High prediction skill of decadal tropical cyclone variability in the North Atlantic and East Pacific in the met office decadal prediction system DePreSys4

npj | climate and atmospheric science

Article

Published in partnership with CECCR at King Abdulaziz University

https://doi.org/10.1038/s41612-025-00919-y

High prediction skill of decadal tropical cyclone variability in the North Atlantic and East Pacific in the met office decadal prediction system DePreSys4

Paul-Arthur Monerie [1] ✉, Xiangbo Feng [1,2], Kevin Hodges [1] & Ralf Toumi [2]

The UK Met Office decadal prediction system DePreSys4 shows skill in predicting the number of tropical cyclones (TCs) and TC track density over the eastern Pacific and tropical Atlantic Ocean on the decadal timescale (up to ACC = 0.93 and ACC = 0.83, respectively, as measured by the anomaly correlation coefficient—ACC). The high skill in predicting the number of TCs is related to the simulation of the externally forced response, with internal climate variability also allowing the improvement in prediction skill. The Skill is due to the model's ability to predict the temporal evolution of surface temperature and vertical wind shear over the eastern Pacific and tropical Atlantic Ocean. We apply a signal-to-noise calibration framework and show that DePreSys4 predicts an increase in the number of TCs over the eastern Pacific and the tropical Atlantic Ocean in the next decade (2023–2030), potentially leading to high economic losses.

Tropical cyclones (TCs) are strong atmospheric weather systems that travel long distances over land and the oceans. TCs are associated with high winds, storm surges, large waves[1] and heavy rainfall, causing casualties and economic losses[2]. TCs mostly affect many regions, including the tropical Atlantic and Caribbean Sea, the tropical Pacific, and the Indian Ocean[3,4].

Predicting the future evolution in the number and intensity of TCs is of high importance for populations and decision-makers, especially in vulnerable areas. Many studies have focused on understanding and improving TC prediction at sub-seasonal to annual timescales, i.e. sub-seasonal and seasonal predictions[4–7]. Some dynamical prediction systems have demonstrated an ability to predict the year-to-year evolution of the number of TCs over the North Atlantic Ocean and the western North Pacific Ocean[4,5]. One source of such prediction is the El Niño Southern Oscillation[4,8], with El-Niño events associated with an increased prediction skill for TCs relative to neutral years.

Evidence shows that anthropogenic activities can modulate the TC activity on multi-decadal to longer timescales. For example, one study has found that changes in anthropogenic aerosol emissions have affected TC activity over the North Atlantic and western North Pacific over the past 40 years through their effects on sea surface temperature and atmospheric circulation[9]. The externally forced response may lead to a future increase in the intensity of TCs[10], leading to a higher risk of TC-related damage over the tropics[11]. Several models project a future decrease in the number of TCs globally but with an increase in the proportion of the strong TCs[12]. However, the future change in TC activity at regional scales remains uncertain[12,13].

While decadal prediction provides useful pre-planning information for stakeholders and policy-makers[14], the performance of decadal prediction systems remains largely unexploited. The decadal timescale bridges the gap between seasonal prediction and climate projection. Studies have shown that dynamical prediction systems have skill for proxied TCs on a decadal timescale in the North Atlantic Ocean[14–16]. They link skill for TC activity to skill in predicting the Atlantic Multidecadal Variability (AMV). Skill has also been reported in predicting TC activity over the western North Pacific[17,18]. However, these aforementioned studies are based on a simplified approach, using proxies of TC (e.g., using the daily minimum of mean sea level pressure[15]), and have relied on statistical (TC predictions based on SSTs only[17,18]) or hybrid statistical-dynamic prediction frameworks[19]. These studies have severe limitations because non-stationarities in the TC-environment proxy relationship can strongly impact the statistical approach. The ability of dynamical prediction systems to explicitly predict TCs on decadal timescales is unknown. Here, we assess the prediction skill of TCs globally, using an explicit tracking algorithm to identify TCs, in the newly developed decadal prediction system of the UK Met Office, DePreSys4 (see Methods).

[1]Department of Meteorology, National Centre for Atmospheric Science, University of Reading, Reading, United Kingdom. [2]Department of Physics, Imperial College, London, United Kingdom. ✉e-mail: p.monerie@reading.ac.uk

We address the following questions:

- Can DePreSys4 predict the evolution of the TC activity up to 10 years ahead?
- Can we define sources of predictive skill for TC activity?
- Can we increase the predictive skill of TCs by addressing the signal-to-noise paradox in climate models?

Results

Representation of the climatology of TC activity

We first assess the ability of DePreSys4 to simulate the TC track density over the 1961–2021 period, focusing on interannual variability to use the largest number of start dates possible. Simulated TCs are validated against TCs tracked in ERA5 using the same TC tracking method. Climatologically, the TC track density is highest over the eastern and western North Pacific Ocean in ERA5 (Fig.1a), as shown in other studies[3]. DePreSys4 simulates the geographical distribution of track density well, with the highest values seen over the Pacific Ocean (Fig. 1b) and resembling the track density pattern in ERA5. However, DePreSys4 underestimates the track density over the North Atlantic Ocean and overestimates it over the Pacific and Indian Oceans (Fig. 1c). These biases in the representation of the track density are consistent with other prediction systems from the Met Office[8,20,21]. Consistent with the track density, DePreSys4 underestimates TC genesis over the North Atlantic Ocean and overestimates it over the Pacific and Indian Oceans (Fig. S1).

We estimate the ability of DePreSys4 to simulate the temporal variability in the TC track density. The DePreSys4 ensemble mean strongly underestimates the interannual variability of TC track density (Fig. 1e, f). We resample the ensemble members to assess the interannual variability of the track density, using a single realisation for each start date[7]. We confirm that the underestimation of the inter-annual variability of the track density is not solely in the ensemble mean but also inherent to individual members of DePreSys4.

Prediction skill

In this subsection, we focus the analysis on the decadal prediction of TCs and show the ability of DePreSys4 to simulate the TC track density on the 2–9-year forecast lead time (Fig. 2a). The 2–9-year forecast lead time is defined as an 8-year average, between the 2nd and 9th year of the hindcasts, and is used to assess the decadal variability of the TCs. DePreSys4 shows a significant prediction skill in simulating TC track density over the North Atlantic Ocean and the eastern North Pacific Ocean (Fig. 2a). At the 2–9-year forecast lead time, no significant prediction skill is found over the southern hemisphere and the western Pacific Ocean. The following analysis then focuses on the Atlantic and eastern Pacific Oceans.

We show that the model has skill in simulating TC genesis density over the eastern and central Atlantic and the northern and central Pacific Ocean (Fig. 2b). We hypothesize that the better skill in track density than in TC genesis is due to the skilful simulation of TC tracks, which are largely controlled by the large-scale steering flow. To better understand the high prediction skill in TC tracks in these two regions, we confine the TC genesis to the eastern Pacific (grey area in Fig. 2e) and the eastern Atlantic (yellow area in Fig. 2e). We show that DePreSys4 has a high and significant skill (ACC > 0.4) in predicting the frequency of TC genesis over both the eastern Pacific and the tropical Atlantic Ocean on the multi-annual to decadal time scales (2–4, 2–5, 3–6, 6–9 and 2–9-year forecast lead time) (Fig. 2c, d). However, the model has low skill in the interannual variability (1-year forecast lead time; Fig. 2c, d). On multi-annual to decadal timescales, the significant skill in simulating track density over the East Pacific and the tropical Atlantic Ocean is related to the significant skill in simulating the TC genesis. DePreSys4 simulates both the frequency of genesis and the trajectory of TCs well in these two regions.

We assess the effect of the externally forced response by removing a trend (quadratic) from the time series of the regional number of TCs. We show that the skill in the eastern Pacific is then much lower and statistically insignificant for most of the lead times (Fig. 2c), demonstrating the dependence of the prediction skill on the long-term evolution of TCs.

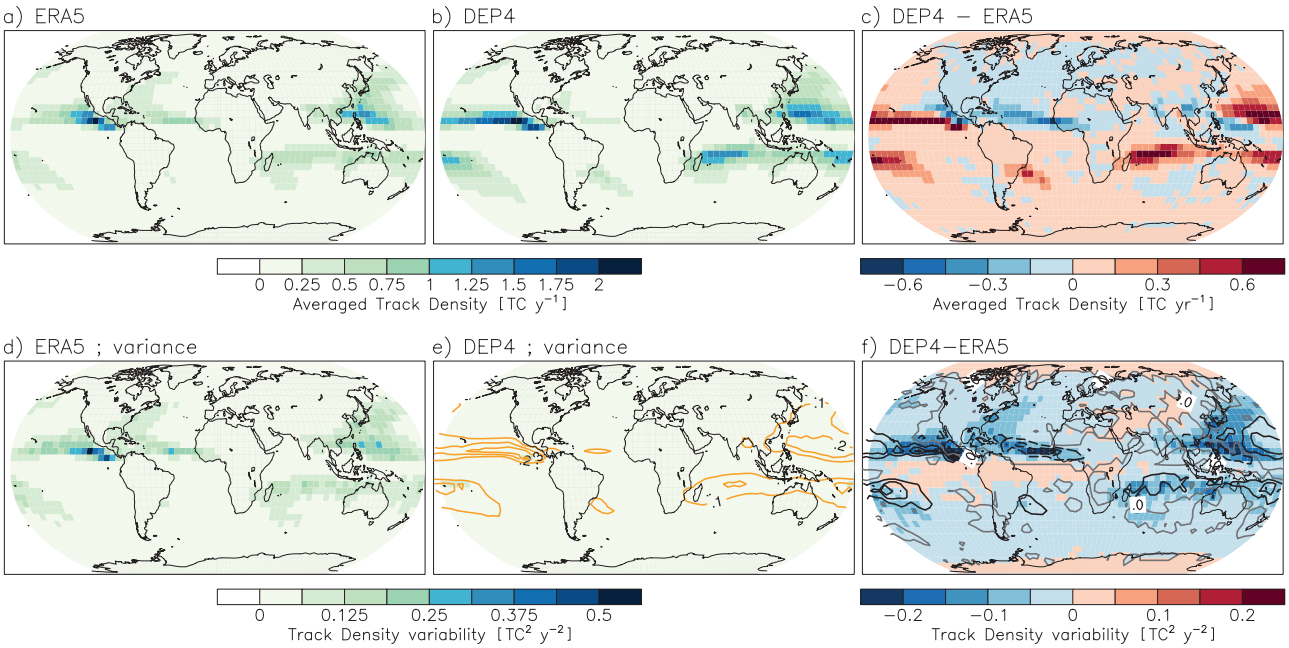

Fig. 1 | Representation of the track density. Track density climatology [1961–2021; TC year^{-1}] for **a** ERA5, **b** DePreSys4, and **c** track density bias (DePreSys4−ERA5). Interannual variability (variance; TC^2 year^{-2}) of TC track density (1961–2021) for **d** ERA5, **e** DePreSys4 ensemble mean, and **f** for the bias (DePreSys4−ERA5). Results are given for JASO in the Northern Hemisphere and DJFM in the Southern Hemisphere. **e**, **f** the contours show the mean value of variance obtained with individual ensemble members, obtained by randomly selecting individual ensemble members for each start date before computing the variance over the period 1961–2021 and with 5000 permutations. **f** Shows positive (negative) anomalies with continuous (dotted) black lines; the grey line shows zero anomalies. The track density climatology and interannual variability are computed from 1-year forecast lead time.

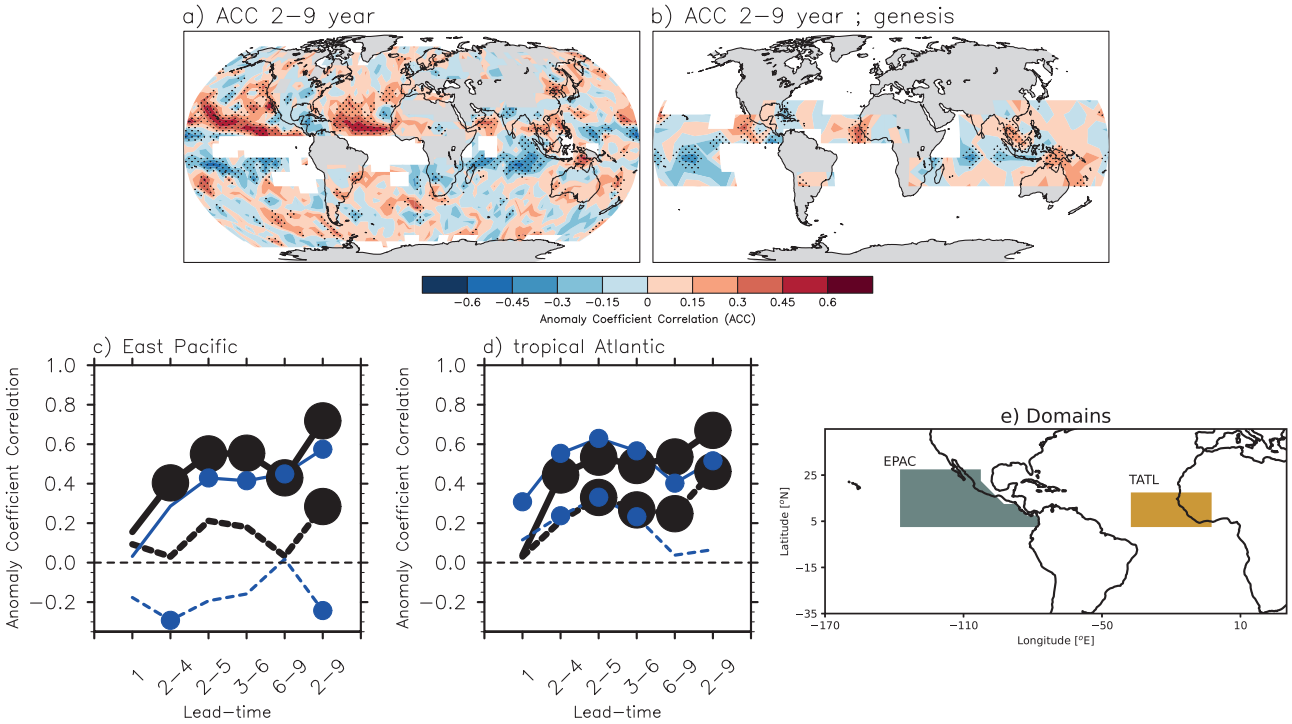

Fig. 2 | Skill in predicting TC activity. Skill (ACC) of DePreSys4 at the **a** 2–9-year forecast lead time (i.e., a decadal time window) for the track density. **b** as in **a** but for the TC genesis density. Stippling indicates that ACC is significantly different from zero according to a Monte–Carlo procedure with 5000 permutations and a 95% confidence level. Skill (ACC) in simulating the number of TC genesis in **c** the eastern Pacific and **d** tropical east Atlantic Ocean for the 1-year, 2–4, 2–5, 3–6, 6–9, and 2–9-year forecast lead times (i.e., in annual-to-decadal time windows). **c**, **d** The continuous (dotted) line shows the skill before (after) removing a quadratic trend, and circles indicate that ACC values are significant according to a Monte-Carlo procedure with 5000 permutations and a 95% confidence level. Black lines and circles show the prediction skill of DePreSys4, while the blue lines and circles show the prediction skill based on persistence only. The domains used to count the number of TC genesis in the East Pacific and tropical East Atlantic Ocean are shown within **e**.

However, over the tropical Atlantic Ocean, the prediction skill of the number of TCs remains high after removing the trend, showing that the long-term evolution and the multi-annual and decadal variability are well predicted in DePreSys4. The residual remains significantly correlated, showing that DePreSys4 can simulate the effect of internal climate variability on the number of TCs (Fig. 2c, d).

Persistence (see Methods) allows for high skill in predicting the number of TCs over both the eastern Pacific and the tropical Atlantic Ocean (Fig. 2c, d). Still, we show that DePreSys4 outperforms the persistence at the 2–9-year forecast lead time, showing added values of the dynamical prediction relative to the persistence.

Prediction skill of a large-scale environment

The development of TCs is, in general, controlled by large-scale environmental factors[4,18,22–29]. We assessed various environmental factors related to TC development. DePreSys4 is not skilful in the decadal variability of the 850 hPa relative vorticity and near-surface relative humidity (not shown). Instead, we find that DePreSys4 has significant skill in surface air temperature and vertical wind shear.

Sea surface temperature and vertical wind shear are important thermal and dynamical drivers for TC generation and development, with anomalously high temperature and low wind shear favouring TC activity[18,28]. We link the high skill in predicting the number of TCs in the above two regions to the high skill in simulating surface air temperature (Fig. 3a) and the vertical wind shear (Fig. 3b) over the Atlantic and the Pacific Ocean at the 2–9-year forecast lead time.

After excluding the long-term trend, the prediction skill in surface air temperature remains high over the North Atlantic, with a pattern of prediction skill reminiscent of the AMV (i.e., a horseshoe pattern) (Fig. 3c). The skill in predicting the vertical wind shear also remains particularly high over the tropical Atlantic Ocean (Fig. 3d). The significant skill in surface air temperature and vertical wind shear agrees with the high skill in predicting the number of TCs over the Atlantic Ocean after removing a quadratic trend (Fig. 2d). We also show that, in the North Atlantic Ocean, the surface air temperature of the North Atlantic Ocean is negatively correlated with the vertical wind shear in both the reanalysis and DePreSys4 (Fig. S2) on the decadal timescale. This is consistent with previous studies showing that the multidecadal variability of North Atlantic SSTs (AMV) can (i) modulate tropical Atlantic surface air temperature[30], (ii) lead to an increase in the number of TCs regionally[31], and (iii) is generally well predicted by climate models[32,33]. The model can predict TCs over the North Atlantic and the eastern Pacific Ocean on the decadal timescale even without skill for low-level relative vorticity and near-surface relative humidity. Therefore, we conclude that a high prediction skill in the North Atlantic air temperature and vertical wind shear contributes to the high skill in the number of the TCs over the tropical Atlantic in DePreSys4 regarding trend and decadal variability.

The skill in predicting the temperature and vertical wind shear is low over the East Pacific Ocean after removing a quadratic trend (Fig. 3c, d). This is consistent with an apparent decrease in the skill in predicting the number of TCs over the East Pacific, when not accounting for the long-term trend in TCs (Fig. 2c). However, the skill remains statistically significant (ACC > 0.3) over the East Pacific. This may be related to the remote effect of the skilfully simulated AMV on the East Pacific atmospheric circulation[34,35], which can modulate TC activity over the Pacific Ocean[36,37].

Calibrating prediction skills and predicting future changes

The number of TCs increases in ERA5 and DePreSys4 over the eastern Pacific and the tropical Atlantic over the hindcast period (Fig. 4a, b). The DePreSys4 ensemble mean underestimates the variability in the number of TCs (in agreement with Fig. 1f) (Fig. 4a, b). The underestimation of the

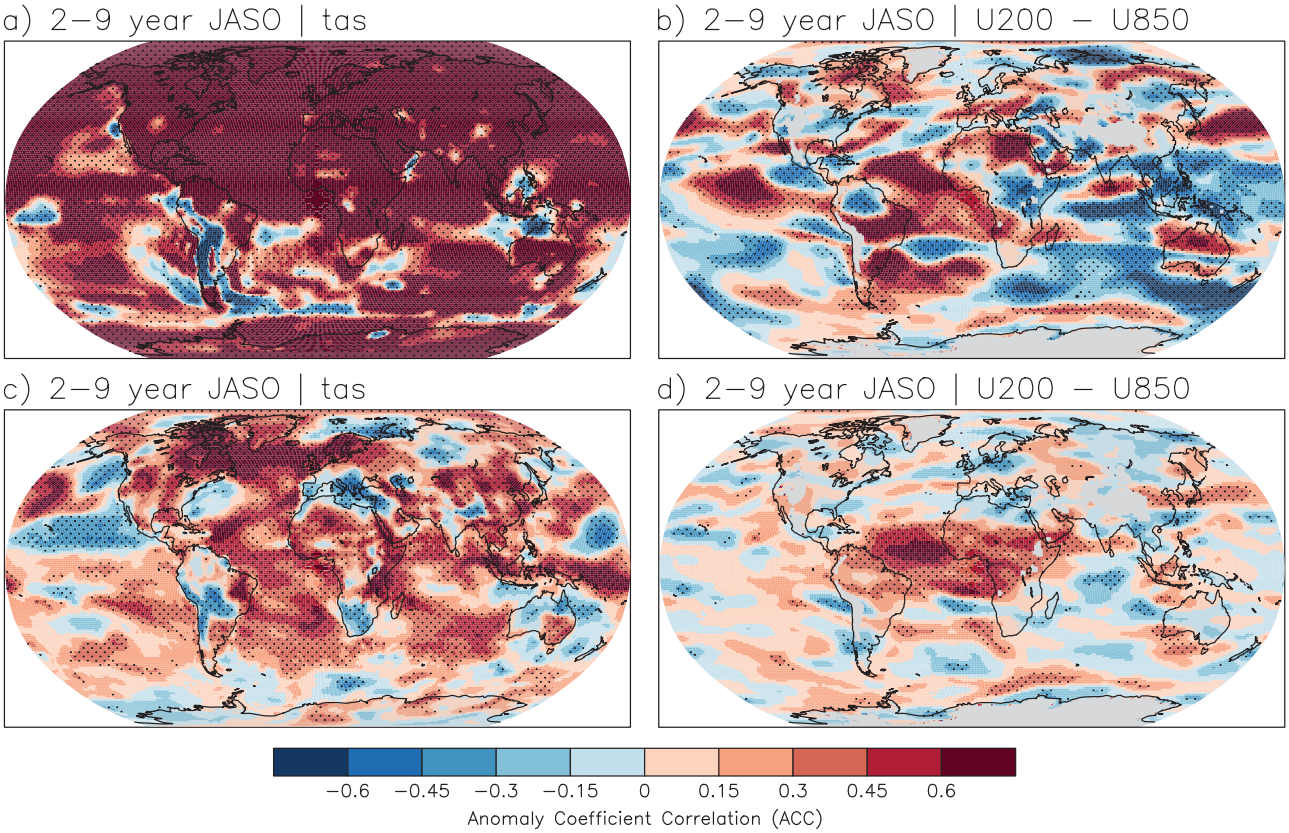

Fig. 3 | Environmental drivers of TC activity. Skill (ACC) of DePreSys4 in predicting **a** surface-air temperature and **b** vertical wind shear (U200−U850) at the 2–9-year forecast lead time (i.e., a decadal time window) in JASO. **c**, **d** as in **a**, **b**, but after removing a quadratic trend. Stippling indicates that ACC is significantly different from zero according to a Monte–Carlo procedure (bootstrap permutations) with 5000 permutations and a 95% confidence level.

decadal variability is related to the large ensemble spread and drastically reduced variability in the ensemble mean. The existence of a large ensemble spread in a prediction system has been studied in a signal-to-noise paradox framework[38]. The ratio of the predictable component (RPC) of TC frequency is greater than unity in both the East Pacific (RPC = 1.65) and tropical Atlantic (RPC = 1.83), indicating that DePreSys4 is better at predicting a version of the 'real world' than predicting itself[38]. The high RPC ratio in DePreSys4 also indicates that the model's skill in predicting the coherent variability of TCs is underestimated. We note that a different evolution of the number of TCs could be obtained in observations (IBTrACS and Best Track), showing a disparity from ERA5[26]. This is further shown in Figs. S3 and S4 and discussed in the supplementary material. Nevertheless, our evidence points out that the calibration method can further improve the prediction skill (Fig. 4). Moreover, we replicate the analysis using a second reanalysis (JRA-3Q[39]) and show a similar skill (ACC = 0.69 in the East Pacific and ACC = 0.63 in the tropical Atlantic) as for ERA5 (Fig. S5).

We apply the lagged ensemble technique[40], to increase the degree of freedom for each start date and combine the ensemble members from four consecutive start dates to 40 ensemble members to reduce the ensemble spread[41]. We also scaled the variance of DePreSys4 back to the variance of ERA5 to capture better the simulated temporal variability in the number of TCs. This post-processing procedure increases the skill for the East Pacific (ACC = 0.93) and the tropical Atlantic (ACC = 0.84). Therefore, we show that a higher skill is achievable after simply calibrating the forecast data. We predict the number of TCs in the recent and following decades using the forecasts initialised in and after November 2013 (red lines in Fig. 4). The lagged ensemble shows that the number of TCs is predicted to be much higher in the next decade (e.g., the value in the year 2026 is for the average of the period 2023–2030) over both the eastern Pacific and the tropical Atlantic (Fig. 4c, d). We note here that one should be cautious when analysing the prediction. We show high skill in predicting TCs over the East Pacific and Atlantic Ocean but acknowledge that the model's skill could be time-dependent and lower for the forecast period.

In addition to the number of TCs, we assess the ability of DePreSys4 to simulate the accumulated cyclone energy (ACE), the energy associated with TC activity. An increase in ACE can lead to strong impacts on the ocean and land. Prediction skill of the lagged ensemble in the decadal variability of ACE is high over the eastern Pacific (ACC = 0.72; Fig. 5a) and over the tropical Atlantic (ACC = 0.81; Fig. 5b). In addition to the tropical Atlantic, we show that DePreSys4 can predict ACE over a larger area, the North Atlantic (from 0°N to 60°N; ACC = 0.63; Fig. 5c), with cyclones and high wind speeds that can reach the Caribbean, Central America, and the eastern US. We show that DePreSys4 generally predicts a future increase in ACE (Fig. 5a–c), indicating that cyclone-related losses could increase in the near future.

Discussion

We assess the ability of a decadal prediction system developed by the UK Met Office, DePreSys4, to predict the number of TCs and TC track density. DePreSys4 consists of a large number of hindcasts, initialised every year from 1960 to 2021, with ten-year hindcasts and ten ensemble members for each start date. We found that DePreSys4 can predict the number of TCs and track density up to a decade ahead. While several studies have shown that prediction systems can predict TC characteristics on a decadal time scale, these studies have been based on proxies for TCs[15,17–19], with some studies based on a statistical approach that uses the decadal modes of SST variability as proxies for the TC activity[17,18]. The downside of these statistical approaches is that there may be no stationarity in the SST-TCs relationship and that other drivers of TC variability are not considered. It is essential to know how well the decadal prediction systems explicitly simulate TCs, as the current model resolution is high enough to resolve the storm dynamically. Here, we bridge this gap, using, for the first time, a TC tracking algorithm to

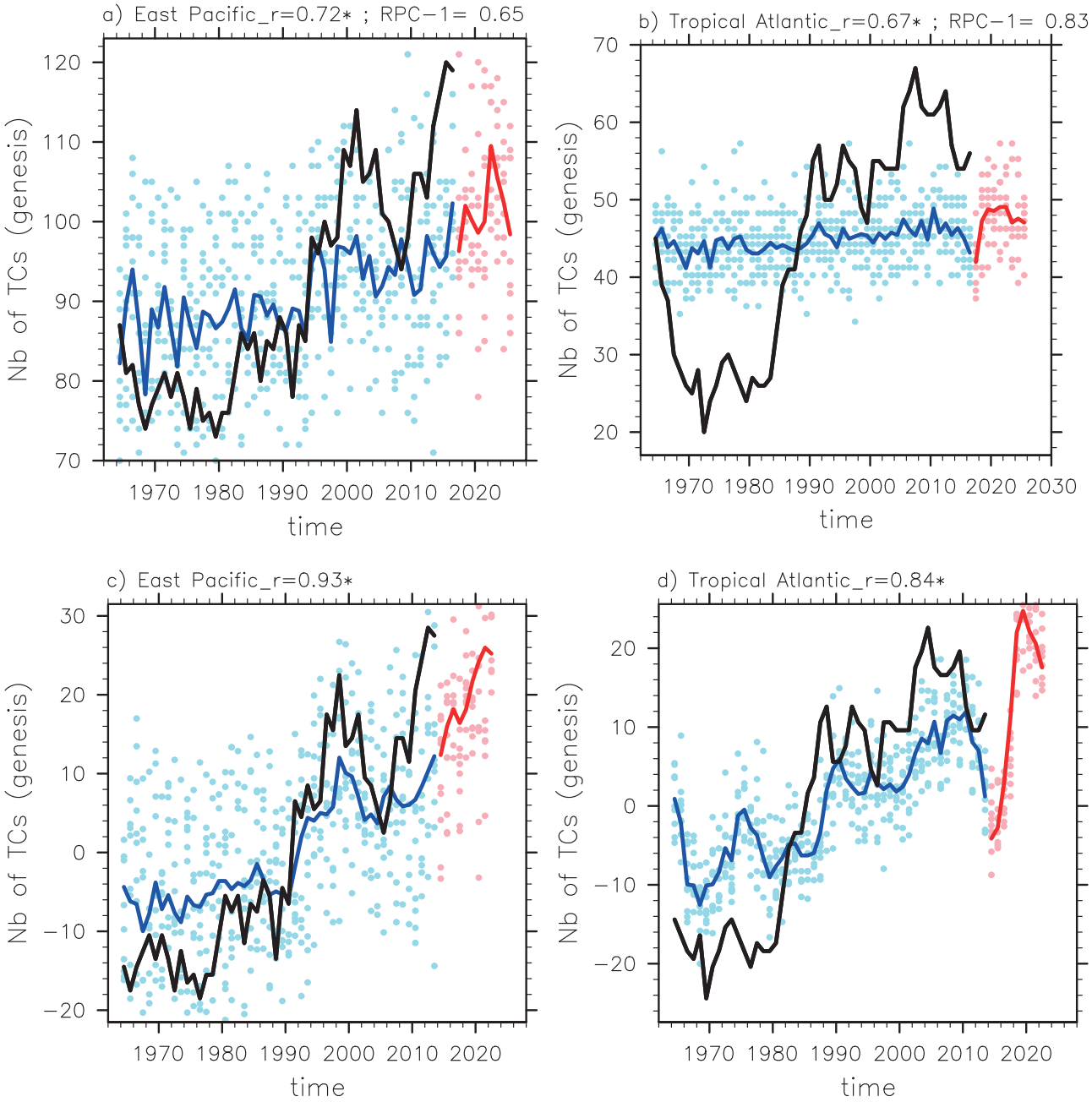

Fig. 4 | Prediction of the overall number of TC genesis in the eastern Pacific and tropical Atlantic oceans. The overall number of TC genesis in ERA5 (black line) and the DePreSys4 ensemble mean (blue and red lines) for the **a** eastern Pacific Ocean and **b** eastern tropical Atlantic Ocean in the 2–9-year forecast lead time. The blue and red lines are used for the hindcast and forecast periods, respectively. All individual ensemble members are shown with a dot. The ACC and RPC values are shown and computed over the hindcast period. **c**, **d** as for **a**, **b**, but for the lagged ensembles with a scaling of the variance. **c**, **d** are shown in an anomaly relative to the whole time series. For DePreSys4, the years represent the mid-point of the 2–9-year forecast period (a decadal time window). The year 2026 represents the average period 2023–2030.

identify TCs in the prediction system on a fine time scale (6 hours) to assess the prediction skill for a decade ahead.

We show that DePreSys4 underestimates the amplitude of long-term variability in the number of TCs and TC track density globally. Still, DePreSys4 is skilful on the multi-annual and decadal lead times (i.e., 2–4, 6–9, 2–9 years forecast lead time) over the eastern Pacific and tropical Atlantic Ocean. DePreSys4 is not skillful over the southern hemisphere and western North Pacific Ocean. We show that the skill in predicting the number of TCs is mainly due to a trend in the evolution of the number of TCs. Still, the simulation of internal climate variability also significantly contributes to the prediction skill. We suggest that the high skill in predicting the surface temperature and the vertical wind shear over the tropical Atlantic and eastern Pacific contributes to the high prediction skill for TC genesis in the Atlantic and the east Pacific.

We show that the calibration method[40] allows a significant increase in prediction skills. Predicting the evolution of the number of TCs on the decadal timescale provides useful information for adaptation strategy and planning management. We show that DePreSys4 predicts the number of TCs, and the energy associated with the TCs to increase over the East Pacific and the Atlantic in the near future. This may increase cyclone-related losses over the Atlantic and the East Pacific Ocean in the near future.

Additional work could be done to better understand the sources of prediction skill for the TC genesis and track density at the decadal timescale, focusing, for example, on specific case studies (decades) and highlighting

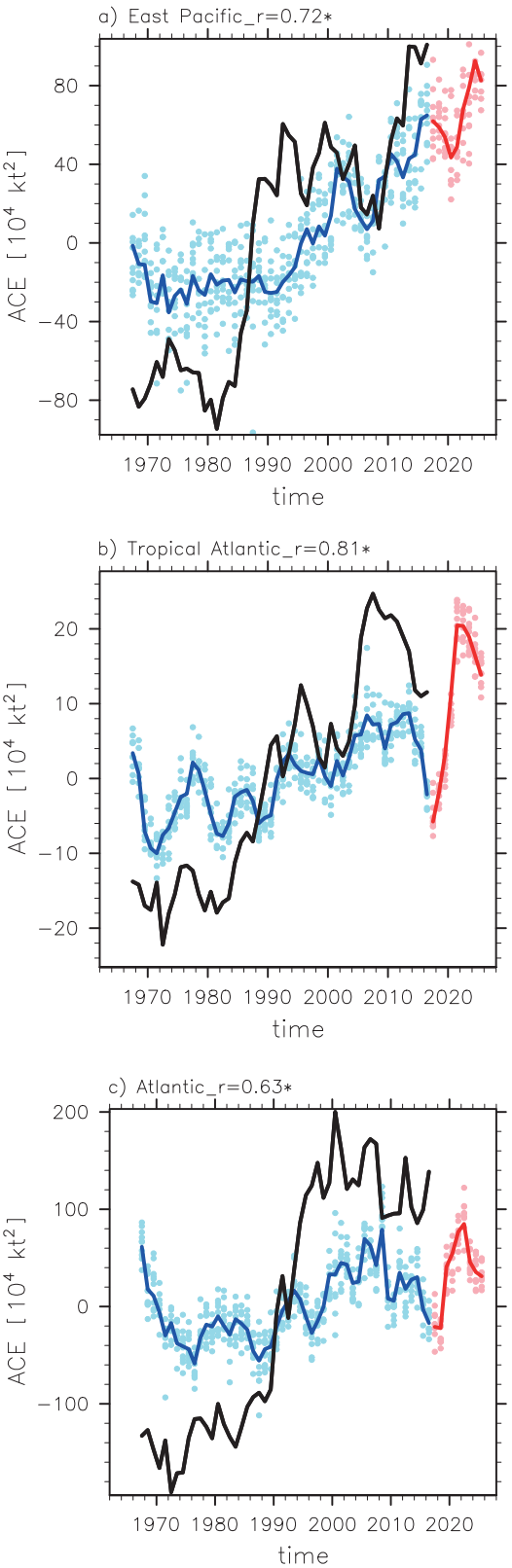

Fig. 5 | Changes in accumulated cyclone energy. As for Fig. 4a but for accumulated cyclone energy (ACE; 10^4 kt^2) and **a** the East Pacific, **b** the tropical Atlantic, and **c** the North Atlantic.

mechanisms at play. Work could also be devoted to investigating the causes of model biases in the simulation of TC genesis and track density[8,21] and understanding their impact on forecast skill, for example, through sensitivity experiments[42] in which key variables are kept close to the reanalyses. In addition, we expect prediction skills to be model-dependent and advocate a multi-model analysis using hindcasts from a large ensemble of prediction systems. Because evaluating multi-model hindcasts requires a large amount of sub-daily (e.g., 6-hourly) field data to be analyzed and stored for TC identification, we suggest modeling groups could provide TC track data as an output for CMIP7. This would require the climate centres to use the same tracking scheme.

Methods

DePreSys4

We assess the ability of a decadal prediction system, DePreSys4, developed by the UK Met Office to predict TC activity up to a decade ahead. DePreSys4 is based on HadGEM3-GC31-MM[43], an ocean-atmosphere general circulation model with a resolution of ~0.5° longitude and ~0.8° latitude and with 36 vertical levels. We use 10-year simulations, initialised each November, from 1960 to 2021. There are ten ensemble members that differ from their initial conditions (initialised from different ocean analyses to sample uncertainties in the initial conditions) for each start, for a total of 6200 years of simulations.

ERA5

We assess the ability of DePreSys4 to predict TC activity by contrasting with the European Centre for Medium-Range Weather Forecasts (ECMWF) 5th-generation reanalysis (ERA5[44]). ERA5 is used at a resolution of 0.25° of latitudes and longitudes. We use data from ERA5 covering the period 1960–2022. The same TC identification criteria are used for both DePreSys4 and ERA5. We do not use the IBTrACS[45] and Best Track (National Hurricane Center)[46] observations, for which different operational procedures are used in different ocean basins, which does not allow a clean comparison with DePreSys4. However, comparisons to observations are shown in the Supplementary Material.

NCEP

Skill at predicting surface air temperature and wind speed is quantified using the NCEP reanalysis[47], given on a 2.5° × 2.5° horizontal resolution and from 1948 to the present.

Tracking algorithm

Our identification of TCs follows previous studies[8,48–51]. The tracking uses the 6-hourly 850 hPa relative vorticity. We reduce noise in the vorticity field before performing the tracking. The data is spectrally filtered using a fit to a spherical harmonic expansion truncated to total wave numbers 42(T42), with the total wave numbers less than or equal to 5 removed. Initially, all systems tracked that exceed an intensity maximum greater than $5.0 \times 0^{-6}\,s^{-1}$ in the NH or a minimum $<-5.0 \times 0^{-6}\,s^{-1}$ in the SH. The tracking first initialises a set of tracks using a nearest neighbour method, which is then refined by minimising a cost function for track smoothness subject to adaptive constraints on the displacement distance and track smoothness. Following the tracking, the T63 vorticity maxima/minima are recursively added to the tracks at the available levels of 850, 500, and 200 hPa, as well as the 10 m wind maxima using a 6° search radius, and the MSLP minima using a 5° search radius. The genesis (first tracked point) must be within the tropics (30°S–30°N). The difference in vorticity between the 850 and 200 hPa levels must be $>6 \times 10^{-5}\,s^{-1}$; the intensity at 850 hPa must be greater than $6 \times 10^{-5}\,s^{-1}$; we must obtain a coherent vertical structure, as defined by the presence of a vorticity center at each vertical level between 850 and 200hPa; these last three criteria must be satisfied for at least 4 consecutive time steps over the ocean.

Track density and genesis

We use a grid with a horizontal resolution of 5° × 5° to remap the track density for DePreSys4 and ERA5. We count the number of TCs in each grid point. We use a larger grid with a horizontal resolution of 15° × 15° to show the TC genesis. TC genesis is defined as the location of a TC at its

first step, and the number of TCs is defined as the number of TCs registered over a given domain. For the East Pacific, we register the number of TCs whose genesis occurred over the East Pacific domain (See Fig. 2e). For the Tropical Atlantic domain, we account for the number of TCs that travel through the Atlantic domain (see Fig. 2e), to also account for TCs whose genesis occurs inland, over West Africa. This allows accounting for ~67% of the TCs as obtained over the full North Atlantic domain in ERA5.

Assessing skill

Prediction skill is estimated using the anomaly correlation coefficient (ACC) metric, calculated between ERA5 and DePreSys4, and for several forecast lead times. We focus on different timescales of variability by using different forecast lead times and by comparing them. The 1 year forecast lead time (here the first winter and the first summer), therefore, allows us to determine the skill in predicting the interannual variability, we also use the 2–4, 2–5, 3–6, and 6–9 year forecast lead times, where 4-year averages allow us to document the skill in predicting the multi-year variability of TCs, and finally we use the 2–9 year forecast lead time (an 8-year average) to document the skill in predicting the decadal variability of TC activity.

We assess skill for the July–October season (JASO) in the Northern Hemisphere from 1960 to 2021 and the December–March season (DJFM) in the Southern Hemisphere from 1960 to 2020.

The significance of the ACC is estimated by randomly resampling the time series of the ensemble means. We use a 5-year block bootstrap to preserve low-frequency variability using 5000 permutations in a Monte Carlo framework. The ACC values are judged significant at the 95% confidence level using a two-sided test.

Persistence

We use persistence as a benchmark to assess the usefulness of DePreSys4. The *n*-year persistence is calculated based on the ERA5 values in the *n-years* before the start date. We calculated 1-year, 4-year, and 8-year persistence.

Drift correction

We remove the model's drift following the recommendations of the World Climate Research Programme[52], which is defined as the lead-time bias relative to ERA5. Note that removing the drift does not affect the prediction's skill.

The RPC and lagged ensemble

We assess confidence in prediction by using the signal-to-noise framework of Scaife et al.[38], which allows a quantification of the inconsistency between the low strength of the predictable signals in a climate model and the relatively high level of agreement it exhibits with the observed variability. The RPC[38], $RPC^2 = r_{om}^2/r_{mm}^2$ is used, where r_{om} is the correlation between the DePreSys4 ensemble mean and ERA5, r_{mm} is the correlation between the ensemble mean and a single ensemble member (obtained here as the average of an ensemble of synthetic time series obtained by randomly a single ensemble for each start date and with 5000 permutations). As r_{om} and r_{mm} indicate the ability to reproduce the predictable component of a signal, the RPC indicates a ratio between the RPC in ERA5 and the predictable component in DePreSys4[53]. An RPC equal to unity indicates a perfect prediction system. RPC greater than unity denotes that the RPC is lower in DePreSys4 than in ERA5.

We expect the prediction skill to increase when increasing the ensemble size[8]. The lagged ensemble allows for increasing the ensemble size by combining the four latest forecasts available at each start date (giving 40 ensemble members instead of 10 ensemble members)[40].

We rescale the variance of the predicted ensemble mean by scaling DePreSys4 by $\sqrt{\frac{var(obs)}{var(model)}}$, where var(obs) is the variance of ERA5 and var(model) is the variance of DePresys4[54]. The variance is computed from the detrended time series on each considered forecast lead time.

Accumulated cyclone energy

A way to estimate the cyclone intensity is to use the ACE[5,55,56]. We estimate ACE using the 10-m maximum wind speed for each region as

$$ACE = 10^{-4} \sum_i \sum_j V_{\max}^2,$$

where $V_{\max}$ is the 6-hourly maximum 10-m wind speed associated with each cyclone and is given as the sum of the square of the wind speed over all tracks i and track points j. ACE is in 10^4 kt^2 (1 kt~0.5 m s^{-1}).

Data availability

DePreSys4 contributes to CMIP6, and outputs are available from the CMIP6 public repositories, including https://esgf-index1.ceda.ac.uk/search/cmip6-ceda/. The ERA5 data are generated by ECMWF and available on their website (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview)/. NCEP data are provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their website at https://downloads.psl.noaa.gov/Datasets/ncep.reanalysis/Monthlies/pressure/.

Code availability

Codes are available upon reasonable request to the corresponding author. The tracking code can be downloaded from https://gitlab.act.reading.ac.uk/track/track/-/releases.

Received: 16 September 2024; Accepted: 14 January 2025;
Published online: 25 January 2025

References

1. Shi, J. et al. Global increase in tropical cyclone ocean surface waves. *Nat. Commun.* **15**, 174 (2024).
2. Pant, S. & Cha, E. J. Wind and rainfall loss assessment for residential buildings under climate-dependent Hurricane scenarios. *Struct. Infrastruct. Eng.* **15**, 771–782 (2019).
3. Emanuel, K. Tropical cyclones. *Annu. Rev. Earth Planet. Sci.* **31**, 75–104 (2003).
4. Takaya, Y. et al. Recent advances in seasonal and multi-annual tropical cyclone forecasting. *Trop. Cyclone Res. Rev.* **12**, 182–199 (2023).
5. Befort, D. J., Hodges, K. I. & Weisheimer, A. Seasonal prediction of tropical cyclones over the North Atlantic and Western North Pacific. *J. Clim.* **35**, 1385–1397 (2022).
6. Vitart, F. & Stockdale, T. N. Seasonal forecasting of tropical storms using coupled GCM integrations. *Mon. Weather Rev.* **129**, 2521–2537 (2001).
7. Vitart, F. Seasonal forecasting of tropical storm frequency using a multi-model ensemble. *Q. J. R. Meteorol. Soc.* **132**, 647–666 (2006).
8. Feng, X., Klingaman, N. P., Hodges, K. I. & Guo, Y.-P. Western North Pacific tropical cyclones in the met office global seasonal forecast system: performance and ENSO teleconnections. *J. Clim.* **33**, 10489–10504 (2020).
9. Murakami, H. Substantial global influence of anthropogenic aerosols on tropical cyclones over the past 40 years. *Sci. Adv.* **8**, eabn9493 (2024).
10. Knutson, T. R. et al. Tropical cyclones and climate change. *Nat. Geosci.* **3**, 157–163 (2010).
11. Mendelsohn, R., Emanuel, K., Chonabayashi, S. & Bakkensen, L. The impact of climate change on global tropical cyclone damage. *Nat. Clim. Chang.* **2**, 205–209 (2012).
12. Walsh, K. J. E. et al. Tropical cyclones and climate change. *WIREs Clim. Chang.* **7**, 65–89 (2016).
13. Wu, L., Zhao, H., Wang, C., Cao, J. & Liang, J. Understanding of the effect of climate change on tropical cyclone intensity: a review. *Adv. Atmos. Sci.* **39**, 205–221 (2022).
14. Dunstone, N. et al. Towards useful decadal climate services. *Bull. Am. Meteorol. Soc.* **103**, E1705–E1719 (2022).

15. Smith, D. M. et al. Skilful multi-year predictions of Atlantic hurricane frequency. *Nat. Geosci.* **3**, 846–849 (2010).
16. Caron, L.-P. et al. How skillful are the multiannual forecasts of atlantic hurricane activity?. *Bull. Am. Meteorol. Soc.* **99**, 403–413 (2018).
17. Xu, Y., Wu, B., Hu, S. & Zhou, T. Skillful decadal prediction for Northwest Pacific tropical cyclone activity. *Clim. Dyn*. https://doi.org/10.1007/s00382-024-07281-4 (2024).
18. Zhang, R. et al. Decadal prediction of location of tropical cyclone maximum intensity over the Western North Pacific. *Geophys. Res. Lett.* **51**, e2023GL106746 (2024).
19. Vecchi, G. A. et al. Multiyear predictions of north Atlantic hurricane frequency: promise and limitations. *J. Clim.* **26**, 5337–5357 (2013).
20. Camp, J. et al. Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. *Q. J. R. Meteorol. Soc.* **141**, 2206–2219 (2015).
21. Feng, X., Toumi, R., Roberts, M., Hodges, K. I. & Vidale, P. L. An approach to link climate model tropical cyclogenesis bias to large-scale wind circulation modes. *Geophys. Res. Lett.* **50**, e2023GL103838 (2023).
22. Emanuel, K., DesAutels, C., Holloway, C. & Korty, R. Environmental control of tropical cyclone intensity. *J. Atmos. Sci.* **61**, 843–858 (2004).
23. Boucharel, J., Jin, F.-F., Lin, I. I., Huang, H.-C. & England, M. H. Different controls of tropical cyclone activity in the Eastern Pacific for two types of El Niño. *Geophys. Res. Lett.* **43**, 1679–1686 (2016).
24. Vannière, B. et al. The moisture budget of tropical cyclones in HighResMIP models: large-scale environmental balance and sensitivity to horizontal resolution. *J. Clim.* **33**, 8457–8474 (2020).
25. Kim, D., Ho, C.-H., Murakami, H. & Park, D.-S. R. Assessing the influence of large-scale environmental conditions on the rainfall structure of Atlantic Tropical cyclones: an observational study. *J. Clim.* **34**, 2093–2106 (2021).
26. Sobel, A. H. et al. Tropical cyclone frequency. *Earth's Future* **9**, e2021EF002275 (2021).
27. Dai, Y., Majumdar, S. J. & Nolan, D. S. Tropical cyclone resistance to strong environmental shear. *J. Atmos. Sci.* **78**, 1275–1293 (2021).
28. Rios-Berrios, R. et al. A review of the interactions between tropical cyclones and environmental vertical wind shear. *J. Atmos. Sci.* **81**, 713–741 (2024).
29. Slocum, C. J., Razin, M. N., Knaff, J. A. & Stow, J. P. Does ERA5 mark a new era for resolving the tropical cyclone environment? *J. Clim.* **35**, 7147–7164 (2022).
30. Monerie, P.-A., Robson, J., Dong, B., Hodson, D. L. R. & Klingaman, N. P. Effect of the atlantic multidecadal variability on the global monsoon. *Geophys. Res. Lett*. **46**, 1765–1775 (2019).
31. Goldenberg, S. B., Landsea, C. W., Mestas-Nuñez, A. M. & Gray, W. M. The recent increase in Atlantic hurricane activity: causes and implications. *Science* **293**, 474–479 (2001).
32. Smith, D. M. et al. Robust skill of decadal climate predictions. *npj Clim. Atmos. Sci.* **2**, 13 (2019).
33. García-Serrano, J., Guemas, V. & Doblas-Reyes, F. J. Added-value from initialization in predictions of Atlantic multi-decadal variability. *Clim. Dyn.* **44**, 2539–2555 (2015).
34. Ruprich-Robert, Y. et al. Impacts of Atlantic multidecadal variability on the tropical Pacific: a multi-model study. *npj Clim. Atmos. Sci.* **4**, 33 (2021).
35. Monerie, P.-A., Robson, J., Dong, B. & Hodson, D. Role of the Atlantic multidecadal variability in modulating East Asian climate. *Clim. Dyn*. https://doi.org/10.1007/s00382-020-05477-y (2020).
36. Hsu, W.-C., Patricola, C. M. & Chang, P. The impact of climate model sea surface temperature biases on tropical cyclone simulations. *Clim. Dyn.* **53**, 173–192 (2019).
37. Zhang, W. et al. Dominant role of Atlantic multidecadal oscillation in the recent decadal changes in western north pacific tropical cyclone activity. *Geophys. Res. Lett.* **45**, 354–362 (2018).
38. Scaife, A. A. & Smith, D. A signal-to-noise paradox in climate science. *npj Clim. Atmos. Sci.* **1**, 28 (2018).
39. Kosaka, Y. et al. The JRA-3Q reanalysis. *J. Meteorol. Soc. Japan. Ser. II* **102**, 49–109 (2024).
40. Smith, D. M. et al. North Atlantic climate far more predictable than models imply. *Nature* **583**, 796–800 (2020).
41. Monerie, P.-A., Wilcox, L. J. & Turner, A. G. Effects of anthropogenic aerosol and greenhouse gas emissions on Northern Hemisphere monsoon precipitation: mechanisms and uncertainty. *J. Clim*. **35**, 1–66 (2022).
42. Feng, X., Klingaman, N. P. & Hodges, K. I. The effect of atmosphere–ocean coupling on the prediction of 2016 western North Pacific tropical cyclones. *Q.J.R. Meteorol. Soc.* **145**, 2425–2444 (2019).
43. Kuhlbrodt, T. et al. The low-resolution version of HadGEM3 GC3.1: development and evaluation for global climate. *J. Adv. Model. Earth Syst.* **10**, 2865–2888 (2018).
44. Hersbach, H. et al. The ERA5 global reanalysis. *Q.J.R. Meteorol. Soc.* **146**, 1999–2049 (2020).
45. Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J. & Neumann, C. J. The International Best Track Archive for Climate Stewardship (IBTrACS): unifying tropical cyclone data. *Bull. Am. Meteorol. Soc.* **91**, 363–376 (2010).
46. Cangialosi, J. P. et al. Recent progress in tropical cyclone intensity forecasting at the national hurricane center. *Weather Forecast* **35**, 1913–1922 (2020).
47. Kanamitsu, M. et al. NCEP–DOE AMIP-II reanalysis (R-2). *Bull. Am. Meteorol. Soc.* **83**, 1631–1643 (2002).
48. Bengtsson, L., Hodges, K. I. & Esch, M. Tropical cyclones in a T159 resolution global climate model: comparison with observations and re-analyses. *Tellus A Dyn. Meteorol. Oceanogr.* **59**, 396–416 (2007).
49. Hodges, K., Cobb, A. & Vidale, P. L. How well are tropical cyclones represented in reanalysis datasets? *J. Clim.* **30**, 5243–5264 (2017).
50. Roberts, M. J. et al. Impact of model resolution on tropical cyclone simulation using the HighResMIP–PRIMAVERA multimodel ensemble. *J. Clim.* **33**, 2557–2583 (2020).
51. Befort, D. J. et al. Combination of decadal predictions and climate projections in time: challenges and potential solutions. *Geophys. Res. Lett.* **49**, e2022GL098568 (2022).
52. ICPO. Data and bias correction for decadal climate predictions. *Int. CLIVAR Proj. Off.* **150**, 5 (2011).
53. Eade, R. et al. Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys. Res. Lett.* **41**, 5620–5628 (2014).
54. Gaitán, C. F. Effects of variance adjustment techniques and time-invariant transfer functions on heat wave duration indices and other metrics derived from downscaled time-series. Study case: Montreal, Canada. *Nat. Hazards* **83**, 1661–1681 (2016).
55. Saunders, M. A. & Lea, A. S. Seasonal prediction of hurricane activity reaching the coast of the United States. *Nature* **434**, 1005–1008 (2005).
56. Zarzycki, C. M., Ullrich, P. A. & Reed, K. A. Metrics for evaluating tropical cyclones in climate data. *J. Appl. Meteorol. Climatol.* **60**, 643–660 (2021).

Acknowledgements

The Singapore Green Finance Centre supported R.T. and X.F. X.F. and K.H. were supported by the UK Met Office Weather and Climate Science for Service Partnership for Southeast Asia as part of the Newton Fund. X.F., R.T., and K.H. were also supported by the UK NERC Large Grant project (NE/W009587/1). X.F. was also supported by the UK National Centre for Atmospheric Science through the NERC National Capability International Programmes Award (NE/X006263/1). R.T. thanks the Vodafone Foundation (DreamLab project). DreamLab users provided their phone computing power and time. P.A.M. was funded by the Natural

Environment Research Council (NERC) through the WISHBONE (NE/T013516/1) and ALPACA (NE/Y005279/1) projects. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies that support CMIP6 and ESGF. We thank the two anonymous reviewers for their comments and suggestions.

Author contributions

P.A.M. conceived the study, performed the analysis, and led the writing. K.H. performed the tracking. X.F., K.H., and R.T. contributed to the design of the study, discussed the results, and contributed to the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-025-00919-y.

Correspondence and requests for materials should be addressed to Paul-Arthur Monerie.