Received: 27 June 2022 | Revised: 30 March 2023 | Accepted: 10 August 2023

DOI: 10.1002/qj.4563

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Multiweek tropical cyclone prediction for the Southern Hemisphere in ACCESS-S2: Maintaining operational skill and continuity of service

J. Camp[1] | **P. Gregory**[1] | **A. G. Marshall**[2] | **J. Greenslade**[1] | **M. C. Wheeler**[1]

[1]Bureau of Meteorology, Melbourne, Victoria Australia

[2]Bureau of Meteorology, Hobart, Tasmania Australia

Correspondence
J. Camp, Bureau of Meteorology, Melbourne, VIC, Australia.
Email: jo.camp@bom.gov.au

Abstract

The skill of subseasonal (multiweek) forecasts of tropical cyclone (TC) occurrence over the Southern Hemisphere is examined in the Australian Bureau of Meteorology's (BoM) multiweek to seasonal prediction system, ACCESS-S2. Relative to its predecessor, ACCESS-S1, ACCESS-S2 shows improved biases in spatial TC frequency in the South Pacific and Southwest Indian Ocean. However, there is no improvement to the known negative bias in TC frequency off the coast of NW Australia. The ability of ACCESS-S2 to provide probabilistic forecasts of TC occurrence for the Southern Hemisphere on multiweek timescales is examined using reliability measures and Brier skill scores. For the period November–February 1990–2012, both ACCESS-S1 and ACCESS-S2 show positive skill relative to climatology for calibrated forecasts out to week 5. However, the skill of ACCESS-S2 is slightly reduced compared to ACCESS-S1 at all lead times, which may be due to the fewer number of ensemble members available. For the full ACCESS-S2 hindcast period, November–April 1981–2018, ACCESS-S2 again shows positive skill of calibrated forecasts over climatology out to week 5. For weeks 1–2, skill is reduced compared to the shorter 1990–2012 period; whereas it is marginally improved for longer lead times (weeks 3–5). Use of lagged ensembles, an alternative linear regression calibration, as well as removing weaker model TCs were examined to potentially improve the skill of ACCESS-S2 forecasts; however, none of these methods were able to significantly increase skill at all lead times. Continued use of the original calibration method is therefore recommended in order to retain skill and continuity of service of the BoM operational and public multiweek TC forecasts.

KEYWORDS

ensembles, Madden–Julian Oscillation, multiweek forecasting, southern hemisphere, subseasonal, tropical cyclones

1 | INTRODUCTION

Tropical cyclones (TCs) and the associated extreme winds, storm surge and heavy rainfall can have devastating impacts on tropical regions. In the Australian region there are around 11 TCs annually with around 40% making landfall. Subseasonal and seasonal forecasts of TC activity can help prepare coastal communities at risk of TC landfall.

These forecasts are of considerable interest to society, as well as risk managers, emergency services, agriculture and the oil and gas industry.

The primary source of predictability for subseasonal TC variability is the Madden–Julian oscillation (MJO; Madden and Julian, 1971, 1972). Globally, the active phase of the MJO can increase TC frequency, as well as periods of TC rapid intensification (Klotzbach, 2014). Accurately forecasting active phases of the MJO is therefore important for predictions of TCs at lead times of 1–5 weeks, which is considered to be the current limit of MJO prediction skill in dynamical forecast systems (Lim *et al.*, 2018). Domeisen *et al.* (2022) showed that forecasts of the exceptionally active MJO in January 2020 played an important role in forecasting severe cyclone *Claudia*, which impacted northwest Australia, with three weeks lead time.

The skill of subseasonal forecasts using the operational forecast system at the Bureau of Meteorology (BoM) has been well documented. Vitart and Robertson (2018) assessed the ability of eight models used as part of the Subseasonal to Seasonal prediction project (S2S; Vitart *et al.*, 2017) to reproduce the observed MJO–TC relationship in the South Pacific and South Indian Ocean. They found that the models were able to reproduce the observed increase in TC activity in the South Indian Ocean, and the decrease in TC activity in the South Pacific, when the convectively active phase of the MJO was in Phase 2 + 3 (Indian Ocean). This included the then operational seasonal forecast system at the BoM, POAMA (Predictive Ocean Atmosphere Model for Australia; Yin *et al.*, 2011), with a 2.5° atmospheric horizontal resolution. This model was further examined by Lee *et al.* (2018) for predictions of TC occurrence for weeks 1–4, with skill shown over climatology for uncalibrated predictions of TC occurrence in the western North Pacific to week 4 (May–December), in the South Pacific and Australian region to week 2 (November–April), and in the South Indian Ocean in week 1 (October–April).

Camp *et al.* (2018) showed the Australian Community Climate and Earth-System Simulator – Seasonal version 1 (ACCESS-S1; Lim *et al.*, 2016; Hudson *et al.*, 2017) to have skill for predictions of the MJO out to a lead time of around 30 days and that the system was able to reproduce the observed eastward propagation of the large-scale environmental conditions conducive for TC development. Specifically, for the Australian region, ACCESS-S1 was able to reproduce the observed increase in TC activity in Phases 2 + 3 and 4 + 5, and a decrease in TC activity in Phases 6 + 7 and 8 + 1. Ensemble forecast verification for start dates of 1 November, December, January and February 1990–2012 showed that ACCESS-S1 could potentially provide skilful predictions of TC occurrence over the Southern Hemisphere for a lead time of 1–5 weeks when a spatial and temporal calibration to the forecast probabilities was applied.

The results of Camp *et al.* (2018) were applied to real-time multiweek forecasts of TC activity for the following 2017/18 (Gregory *et al.*, 2019) and 2018/19 (Camargo *et al.*, 2019) Southern Hemisphere cyclone seasons and showed positive results, with ACCESS-S1 providing useful guidance for the development of major TCs, including cyclones *Gita* in the South Pacific and *Hilda* off the Australian west coast, at more than two weeks lead time. For the full season, November–April, verification showed that applying a 850-hPa wind speed threshold of 14 $m \cdot s^{-1}$ to the model TCs provided the greatest skill for forecasts of TC occurrence for weeks 1–4, with Brier skill scores (BSSs) exceeding that of the calibrated forecasts. However, for the shorter period November–February, as analysed in Camp *et al.* (2018), calibration still outperformed the raw forecasts and those with a wind speed threshold at weeks 3 and 4 (Gregory *et al.*, 2019). Following these results, a wind speed threshold and updated calibration for the full season (November–April) was implemented for the 2018/19 TC season.

Gregory *et al.* (2020) examined the skill of a multimodel ensemble to predict TC occurrence during the 2017/18 and 2018/19 TC seasons, by combining real-time forecasts from the ECMWF's Medium- and Extended-Range Ensemble Integrated Forecasting System (IFS) with ACCESS-S1. They found the multimodel skill to be superior to that of the individual component models, due to complimentary biases in TC tracks, as well as the larger ensemble size. Improvements were made to ACCESS-S1 forecasts by applying a lagged ensemble, with a two-day lag found to be optimal for weeks 2–3 and a five-day lag for week 4. Applying an 850-hPa wind speed threshold of 14 $m \cdot s^{-1}$ to the model TCs, as in Gregory *et al.* (2019), also improved skill for weeks 2–4. Finally, Gregory *et al.* (2020) showed that atmospheric data assimilation schemes had important implications for forecast skill for days 8–10.

Lee *et al.* (2020) examined the skill of calibrated forecasts of TC occurrence from six S2S models over the period 1981–2013 using various calibration techniques, including that used in Camp *et al.* (2018). They found that, for the S2S models, the basinwide skill of TC occurrence forecasts could vary by basin and lead time. In other words, using the calibration method, as in Camp *et al.* (2018), does not guarantee consistently improved skill across all basins and forecast lead times. This was demonstrated in Camargo *et al.* (2019) for real-time forecasts for the Southern Hemisphere 2018/19 TC season using ACCESS-S1. Indeed, Lee *et al.* (2020) showed that skill of forecasts of TC occurrence could be improved through calibration in regions where the mean forecast probability is lower than the observed probability. In areas where the forecast

probability is higher than observed, calibration could potentially result in lower values of BSS, resulting in lower basinwide BSS overall (Lee *et al.*, 2020). This suggests that solutions to calibrating subseasonal forecasts may not only be model-dependent, but also require basin-by-basin or more regional calibration techniques.

During the 2017/18–2018/19 TC seasons, the BoM made available trial multiweek TC probability forecasts for use by National Meteorological Services. These became routine operational products for the 2019/20 and 2020/21 TC seasons. In April 2021, ACCESS-S1 provided good guidance on severe tropical cyclone *Seroja*, which became the strongest southernmost TC to make landfall in Western Australia since 1956 (WMO, 2021), with three weeks lead time. Following the successful trials for the Southern Hemisphere, research was then extended to the NW Pacific basin, and skill was found over climatology for calibrated forecasts of TC occurrence out to week 4 (BoM, 2020).

In 2021 the BoM multiweek TC forecasts were made freely available to the public (see http://www.bom.gov.au/climate/pacific/outlooks/). Forecasts are currently issued for the South Pacific and Northwest Pacific Ocean basins for weeks 2–3 and 2–4, respectively. Forecasts are updated daily during a region's TC season and a two-week archive is also made available. Three products are provided: raw modekl probabilities of TC occurrence, calibrated probabilities (following Camp *et al.*, 2018) and the calibrated probability relative to observed climatology.

The BoM is required to deliver subseasonal to seasonal forecasts of TC activity for the Australian region and South Pacific, as part of improving national Climate and Disaster Resilience. It is therefore important that a full assessment of the skill of TC forecasts for the new version of ACCESS-S (ACCESS-S2; Wedd *et al.*, 2022), as well as a comparison with ACCESS-S1, is performed in order to retain operational skill and continuity of service.

In this study we examine the 11-member hindcast ensemble from ACCESS-S1 initialised on 1 November, December, January and February for the 23-year period 1990–2012, as in Camp *et al.* (2018), and directly compare these results to the nine-member ensemble available from the new ACCESS-S2 system. Section 2 outlines the models, data and method used. In Section 3 we compare model TC characteristics between ACCESS-S1 and ACCESS-S2, such as track frequency and location, and biases in key TC fields, such as vertical wind shear and mid-level relative humidity. In Section 4 we examine the MJO–TC relationship. In Section 5 we look at the skill of multiweek forecasts of TC occurrence over the Southern Hemisphere for ACCESS-S1 and ACCESS-S2. We then extend this analysis for ACCESS-S2 for the longer period, 1981–2018, in Section 6. The real-time performance of ACCESS-S2 to predict TC activity in the southwest Indian Ocean in February 2022 is investigated in Section 7. Conclusions are provided in Section 8.

2 | MODELS, DATA AND METHOD

2.1 | ACCESS-S

The BoM subseasonal (multiweek) to seasonal forecasts are produced operationally using the ACCESS-S system. ACCESS-S is a coupled ocean–atmosphere global ensemble prediction system, built around UK Met Office's Global Seasonal forecast system version 5 (GloSea5; MacLachlan *et al.*, 2015) with the Global Coupled model 2.0 (GC2; Williams *et al.*, 2015) configuration. ACCESS-S1 was used operationally from April 2018 to September 2021, including for forecasts of TC activity (Camp *et al.*, 2018; Camargo *et al.*, 2019; Gregory *et al.*, 2019; Gregory *et al.*, 2020). ACCESS-S1 has an atmospheric horizontal resolution of 0.83° longitude × 0.55° latitude (~60 km in the mid-latitudes) and 85 vertical levels extending into the stratosphere. This is coupled to an ocean model with a global resolution of 0.25° on a tripolar grid and 75 vertical levels.

In October 2021 the operational system was upgraded to ACCESS-S2, which has the same model configuration and resolution as its predecessor, but with a locally developed 3D-VAR data assimilation system for the ocean and improved land surface initialisation for soil moisture. It also has a longer hindcast period (38 years, 1981–2018) and is configured such that forecasts can be run out to lead times of several years to facilitate multiyear prediction. Relative to ACCESS-S1, ACCESS-S2 shows improved forecasts of the El Niño–Southern Oscillation (ENSO), which is important for TC prediction on seasonal timescales. Specifically, for multiweek TC forecasts, ACCESS-S2 maintains the high skill of forecasts of the MJO out to day 30 and shows reduced biases for sea surface temperature (SST) and subsurface temperature over the tropics (Wedd *et al.*, 2022).

2.2 | Data

Model TCs in both ACCESS-S1 and ACCESS-S2 are located and tracked in the atmospheric component of the model using the Okubo–Weiss–Zeta (OWZ) parameter TC detection scheme, which is fully discussed in Tory *et al.* (2013). This algorithm has been used operationally at the BoM for TC detection on multiweek timescales since the 2017/18 TC season (Gregory *et al.*, 2020). Briefly, OWZ identifies favourable environmental conditions for TC formation, including solid body rotation at 850 hPa (50×10^{-6} s^{-1})

and 500 hPa ($40 \times 10^{-6}\ s^{-1}$), positive relative humidity at 950 hPa (>65%) and 700 hPa (>50%), positive specific humidity at 950 hPa ($>10\ g \cdot kg^{-1}$), and low wind shear between 850 and 200 hPa ($<25\ m \cdot s^{-1}$). The initial thresholds for TC formation were determined through monitoring of pre-formation TC circulations in ERA-Interim reanalysis in coarse ($1° \times 1°$ degree)-resolution data (Tory *et al.*, 2013). When applied to ERA-Interim data for the 20-year period 1989–2008, the algorithm identified 95% of observed TCs globally. Favourable conditions must be satisfied for at least 48 h (THREE daily timesteps) before a TC is considered to have formed, thus TC tracking results are available from day 4 onwards. The thresholds for OWZ are model- and grid-resolution-independent; the same criteria are used for both models and in all ocean basins. Relative to Camp *et al.* (2018), the OWZ TC tracks data for ACCESS-S1 include one additional TC in December 1990.

For analysis of the observed MJO, we use the same methodology as in Camp *et al.* (2018). That is, the observed amplitude and phase of the MJO are defined as in Wheeler and Hendon (2004) using the two leading modes of an Empirical Orthogonal Function (EOF) analysis of outgoing longwave radiation (OLR) and zonal winds at 850-hPa and 200-hPa levels. Data are available in near real-time from the BoM (http://www.bom.gov.au/climate/mjo/). Observed MJO results for the period 1990–2012 have been updated relative to Camp *et al.* (2018) following revisions to the observational database from March 2003.

For the models, MJO amplitude and phase is computed as in Camp *et al.* (2018) by projecting the model zonal winds and OLR onto the observed EOF patterns following the procedure outlined in Gottschalck *et al.* (2010) and Rashid *et al.* (2011). For both the model and observations, we classify the MJO into eight phases, as in Wheeler and Hendon (2004), representing the daily location of peak convection as it moves eastwards from the Indian Ocean (Phases 2 + 3) through to the Maritime Continent (Phases 4 + 5), western Pacific (Phases 6 + 7) and the Western Hemisphere and Africa (Phases 8 + 1). To be classified as an MJO event we use an amplitude threshold of 1 or more and classify periods of smaller amplitude as 'weak MJO'.

Observed TCs for the Southern Hemisphere are obtained from the US Navy's Joint Typhoon Warning Centre (JTWC) best-track database (Chu *et al.*, 2002). There have been no changes to the JTWC observational dataset for 1990–2012. Therefore, the observed TC tracks used for this period are identical to those in Camp *et al.* (2018).

To assess the large-scale environment, we use global daily atmospheric and oceanic data from ERA-Interim (Dee *et al.*, 2011). Atmospheric data are retrieved at a spatial resolution of $0.75° \times 0.75°$ (~80 km) and ocean data at a resolution of 1°. Data are retrieved for both model hindcast periods: 1990–2012 and 1981–2018. For creation of model biases, the reanalysis datasets are regridded to match the model data prior to anomalies being calculated.

2.3 | Method

Daily TC track frequencies are computed by gridding the model and observed TC tracks into $4° \times 4°$ boxes and then dividing by the number of days in a season. For the MJO–TC anomalies, we use the same approach, but instead divide by the number of days in each particular MJO phase. We note that as a TC can span more than one MJO phase during its lifetime, a TC track may be split between multiple MJO phases. TC track anomalies are then obtained by subtracting climatology from the computed TC frequencies. Spatial anomalies in tropical cyclogenesis are examined using the same methodology using the first point of each TC track. For ACCESS-S1 and ACCESS-S2, this is the first point at which thresholds on the environmental conditions have been satisfied for at least 48 hr (three daily timesteps).

To examine multiweek forecast skill, we compute reliability diagrams and BSS for the raw model TC tracks for forecast days 4–7 (week 1), 8–14 (week 2), 15–21 (week 3), 22–28 (week 4) and 29–35 (week 5). The method used is the same as that in Camp *et al.* (2018). We start by binning daily TC track locations over the Southern Hemisphere (0°–30° S, 30°–240° E) into 60 boxes, each with size 15° latitude and 20° longitude and overlapping by 7.5° latitude and 10° longitude (see also Vitart *et al.*, 2010). For each start date (1 November, December, January and February) and each year (1990–2012) we then calculate the forecast probability of a TC occurrence for each region. This provides a total of 5520 (4 start dates × 23 years × 60 regions) forecasts. The weekly probability of a TC occurrence is given by the proportion of ensemble members which predict a TC occurring in that region. These are verified against observations (given by a zero or one).

To account for spatial biases in model TC track frequency, we also apply a calibration technique to the model forecasts. This scheme (CAL2-CV) is fully documented in Camp *et al.* (2018) and briefly described here. To create the calibrated forecast, the raw model forecast probabilities for each hindcast year are scaled by the ratio of the observed climate mean to the forecast mean over the entire hindcast excluding the year of interest (leave-one-out cross-validation). This scaling factor is computed for each of the 60 regions for each week of the forecast. Probabilistic forecast verification is then performed for the resulting calibrated probabilities.

For the full ACCESS-S2 hindcast assessment, we compute the skill of the raw and calibrated forecasts as above;

however, a greater number of years (1981–2018) and start dates (1 and 16th month from November–April) are used, providing a total of 27,360 (12 start dates × 38 years × 60 regions) forecasts.

3 | MODEL TC CLIMATOLOGY (1990–2012)

We first compare the skill of ACCESS-S1 and ACCESS-S2 to reproduce observed TC variability, tracks and intensity for forecasts starting 1 November, December, January and February over the period 1990–2012.

3.1 | TC tracks

The spatial distribution of TC genesis and tracks in ACCESS-S1 and ACCESS-S2 are compared with observations for forecast days 4–40 in Figure 1. The location of TC tracks (Figure 1a–e) are very similar between the two systems and correspond well with observations.

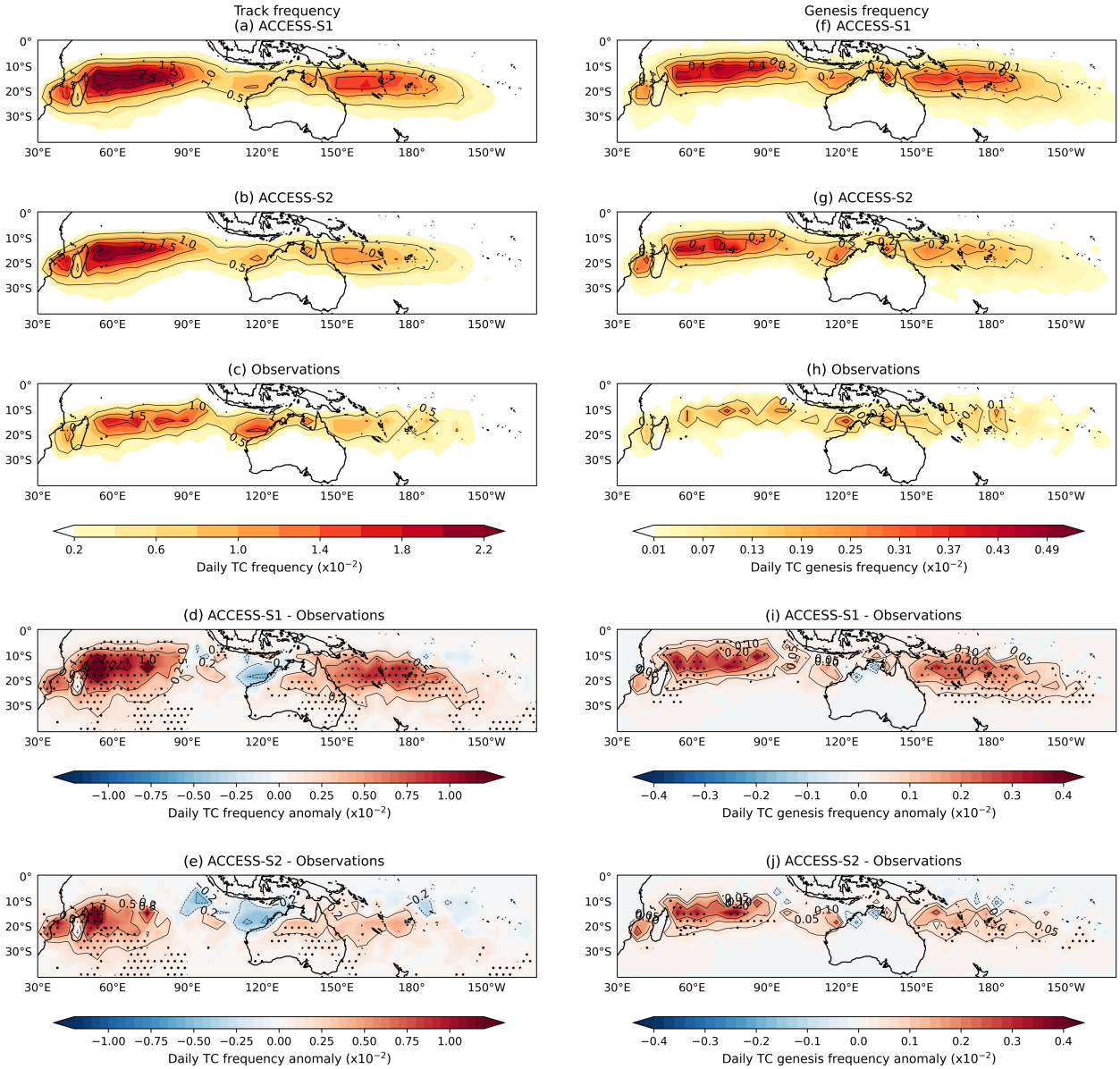

FIGURE 1 Climatology and difference in daily (a–e) tropical cyclone (TC) track and (f–j) genesis frequency for ACCESS-S1 and ACCESS-S2 relative to observations for forecast days 4–40 starting 1 November–February, 1990–2012. The total number of days used for ACCESS-S1 is 37,444 (23 years × 11 members × 4 start dates × 37 days), for ACCESS-S2 is 30,636 (23 years × 9 members × 4 start dates × 37 days) and for observations is 3,404 (23 years × 4 start dates × 37 days). Black dots indicate where the anomaly is significant at the 95% level. [Colour figure can be viewed at wileyonlinelibrary.com]

The significant positive bias in model TC tracks in the Southwest Indian Ocean and South Pacific is improved in ACCESS-S2, particularly in the South Pacific and off East Coast Australia. However, both systems undersimulate TC frequency off NW Australia, with ACCESS-S2 showing no improvement compared to ACCESS-S1, with a slightly larger region of this bias (not significant). In ACCESS-S1 this was attributed to a dry rainfall bias in this region (Lim *et al.*, 2016). The bias in TC tracks can be partly attributed to biases in TC genesis (Figure 1f–j), with both ACCESS-S1 and ACCESS-S2 showing significant positive anomalies in both the Southwest Indian Ocean and South Pacific. Nevertheless, the magnitude in TC genesis anomalies is reduced and improved relative to observations in ACCESS-S2.

The main difference between the two systems is in the total number of model TCs, with ACCESS-S2 showing fewer TCs in the Southern Hemisphere overall compared to ACCESS-S1. In ACCESS-S1, the total number of tracked TCs during the 1990–2012 period is 12,513 (average 12.4 TCs per member per start date) and in ACCESS-S2 it is 7,612 (average 9.2 TCs). The total number of observed storms is 435 (average 4.7 per start date). Therefore, despite the decrease in TC frequency, ACCESS-S2 is still simulating around twice the number of TCs than observed, the majority of these in the SW Indian Ocean.

Example TC tracks from ACCESS-S2 are compared against observations for forecast days 4–40 starting 1 November, December, January and February 1990 in Figure 2. A positive feature of ACCESS-S1 was its ability to reproduce a realistic representation of the observed location and direction of TC tracks (Camp *et al.*, 2018, their figure 3). This is also replicated in ACCESS-S2, with forecast tracks generally clustered in regions where an observed TC occurred. A noticeable difference between the two systems, however, is the reduced frequency of TC tracks in the eastern South Pacific in ACCESS-S2 from all start dates, due to an improvement in the model bias here. Both systems are able to simulate changes in TC track frequency between each month, with a greater number of TC tracks in January and February compared to November and December, in agreement with observations.

To investigate whether the seasonal cycle is captured across the full ensemble, we calculate the ensemble mean TC frequency per member per start date for the period 1990–2012 (Figure 3). The seasonal cycle of TC counts shows that both ACCESS-S1 and ACCESS-S2 produce too many TCs each month compared to observations, although this bias is improved in ACCESS-S2, with the greatest difference generally during peak season from 1 January and 1 February. Nevertheless, the observed progressive increase in TC frequency from 1 November to

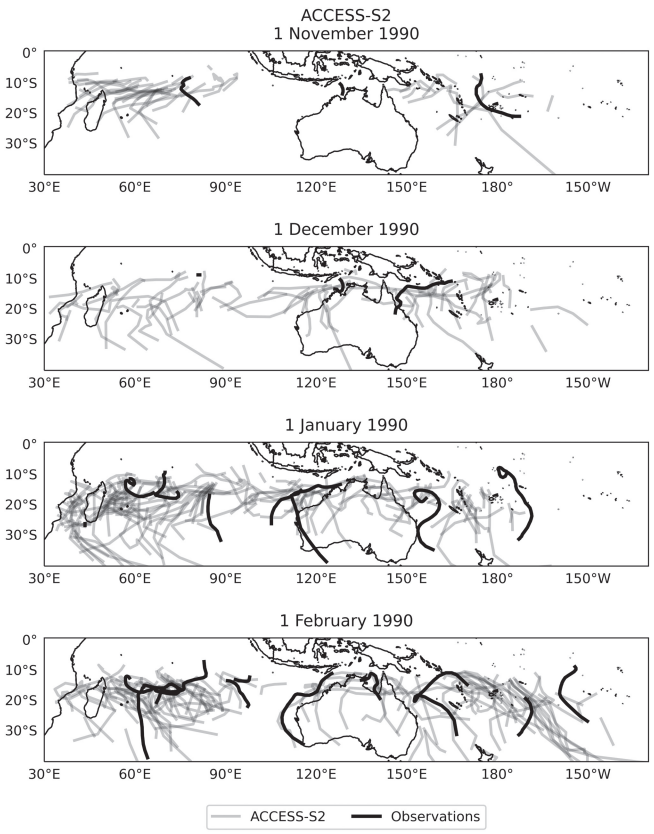

FIGURE 2 Observed (black lines) and forecast (grey lines) tropical cyclone (TC) tracks for ACCESS-S2 for days 4–40 starting 1 November–February, 1990. Note that there are more forecast tracks compared to observations due to the availability of a nine-member ensemble.

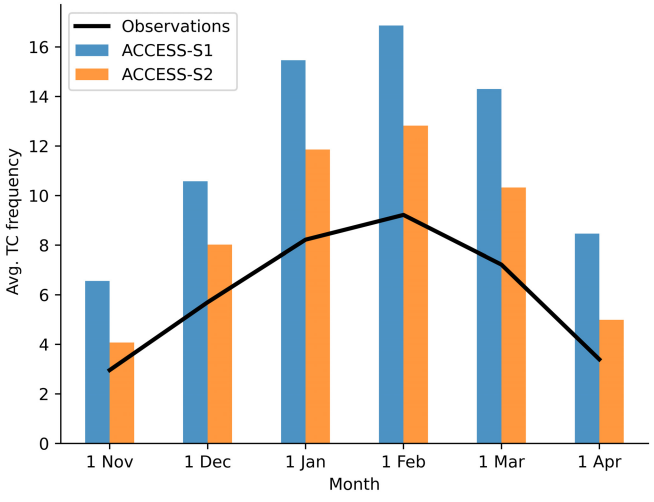

FIGURE 3 Seasonal cycle of tropical cyclone (TC) frequency over the Southern Hemisphere for ACCESS-S1, ACCESS-S2 and observations (black line) for days 4–40 starting 1 November–April, 1990–2012. Additional ACCESS-S1 data after Camp *et al.* (2018) have been produced in order to extend this analysis beyond 1 February. [Colour figure can be viewed at wileyonlinelibrary.com]

1 February, then a decrease from 1 February to 1 April is well captured by the models, showing the seasonal cycle is well simulated.

3.2 | TC intensity

A further characteristic of the model TCs we can investigate is intensity. A probability density function (PDF) of TC lifetime maximum intensity, as measured by maximum wind speed (v_{max}), is provided for ACCESS-S1, ACCESS-S2 and observations for forecast days 4–40 for the period 1990–2012 in Figure 4. The PDF shows a similar distribution of TC v_{max} for both ACCESS-S1 and ACCESS-S2, but with ACCESS-S2 *slightly* shifted to more intense TCs. For ACCESS-S1 the median v_{max} of the TCs is 15.1 ms^{-1} (with a minimum of 1.1 ms^{-1} and maximum of 45.1 ms^{-1}) and for ACCESS-S2 the median is 15.8 ms^{-1} (minimum of 2.2 ms^{-1} and a maximum 48.1 ms^{-1}). As to be expected, the models are unable to reproduce the full range of observed wind speeds, which have a median of 33.4 ms^{-1} (minimum of 10.3 ms^{-1} and maximum of 79.7 ms^{-1}), primarily because of the $1° \times 1°$ coarse resolution of the model wind speed data output from the OWZ tracking algorithm. It is also known (e.g., Williams *et al.*, 2015) that the Met Office GloSea5 system, from which ACCESS-S is based, is unable to reproduce TC wind speeds as high as in the real world, both at its native resolution (~55 km in the midlatitudes) and at higher resolution (~25 km in the midlatitudes; Scaife *et al.*, 2019).

We can further examine differences in characteristics between TCs in ACCESS-S1 and ACCESS-S2 by comparing parameters used as part of the OWZ tracking algorithm, which are associated with TC formation conditions. These include 850-hPa OWZ, 950-hPa specific humidity, 950-hPa relative humidity and 850-hPa maximum wind speed (Tory *et al.*, 2013). We examine each of these across the full lifetime of the model TCs. For ACCESS-S1 there are a total of 12,513 TCs with 78,372 daily track positions (average TC lifetime 6.3 days); for ACCESS-S2 there are 7,612 TCs with 48,984 daily track positions (average lifetime 6.4 days). Both ACCESS-S1 and ACCESS-S2 show very similar characteristics across each of these parameters (Figure 5), suggesting that the OWZ tracking algorithm has consistently identified similar types of TCs between the two systems. This is somewhat expected as the thresholds used to identify the TCs are identical for both ACCESS-S1 and ACCESS-S2. Both systems show a mean conducive environment for model TCs with an 850-hPa maximum wind speed of approximately 12.5–13 $m \cdot s^{-1}$, an average 950-hPa relative humidity of approximately 94%, 950-hPa specific humidity of approximately 16 $g \cdot kg^{-1}$ and 850-hPa OWZ of approximately 200×10^{-6} s^{-1}. The results therefore suggest that it is not differences in TC characteristics which are resulting in the overall lower TC count in ACCESS-S2 compared to ACCESS-S1.

3.3 | Model biases

To examine whether the reduced frequency of TC tracks in ACCESS-S2 relative to ACCESS-S1 is due to differences in model biases, we compare key components of the OWZ tracking algorithm – 850–200-hPa wind shear, 600-hPa

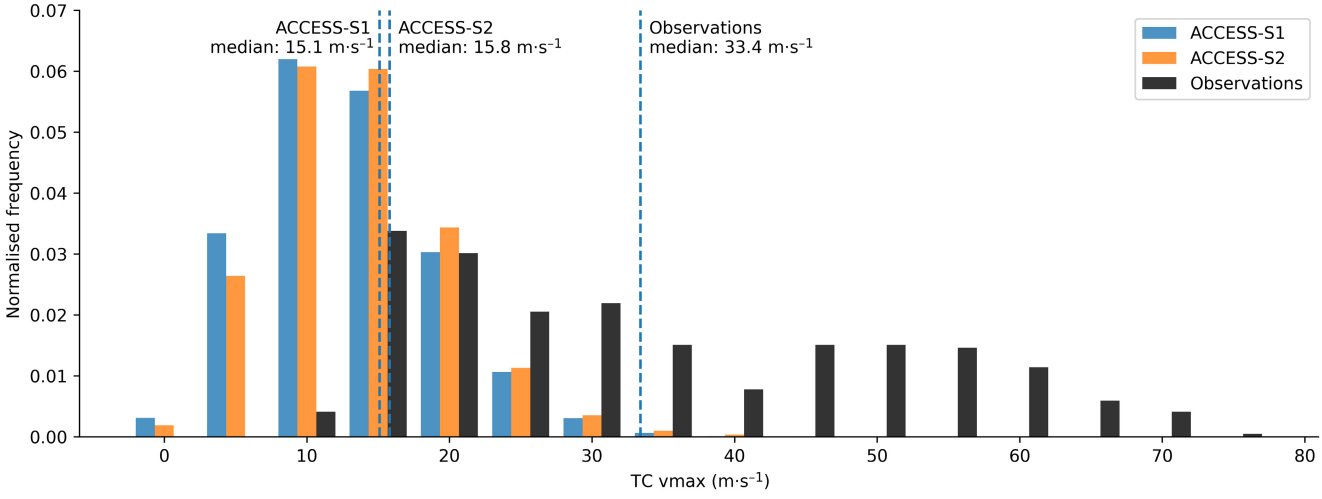

FIGURE 4 Lifetime maximum intensity (as measured by wind speed; ms^{-1}) for ACCESS-S1, ACCESS-S2 and observed tropical cyclones (TCs) over the Southern Hemisphere for days 4–40 starting 1 November–February, 1990–2012. Model winds are $1° \times 1°$ at 850 hPa. Observations are 1-min maximum sustained winds reported at 00Z from JTWC. Observed winds are converted from kts to ms^{-1} using the conversion 1 kt = 0.514 ms^{-1}. [Colour figure can be viewed at wileyonlinelibrary.com]

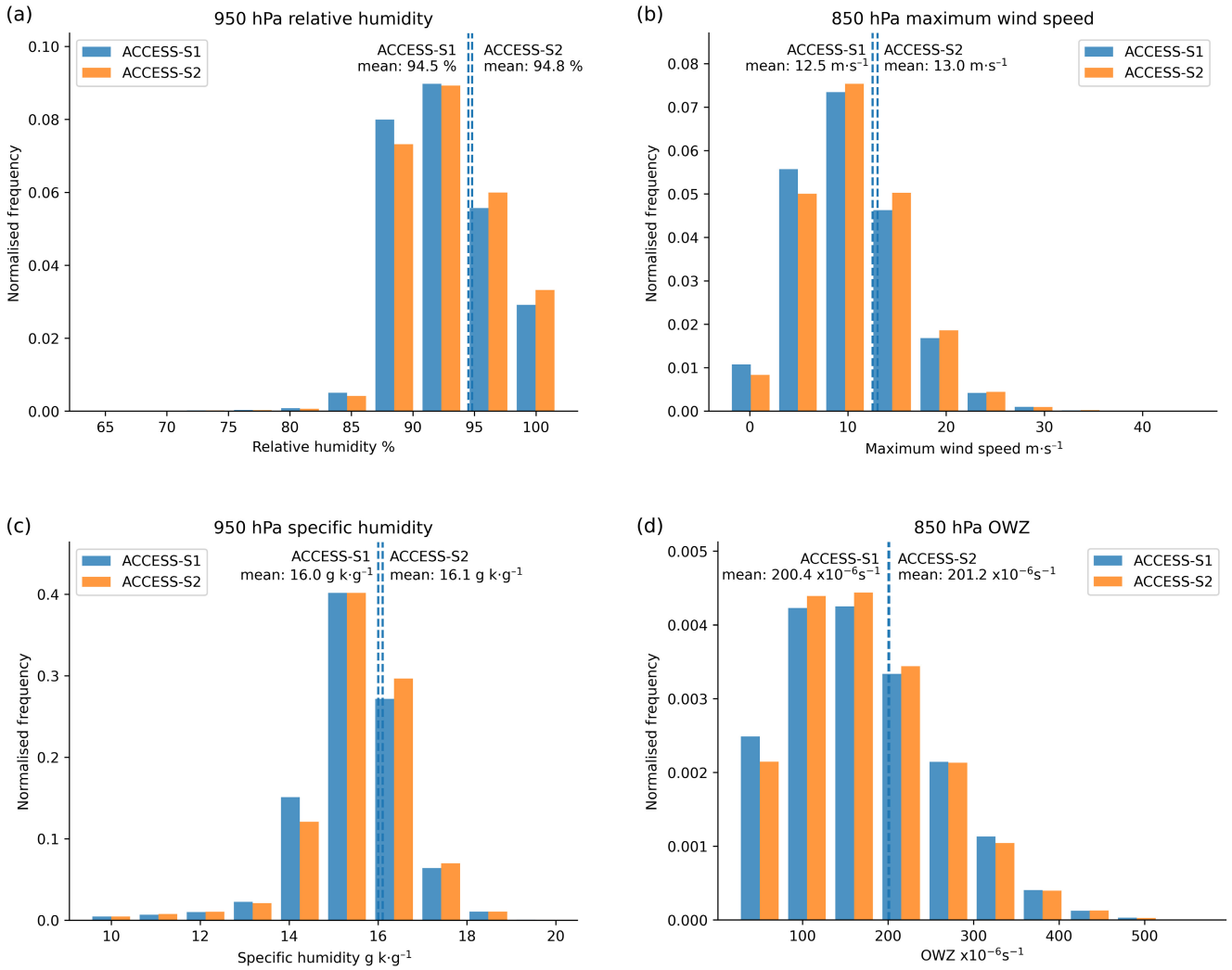

FIGURE 5 Tropical cyclone (TC) average (a) 950-hPa relative humidity (%), (b) 850-hPa maximum wind speed ($m{\cdot}s^{-1}$), (c) 950-hPa specific humidity ($g{\cdot}kg^{-1}$) and (d) maximum 850-hPa Okubo–Weiss–Zeta (OWZ) parameter ($\times 10^{-6}\ s^{-1}$) for ACCESS-S1 and ACCESS-S2 for days 4–40 starting 1 November–February, 1990–2012. All model fields are at $1° \times 1°$ resolution. The total number of daily track positions is 80,509 for ACCESS-S1 and 50,273 for ACCESS-S2. [Colour figure can be viewed at wileyonlinelibrary.com]

relative humidity and 850-hPa absolute vorticity, as well as daily precipitation and SSTs – between ACCESS-S1 and ACCESS-S2 relative to ERA-Interim reanalysis for days 4–40, 1990–2012.

Both ACCESS-S1 and ACCESS-S2 show very similar biases for each parameter relative to ERA-Interim, both in terms of location and magnitude (Figure 6). In particular, regions of significantly enhanced TC formation and tracks relative to observations (Figure 1) generally coincide with regions of significantly enhanced 600-hPa relative humidity (South Pacific and Southwest Indian Ocean) and negative 850-hPa absolute vorticity (South Pacific). This suggests that the moist atmosphere may be a main contributor to the significantly enhanced TC genesis and tracks in ACCESS-S1 and ACCESS-S2 in these regions. Nevertheless, both systems show significant biases in the deep tropics and far northern Australia relative to ERA-Interim. Here, both a negative precipitation (Figure 6g,h) and positive wind shear (Figure 6c,d) bias may be acting to reduce TC formation rates, particularly over northern and NW Australia. Reasons for the reduced TC count in ACCESS-S2 relative to ACCESS-S1 may be due to the increase in the magnitude and extent of the significant negative wind shear bias around 20–30° S (Figure 6d), as well as a small reduction in the positive SST bias (not significant) around the South Pacific Islands (Figure 6j), both of which may act to reduce TC frequency in the SW Indian and South Pacific Oceans.

In addition to model biases versus ERA-Interim, we examine differences in the large-scale fields between ACCESS-S1 and ACCESS-S2 directly. For each field we compare the mean for days 4–40 starting 1 November,

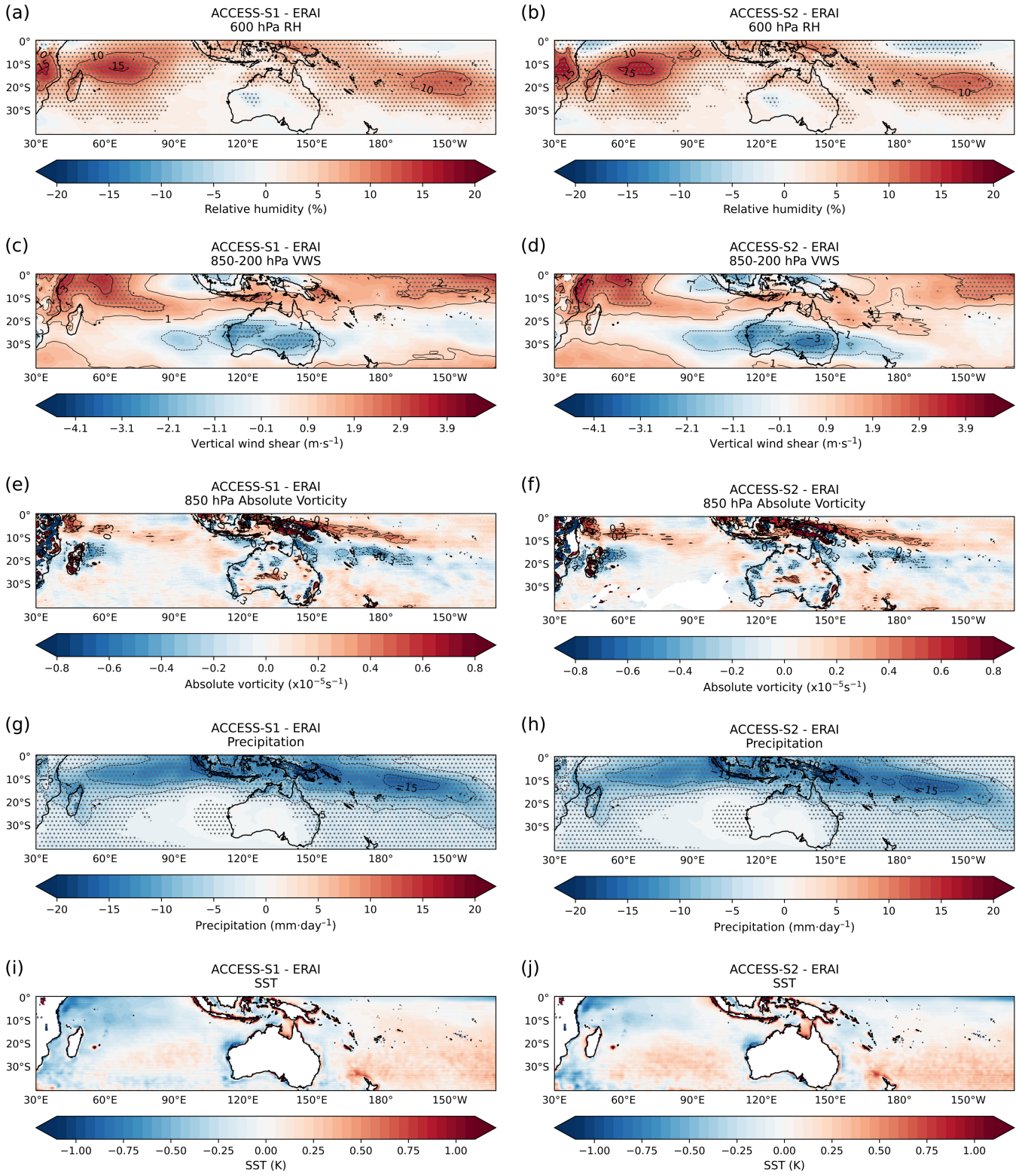

FIGURE 6 Daily mean difference in (a, b) 600-hPa relative humidity (%), (c, d) 850–200-hPa wind shear ($m \cdot s^{-1}$), (e, f) 850-hPa absolute vorticity ($\times 10^{-5}\ s^{-1}$), (g, h) precipitation ($mm \cdot day^{-1}$) and (i, j) sea-surface temperatures (SSTs) (K) for (left) ACCESS-S1 and (right) ACCESS-S2 versus ERA-Interim climatology for days 4–40 starting 1 November–February, 1990–2012. The total number of days used for ACCESS-S1 is 37,444 (23 years × 11 members × 4 start dates × 37 days), for ACCESS-S2 is 30,636 (23 years × 9 members × 4 start dates × 37 days) and for ERA-Interim is 3,404 (23 years × 4 start dates × 37 days). Black dots indicate where the anomaly is significant at the 95% level. [Colour figure can be viewed at wileyonlinelibrary.com]

December, January and February 1990–2012 and compute anomalies by subtracting the mean of ACCESS-S1 from ACCESS-S2 (Figure 7). Relative to ACCESS-S1, ACCESS-S2 shows a small reduction in 600-hPa relative humidity, precipitation and SST, and a small increase in 850–200-hPa wind shear in the South Pacific, all of which may be contributing to the reduced TC count in this region. Around NW Australia, neutral to slightly negative SSTs (Figure 7e), and in the southwest Indian Ocean a small increase in wind shear may be contributing to the decrease in TC genesis and tracks in ACCESS-S2 in these regions. Overall, however, we find no significant differences between ACCESS-S1 and ACCESS-S2 in the Southern Hemisphere, except for 600-hPa relative humidity in two regions close to the equator in the southwest Indian Ocean (centred ~60° E) and South Pacific (centred ~180°), and in precipitation off the west coast of Australia. Whether the significant negative biases in relative humidity contribute to the reduced TC genesis and tracks in the Southwest Indian Ocean and South Pacific in ACCESS-S2 warrants further research.

4 | MJO–TC RELATIONSHIP

The MJO is a key driver of TC activity on subseasonal timescales. Thus, the ability of the models to simulate the observed MJO–TC relationship is extremely important for reliable forecasts on this timescale. Here we compare the models' MJO–TC relationship, as well as the MJO

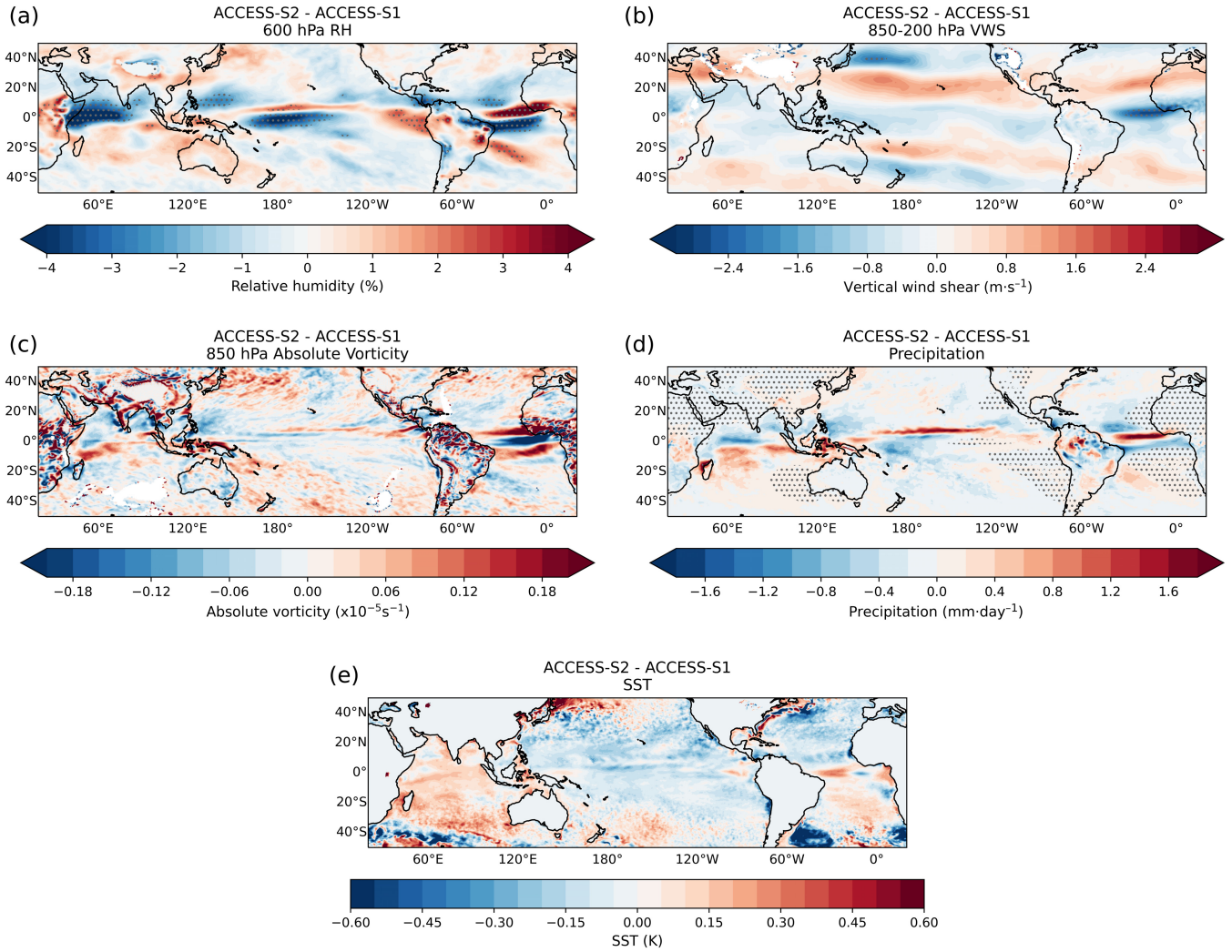

FIGURE 7 Daily mean difference in model (ACCESS-S2 minus ACCESS-S1) (a) 600-hPa relative humidity (%), (b) 850–200-hPa wind shear, (c) 850-hPa absolute vorticity, (d) precipitation, and (e) sea-surface temperatures (SSTs) for days 4–40 starting 1 November–February, 1990–2012. The total number of days used for ACCESS-S1 is 37,444 (23 years × 11 members × 4 start dates × 37 days) and for ACCESS-S2 is 30,636 (23 years × 9 members × 4 start dates × 37 days). Black dots indicate where the anomaly is significant at the 95% level. [Colour figure can be viewed at wileyonlinelibrary.com]

modulation of large-scale fields: 600-hPa relative humidity, 850-hPa absolute vorticity, 850–200-hPa vertical wind shear, SST and precipitation, for forecast days 15–40 from 1 November, December, January and February, 1990–2012.

4.1 | MJO-MODULATED TC TRACKS

Daily average TC track frequency anomalies for each phase of the MJO are shown for ACCESS-S1, ACCESS-S2 and observations in Figure 8. A key feature of ACCESS-S1 was its ability to reproduce the change in TC frequency with each MJO phase (Camp *et al.*, 2018). The eastward propagation of regions of significantly enhanced and suppressed TC activity, coinciding with the eastward advancement of the MJO, is also well captured by ACCESS-S2, and the pattern of anomalies is very similar between the systems with a couple of exceptions. For example, regions of significantly suppressed TC activity are slightly reduced in magnitude in ACCESS-S2 relative to ACCESS-S1, apart from in Phase 2 + 3. Here a stronger suppression of TC frequency occurs off East Coast Australia, as seen in observations, showing an improvement over ACCESS-S1. Around NW Australia (Phase 4 + 5), the South Pacific Islands (Phase 8 + 1) and in the southwest Indian Ocean (Phase 2 + 3), the magnitude of the significant positive anomalies also appear to be reduced in ACCESS-S2 relative to ACCESS-S1 and observations. A weaker MJO–TC relationship may be one of the reasons why ACCESS-S2 has fewer TC tracks in these regions compared to ACCESS-S1 (Figure 1). Nevertheless investigating the role of the TCs on these anomalies is necessary to understand these biases further.

4.2 | MJO modulated large-scale fields

We now examine the modulation of large-scale fields in the ocean and atmosphere with the phase of the MJO. These include 600-hPa relative humidity, 850-hPa absolute vorticity, SST, 850–200-hPa wind shear and precipitation. Model results are compared to ERA-Interim in Figures 9–11.

The observed pattern and magnitude of 600-hPa relative humidity anomalies are well captured by ACCESS-S1 and ACCESS-S2 (Figure 9a–c). The eastward propagation of anomalously high relative humidity, coinciding with regions of significantly enhanced TC frequency, was shown to be particularly well captured by ACCESS-S1 in Camp *et al.* (2018), albeit with slightly reduced magnitude compared to ERA-Interim. ACCESS-S2 produces a similar pattern and magnitude of significant relative humidity

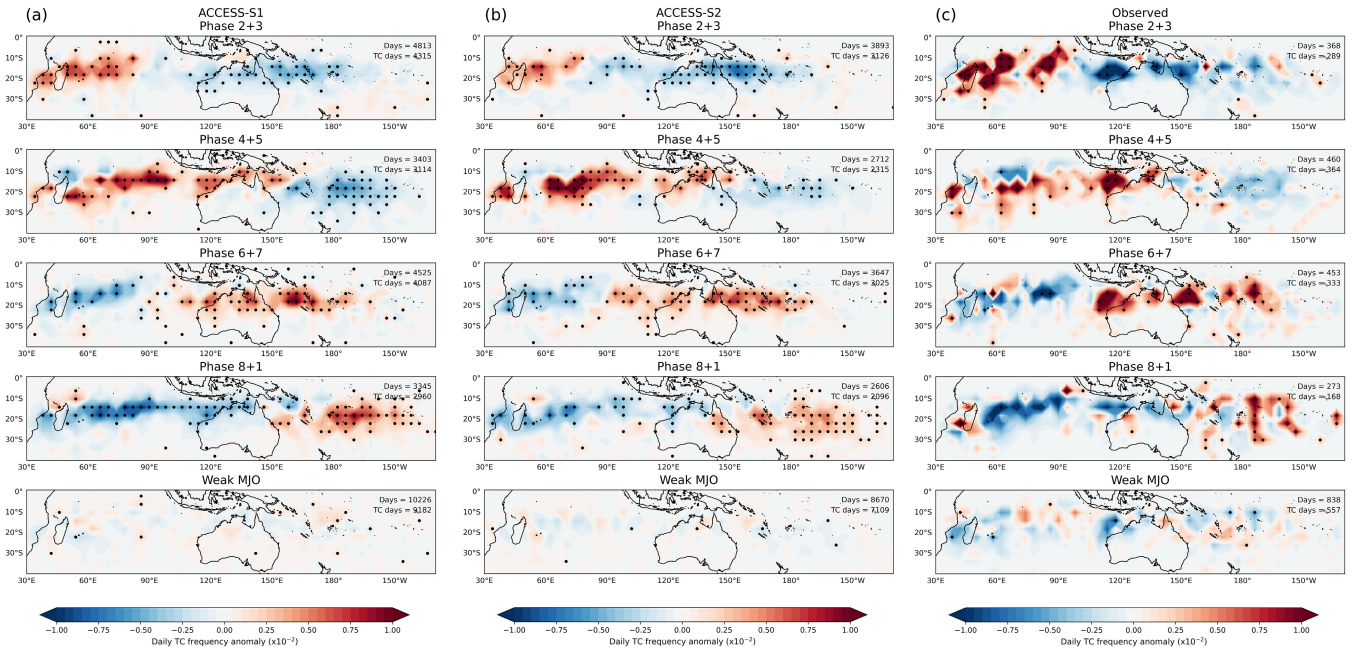

FIGURE 8 Daily tropical cyclone (TC) track frequency anomalies for (a) ACCESS-S1, (b) ACCESS-S2 and (c) observations for (top to bottom) Madden–Julian Oscillation (MJO) Phases 2 + 3 (Indian Ocean), 4 + 5 (Maritime Continent), 6 + 7 (western Pacific), 8 + 1 (Western Hemisphere and Africa) and weak MJO (amplitude <1) for forecast days 15–40 starting 1 November–February, 1990–2012. Black dots indicate where the TC track frequency anomaly is significant at the 95% level. The total number of days used for ACCESS-S1 is 26,312 (23 years × 11 members x 4 start dates x 26 days), for ACCESS-S2 is 21,528 (23 years × 9 members × 4 start dates × 26 days) and for observations is 2,392 (23 years × 4 start dates × 26 days). [Colour figure can be viewed at wileyonlinelibrary.com]

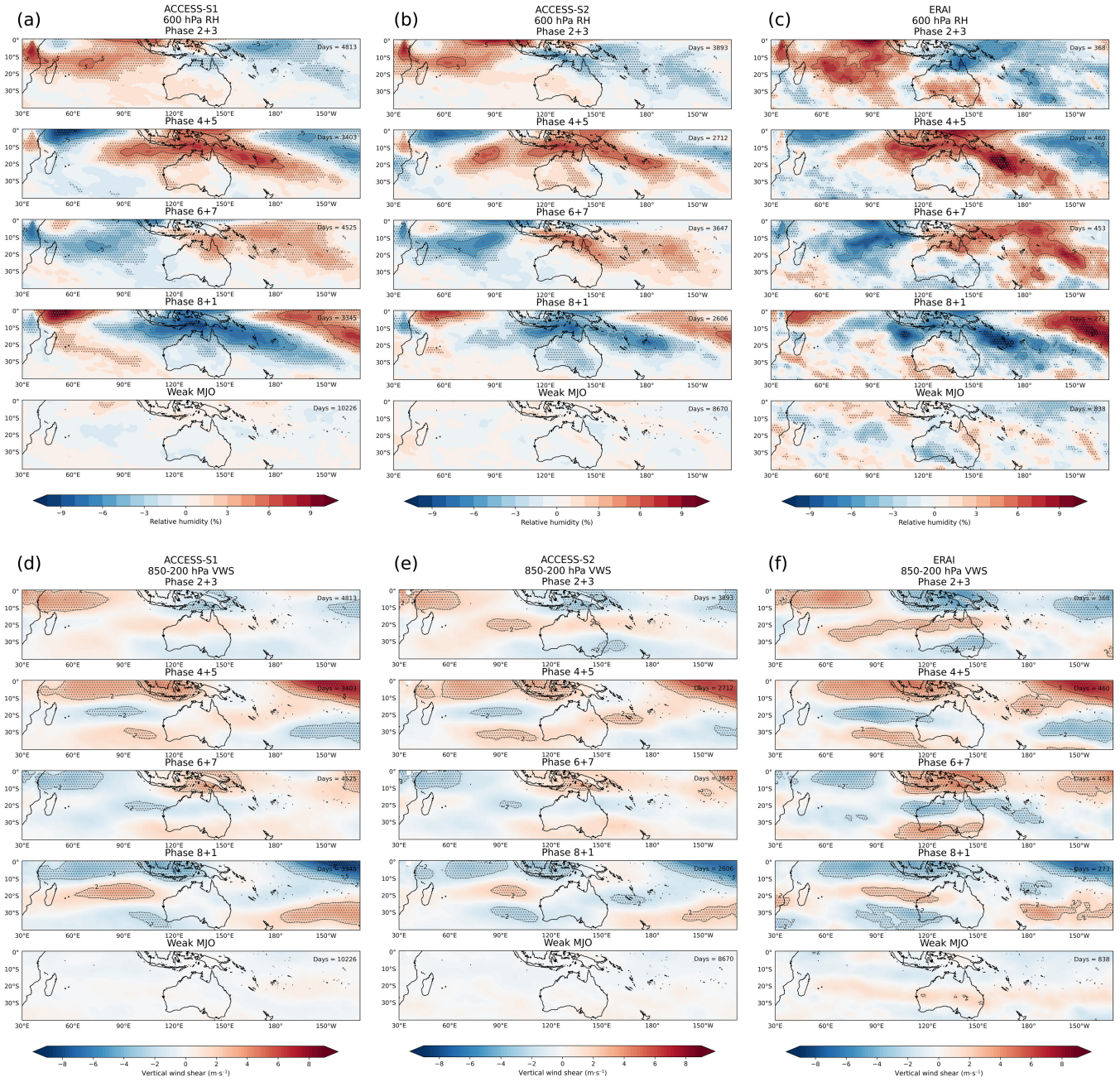

FIGURE 9 Daily mean (a–c) 600-hPa relative humidity (%) and (d–f) 850–200-hPa wind shear ($m \cdot s^{-1}$) anomalies for (left to right) ACCESS-S1, ACCESS-S2 and ERA-Interim for (top to bottom) Madden–Julian Oscillation (MJO) Phases 2 + 3, 4 + 5, 6 + 7, 8 + 1 and weak MJO (amplitude <1) for days 15–40 starting 1 November–February, 1990–2012. The total number of days used for ACCESS-S1 is 26,312 (23 years × 11 members × 4 start dates × 26 days), for ACCESS-S2 is 21,528 (23 years × 9 members × 4 start dates × 26 days) and for ERA-Interim is 2,392 (23 years × 4 start dates × 26 days). Black dots indicate where the anomaly is significant at the 95% level. [Colour figure can be viewed at wileyonlinelibrary.com]

anomalies to ACCESS-S1 (Figure 9b). However, during Phase 4 + 5 and Phase 6 + 7, the magnitude of the anomaly to the north and NW of Australia is reduced compared to ERA-Interim (and is weaker compared to ACCESS-S1 in Phase 4 + 5). This may be the primary reason why ACCESS-S2 produces fewer TCs around NW Australia compared to ACCESS-S1, as it is during Phase 4 + 5 and 6 + 7 that the greatest positive anomalies in observed TC frequency occur.

The pattern of 850–200-hPa wind shear was shown in Camp *et al.* (2018) to be more complex and not exhibit a strong eastward progression coinciding with enhanced TC frequency. Indeed, for some phases (e.g., Phase 2 + 3) the wind shear can act to reduce rather than enhance TC

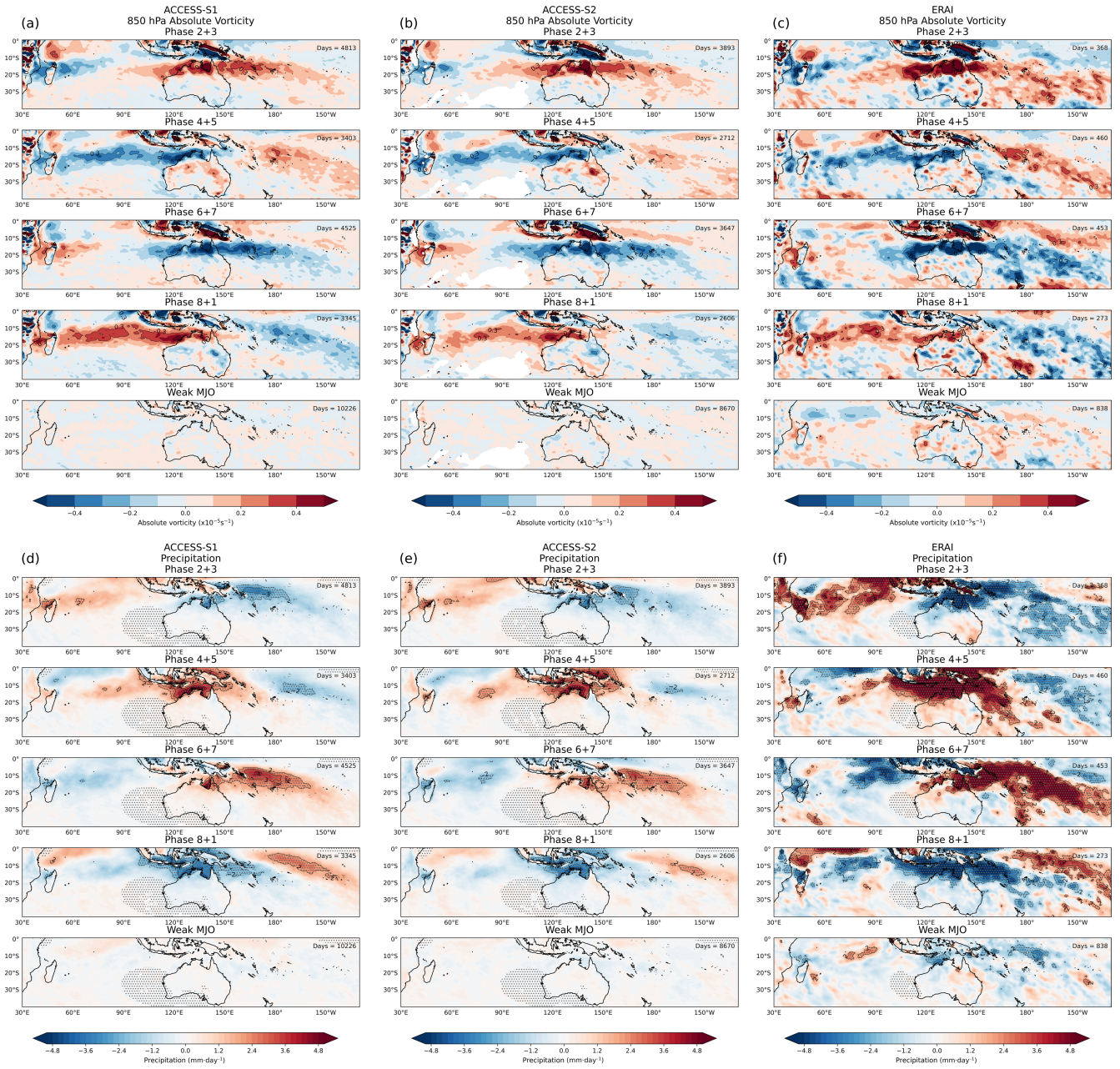

FIGURE 10 As in Figure 9, but for daily mean (a–c) 850-hPa absolute vorticity ($\times 10^{-5}$ s^{-1}) and (d–f) precipitation (mm·day^{-1}) anomalies. [Colour figure can be viewed at wileyonlinelibrary.com]

activity (e.g., Camargo *et al.*, 2009). The pattern of anomalies as seen in ERA-Interim is generally well captured by the models (Figure 9d–f). However, the magnitude is often reduced, particularly around Australia. In Phase 8 + 1, ACCESS-S2 has a smaller region of significant positive wind shear across the SW Indian Ocean relative to both ACCESS-S1 and ERA-Interim. This could be one of the reasons why ACCESS-S2 retains the positive anomaly in TC frequency over this region, as it is during Phase 8 + 1 that the greatest observed suppression in TC activity occurs.

The location and magnitude of 850-hPa absolute vorticity (Figure 10a–c) and precipitation (Figure 10d–f) anomalies are also similar between ACCESS-S1 and ACCESS-S2, with both simulating well-observed regions of significantly enhanced and reduced anomalies with the phase of the MJO (Figure 10). For example, in both systems, negative absolute vorticity anomalies and positive precipitation anomalies coincide with increased TC track frequency. However, precipitation anomalies are generally reduced in magnitude in the models compared to ERA-Interim. This may in part be due to the dry

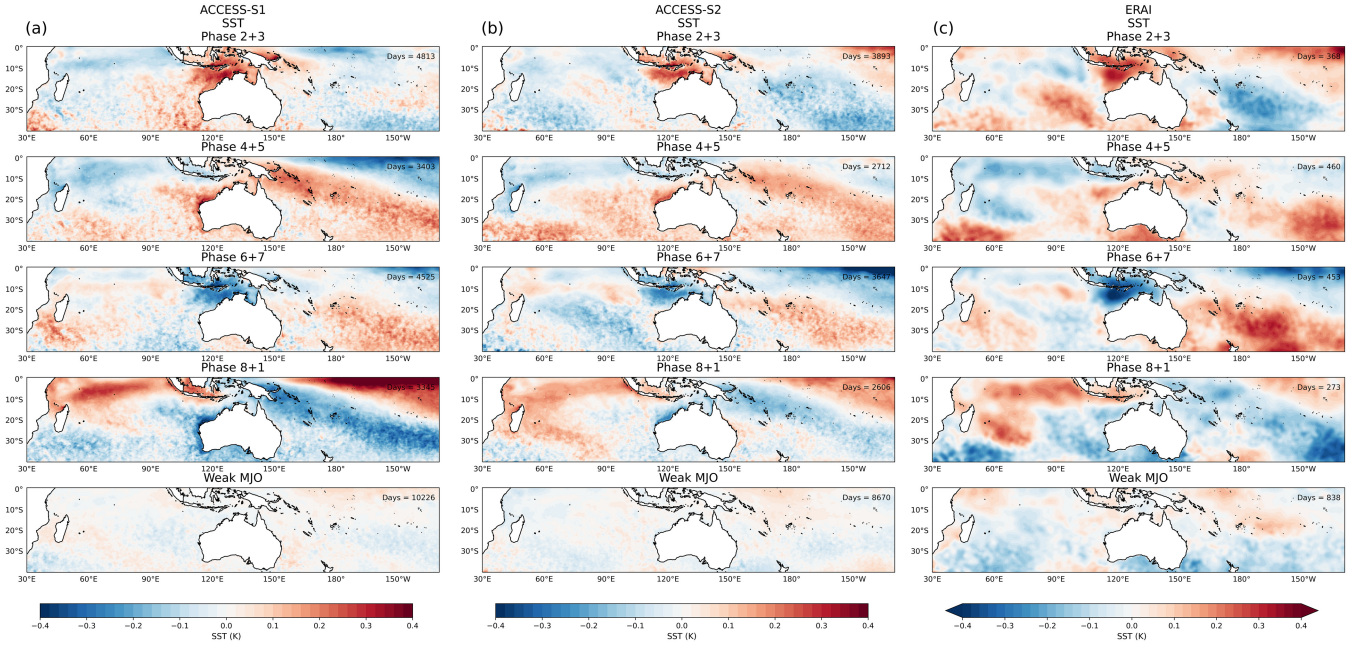

FIGURE 11 As in Figure 9, but for daily mean (a–c) sea surface temperature (SST, K) anomalies. [Colour figure can be viewed at wileyonlinelibrary.com]

precipitation bias (Figure 6g,h), which is present in both ACCESS-S1 and ACCESS-S2 (e.g. Lim *et al.*, 2016; Wedd *et al.*, 2022).

The MJO modulation of SSTs is shown in Figure 11. SST anomalies do not show a strong eastward progression with the phase of the MJO; nevertheless, the pattern and location of anomalies is well captured by both ACCESS-S1 and ACCESS-S2. For example, around NW Australia, both systems are able to capture the strong positive SST anomalies (not significant) during Phase 2 + 3 and negative anomalies during Phase 6 + 7. Similarly, in the South Pacific, the negative (positive) SST anomalies during Phase 2 + 3 and 8 + 1 (4 + 5 and 6 + 7) are also reasonably well captured. Indeed, ACCESS-S2 shows an improvement in the magnitude of the negative SST anomaly in the South Pacific during Phase 2 + 3. This may be one of the factors contributing to the reduced bias in TC frequency in this region.

5 | RELIABILITY AND BSS

We now compare the skill of ACCESS-S1 and ACCESS-S2 to provide multiweek probabilistic forecasts of TC occurrence for weeks 1–5. We first investigate the performance of the raw uncalibrated forecasts (UNCAL), before applying calibration method CAL2-CV, as in Camp *et al.* (2018). This calibration technique takes each forecast probability and scales it by the ratio of the climate mean to the forecast mean for each region individually, thus providing a more regional calibration technique. BSSs and reliability are computed for raw model tracks (UNCAL) and CAL2-CV forecasts for 1 November–February 1990–2012.

5.1 | Raw model tracks

The reliability and skill of the raw model forecasts (UNCAL) for predictions of TC occurrence in ACCESS-S1 and ACCESS-S2 are provided for weeks 1–5 in Figure 12 and Table 1. The raw forecasts for ACCESS-S1 exhibit positive BSSs for weeks 1–2, showing skill over climatology for this timerange; for weeks 3–5, the BSSs of the forecasts are negative, possibly due to the overprediction of TCs contributing to a high false alarm rate (e.g., Gregory *et al.*, 2020). ACCESS-S2 shows improved skill over ACCESS-S1, with positive BSSs for all weeks (1–5) and greater positive skill in weeks 1 and 2. Overall, ACCESS-S2 shows good reliability in week 1 (with the majority of forecast probabilities lying close to the 45° line) and sharpness out to week 5. In comparison to ACCESS-S1, which predominantly showed underconfidence in the forecasts across all forecast probabilities and lead times, ACCESS-S2 shows overconfidence at low probabilities during weeks 1–2. Similarly to Camp *et al.* (2018), we conclude that this may also be improved with appropriate calibration.

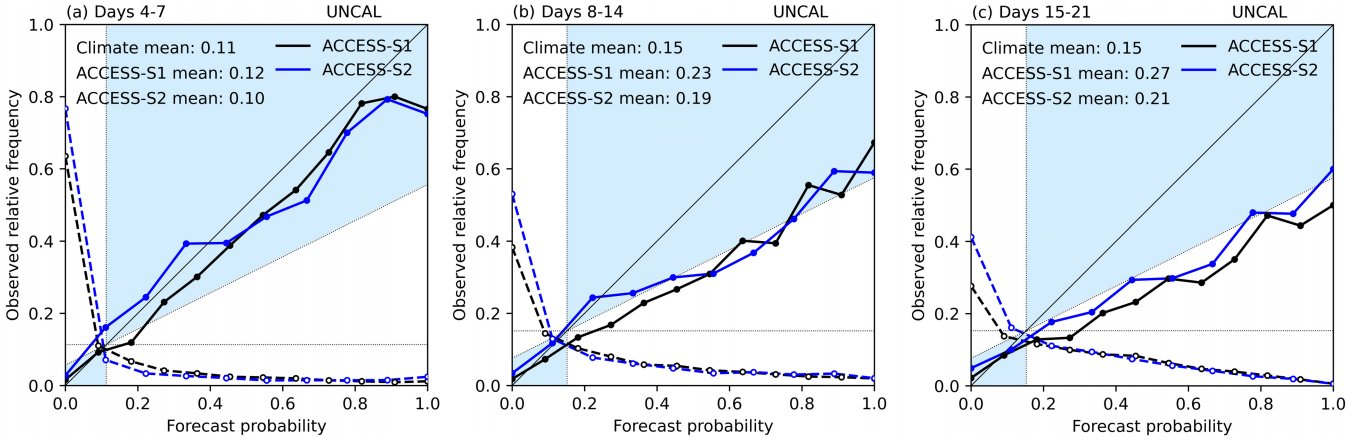

FIGURE 12 Reliability diagrams showing the probability of tropical cyclone (TC) occurrence in the Southern Hemisphere in ACCESS-S1 and ACCESS-S2 for raw modekl forecasts (UNCAL) for forecast days (a) 4–7 (week 1), (b) 8–14 (week 2) and (c) 15–21 (week 3), starting 1 November–February, 1990–2012. The 45° diagonal line indicates perfect reliability. The horizontal and vertical black dotted lines denote the observed climatological frequency; the diagonal dotted line between the 45° diagonal and the observed climatological frequency is the no-skill line. Points located in the grey-shaded region contribute positively to the Brier skill score. The dashed lines show the relative frequency with which each probability bin was forecast (sharpness). [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Brier skill scores (BSSs) for the uncalibrated (UNCAL) probability of tropical cyclone (TC) occurrence in the Southern Hemisphere in ACCESS-S1 and ACCESS-S2 for 60 regions over the Southern Hemisphere for forecast days 4–7 (week 1), 8–14 (week 2), 15–21 (week 3), 22–28 (week 4) and 29–35 (week 5) starting 1 November–February 1990–2012

	Week 1	Week 2	Week 3	Week 4	Week 5
ACCESS-S1	0.326	0.057	−0.137	−0.162	−0.076
ACCESS-S2	0.327	0.070	−0.031	−0.038	−0.034

5.2 | Calibrated data (CAL2-CV)

For calibration method CAL2-CV (Figure 13 and Table 2), ACCESS-S1 shows greater BSSs for all weeks compared to UNCAL, with skill of forecasts of TC occurrence above climatology out to week 5, mirroring the conclusions of Camp *et al.* (2018). For ACCESS-S2, improvement is also found with method CAL2-CV in all weeks apart from week 1. Here, the BSS is slightly reduced after calibration, although it is still positive showing skill over climatology. Nevertheless, even after calibration, the skill of ACCESS-S2 is reduced relative to ACCESS-S1 for all weeks. Therefore, compared to UNCAL, for which ACCESS-S2 showed superior skill compared to ACCESS-S1, the opposite is the case after calibration, with ACCESS-S2 showing reduced skill compared to ACCESS-S1 for lead time weeks 1–5. This may, in part, be due to the smaller number of ensemble members available for ACCESS-S2. It may also be that this system would benefit from removing weaker model TCs (with a lifetime maximum wind speed $<14\,\mathrm{m\cdot s^{-1}}$) as in Gregory *et al.* (2019, 2020) and use of a lagged ensemble (Gregory *et al.*, 2020). Both of these will be investigated in the following section, where we explore the skill of the full ACCESS-S2 ensemble from 1981 to 2018.

6 | FORECAST SKILL FOR ACCESS-S2 (1981–2018)

We now examine the skill of forecasts of TC occurrence for ACCESS-S2 for the longer hindcast period using start dates of 1st and 16th of the month for November–April 1981–2018. Skill is compared for (1) raw model forecasts, (2) raw model forecasts with a wind speed threshold, (3) calibrated forecasts using method CAL2-CV, and (4) a linear regression-based calibration for lead time weeks 1–5. For (2), any model TC with a lifetime maximum 850-hPa wind speed lower than $14\,\mathrm{m\cdot s^{-1}}$ is discarded, as applied in Gregory *et al.* (2019, 2020). Reliability diagrams compare the results from each method for weeks 1–3 (Figure 14); corresponding BSSs for weeks 1–5 are shown in Table 3.

6.1 | Raw model skill

For the raw model forecasts (UNCAL), results show that ACCESS-S2 exhibits skill over climatology for weeks 1–2. For all other lead times the BSSs are negative, showing climatology provides a more skilful forecast. When compared to the shorter period (November–February 1990–2012), as examined previously, the UNCAL skill is reduced in ACCESS-S2 for all lead times (weeks 1–5). When a wind speed threshold is applied to the model TCs, ACCESS-S2

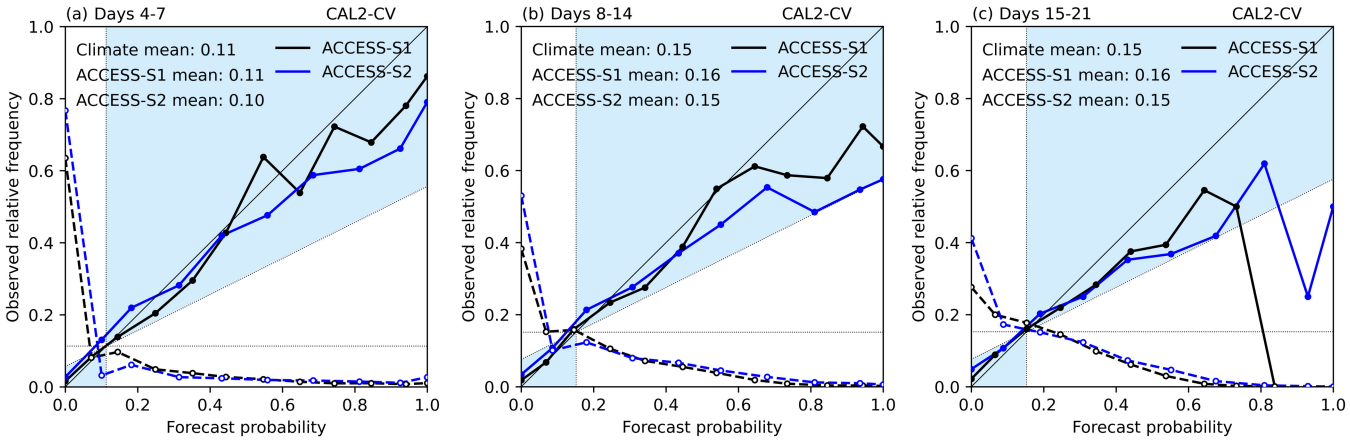

FIGURE 13 As in Figure 12, but for calibrated forecasts using method CAL2-CV, 1990–2012. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 As Table 1, but for calibrated (CAL2-CV) probabilities of tropical cyclone (TC) occurrence

	Week 1	Week 2	Week 3	Week 4	Week 5
ACCESS-S1	0.335	0.197	0.095	0.060	0.087
ACCESS-S2	0.315	0.151	0.063	0.040	0.016

shows skill for forecasts of TC occurrence out to week 4. In week 5, climatology still provides a more skilful forecast. Applying the wind speed threshold at week 1 degrades forecast skill compared to using the raw model tracks, possibly due to the forecasts being more overconfident at lower probabilities (Figure 14).

6.2 | Calibrated skill

For calibrated forecasts (CAL2-CV), ACCESS-S2 is shown to have skill for forecasts of TC occurrence over climatology out to week 5. For lead time weeks 2–5, the calibrated forecasts show the greatest skill; use of the raw model probabilities is still more beneficial at week 1. Compared to the shorter period 1990–2012, the skill of calibrated forecasts from ACCESS-S2 is reduced for weeks 1–2, whereas it is marginally improved for longer lead times (weeks 3–5). The skill of calibrated forecasts with a wind speed threshold were also examined; however, it did not improve forecast skill for any lead time (not shown).

6.3 | Lagged ensembles

Gregory *et al.* (2020) showed that forecast skill for ACCESS-S1 could be improved by using a lagged ensemble. To see whether this also applies to ACCESS-S2, we combine older hindcast start dates with those starting on the 1st and 16th of each month. Here we use a lag of one and two days, where lag 1 is the addition of ensemble members from the day before (total 18 members per forecast start date per year), and lag 2 is the addition of ensemble members from one and two days before (total 27 members per forecast start date per year).

We find that using a lagged ensemble of one or two days does not increase the skill of the raw forecasts or those with a wind speed threshold applied, except at week 5; however, skill does not exceed that of climatology (not shown). For the calibrated forecasts, a lagged ensemble of one or two days marginally improves forecast skill at week 5 (not shown). This is somewhat expected, as the skill of short-range forecasts (particularly those to week 4) would be more greatly impacted by using older forecasts in a lagged ensemble than those at longer lead times (Vitart and Takaya, 2021). Nevertheless, the improvements seen in Gregory *et al.* (2020) are not replicated here. This could be for a number of reasons: (1) the number of ensemble members is substantially reduced compared to those used in Gregory *et al.* (2020); in this study, a total of 66 members were available for lag-1 forecasts and 99 members for lag-2 forecasts, since they utilised the operational forecasts (33 members per day) as opposed to the hindcasts, as examined here. (2) We also only examine a lag of one and two days, since these dates have the highest number of ensemble members available (nine members per day), whereas a lagged ensemble of up to five days was examined in Gregory *et al.* (2020), as the real-time forecasts are produced daily. Finally, the atmospheric initialisation of the real-time forecasts differ from the hindcasts (ACCESS-G vs. ERA-Interim), which may offer improved skill of real-time forecasts at short lead times compared to the hindcasts (Gregory *et al.*, 2020).

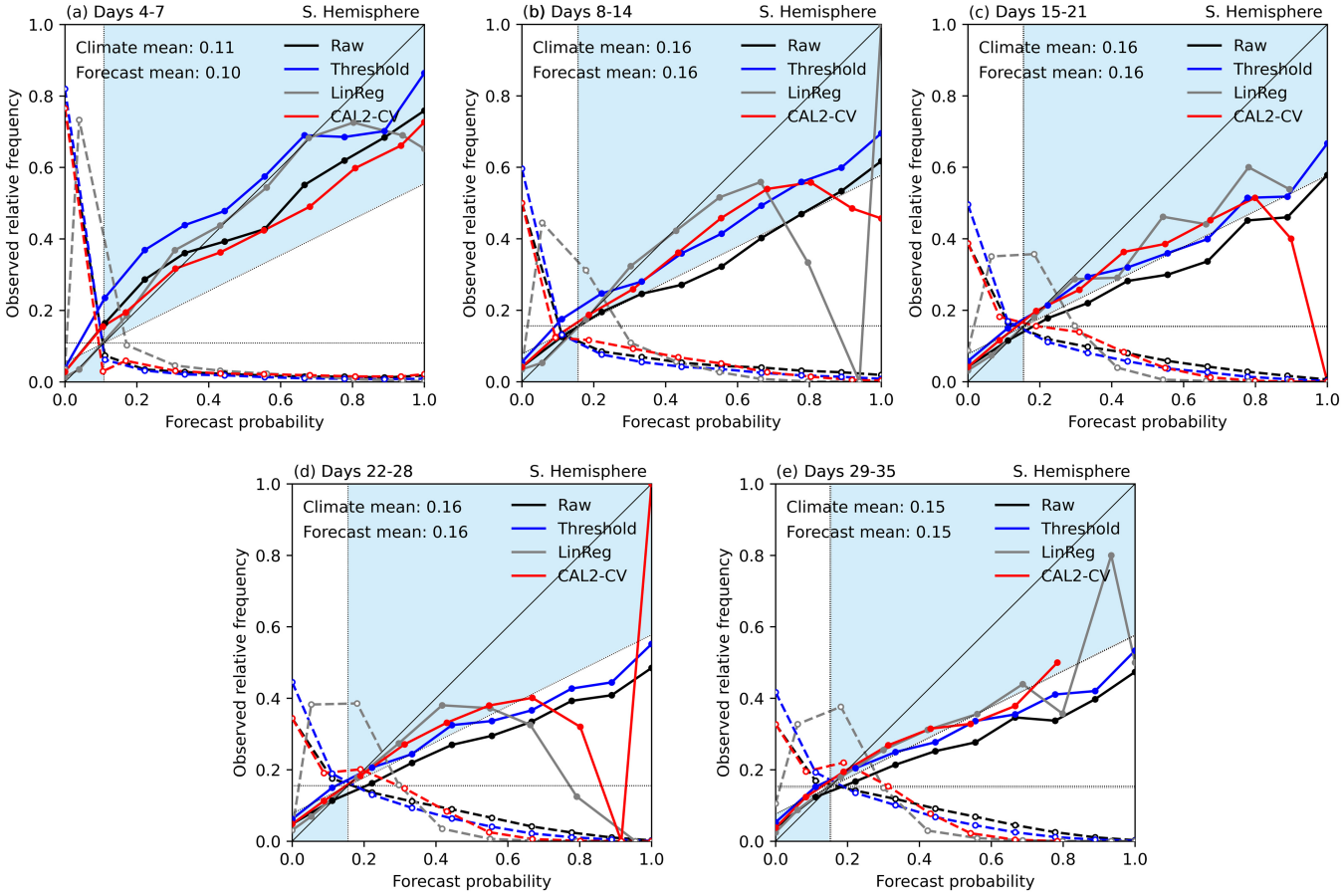

FIGURE 14 Reliability diagrams for the Southern Hemisphere for ACCESS-S2 using four different processing methods (raw modekl tracks, raw modekl tracks with wind speed threshold, calibrated-probabilities CAL2-CV, and the linear regression-based calibration LinReg) for forecast days (a) 4–7 (week 1), (b) 8–14 (week 2) and (c) 15–21 (week 3), (d) 22–28 (week 4) and (e) 29–35 (week 5). Hindcasts starting 1st and 16th of the month, November–April 1981–2018. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Brier skill scores (BSSs) for ACCESS-S2 raw tracks, raw tracks with wind speed threshold, calibrated probabilities CAL2-CV, and the linear regression-based calibration for the Southern Hemisphere for forecast days 4–7 (week 1), 8–14 (week 2), 15–21 (week 3), 22–28 (week 4) and 29–35 (week 5) starting monthly from 1 and 16 November–April 1981–2018.

	Week 1	Week 2	Week 3	Week 4	Week 5
Raw tracks	0.272	0.040	−0.039	−0.079	−0.107
Raw tracks w/wind speed threshold	0.256	0.117	0.047	0.003	−0.009
Calibrated probabilities (CAL2-CV)	0.251	0.133	0.081	0.054	0.048
Linear regression	0.267	0.136	0.065	0.062	0.043

6.4 | Linear regression calibration

Finally, we examine the skill of forecasts of TC occurrence for the Southern Hemisphere using a linear regression-based calibration following Lee *et al.* (2020). In their study Lee *et al.* (2020) showed that the CAL2-CV calibration technique had limitations that could potentially result in lower basinwide BSSs (e.g., Camargo *et al.*, 2019). This is because CAL2-CV only corrects the mean forecast bias, not the mean square error, which the BSS measures (Lee *et al.*, 2020). Thus, a linear regression-based calibration, which minimises the mean square error, could offer a more skilful alternative (Lee *et al.*, 2020). We apply this method to the ACCESS-S2 hindcast to examine whether this technique could further improve on the forecast skill of CAL2-CV.

For each week, a linear regression is created between the mean forecast probability X and the mean observation Y (where the observation is either 1 or 0 for each day within the forecast window):

$$Y = mX + b. \quad (1)$$

Y and X have one data point per year within the hindcast, allowing a least-squares fit of m and b. The fitted m and b are then applied to each model probability X to produce the calibrated forecast probability X_{linear}:

$$X_{\text{linear}} = mX + b. \quad (2)$$

For each year in the hindcast, the linear regression is produced using all other years as training data: the regression for the BSSs (Table 3) and reliability (Figure 14) is based on the m and b applied each year produced through this leave-one-out cross-validation technique. A regression for each of the 60 (10° × 15° degree) boxes is created and applied this way. As noted in Lee *et al.* (2020), the resulting calibrated probability (Equation 2) can be negative or greater than 1, and are therefore restricted to be between 0 and 1. Corresponding BSSs for weeks 1–5 are provided in Table 3, and reliability diagrams in Figure 14, respectively.

For week 1, the linear regression-based calibration does not outperform the skill of CAL2-CV, except for week 4. Here, the linear regression method has the greatest BSS, exceeding that of the raw forecasts (both with and without the wind speed threshold). In week 1 the greatest skill remains using the raw probabilities; for weeks 2, 3 and 5 the BSSs of the linear regression-based calibration are comparable to, or lower than, CAL2-CV. The effect of not calibrating the forecast regions with poor regressions (p-values greater than 0.05 [Equation 1]) was a reduction in BSSs for all weeks, with weeks 3–5 dropping below zero and being outperformed by CAL2-CV. It is also of note that the linear regression-based calibration results in many fewer forecasts of zero TC occurrence probability since the y-intercept (b in Equation 1) is positive for about 75% of cases. It is therefore recommended that the current calibration technique (CAL2-CV) is retained for the operational TC forecasts.

7 | CASE STUDY

The real-time forecast performance of ACCESS-S2 to predict the probability of TC occurrence in the Southern Hemisphere is examined for cyclone *Emnati* and tropical storms *Fezile* and *Dumako* in the southwest Indian Ocean in February 2022. Tropical storm *Dumako* (10–18 February 2022) made landfall on the northeast coast of Madagascar on 15 February with 10-minute maximum sustained winds of 65 mph (~30 ms^{-1}) and resulted in 14 casualties; it later made landfall in Mozambique as a tropical depression, causing flooding and impacting more than 23,000 people (OCHA, 2022a, 2022b). Cyclone *Emnati* (15–26 February) was the second-most intense TC of the 2021/22 southwest Indian Ocean TC season, reaching 10-min maximum sustained wind speeds of 110 mph (~50 ms^{-1}), and the fourth TC to make landfall in Madagascar (23 February) within a month (OCHA, 2022c). *Emnati* caused significant crop and infrastructure damage, resulted in 15 casualties and affected around 170,000 people (OCHA, 2022c, 2022d). Tropical storm *Fezile* (16–18 February 2022) formed in the eastern part of the basin, but did not impact any land areas.

Real-time forecasts for cyclone *Emnati* and tropical storms *Fezile* and *Dumako* are shown for the period 16–22 February 2022 in Figure 15. Forecasts were issued for lead time week 4 (initialised 25 January 2022), week 3 (initialised 1 February 2022) and week 2 (initialised 8 February 2022). Corresponding MJO forecasts from ACCESS-S2 for each of these initialisation dates are provided in Figure 16.

ACCESS-S2 showed a 30%–40% chance of TC formation in the southwest Indian Ocean at four weeks lead time, which was close to the observed location of tropical storm *Fezile*. The MJO forecast during this time (16–22 February) showed a preference, albeit weak, for the MJO in Phase 3 (Indian ocean), which would favour enhanced TC activity in the region predicted. The MJO forecast matched the observed MJO phase well, but was weaker in magnitude.

In week 3, ACCESS-S2 showed two regions of enhanced TC probabilities, located close to the observed tracks of tropical storm *Dumako* (Mozambique Channel) and tropical storm *Fezile*. The corresponding MJO forecast from ACCESS-S2 showed a moderate MJO in Phase 3; both the magnitude and phase of the MJO matched observations well.

At two weeks' lead time the probability of TC formation had increased to 60%–70% around 60° E and 90° E, which coincided well with the observed tracks of cyclone *Emnati* and tropical storm *Fezile*. The increase in forecast TC probabilities was likely due to a strengthening of the forecast amplitude of the MJO in Phase 3; the forecast amplitude and the phase of the MJO were both close to observations. We also note that during this time (weeks 2–4) there were also very few false alarms around Australia and in the S Pacific, likely due to the persistent MJO forecast being in Phase 3 which would favour reduced activity across northern Australia and in the South Pacific. Examining additional forecasts before these valid times (e.g., for

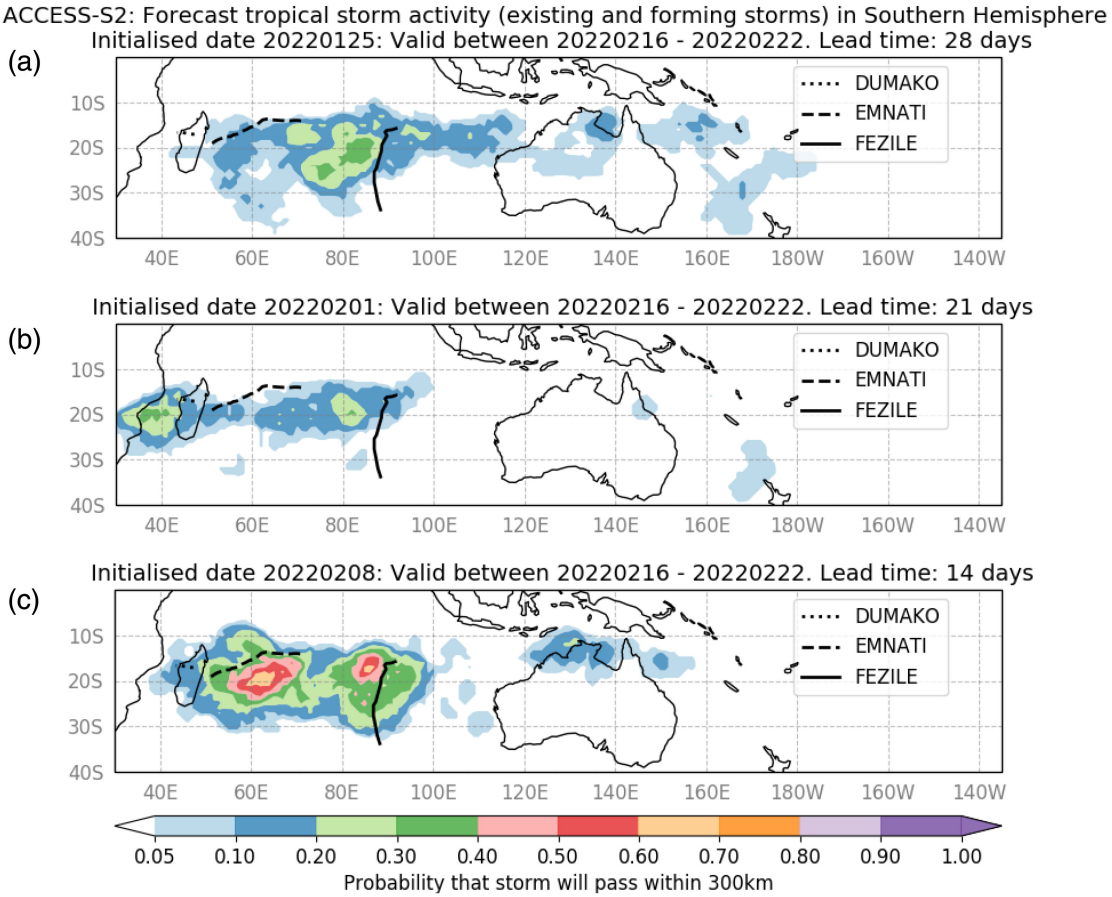

FIGURE 15 ACCESS-S2 real-time forecasts and corresponding observations for cyclone *Emnati* and tropical storms *Fezile* and *Dumako* in the South Indian Ocean (16–22 February 2022) for lead time (a) week 4 (initialised 25 January 2022), (b) week 3 (initialised 1 February 2022) and (c) week 2 (initialised 8 February 2022). All forecasts cover the period 16–22 February 2022. [Colour figure can be viewed at wileyonlinelibrary.com]

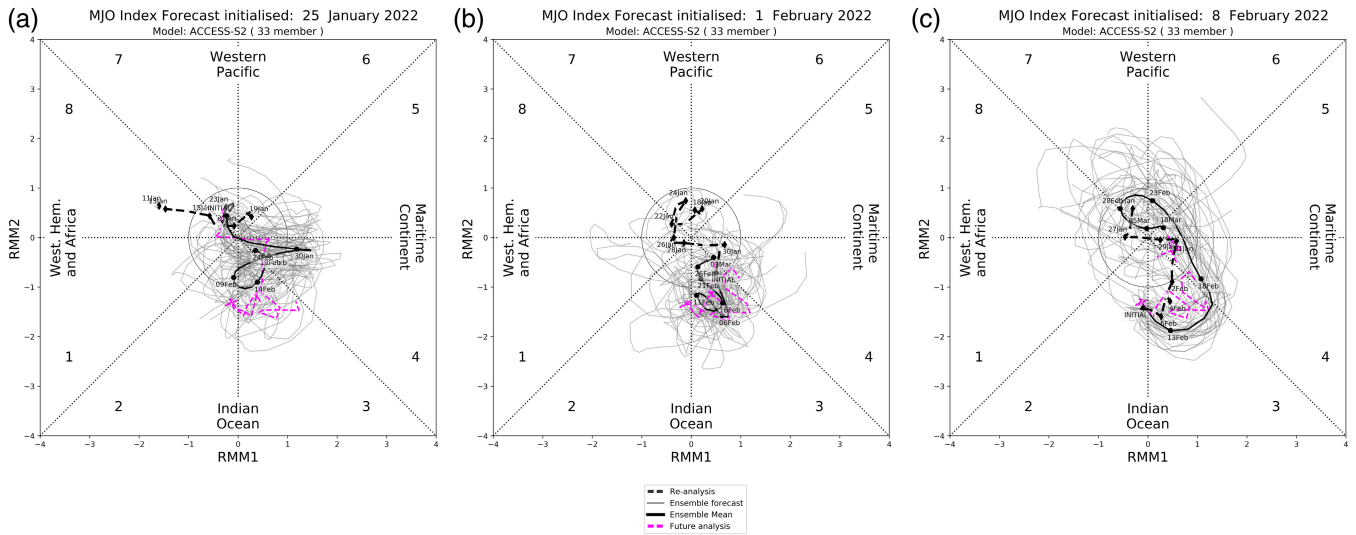

FIGURE 16 ACCESS-S2 daily Madden–Julian Oscillation (MJO) indices (real-time multivariate [RMM1], RMM2) for forecast days 1–30 initialised (a) 25 January, (b) 1 February and (c) 8 February 2022. Model ensemble (33 members; grey) and ensemble mean (black); ACCESS-S2 prior analysis (black dashed) and verified forecast (pink dashed). Eight regions of the phase space are labelled, alongside the approximate locations of the enhanced convective signal of the MJO (e.g. Indian Ocean). Points inside the inner circle signify weak MJO activity (amplitude <1). [Colour figure can be viewed at wileyonlinelibrary.com]

periods 15–21 February, and 14–20 February) showed the false alarms located near Australian and the South Pacific were spatially random with weak probabilities, in contrast to the forecasts in the Indian Ocean which demonstrated a persistently strong signal.

8 | CONCLUSIONS

The skill of multiweek (subseasonal) TC forecasts is examined for the Southern Hemisphere in the Australian BoM multiweek to seasonal prediction system, ACCESS-S2 (Wedd *et al.*, 2022). Results are compared to the predecessor system, ACCESS-S1, for the period 1990–2012. An additional assessment for ACCESS-S2 for the full 38-year hindcast period, 1981–2018, is also provided.

We examine the characteristics and frequency of TCs in both ACCESS-S1 and ACCESS-S2 for days 4–40 over the period 1 November–February, 1990–2012. Both systems have TC genesis and tracks in similar locations and these verify well against observations. Characteristics of the model TCs, such as maximum wind speed and relative humidity, which are identified as part of the OWZ tracking algorithm, also look very similar. Both systems are able to simulate the observed seasonal cycle of TC frequency from November to April, despite more TCs detected than observed. ACCESS-S2 shows fewer TCs on average per member compared to ACCESS-S1, through improvements in TC frequency biases in the South Pacific and Southwest Indian Ocean. Possible reasons for the reduction include an increase in 850–200-hPa wind shear around 20–30° S, which may act to reduce TC activity in the South Pacific and Southwest Indian Ocean, as well as a small decrease in positive SST anomalies around the South Pacific Islands. We note, however, that these anomalies are also influenced by the TCs themselves and this has not been examined here. The relative contribution of TCs to the large-scale fields will be explored in a future study.

The MJO strongly modulates TC activity on multiweek timescales. We examine the relationship between the MJO and TC frequency in ACCESS-S1 and ACCESS-S2 and compare this to ERA-Interim reanalysis for days 15–40 for the period 1990–2012. The eastward progression of TC track frequency with the phase of the MJO is generally well captured and similar between the two models. However, the known negative bias in TC frequency off the coast of NW Australia is not improved in ACCESS-S2. Examining the large-scale fields suggests that the observed increase in 600-hPa relative humidity during MJO Phases 4 + 5 and 6 + 7 is not as strong as in ERA-Interim (and is weaker relative to ACCESS-S1 in Phase 4 + 5), which may be contributing to the lower TC frequency in this region compared to observations. However, this bias may be in part due to the reduced TC genesis in this region and warrants further research. Further analysis of TC tracks and steering flow in ACCESS-S2 would also be beneficial to understand model biases further.

Reliability and BSSs are used to assess the skill of multiweek TC forecasts for lead time weeks 1–5. Raw model tracks are compared against those calibrated using method CAL2-CV (Camp *et al.*, 2018), which provides a regional adjustment to the forecast probabilities to more closely matched observed climatology. For the raw model forecasts, skill is greatest at week 1 and degrades gradually with lead time, as expected. For the period 1990–2012, ACCESS-S2 shows similar or greater skill to ACCESS-S1 for UNCAL forecasts for weeks 1–2. Calibration improves ACCESS-S2 forecast skill for weeks 2–5; however, at all lead times, the skill of the calibrated forecasts from ACCESS-S1 outperforms those from ACCESS-S2.

For the period 1981–2018, skill is found over climatology for raw model forecasts of TC occurrence for weeks 1–2, which can be increased to weeks 1–4 if a wind speed threshold is applied to the model TCs, as in Gregory *et al.* (2019, 2020). For calibrated forecasts (CAL2-CV), ACCESS-S2 shows skill over climatology out to week 5. A linear regression-based calibration, as performed in Lee *et al.* (2020), does not improve forecast skill beyond CAL2-CV, except at week 4. At week 1, skill is greatest using the raw model TC tracks with no wind speed threshold applied.

Gregory *et al.* (2020) found the skill of ACCESS-S1 forecasts could be improved using a lagged ensemble. We find no benefit to using a lagged ensemble of one and two days to the ACCESS-S2 forecast probabilities, except at week 5; however, skill is still lower than that of climatology. As the Southern Hemisphere multiweek forecasts issued by the BoM are for weeks 2–3, these results suggest that calibrated forecasts using method CAL2-CV, as opposed to raw forecasts with a wind speed threshold applied, or a linear-regression-based calibration technique, would still be the most skilful for use operationally. However, an assessment of the skill of real-time forecasts, which are run daily with a much larger number of ensemble members and different atmospheric initialisation, would be beneficial.

Real-time forecasts of TC occurrence from ACCESS-S2 were examined for cyclone *Emnati* and tropical storms *Dumako* and *Fezile* in the southwest Indian Ocean from 16 to 22 February 2022. ACCESS-S2 provided good guidance of possible TC activity out to week 4, with probabilities of TC formation greater than 60% at week 2, coinciding with the observed tracks of cyclone *Emnati* and tropical storm

Fezile. During this time the MJO was well predicted, with the model amplitude and phase of the MJO (Phase 3) matching observations well. The enhanced signal for TC activity in the southwest Indian Ocean out to week 4 was likely due to the consistent forecast for enhanced MJO activity in Phase 3, which would favour TC activity in this region. The real-time performance of ACCESS-S2 to predict other major TC events in the Southern Hemisphere during the 2021/22 season, as well as the multiweek prediction skill of ACCESS-S2 in the Northern Hemisphere, is now being examined.

AUTHOR CONTRIBUTIONS

J. Camp: Formal analysis; methodology; software; visualization; writing – original draft. **P. Gregory:** Conceptualization; formal analysis; software; visualization; writing – review and editing. **A. G. Marshall:** Data curation; writing – review and editing. **J. Greenslade:** Formal analysis; investigation; writing – review and editing. **M. C. Wheeler:** Data curation; methodology; writing – review and editing.

ACKNOWLEDGEMENTS

Madden–Julian Oscillation (MJO) real-time multivariate (RMM) index data was produced with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

ORCID

J. Camp https://orcid.org/0000-0002-4567-9622
M. C. Wheeler https://orcid.org/0000-0002-9769-1973

REFERENCES

BoM. (2020) Bureau of meteorology ACCESS-S multi-week tropical cyclone guidance: extended skill. Accessed 25 April 2022. http://access-s.clide.cloud/files/guidance_documents/About_weekly_ACCESS-S_TC_forecasts_extended_skill.pdf.

Camargo, S.J., Wheeler, M.C. and Sobel, A.H. (2009) Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *Journal of the Atmospheric Sciences*, 66(10), 3061–3074. https://doi.org/10.1175/2009JAS3101.1.

Camargo, S.J., Camp, J., Elsberry, R.L., Gregory, P.A., Klotzbach, P.J., Schreck, C.J., Sobel, A.H., Ventrice, M.J., Vitart, F., Wang, Z., Wheeler, M.C., Yamaguchi, M. and Zhan, R. (2019) Tropical cyclone predictability on sub-seasonal timescales. *Tropical Cyclone Research and Review*, 8(3), 150–165. https://doi.org/10.1016/j.tcrr.2019.10.004.

Camp, J., Wheeler, M.C., Hendon, H.H., Gregory, P.A., Marshall, A.G., Tory, K.J., Watkins, A.B., MacLachlan, C. and Kuleshov, Y. (2018) (2018). Skilful multiweek tropical cyclone prediction in ACCESS-S1 and the role of the MJO. *Quarterly Journal of the Royal Meteorological Society*, 144, 1337–1351. https://doi.org/10.1002/qj.3260.

Chu, J.H., Sampson, C.R., Levine, A.S. and Fukada, E. (2002) The joint typhoon warning center tropical cyclone best-tracks, 1945-2000. Technical report. Naval Research Laboratory Tech. Rep. NRL/MR/7540-02-16. [Available online at https://www.metoc.navy.mil/jtwc/products/best-tracks/tc-bt-report.html].

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N. and Vitart, F. (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. https://doi.org/10.1002/qj.828.

Domeisen, D.I.V., White, C.J., Afargan-Gerstman, H., Muñoz, Á.G., Janiga, M.A., Vitart, F., Wulff, C.O., Antoine, S., Ardilouze, C., Batté, L., Bloomfield, H.C., Brayshaw, D.J., Camargo, S.J., Charlton-Pérez, A., Collins, D., Cowan, T., del Mar Chaves, M., Ferranti, L., Gómez, R., González, P.L.M., González Romero, C., Infanti, J.M., Karozis, S., Kim, H., Kolstad, E.W., LaJoie, E., Lledó, L., Magnusson, L., Malguzzi, P., Manrique-Suñén, A., Mastrangelo, D., Materia, S., Medina, H., Palma, L., Pineda, L.E., Sfetsos, A., Son, S.W., Soret, A., Strazzo, S. and Tian, D. (2022) Advances in the subseasonal prediction of extreme events: relevant case studies across the globe. *BAMS*, 103, E1473–E1501. https://doi.org/10.1175/BAMS-D-20-0221.1.

Gottschalck, J., Wheeler, M., Weickmann, K., Vitart, F., Savage, N., Lin, H., Hendon, H., Waliser, D., Sperber, K., Nakagawa, M., Prestrelo, C., Flatau, M. and Higgins, W. (2010) A framework for assessing operational Madden-Julian oscillation forecasts: a CLIVAR MJO working group project. *American Meteor Society*, 91(9), 1247–1258. https://doi.org/10.1175/2010BAMS2816.1.

Gregory, P.A., Camp, J., Bigelow, K. and Brown, A. (2019) Sub-seasonal predictability of the 2017–2018 southern hemisphere tropical cyclone season. *Atmospheric Science Letters*, 20, e886. https://doi.org/10.1002/asl.886.

Gregory, P., Vitart, F., Rivett, R., Brown, A. and Kuleshov, Y. (2020) Subseasonal forecasts of tropical cyclones in the southern hemisphere using a dynamical multimodel ensemble. *Weather and Forecasting*, 35(5), 1817–1829. https://doi.org/10.1175/WAF-D-20-0050.1.

Hudson, D., Alves, O., Hendon, H.H., Lim, E., Liu, G., Luo, J.-J., MacLachlan, C., Marshall, A.G., Shi, L., Wang, G., Wedd, R., Young, G., Zhao, M. and Zhou, X. (2017) ACCESS-S1: the new Bureau of Meteorology multi-week to seasonal prediction system. *Journal of Southern Hemisphere Earth Systems Science*, 67(3), 132–159. https://doi.org/10.22499/3.6703.001.

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

Lee, C.-Y., Camargo, S.J., Vitart, F., Sobel, A.H. and Tippett, M.K. (2018) Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Weather Forecasting*, 33, 967–988. https://doi.org/10.1175/WAF-D-17-0165.1.

Lee, C.-Y., Camargo, S.J., Vitart, F., Sobel, A.H., Camp, J., Wang, S., Tippett, M.K. and Yang, Q. (2020) Subseasonal predictions of tropical cyclone occurrence and ACE in the S2S dataset. *Weather*

and Forecasting, 35(3), 921–938. https://doi.org/10.1175/WAF-D-19-0217.1.

Lim, E.P., Hendon, H.H., Hudson, D., Zhao, M., Shi, L., Alves, O. and Young, G. (2016) *Evaluation of the ACCESS-S1 Hindcasts for Prediction of Victorian Seasonal Rainfall*. Technical report. Bureau Research Report No. 019. Docklands: Bureau of Meteorology.

Lim, Y., Son, S.-W. and Kim, D. (2018) MJO prediction skill of the subseasonal-to-seasonal prediction models. *Journal of Climate*, 31, 4075–4094. https://doi.org/10.1175/JCLI-D-17-0545.1.

MacLachlan, C., Arribas, A., Peterson, K.A., Maidens, A., Fereday, D., Scaife, A.A., Gordon, M., Vellinga, M., Williams, A., Comer, R.E., Camp, J., Xavier, P. and Madec, G. (2015) Global seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, 141, 1072–1084. https://doi.org/10.1002/qj.2396.

Madden, R.A. and Julian, P.R. (1971) Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *Journal of Atmospheric Sciences*, 28(5), 702–708. https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

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

OCHA. (2022a) Southern Africa: Cyclone Season Flash Update No. 8 (22 February 2022). Accessed 15 June 2022. https://reliefweb.int/report/madagascar/southern-africa-cyclone-season-flash-update-no-8-22-february-2022.

OCHA. (2022b) Mozambique: Tropical Storms Ana and Dumako Flash Update No.10 (As of 2 March 2022). Accessed 15 June 2022. https://reliefweb.int/report/mozambique/mozambique-tropical-storms-ana-and-dumako-flash-update-no10-2-march-2022/.

OCHA. (2022c) Tropical Cyclone Emnati-Feb 2022. Accessed 15 June 2022. https://reliefweb.int/disaster/tc-2022-000175-mdg.

OCHA. (2022d) WFP Madagascar Cyclone Response Update (As of 15 April 2022). Accessed 15 June 2022. https://reliefweb.int/report/madagascar/wfp-madagascar-cyclone-response-update-15-april-2022.

Rashid, H., Hendon, H., Wheeler, M. and Alves, O. (2011) Prediction of the Madden-Julian oscillation with the POAMA dynamical prediction system. *Climate Dynamics*, 36(3–4), 649–661. https://doi.org/10.1007/s00382-010-0754-x.

Scaife, A.A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., MacLachlan, C., Martin, N., Nie, Y., Ren, H.L., Roberts, M., Robinson, W., Smith, D. and Vidale, P.L. (2019) Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, 20, e922. https://doi.org/10.1002/asl.922.

Tory, K.J., Dare, R.A., Davidson, N.E., McBride, J.L. and Chand, S.S. (2013) The importance of low-deformation vorticity in tropical cyclone formation. *Atmospheric Chemistry and Physics*, 13, 2115–2132. https://doi.org/10.5194/acp-13-2115-2013.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H.-S., Kumar, A., Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A.W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D.-J., Xiao, H., Zaripov, R. and Zhang, L. (2017) The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, 98(1), 163–173. https://doi.org/10.1175/BAMS-D-16-0017.1.

Vitart, F., Leroy, A. and Wheeler, M. (2010) A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the southern hemisphere. *Monthly Weather Review*, 138, 3671–3682. https://doi.org/10.1175/2010MWR3343.1.

Vitart, F. and Robertson, A.W. (2018) The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate and Atmospheric Science*, 1, 3. https://doi.org/10.1038/s41612-018-0013-0.

Vitart, F. and Takaya, Y. (2021) Lagged ensembles in sub-seasonal predictions. *Quarterly Journal of the Royal Meteorological Society*, 147(739), 3227–3242. https://doi.org/10.1002/qj.4125.

Wedd, R., Oscar, A., de Catherine, B.-D., Christopher, D., Morwenna, G., Hendon Harry, H., Debra, H., Shuhua, L., Eun-Pa, L., Marshall Andrew, G., Li, S., Paul, S., Grant, S., Spillman Claire, M., Guomin, W., Wheeler Matthew, C., Hailin, Y., Yonghong, Y., Griffith, Y., Mei, Z., Yi, X. and Xiaobing, Z. (2022) ACCESS-S2: the upgraded Bureau of Meteorology multi-week to seasonal prediction system. *Journal of Southern Hemisphere Earth Systems Science*, 72, 218–242. https://doi.org/10.1071/ES22026.

Wheeler, M.C. and 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. https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

Williams, K.D., Harris, C.M., Bodas-Salcedo, A., Camp, J., Comer, R.E., Copsey, D., Fereday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S.F., Roberts, M.J., Rowell, D.P., Sanchez, C., Shelly, A., Sinha, B., Walters, D.N., West, A., Woollings, T. and Xavier, P.K. (2015) The met Office global coupled model 2.0 (GC2) configuration. *Geoscientific Model Development*, 8, 1509–1524. https://doi.org/10.5194/gmd-8-1509-2015.

WMO. (2021) State of the Global Climate 2021: WMO Provisional report. Accessed 25 April 2022. https://library.wmo.int/index.php?lvl=notice_display&id=21982#.YmZS2zdByWA.

Yin, Y., Alves, O. and Oke, P.R. (2011) An Ensemble Ocean data assimilation system for seasonal prediction. *Monthly Weather Review*, 139, 786–808. https://doi.org/10.1175/2010MWR3419.1.

How to cite this article: Camp, J., Gregory, P., Marshall, A.G., Greenslade, J. & Wheeler, M.C. (2023) Multiweek tropical cyclone prediction for the Southern Hemisphere in ACCESS-S2: Maintaining operational skill and continuity of service. *Quarterly Journal of the Royal Meteorological Society*, 149(757), 3401–3422. Available from: https://doi.org/10.1002/qj.4563