or running models at local scales (Archer et al., 2024; Hooker et al., 2026).

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

1.2 Aim and Research Questions

We believe the plethora of existing tropical cyclone modelling approaches present untapped potential to explore scenarios of unprecedented events. Calls have been made to combine multiple lines of evidence and methods for studying the risks of unprecedented extremes (Kelder et al., 2025), and intercomparisons have been conducted to compare tropical cyclone risk within these different datasets (Bourdin et al., 2022; Meiler et al., 2022). However, these intercomparisons have not addressed how different tropical cyclone (TC) datasets specifically characterise unprecedented extremes. In this study, we conduct a multi-method investigation into existing statistical and physical-based tropical cyclone models, specifically setting out to answer three research questions:

1. What are some existing TC datasets that can be repurposed to derive different types of unprecedented tropical cyclone scenarios?
2. How well do these datasets represent the observational record for different dimensions of unprecedented tropical cyclones, and are their most extreme scenarios plausible?
3. What do these datasets reveal about plausible risks of unprecedented tropical cyclones in different ocean basins?

From these questions, we illustrate the wealth of opportunity in a selection of already existing open-source TC datasets – each with their own advantages and limitations – and explore how these can be used to extract robust and imaginative scenarios for disaster management, with the aim to enhance disaster preparedness.

2 Methods

Lin and Emanuel (2016) call "grey swans" the high-impact tropical cyclones that "would not be predicted based on history but may be foreseeable." The 1900 Galveston hurricane is one of the starker examples of a grey swan storm, one that hit an unprecedented location, but there are many other ways that a storm can be "unprecedented", all which would have different implications for disaster risk management (Heinrich et al., 2024). A growing amount of tropical cyclone research has been conducted on the question of unprecedented or record-breaking tropical cyclone risks, often analysing one or several observed events that were unusual. A cursory search of the literature found over a thousand articles that explored "unprecedented", "record-breaking" or "unseen" tropical cyclones. Within this group of literature, a large amount of work focuses on unprecedented rainfall (e.g. Menemenlis et al., 2024; Paerl et al., 2019; Tam et al., 2025) or windspeeds (e.g., (Aberson et al., 2006; Biggerstaff et al., 2022) Some studies focus on risks of unprecedented numbers of storms in a season (e.g., Ciullo et al., 2021; Schmitt et al., 2025), and others look at the shifting location of tropical cyclone tracks beyond usual regions (e.g., (Tao et al., 2021). Recently, many studies have delved into the question of rapid intensification (e.g. Radfar et al., 2024; Wang et al., 2025).

As an analytical framework, we adapt a co-produced typology of unprecedented weather developed in Heinrich et al. (2024) which argues that scientists and disaster practitioners must look beyond magnitude and intensity when developing scenarios of unprecedented weather events. The typology categorizes key dimensions in which extreme weather events can be unprecedented: in degree (magnitude and intensity), time (duration, timing, frequency, speed), space (location, extent, and pattern), and compound events (including in contexts of heightened vulnerability). We apply this typology to define what unprecedented could mean for a tropical cyclone (Table 1).

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

DIMENSION	RAINFALL	TRACK AND WIND	STORM SURGE
Magnitude and intensity	Unprecedented rainfall amount	Unprecedented lifetime maximum intensity (LMI)	Unprecedented surge height
Duration	Unprecedented period of rainfall	Unprecedented amount of time active (e.g. from cyclogenesis to decay)	Unprecedented amount of time above normal tide levels
Timing	Unprecedented time of the year (i.e. earlier or later than record)		
Frequency		Unprecedented number of tropical cyclones of a given intensity	
Speed	Unprecedented rate of horizontal movement	Unprecedented 24h intensification	Unprecedented speed of onset of storm surge
Location		Unprecedented landfall location	Unprecedented storm surge location
Extent	Unprecedented rainfall and flooding footprint	Unprecedented radius of maximum windspeeds	
Pattern	Shortest amount of time between storm and other rainfall events		

Table 1. Examples of unprecedented dimensions of tropical cyclones

2.1 Dataset Selection

We select five different open-source datasets which span a breadth of the existing tropical cyclone track modelling approaches described above. Four are global: the observed record as represented by IBTrACS; two datasets from fully-statistical synthetic approaches, STORM (Bloemendaal et al., 2020) and IRIS (Sparks & Toumi, 2024); and a NWP dataset called WATTCH, derived from hindcast runs from ECMWF (Hooker et al., 2026). Datasets from coupled approaches combining track, rainfall, and flood modelling were not available globally, so we illustrate these approaches through a TC flood event catalogue for Puerto Rico developed by Archer et al. (2024).

A summary of these five datasets can be found in Table 2 below.

Type of approach	**Observational recording**	**Statistical approaches**		**Physical-based counterfactual approaches**	
Dataset	**IBTrACS**	**STORM**	**IRIS**	**WATTCH**	**Puerto Rico TC flood event catalogue**
Dataset size (n)	3,601 events (Cat 1 and above)	712,800 events	472,162 events	36,793 events (matched to 1,437 observed events)	4,909 events

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Timescales covered	Historical (1851-present)	Atemporal (representative of 1980-2017 climate conditions)	Atemporal (representative of 1980-2023 climate conditions)	Near-present day (2003-2025)	Near-present day (2005-2016) and future (~2100)
Spatial scales covered	Global	Global	Global	Global	National (Puerto Rico)
Variables	Over 150 from different regional centres. Variables used for this analysis include: Basin Date and Time SID Windspeed Latitude Longitude Landfall code	Year Month tropical cyclone number Timestep Basin ID Latitude and Longitude Minimum pressure Maximum wind Radius of max wind Saffir Simpson Category Landfall code Distance to land	Unique ID Year tropical cyclone number Month Timestep Latitude and Longitude Maximum wind Minimum pressure Radius of max wind Radius of gale force winds	Unique ID Ensemble member Latitude and Longitude of track Date and Time Forecast initialisation date Minimum pressure Windspeed Timestep Lead time to landfall Basin Overland code Rapid intensification codes Stalling codes	Year Latitude and Longitude Rainfall amounts Flood maps Estimated population exposure
Dataset access	Gahtan et al. (2024) ; Knapp et al. (2010)	Bloemendaal et al. (2020)	Sparks and Toumi ((2024)	Hooker et al. (2026)	Archer et al. (2024)

Table 2. Description of the datasets used to characterize unprecedented tropical cyclone risks in this analysis

2.1.1 Observational record: IBTrACS dataset

For observational data, we use the International Best Track Archive for Climate Stewardship (IBTrACS) dataset version v04r01 accessed last on August 15th, 2025 (Gahtan et al., 2024; Knapp et al., 2010). This is publicly available from the NCEI website and includes a technical document with a description of over 150 variables, the methodology for assessing these, caveats and limitation, and other pieces of valuable information. IBTrACS is updated three times a week with provisional data, with the final quality-controlled data only available post-season.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

2.1.2 Statistical approaches: STORM and IRIS datasets

As statistical approaches, we use the datasets from the Synthetic Tropical cyclOne geneRation Model (STORM) and Imperial College Storm (IRIS) model, two fully statistical synthetic models of tropical cyclone tracks and windspeeds both published open source within the last five years.

STORM is a statistical algorithm developed by Bloemendaal et al. (2020) which uses Markov chains to generate storm tracks and stimulate tropical cyclone intensity. The track and intensity of a tropical cyclone is statistically resampled and modelled from tropical cyclones in the IBTrACS observational dataset (from 1980–2017) creating an event with similar characteristics, and then rerun many times to obtain a larger dataset which includes higher return period events than are present in the short historical record (Bloemendaal et al., 2020). The end-product is a dataset of 10,000 years of tropical cyclones' track (latitude and longitude), intensity (windspeeds and pressure), and size. The stated aim of the dataset is to serve tropical cyclone risk assessments of various kinds. More details about the methodology can be found in a comprehensive Data Descriptor in Bloemendaal et al. (2020).

IRIS also uses IBTrACS tropical cyclones (from 1980 to 2021) as the underlying dataset from which to statistically simulate 10,000 years of tropical cyclone tracks and intensities (Sparks and Toumi, 2024). A key difference with STORM, however, is that IRIS starts the simulation after the lifetime maximum intensity (LMI) of the storm (i.e. not the whole lifecycle of the storm), which enables the model to focus on the "most hazardous and critical, post-LMI, stage of the tropical cyclone, including landfall" (Sparks & Toumi, 2024). The model has been applied for attribution and climate change projections (Sparks & Toumi, 2025a, 2025b)). The variables available in the open-source dataset include track location and intensity as both windspeeds and pressure. Again, a detailed methodology can be found in a comprehensive Data Descriptor published in Sparks and Toumi (2024).

2.1.3 Physical-based approaches: WATTCH dataset and TC flood event catalogue

Numerical Weather Prediction (NWP) archive: WATTCH dataset

The European Centre for Medium-Range Weather Forecasts (ECMWF) produces ensemble hindcasts (reforecasts) using the same IFS model configuration as real-time forecasts but initialised on past dates. The hindcast system comprises 11 ensemble members (1 control plus 10 perturbations) run twice weekly at 9 km resolution with 6-hourly output out to 15 days over a rolling 20-year period (ECMWF, 2016; Vitart et al., 2022).

The World Archive of Tracked Tropical Cyclone Hindcasts (WATTCH) contains tropical cyclone tracks and intensities from these IFS hindcast ensemble members (cycle 49r1), matched to observational records (Hooker et al., 2026). tropical cyclones are identified using the TRACK algorithm (Hodges, 2017), which detects systems based on spatially filtered, vertically averaged relative vorticity (850–600 hPa) that persist for at least two days. Intensity is characterised by maximum 10-m wind speed and minimum mean sea level pressure.

To establish a verification baseline, ERA5 reanalysis is tracked first using the same algorithm, then matched to IBTrACS best-track observations. This approach extends trajectories beyond the observational record, capturing early genesis and late decay stages often absent from IBTrACS. Finally, hindcast ensemble tracks are matched to these verified ERA5-IBTrACS events using a mean spatial separation threshold of 4° (Emerton et al., 2024; K. I. Hodges & Emerton, 2015). This dataset was processed during the research phase of this work and is now available at Hooker et al., (2026).

Coupled approach: Puerto Rico TC flood event catalogue

To illustrate opportunities provided by coupled approaches, we explore a TC flood event catalogue for Puerto Rico developed by Archer et al. (2024) comprised of 4,909 synthetic events generated by Vosper et al. (2020). This uses a subset of four global

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

climate model ensemble members from the half a degree additional warming, projections, prognosis and impacts (HAPPI) ensemble (Mitchell et al., 2017), under a recent-past (2005-2016), as well as a 1.5° warmer and 2° warmer world compared to pre-industrial levels. All storm tracks and their windspeeds are then run through a physics-based rainfall model (Vosper et al., 2020) which are then simulated spatiotemporally (~10km, two-hourly) in the LISFLOOD-FP hydrodynamic flood model to generate 4,909 corresponding flood footprint maps. From this, estimates of the total number of people exposed to current and future flooding are calculated from each of these events using World Pop data (Archer et al., 2024). As a set, the catalogue thus connects an event in one climate scenario with wind and track information, as well as a corresponding rainfall footprint, flood map, and a population exposure estimate.

2.2 Workflow

2.2.1 Data processing

To process these datasets, we first define a spatial and temporal boundary in which we are interested. For the purposes of this paper, we take the spatial boundary as the six tropical cyclone basins defined in IBTrACS (Figure 2). We take as a temporal boundary the whole IBTrACS record since the 1870s – we acknowledge gaps in 19th century reporting but believe that shortening the record to the 1960s as often done in modelling would limit unnecessarily our analysis as we are interested in as long a record as possible, whether or not it may be missing some events. For comparability, we subset all datasets events that reach Tropical Cyclone windspeeds (Category 1 on the Saffir Simpson scale). For each of these datasets, we follow a workflow illustrated in Figure 1. AI-based tools were used to assist with code-debugging.

Figure 1. Generalised workflow for data analysis. We first pre-processed and prepared each dataset, then summarized and applied bootstrapping with replacement. We then computed fidelity tests on these results and generate boxplots of variables for each dataset, identifying events beyond the observational record. We ended with extracting illustrative unprecedented TC scenarios from datasets which pass the fidelity tests for the basin.

2.2.2 Plausibility testing

Data preprocessing is needed to ensure that all variables are in comparable units, that all events have unique identifiers, and derive key variables such as landfall codes.

Since the sizes of the model datasets are orders of magnitude larger than those of the observations, any robust statistical comparisons regarding extreme values require adequate resampling techniques. For this, we apply non-parametric bootstrapping to estimate the distributions of different variables in the model datasets and compare the distributions with those of the observations. To do so, we count the number of tropical cyclones recorded in IBTrACS by basin over the whole period, which determines the sample size (n) for the bootstrapping. For each variable, we perform the bootstrap with replacement with 1,000 iterations. Each bootstrap sample contains (n) modelled events which provided us with comparable samples with which to compare the distribution of variables (Table 3). Note that for all frequency metrics, we bootstrap years instead of individual events.

BASIN	IBTrACS event count (n)
Eastern Pacific (EP)	586
North Atlantic (NAT)	971
North Indian (NI)	97
South Indian (SI)	408
South Pacific (SP)	234
Western Pacific (WP)	1,327

Table 3. Number of observed tropical cyclones in IBTrACS, by basin. This corresponds to the basin's bootstrap sample size.

We then conduct a series of fidelity tests on these bootstrapped replicate samples, comparing the kurtosis, mean, standard deviation, and skewness of the values within each sample to identify where the characteristic of the observations falls within the 99% CI of the bootstrapped model distribution characteristics. These tests replicate the methodology of the UNSEEN protocol first developed in Thompson et al. (2017) to evaluate whether models represent observations with enough accuracy for their outliers (or unprecedented events compared to the observational record) to be trusted as plausible. By 'plausible' we mean, as Thompson et al. (2017) to put it, to "ensure the behaviour in the model tails is indistinguishable from the observations". As such, we check whether the four statistics fall within a realistic range. If the statistics are uncharacteristically high compared to observations, then the models are producing too many unprecedented events and may be overestimating their likelihood.

2.2.3 Distribution visualization and extraction of unprecedented event- scenario

We visualise these distributions with side-by-side boxplots, comparing the observed record from IBTrACS observations to the datasets from our models by basin. Each boxplot includes the maximum value of the corresponding IBTrACS observation, allowing a visualization of unprecedented events (i.e. above the record) in the model datasets. Selecting a dataset with good fidelity test results, we then extract one event-scenario that would be unprecedented for a particular basin. We illustrate this event with a map of its track and a short description.

2.3 Scale of analysis

All analysis is run at ocean basin scales as defined in Figure 2, but we acknowledge that future researchers and practitioners may want to replicate a similar process for more local scales as well. For comparability, we note that IRIS, STORM,

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

and WATTCH all the following 6 basins: EP, NA, NI, SI, SP, WP), omitting the South Atlantic basin used by IBTrACS, and therefore homogenize in this same way.

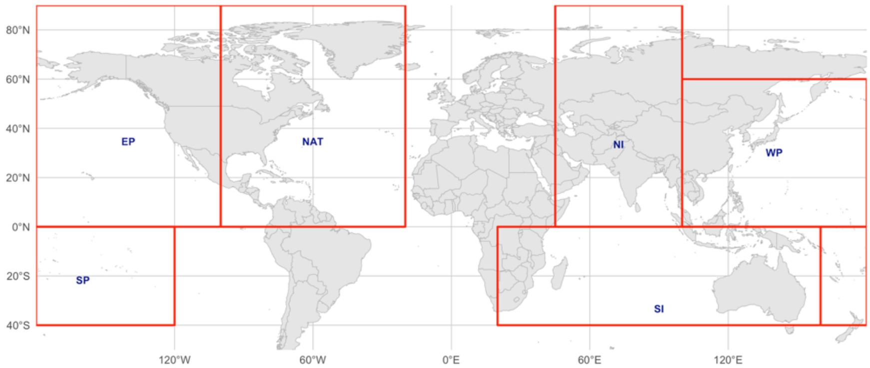


Figure 2. Basin bounding boxes defined for this paper based on IBTrACS v04r0. Note that the Southern Pacific (SP) spans across the date line hence is shown as two sub-basins on this figure.

3 Results

3.1 Which existing TC datasets can be repurposed to explore unprecedented tropical cyclone scenarios?

As they currently are, all four datasets can be used to uncover some but not all types of tropical cyclone scenarios that would be unprecedented. Certain gaps are fundamental to the way the datasets were created, but there are also incidental which could be filled by additional modelling. Table 4 shows a summary of this review and could be used as a reference to identify which dataset (in its current form) may contain each type of unprecedented tropical cyclone scenario.

DIMENSION	RAINFALL	TRACK AND WIND	STORM SURGE
Magnitude and intensity	Rainfall amount TC flood event catalogue* **Sect. 3.2.5.**	Lifetime maximum intensity IRIS* STORM* WATTCH* **Sect 3.2.1.**	Surge height
Duration	Period of rainfall TC flood event catalogue*	Amount of time active, e.g. from cyclogenesis to decay	Amount of time above normal tide levels
Timing	Time of the year (i.e. earlier or later than record) **Figure 3**		
Frequency		Number of events IRIS* STORM* WATTCH* **Sect 3.2.2.**	
Speed	Rate of horizontal movement	24h change in wind speed IRIS* STORM* WATTCH* **Sect. 3.2.3.**	Speed of onset of storm surge
Location		Landfall location IRIS* STORM* WATTCH* **Sect 3.2.4.**	Storm surge location
Extent	Rainfall and flooding footprint TC flood event catalogue* **Sect 3.2.5**	Radius of maximum windspeeds IRIS* STORM* WATTCH*	

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Pattern	Amount of time between storm and other rainfall events		

*Table 4. Summary table. What dimensions of unprecedented tropical cyclones can these different datasets (in their current form) help uncover? Each * implies that this dimension can be found in IRIS (*), STORM (*) WATTCH (*), and TC flood event catalogue (*). In **bold**, we give the paper sections in which these dimensions are explored.*

3.1.1 Statistical approaches: STORM and IRIS datasets

The datasets derived from STORM and IRIS provide tropical cyclone variables related to the location and timestep of the track, maximum wind speeds, and minimum pressure. Both datasets can be used to derive speed and location of tropical cyclone tracks. For both, the timing, duration, and frequency dimensions are constrained by the model parameters and training data. Additionally, both models have a variable for the radius of maximum winds which could be used to derive a measure of tropical cyclone extent – however, the corresponding variable in IBTrACS is inconsistent between basins and therefore, at a global level, we cannot compare the models with observations. Finally, since both STORM and IRIS currently only model tracks and windspeeds, in their current form, they cannot be used to look at any dimensions involving rainfall or storm surge.

3.1.2 Physical-based approaches: WATTCH dataset and TC flood event catalogue

Numerical Weather Prediction (NWP) archive: WATTCH dataset

Tropical cyclone ensemble forecasts can provide a window into plausible alternatives of real-lived storms. Whilst a particular tropical cyclone only occurs once in the real world, the ensemble members from any given forecast run can be taken as a counterfactual of that storm. WATTCH, the NWP dataset we explore here, provides information about track, windspeeds, and pressure, but currently has no associated rainfall. As such, it can be used to uncover counterfactual events that would have been different from the observed one in track location and windspeeds. The duration of the track is limited by the current length of the simulation (15 days) and, as the current dataset is matched only with storms that materialised, there are no storms outside seasons, or seasons with above-record tropical cyclone activity (Figure 3).

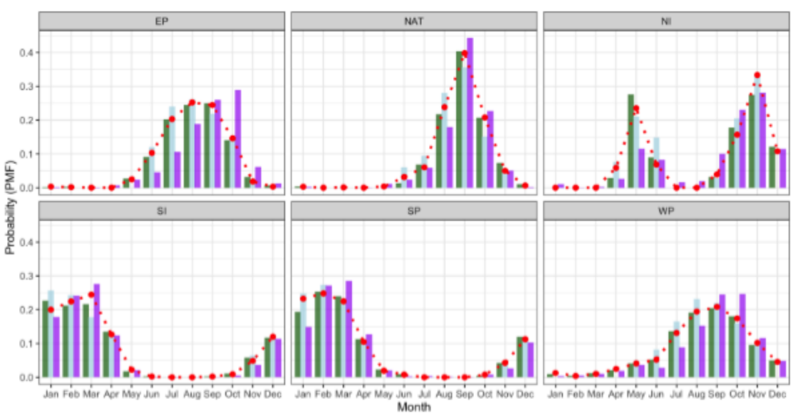


Figure 3. Probability Density Function representing the frequency of tropical cyclones, by month and by basin, in IRIS (blue), STORM (green), WATTCH (purple) and IBTrACS (red dotted). This shows how the observed seasonality of tropical cyclones is replicated by design in all our datasets and therefore there are no events outside of observed seasons.

Puerto Rico TC flood event catalogue

Climate-perturbed flood inundation models can provide a window into tropical cyclone risks in a changing climate. These methods can help us answer the question of what storms might be plausible, currently and in the future, under different climate

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

change trajectories. To illustrate this approach, the Puerto Rico TC flood event catalogue is quite comprehensive at a local scale. It can be used to obtain more detailed cascading hazard scenarios that include not only windspeeds but also rainfall, flooding, and even be extended to impacts.

3.2 What can we learn from these datasets about plausible scenarios of unprecedented tropical cyclones?

To assess the plausibility of the unprecedented tropical cyclone scenarios that can be extracted from each of these datasets, we compare the statistical distribution from the real-world observational record by ocean-basin and model. Fidelity tests can help answer whether we could confidently extract unprecedented scenarios from these models for these basins or whether we would urge caution. When the IBTrACS statistics (i.e. real world) fall within the 99% two-sided confidence interval (CI) of the model for this basin, it suggests that the model reproduces well the statistics of the real world, and therefore we would trust the plausibility of outliers in the model. When the observed value does not fall within the 99% CI of the model, this does not mean that we can accept the alternative hypothesis (i.e. that the outliers are implausible). There may be many explanations for this, among which:

- If the IBTrACS statistic exceeds the upper limit of the 99CI of the model, the model could be *under-imaginative*. In other words, extreme values produced by the model may very well be plausible but do not represent the full space of possibilities. We may not find events that are very different from the limits already experienced in the real world.

If the IBTrACS statistic is below the lower limit of the 99CI of the model, the model could be *over-imaginative.* In other words, the model may be producing values that are too extreme and could be implausible in the current climate. In this case, this would give us a red flag for the purposes of sampling outliers from the models. Specifically, 'over-imaginativeness' of the mean could be interpreted as the model producing values that are too high. 'Over-imaginativeness' of the standard deviation, skewness, and kurtosis would indicate the presence of too many outliers, or tails that may be implausibly heavy

3.2.1 Unprecedented intensity

Table 5 shows the results of fidelity tests for lifetime maximum intensity (LMI). It identifies datasets and basins where we would exercise more caution in extracting unprecedented scenarios (i.e. where the model may be over-imaginative) or indicates where the dataset may not have many unprecedented events.

Dataset	Statistic	EP	NAT	NI	SI	SP	WP
IRIS	Mean	[89, 94.5] (94.3) *	[93.1, 97.5] (92.3)	[95.7, 112.6] (85.6)	[94.1, 100.9] (86.1)	[97, 106.5] (85.4)	[95.1, 99.8] (82.4)
IRIS	SD	[23.8, 26.7] (22.7)	[25.1, 27.4] (21.4)	[23.7, 31.3] (25) *	[24.5, 28.2] (21.6)	[25, 29.8] (21.4)	[24.4, 26.6] (18.2)
IRIS	Kurtosis	[1.9, 2.6] (2.4) *	[1.9, 2.2] (3.2) ^	[1.6, 2.5] (2) *	[1.8, 2.4] (2.3) *	[1.7, 2.5] (2.9) ^	[1.8, 2.1] (2.3) ^
IRIS	Skewness	[0.4, 0.7] (0.5) *	[0.3, 0.6] (0.9) ^	[-0.3, 0.6] (0.1) *	[0.2, 0.6] (-0.1)	[0, 0.5] (0.1) *	[0.2, 0.4] (0.1)
STORM	Mean	[78.6, 84.6] (94.3) ^	[74.3, 78.9] (92.3) ^	[64.9, 83.4] (85.6) ^	[71.8, 77.7] (86.1) ^	[70, 77.6] (85.4) ^	[80.8, 85.6] (82.4) *
STORM	SD	[24.8, 28] (22.7)	[29.9, 33.3] (21.4)	[23.8, 39.2] (25) *	[20.6, 24.7] (21.6) *	[19, 24.8] (21.4) *	[25.5, 27.8] (18.2)
STORM	Kurtosis	[2.1, 2.8] (2.4) *	[2.3, 3] (3.2) ^	[2, 7.1] (2)	[2.3, 4] (2.3) *	[2.5, 5.3] (2.9) *	[1.7, 2] (2.3) ^
STORM	Skewness	[0.3, 0.7] (0.5) *	[0.7, 1] (0.9) *	[0.5, 1.9] (0.1)	[0.5, 1] (-0.1)	[0.5, 1.3] (0.1))	[0.1, 0.4] (0.1)
WATTCH	Mean	[56.9, 62] (94.3) ^	[78.4, 82.4] (92.3) ^	[47.4, 68.2] (85.6) ^	[78.3, 83.1] (86.1) ^	[71.4, 78.1] (85.4) ^	[84.4, 88.8] (82.4)
WATTCH	SD	[22.5, 25.4] (22.7) *	[24.3, 27.2] (21.4)	[29.7, 38] (25)	[16.3, 20.7] (21.6) ^	[16.3, 21.6] (21.4) *	[20.5, 24.2] (18.2)
WATTCH	Kurtosis	[2, 2.7] (2.4) *	[2.7, 3.2] (3.2) *	[1.4, 2.8] (2) *	[3.2, 5.6] (2.3)	[2.8, 4.3] (2.9) *	[3.4, 4.3] (2.3)
WATTCH	Skewness	[0.1, 0.5] (0.5) ^	[-0.5, -0.2] (0.9) ^	[-0.3, 0.7] (0.1) *	[-0.3, 0.8] (-0.1) *	[-0.5, 0.3] (0.1) *	[-0.4, 0.1] (0.1) *

Table 5. LMI fidelity test results by dataset and basin [range of the bootstrapped sample values from the models] and (the corresponding IBTrACS statistic). Dark green (*) indicates where IBTrACS statistic falls within the 99 CI of the models. Light

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

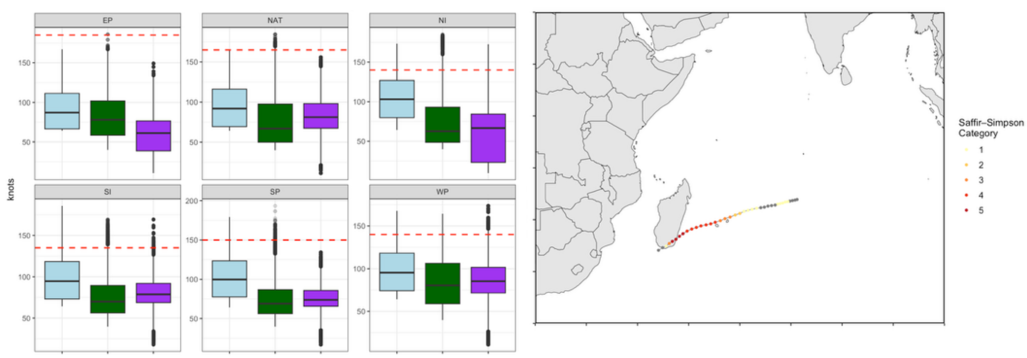
green (^) indicates where the model may be under-imaginative, not containing many unprecedented events. Red flags ( ) indicate where over-imaginativeness of the models would suggest caution for the purpose of extracting unprecedented scenarios.

- In the Eastern Pacific basin, WATTCH passes all tests. IRIS and STORM pass three tests but overestimate standard deviation.
- In the North Atlantic, IRIS overestimates mean and standard deviation. STORM and WATTCH pass three tests, but overestimate standard deviation.
- In the North Indian ocean, IRIS overestimates mean values. We also have red flags for STORM which overestimate skewness/kurtosis metrics. WATTCH underestimates mean values but overestimates standard deviation.
- In the South Indian ocean, IRIS overestimates mean values and skewness. STORM is slightly skewed, and WATTCH has high kurtosis.
- In the South Pacific, STORM and WATTCH pass all the tests whilst IRIS overestimates the mean and standard deviation.
- In the Western Pacific, we have a lot of red flags. We see overestimation of the mean, standard deviation, and skewness for IRIS, overestimated standard deviation and skewness in STORM and red flags for WATTCH for all tests apart from skewness.

Table 5 shows indication that some models may be over or under-imaginative with regards to extreme values. With these tests completed, we can identify plausible events in the models that exceed the basin-level observational LMI record (Figure 4).

- In the Eastern Pacific, none of the datasets contain unprecedented storms.
- In the North Atlantic, North Indian, and South Indian all datasets have unprecedented events, but we would exercise caution in extracting the most extreme ones because the models may be over-imaginative.
- In the South Pacific, we would trust the plausibility of the unprecedented intensities from STORM and WATTCH but would be cautious in extracting extreme events from IRIS.
- In the Western Pacific, we would offer caution in extracting the more extreme scenarios from all the datasets.

To illustrate a use case, Figure 4b extracts an example of an unprecedented scenario from STORM in the South Indian basin that we would consider plausible according to the fidelity test results. In this scenario, a tropical cyclone of Category 5 would make landfall in Southern Madagascar with an LMI of 143.22 knots, exceeding the observed record for the basin of 135 knots (Monica, 2006).



https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Figure 4 (a) Distribution of maximum windspeeds (LMI) as 1-minute windspeeds in knots, by asin in bootstrapped samples from IRIS (blue), STORM (green), WATTCH (purple), with the observed maximum value from IBTrACS for that basin (red dashed line).(b) Example event from STORM exceeding the observed LMI record for the South Indian basin (ID: 13750).

3.2.2 Unprecedented frequency

As a metric of frequency, we assess the distribution of the monthly occurrence of high intensity (Category 4 and 5) tropical cyclones by model and basin (Table 6). As detailed in 3.1, only STORM and IRIS can be used to extract frequency metrics. We note first that, because the frequency of high intensity storms is quite low, the mean and standard deviation metrics are likely more relevant measures.

Dataset	Statistic	EP	NAT	NI	SI	SP	WP
IRIS	Mean	[0.17, 0.27] (1.36) ^	[0.15, 0.23] (1.2) ^	[0.05, 0.18] (1) ^	[0.16, 0.31] (1.15) ^	[0.1, 0.26] (1.14) ^	[0.3, 0.51] (1.06) ^
IRIS	SD	[0.42, 0.59] (0.64) ^	[0.43, 0.58] (0.45) *	[0.22, 0.46] (0)	[0.42, 0.66] (0.44) *	[0.32, 0.59] (0.36) *	[0.6, 0.87] (0.25)
IRIS	Kurtosis	[3.34, 17.25] (2.39)	[6.54, 23.94] (4.08)	[2.04, 36.8] (NA)	[2.71, 20.54] (7.93) *	[2.56, 26.73] (1.69)	[1.21, 23.69] (9.71) *
IRIS	Skewness	[2.03, 3.47] (1.7)	[2.57, 4.05] (2.18)	[2, 5.32] (NA)	[1.9, 3.85] (2.89) *	[1.85, 4.42] (1.9) *	[1.4, 3.72] (3.38) *
STORM	Mean	[1.49, 2.04] (1.36)	[1.29, 1.7] (1.2)	[0.12, 0.35] (1) ^	[0.57, 0.92] (1.15) ^	[0.31, 0.69] (1.14) ^	[2.8, 4.26] (1.06)
STORM	SD	[2.37, 2.93] (0.64)	[2.62, 3.17] (0.45)	[0.37, 0.81] (0)	[1.05, 1.56] (0.44)	[0.75, 1.36] (0.36)	[4.1, 5.23] (0.25)
STORM	Kurtosis	[0.55, 4.02] (2.39) *	[3.22, 6.98] (4.08) *	[3.1, 29.33] (NA)	[1.35, 11.72] (7.93) *	[2.27, 15.48] (1.69)	[-0.64, 1.96] (9.71) ^
STORM	Skewness	[1.29, 1.92] (1.7) *	[1.98, 2.57] (2.18) *	[1.99, 4.85] (NA)	[1.51, 2.9] (2.89) *	[1.78, 3.55] (1.9) *	[0.84, 1.59] (3.38) ^

Table 6. As in Table 5 but for annual frequency of Category 4 and 5 storms.

- In the Eastern Pacific, IRIS underestimates the frequency of high intensity storms compared to the observations whilst STORM shows over-estimation for the mean and standard deviation.
- In the North Atlantic, IRIS shows good fidelity with the observed mean and standard deviation whilst STORM shows overestimation for these statistics.
- In the North Indian basin, high intensity tropical cyclones are very rare (maximum one event per month) and therefore it is not possible to conclude much about the plausibility of unprecedented monthly frequency.
- In the South Indian, IRIS passes all the tests and STORM has some overestimation flagged in its standard deviation.
- In the Western Pacific, IRIS shows slight underestimation of the mean and overestimation in the standard deviation. We have high red flags for the mean and standard deviation of STORM which overestimates frequency.

Figure 5a shows the distribution of the monthly number of Category 4-5 storms by basin and models. This helps us identify plausible months in which high intensity tropical cyclone frequency would be higher than observed.

- In the Eastern Pacific, North Atlantic, South Indian, and South Pacific basin, we would trust the plausibility of record-breaking months from IRIS but would exercise caution in those from STORM.
- For the North Indian, we would offer caution in extracting any unprecedented tropical cyclone months from the models because of limits in the observational record.
- In the Western Pacific, unprecedented frequency months in STORM and IRIS should offer pause as both models may be overimaginative, although the latter slightly less.

To illustrate a use case, Figure 5b. shows a month in IRIS where the Eastern Pacific would have a higher number of high-intensity storms than previously recorded. In this scenario, five tropical cyclones would reach Category 4 and 5 windspeeds,

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

exceeding the latest record of September/October 2023 where the region saw four Category 4 and 5 tropical cyclones (Hurricanes Jova, Lidia, Norma, and Otis).

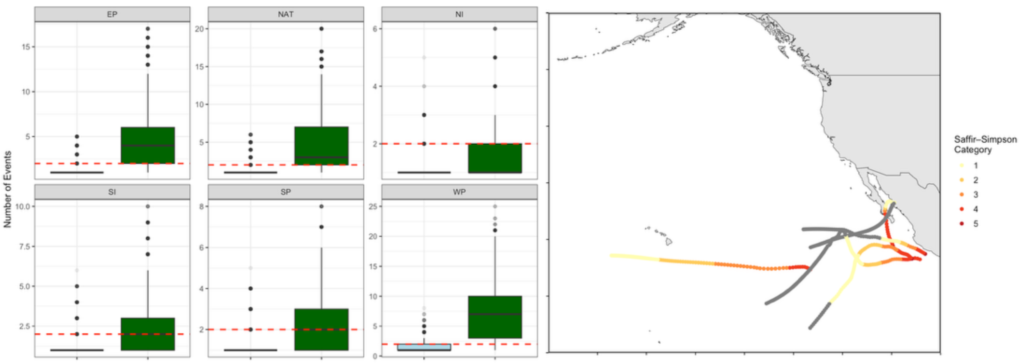


Figure 5. (a) Distribution of the monthly counts of tropical cyclones reaching Category 4 and 5 winds, by basin from IRIS (blue), STORM (green), WATTCH (purple), with the observed maximum number observed in IBTrACS for that basin (red dashed line). (b) Example year from IRIS exceeding the IBTrACS monthly frequency of Cat 4-5 record for Eastern Pacific basin ('Year': 306695).

3.2.3 Unprecedented intensification

To explore storms of unprecedented intensification, we compare 24h change in wind speeds from the models with IBTrACS observations and test whether IBTrACS values fall within 99% CI of the models. Results can be found in Table 7.

Dataset	Statistic	EP	NAT	NI	SI	SP	WP
IRIS	Mean	[27.2, 31.3] (26.8)	[34.6, 38.4] (24.3)	[43.1, 58] (34.5)	[29.5, 34.7] (31.6) *	[28.7, 35.3] (30.3) *	[34.2, 36.8] (34.5) *
IRIS	SD	[15.8, 20] (19.9) *	[19.6, 22.8] (16.5)	[20.7, 29.3] (22.2) *	[16.9, 22.1] (17.2) *	[15.7, 22.2] (18.5) *	[16.8, 18.8] (19.8) ^
IRIS	Kurtosis	[3.9, 7.1] (2.5)	[3.3, 4.7] (3.8) *	[1.9, 4.1] (2) *	[3.6, 7.4] (2.3)	[3.2, 8.9] (2.3)	[3.1, 4.2] (2.7)
IRIS	Skewness	[1, 1.7] (0.3)	[0.9, 1.3] (0.9)	[0.1, 1.1] (0.2) *	[1, 1.7] (0)	[0.8, 1.9] (0.2)	[0.7, 1] (0.4)
STORM	Mean	[32.6, 36.3] (26.8)	[33.3, 37.2] (24.3)	[35.8, 54.8] (34.5)^	[28.3, 32.6] (31.6) *	[25.4, 30.6] (30.3) *	[32.1, 34.5] (34.5) *
STORM	SD	[16.4, 18.7] (19.9) ^	[20.7, 24.1] (16.5)	[22, 37.1] (22.2) *	[15.5, 18.9] (17.2) *	[13.1, 17.5] (18.5) ^	[16.3, 18.2] (19.8) ^
STORM	Kurtosis	[2.1, 3.5] (2.5) *	[3, 5] (3.8) *	[1.9, 5.6] (2) *	[2.4, 6.1] (2.3)	[2.3, 7.5] (2.3) *	[2.8, 3.8] (2.7)
STORM	Skewness	[0, 0.4] (0.3) *	[0.6, 1.1] (0.9) *	[0.1, 1.5] (0.2) *	[0.3, 1.1] (0)	[0.1, 1.3] (0.2) *	[0.3, 0.7] (0.4) *
WATTCH	Mean	[16.6, 19.2] (26.8) ^	[25.4, 27.4] (24.3)	[16.4, 25.7] (34.5) ^	[26, 29] (31.6) ^	[21.8, 25.7] (30.3) ^	[26.8, 28.4] (34.5) ^
WATTCH	SD	[12, 14.2] (19.9) ^	[11.7, 13.5] (16.5) ^	[13.2, 19.8] (22.2) ^	[10.9, 13.6] (17.2) ^	[10.5, 13.5] (18.5) ^	[12, 13.4] (19.8) ^
WATTCH	Kurtosis	[2.3, 6.3] (2.5) *	[2.8, 8] (3.8) *	[1.6, 10.2] (2) *	[2.7, 6.9] (2.3)	[2.3, 5.2] (2.3)	[2.8, 4.3] (2.7)
WATTCH	Skewness	[0.4, 1.2] (0.3)	[0.2, 1.1] (0.9) *	[0, 1.8] (0.2) *	[0.2, 1.2] (0)	[0, 0.9] (0.2) *	[0.2, 0.6] (0.4) *

Table 7. As in Table 5 but for 24h change in wind speed.

- In the Eastern Pacific basin, WATTCH passes all the fidelity tests apart from a slight overestimation of skewness. STORM overestimates the mean. IRIS shows many red flags.
- For the North Atlantic, IRIS, STORM, and WATTCH show red flags for their means.
- For the North Indian, we have a red flag in IRIS which overestimates mean values.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

- In the South Indian, all three datasets have a red flag for an overestimation of their skewness and kurtosis.
- For the South Pacific, IRIS and WATTCH overestimate skewness and kurtosis, but all other tests are passed.
- For the Western Pacific, all three datasets overestimated kurtosis, but the other tests were passed.

In Figure 6a, we present the distribution of 24h change in wind speeds by basin and models, with events above the observed IBTrACS value considered unprecedented.

- In the Eastern Pacific, we would suggest caution in extracting unprecedented intensification scenarios from IRIS and STORM. Although we would have greater confidence in WATTCH values, none are above the observational record.
- In the North Atlantic, there are many unprecedented scenarios in all three datasets but would be cautious in using any of these because of red flags in their mean test.
- For the North Indian, we could extract unprecedented events that are available from WATTCH and STORM.
- In the South Indian, we would exercise some caution in extracting storms in the tail ends of the distribution because of the red flag in skewness for all three datasets. IRIS has the smallest overestimation of the three datasets.
- In the South Pacific, we would feel confident in the plausibility of unprecedented scenarios from STORM. WATTCH has no events above the observational record for this dimension.
- In the Western Pacific, we would have confidence in extracting rapidly intensifying tropical cyclone scenarios from all datasets but offer caution in the ones that appear most extreme because of kurtosis overestimation.

To illustrate a use case, Figure 6b presents a storm from WATTCH that exceeds the record 24h intensification observed in the South Indian ocean. In this scenario, a track hindcast for tropical cyclone Kenneth (2019), a tropical cyclone would intensify from a tropical storm to a Category 5 tropical cyclone less than 100 km off the southern coast of Tanzania.

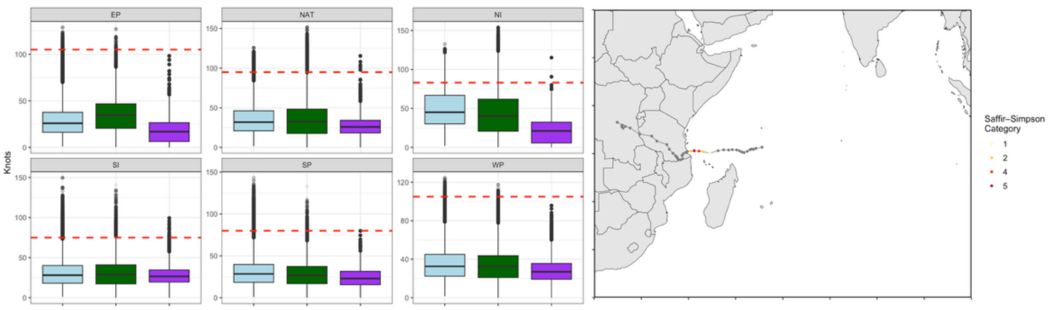


Figure 6. (a) Distribution of 24h change in wind speeds by basin from IRIS (blue), STORM (green), WATTCH (purple), with the observed maximum from IBTrACS for that basin (red dashed line). (b) Example event from WATTCH exceeding the IBTrACS record for the South Indian basin (ID: 2019_tr0022_2019042100_10)

3.2.4 Unprecedented location

The location of tracks taken by tropical cyclones are critical drivers of their impacts, and places that have never experienced a tropical cyclone landfall may be underprepared. Figure 7 shows the locations of observed first landfalls from IBTrACS. In addition, the characteristics of the basins and their coastlines matter for landfall location – for example, the track of a tropical cyclone moving across the South Indian ocean will encounter less coastline than one moving across the Eastern Pacific, and the South Indian tropical cyclone are also more likely to have multiple landfalls as they often move across Madagascar and into the Mozambique channel. As such, there are many legitimate ways to define an unprecedented location for a tropical cyclone.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Figure 7. World map showing location of observed first landfalls in IBTrACS (red)

For purposes of choosing one, we looked here at latitude at first landfall, looking specifically for poleward landfalls. Table 8 reports the fidelity test results for latitudes at first landfall for the 99% CI of the models and IBTrACS observation, by basin.

For this dimension, although all metrics are helpful to assess fidelity, we are particularly interested in skewness: large deviations would lead us to be cautious in overinterpreting the most poleward landfall locations. We expect basins in the Southern Hemisphere to have negative signed skewness (i.e. landfalls south of the Equator) and basins in the Northern Hemisphere to have positive signed skewness (i.e. landfalls north of the Equator). Where the real-world skewness values are lower than the model range, this indicates that the models are producing too many poleward storms and gives us a red flag in interpreting extreme poleward landfalls.

Dataset	Statistic	EP	NAT	NI	SI	SP	WP
IRIS	Mean	[22.3, 23.1] (15.5)	[25.6, 26.8] (23.3)	[17.4, 19.8] (14.3)	[-20.9, -19.7] (-15.3)	[-20.9, -18.5] (-15.7)	[22.7, 23.9] (17.2)
IRIS	SD	[3.9, 4.7] (3.1)	[7.3, 8.2] (7.4) *	[3.4, 5] (3.9) *	[4.1, 5.1] (3.6)	[6.1, 8.3] (3.6)	[7.9, 8.5] (5.7)
IRIS	Kurtosis	[3.4, 5.7] (12.2)^	[2.8, 3.3] (2.3) 🚩	[1.9, 4.6] (3) *	[3.7, 5.2] (3.9) *	[3.5, 6.7] (3.4) 🚩	[2, 2.5] (3.5)^
IRIS	Skewness	[-0.3, 0.6] (1.7) 🚩	[0.2, 0.6] (0.3) *	[-1.4, -0.4] (-0.1) 🚩	[-0.9, -0.1] (-0.6) *	[-1.7, -1] (-0.6) 🚩	[0.2, 0.4] (0.5) ^
STORM	Mean	[22.2, 23.6] (15.5)	[23.8, 25.4] (23.3)	[14.5, 17.2] (14.3)	[-20.2, -19.2] (-15.3)	[-22.6, -20] (-15.7) ^	[17.8, 18.9] (17.2)
STORM	SD	[5.5, 7.8] (3.1)	[9.2, 10.4] (7.4)	[4.5, 5.7] (3.9)	[3.5, 4.7] (3.6) *	[6.3, 8.9] (3.6)	[7.2, 8.1] (5.7)
STORM	Kurtosis	[10.4, 16.2] (12.2) *	[3, 3.7] (2.3) 🚩	[1.7, 2.4] (3)^	[3.3, 13.4] (3.9) *	[3.5, 7.2] (3.4) 🚩	[3.3, 5] (3.5) *
STORM	Skewness	[1.8, 2.9] (1.7)	[0.6, 0.8] (0.3) 🚩	[-0.4, 0.3] (-0.1) *	[-1.8, 0.3] (-0.6) *	[-2, -1.2] (-0.6) ^	[0.9, 1.3] (0.5) 🚩
WATTCH	Mean	[39.8, 43.5] (15.5)	[28.1, 30.2] (23.3)	[16.8, 22.6] (14.3)	[-17.7, -16.4] (-15.3)	[-26.8, -23.1] (-15.7)	[25, 26.5] (17.2)
WATTCH	SD	[17.1, 18] (3.1)	[12.3, 13.8] (7.4)	[6.3, 15.5] (3.9)	[4.8, 6] (3.6)	[9.9, 11.1] (3.6)	[10.1, 11.2] (5.7)
WATTCH	Kurtosis	[1.2, 1.3] (12.2)^	[2.8, 3.8] (2.3) 🚩	[5.2, 24.3] (3) 🚩	[3.4, 5.2] (3.9) *	[1.3, 1.9] (3.4) ^	[2.8, 3.5] (3.5) *
WATTCH	Skewness	[-0.2, 0.2] (1.7) 🚩	[0.7, 1] (0.3) 🚩	[1.7, 3.7] (-0.1) 🚩	[-1, -0.2] (-0.6) *	[-0.6, 0] (-0.6) *	[0.3, 0.7] (0.5) *

Table 8. As in Table 5 but for latitude at first landfall.

- In the Eastern Pacific, STORM seems to best capture the skewness of IBTrACS with values nearest to the 99CI, but both IRIS and WATTCH show poleward skewness.
- In the North Atlantic, IRIS shows good fidelity with observed latitudes, but STORM and IRIS show more poleward skewness.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

- In the North Indian, IRIS shows more poleward skewness than observations. STORM and WATTCH have a range that crosses the equator which gives us a red flag.
- In the South Indian, observations fall within the confidence interval of all three model datasets.
- In the South Pacific, IRIS and STORM have more poleward skewness than observations, but WATTCH shows good fidelity with observations.
- In the Western Pacific, observations fall within the confidence interval for WATTCH - IRIS has smaller skewness than observations. STORM seems to overestimate poleward skewness.

All the approaches we review contain landfall locations that would be unprecedented in the observed record. Figure 8a shows the distribution of latitude at first landfall by basin and model, with poleward landfall records recorded in IBTrACS for that basin. As seen, all datasets show a high number of outliers, and a large amount of these events would make landfall more north or south than experienced in the past.

- In the Eastern Pacific basin, we could trust that the most poleward landfalls from STORM would be plausible but would offer caution in extracting events from IRIS and WATTCH and would not trust the plausibility of those which make landfall south of the Equator.
- In the North Atlantic, we could extract unprecedented poleward landfalls from IRIS rather than from STORM and WATTCH. In particular, the outlier events in WATTCH that make landfall north of the equator may be implausible.
- For the North Indian, we would have confidence in using poleward landfalls from STORM and WATTCH but would question the plausibility of the poleward storms in IRIS.
- In the South Indian, we would trust the plausibility of poleward landfalls from all three datasets.
- In the South Pacific and Western Pacific, WATTCH produces landfalls we would consider plausible whilst we would question the more extreme events STORM. In the Western Pacific, however, IRIS shows less poleward skewness than observations and so, for the purposes of extracting poleward storms, IRIS could be used but noting that these events may not be very different from past ones.

In Figure 8b, we extract an event from the STORM dataset that would make landfall south of the southernmost recorded landfall in Mozambique. In this scenario, a tropical cyclone would make landfall near Xai-Xai, the 10th largest city by population in the country.

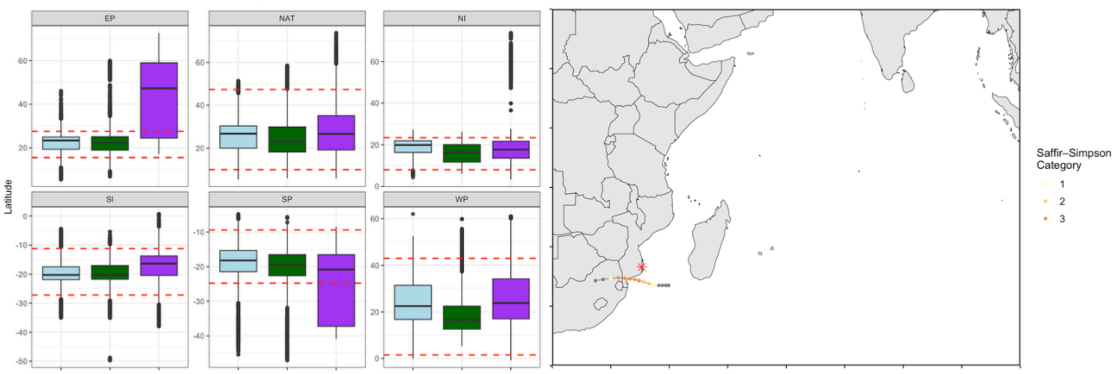


Figure 8. (a) Distribution of latitudes at first landfall by basin from IRIS (blue), STORM (green), WATTCH (purple), with the observed southernmost and northernmost first landfalls from IBTrACS for that basin (red dashed line). (b) Example event from STORM making landfall south of the southernmost landfall recorded in Mozambique (ID: 56882)

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

3.2.5 Unprecedented rainfall intensity and flood extent

A key limitation in using STORM, IRIS, and WATTCH datasets (in their present form) to explore unprecedented tropical cyclones is that they only provide variables related to track, windspeed, and pressure. As detailed in Table 1, tropical cyclones can also be unprecedented by their rainfall and flood extent, and our story would be incomplete without illustrating these.

As an example of a coupled-approach, we present here work done by Archer et al. (2024) who model extreme tropical cyclones counterfactual to Hurricane Maria (2017) which made landfall as a high-end Category 4 tropical cyclone in Puerto Rico in 2017, and had both unprecedented rainfall (Keellings & Hernández Ayala, 2019; Ramos-Scharrón & Arima, 2019) as well as a devastating impact on lives and livelihoods unprecedented in Puerto Rico's living memory (Audi et al., 2018; Pasch et al., 2023)).

Their Puerto Rico TC flood event catalogue contains 4,909 events in a present-day or 2° warmer climate. 264 (5%) of these events would bring unprecedented rainfall (Figure 9).

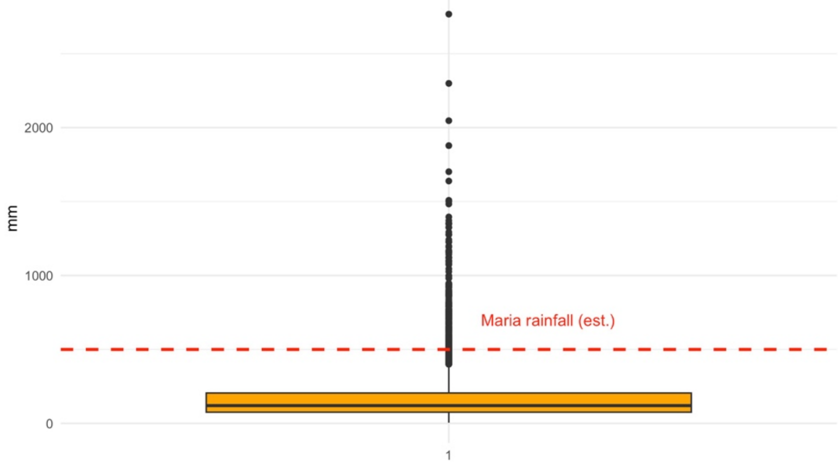


Figure 9. Distribution of total rainfall by event in the Puerto Rico TC flood event catalogue, compared to the total rainfall of Hurricane Maria used to calculate the RMSE of rainfall for each event (dashed red line).

With regards to flooding, Archer et al. (2024) identify 20 plausible events that would cause flood extents worse than Hurricane Maria. These unprecedented tropical cyclone events were identified in the event set by selecting tropical cyclone tracks that had a larger maximum total rainfall than Hurricane Maria, but exhibited similar spatial characteristics i.e. track trajectory and landfall location. To illustrate a use case, Figure 10 shows the flood footprint map of one such event and its corresponding track.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

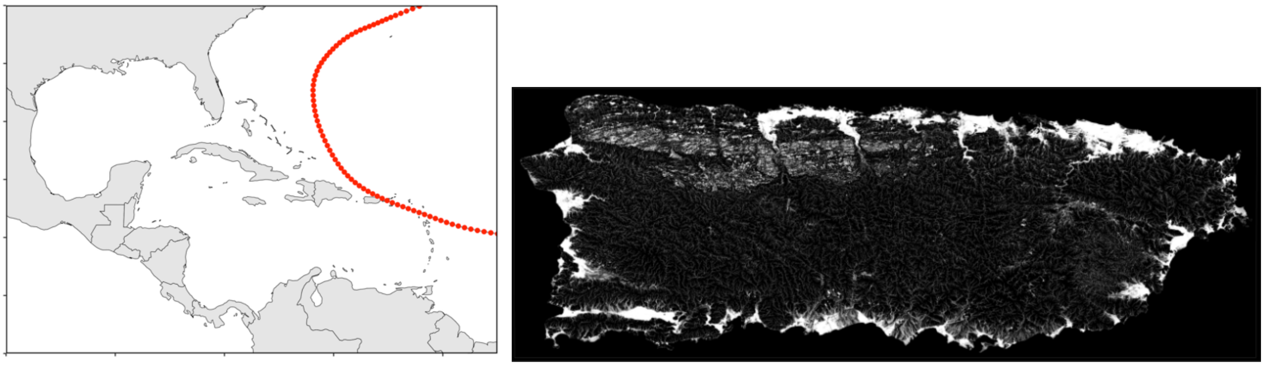


Figure 10. (a) Example scenario of tropical cyclone track and (b) associated flood footprint map from the Puerto Rico TC flood event catalogue (ID: event canam_hist_413).

4 Discussion

To frame this work, we applied a multidimensional definition of "unprecedented" to tropical cyclones. We found that the current literature about record-breaking tropical cyclone risk focuses mainly on the intensity of storms, and relatively less on risks of unprecedented location, seasonality, frequency, and duration. But these other dimensions matter significantly for disaster preparedness: strong cyclones occurring early in the season, like Hurricane Beryl in 2024, can catch disaster managers by surprise before their seasonal preparedness efforts have been conducted; a tropical cyclone of unprecedented duration, like Tropical Cyclone Freddy (2023), requires longer lasting emergency management activities; a warning of a tropical cyclone landfall in a new location, like Galveston (1900) can be confronted by scepticism and a lack of disaster management plans.

Disaster managers can make use of event-scenarios extracted from various datasets derived from different tropical cyclone modelling approaches. However, the added value of these scenarios will depend on how different they are from what has already been experienced and whether they are considered plausible by decision-makers. Our results indicated clear basin-level differences in both fidelity tests and in the range of simulated extreme events in all three datasets - no one dataset always has higher fidelity or range than the others - results range between 1/4 tests passed (e.g. IRIS-LMI- NA basin) to 4/4 tests passed (e.g. STORM-Rate of intensification-NI basin). As such, the choice of appropriate dataset from which to extract salient scenarios should be determined by the ocean basin and the type of scenario of interest.

4.1 Accessing and repurposing datasets: opportunities and gaps

The first aim of our work was to identify existing datasets that could be repurposed to derive unprecedented tropical cyclones scenarios. Through our review of the literature, we identified a plethora of different approaches used to characterize tropical cyclone risk. From these, we selected four open-source peer-reviewed datasets which represented a spectrum of existing approaches. In each dataset, we identified several hundreds of plausible TC scenarios which would be unprecedented in the basin's historical record for LMI, 24h change in windspeed, monthly Cat4 and 5 frequency, and latitude at first landfall. We highlighted the importance of breaking-down the analysis by model and by basin of interest as there are differences in dataset fidelity with observations and in the range of events available.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

As seen in the sections above, these models are run at global scales and contain information about TC track, windspeeds, and pressure. Critically, the absolute limits of their values will grow with sample size, the datasets lengthened if the models rerun. All datasets can be extended with each new tropical cyclone that occurs, creating new plausible hazard scenarios to explore and new runs of the models would produce more simulated storms.

Some gaps are intrinsic to each of the approaches and would require modelling investments to better characterize a fuller range of unprecedented storms. Notably, there are gaps in terms of including multi-risk and cascading hazards related to tropical cyclones. WATTCH, IRIS, and STORM do not currently include estimates of rainfall or event-based information on storm surge. The WATTCH dataset could be enhanced to include additional variables, such as rainfall, or from the river flow data from the hydrological models which are already coupled to the NWP track approach for operational forecasting purposes. The Puerto Rico TC flood event catalogue may be the most comprehensive reviewed in terms of the variables it directly offers (rainfall, flood, and population exposure footprints) which could be more readily used in multi-risk scenario exercises, but it is only run at local scales.

Other gaps are incidental to the datasets themselves and could be addressed through small adjustments to the decisions and methods used to generate each dataset. For example, we found limited ability to compute scenarios of tropical cyclones of unprecedented duration across its lifecycle: IRIS starts modelling at max LMI and stops when the storm reaches landfall, and storms within WATTCH have a simulation length of up to 15 days. Future research could homogenize the recording of time in the observations and adjust the dataset's generation restrictions to include storms before they develop into tropical cyclones, broadening to 'unmatched' storms. Similarly, datasets we explored cannot be used to explore unprecedented seasonality in a meaningful way. The hindcasts in WATTCH are all counterfactuals to observed storms, and IRIS and STORM constrain the months in which they allow tropical cyclones to develop by using observed storms as statistical 'seeds'. Future work on hazard scenarios of unprecedented seasonality could extend WATTCH with ensemble members that did not materialize into tropical cyclone intensities or change the seasonal boundaries of the synthetic models. In relation to tropical cyclone track, post-landfall track can be a critical piece of information for the rainfall and flood damage it may bring - WATTCH contains tracks that go inland, but neither IRIS nor STORM do and therefore they will miss events where a storm may move over land before reaching LMI (for example, like cyclones Chedza in 2015 and Idai in 2021). Continuing to model synthetic tracks once the storms make landfall could be a key development as this could then be coupled with rainfall and flood modelling for a more comprehensive scenario. In the Puerto Rico TC flood event catalogue, windspeeds are available for each event and could be used to generate local multi-risk scenarios.

Our analysis supports a call by Kelder et al. (2025) which advocates for the combination of multiple lines of evidence and methods when studying risks of unprecedented extremes. We would like to re-emphasize this point here – although none of the approaches reviewed here capture all TC dimensions, these datasets together provide many pieces of the puzzle. Across all our datasets, we identified several hundreds of plausible TC scenarios unprecedented in LMI, rate of intensification, frequency of Cat 4/5 storms, landfall location, and even flood extent. Taken as complementary, these datasets provided us a more diverse range of scenarios than any single approach did.

4.2 Assessing plausibility of unprecedented tropical cyclone scenarios

To derive unprecedented tropical cyclone scenarios from each of these datasets, it is not enough to highlight that unprecedented events *can be found* in these datasets: whether these are *plausible* critically matter for their value for decision-making: researchers and practitioners need to know whether these events could happen in the real world although the level of trust in their plausibility required varies depending on the use one wants to make of this. The second aim of this piece was to provide an assessment of the plausibility of record-breaking events for the use in scenario planning by evaluation performance with

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

regards to fidelity tests (mean, standard deviation, kurtosis, and skewness). Our analysis showed that we would trust the plausibility of outlier scenarios in certain model-basin combinations more than others and that no dataset shows consistently closer fidelity than the others. Across our four variables, between 50% and 89% of dataset-basin combinations pass at least 2/4 of fidelity tests if we include where models indicate potential underestimation.

These tests indicate whether decision-makers could trust the plausibility of an unprecedented TC scenario found in these datasets. Our results show that the answer depends on ocean basin and variable of interest. For example, for LMI in the Eastern Pacific, WATTCH would be considered the best performing of the datasets if solely based on fidelity test results (4/4 tests passed) but contains no events above the basin's LMI record from IBTrACS. From LMI in the South Indian, WATTCH also passes all the fidelity tests but it contains many outlier events that exceed the IBTrACS record. And so on.

The full space of possible extreme values presented by each of the datasets is partially determined by how strongly each of the models are conditioned. Indeed, estimates of plausibility and uncertainty will depend on the inputs and assumptions under which each of the models are designed and run. All these approaches condition on the historical record and therefore constrain the plausibility space: for instance, generating statistical distributions from observations as do IRIS and STORM requires selecting variables of interest and therefore reduces the number of variables available to model and increases the uncertainty: the models are trained on observed datasets that are biased and may not reflect the full range of variability. WATTCH uses a physical model and therefore could theoretically provide a fuller set of events but the current version filtered to include only storms which have materialised into observed cyclones. The strongest conditioning is seen in the Puerto Rico coupled approach which brings together track, windspeed, rainfall, and flood modelling approaches. Diving deeper into the reasons behind the statistical tests results falls outside the scope of this piece but a few elements are noteworthy. For example, the poleward skewness of WATTCH may be because it is the only dataset that continues the track into the post-tropical stages. Similarly, STORM does not model extratropical transition, making it possible for events to go further northwards when they would usually transition or dissipate in the real world. More work into the plausibility of unprecedented track and landfall locations in all the models could resemble that conducted by Li and Toumi (2026) using IRIS to look at the plausibility of tropical cyclones forming near the Equator, for example.

The scope of this work is limited to unprecedented tropical cyclones in the current climate. However, given that the climate is changing, so is the boundary of what might be plausible both today and in the future. Exploring this question falls within a growing body of research on shifting TC track locations, changes to intensity, seasonal frequency and much more (e.g. Sparks and Toumi, 2025; Perez-Alarcòn et al. 2023; Chand et al. 2022; Camargo and Wing, 2021) . Where Global Climate Models may be limited in their ability to represent storm intensity accurately, considerable conditioning (e.g. downscaling, bias-correction) is needed (as with the coupled-approach used within this study) to generate plausible unprecedented cyclones from these models. Convection-permitting models may provide a new source of simulated storms, but the computational power required is a challenge for producing the sample sizes required to robustly test plausibility. Harnessing AI approaches could be one emerging way to generate a large number of events (Baño-Medina et al., 2025; Jiménez-Esteve et al., 2025).

4.3 Revealing potential for unprecedented tropical cyclones at regional and national level

When they occur, extreme weather events are often catalysts for adaptation and enhanced systems (S. Lee & Zhao, 2021) but preparing for extreme weather events before they occur is much more difficult. This is, in part, a challenge of imagination – a major barrier to awareness and preparedness for disasters (Ommer et al., 2024) and scenario-planning has been argued one way to address this challenge (Jeffries et al. Submitted).

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

The third aim of this piece was to illustrate how plausible unprecedented scenarios could be extracted from these datasets. We identify several hundreds scenarios that would be unprecedented at an ocean-basin scale but there are likely many more that would be unprecedented at a country-scale or an even smaller administrative unit. These datasets can therefore provide a large and underutilised information source for disaster managers. Research on how these scenarios could be communicated to resonate with disaster management practitioners, linking with existing research on foresight workshops and simulation exercises (Jeffries et al., Submitted), would be highly relevant. Opportunities for co-production between researchers and practitioners (Djenontin & Meadow, 2018; Meadow et al., 2015; Visman et al., 2022) could see researchers designing and communicating scenarios for practitioners to access and use to design and test robust early warning, anticipatory action, and contingency plans.

5 Conclusions

What would happen if southern Madagascar was hit with tropical cyclone winds above 140 knots? How prepared are communities living on Mexico's Pacific coast to a season of five consecutive Category 4 tropical cyclones? Who would warn towns in southern Tanzania of a rapidly intensifying tropical storm? Is the city of Xai-Xai in Mozambique ready for a tropical cyclone making landfall? In Puerto Rico, what would be the impact of unprecedented tropical cyclone-driven flooding?

Research has argued that preparedness to unprecedented events can be enhanced through access to robust and diverse event-scenarios (Jeffries et al., Submitted). Across four variables (LMI, 24h change in windspeeds, monthly frequency of Cat 4/5 storms, and latitude at first landfall), we find variation in fidelity test results across datasets and ocean basins, and no single dataset emerges with consistently better results than others: between 50% and 89% of all dataset-basin combinations pass 2/4 tests. From this, IRIS, STORM, and WATTCH contain several hundreds of plausible TC scenarios that would exceed the IBTrACS record in different dimensions.

For this work, we selected datasets that could be easily accessed by disaster managers tasked with tropical cyclone preparedness. We encourage these practitioners to recognize that tropical cyclones can be unprecedented in diverse dimensions and to use these datasets to design and test disaster preparedness plans for different scenarios. To do so, we recommend that they first select the dataset(s) that passes fidelity tests in their specific basin and their dimension of interest. From this, they can review the hundreds of scenarios identified and select the most salient ones to design and test preparedness plans through simulation exercises.

Addressing unprecedented tropical cyclone risks requires strong connections between scientists and disaster managers. To tropical cyclone researchers, this paper calls for further development of tropical cyclones to capture all their dimensions and cascading risks and for the datasets from tropical cyclone risk modelling to be published open-source and accessible. To disaster management practitioners, this paper hopes to encourage access and use of unprecedented tropical cyclone scenarios that push imagination away from a reliance on the past towards better preparedness for future events.

Code and data availability

All data used for this piece is publicly available. IBTrACS data is available from ncei.noaa.gov/products/international-best-track-archive. All other datasets are available from the cited sources. Code available on request.

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Author contributions

Dorothy Heinrich designed the study and conducted the analysis with supervision from Elisabeth Stephens and Erin Coughlan de Perez. Helen Hooker and Dorothy Heinrich curated the WATTCH dataset with supervision support from Kevin I. Hodges and Elisabeth Stephens. Dorothy Heinrich wrote the manuscript with supervision from Elisabeth Stephens and Erin Coughlan de Perez. All co-authors provided feedback on the analysis and writing.

Competing interests

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.

Disclaimer

"Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher."

Acknowledgements

The authors would like to thank all colleagues who have thought along, proposed feedback, and offered support to this work including Sophie Cuckow, Andrew Tabas, and many others.

Financial support

This work was supported by NASA under award Number 80NSSC25K7503; the REPRESA project which was funded through the Climate Adaptation and Resilience (CLARE) programme, with support from the UK Foreign, Commonwealth & Development Office (FCDO) and the International Development Research Centre (IDRC), Canada, and additionally supported by the Natural Environment Research Council (NERC) via NCAS NC-International funding; and the Hazard Impact Tracker (HIT): Leveraging New Global Cyclone Data to create a Catastrophe Portfolio Management Platform (project number: 10105861) funded by Innovate UK (UKRI).

715 **References**

Aberson, S. D., Montgomery, M. T., Bell, M., & Black, M. (2006). Hurricane Isabel (2003): New insights into the physics of intense storms. Part II: Extreme localized wind. *Bulletin of the American Meteorological Society*, *87*(10), 1349–1354. https://doi.org/10.1175/BAMS-87-10-1349

Archer, L., Hatchard, S., Devitt, L., Neal, J. C., Coxon, G., Bates, P. D., Kendon, E. J., & Savage, J. (2024). Future Change in Urban Flooding Using New Convection-Permitting Climate Projections. *Water Resources Research*, *60*(1), e2023WR035533. https://doi.org/10.1029/2023WR035533

Archer, L., Neal, J., Bates, P., Vosper, E., Carroll, D., Sosa, J., & Mitchell, D. (2024). Current and future rainfall-driven flood risk from hurricanes in Puerto Rico under 1.5 and 2°C climate change. *Natural Hazards and Earth System Sciences*, *24*(2), 375–396. https://doi.org/10.5194/nhess-24-375-2024

Audi, C., Segarra, L., Irwin, C., Craig, P., Skelton, C., & Bestul, N. (2018). *Ascertainment of the Estimated Excess Mortality from Hurricane María in Puerto Rico*. Milken Institute of Public Health, The George Washington University. https://publichealth.gwu.edu/sites/g/files/zaxdzs4586/files/2023-06/acertainment-of-the-estimated-excess-mortality-from-hurricane-maria-in-puerto-rico.pdf

Baker, A. J., Vannière, B., & Vidale, P. L. (2024). On the Realism of Tropical Cyclone Intensification in Global Storm-Resolving Climate Models. *Geophysical Research Letters*, *51*(17), e2024GL109841. https://doi.org/10.1029/2024GL109841

Ballard, C., Wilson, M., Nojima, Y., Matanik, R., & Shing, R. (2020). Disaster as Opportunity? Cyclone Pam and the Transmission of Cultural Heritage. *Anthropological Forum*, *30*(1–2), 91–107. https://doi.org/10.1080/00664677.2019.1647825

Baño-Medina, J., Sengupta, A., Michaelis, A., Monache, L. D., Kalansky, J., & Watson-Parris, D. (2025). Harnessing AI Data-Driven Global Weather Models for Climate Attribution: An Analysis of the 2017 Oroville Dam Extreme Atmospheric River. *Artificial Intelligence for the Earth Systems*, *4*(3). https://doi.org/10.1175/AIES-D-24-0090.1

Benito, I., Eilander, D., Kelder, T., Ward, P. J., Aerts, J. C. J. H., & Muis, S. (2025). Pooling Seasonal Forecast Ensembles to Estimate Storm Tide Return Periods in Extra-Tropical Regions. *Journal of Geophysical Research: Oceans*, *130*(8), e2025JC022614. https://doi.org/10.1029/2025JC022614

Berghald, S., Mayer, S., & Bohlinger, P. (2024). Revealing trends in extreme heatwave intensity: Applying the UNSEEN approach to Nordic countries. *Environmental Research Letters*, *19*(3), 034026. https://doi.org/10.1088/1748-9326/ad2893

Biggerstaff, M. I., Alford, A. A., Carrie, G. D., & Stevenson, J. A. (2022). Hurricane Florence (2018): Long duration single- and dual-Doppler observations and wind retrievals during landfall. *Geoscience Data Journal*, *9*(2), 273–287. https://doi.org/10.1002/gdj3.137

Bloemendaal, N., de Moel, H., Muis, S., Haigh, I. D., & Aerts, J. C. J. H. (2020). Estimation of global tropical cyclone wind speed probabilities using the STORM dataset. *Scientific Data*, *7*(1), 377. https://doi.org/10.1038/s41597-020-00720-x

Bloemendaal, N., Haigh, I. D., de Moel, H., Muis, S., Haarsma, R. J., & Aerts, J. C. J. H. (2020). Generation of a global synthetic tropical cyclone hazard dataset using STORM. *Scientific Data*, *7*(1), 40. https://doi.org/10.1038/s41597-020-0381-2

Bourdin, S., Camargo, S. J., Lee, C.-Y., Lin, J., Vrac, M., Vaittinada Ayar, P., & Faranda, D. (2025). Improving analogues-based detection & attribution approaches for hurricanes. *Environmental Research Letters*, *20*(2), 024042. https://doi.org/10.1088/1748-9326/adaa8d

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

Brunner, M. I., & Slater, L. J. (2022). Extreme floods in Europe: Going beyond observations using reforecast ensemble pooling. *Hydrology and Earth System Sciences*, *26*(2), 469–482. https://doi.org/10.5194/hess-26-469-2022

Buonomo, E., Savage, N., Dong, G., Becker, B., Jones, R. G., Tian, Z., & Sun, L. (2024). Tropical Cyclone Changes in Convection-Permitting Regional Climate Projections: A Study Over the Shanghai Region. *Journal of Geophysical Research: Atmospheres*, *129*(5), e2023JD038508. https://doi.org/10.1029/2023JD038508

Carozza, D. A., Boudreault, M., Grenier, M., & Caron, L.-P. (2024). UQAM-TCW: A Global Hybrid Tropical Cyclone Wind Model Based Upon Statistical and Coupled Climate Models. *Journal of Advances in Modeling Earth Systems*, *16*(3), e2023MS003799. https://doi.org/10.1029/2023MS003799

Chenoweth, M., & Landsea, C. (2004). The San Diego Hurricane of 2 October 1858. *Bulletin of the American Meteorological Society*, *85*(11), 1689–1698. https://doi.org/10.1175/BAMS-85-11-1689

Ciullo, A., Martius, O., Strobl, E., & Bresch, D. N. (2021). A framework for building climate storylines based on downward counterfactuals: The case of the European Union Solidarity fund. *Climate Risk Management*, *33*, 100349. https://doi.org/10.1016/j.crm.2021.100349

Coughlan de Perez, E., Artur, L., Heinrich, D., Lagmay, A. M. F. A., Masukwedza, G. I. T., Nhantumbo, B., & Stephens, L. (2025). *Unseen but increasing: Recent changes in risk of extreme precipitation over Southern Africa and Southeast Asia* (SSRN Scholarly Paper No. 5812376). Social Science Research Network. https://doi.org/10.2139/ssrn.5812376

Davis, C. A. (2018). Resolving Tropical Cyclone Intensity in Models. *Geophysical Research Letters*, *45*(4), 2082–2087. https://doi.org/10.1002/2017GL076966

Djenontin, I. N. S., & Meadow, A. M. (2018). The art of co-production of knowledge in environmental sciences and management: Lessons from international practice. *Environmental Management*, *61*(6), 885–903. https://doi.org/10.1007/s00267-018-1028-3

ECMWF. (2016). *Part V: Ensemble Prediction System* (IFS DOCUMENTATION – Cy41r2 Operational Implementation). European Centre for Medium-Range Weather Forecasts. https://www.ecmwf.int/sites/default/files/elibrary/2016/79698-ifs-documentation-cy41r2-part-v-ensemble-prediction-system_1.pdf?utm_source=chatgpt.com

Emerton, R., Hodges, K. I., Stephens, E., Amelie, V., Mustafa, M., Rakotomavo, Z., Coughlan de Perez, E., Magnusson, L., & Vidale, P.-L. (2024). How well can global ensemble forecasts predict tropical cyclones in the southwest Indian Ocean? *Meteorological Applications*, *31*(3), e2195. https://doi.org/10.1002/met.2195

Fischer, E., Beyerle, U., Bloin-Wibe, L., Gessner, C., Humphrey, V., Lehner, F., Pendergrass, A., Sippel, S., Zeder, J., & Knutti, R. (2023). *Probing the Unfathomable: Ensemble Boosting for Physical Climate Storylines of Unseen Extremes* (Nos. EGU23-11466). Copernicus Meetings. EGU23. https://doi.org/10.5194/egusphere-egu23-11466

Frank, N. L. (2003). The Great Galveston Hurricane of 1900. In *Hurricane! Coping with Disaster: Progress and Challenges Since Galveston, 1900* (pp. 128–140). American Geophysical Union (AGU). https://doi.org/10.1029/SP055p0128

Frew, E., & White, L. (2020). Commemoration, communication and dark tourism: The cyclone tracy exhibit at the Northern Territory museum. In *CAUTHE 2013: Tourism and Global Change: On the Edge of Something Big* (pp. 227–230). https://doi.org/10.3316/informit.512819190983056

Gahtan, J., Knapp, K. R., Schreck, C. J., Diamond, H. J., Kossin, J. P., & Kruk. (2024). *International Best Track Archive for Climate Stewardship (IBTrACS) Project* (Version Version 4r01) [Dataset]. https://doi.org/10.25921/82ty-9e16

Hazeleger, W., van den Hurk, B. J. J. M., Min, E., van Oldenborgh, G. J., Petersen, A. C., Stainforth, D. A., Vasileiadou, E., & Smith, L. A. (2015). Tales of future weather. *Nature Climate Change*, *5*(2), Article 2. https://doi.org/10.1038/nclimate2450

Heinrich, D., Stephens, E., & Coughlan de Perez, E. (2024). More than Magnitude: Towards a multidimensional understanding of unprecedented weather to better support disaster management. *Water Security*, *23*, 100181. https://doi.org/10.1016/j.wasec.2024.100181

Hodges, K., Cobb, A., & Vidale, P. L. (2017). How Well Are Tropical Cyclones Represented in Reanalysis Datasets? *Journal of Climate*, *30*(14), 5243–5264. https://doi.org/10.1175/JCLI-D-16-0557.1

Hodges, K. I., & Emerton, R. (2015). The Prediction of Northern Hemisphere Tropical Cyclone Extended Life Cycles by the ECMWF Ensemble and Deterministic Prediction Systems. Part I: Tropical Cyclone Stage. *Monthly Weather Review*, *143*(12), 5091–5114. https://doi.org/10.1175/MWR-D-13-00385.1

Hooker, H., Dance, S. L., Mason, D. C., Bevington, J., & Shelton, K. (2023). Assessing the spatial spread–skill of ensemble flood maps with remote-sensing observations. *Natural Hazards and Earth System Sciences*, *23*(8), 2769–2785. https://doi.org/10.5194/nhess-23-2769-2023

Hooker, H., Heinrich, D., Cuckow, S., Hodges, K., & Stephens, L. (2026). *WATTCH: World Archive of Tracked Tropical Cyclone Hindcasts Preview* [Dataset]. https://doi.org/10.5281/zenodo.18608958

Hooker, H., Steinkopf, J., Vanya, C. L., Maure, G., Nhantumbo, B., Engelbrecht, F., Cloke, H., & Stephens, E. (2026). Extreme rainfall from tropical cyclones is revealed by km-scale downscaling in Southeast Africa. *Journal of Hydrometeorology*, *1*(aop). https://doi.org/10.1175/JHM-D-25-0179.1

Jeffries, I., Peters, L. E. R., Aljoda, A., Van Den Hoek, J., & Coughlan de Perez, E. (Submitted). What Makes Scenario Exercises Successful? Insights from Disaster Management Practice. *International Journal of Disaster Risk Reduction*.

Jiménez-Esteve, B., Barriopedro, D., Johnson, J. E., & García-Herrera, R. (2025). AI-Driven Weather Forecasts to Accelerate Climate Change Attribution of Heatwaves. *Earth's Future*, *13*(8), e2025EF006453. https://doi.org/10.1029/2025EF006453

Kay, A. L., Dunstone, N., Kay, G., Bell, V. A., & Hannaford, J. (2024). Demonstrating the use of UNSEEN climate data for hydrological applications: Case studies for extreme floods and droughts in England. *Natural Hazards and Earth System Sciences*, *24*(9), 2953–2970. https://doi.org/10.5194/nhess-24-2953-2024

Keellings, D., & Hernández Ayala, J. J. (2019). Extreme Rainfall Associated With Hurricane Maria Over Puerto Rico and Its Connections to Climate Variability and Change. *Geophysical Research Letters*, *46*(5), 2964–2973. https://doi.org/10.1029/2019GL082077

Kelder, T., Heinrich, D., Klok, L., Thompson, V., Goulart, H. M. D., Hawkins, E., Slater, L. J., Suarez-Gutierrez, L., Wilby, R. L., Coughlan de Perez, E., Stephens, E. M., Burt, S., van den Hurk, B., de Vries, H., van der Wiel, K., Schipper, E. L. F., Carmona Baéz, A., van Bueren, E., & Fischer, E. M. (2025). How to stop being surprised by unprecedented weather. *Nature Communications*, *16*(1), 2382. https://doi.org/10.1038/s41467-025-57450-0

Kelder, T., Müller, M., Slater, L. J., Marjoribanks, T. I., Wilby, R. L., Prudhomme, C., Bohlinger, P., Ferranti, L., & Nipen, T. (2020). Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes. *Npj Climate and Atmospheric Science*, *3*(1), 47. https://doi.org/10.1038/s41612-020-00149-4

Kelder, T., Wanders, N., Wiel, K. van der, Marjoribanks, T. I., Slater, L. J., Wilby, R. l, & Prudhomme, C. (2022). Interpreting extreme climate impacts from large ensemble simulations—Are they unseen or unrealistic? *Environmental Research Letters*, *17*(4), 044052. https://doi.org/10.1088/1748-9326/ac5cf4

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 Data. *Bulletin of the American Meteorological Society*, *91*(3), 363–376. https://doi.org/10.1175/2009BAMS2755.1

Lee, M., Min, S.-K., & Cha, D.-H. (2023). Convection-permitting simulations reveal expanded rainfall extremes of tropical cyclones affecting South Korea due to anthropogenic warming. *Npj Climate and Atmospheric Science*, *6*(1), 1–12. https://doi.org/10.1038/s41612-023-00509-w

Lee, S., & Zhao, J. (2021). Adaptation to climate change: Extreme events versus gradual changes. *Journal of Economic Dynamics and Control*, *133*, 104262. https://doi.org/10.1016/j.jedc.2021.104262

Li, M., & Toumi, R. (2026). Can tropical cyclones exist near the Equator? *Quarterly Journal of the Royal Meteorological Society*, *152*(774), e70014. https://doi.org/10.1002/qj.70014

Lin, N., & Emanuel, K. (2016). Grey swan tropical cyclones. *Nature Climate Change*, *6*(1), 106–111. https://doi.org/10.1038/nclimate2777

Mahdi, S. S., Jafri, H. A., Allana, R., Battineni, G., Khawaja, M., Sakina, S., Agha, D., Rehman, K., & Amenta, F. (2023). Systematic review on the current state of disaster preparation Simulation Exercises (SimEx). *BMC Emergency Medicine*, *23*, 52. https://doi.org/10.1186/s12873-023-00824-8

Meadow, A. M., Ferguson, D. B., Guido, Z., Horangic, A., Owen, G., & Wall, T. (2015). Moving toward the Deliberate Coproduction of Climate Science Knowledge. *Weather, Climate, and Society*, *7*(2), 179–191. https://doi.org/10.1175/WCAS-D-14-00050.1

Meiler, S., Vogt, T., Bloemendaal, N., Ciullo, A., Lee, C.-Y., Camargo, S. J., Emanuel, K., & Bresch, D. N. (2022). Intercomparison of regional loss estimates from global synthetic tropical cyclone models. *Nature Communications*, *13*(1), 6156. https://doi.org/10.1038/s41467-022-33918-1

Menemenlis, S., Vecchi, G. A., Gao, K., Smith, J. A., & Cheng, K.-Y. (2024). Extreme Rainfall Risk in Hurricane Ida's Extratropical Stage: An Analysis with Convection-Permitting Ensemble Hindcasts. *Journal of the Atmospheric Sciences*, *81*(7), 1165–1179. https://doi.org/10.1175/JAS-D-23-0160.1

Mitchell, D., AchutaRao, K., Allen, M., Bethke, I., Beyerle, U., Ciavarella, A., Forster, P. M., Fuglestvedt, J., Gillett, N., Haustein, K., Ingram, W., Iversen, T., Kharin, V., Klingaman, N., Massey, N., Fischer, E., Schleussner, C.-F., Scinocca, J., Seland, Ø., … Zaaboul, R. (2017). Half a degree additional warming, prognosis and projected impacts (HAPPI): Background and experimental design. *Geoscientific Model Development*, *10*(2), 571–583. https://doi.org/10.5194/gmd-10-571-2017

Msemo, H. E., Finney, D. L., & Mbuya, S. I. (2022). Forgotten accounts of tropical cyclones making landfall in Tanzania. *Weather*, *77*(4), 127–131. https://doi.org/10.1002/wea.3921

Ommer, J., Neumann, J., Kalas, M., Blackburn, S., & Cloke, H. L. (2024). Surprise floods: The role of our imagination in preparing for disasters. *EGUsphere*, 1–22. https://doi.org/10.5194/egusphere-2024-296

Paerl, H. W., Hall, N. S., Hounshell, A. G., Luettich, R. A., Rossignol, K. L., Osburn, C. L., & Bales, J. (2019). Recent increase in catastrophic tropical cyclone flooding in coastal North Carolina, USA: Long-term observations suggest a regime shift. *Scientific Reports*, *9*(1), 10620. https://doi.org/10.1038/s41598-019-46928-9

Pasch, R. J., Penny, A. B., & Berg, R. (2023). *Hurricane Maria* (Tropical Cyclone Report No. AL152017). National Hurricane Centre. https://www.nhc.noaa.gov/data/tcr/AL152017_Maria.pdf

Philp, T. J., Champion, A. J., Hodges, K. I., Pigott, C., MacFarlane, A., Wragg, G., & Zhao, S. (2022). Identifying Limitations when Deriving Probabilistic Views of North Atlantic Hurricane Hazard from Counterfactual Ensemble NWP Re-forecasts. In J. M. Collins & J. M. Done (Eds.), *Hurricane Risk in a Changing Climate* (pp. 233–254). Springer International Publishing. https://doi.org/10.1007/978-3-031-08568-0_10

Radfar, S., Foroumandi, E., Moftakhari, H., Moradkhani, H., Foltz, G. R., & Sen Gupta, A. (2024). Global Predictability of Marine Heatwave Induced Rapid Intensification of Tropical Cyclones. *Earth's Future*, *12*(12). https://doi.org/10.1029/2024EF004935

Ramos-Scharrón, C. E., & Arima, E. (2019). Hurricane María's Precipitation Signature in Puerto Rico: A Conceivable Presage of Rains to Come. *Scientific Reports*, *9*(1), 15612. https://doi.org/10.1038/s41598-019-52198-2

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vanniere, B., Mecking, J., Haarsma, R., Bellucci, A., Scoccimarro, E., Caron, L.-P., Chauvin, F., Terray, L., Valcke, S., Moine, M.-P., Putrasahan, D., Roberts, C., Senan, R., Zarzycki, C., & Ullrich, P. (2020). *Impact of Model Resolution on Tropical Cyclone Simulation Using the HighResMIP–PRIMAVERA Multimodel Ensemble*. https://doi.org/10.1175/JCLI-D-19-0639.1

Russotto, R. D., Strong, J. D. O., Camargo, S. J., Sobel, A., Elsaesser, G. S., Kelley, M., Del Genio, A., Moon, Y., & Kim, D. (2022). Evolution of Tropical Cyclone Properties Across the Development Cycle of the GISS-E3 Global Climate Model. *Journal of Advances in Modeling Earth Systems*, *14*(1), e2021MS002601. https://doi.org/10.1029/2021MS002601

Rye, C. J., & Boyd, J. A. (2022). Downward Counterfactual Analysis in Insurance Tropical Cyclone Models: A Miami Case Study. In J. M. Collins & J. M. Done (Eds.), *Hurricane Risk in a Changing Climate* (pp. 207–232). Springer International Publishing. https://doi.org/10.1007/978-3-031-08568-0_9

Schmitt, D., Gischler, E., Melles, M., Wennrich, V., Behling, H., Shumilovskikh, L., Anselmetti, F. S., Vogel, H., Peckmann, J., & Birgel, D. (2025). An annually resolved 5700-year storm archive reveals drivers of Caribbean cyclone frequency. *Science Advances*, *11*(11). https://doi.org/10.1126/sciadv.ads5624

Schreck, C. J., Knapp, K. R., & Kossin, J. P. (2014). The Impact of Best Track Discrepancies on Global Tropical Cyclone Climatologies using IBTrACS. *Monthly Weather Review*, *142*(10), 3881–3899. https://doi.org/10.1175/MWR-D-14-00021.1

Sloan, S. (2008). Oral History and Hurricane Katrina: Reflections on Shouts and Silences. *The Oral History Review*, *35*(2), 176–186. https://doi.org/10.1093/ohr/ohn027

Sobel, A. H., Lee, C.-Y., Camargo, S. J., Mandli, K. T., Emanuel, K. A., Mukhopadhyay, P., & Mahakur, M. (2019). Tropical Cyclone Hazard to Mumbai in the Recent Historical Climate. *Monthly Weather Review*, *147*(7), 2355–2366. https://doi.org/10.1175/MWR-D-18-0419.1

Sparks, N., & Toumi, R. (2024). The Imperial College Storm Model (IRIS) Dataset. *Scientific Data*, *11*(1). Scopus. https://doi.org/10.1038/s41597-024-03250-y

Sparks, N., & Toumi, R. (2025a). Climate change attribution of Typhoon Haiyan with the Imperial College Storm Model. *Atmospheric Science Letters*, *26*(1), e1285. https://doi.org/10.1002/asl.1285

Sparks, N., & Toumi, R. (2025b). The impact of global warming on U.S. hurricane landfall: A storyline approach. *Environmental Research Letters*, *20*(11), 114006. https://doi.org/10.1088/1748-9326/ae0956

Tam, H. C., He, Y.-H., Chan, P. W., Yu, S., Mo, H., Su, H., Hsiao, L.-F., & Gong, Y. (2025). Historical rainstorm in Hong Kong on 7–8 September 2023: Diagnosis, forecasting and nowcasting. *Atmospheric Science Letters*, *26*(1). https://doi.org/10.1002/asl.1284

Tao, S., Liu, K., Yu, K., Shi, Q., Yan, H., Zhang, H., Wang, L., Huang, Z., & Chen, T. (2021). Poleward Shift in Tropical Cyclone Tracks in the Northwest Pacific During Warm Periods: Past and Future. *Paleoceanography and Paleoclimatology*, *36*(12), e2021PA004367. https://doi.org/10.1029/2021PA004367

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., & Belcher, S. E. (2017). High risk of unprecedented UK rainfall in the current climate. *Nature Communications*, *8*(1), 107. https://doi.org/10.1038/s41467-017-00275-3

Thompson, V., Mitchell, D., Hegerl, G. C., Collins, M., Leach, N. J., & Slingo, J. M. (2023). The most at-risk regions in the world for high-impact heatwaves. *Nature Communications*, *14*(1), 2152. https://doi.org/10.1038/s41467-023-37554-1

Titley, H. A., Bowyer, R. L., & Cloke, H. L. (2020). A global evaluation of multi-model ensemble tropical cyclone track probability forecasts. *Quarterly Journal of the Royal Meteorological Society*, *146*(726), 531–545. https://doi.org/10.1002/qj.3712

https://doi.org/10.5194/egusphere-2026-2758
Preprint. Discussion started: 23 June 2026

Vickery, P. J., Skerlj, P. F., & Twisdale, L. A. (2000). Simulation of Hurricane Risk in the U.S. Using Empirical Track Model. *Journal of Structural Engineering*, *126*(10), 1222–1237. https://doi.org/10.1061/(ASCE)0733-9445(2000)126:10(1222)

Visman, E., Vincent, K., Steynor, A., Karani, I., & Mwangi, E. (2022). Defining metrics for monitoring and evaluating the impact of co-production in climate services. *Climate Services*, *26*, 100297. https://doi.org/10.1016/j.cliser.2022.100297

Vitart, F., Balmaseda, M. A., Ferranti, L., & Fuentes, M. (2022). The next extended-range configuration for IFS Cycle 48r1. *ECMWF*, *Number 173*(Autumn 2022). https://www.ecmwf.int/en/newsletter/173/earth-system-science/next-extended-range-configuration-ifs-cycle-48r1

Vosper, E. L., Mitchell, D. M., & Emanuel, K. (2020). Extreme hurricane rainfall affecting the Caribbean mitigated by the paris agreement goals. *Environmental Research Letters*, *15*(10), 104053. https://doi.org/10.1088/1748-9326/ab9794

Wang, Q., Guan, S., Luo, C., Zhao, W., Zhang, L., & Yang, S. (2025). Record-breaking lifespan, rapid intensification, and long-lasting coastal activity of tropical cyclone Freddy (2023) in the South Indian Ocean. *Environmental Research Letters*, *20*(6). https://doi.org/10.1088/1748-9326/add7ec

WMO. (2021). *Atlas of Mortality and Economic Losses from Weather, Climate and Water Extremes (1970–2019). Geneva: WMO.* World Meteorological Organization. https://doi.org/10.1163/9789004322714_cclc_2023-0268-1218

Xu, W., Balarugu, K., Judi, D., Rice, J., Leung, R., & Lipari-DiLeonardo, S. (2023). *North Atlantic synthetic tropical cyclone track, intensity, and rainfall dataset from RAFT* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.10392725