How well can global ensemble forecasts predict tropical cyclones in the southwest Indian Ocean?

Received: 20 July 2023 | Revised: 18 February 2024 | Accepted: 7 March 2024

DOI: 10.1002/met.2195

RESEARCH ARTICLE

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

How well can global ensemble forecasts predict tropical cyclones in the southwest Indian Ocean?

R. Emerton[1,2] | K. I. Hodges[2] | E. Stephens[3,4] | V. Amelie[5] | M. Mustafa[6] | Z. Rakotomavo[7] | E. Coughlan de Perez[4,8] | L. Magnusson[1] | P.-L. Vidale[2,3]

[1]European Centre for Medium-Range Weather Forecasts, Reading, UK

[2]National Centre for Atmospheric Science, University of Reading, Reading, UK

[3]Department of Meteorology, University of Reading, Reading, UK

[4]Red Cross Red Crescent Climate Centre, The Hague, The Netherlands

[5]Seychelles Meteorological Authority, Mahe, Seychelles

[6]Instituto Nacional de Meteorologia, Maputo, Mozambique

[7]Météo Madagascar, Antananarivo, Madagascar

[8]Feinstein International Center, Tufts University, Medford, Massachusetts, USA

Correspondence

R. Emerton, European Centre for Medium-Range Weather Forecasts, Reading, UK.
Email: rebecca.emerton@ecmwf.int

Funding information

Natural Environment Research Council, Grant/Award Number: NE/S005897/1

Abstract

The southwest Indian Ocean (SWIO) recently experienced its most active, costliest and deadliest cyclone season on record (2018–2019). The anticipation and forecasting of natural hazards, such as tropical cyclones, are crucial to preparing for their impacts, but it is important to understand how well forecasting systems can predict them. Despite the vulnerability of the SWIO to tropical cyclones, comparatively little research has focused on this region, including understanding the ability of numerical weather prediction systems to predict cyclones and their impacts in southeast Africa. In this study, we evaluate ensemble probabilistic and high-resolution deterministic forecasts of tropical cyclones in the SWIO from 2010 to 2020, using two state-of-the-art global forecasting systems: one from the European Centre for Medium-Range Weather Forecasts (ECMWF) and the other from the U.K. Met Office. We evaluate predictions of the track, assessing the location of the centre of each storm and its speed of movement, as well as its intensity, looking at maximum wind speeds and minimum central pressure, and discuss how the forecasts have evolved over the 10-year period. Overall, ECMWF typically provides more accurate forecasts, but both systems tend to underestimate translation speed and intensity. We also investigate the impact of the Madden-Julian Oscillation (MJO) on tropical cyclones and their forecasts. The MJO impacts where and when tropical cyclones form, their tracks and intensities, which in turn impacts forecast skill. These results are intended to provide an increased understanding of the ability of global forecasting systems to predict tropical cyclones in the SWIO, for the purpose of decision making and anticipatory action.

KEYWORDS

decision-making, ensemble forecasting, forecast skill, high impact weather, Indian Ocean, MJO, numerical weather prediction, tropical cyclones

1 | INTRODUCTION

The southwest Indian Ocean (SWIO) recently experienced its most active, costliest and deadliest tropical cyclone (TC) season on record (2018–2019). TCs in this region can impact countries along the eastern coast of Africa and island nations in the southern Indian Ocean. In 2019, two devastating TCs, Idai and Kenneth, impacted Mozambique just 6 weeks apart, leading to catastrophic flooding and loss of life, and leaving millions of people in need of humanitarian assistance (Emerton et al., 2020). Two cyclone landfalls in Madagascar in 2018 displaced tens of thousands of people. Cyclones in the region can also affect the Seychelles archipelago, such as the category 5 Cyclone Fantala in 2016, and the Mascarene islands including La Réunion and Mauritius. Many more TCs have impacted these countries in recent years.

The anticipation and forecasting of natural hazards such as TCs are crucial to preparing for their impacts. But it is important to understand how well forecast models are able to predict these events, and the limitations of the forecasts. Despite the vulnerability of this region to TCs and their related hazards, and despite having similar TC activity to the north Atlantic Ocean basin, comparatively little research has focussed on TCs in the SWIO, including understanding the ability of numerical weather prediction systems to predict cyclones and their impacts in southeast Africa (Leroux et al., 2018; WMO, 2017). Although numerical weather prediction systems are continuously improving thanks to ongoing model development efforts, the use of ensemble forecasts is important in capturing and communicating uncertainty in TC predictions. Rather than producing a single forecast scenario, ensemble forecasts produce a range of possible forecast scenarios in order to capture the uncertainty stemming from the chaotic nature of the atmosphere, and from imperfections in models of the atmosphere and Earth System.

In the SWIO, the Regional Specialised Meteorological Centre (RSMC) responsible for disseminating TC forecasts and warnings is Météo France in La Réunion (RSMC-LR). Recent research by Bousquet et al. (2020) and Bonnardot et al. (2018) has highlighted and evaluated new forecasting approaches specific to the region, such as a convection-permitting, limited-area, short-range, high-resolution model which is capable of providing improved intensity forecasts, and a prototype ensemble prediction system (AROME-IO; Bousquet et al., 2020). Bonnardot et al. (2018) describe a methodology that aims to improve the representation of forecast uncertainty, combining the RSMC-LR's official deterministic forecast with associated probabilities based on scenarios from climatology and the European Centre for Medium-Range Weather Forecasts' (ECMWF) ensemble prediction system (EPS). While the RSMC-LR forecasts present the official forecast guidance and warnings, these forecasts take into consideration the model output from global numerical weather prediction models, and forecasters at the national meteorological and hydrological services often follow and utilise the TC forecasts provided by global models alongside the guidance from the RSMC-LR. Two of the most widely used forecast models for this purpose are those from ECMWF and the U.K. Met Office (UKMO).

At the UKMO, ensemble probabilistic forecasts from three global forecasting centres (the UKMO global and regional models, ECMWF and the National Centre for Environmental Prediction [NCEP] global forecast system [GEFS]) are combined to produce a multi-model forecast. Titley et al. (2020) evaluated these forecasting systems and their combined multi-model forecast for a 2-year period across all ocean basins globally, and highlighted that the best performing individual EPS varies from basin to basin (and storm to storm). The Indian Ocean was found to have poorer skill than other basins in both the ECMWF and UKMO forecasts, and the multi-model ensemble (Titley et al., 2020).

In this study, we present a longer term evaluation of both the high-resolution deterministic (HRES) and ensemble probabilistic (ENS) TC forecasts from the UKMO and ECMWF state-of-the-art numerical weather prediction systems, over a 10-year period from 2010 to 2020 in the SWIO. The official forecast guidance from the RSMC-LR is also discussed and compared where data were available. We focus on the ability to predict the track, in terms of both the location of the centre of the storm and the speed of movement of the storm, and the intensity, looking at both the maximum wind speeds and the minimum central pressure. The evaluation was developed in collaboration with meteorologists at national meteorological services operating in the SWIO region, and with those working in humanitarian anticipatory action, in order to provide information that is useful for decision makers tasked with providing forecasts and warnings and preparing for TCs. We aim to provide an increased understanding of forecast accuracy in predicting TC track and intensity at lead times up to a week ahead, based on a large sample of 94 TCs and thousands of forecasts over the past decade (up to ~1400 deterministic forecasts and ~60,000 ensemble forecast members per model, depending on the model and lead time; see Figure 4c,f). This large sample of forecasts from state-of-the art numerical weather prediction systems is investigated using a tracking and evaluation methodology that allows robust results that are fully comparable across forecasting systems.

We also investigate the impact of the Madden-Julian Oscillation (MJO) on TCs and their forecasts. The MJO is a major fluctuation in tropical weather on weekly to monthly timescales. It is a key driver of TC formation in the SWIO, and therefore may provide 'windows of opportunity' for more accurate forecasts, and longer lead times for useful predictions, depending on the large-scale tropical conditions. We discuss the modulation of TC occurrence by the MJO, and present a conditional verification of the ECMWF and UKMO forecasting systems.

Information on TC forecast accuracy, as well as factors influencing this accuracy, is key for operational meteorologists, humanitarian organisations and other decision makers tasked with using such information to take early action ahead of an event—when should we trust the forecasts, and when should we not?

2 | CLIMATOLOGY OF TROPICAL CYCLONES IN THE SOUTHWEST INDIAN OCEAN

Over the 10-year period from July 2010 to June 2020, a total of 94 TCs moved through the SWIO (Figure 1a), of which 31 impacted Madagascar, Mozambique or the Seychelles (Table 1). Between 5 and 15 cyclones occurred in any given year, and cyclones in this region typically occur between November and April, with the majority of storms forming between December and February. Of the 94 TCs, 21 were very intense TCs (Figure 2a,b), with a maximum wind speed exceeding 212 km/h. Almost half of these (10) occurred in the latter three seasons of the study period, between 2017 and 2020, with eight very intense TCs during the 2018–2019 season alone. Aside from the potentially dangerous impacts of TCs from their associated hazards, including wind, rainfall, flooding and storm surge, these cyclones are also an important source of precipitation for the region.

Over much of Mozambique, up to 10% of the annual total rainfall comes from TCs, and in parts of Madagascar including the southwest and northeast the total contribution of TCs to the total annual rainfall is up to 30% in some small regions and up to 25% more widely (Figure 1b). In February 2000, Cyclone Eline led to severe flooding in Mozambique, Zimbabwe and South Africa, but it also 'contributed 25% of the semi-arid Namibian rainfall during that summer season' (Muthige et al., 2018).

The MJO is a mode of intra-seasonal tropical climate variability (i.e., it varies on a week-to-week basis) (Madden & Julian, 1971, 1972). It is a dipole of enhanced and suppressed convection that moves eastward through the tropics, taking ~30–90 days to traverse the globe. The MJO is considered to be active when this dipole of convection is present and moving eastward with time.

Eight phases are used to describe the location of the MJO, based on the location of the area of active convection (Wheeler & Hendon, 2004). Many aspects of tropical and extra-tropical weather can be influenced by the MJO, including the strength and timing of monsoons and changes in jet streams that can result in MJO teleconnections to the mid-latitudes. The MJO also impacts the genesis, timing and intensity of TCs in most ocean basins. This link with TCs was first highlighted in 1963 in a paper that described what is now known as the MJO (Xie et al., 1963; (in Chinese); Xie et al., 2018 (in English); see Li et al., 2018 for more information).

The majority of cyclones in the SWIO form when the MJO is active in phases 2–5, but they can occur in any MJO phase (Figure 3) or when there is no active MJO. TCs in this region are also more likely to be intense, or to rapidly intensify, when they form in phases 2–6. However, in phases 3–5, many TCs tend to occur further east, posing less of a threat to land. It is important to note that this is not always the case, however, and Cyclone Idai was the first storm to reach TC strength in the Mozambique Channel in phases 3–5 (Idai formed during phase 4). While TCs in the SWIO generally move from east to west/southwest, TCs that form in phase 6 tend to move to the southeast (Figure 3).

3 | FORECASTING SYSTEMS

This section provides an overview of the forecasting systems evaluated in this study: the HRES and ENS forecasts from the ECMWF and the UKMO, respectively. Table 2 summarises the key characteristics of these forecasting systems, including significant changes made to the systems in terms of their horizontal resolution, during the 2010–2020 period of the study. The forecasts available for the region from the RSMC-LR are briefly described, followed by information about the ECMWF and UKMO forecasting systems.

3.1 | La Réunion Regional Specialised Meteorological Centre

RSMCs have the World Meteorological Organisation (WMO)-mandated responsibility to monitor and name TCs in their region and provide forecasts to national hydro-meteorological services. RSMC-LR provides daily updates on the meteorological situation and potential for cyclogenesis. They issue technical bulletins and graphical warning products every 6 h during a TC. The graphical

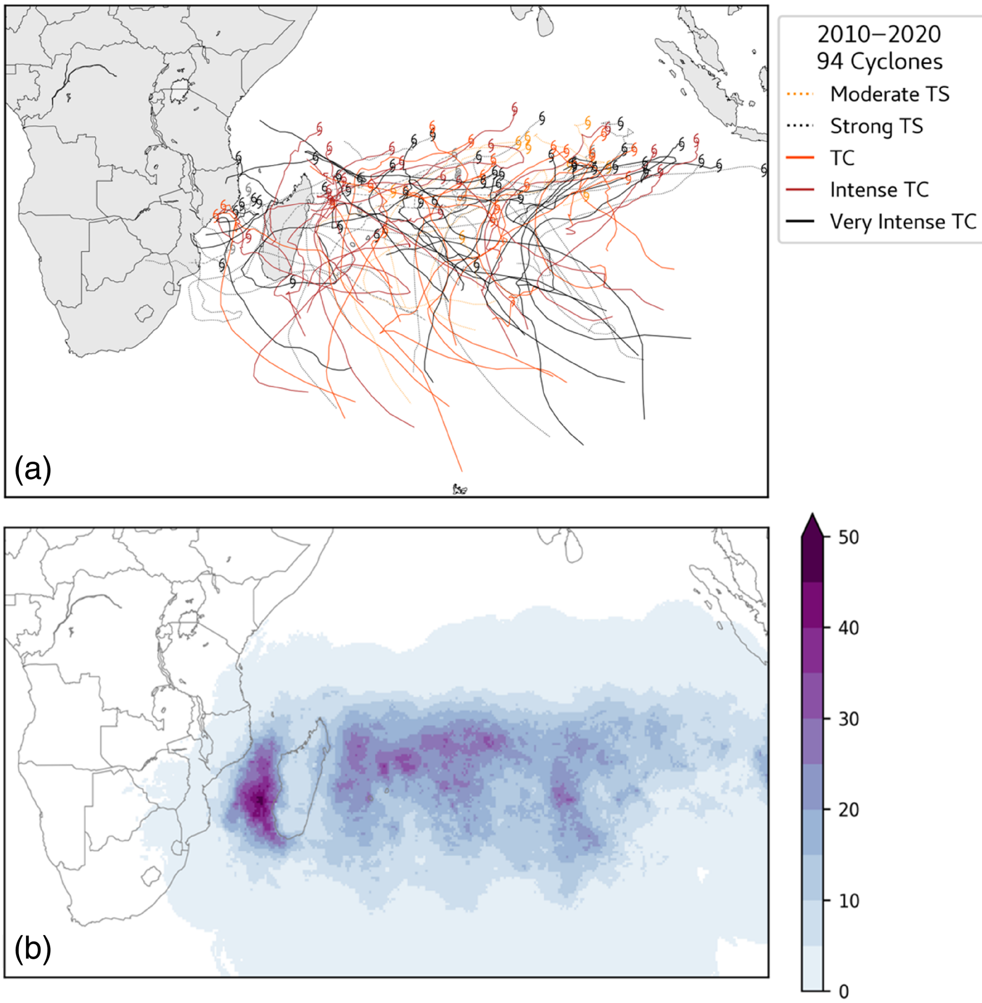

FIGURE 1 (a) Tracks and maximum intensity of 94 tropical cyclones in the southwest Indian Ocean from 2010 to 2020, using the International Best Track Archive for Climate Stewardship (IBTrACS) track database. Intensity categories are based on the cyclone warning scale used by the La Réunion Regional Specialised Meteorological Centre, where TS indicates tropical storm and TC tropical cyclone. Cyclone symbols indicate the start of the track. (b) Percentage contribution of tropical cyclones to the total annual rainfall, based on TRMM satellite data from 2006 to 2018.

TABLE 1 Number of tropical cyclones in each intensity category from July 2010 to June 2020 (based on the maximum wind speed for each tropical cyclone in IBTrACS, and using the cyclone warning scale used by the La Réunion Regional Specialised Meteorological Centre) and number of tropical cyclones in each intensity category that made landfall in Madagascar, Mozambique and the Seychelles from 2010 to 2020.

Intensity category	Total	Impacted Madagascar	Impacted Mozambique	Impacted Seychelles
Moderate TS	9	0	0	0
Strong TS	21	4	4	1
TC	18	1	1	0
Intense TC	17	7	1	2
Very Intense TC	21	1	4	2
Total	94	13	10	5

The colours used in Table 1 correspond to the colours used by the La Réunion Regional Specialised Meteorological Centre in their warnings and communications, for each TC intensity category.
Abbreviations: TC, tropical cyclone; TS, tropical storm.

warning products are issued through the Météo-France website (www.meteofrance.re/cyclone/) and provide maps of the predicted track of the centre of the tropical system out to 5 days ahead, including a 'potential track area' (sometimes known as a 'cone of uncertainty'). They also indicate the predicted intensity of the storm. The technical bulletins contain detailed information on the location, size and intensity of the tropical system in text format designed for the use of operational forecasters at the national authorities. These forecasts are produced

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

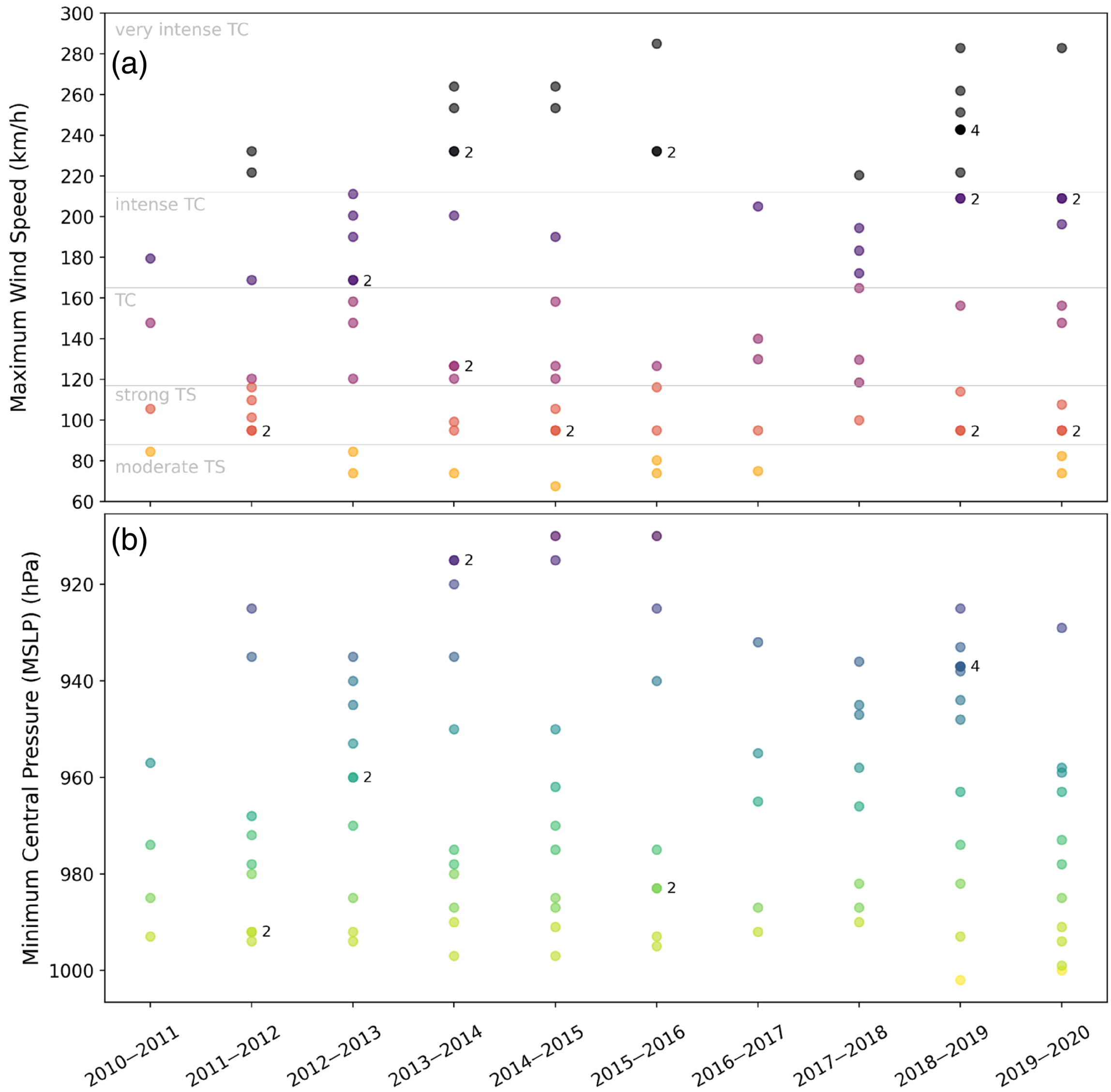

FIGURE 2 Maximum intensity of each tropical cyclone included in the study, per season from 2010 to 2020, in terms of (a) maximum wind speed and (b) minimum central pressure. Data are taken from the IBTrACS dataset (see Section 4). Numbers are used to highlight where multiple storms were recorded at the same maximum intensity; for example, four storms occurred with a maximum wind speed of ~243 km/h in 2018–2019. Horizontal lines shown in (a) indicate the thresholds between the tropical cyclone intensity categories used by the La Réunion Regional Specialised Meteorological Centre. MSLP, mean sea-level pressure; TC, tropical cyclone; TS, tropical storm.

by operational forecasters, based on forecasts from a range of numerical weather prediction models, and include the expertise and interpretation of the forecasters with local knowledge of the region, its weather and TCs.

3.2 | European Centre for Medium-Range Weather Forecasts

ECMWF runs several global-scale forecasting system components as part of their operational Integrated Forecasting System (IFS), providing high-resolution deterministic forecasts (up to 10 days ahead) and ensemble medium-range (up to 15 days ahead), extended-range (up to 46 days ahead) and seasonal (up to 7 and 13 months ahead) forecasts. For this study, we evaluated forecasts from the HRES model and the medium-range ensemble (ENS).

Both the HRES and the ENS are run four times per day (at 00, 06, 12, and 18 UTC). The HRES provides a single forecast solution, and the ENS has 51 ensemble members (1 control and 50 perturbed members), allowing for probabilistic forecasts representing the uncertainty. The ENS has typically been run at a lower resolution than the HRES, although the latest version of the IFS (implemented on 28 June 2023) sees the ENS and HRES forecasts

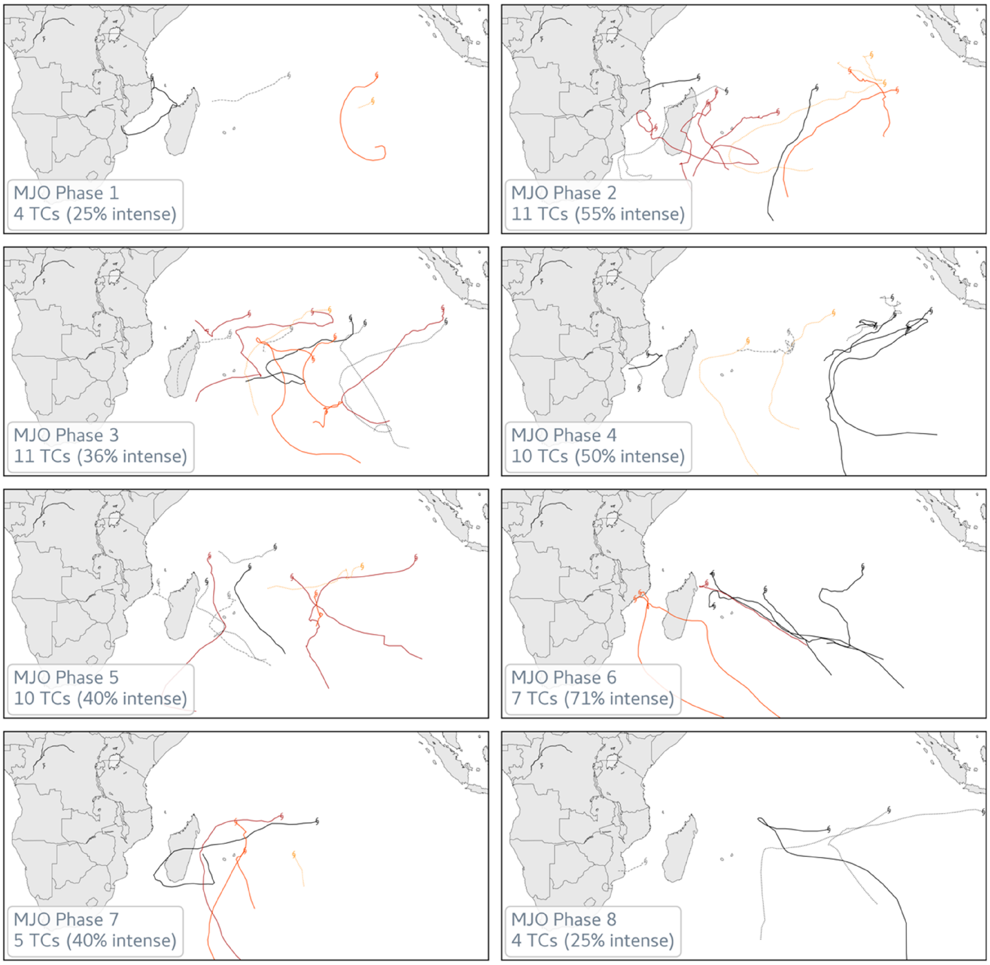

FIGURE 3 Tracks of observed tropical cyclones ('best track' data from IBTrACS) in the southwest Indian Ocean for tropical cyclones (TCs) with cyclogenesis occurring during each phase of the Madden-Julian Oscillation (MJO) (based on the MJO phase on the first day of the IBTrACS best track) for the 10-year period 2010–2020. Indicated on each map is the MJO phase, the number of TCs that formed during that phase and the percentage of those storms that went on to become intense or very intense TCs, although it is noted that these percentages are based on small sample sizes in some cases. Tracks are colour-coded according to their maximum intensity category, using the intensity scale of the La Réunion Regional Specialised Meteorological Centre: tropical depression (yellow dotted), moderate tropical storm (orange dotted), strong tropical storm (black dotted), tropical cyclone (orange), intense tropical cyclone (red), and very intense tropical cyclone (black).

run at the same ∼9 km horizontal resolution for the first time. During the 2010–2020 period of this study, ECMWF upgraded the horizontal resolution of the IFS once, in March 2016 (Table 2). This changed the HRES from ∼16 km horizontal resolution to ∼9 km, and the ENS from ∼32 to ∼18 km. A number of other changes have been made to the IFS during the study period, as the model is continuously being developed and improved. A breakdown of the evolution of the IFS is available via the ECMWF website (https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model).

A range of TC forecast products are openly available from ECMWF (https://www.ecmwf.int/en/forecasts/charts/latest-tropical-cyclones-forecast), including ensemble track and intensity forecasts for all current TCs, TC activity forecasts indicating potential TC activity at different time ranges in the medium- and extended-range forecasts, as well as seasonal forecasts of TC number, frequency and accumulated cyclone energy.

In this study, we analyse between 1200 and 1400 forecasts of TCs (including up to 60,000 individual ensemble members) produced between July 2010 and June 2020 from the 00 and 12 UTC HRES and ENS runs. The exact sample size varies by lead time and forecast type, depending on whether and when an individual forecast or ensemble member picked up the existence of a TC at different times. We also consider separately the control forecast, which is equivalent to the HRES forecast but run at the lower resolution of the ENS forecasting system. Although the operational forecasts are available out to

Meteorological Applications Science and Technology for Weather and Climate Open Access RMetS

TABLE 2 Key characteristics of the European Centre for Medium-Range Weather Forecasts (ECMWF) and U.K. Met Office (UKMO) high-resolution deterministic and ensemble forecasting systems during the 2010–2020 period of this study.

Forecasting system	Time period	Grid	Horizontal resolution at equator (~km)	Ensemble members
ECMWF deterministic	July 2010–March 2016	T_L1279	16	
	March 2016–July 2020	$T_{CO}1279$	9	
ECMWF ensemble	July 2010–March 2016	T_L639	32	51
	March 2016–July 2020	$T_{CO}639$	18	51
UKMO deterministic	July 2010–June 2014	N512	40	
	July 2014–June 2017	N768	26	
	July 2017–July 2020	N1280	16	
UKMO ensemble	July 2010–June 2014	N216	93	24
	July 2014–June 2017	N400	50	12
	July 2017–July 2020	N640	31	18

Note: Grid notation NX denotes a regular model grid with X latitude bands between the equator and poles. In the ECMWF Integrated Forecasting System (IFS), many calculations are not computed on a regular grid but are computed in 'spectral space', representing meteorological fields as a sum of wave functions, and then transformed back to grid-point space. T_LX represents the spectral transform grid with truncation at wavenumber X. $T_{CO}X$ represents the newer cubic-octahedral grid, which moved from a linear ($_L$) grid to a cubic ($_C$) grid. This increases the number of grid points representing each wavelength and therefore increases the effective resolution while keeping the number of spherical harmonics (X) constant (Malardel et al., 2016). Approximate horizontal resolution (km) is given at the equator. The resolution of ECMWF's ensemble prediction system is reduced from day 16, but only the first 7 days are considered in this study. The number of ensemble members shown includes one unperturbed control member alongside a number of perturbed members.

10 and 15 days, we analyse the first 7 days only, as this is the lead time for which the UKMO forecasts are also available.

3.3 | U.K. Met Office

The UKMO runs their forecasting system, called the Unified Model (UM), at global and regional scales at a range of timescales from medium-range to climate modelling, and with both deterministic and ensemble configurations. The global ensemble forecasting system (MOGREPS, referred to in this study as the UKMO ENS) and deterministic forecasting system (UKMO HRES) both provide forecasts out to 7 days ahead and are run four times per day, at 00, 06, 12 and 18 UTC. Other configurations provide predictions on seasonal (~7 months), decadal (~5 years) and climate timescales.

In this study, we use the UKMO ENS and HRES forecasts produced at 00 and 12 UTC from July 2010 to June 2020. This provides a sample of between 1200 and 1400 forecasts of TCs (including >20,000 individual ensemble members). As with the ECMWF forecasts, the sample size can vary depending on whether and when an individual forecast or ensemble member picked up the existence of a TC. During this period, the UKMO model also underwent a series of model upgrades and developments, including two upgrades to the horizontal resolution of the models, in July 2014 and July 2017. This took the ENS resolution from ~93 km in 2010 to ~31 km in 2020, and the HRES from ~40 to ~16 km. The number of ensemble members used in the ENS forecasting system also varies during the period of the study, between 12 and 24 (Table 2). We also consider separately the control forecast from the UKMO ENS, as with ECMWF.

The UKMO provide TC guidance messages twice per day based on the global forecasts, provided for information only to complement official forecasts and warnings from RSMCs (https://www.metoffice.gov.uk/research/weather/tropical-cyclones/warnings).

4 | TROPICAL CYCLONE TRACKING

This section describes the TC track database used to verify the forecasts (IBTrACS) and the tracking software used to identify TC tracks in the forecast model output from both ECMWF and UKMO.

4.1 | International Best Track Archive for Climate Stewardship

The IBTrACS dataset (Knapp et al., 2010, 2018) merges TC data from multiple agencies around the globe to create an openly available, unified global archive of TC 'best tracks'. The dataset provides maximum sustained windspeeds, minimum central pressure and the storm centre of circulation at 3-hourly intervals (mostly interpolated from 6-hourly reported data) for TCs in all ocean basins since 1841. 'Working best track' data is often provided by RSMCs in near-real time, but after a storm, it is possible to gather all available data and information about each TC and produce a best estimate of the track and intensity, that is, the 'best track'. IBTrACS provides working best tracks as provisional data, and the final best track data becomes available after a delay. Uncertainties in the data, due to changes in methodologies and the introduction of satellites, are larger prior to 1980. Since 2000, uncertainty in the intensity (in terms of wind speed) is typically $\leq \pm 10$ kn (~ 5 m/s, 18.5 km/h), and uncertainty in TC position can be 10–40 km depending on the intensity (NOAA, 2019). The position uncertainty is smaller for stronger storms, which have more well-defined eyes, whereas weaker storms with larger centres of circulation can be more difficult to identify resulting in larger uncertainty. The final best track data from IBTrACS are used to verify the forecasts in this study.

4.2 | Identifying the forecast tracks

Although the UKMO and ECMWF have their own TC tracking schemes used to provide their operational forecast products, in this study we use the TC tracking software of Hodges (1994, 1995, 1999) to identify TCs in the raw forecast data from each of the forecasting systems outlined in Section 3. This allows us to use a consistent track identification method across all the forecasts.

The method for identifying TCs in the forecast data is described in detail by Hodges and Klingaman (2019) and Hodges and Emerton (2015) and summarised here. For each forecast (twice per day from July 2010 to June 2020 for each of the ECMWF and UKMO ENS and HRES forecasting systems), the forecast relative vorticity (at 6-hourly intervals out to 7 days ahead) is vertically averaged between 850 and 600 hPa, and then spatially filtered to T63 resolution to remove high-spatial-frequency noise. All variability at wavenumbers ≤ 5 is also removed, to remove the large-scale background flow. Vorticity maxima (exceeding $5 \times 10^{-6}\ \mathrm{s}^{-1}$) are identified and tracked in the filtered forecast data, and retained if they last for at least 2 days. This provides the latitude and longitude of the centre of the TC throughout the forecast horizon at 6-hourly intervals. The full-resolution maximum 10-m wind within a 6° radius of each track point and the minimum sea-level pressure within a 5° radius are added to the dataset to measure the intensity of the storms.

To evaluate the forecasts, the forecast tracks are matched against IBTrACS. If the tracks have a mean separation of <4° for the first day of the forecast track (which might not be the first day of the forecast itself, as the storm may form later in the forecast horizon), they are confirmed to be the same TC. This ensures that we are identifying the same TCs in the forecasts as those that were observed and exist in the IBTrACS dataset and that we are always evaluating the forecasts against the correct identical observed TC.

5 | TRACK AND INTENSITY PREDICTION SKILL

In this section, we evaluate and compare the skill of the UKMO and ECMWF prediction systems over the 10-year period from July 2010 to June 2020. We focus on the ability to predict the track, in terms of both the location of the centre of the storm and the speed of movement of the storm, as well as the intensity, looking at both the maximum wind speeds and the minimum central pressure. For both prediction systems, skill information is calculated and shown for the HRES forecasts, and for the ensemble prediction systems it is further broken down to provide statistics for the control (unperturbed) forecast, the ensemble mean forecast (where a deterministic forecast is produced by taking the mean of all the ensemble members) and all individual ensemble members. The ensemble mean is an oft-used tool for assessing the most likely outcome of an ensemble forecast; however, it may not represent a physically likely state of the atmosphere. The ensemble mean highlights, instead, the predictable elements of the forecast and smooths out the less predictable details, meaning it may not capture the risk of extreme events.

This evaluation is also compared, where data were available, to the track forecast skill of the RSMC-LR's official track forecast guidance for the region. The evaluation, including the metrics used and the presentation of the results, was developed in collaboration with meteorologists at national meteorological services operating in the SWIO region, and with those working in humanitarian anticipatory action, in order to provide information that is useful for decision makers tasked with providing forecasts and warnings, and preparing for TCs.

We further break down this analysis to look at changes in the skill of the ECMWF and UKMO forecasting systems over the 10-year period, using the same evaluation metrics. Both forecasting centres have increased the horizontal resolution during this time, alongside various other model changes, representing a step change in being able to better resolve features such as TCs, which are often small compared to the resolution of models and therefore challenging to accurately simulate.

Meteorological Applications
Science and Technology for Weather and Climate
Open Access
RMetS

5.1 | Model comparison

Although there are similarities in the forecast performance of the models for predicting TCs in the SWIO, on average ECMWF forecasts of track and intensity tend to be more accurate, particularly in terms of track forecasting (Figure 4). The UKMO track forecasts are slightly more accurate for the first day of lead time, but beyond ~1.5 days ahead, ECMWF track forecasts perform better (Figure 4a). This is likely due to differences in the data used in the forecast initialisation. In the UKMO model, when a TC already exists at the time the forecast is initialised (as opposed to a TC that forms later in the forecast horizon), the working Best Track (see Section 4.1) estimates are assimilated. This assimilation scheme (Heming, 2016a) has been used by the UKMO since 2015, and was found to reduce forecast errors at all lead times. These estimates include their own uncertainties given their subjective nature, and recent research at ECMWF following the method of Heming (2016a, 2016b) found no statistically significant and consistent results from including the working Best Track estimates in the IFS's data assimilation, except for improvements for 'a small sample of strong TCs in the north-west Pacific' (Magnusson et al., 2021).

At 1 day ahead, typical track location errors (i.e., the difference in km between the location of the centre of the TC in the forecast and the eventual observed location of the TC) for both systems are ~80 km (Figure 4a). This rises to ~200 km (ECMWF)/~250 km (UKMO) by 3 days ahead, and ~600 km (ECMWF)/~750 km (UKMO) by 7 days ahead, in the HRES forecasts (Figure 4a). At lead times beyond ~4 days ahead, the ENS mean becomes more accurate than the HRES forecasts, but these deterministic forecasts are always most useful in combination with an ensemble of forecasts, providing a representation

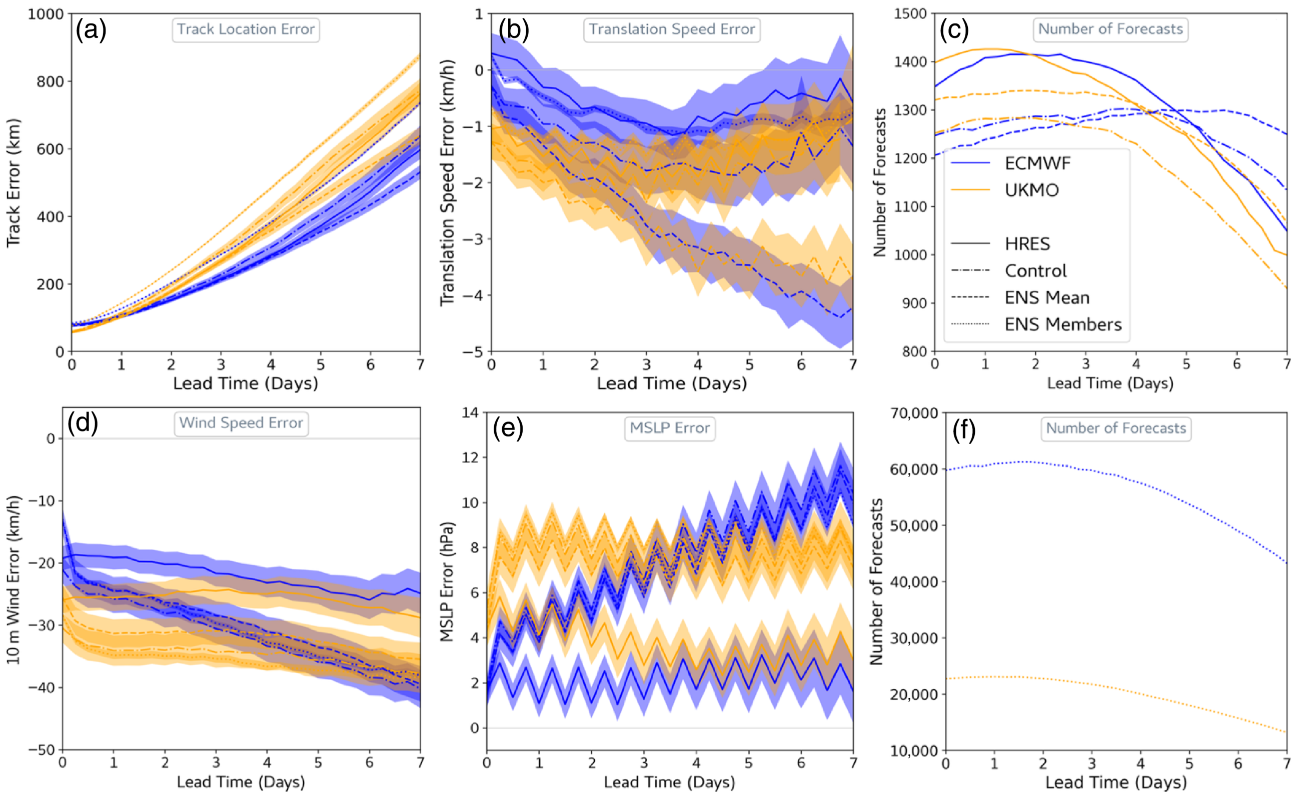

FIGURE 4 2010–2020 comparison of the European Centre for Medium-Range Weather Forecasts (ECMWF) (blue) and U.K. Met Office (UKMO) (yellow) tropical cyclone forecasts at lead times 0 to 7 days ahead, verified for the high-resolution deterministic (HRES) (solid lines), control (dot-dashed), ensemble mean (dashed), and ensemble members (dotted). Shading indicates the 95% confidence interval. (a) Track location error, that is, the distance in km between the forecast location of the centre of the tropical cyclone (TC), and the observed location; (b) translation speed error, that is, the difference between the predicted speed of movement of the TC and the observed speed of movement; (c) number of forecasts included in the statistics, for the HRES, control, and ensemble mean forecasts; (d) wind speed error, that is, the signed difference between the predicted maximum wind speed of the TC and the observed maximum wind speed, where negative values indicate that the wind speeds were under-predicted by the forecasts; (e) mean sea-level pressure (MSLP) error, that is, the signed difference between the predicted minimum pressure and the observed minimum pressure, where positive values indicate that the intensity of the TC was too weak in the forecasts; and (f) number of forecasts included in the statistics, for the ensemble member forecasts.

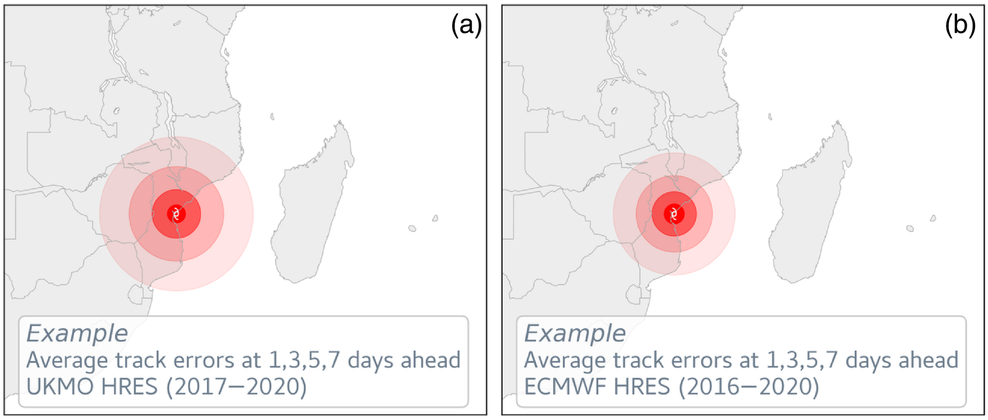

FIGURE 5 Example of the average track location error distances for the most recent versions of the (a) U.K. Met Office (UKMO) (2017–2020) and (b) European Centre for Medium-Range Weather Forecasts (ECMWF) (2016–2020) high-resolution deterministic (HRES) forecasts evaluated in this study, shown against the Mozambique coastline. The white cyclone symbol is representative of the 'observed' location, the red circles indicate the 1-, 3-, 5-, and 7-day-ahead errors with increasing radius from the centre and from darkest to lightest shading.

of the uncertainty in the predictions. It is also important to note that the ENS mean does not represent a forecast scenario itself but highlights the predictable signal. To better provide a visual representation of the scale of these track errors, the average errors in the HRES forecasts (using a sample of forecasts from only the most recent model resolutions (~16 km for UKMO, ~9 km for ECMWF)) at 1, 3, 5 and 7 days ahead is shown relative to the coastline of Mozambique (Figure 5). This highlights the potential implications for decision making around the landfall location and inland track of TCs at various lead times and for communication of forecast uncertainty.

Both forecasting systems underestimate the translation speed of TCs in this region, that is, predictions tend to indicate the TCs to be moving too slowly along their tracks (Figure 4b). Errors in the translation speed for TCs can have implications for planning for the timing of landfall and impact, and for predicting TC hazards such as rainfall and flooding. Storms that are slower moving tend to remain in one place for longer; therefore, more rain falls in one location resulting in a higher likelihood of flooding than storms that move faster (Titley et al., 2021). Beyond the implications for decision making related directly to TCs, the slow translation speed bias can also affect forecasts of extra-tropical transition of TCs (Froude, 2010, 2011) and downstream impacts on the mid-latitudes in some basins.

ECMWF HRES forecasts also provide a slightly more accurate representation of translation speed, but there is still an underestimation beyond 1 day ahead. The UKMO forecasts have a similar translation speed error as the ECMWF ensemble control forecast. The translation speed errors of the ensemble mean forecasts appear worse than the deterministic, control and individual ensemble members, particularly at longer lead times, due to the effect of taking the mean from a large ensemble spread before computing the errors. Averaging the movement of an ensemble of tracks, if these diverge significantly in direction, can make the translation speed (i.e., along-track movement) appear to be very small or even zero, thereby giving large errors.

The intensity of TCs in this region, in terms of both maximum wind speed and minimum central pressure, is also underestimated in both forecasting systems. On average, maximum wind speeds are too low, and minimum central pressures are too high (Figure 4d,e). High-resolution forecasts tend to be more accurate at predicting TC intensity than ensemble forecasts because they are better able to resolve the smaller structures. This is reflected in the results, with the ECMWF HRES forecast providing the most accurate intensity forecasts, followed by the UKMO HRES beyond day 2 (Figure 4d,e). Again, these deterministic forecasts are most useful in combination with an ensemble forecast, allowing assessment of the intensity from a higher resolution system, alongside representation of the uncertainty in the forecasts.

The semi-diurnal oscillations apparent in the MSLP errors (Figure 4e) are due to the handling of the atmospheric tide. This is a tidal effect 'associated with the absorption of solar radiation by ozone and water vapour, and heating from the surface' (Hodges & Klingaman, 2019), which shows up in the SLP as internal gravity waves (Dai & Wang, 1999). The magnitude of these oscillations, ~1 hPa, is similar to those found from direct surface observations (Dai & Wang, 1999). This effect is discussed further in a recent study by Hodges and Klingaman (2019).

Beyond statistics for a large number of TCs over the 10-year period, the skill of forecasts can vary from storm

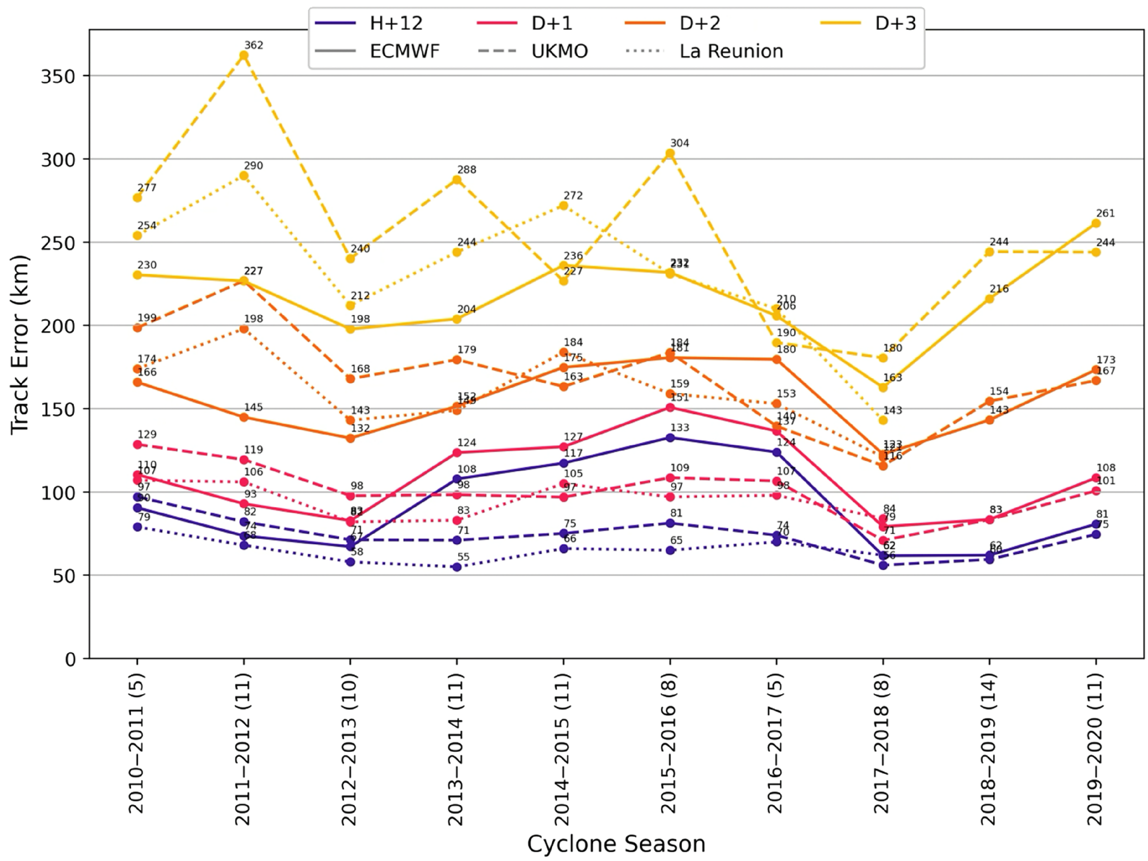

FIGURE 6 2010–2020 comparison of the track location errors during each cyclone season, for the European Centre for Medium-Range Weather Forecasts (ECMWF) and U.K. Met Office (UKMO) high-resolution deterministic forecasts and the La Réunion Regional Specialised Meteorological Centre (RSMC-LR) track forecasts, at lead times of 12 h (purple), 1 day (pink), 2 days (orange), and 3 days (yellow) ahead. Reproduced based on skill information available from RSMC-LR (available up to the 2017–2018 season, but no data were available for the 2018–2019 and 2019–2020 seasons).

to storm and year to year. In general, forecast models are becoming more accurate with time as they are constantly developed, updated, and improved (Figure 6). However, there are also occasions where the skill of one or more models may decrease for a particular year or cyclone season. Where multiple models experience a reduction in forecast skill, this can be due to a lack of predictability or challenging forecast situation, or even influenced by a particularly bad forecast for one storm in cases where sample sizes for a season are small. For example, the track errors of both the ECMWF and UKMO high-resolution forecasts were larger in the 2018–2019 and 2019–2020 seasons, following a period where errors had been reducing, particularly at longer lead times (Figure 6). The seasons from 2015–2016 to 2017–2018 were all below average for the number of cyclones that occurred (8, 5 and 8, respectively), while 2018–2019 and 2019–2020 were above average, with 14 and 11 TCs (Figure 2). During these two later seasons, alongside a larger number of TCs, there were also several very intense TCs (nine, compared to four during the previous three seasons; Figure 2a) and cyclones with challenging tracks (e.g., Cyclone Idai, March 2019; Emerton et al., 2020). TCs with tracks that recurve have been shown to be more challenging to predict than those with more zonal tracks (Hodges & Emerton, 2015); and for the period evaluated in this study, intense and very intense TCs showed some of the largest track errors, particularly at longer lead times (although with smaller sample sizes; not shown). The 2018–2019 season was the most active and deadliest cyclone season on record in this region, with 10 intense or very intense TCs. Research by Terry et al. (2013) found that the early months of the SWIO cyclone season (September to December) tend to produce more zonal, and therefore more predictable, tracks, while tracks that are more sinuous/meandering, and therefore more challenging to forecast, are more common in January.

Over the 10 seasons, the official forecast track guidance provided by the RSMC-LR has typically been more accurate at shorter lead times than the individual forecast models (Figure 6). The official track guidance is produced

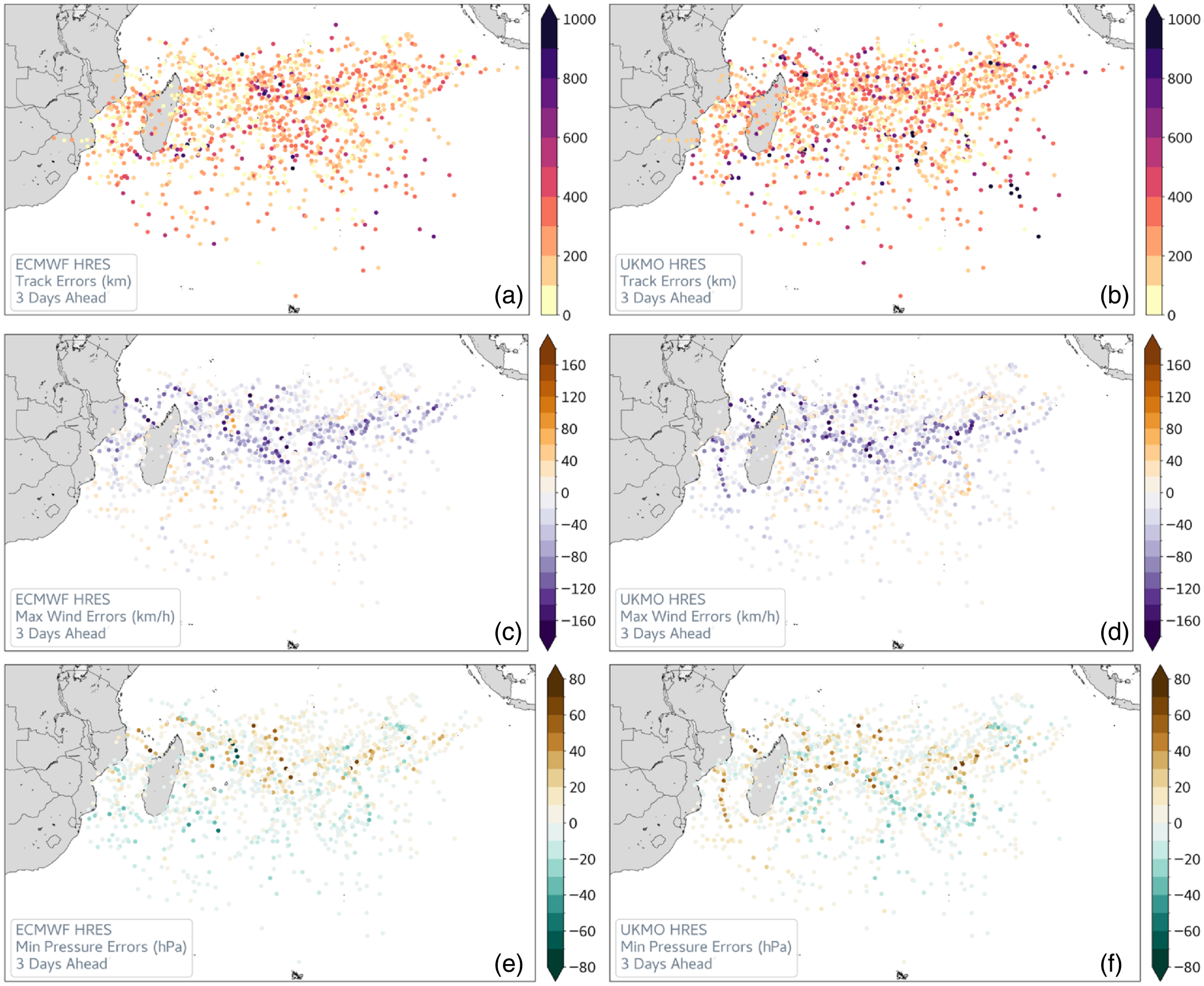

FIGURE 7 Spatial representation of the tropical cyclone track and intensity errors for all 94 storms between 2010 and 2020, in the European Centre for Medium-Range Weather Forecasts (ECMWF) (left) and U.K. Met Office (UKMO) (right) high-resolution deterministic forecasts at a lead time of 3 days ahead. Each dot indicates the error of one forecast, and the location shown is the observed location of the tropical cyclone at the time the forecast is valid. (a, b) Track location errors, km; (c, d) maximum wind speed errors (km/h), where purple indicates that the intensity is underestimated, that is, the wind speeds are too low in the forecast, and orange indicates the intensity is overestimated, that is, the wind speeds are too high in the forecast; (e, f) minimum mean sea-level pressure errors (hPa), where green indicates that the intensity is overestimated, that is, the minimum pressure is too deep in the forecast, and brown indicates the intensity is underestimated, that is, the pressure is too high in the forecast.

by forecasters based on model output from several different forecast models, with the expertise and local knowledge of the forecasters producing the guidance and warnings. At lead times beyond 1 day ahead, the error of the RSMC-LR's forecast tracks tend to fall between those of the ECMWF and UKMO models (Figure 6).

Forecast skill can also vary spatially, which can be important for decision makers to be aware of. As highlighted above, TCs that meander or recurve have been shown to be more challenging to forecast (Hodges & Emerton, 2015). In the SWIO, the largest track forecast errors occur towards the northern and southern boundaries of the basin, where TCs may be more likely to recurve (an example for 3 days ahead is shown in Figure 7a,b). This includes to the north and south of Madagascar, while track errors may be lower to the east of Madagascar and in the Mozambique Channel, although there is no significant difference in the track errors in the Channel compared to much of the rest of the SWIO.

Wind speed errors are typically largest in the north and centre of the SWIO and in the Mozambique Channel, and smaller towards the southern part of the basin (Figure 7c,d). Any overestimation of the wind speeds tends to occur in the southern part of the basin, where storms tend to be weaker, and in some cases along the east coast of Madagascar (Figure 7c,d). The intensity in terms of the minimum pressure is typically underestimated (Figure 7e,f), and this is most evident across the northern and central parts of the SWIO, and in the UKMO HRES forecasts in the Mozambique Channel (Figure 7f). MSLP errors appear to be slightly smaller, and slightly overestimated, in much of the Mozambique Channel in the ECMWF HRES forecasts, and, as with the UKMO HRES forecasts, the pressure tends to be overestimated across central and northern parts of the basin and too deep towards the southern edges of the basin, where TCs are typically weaker. Variations are shown for 3 days ahead in Figure 7, and other lead times are shown in the Supplementary Material (Figures S1–S3).

5.2 | Changes in skill with updates to model resolution

Numerical weather prediction models are constantly under development to further improve their representation of the Earth System and the accuracy of the forecasts. Over the 10-year period of this study, both the ECMWF and UKMO forecasting systems have undergone several significant model upgrades, outlined in Section 3 and discussed further in the following subsections. The impact of increasing model resolution, in comparison to other model upgrades, is an active area of research within the scientific community. Another key area of research is the use of stochastic physics as an additional alternative approach to improve the systematic underestimation of TC intensity. Stochastic physics parameterisation schemes aim to account for the missing non-linear interactions between processes that are not resolved in current numerical weather prediction and climate models, as well as the associated interactions between scales (Vidale et al., 2021). Both higher resolutions and the use of stochastic physics can be used to reduce model uncertainty and improve reliability, but the latter may be able to provide some of the benefits of increased resolution for a fraction of the resources (Vidale et al., 2021). Particularly with reduced precision, this has provided the opportunity to increase ensemble sizes, which is further beneficial for accurately representing forecast uncertainty. Vidale et al. (2021) examined the impact of stochastic physics for simulating TCs in general circulation models and found that stochastic physics was able to provide the same benefit as a 50% increase in resolution.

Although it is not possible to diagnose the impact of every model change over the past decade on TC skill, here we quantify the overall improvement in skill for the two forecasting systems and focus on the periods over which the models were run at different resolutions.

5.2.1 | U.K. Met Office

The UKMO forecasting system's horizontal resolution was upgraded twice during the 2010–2020 period of this study, in 2014 and 2017 (Table 2). The forecast dataset was split into three distinct periods (the 2010/2011–2013/2014 cyclone seasons, 2014/2015–2016/2017 seasons and 2017/2018–2019/2020 seasons) in order to evaluate any changes in skill with increases in the horizontal resolution of the models. In 2014, alongside the increased horizontal resolution, the model upgrade also included a new dynamical core, changes to the model physics and introduction of new satellite data. This combination of changes was shown to result in major improvements to TC track and intensity predictions (Heming, 2016a).

The track errors have improved for both the HRES and ENS mean forecasts (Figure 8a,e). The track forecasts have effectively gained 1–1.5 days of lead time throughout the 10-year period; the errors that we see at 5 days ahead in the 2017–2020 seasons are the same as the errors we saw at 3.5 days ahead in the system that was operational in 2010–2014 (Figure 8a,e). Translation speed errors are variable and the speed is still generally underestimated, but both have improved at longer lead times (Figure 8c,g).

In general, the intensity errors improved following the resolution upgrade in 2014, but were worse again, particularly for the maximum wind speed, after the upgrade in 2017 (Figure 8b,d,f,h). In the HRES forecasts, the maximum wind speed errors indicated an average underestimation of the maximum wind speeds of ~30 km/h during the 2010–2014 period, which was reduced to ~15 km/h during 2014–2017. In the more recent period (2017–2020), this error was again ~30 km/h, and worsening at longer lead times. It is important to note, however, as discussed in the previous section, that there have also been several seasons with many very intense and challenging storms during this more recent time period, which is likely to influence the results for TC intensity, given that more intense storms tend to see larger errors (Hodges & Emerton, 2015). This is discussed further in the following section.

The exception is the intensity in terms of the minimum central pressure in the HRES forecasts, which has greatly improved at shorter lead times, including during the 2017–2020 period, with average errors around 0 in the first 2 days of the forecast, although there is now a tendency towards overestimating the minimum pressure beyond 2 days ahead. For the ENS, the 2017–2020

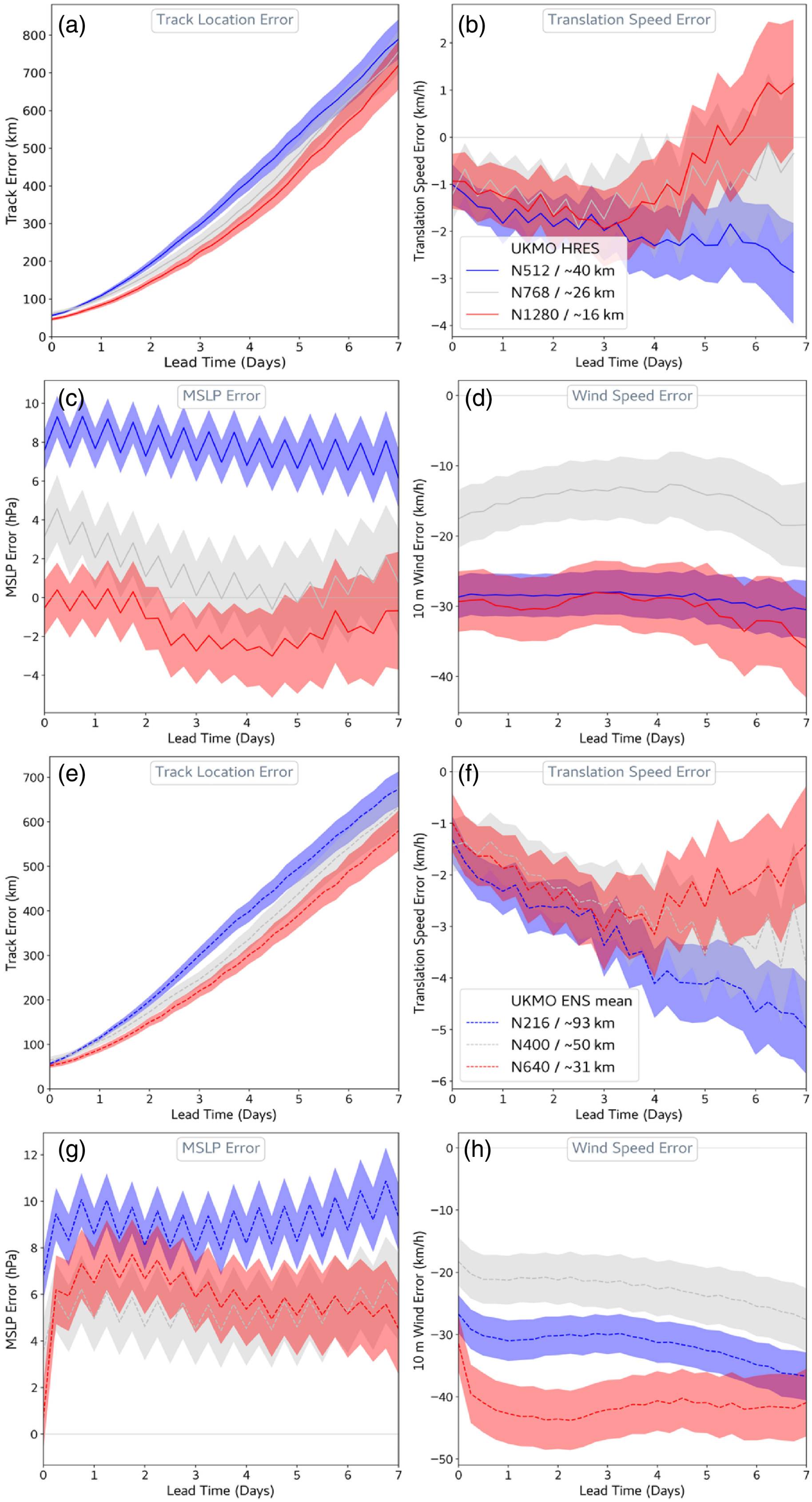

FIGURE 8 U.K. Met Office (UKMO) high-resolution deterministic (HRES) (a–d) and ensemble probabilistic (ENS) mean (e–h) tropical cyclone (TC) forecast errors at lead time days 0 to 7, for each model resolution used over the 10-year period from 2010 to 2020. This represents three distinct periods: July 2010 to June 2014 (blue), July 2014 to June 2017 (grey), and July 2017 to June 2020 (red). See Table 2 for details. Shading indicates the 95% confidence interval. (a, e) Track error (b, f) translation speed error, (c, g) wind speed error, where negative values indicate that the wind speeds were under-predicted by the forecasts, and (d, h) minimum pressure error, where positive values indicate that the intensity of the TC was too weak in the forecasts. MSLP, mean sea-level pressure.

seasons saw worse minimum pressure errors in the first 5 days of the forecasts than 2014–2017 but a slight improvement at days 6–7.

Results for the ENS control forecasts and the mean of the individual ENS members (not shown) are similar to those shown for the ensemble mean (Figure 8e–h), with smaller confidence intervals for the individual ENS members given the much larger sample size.

5.2.2 | European Centre for Medium-Range Weather Forecasts

The ECMWF forecasting system's horizontal resolution was upgraded once during the 2010–2020 period, in March 2016 (Table 2). The forecast dataset was split into two distinct periods to reflect this change. The track forecasts of both the HRES and ENS forecasting systems improved with the increase in resolution, particularly at lead times <2 days ahead and beyond 4 days ahead (Figure 9a,e). This equates to a ~1 day of gained lead time through improved forecast skill, at short lead times (<1 day) and longer lead times (>5 days); the skill at 6 days ahead for forecast track in 2016–2020 is similar to the skill at day 5 in 2010–2016 (Figure 9a,e). The improvement at lead times of 2–5 days ahead is smaller. While the UKMO forecasts have seen a slightly larger improvement, the maximum errors in the ECMWF forecasts at day 7 for both periods are smaller than those seen for all three UKMO periods. The translation speed errors (Figure 9b,f) do not significantly change in either forecasting system over the period of the study, but are better in the HRES forecasts than the ENS.

The ECMWF intensity forecast errors also show more mixed results. In the HRES forecasts, the average minimum pressure error (Figure 9c) is similar during both periods up to 5 days ahead, beyond which the sample size is smaller for the 2016–2020 period, but the errors are reduced and sometimes overestimated the intensity (minimum pressure too low). For the ENS (Figure 9g), the minimum pressure errors are slightly worse at all lead times. In terms of the maximum wind speeds (Figure 9d,h), the underestimation appears significantly worse during the 2016–2020 period (~35 km/h HRES, ~40–60 km/h ENS) than 2010–2016 (~15 km/h HRES, ~15–25 km/h ENS). As mentioned in the two previous sections, this latter period has experienced several seasons with very intense and challenging storms, which may be a factor in the intensity forecasts results given the sample periods used. The first period saw a total of 9 very intense and 9 intense TCs, while the second period saw 12 very intense and 10 intense TCs (Figure 2a; it is noted here that two of the very intense TCs shown in the 2015–2016 season occurred during March and April 2016, and therefore fall in the second period of study here). Of the very intense TCs, the maximum wind speeds were generally slightly higher during the latter period (Figure 2a; an average of 251 km/h maximum wind speed across the very intense TCs from March 2016 to July 2020, compared to an average of 243 km/h from July 2010 to March 2016). Overall, the average TC intensity across all storms was higher during the latter period, with a mean maximum wind speed across all TCs from March 2016 to July 2020 of 174 km/h, compared to a mean of 146 km/h from July 2010 to March 2016.

Results for the ENS control (not shown) are similar to those shown for the ENS mean (Figure 9e–h), although the translation speed errors are smaller for the control. This effect is explained in Section 5.1. The ENS mean typically shows slightly smaller track errors than the mean of the individual ENS members (not shown), with a maximum average track error at day 7 of ~750 km during the 2010–2016 period and ~675 km during the 2016–2020 period. The intensity error results for the individual ENS members (not shown) are similar to those shown for the ENS mean (Figure 9g,h) but with much smaller confidence intervals given the larger sample size for the individual ensemble members.

Although the resolution of forecasting systems is a key factor that can result in significant changes and improvements to forecast errors, a key area of research in the scientific community is whether increasing model resolution improves TC intensity forecasts compared to other potential model changes. Some higher resolution (~4 km) experiments using the ECMWF IFS have resulted in TCs that are too intense (Magnusson et al., 2021); however, experiments using the new model physics package that was introduced to the IFS in 2021 resulted in a decrease in intensity (on average). This means that at model resolutions running operationally in the early 2020s, the newly introduced model physics could be expected to see underestimated TC intensities, but should lead to more accurate intensity forecasts at future higher model resolutions.

Other recent model developments have also led to improvements in ECMWF's TC forecasts. For example, introduction of ocean coupling (in 2013 for the ENS, and in 2018 for the HRES forecasts; Buizza et al., 2018) improved intensity forecasts by avoiding over-deepening TCs (Mogensen et al., 2018). This is likely to be another factor in the differences between the skill of the ECMWF and UKMO models, as the UKMO model at the time of this study is not fully coupled. In 2020, changes to the ocean drag in the ECMWF IFS led to improvements in the wind-pressure relationship and maximum wind speed forecasts (Bidlot et al., 2020). While various aspects of the

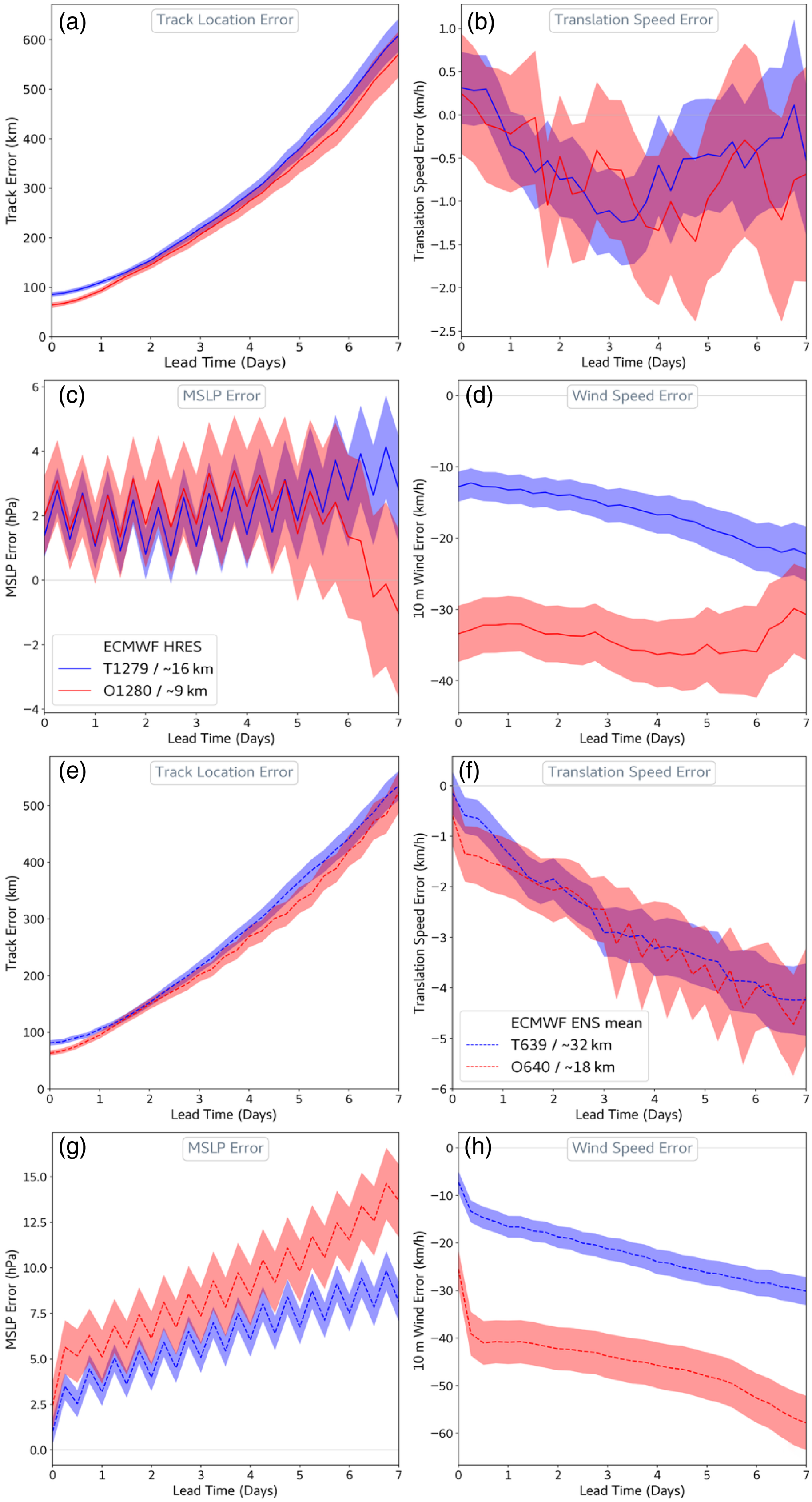

FIGURE 9 European Centre for Medium-Range Weather Forecasts (ECMWF) high-resolution deterministic (HRES) (a–d) and ensemble probabilistic (ENS) mean (e–h) tropical cyclone (TC) forecast errors at lead time days 0 to 7, for each model resolution used over the 10-year period from 2010 to 2020. This represents two distinct periods: July 2010 to 7 March 2016 (blue), and 8 March 2018 to June 2020 (red). See Table 2 for details. Shading indicates the 95% confidence interval. (a, e) Track location error; (b, f) translation speed error; (c, g) wind speed error, where negative values indicate that the wind speeds were under-predicted by the forecasts; and (d, h) minimum pressure error, where positive values indicate that the intensity of the TC was too weak in the forecasts. MSLP, mean sea-level pressure.

model are important for improving TC forecasts, the higher resolution 4-km experiments also demonstrated that model resolution is key for predicting not only the maximum wind speeds and minimum pressures of TCs but also the relationship between the two (Magnusson et al., 2021; Majumdar et al., 2023).

6 | IMPACT OF THE MJO ON PREDICTION SKILL

Most research investigating the impact of tropical modes of variability on TCs has focussed on TC occurrence on seasonal timescales. At shorter timescales, recent work by Hodges and Klingaman (2019) found that the MJO and its counterpart, the Boreal Summer Intraseasonal Oscillation (BSISO; Lee et al., 2013), significantly impact the skill of short- and medium-range TC intensity forecasts in the western North Pacific. This impact on forecast skill is mainly due to variations in the intensity of TCs that occur in different phases of the MJO/BSISO and impacts on the initial states of the forecasts (Hodges & Klingaman, 2019).

Here, we evaluate the impact of the MJO on TC forecast skill in the SWIO, given the strong influence of the MJO on TCs in this region (see Section 2). If there are phases in which forecasts tend to be more or less accurate, this could provide windows of opportunity for enhanced predictability and better inform decision making at longer lead times. The BSISO is not considered, as it is typically more active from May–October, and therefore does not coincide with the SWIO cyclone season.

A daily timeseries of the RMM indices of Wheeler and Hendon (2004), which describes the phase (position) and amplitude the MJO, was used to identify the MJO phase on the day that each TC forecast in the sample was initialised. Only those forecasts produced on days when the MJO was active, with an amplitude ≥1, were

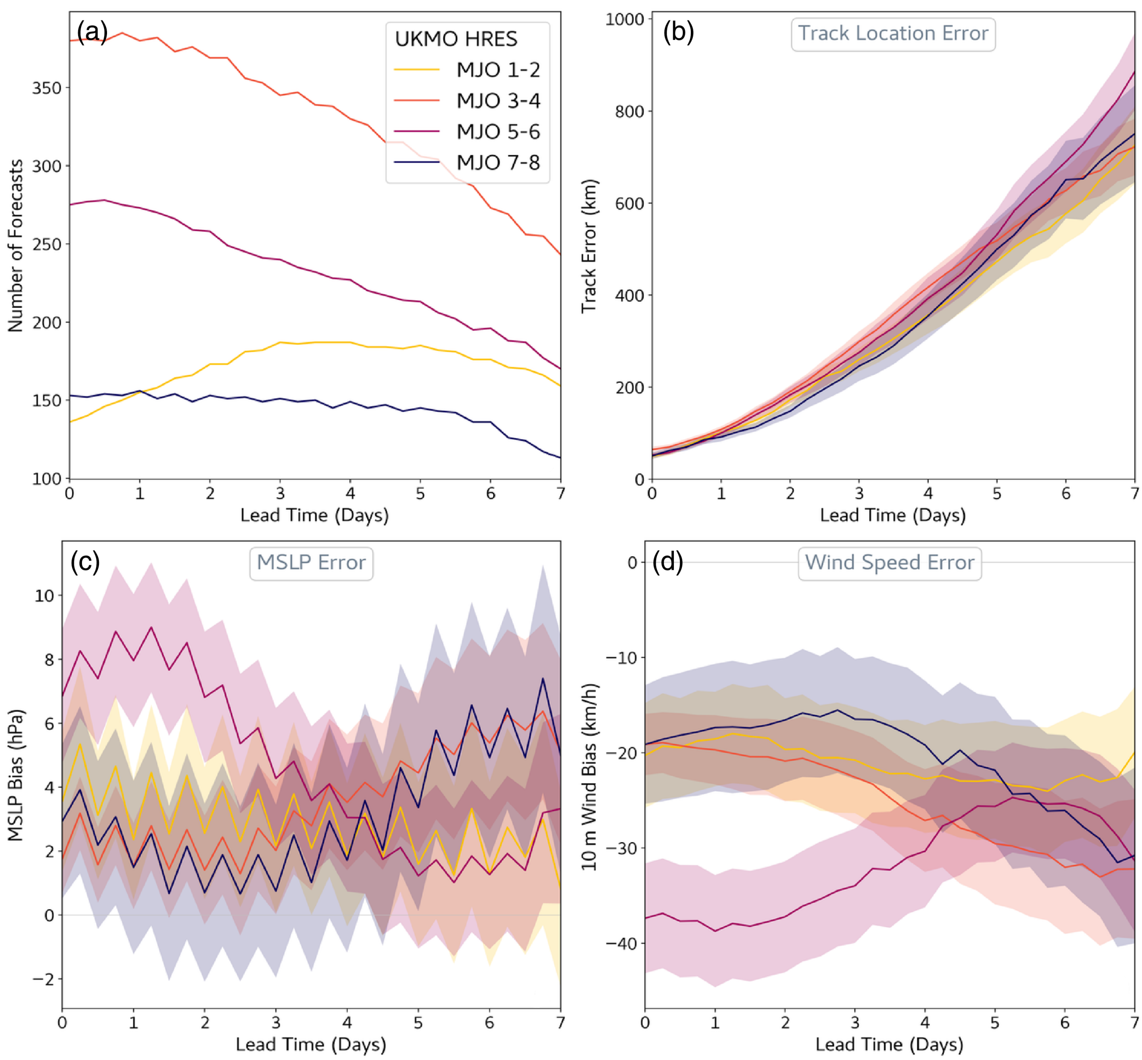

FIGURE 10 Conditional prediction skill for U.K. Met Office (UKMO) high-resolution deterministic (HRES) forecasts of tropical cyclone track and intensity from 2010 to 2020, based on the Madden-Julian Oscillation (MJO) phase when each forecast was initialised. Only forecasts initialised when the MJO was considered active (amplitude ≥ 1) are included, and results are shown for pairs of MJO phases. (a) Number of forecasts in the sample for each MJO phase pair, (b) mean track forecast errors, (c) mean minimum mean sea-level pressure (MSLP) errors, and (d) mean maximum wind speed errors.

retained. These RMM data can be obtained from the Bureau of Meteorology (http://www.bom.gov.au/climate/mjo/graphics/rmm.74toRealtime.txt).

It is common for MJO composites to use pairs of MJO phases to increase sample size. Typically, geographical phase pairs are used corresponding to distinct regions: phases 2–3 (Indian Ocean), 4–5 (Maritime Continent), 6–7 (western Pacific Ocean) and 8–1 (MJO transition, where the MJO dissipates near the date line and re-emerges over Africa). Here, we follow the method of Peatman et al. (2019) and use alternative pairings that are relevant for TC genesis due to the finding that TC activity is 'enhanced in the MJO phases associated with and immediately following the convective maximum in a specific basin' (Klotzbach, 2014; Klotzbach & Oliver, 2015). This lag causes a shift of one phase from the maximum convection. The phase pairs used are therefore 1–2, 3–4, 5–6, and 7–8. The results are summarised here for the HRES forecasts (shown in Figures 10 and 11), but the same conclusions arise for the ENS forecasts (not shown). The sample size of forecasts in each phase pair ranges from ∼140 to ∼375 at day 0, with phases 7–8 having the smallest sample size.

An interesting result from the conditional verification is that the phases in which the two forecasting systems provide less accurate forecasts differ between the two systems, and with lead time. There is no clear MJO phase pair in which forecasts can always be expected to be more or less accurate. The MJO appears to have a larger impact on intensity errors than track errors.

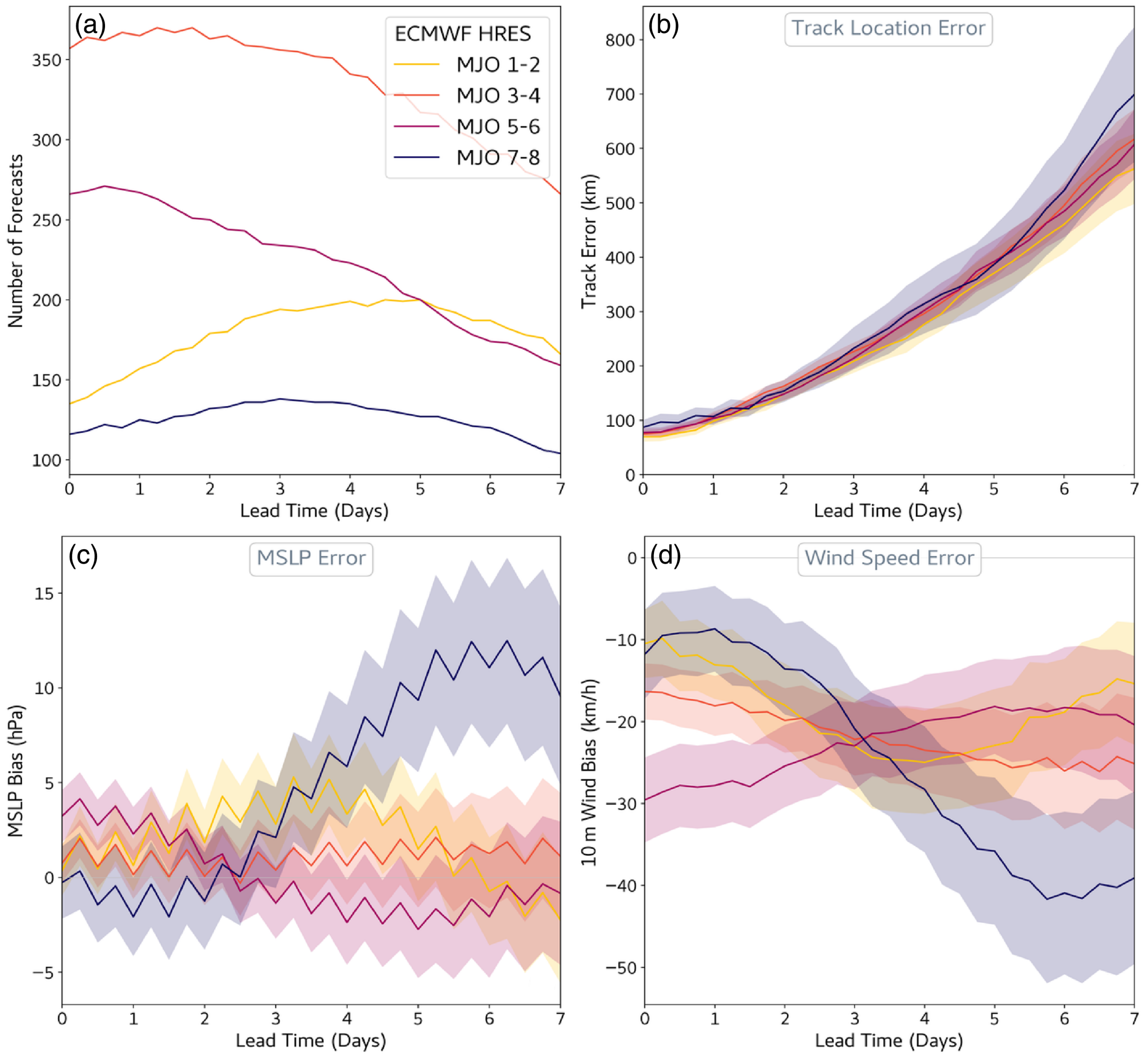

FIGURE 11 Conditional prediction skill for European Centre for Medium-Range Weather Forecasts (ECMWF) high-resolution deterministic (HRES) forecasts of tropical cyclone track and intensity from 2010 to 2020, based on the Madden-Julian Oscillation (MJO) phase when each forecast was initialised. Only forecasts initialised when the MJO was considered active (amplitude ≥ 1) are included, and results are shown for pairs of MJO phases. (a) Number of forecasts in the sample for each MJO phase pair, (b) mean track forecast errors, (c) mean minimum mean sea level pressure (MSLP) errors, and (d) mean maximum wind speed errors.

In the UKMO forecasts (Figure 10a–d), the track errors tend to be similar in all MJO phases at very short lead times (Figure 10b). From 1.5 to 4 days ahead, forecasts initialised when the MJO is in phases 7–8 are more accurate, and beyond day 4, forecasts initialised in phases 1–2 are more accurate. Beyond 1 day ahead, forecasts produced in phases 3–6 are less accurate, and this is particularly the case for phases 5–6 after 5 days ahead. By day 7, 1 day of lead time is lost for forecast produced in phases 5–6 compared to all other phases (the errors at day 7 for phases 1–4 and 7–8 are similar to those for phases 5–6 at day 6). Beyond 5 days ahead, the difference in track error between phases 5–6 and 1–2 is statistically significant, but at other lead times and for other phases, the overlapping confidence intervals suggests that the differences are not statistically significant.

For ECMWF (Figure 11a–d), forecasts produced in all phases have similar track errors out to ~day 3, and the differences in errors between MJO phases are not statistically significant (Figure 10b). Beyond day 3, on average, the forecasts produced in phases 1–2 have more accurate tracks, and beyond day 5, forecasts produced in phases 7–8 are least accurate.

For TC intensity, the UKMO forecasts produced in phases 5–6 are significantly less accurate than those produced in all other phases of the MJO, at lead times of 0–3 days ahead, for both maximum wind speed (Figure 10d) and minimum pressure (Figure 10c). Differences in intensity errors between the other phases and lead times are not statistically significant and vary with lead time (Figure 10c,d). Forecasts produced in phases 1–2 appear to be more consistent in the underestimation of intensity, while forecasts produced in phases 2–4 and 7–8 worsen with lead time, and those produced in phases 5–6 become more accurate beyond day 4.

In the ECMWF forecasts (Figure 11c,d), forecasts produced in phases 7–8 have the smallest errors out to day 2, but the skill rapidly declines and these forecasts have worse errors than other phases at lead times beyond 4 days. Inspection of the individual phases (not shown) indicates that this primarily comes from the errors for forecasts produced in phase 8, for which the sample size is very small beyond day 4 (~10 forecasts). On the other

TABLE 3 Summary of the Madden-Julian Oscillation (MJO) phase pairs in which the tropical cyclone forecast track, maximum wind, and minimum pressure forecasts tend to be better or worse, for various lead times in both the European Centre for Medium-Range Weather Forecasts (ECMWF) and U.K. Met Office (UKMO) high-resolution deterministic forecasts.

		Track		Wind		Pressure	
MJO phase	**Lead time**	**ECMWF**	**UKMO**	**ECMWF**	**UKMO**	**ECMWF**	**UKMO**
1–2	Day 1			↑	↑		↑
	Day 3	↑				↓	
	Day 5	↑	↑	↑		↑	↑
	Day 7	↑	↑	↑	↑	↑	↑
3–4	Day 1				↑		↑
	Day 3		↓			↑	
	Day 5				↓	↑	↓
	Day 7		↑			↑	↓
5–6	Day 1			↓	↓	↓	↓
	Day 3	↑			↓		↓
	Day 5		↓	↑		↑	↑
	Day 7		↓	↑		↑	↑
7–8	Day 1			↑	↑	↑	↑
	Day 3	↓	↑		↓	↓	↓
	Day 5			↓	↑	↓	↓
	Day 7	↓		↓		↓	↓

Note: The MJO phase refers to the phase when the forecast is initialised, when the MJO is active only (amplitude ≥ 1), therefore providing a look-up table indicating whether each forecasting system is likely to provide a more or less accurate forecast than usual/during other phases, based on the current MJO phase at the time the forecast is produced. Purple up arrows indicate that the forecast skill is better at the lead time indicated, when initialised in the MJO phase indicated, and red down arrows indicates that it is worse. Bold arrows highlight that the difference is statistically significant, and no entry implies that the skill is similar to other phases and there is no clear difference. Alternative layouts of this table are provided in Tables S1 and S2.

hand, forecasts produced in phases 5–6 have less accurate intensity forecasts in the first 2 days, but smaller errors at longer lead times (particularly phase 6, not shown), although this is less significant for wind speed than minimum pressure. Forecasts produced in phases 3–4 for pressure and 1–2 for wind speed are more consistent with lead time.

For both forecasting systems, the large and significant differences in the intensity errors for different MJO phase pairs suggests that the MJO has a strong influence at the start of the forecasts, and for some phases this continues to affect the forecast (e.g., phases 1–4); or, in some cases this influence quickly disappears and rapid error growth is observed (e.g., phases 5–6 and 8–1). This was also observed in the intensity errors in the western North Pacific basin (Hodges & Klingaman, 2019).

In the SWIO, TCs are more likely to be intense, or to rapidly intensify, when they form in phases 2–6, and phase 6 sees the largest percentage of intense or very intense TCs (71% of TCs that formed in phase 6 became intense TCs, Figure 3). This is followed by phases 2–5, with between 36% and 55% of the TCs that form in these phases becoming intense or very intense (Figure 3). Phases with more intense TCs at the start of the forecasts may be more likely to have higher intensity errors for the first few days of lead time, while phases with less intense TCs typically have smaller errors. Inspection of the errors for individual phases (not shown) suggests that the difference in typical track direction observed for TCs that form in phase 6 (Figure 3) does not significantly impact the track forecast skill.

Table 3 provides a summary of the phases that provide better or worse errors for the track and intensity for each forecasting system.

7 | CONCLUSIONS

The anticipation and forecasting of natural hazards, such as tropical cyclones (TCs), is crucial to preparing for their impacts. It is important, therefore, to understand how well forecast models are able to predict these events as well as the limitations of the forecasts. In this study, we have evaluated TC forecasts from two state-of-the-art global ensemble weather prediction systems over a 10-year period from 2010 to 2020, with a focus on the southwest Indian Ocean (SWIO). This large sample of forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) and U.K. Met Office (UKMO) forecasting systems has allowed us to investigate their strengths and weaknesses, aiming to identify changes in forecast skill at various lead times ahead of a TC, and evaluating how the two forecasting systems have changed and improved with time. Although there are similarities between the skill of the ECMWF and the UKMO forecasts, ECMWF tend to provide, on average, more accurate TC track (beyond 1.5 days ahead) and intensity (in terms of both the minimum pressure and maximum wind) forecasts. Both forecasting systems typically underestimate the translation speed and the intensity of TCs, which has implications for planning, decision making, hazards, and impacts. The high-resolution deterministic models from both ECMWF and the UKMO provide more accurate intensity forecasts than their ensemble counterparts, and can therefore be useful to consider in conjunction with an ensemble forecast that represents the full range of uncertainty.

Over the 10-year period of the study, both forecasting systems have undergone a range of model changes and upgrades, including large improvements to the resolution of the forecasts, which represents a step change towards better resolving features such as TCs. There has been a clear improvement in the forecast skill for TC tracks in the SWIO, with the UKMO forecasts gaining ∼1.5 days of lead time (i.e., the errors we see at day 5 in recent years are equivalent to the errors we saw at day 3.5 in 2010), and the ECMWF forecasts gaining ∼1 day of lead time. The changes in intensity forecast skill for both systems are more variable because of combination of model changes and year-to-year variability in the number and intensity of TCs. Comparison of the track forecast skill of these two systems with the official forecast guidance provided by the La Réunion Regional Specialised Meteorological Centre (RSMC-LR) also shows that at short lead times, the official guidance from the RSMC-LR provides more accurate forecasts, but tends to fall between the ECMWF and UKMO forecasts at longer lead times. Future model changes are likely to lead to further improvements in TC skill as we move towards better physical representations of the atmosphere and Earth system, and higher resolutions.

Ongoing and future work should consider a range of sensitivity analyses to determine which model components contribute to the errors and improvements, to further guide model development towards improving the accuracy of TC forecasts. Magnusson et al. (2021) provide a detailed overview of TC research and diagnostics at ECMWF, including observations and tracking, verification, forecast challenges, data assimilation, and model development. This includes experiments comparing high-resolution forecasts at 9 and 4 km to investigate the model's ability to represent TC intensity at different resolutions and in combination with different model physics. Magnusson et al. (2021) also mention avenues for future improvement of TC forecasting in numerical weather prediction, from data assimilation experiments to investigation of propagation speed and intensification challenges, to the use of machine learning.

This study also looked at the impact of the Madden-Julian Oscillation (MJO) on TCs and their forecasts in this region. The majority of TCs in the SWIO form when the MJO is active in phases 2–5, although they can occur during any phase of the MJO. Those that form when the MJO is active in phases 3–5 tend to occur further east, posing less of a threat to land. While fewer storms typically form during phase 6, a larger percentage of those that do form go on to become intense or very intense cyclones, and they can have the tendency to follow a more eastward-propagating track than is typical for storms in this region. Given the influence the MJO exerts on TCs in the SWIO, we further investigated the impact of the MJO on forecast skill. The MJO has a larger impact on intensity errors than track errors, with a strong influence on intensity seen at the start of the forecasts, which for some phases continues to affect the forecasts; for others, the influence disappears with lead time and leads to more rapid error growth. This finding is similar to that of Hodges and Klingaman (2019) for impacts of the MJO on TCs on the northwest Pacific. There is no clear MJO phase that leads to more, or less, accurate forecasts; the impact varies between forecasting systems and with lead time.

A known challenge for TC forecasting is the prediction of storms that undergo rapid intensification. In the North Atlantic, it has been found that hurricane seasons with more frequent rapid intensification cases tend to have larger annual average forecast errors (DeMaria et al., 2021). Decision makers involved in activating anticipatory humanitarian action in Madagascar have also highlighted the challenges of taking early action for rapidly intensifying cyclones due to forecast limitations (Start Network, 2022). Future work should consider using reforecast datasets produced using the latest versions of forecasting systems to allow more robust evaluation of the forecast skill in this region for rapidly intensifying TCs, as they provide a larger sample size of forecasts at the latest model resolutions. Similarly, the combined impacts of the MJO and other modes of variability, such as the Indian Ocean Dipole and El Niño Southern Oscillation could be considered with the use of reforecast datasets.

The results presented here are intended to provide an increased understanding of the ability of global ensemble forecasting systems to predict TCs in the SWIO, for the purpose of decision making and anticipatory forecast-based action. However, while forecasting the track, landfall location, and intensity of TCs is an important aspect of decision making and preparedness, hazard-focused verification (including the wider wind fields, precipitation, storm surge, and flooding) is also key in terms of understanding prediction skill. Particularly with an increasing move towards impact-based forecasting, further research should extend this analysis by evaluating forecasts for TC hazards, providing further information that is useful for decision making and disaster risk reduction efforts for TC impacts, both in the SWIO and in other basins worldwide.

AUTHOR CONTRIBUTIONS

R. Emerton: Conceptualization (equal); data curation (lead); formal analysis (lead); investigation (lead); methodology (equal); software (equal); validation (lead); visualization (lead); writing – original draft (lead). **K. I. Hodges:** Conceptualization (equal); data curation (equal); investigation (equal); methodology; software (lead); supervision (equal); writing – review and editing (equal). **E. Stephens:** Conceptualization (equal); investigation (equal); methodology; supervision (equal); writing – review and editing (equal). **V. Amelie:** Conceptualization (supporting); investigation (supporting); methodology (equal); writing – review and editing (supporting). **M. Mustafa:** Conceptualization (supporting); investigation (supporting); methodology (equal); writing – review and editing (supporting). **Z. Rakotomavo:** Conceptualization (supporting); investigation (supporting); methodology (equal); writing – review and editing (supporting). **E. Coughlan de Perez:** Conceptualization (supporting); investigation (supporting); methodology (equal); writing – review and editing (supporting). **L. Magnusson:** Data curation (supporting); formal analysis (supporting); investigation (supporting); methodology (supporting); writing – review and editing (supporting). **P.-L. Vidale:** Conceptualization (equal); data curation (supporting); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); supervision (equal); writing – review and editing (equal).

ACKNOWLEDGEMENTS

The authors thank Dr Nick Klingaman for his role in facilitating the PICSEA project, fostering the collaboration of the authors of this paper and supporting the work and ideas throughout the project. The authors acknowledge support from the UK's Natural Environment Research Council (NERC) and Department for International Development (DFID) under the Science for Humanitarian Emergencies and Resilience (SHEAR) research programme project PICSEA (Predicting Impacts of Cyclones in South East Africa, grant number NE/S005897/1). R.E. is also grateful for support from ECMWF by providing time and resources towards the continuation and publication of this work following the end of the PICSEA project. More information on tropical cyclones, the MJO and their forecasts, can be found in the PICSEA project's free online training course at https://the-iea.github.io/picsea/. This training course is

aimed at forecasters and decision makers in the southwest Indian Ocean, but contains more widely applicable information on forecasting tropical cyclones, with videos, interviews and a podcast. Transcripts are available in French and Portuguese. The authors thank those involved in the creation of these resources.

DATA AVAILABILITY STATEMENT

ECMWF and UKMO ENS forecast data are available through the TIGGE archive after a short delay (https://www.ecmwf.int/en/research/projects/tigge). ECMWF HRES data can be accessed through ECMWF's Meteorological Archival and Retrieval System (MARS), subject to licensing. The TC tracks identified from the raw forecast data for this study using the TC tracking software of Hodges (1994, 1995, 1999) can be made available on request. The tracking software is available online at https://gitlab.act.reading.ac.uk/track/track/-/releases. The IBTrACS TC best track data are openly available at https://www.ncei.noaa.gov/products/international-best-track-archive. Operational forecasts from the RSMC in La Réunion can be seen at www.meteofrance.re/cyclone/, UKMO operational TC guidance messages at https://www.metoffice.gov.uk/research/weather/tropical-cyclones/warnings and ECMWF operational TC forecasts at https://www.ecmwf.int/en/forecasts/charts/latest-tropical-cyclones-forecast.

ORCID

R. Emerton https://orcid.org/0000-0003-0707-3993
L. Magnusson https://orcid.org/0000-0003-4707-2231

REFERENCES

Bidlot, J.-R., Prates, F., Ribas, R., Mueller-Quintino, A., Crepulja, M. & Vitart, F. (2020) Enhancing tropical cyclone wind forecasts. *ECMWF Newsletter*, 164. Available at: https://www.ecmwf.int/en/newsletter/164/meteorology/enhancing-tropical-cyclone-wind-forecasts [Accessed 12th May 2023].

Bonnardot, F., Quetelard, H., Jumaux, G., Leroux, M.-D. & Bessafi, M. (2018) Probabilistic forecasts of tropical cyclone tracks and intensities in the southwest Indian Ocean basin. *Quarterly Journal of the Royal Meteorological Society*, 145(719), 675–686. Available from: https://doi.org/10.1002/qj.3459

Bousquet, O., Barbary, D., Bielli, S., Kebir, S., Raynaud, L., Malardel, S. et al. (2020) An evaluation of tropical cyclone forecast in the Southwest Indian Ocean basin with AROME-Indian Ocean convection-permitting numerical weather predicting system. *Atmospheric Science Letters*, 21(3), e950. Available from: https://doi.org/10.1001/asl.950

Buizza, R., Balsamo, G. & Haiden, T. (2018) IFS upgrade brings more seamless coupled forecasts. *ECMWF Newsletter*, 156 (Summer 2018). Available at: https://www.ecmwf.int/en/newsletter/156/meteorology/ifs-upgrade-brings-more-seamless-coupled-forecasts [Accessed 14th December 2022].

Dai, A. & Wang, J. (1999) Diurnal and semidiurnal tides in global surface pressure fields. *Journal of the Atmospheric Sciences*, 56, 3874–3891. Available from: https://doi.org/10.1175/1520-0469(1999)056<3874:DASTIG>2.0.CO;2

DeMaria, K., Franklin, J.L., Onderlinde, M. & Kpalan, J. (2021) Operational forecasting of tropical cyclone rapid intensification at the National Hurricane Center. *Atmosphere*, 12(6), 683. Available from: https://doi.org/10.3390/atmos12060683

Emerton, R., Cloke, H., Ficchi, A., Hawker, L., de Wit, S., Speight, L. et al. (2020) Emergency flood bulletins for cyclones Idai and Kenneth: a critical evaluation of the use of global flood forecasts for international humanitarian preparedness and response. *International Journal of Disaster Risk Reduction*, 50, 101811. Available from: https://doi.org/10.1016/j.ijdrr.2020.101811

Froude, L.S.R. (2010) TIGGE: comparison of the prediction of northern hemisphere extratropical cyclones by different ensemble prediction systems. *Weather and Forecasting*, 25(3), 819–836. Available from: https://doi.org/10.1175/2010WAF2222326.1

Froude, L.S.R. (2011) TIGGE: comparison of the prediction of southern hemisphere extratropical cyclones by different ensemble prediction systems. *Weather and Forecasting*, 26(3), 388–398. Available from: https://doi.org/10.1175/2010WAF2222457.1

Heming, J.T. (2016a) Met Office unified model tropical cyclone performance following major changes to the initialization scheme and a model upgrade. *Weather and Forecasting*, 31(5), 1433–1449. Available from: https://doi.org/10.1175/WAF-D-16-0040.1

Heming, J.T. (2016b) Tropical cyclone tracking and verification techniques for the Met Office numerical weather prediction models. *Meteorological Applications*, 24(1), 1–8. Available from: https://doi.org/10.1002/met.1599

Hodges, K.I. (1994) A general method for tracking analysis and its application to meteorological data. *Monthly Weather Review*, 122, 2573–2586. Available from: https://doi.org/10.1175/1520-0493(1994)122<2573:AGMFTA>2.0.CO;2

Hodges, K.I. (1995) Feature tracking on the unit sphere. *Monthly Weather Review*, 123, 3458–3465. Available from: https://doi.org/10.1175/1520-0493(1995)123<3458:FTOTUS>2.0.CO;2

Hodges, K.I. (1999) Adaptive constraints for feature tracking. *Monthly Weather Review*, 127, 1362–1373. Available from: https://doi.org/10.1175/1520-0493(1999)127<1362:ACFFT>2.0.CO;2

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 1: tropical cyclone stage. *Monthly Weather Review*, 143(12), 5091–5114. Available from: https://doi.org/10.1175/MWR-D-13-00385.1

Hodges, K.I. & Klingaman, N.P. (2019) Prediction errors of tropical cyclones in the Western North Pacific in the Met Office global forecast model. *Weather and Forecasting*, 34(5), 1189–1209. Available from: https://doi.org/10.1175/WAF-D-19-0005.1

Klotzbach, P.J. (2014) The Madden-Julian Oscillation's impacts on worldwide tropical cyclone activity. *Journal of Climate*, 27(6), 2317–2330. Available from: https://doi.org/10.1175/JCLI-D-13-00483.1

Klotzbach, P.J. & Oliver, E.C.J. (2015) Variations in global tropical cyclone activity and the Madden-Julian Oscillation since the mid-twentieth century. *Geophysical Research Letters*, 42(10), 4199–4207. Available from: https://doi.org/10.1002/2015GL063966

Knapp, K.R., Diamond, H.J., Kossin, J.P., Kruk, M.C. & Schreck, C.J. (2018) International Best Track Archive for Climate Stewardship (IBTrACS) Project, Version 4. *NOAA National Centers for Environmental Information*. Available from: https://doi.org/10.25921/82ty-9e16

Knapp, K.R., Kruk, M.C., Levinson, D.H., Diamond, H.J. & Neumann, C.J. (2010) The International Best Track Archive for Climate Stewardship (IBTrACS): unifying tropical cyclone best track data. *Bulletin of the American Meteorological Society*, 91, 363–376. Available from: https://doi.org/10.1175/2009BAMS2755.1

Lee, J.-Y., Wang, B., Wheeler, M.C., Fu, X., Waliser, D.E. & Kang, I.-S. (2013) Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. *Climate Dynamics*, 40, 493–509. Available from: https://doi.org/10.1007/s00382-012-1544-4

Leroux, M.-D., Meister, J., Mekies, D., Dorla, A.-L. & Caroff, P. (2018) A climatology of southwest Indian Ocean tropical systems: their number, tracks, impacts, sizes, empirical potential intensity, and intensity changes. *Journal of Applied Meteorology and Climatology*, 57(4), 1021–2041. Available from: https://doi.org/10.1175/JAMC-D-17-0094.1

Li, T., Wang, L., Peng, M., Wang, B., Zhang, C., Lau, W. et al. (2018) A paper on the tropical intraseasonal oscillation published in 1963 in a Chinese journal. *Bulletin of the American Meteorological Society*, 99(9), 1765–1779. Available from: https://doi.org/10.1175/BAMS-D-17-0216.1

Madden, R.A. & Julian, P. (1971) Detection of a 40-50 day oscillation in the zonal wind. *Journal of the Atmospheric Sciences*, 28, 1109–1123. Available from: https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2

Madden, R.A. & Julian, P.R. (1972) Description of global-scale circulation cells in the tropics with a 40-50 day period. *Journal of the Atmospheric Sciences*, 29(6), 1109–1123. Available from: https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2

Magnusson, L., Majumdar, S., Emerton, R., Richardson, D., Alonso-Balmaseda, M., Baugh, C. et al. (2021) Tropical cyclone activities at ECMWF. *ECMWF Technical Memorandum*, 888. Available from: https://doi.org/10.21957/zzxzzygwv

Majumdar, S.J., Magnusson, L., Bechtold, P., Bidlot, J.-R. & Doyle, J.D. (2023) Advanced tropical cyclone prediction using the experimental global ECMWF and operational regional COAMPS-TC systems. *Monthly Weather Review*, 151, 2029–2048. Available from: https://doi.org/10.1175/MWR-D-22-0236.1

Malardel, S., Wedi, N., Deconinck, W., Diamantakis, M., Kühnlein, C., Mozdzynski, G. et al. (2016) A new grid for the IFS. *ECMWF Newsletter*, 146(Winter 2015/16), 23–28. Available at: https://www.ecmwf.int/node/15041 [Accessed 10th November 2022].

Mogensen, K.S., Magnusson, L., Bidlot, J.-R. & Prates, F. (2018) Ocean coupling in tropical cyclone forecasts. *ECMWF Newsletter*, 154(Winter 2018). Available at: https://www.ecmwf.int/en/newsletter/154/meteorology/ocean-coupling-tropical-cyclone-forecasts [Accessed 14th December 2022].

Muthige, M.S., Malherbe, J., Englebrecht, F.A., Grab, S., Beraki, A., Maisha, T.R. et al. (2018) Projected changes in tropical cyclones over the south West Indian Ocean under different extents of global warming. *Environmental Research Letters*, 13, 065019. Available from: https://doi.org/10.1088/1748-9326/aabc60

NOAA. (2019) *International Best Track Archive for Climate Stewardship (IBTrACS version 04) technical documentation*. Available at: https://www.ncei.noaa.gov/products/international-best-track-archive [Accessed 9th May 2023].

Peatman, S.C., Klingaman, N.P. & Hodges, K.I. (2019) Tropical cyclone-related precipitation over the northwest tropical Pacific in Met Office global operational forecasts. *Weather and Forecasting*, 34(4), 923–941. Available from: https://doi.org/10.1175/WAF-D-19-0017.1

Start Network. (2022) *Global community meeting FOREWARN Q4 2022: cyclones and topical storms*. Available at: https://www.youtube.com/watch?v=u39OK6IAQ-I [Accessed 21st December 2022].

Terry, J.P., Kim, I.-H. & Jolivet, S. (2013) Sinuosity of tropical cyclone tracks in the south West Indian Ocean: spatio-temporal patterns and relationships with fundamental storm attributes. *Applied Geography*, 45, 29–40. Available from: https://doi.org/10.1016/j.apgeog.2013.08.006

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. Available from: https://doi.org/10.1002/qj.3712

Titley, H.A., Cloke, H.L., Harrigan, S., Pappenberger, F., Prudhomme, C., Robbins, J.C. et al. (2021) Key factors influencing the severity of fluvial flood hazard from tropical cyclones. *Journal of Hydrometeorology*, 22(7), 1801–1817. Available from: https://doi.org/10.1175/JHM-D-20-0250

Vidale, P.L., Hodges, K., Vannière, B., Davini, P., Roberts, M.J., Strommen, K. et al. (2021) Impact of stochastic physics and model resolution on the simulation of tropical cyclones in climate GCMs. *Journal of Climate*, 34(11), 4315–4341. Available from: https://doi.org/10.1175/JCLI-D-20-0507.1

Wheeler, M.C. & Hendon, H.H. (2004) An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Monthly Weather Review*, 132(8), 1917–1932. Available from: https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2

WMO. (2017) *Global guide to tropical cyclone forecasting*. World Meteorological Organization, WMO-No. 1194. Available at: https://cyclone.wmo.int/pdf/Global-Guide-to-Tropical-Cyclone-Forecasting.pdf [Accessed 8th November 2022].

Xie, Y.-B., Chen, S.-J., Chang, I.-L. & Huang, Y.L. (1963) A preliminary statistic and synoptic study about the basic currents over southeastern Asia and the initiation of typhoons. *Acta Meteorologica Sinica*, 33, 206–217. Available from: https://doi.org/10.11676/qxxb1963.020 (in Chinese).

Xie, Y.-B., Chen, S.-J., Chang, I.-L. & Huang, Y.L. (2018) A preliminary statistic and synoptic study about the basic currents over southeastern Asia and the initiation of typhoons. *Journal of Meteorological Research*, 32, 661–669. Available from: https://doi.org/10.1007/s13351-018-8888-6 (in English).

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: 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