

Subseasonal Forecasts of Tropical Cyclones in the Southern Hemisphere Using a Dynamical Multimodel Ensemble

PAUL GREGORY

Bureau of Meteorology, Melbourne, Victoria, Australia

FREDERIC VITART

ECMWF, Reading, United Kingdom

RABI RIVETT

Bureau of Meteorology, Perth, Western Australia, Australia

ANDREW BROWN AND YURIY KULESHOV

Bureau of Meteorology, Melbourne, Victoria, Australia

(Manuscript received 8 April 2020, in final form 19 June 2020)

ABSTRACT

Subseasonal tropical cyclone forecasts from two operational forecast models are verified for the 2017/18 and 2018/19 Southern Hemisphere cyclone seasons. The forecasts are generated using the ECMWF's Medium- and Extended-Range Ensemble Integrated Forecasting System (IFS), and the Bureau of Meteorology's seasonal forecasting system ACCESS-S1. Results show the IFS is more skillful than ACCESS-S1, which is attributed to the IFS's greater ensemble size, increased spatial resolution, and data assimilation schemes. Applying a lagged ensemble with ACCESS-S1 increases forecast reliability, with the optimum number of lagged members being dependent on forecast lead time. To investigate the impacts of atmospheric assimilation at shorter lead times, comparisons were made between the Bureau of Meteorology's ACCESS-S1 and ACCESS-GE2 systems, the latter a global Numerical Weather Prediction system running with the same resolution and model physics as ACCESS-S1 but using an ensemble Kalman filter for data assimilation. This comparison showed the data assimilation scheme used in the GE2 system gave improvements in forecast skill for days 8–10, despite the smaller ensemble size used in GE2 (24 members per forecast compared to 33). Finally, a multimodel ensemble was created by combining forecasts from both the IFS and ACCESS-S1. Using the multimodel ensemble gave improvements in forecast skill and reliability. This improvement is attributed to complementary spatial errors in both systems occurring across much of the Southern Hemisphere as well as an increase in the ensemble size.

SIGNIFICANCE STATEMENT

Advances in model development allow skillful forecasts of tropical cyclone activity beyond the normal limit of weather prediction (typically 14 days, or a two-week forecast) and into the “subseasonal” time frame. This is achieved by coupling high-resolution ensemble global forecast models to global ocean models. These subseasonal forecasts fill the gap between traditional weather forecasts and monthly climate outlooks. This study verifies two separate subseasonal cyclone forecasting system for the 2017/18 and 2018/19 cyclone seasons over the Southern Hemisphere. Both systems showed good skill in forecasting cyclone activity out to three weeks in advance. By combining the results of both models, forecast skill is further improved.

1. Introduction

Subseasonal tropical cyclone forecasts at the ECMWF were first verified by Vitart et al. (2010). This study analyzed 46-day hindcasts issued on the 15th of each

Corresponding author: Paul Gregory, paul.gregory@bom.gov.au

DOI: 10.1175/WAF-D-20-0050.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

month from 1989 to 2008 with ECMWF Integrated Forecasting System (IFS), version Cy32r3. Forecasts issued over the Southern Hemisphere were compared against a statistical model developed by Leroy and Wheeler (2008) that used Madden–Julian oscillation (MJO) real-time multivariate indices and empirical orthogonal functions of sea surface temperature anomalies as predictors. This IFS model forecast a large number of false alarms that gave the statistical model better skill after week two. Applying a simple calibration method to the dynamical model improved forecast reliability significantly, at the expense of sharpness. A multimodel combination of both models gave the best Brier skill scores (BSS), but with less sharpness than the calibrated dynamical model.

The ECMWF began issuing twice-weekly global multiweek cyclone forecasts in 2012 (Vitart et al. 2012). Forecast skill was assessed using a research version of the 32-day IFS with model version Cy36r4 featuring 15 ensemble members. This assessment used forecasts initialized on the first day of every month from 1989 to 2008. BSS showed skill out to 25 days, with improved skill in the Southern Hemisphere.

The ECMWF remains the only center to issue operational subseasonal cyclone forecasts based on a dynamical model. A recent review by Camargo et al. (2019) highlighted that only four agencies currently issue subseasonal forecasts. They are as follows:

Colorado State University issues a two-week forecast of accumulated cyclone energy (ACE) for the North Atlantic based on a statistical model that uses a variety of inputs, including current MJO indices, current storm activity, and numerical weather prediction (NWP) inputs.

ECMWF issues predicted numbers of storms and ACE, as well as strike probability maps for weeks 1–4 over the globe based on IFS forecasts.

The China Meteorological Administration (CMA) issues a variety of forecast products (ACE, landfall in China, storm numbers) for the northwest Pacific for weeks 1–4. Their system uses inputs from the Canadian CFSv2 model to provide statistical forecasts (see Gao and Li 2011).

BoM computes predicted storm tracks and strike probabilities for weeks 1–4 over the globe using ACCESS-S1 forecasts. Calibrated products are also generated over the Southern Hemisphere and northwest Pacific. This is a nonoperational product currently used only for research purposes.

Currently, only the Colorado State University forecasts are freely available to the general public.

The main aim of this study is to validate the ECMWF and BoM dynamical model-based forecasts of subseasonal

cyclone activity over the Southern Hemisphere for the 2017/18 and 2018/19 seasons. This will provide accurate verification statistics of the latest ECMWF forecasts, which were last verified in 2012. It will also extend the prior work of Gregory et al. (2019) to validate the ACCESS-S1 real-time forecasts over a larger sample space. In addition to aggregate statistics, this study includes two case studies for illustration.

An additional comparison between ACCESS-S1 and ACCESS-GE2 will determine the impact of atmospheric data assimilation on forecast skill for days 8–10. ACCESS-GE2 was a prototype NWP system that used the same atmospheric model as ACCESS-S1 at the same resolution but used an ensemble Kalman filter assimilation scheme.

The secondary aim is to determine whether a multimodel ensemble of both forecast systems (IFS and ACCESS-S1) provides improvements in forecast skill.

2. Southern Hemisphere tropical cyclone activity from 2017 to 2019

In the Southern Hemisphere, the tropical cyclone (TC) season officially begins on 1 November and ends on 30 April; however, a TC can form at any time during the year. To ensure consistency with results published earlier by Vitart et al. (2012), the verification period was defined as 1 November 2017–30 April 2018, and 1 November 2018–30 April 2019. TCs outside of this period are excluded from the analysis.

Real-time track data were collated from NOAA.¹ For these 2017/18 and 2018/19 seasons, the number of storms recorded during the verification period in the Southern Hemisphere was 21 and 24, respectively. These represent slightly below-average amounts of storm activity when compared to the long-term averages of 24.3 storms per season. This long-term average was computed by Schreck et al. (2014) using data between 1981 and 2010 from International Best Track Archive for Climate Stewardship–World Meteorological Organization (IBTrACS–WMO).

Observed tracks for both seasons are given in Fig. 1. In terms of TC impact, both seasons featured several destructive storms, including Gita in the South Pacific Ocean in February 2018 and Kenneth in the southwest Indian Ocean in April 2019. Severe TC Gita was the most intense storm recorded in the South Pacific region (160°E–120°W) in the 2017/18 TC season with a maximum sustained wind (10-min average) of 205 km h^{−1} and lifetime minimum central pressure of 927 hPa. Gita impacted many Pacific island nations with the most

¹ <https://www.ssd.noaa.gov/PS/TROP/DATA/ATCF/JTWC>.

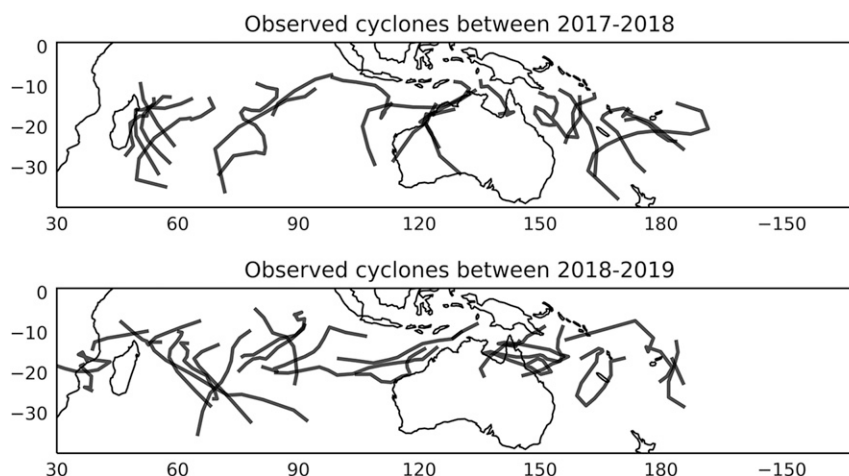


FIG. 1. Tracks used in the verification study (top) from November 2017 to April 2018 and (bottom) from November 2018 to April 2019. Data obtained from NOAA.

significant damage being reported in Tonga and the Samoan Islands. Intense TC Kenneth was the strongest storm in the southwest Indian Ocean region (30° – 90° E) with maximum sustained wind (10-min average) of 215 km h^{-1} and lifetime minimum central pressure of 934 hPa. Kenneth caused widespread destruction and several deaths in Comoros and Mozambique; it was the most intense landfalling storm in Mozambique's recorded history.

To reduce TC impacts on populations and infrastructure, early warning systems can be strengthened by providing advanced forecasts (2–4 weeks ahead) of storm development and impacts. Note that over 30 000 people were evacuated in Mozambique ahead of Kenneth's path. Improved subseasonal forecasts will be invaluable in such situations.

3. Model configurations

a. ECMWF forecast system

The ECMWF IFS underwent one model change during the verification period. Cycle 43r3 was active from 11 July 2017 until 5 June 2018 when it was replaced with Cycle 45r1.

All IFS forecasts in this study (both medium and extended range) were coupled to a 0.25° ocean model (NEMO).

Both the medium and extended-range IFS use the same 51-member ensemble prediction system. The 51-member medium-range forecasts are produced twice daily (0000 and 1200 UTC) at 16-km resolution with lead time out to 15 days. The extended-range forecasts are an extension of the same 51-member medium-range forecasts with a lead time extended to 46 days. These forecasts are run twice a week (every Monday and Thursday) at a 32-km resolution after day 15. The 51-member ensemble is initialized

from the operational 4D-VAR analysis. There was no change in model resolution between IFS versions C43r4 and C45r1.

Initial perturbations are generated using a combination of singular vectors, and perturbations generated using the ECMWF ensemble of data assimilations (see [Leutbecher and Palmer 2008](#); [Buizza et al. 2008](#)). A stochastic physics scheme [stochastically perturbed parametrization tendencies (SPPT)] is used to represent model uncertainty ([Lock et al. 2019](#)).

The total number of medium-range forecasts verified in this study was 383, while 112 extended-range forecasts were verified.

b. Bureau ACCESS forecast systems

The Australian Bureau of Meteorology's ACCESS-S1 model is a coupled ocean–atmosphere model used primarily for seasonal prediction. The system became operational in August 2018, with daily forecasts out to 210 days. This system is based on the Met Office Global Seasonal forecast system, version 5 (GloSea5; [MacLachlan et al. 2015](#)), and differs only in the initialization of the atmospheric ensemble spread (see [Hudson et al. 2017](#)).

ACCESS-S1 uses scaled random perturbations to generate an atmospheric ensemble spread. Eleven ensemble members are forecast out to 210 days. To provide a better skill in the subseasonal lead times, 22 additional ensembles are generated with forecast leads of 42 days.

The atmospheric resolution of ACCESS-S1 is N216 (approximately 60 km) and it uses the Met Office Unified Model, version 8.3. This version of the Unified Model (UM) was released as part of the Global Atmosphere 6.0 (GA6.0) suite; full details can be found in [Walters et al. \(2017\)](#). This atmospheric model is coupled

TABLE 1. BSSs for ACCESS-S1 raw tracks, raw tracks with wind speed threshold, and calibrated probabilities.

Forecast period	Raw tracks	Raw tracks w/wind speed threshold	Calibrated probabilities
8–14	0.132	0.193	0.187
15–21	0.029	0.093	0.100
22–28	−0.004	0.075	0.081

to the NEMO ocean model at 0.25° resolution with 75 vertical levels.

ACCESS-GE2 was a prototype NWP system that used an ensemble Kalman filter to generate a 24-member ensemble forecast out to 10 days. It used the same UM version as ACCESS-S1 at the same resolution. Forecasts were generated slightly behind real time throughout the verification period. Only the 0000 UTC forecasts were analyzed to provide a like-for-like comparison with the corresponding ACCESS-S1 forecasts.

c. Tracking schemes

Since 2012, the ECMWF has used the same cyclone tracker for all forecast products. As described in Vitart et al. (1997), vorticity maximums (or minimums in the Southern Hemisphere) that lie within 8° latitude of an MSLP minimum are identified as possible candidates. Additional thresholds are tested for geopotential height and upper-air temperature to ensure the cyclone has a warm core (and hence is not an extratropical storm). Once these points have been identified, the algorithm presented in Van der Grijn et al. (2005) is used to compute the trajectory. A 10-m wind speed threshold of 17 m s^{−1} is applied to the trajectory; storms that fail to reach this threshold are discarded.

For ACCESS-S1 and GE2 forecasts, tracks were extracted for every daily 0000 UTC forecast. The Okubo–Weiss–Zeta

parameter (OWZ) tropical cyclone detection scheme developed by Tory et al. (2013) was used to compute the location of possible cyclones based on ACCESS-S1 atmospheric fields.

The OWZ tracker was designed for use with coarse-resolution coupled ocean–atmospheric models. Instead of identifying storms using MSLP minima and vorticity maximum criteria (which are grid dependent and often not suitable for coarse-resolution coupled models) it searches for conditions conducive to cyclogenesis. These conditions are identified with the following atmospheric parameters: solid body rotation at 850 and 500 hPa, relative humidity at 950 and 700 hPa, specific humidity at 950 hPa, and low wind shear between 850 and 200 hPa. Points that satisfy certain thresholds for these parameters are identified and trajectories are computed that join each point in time. In this study, threshold points were identified at every daily time step. As detailed in Gregory et al. (2019), an 850-hPa wind speed threshold of 14 m s^{−1} was applied to every storm track. Storms that failed to reach this threshold were discarded.

d. Calibration

A calibration scheme was also applied to OWZ storm tracks computed from ACCESS-S1 forecasts. In this scheme, the 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 the forecast (leave-one-out cross validation). The ACCESS-S1 hindcast spans 1990–2012 with four start dates per month. The calibration scaling factors were averaged across the entire season (November–April) from 1990 to 2012. Details of this scheme are found in Camp et al. (2018), where it is denoted as CAL2-CV. This scaling factor is computed for each forecast lead time (i.e., days 8–14, 15–21 and 22–28)

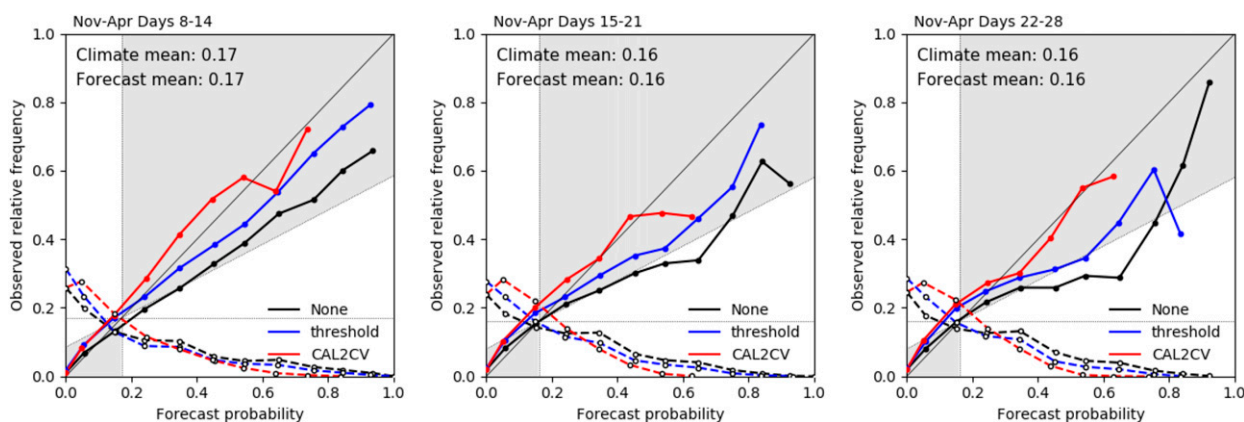


FIG. 2. Reliability plots for ACCESS-S1 using three different processing methods (raw tracks, raw tracks with wind speed threshold, and calibrated probabilities) for varying lead time: (left) 8–14, (center) 15–21, and (right) 22–28 days.

TABLE 2. BSS for ACCESS-S1 raw tracks with wind speed threshold and increasing ensemble size using additional lagged forecasts.

Forecast period	Lag 0	Lag 2	Lag 3	Lag 4	Lag 5
8–14	0.193	0.198	0.193	0.194	0.175
15–21	0.093	0.103	0.105	0.106	0.106
22–28	0.075	0.089	0.097	0.103	0.108

within each verification region. The size of each verification region is 15° latitude \times 20° longitude, so the calibration scheme degrades the spatial resolution of the forecast tracks. Further details of the verification regions are provided in section 3e.

e. Verification methodology

Using the process detailed in Camp et al. (2018), the Southern Hemisphere from 0° to 30° S latitude, and from 30° to 240° E longitude, is divided into 60 overlapping regions, each with size 15° latitude \times 20° longitude. For every available forecast and lead time, the forecast and observed storm probabilities are computed in each region. Observed probabilities in each region are either zero or one. In the case of ACCESS-S1 calibrated forecasts, the raw forecast probability in each verification region is multiplied by the corresponding scaling factor for a given forecast lead time. From this data, reliability diagrams and BSS are computed. For each forecast lead time, the BSS is computed over all forecasts for all regions and is defined as

$$\text{BSS} = 1 - \frac{\text{BS}}{\text{BS}_{\text{reference}}}, \quad (1)$$

that is, the relative skill of a probabilistic forecast over climatology. It ranges from $-\infty$ to one, with scores greater than zero indicating skill over the reference

forecasts (climatology in our case). See Wilks (1995) for further information.

4. Forecast reliability and skill

a. ACCESS-S1

BSS for the ACCESS-S1 forecasting system are shown in Table 1. Scores are shown for 1) raw tracks, 2) tracks with the wind speed threshold applied, and 3) applying the calibration factors computed using the 1990–2012 hindcast. Corresponding reliability plots and sharpness diagrams are shown in Fig. 2.

These results are consistent with Gregory et al. (2019). The wind speed threshold improves skill by filtering out low-intensity tropical storms from the model outputs. Likewise, using the hindcast calibration factors improves reliability for longer lead times at the expense of sharpness.

To take advantage of daily ACCESS-S1 forecasts, lagged ensemble forecasts were also generated. BSS for the lagged ensemble forecasts using the wind speed threshold are shown in Table 2. Here, “lag 0” refers to the default forecasts, “lag 2” includes ensembles from the previous day (hence forecasts include two days of data), “lag 3” includes forecasts from the two previous days (hence forecasts include three days of data) and so on. So, the lag 2 system is a 66-member ensemble, the lag 3 is a 99-member ensemble, and so on.

For the fourth week forecasts (days 22–28), the lag only included forecasts out to 30 days, as this was the maximum lead time of ACCESS-S1 model data retrieved from the operational output. This limits the amount of lower skill forecasts added to the ensemble.

Results shown in Table 2 and Fig. 3 show the optimal amount of lag depends on the forecast lead time.

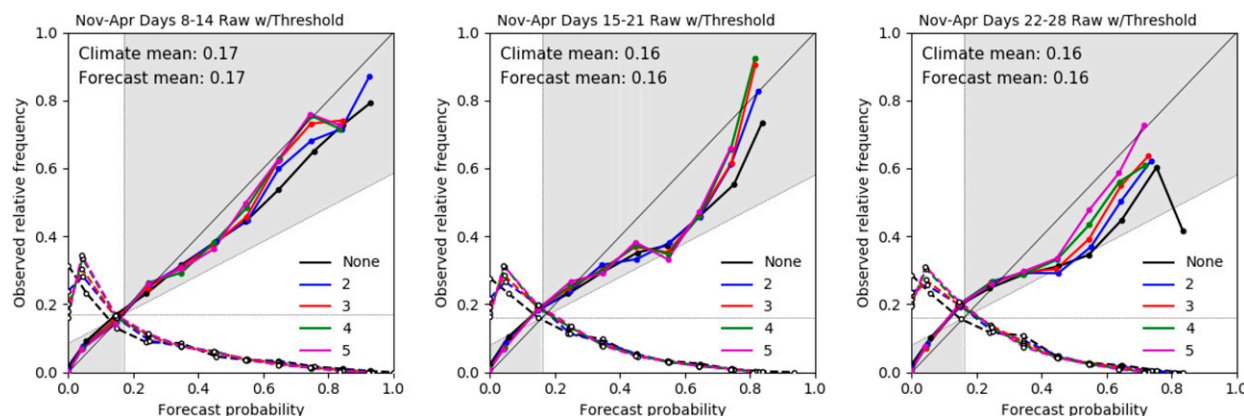


FIG. 3. Reliability plots for ACCESS-S1 tracks using the wind speed threshold for varying forecast lag and lead time: (left) 8–14, (center) 15–21, and (right) 22–28 days.

TABLE 3. BSSs for ACCESS-S1 and ACCESS-GE2 tracks for 8–10-day forecasts with wind speed thresholds and increasing ensemble size using additional lagged forecasts.

Model	Lag 0	Lag 2	Lag 3
ACCESS-S1	0.185	0.187	0.183
ACCESS-GE2	0.201	0.207	0.205

For 8–14-day lead times, using a lagged forecast improves reliability of higher probability events. However, for lag greater than 2, sharpness is reduced. For 15–21-day lead times, reliability is again improved for higher probability events. However, this comes at the expense of lower resolution and loss of skill for forecast events with probabilities equal to 0.5. Finally, for 22–28-day lead times, reliability consistently increases with lag, although this is due to the maximum lead time of available forecasts being restricted to 30 days.

Based on these results, the optimal lag for 8–14- and 15–21-day forecasts was lag 2. For 22–28-day forecasts, reliability continues with increasing lag (subject to the maximum lead time of 30 days). Hence the lag 5 forecasts gave the best reliability.

b. ACCESS-GE2

Cyclone tracks were also extracted from 8–10-day ACCESS-GE2 forecasts over the same period. BSS for the GE2 forecasts are shown in Table 3 with the corresponding reliability plots shown in Fig. 4. The GE2 system consistently provides better skill as measured by BSS, with improved reliability. For lag 0 forecasts, GE2 provides better reliability except for higher probability events (although GE has improved sharpness for events with forecast probabilities greater than 0.8). With increasing ensemble lag, GE2 maintains its superior reliability although it does trend to overforecast low probability events.

GE2 provides consistently more reliable forecasts, despite using fixed SSTs. These results show that atmospheric data assimilation is more important than coupling to an ocean for 8–10-day lead times.

c. ECMWF/IFS

Brier skill scores for the IFS are shown in Table 4. These scores are much higher (and therefore more skillful) than the ACCESS-S1 equivalents. However, examination of the reliability plots in Fig. 5 show that this skill is mainly associated with lower probability events. The IFS is overconfident for forecast probabilities greater than 0.5 and this bias becomes significant for lead times beyond 15 days. However, this lack of reliability may also be a factor of sample size; recall that the IFS longer-range forecasts are made only twice per week. Hence there are only 112 forecasts with lead times of 15–21 and 22–28 days in the sample period, compared with 383 forecasts with 8–14-day lead times.

d. A multimodel ensemble

A multimodel ensemble (MME) was created by adding the subset of ACCESS-S1 forecasts that shared the same forecast dates as the IFS forecast dates. Hence the MME forecasts for weeks three and four are only computed twice per week. This was done in preference to assessing skill for every forecast date, which would have otherwise contaminated the MME sample space with forecasts using only ACCESS-S1 results. For forecasts with a maximum of 10-day lead times, the MME can also include ACCESS-GE2 forecasts.

Figure 6 plots the reliability of an MME, with Table 5 showing the corresponding BSS. The MME has slightly higher scores than the IFS model for all lead times. Using lagged ACCESS-S1 forecasts in the MME degraded the skill. Presumably, this is due to changing

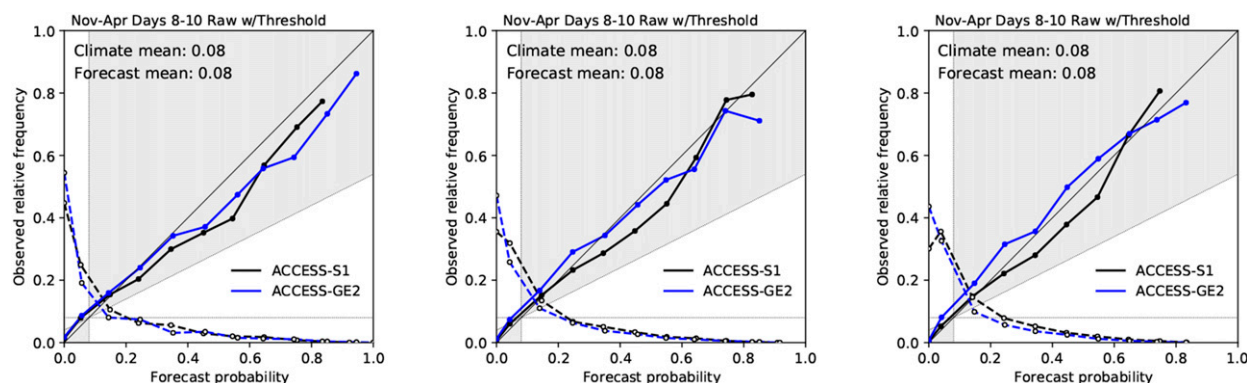


FIG. 4. Reliability plots for ACCESS-S1 and GE2 tracks using the wind speed threshold for 8–10-day forecasts and varying forecast lag: (left) lag 0, (center) lag 2, and (right) lag 3 forecasts.

TABLE 4. BSSs for IFS tracks with default ECMWF storm thresholds.

Forecast period	8–10	8–14	15–21	22–28
	0.294	0.299	0.179	0.149

the weighting between the IFS ensembles (more skillful) and the ACCESS-S1 ensembles (less skillful). The “default” subseasonal MME (with lag 0) is an 84-member ensemble, with 51 members from the IFS and 33 from ACCESS-S1, giving a ratio between the two systems of 1.54:1. Using lag 2 or lag 3 systems changes this ratio to 0.77:1 and 0.52:1, respectively. For 10-day lead times, the MME is a 108 member ensemble, with an additional 24 members sourced from ACCESS-GE2. The skill score of the 108-member MME forecasts shows the same behavior as the 84-member MME, that is, scores degrade as the ratio of ACCESS-GE2/S1 ensembles increases.

The MME improves reliability over the standard IFS forecasts (shown in Fig. 5), particularly for high forecast probabilities. The reliability improves due to the larger ensemble size. For 8–10-day forecasts, the addition of ACCESS-GE2 forecasts to the MME (shown in magenta) shows a useful improvement in reliability compared to the 84-member MME (shown in red).

Further explanation regarding causes of the improved reliability of the MME can be found by analyzing the spatial anomalies of each forecast system over the two cyclone seasons (see Fig. 7). These spatial anomalies are created by first computing the forecast storm density over all available forecasts for a given lead time. The observed storm density corresponding to these forecast valid times is then subtracted from the forecast density. Hence, positive anomalies correspond to a region where the model is overconfident (i.e., forecasting false alarms).

The reliability plots of both systems showed they both overforecast cyclone events. However, the spatial distribution of these forecast biases differs for both

systems across the Southern Hemisphere. Figure 7 shows that while the ACCESS-S1 is overconfident, most of the false alarms are centered over the Indian Ocean between 60° and 75°E longitude. This is most likely related to the well-known wet bias in the Met Office atmospheric model in this region. See Bush et al. (2015) for analysis of this bias using Global Atmosphere 3.0 (GA3.0) models at N96 resolution, Levine and Martin (2018) for GA3.0 analysis at N512 resolutions and Jin et al. (2019) for analysis of GA6.0 at N96 and N216 resolutions. Hence in this region, the wet bias will excessively trigger the humidity thresholds of the OWZ tracking scheme. This bias may contribute to the findings of Titley et al. (2020) who showed that Met Office 7-day global cyclone forecasts [generated from the Met Office Global and Regional Ensemble Prediction System (MOGREPS) system] had their worst skill in the southwest Indian Ocean. Over the remainder of the Southern Hemisphere, the ACCESS-S1 system tends to underforecast cyclone events.

In contrast, the IFS consistently overforecasts cyclone activity throughout the Southern Hemisphere. The exception occurs in the Indian Ocean, where it underpredicts activity between 75° and 90°E (a feature it shares with the ACCESS-S1 system). Titley et al. (2020) also verified an MME for 7-day lead times (using MOGREPS, IFS, and the U.S. National Center for Environment Prediction’s Global Ensemble Forecast System), which showed the IFS cyclone forecasts had their lowest skill in the southwest Indian Ocean. Clearly this region of the Southern Hemisphere contains challenges for cyclone forecasting.

For our subseasonal MME, combining the two model forecasts together reduces forecast bias over much of the Southern Hemisphere, which improves forecast performance. For 8–10-day forecasts, the spatial forecast anomaly distributions are similar. Figure 8 shows the spatial errors for the two ACCESS system for 8–10-day forecasts. This plot suggests the improved skill recorded by ACCESS-GE2 over S1 is attributed to improvements in south Indian

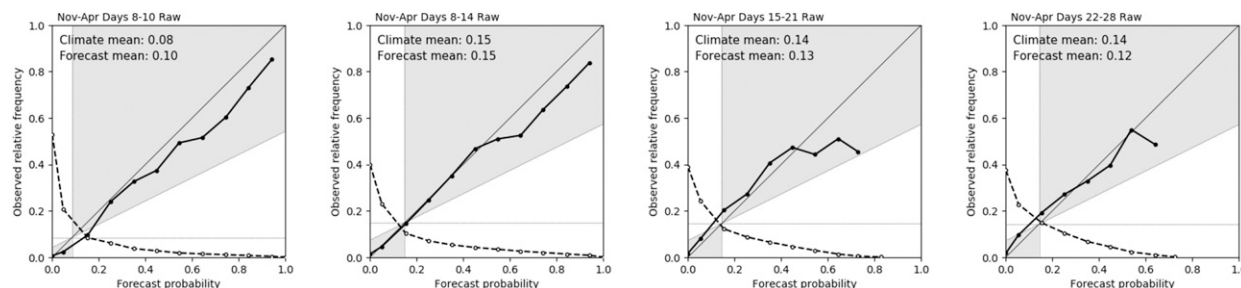


FIG. 5. Reliability plots for IFS tracks for varying lead times (shown from left to right): 8–10, 8–14, 15–21, and 22–28 days.

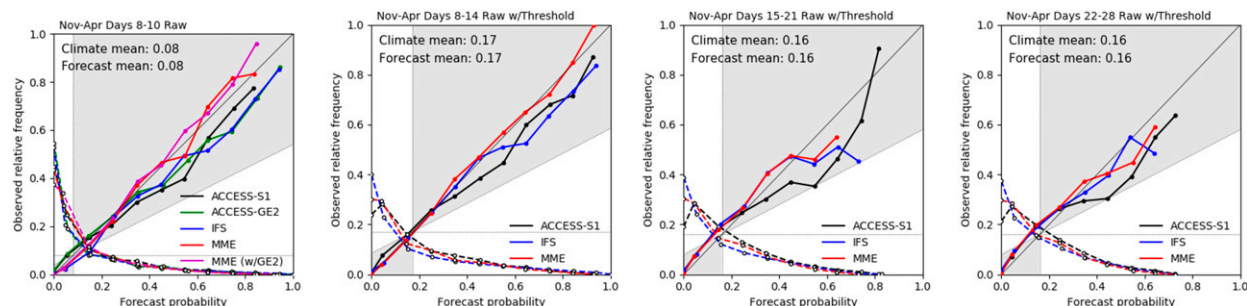


FIG. 6. Reliability plots for an MME, combining IFS, ACCESS-S1, and ACCESS-GE2 tracks for varying lead times (shown from left to right): 8–10, 8–14, 15–21, and 22–28 days.

Ocean near 70°E and in the vicinity of Noumea (160°E–180°).

5. Case studies

Two cases studies show forecasts of the two systems and examples of MME forecasts. The first case study was Cyclone Gita, which impacted Tonga as a category 4 cyclone in February 2018. The second is Cyclone Kenneth, which struck Mozambique in April 2019, also as a category 4 storm.

a. Cyclone Gita

Cyclone Gita formed in the South Pacific Ocean on 9 February and made landfall as a category 4 cyclone in Tonga two days later. Gita was the strongest cyclone of the 2017/18 South Pacific season and its track exhibited significant recurvature. At the same period, Cyclone Cebile (which formed in the southwest Indian Ocean on 27 January at 85°E) weakened to a tropical low on 9 February after earlier reaching category 4 intensity. Figure 9 shows the ACCESS-S, IFS, and MME forecasts issued on 22 January and 29 for 15–21- and 8–14-day forecasts, valid between 6 and 12 February 2018.

The ACCESS-S1 three-week forecast for 22 January incorrectly forecast cyclone activity in the Coral Sea and east of the date line. Good guidance of storm activity was forecast in the Indian Ocean, while the model suggested two likely centers of storm formation in the South Pacific at 175°E and 170°W, respectively. The IFS forecast for this date created false alarms throughout northern Australia in addition to the Coral Sea, as well as near Madagascar and Mozambique. However, it correctly clustered most tracks around Fiji. The MME reduced the probabilities associated with the false alarms across northern Australia, while centering the chance of cyclone activity in the coming three weeks around Fiji at 180°. MME guidance for storm activity in the Indian Ocean lacked a coherent signal.

Moving to the two-week forecasts issued on 29 January, both systems moved the location of most likely storm formation farther east of the date line at around 170°W. ACCESS-S1 no longer forecast any storm activity in the Indian Ocean. This is most likely attributable of the OWZ tracking algorithm, as forecast storm tracks south of 25°S would be unlikely to meet the sustained humidity thresholds for cyclogenesis. Forecast storm probabilities in the Coral Sea have lowered significantly. The IFS forecast in contrast gives excellent guidance for Cebile, while no longer forecasting false alarms in northwest Australia (although it still suggests a strong chance of cyclone formation in the Coral Sea). The MME forecast concentrates storm activity in the Pacific at around 165°W. The MME signal surrounding Cebile is still strong (due to the IFS forecast) and false alarms were still forecast near Madagascar/Mozambique and in the Coral Sea.

Examining 10-day forecasts issued on 2 February, Fig. 10 shows that ACCESS-GE2 forecast storm formation in the Pacific north of 20°S, while ACCESS-S1 shifts the forecast storm location southward. Both ACCESS-S systems produced small false alarms of northwest Australia. The IFS forecast strike probability in the Pacific is much more coherent than the corresponding ACCESS systems, which is most likely attributable to its finer spatial resolution. Both the ACCESS and IFS 10-day forecasts are more spatially coherent than the 14-day equivalent, which is to be expected with shorter lead times. The IFS continued to forecast false alarms in the Coral Sea, although this signal is lower in the MME.

TABLE 5. BSSs for MME tracks. These statistics are computed for available IFS forecasts using ACCESS-S1 tracks with a variety of forecast lag. Scores computed for days 8–10 also include ACCESS-GE2 forecasts.

Forecast period	Lag 0	Lag 2	Lag 3
8–10	0.311	0.295	0.283
8–14	0.314	0.298	0.286
15–21	0.181	0.149	0.130
22–28	0.143	0.117	0.112

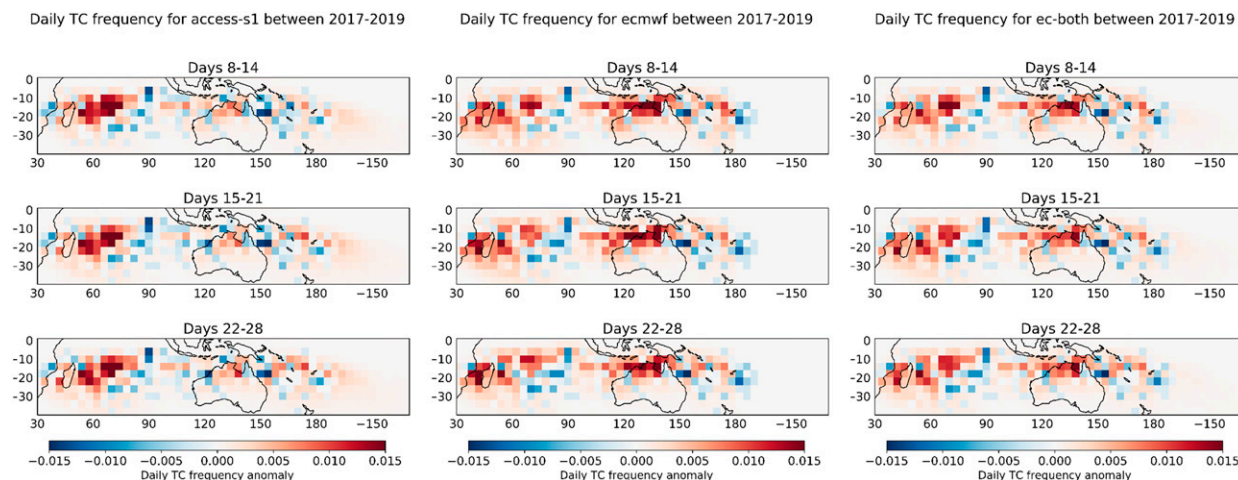


FIG. 7. Spatial anomaly plots computed for the (left) ACCESS-S1 system, (center) IFS system, and (right) MME for the 2017/18 and 2018/19 cyclone seasons. The anomalies are computed in 4° boxes. Lead time increases for each row: (top) 8–14, (middle) 15–21, and (bottom) 22–28 days.

b. Cyclone Kenneth

Cyclone Kenneth formed off the coast of Mozambique on 21 April 2019. It made landfall as a category 4 cyclone over Mozambique on 25 April, making it the strongest storm to do so in recorded history. Kenneth was the strongest storm of the 2018/19 south Indian Ocean season. Around the same time, Cyclone Lorna formed southeast of the Maldives on 22 April and continued to be active around 90°E until 30 April. Figure 11 shows the ACCESS-S, IFS and MME forecasts issued on 4 and 10 April for 15–21- and 8–14-day forecasts, valid between 19 and 25 April 2019.

The ACCESS-S1 3-week forecast for 4 April gave good guidance for storm activity in the south Indian Ocean from 50° to 90°E . The IFS forecast was less emphatic in this region while also forecasting low chances of cyclone formation south of Indonesia and into the Coral Sea. Hence the MME in this case performs worse than the single ACCESS-S1 forecast.

This trend continued for the two-week forecasts issued on 10 April. The ACCESS-S1 forecast now predicted the formation two separate storm systems, one centered around 50°E (with a chance of landfall in Mozambique) and another centered on 75°E . The IFS forecast had a similar spatial distribution throughout the south Indian Ocean but with lower forecast probabilities near Madagascar. Very low probabilities of cyclone formation were forecast throughout Indonesia, Papua New Guinea, and the Solomon Islands. As with the three-week forecast, the 2-week MME forecast is inferior to the ACCESS-S1 forecast.

The 10-day forecast issued on 14 April shown in Fig. 10 shows little differences between the two ACCESS systems,

although the GE2 ensemble favored a “split track” scenario for Kenneth, forecasting both westward and southward trajectories. The IFS 10-day forecast showed a similar spread, whereas the ACCESS-S1 system maintained a more zonal distribution of forecast trajectories. The MME forecast at the shorter lead time showed slight

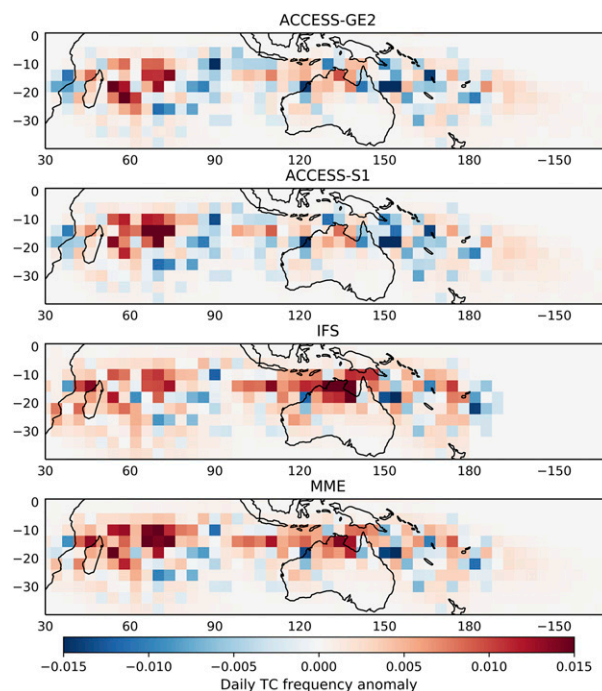


FIG. 8. Spatial anomaly plots computed for various models with 8–10-day lead times. (top) ACCESS-GE2, (top middle) ACCESS-S1, (bottom middle) IFS, and (bottom) MME for the 2017/18 and 2018/19 cyclone seasons. The anomalies are computed in 4° boxes.

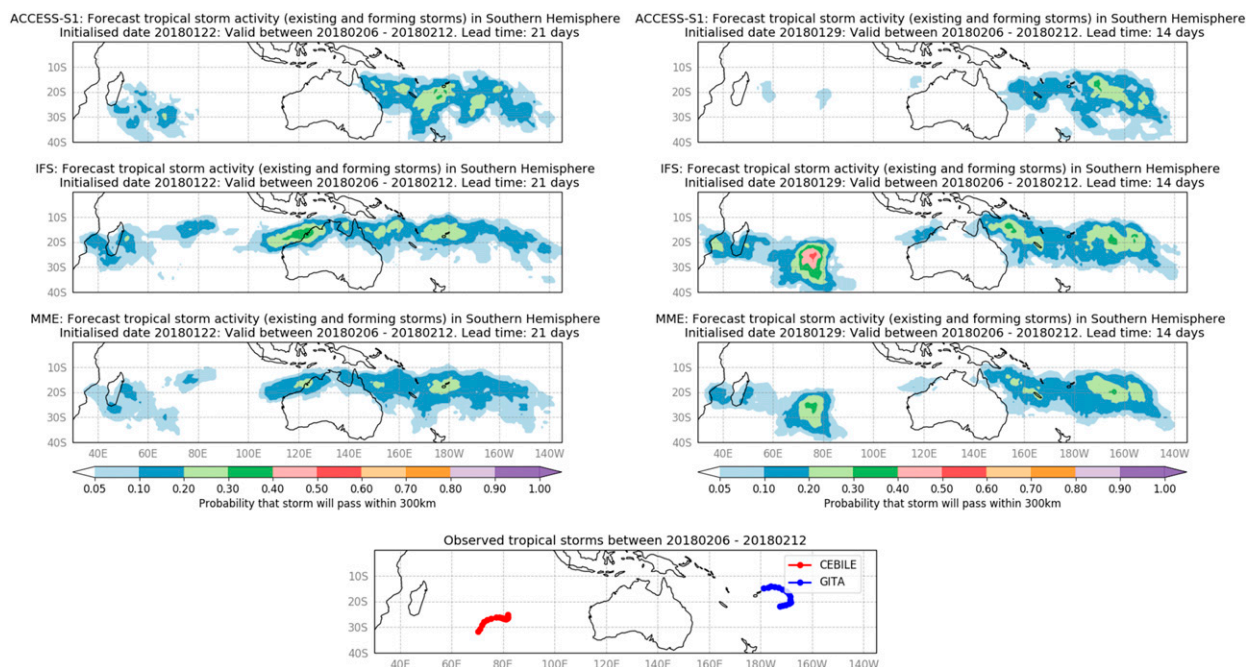


FIG. 9. ACCESS-S1, IFS, and MME forecasts for Cyclones Gita and Cebile with (left) 15–21-day lead times initialized on 22 Jan 2018 and (right) 8–14-day lead times initialized on 29 Jan 2018. All forecasts cover the period between 6 and 12 Feb 2018. Observed storm locations for Gita and Cebile for this period are shown in the bottom panel.

changes to the 8–14-day forecast, with increased chances of storm formation at the northern tip of Madagascar.

6. Discussion and conclusions

Evidence collated over the past two Southern Hemisphere cyclone seasons shows that skillful, sub-seasonal forecasts are now possible with the latest generation of coupled dynamic ocean–atmosphere models. However, this skill has been assessed at large spatial scales (i.e., $15^\circ \times 20^\circ$ latitude/longitude). So, while these model forecasts can predict cyclogenesis locations in subregions of the Southern Hemisphere, forecasts of specific storm impacts at such lead times is still beyond reach.

Determining why the IFS has superior skill than both ACCESS systems is beyond the scope of this study, as both systems used different storm trackers. Hence it is impossible to decouple model forecast errors from storm tracker errors. Both trackers use different meteorological variables to identify storms (or proxies of storm location). The algorithms used to generate storm tracks also differ (i.e., joining together in time the locations of points that exceed the predefined meteorological thresholds). For example, while the ECMWF algorithm uses the steering winds at a specific point to look into the future for the next candidate point along a track, the OWZ algorithm also searches backward in

time, that is, points that are matched by integrating forward in time are checked by integrating the steering winds in the later location backward in time to ensure the two points match.

Nevertheless, there are important differences between the two models. ACCESS-S1 has a fixed resolution of N216 (60 km), whereas the IFS changes from N639 (days 1–15) to N319 (days 16–46). It is expected that higher resolution of the IFS gives it an advantage in resolving flow gradients associated with small storms, especially those that form near the coast. In addition, the OWZ tracking scheme interpolates all data onto a $1^\circ \times 1^\circ$ grid, so as not to track individual thunderstorm systems. Hence applying the OWZ scheme to higher-resolution forecasts may not improve tracker performance.

As mentioned previously, the IFS uses a 51-member ensemble of 4D-VAR forecasts, where each forecast has perturbations to the observation and background error covariances. ACCESS-S1 uses only scaled random perturbations to generate its ensemble spread. The impact of data assimilation on second-week forecast skill was examined by comparing results between ACCESS-S1 and ACCESS-GE2. Recall that both systems used the same atmospheric model at the same resolution; however, the GE2 system generated 24 ensemble members using an ensemble Kalman filter. While GE2 uses fixed ocean sea surface temperatures (compared to S1, which

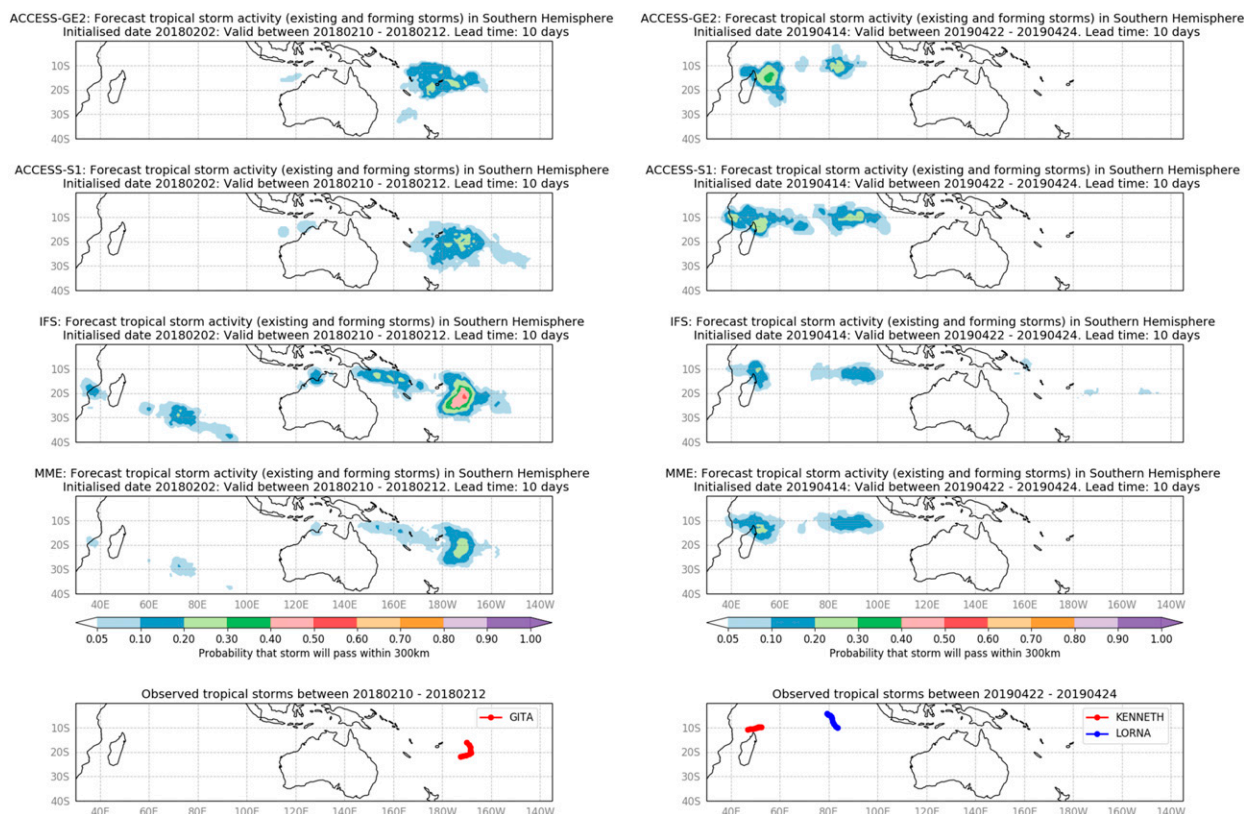


FIG. 10. ACCESS-GE2, ACCESS-S1, IFS, and MME 8–10-day forecasts for (left) Cyclones Gita and Cebile initialized on 2 Feb 2018 and (right) Cyclones Kenneth and Lorna initialized on 12 Apr 2019. Observed storm locations for Gita, Cebile, Kenneth, and Lorna for these periods are shown in the bottom panels.

is coupled to a full ocean model), GE2 was consistently more skillful across 8–10-day lead times. This is additionally impressive when considering the GE2 ensemble size is smaller than S1. This proves that data assimilation is still an important source of second-week forecast skill of cyclogenesis.

Ensemble size is also an important factor in forecast skill, as demonstrated by

- 1) ACCESS-S1 lagged ensembles forecasts having consistently more reliability than the default forecasts,
- 2) IFS forecasts having better skill than ACCESS-S1, and
- 3) the MME having better skill and reliability than both systems.

However, points 1 and 2 are made on the assumption that storm tracker errors are not significant.

The impact of the model's MJO forecasts on cyclogenesis is another avenue for investigation. Extra seasons are required to record sufficient MJO events to generate robust statistics. Based on the average periodicity of the MJO (typically defined as 30–60 or 30–90 days) there may only be three MJO cycles per season. This is an important avenue of further

research as many statistical forecasting schemes use MJO indices as predictors. If the dynamic models show the ability to forecast cyclogenesis at subseasonal time frames with a weak/indeterminate MJO signal, this represents a significant improvement to forecast products and services in the region.

Another point of difference is the two different methodologies used at ECMWF and BoM to deliver “seamless forecasts” between weather and seasonal forecast time scales. The IFS subseasonal forecasting system is based on the ECMWF ensemble numerical weather prediction system. It is the same model (and initial conditions), iterated for longer lead times and coarser resolution after 15 days. Hence it can be considered as an “extended” numerical weather prediction system.

In contrast, ACCESS-S1 is designed primarily for seasonal prediction, with 11 ensemble members integrated to 210 days. The atmospheric ensemble generation (scaled random perturbations) is designed more to generate skillful spread at seasonal lead times. An additional 22 members (integrated to a lead time of 42 days) is required to produce forecast skill at

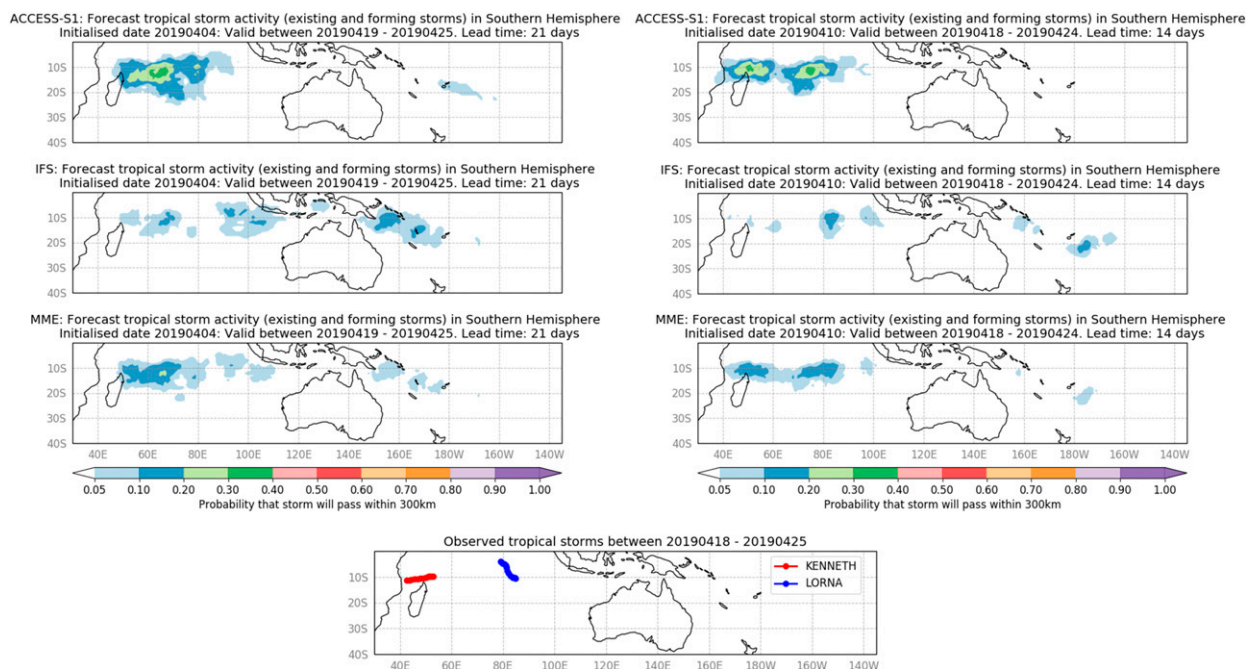


FIG. 11. ACCESS-S1, IFS, and MME forecasts of Cyclones Kenneth and Lorna with (left) 15–21-day lead times initialized on 4 Apr 2019 and (right) 8–14-day lead times initialized on 10 Apr 2019. All forecasts cover the period between 18 and 25 Apr 2019. Observed storm locations for Kenneth and Lorna for this period are shown in the bottom panel.

subseasonal time frames. Hence, the subseasonal forecasts generated by ACCESS-S1 can be considered as by-product of a seasonal forecasting system.

An alternative way forward for the Bureau may be to leverage off ACCESS-GE3, the recently operational version of GE2. This uses an updated version of the UKMO GA6.0 at higher resolution (N400, 33 km), using improved assimilation (hybrid 4DVAR), to generate 10-day forecasts with 18 ensemble members every 6 h. A lagged ensemble forecast of this system (coupled to an ocean model) integrated out to 28 days could provide the best way to deliver seamless forecasts in the multi-week space.

Determining the factors important for subseasonal forecast skill of cyclogenesis requires decoupling forecast errors from tracker errors. Such studies are already ongoing using the S2S database (s2sprediction.net), where the ECMWF tracking algorithm is applied to forecasts from several national agencies. Lee et al. (2018) showed that cyclone forecasting skill is dependent on model MJO forecast skill and climatology biases, both of which can vary from basin to basin.

The next step in this process would be to determine forecast uncertainties due to the tracking algorithm and apply several different trackers to the same model outputs.

Further progress could be made by creating an MME of forecast anomalies relative to each model's climatology. The ECMWF already issues subseasonal

forecast anomalies. While the BoM does issue calibrated cyclone forecasts (which rely on an 11-member 1990–2012 hindcast), it has yet to compute a complete cyclone climatology.

Skillful subseasonal forecasts of cyclone formation have the potential to provide massive benefits to the emergency services community by helping with disaster preparation and planning. There are also benefits to the commercial sectors and general public to help cope with supply-chain and transport disruptions. The study showed that two separate forecast models using two separate tracking algorithms are both capable of producing skillful subseasonal forecasts, and forecast skill was further improved by combining both into a multimodel ensemble. The authors believe that such results should spur the various national and international agencies responsible for issuing cyclone forecasts and warnings to commit to further development and dissemination of such products to protect lives and infrastructure throughout the region.

REFERENCES

- Buizza, R., M. Leutbecher, and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**, 2051–2066, <https://doi.org/10.1002/qj.346>.
- Bush, S. J., A. G. Turner, S. J. Woolnough, G. M. Martin, and N. P. Klingaman, 2015: The effect of increased convective entrainment on Asian monsoon biases in the MetUM general

- circulation model. *Quart. J. Roy. Meteor. Soc.*, **141**, 311–326, <https://doi.org/10.1002/qj.2371>.
- Camargo, S. J., and Coauthors, 2019: Tropical cyclone prediction on subseasonal time-scales. *Trop. Cyclone Res. Rev.*, **8**, 150–165, <https://doi.org/10.1016/j.tcr.2019.10.004>.
- Camp, J., and Coauthors, 2018: Skillful multi-week tropical cyclone prediction in ACCESS-S1 and the role of the MJO. *Quart. J. Roy. Meteor. Soc.*, **144**, 1337–1351, <https://doi.org/10.1002/qj.3260>.
- Gao, J., and T. Li, 2011: Factors controlling multiple tropical cyclone events in the western North Pacific. *Mon. Wea. Rev.*, **139**, 885–894, <https://doi.org/10.1175/2010MWR3340.1>.
- Gregory, P., J. Camp, K. Bigelow, and A. Brown, 2019: Subseasonal predictability of the 2017–2018 Southern Hemisphere tropical cyclone season. *Atmos. Sci. Lett.*, **20**, e886, <https://doi.org/10.1002/asl.886>.
- Hudson, D., and Coauthors, 2017: ACCESS-S1: The new Bureau of Meteorology multi-week to seasonal prediction system. *J. South. Hemisphere Earth Syst. Sci.*, **67**, 132–159, <https://doi.org/10.22499/3.6703.001>.
- Jin, L., H. Zhang, A. Moise, G. Martin, S. Milton, and J. Rodriguez, 2019: Australia-Asian monsoon in two versions of the UK Met Office Unified Model and their impacts on tropical–extratropical teleconnections. *Climate Dyn.*, **53**, 4717–4741, <https://doi.org/10.1007/s00382-019-04821-1>.
- Lee, C.-Y., S. J. Camargo, F. Vitart, A. H. Sobel, and M. K. Tippett, 2018: Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Wea. Forecasting*, **33**, 967–988, <https://doi.org/10.1175/WAF-D-17-0165.1>.
- Leroy, A., and M. C. Wheeler, 2008: Statistical prediction of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.*, **136**, 3637–3654, <https://doi.org/10.1175/2008MWR2426.1>.
- Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227**, 3515–3539, <https://doi.org/10.1016/j.jcp.2007.02.014>.
- Levine, R. C., and G. M. Martin, 2018: On the climate model simulation of Indian monsoon low pressure systems and the effect of remote disturbances and systematic biases. *Climate Dyn.*, **50**, 4721–4743, <https://doi.org/10.1007/s00382-017-3900-x>.
- Lock, S.-J., S. Lang, M. Leutbecher, R. Hogan, and F. Vitart, 2019: Treatment of model uncertainty from radiation by the Stochastically Perturbed Parametrization Tendencies (SPPT) scheme and associated revisions in the ECMWF ensembles. *Quart. J. Roy. Meteor. Soc.*, **145**, 75–89, <https://doi.org/10.1002/qj.3570>.
- MacLachlan, C., and Coauthors, 2015: Global Seasonal Forecast System Version 5 (GloSea5): A high-resolution seasonal forecast system. *Quart. J. Roy. Meteor. Soc.*, **141**, 1072–1084, <https://doi.org/10.1002/qj.2396>.
- Schreck, C. J., K. R. Knapp, and J. P. Kossin, 2014: The impact of best track discrepancies on global tropical cyclone climatologies using IBTrACS. *Mon. Wea. Rev.*, **142**, 3881–3899, <https://doi.org/10.1175/MWR-D-14-00021.1>.
- Titley, H. A., R. L. Bowyer, and H. L. Cloke, 2020: A global evaluation of multi-model ensemble tropical cyclone track probability forecasts. *Quart. J. Roy. Meteor. Soc.*, **146**, 531–545, <https://doi.org/10.1002/qj.3712>.
- Tory, K. J., R. A. Dare, N. E. Davidson, J. L. McBride, and S. S. Chand, 2013: The importance of low-deformation vorticity in tropical cyclone formation. *Atmos. Chem. Phys.*, **13**, 2115–2132, <https://doi.org/10.5194/acp-13-2115-2013>.
- Van der Grijn, G., J.-E. Paulsen, F. Lalaurett, and M. Leutbecher, 2005: Early medium-range forecasts of tropical cyclones. *ECMWF Newsletter*, No. 102, ECMWF, Reading, United Kingdom, 7–14, <https://www.ecmwf.int/node/14623>.
- Vitart, F., J. L. Anderson, and W. F. Stern, 1997: Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations. *J. Climate*, **10**, 745–760, [https://doi.org/10.1175/1520-0442\(1997\)010<0745:SOIVOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<0745:SOIVOT>2.0.CO;2).
- , A. Leroy, and M. C. Wheeler, 2010: A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.*, **138**, 3671–3682, <https://doi.org/10.1175/2010MWR3343.1>.
- , F. Prates, A. Bonet, and C. Sahin, 2012: New tropical cyclone products on the web. *ECMWF Newsletter*, No. 130, ECMWF, Reading, United Kingdom, 17–23, <https://www.ecmwf.int/node/17427>.
- Walters, D., and Coauthors, 2017: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations. *Geosci. Model Dev.*, **10**, 1487–1520, <https://doi.org/10.5194/gmd-10-1487-2017>.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Cambridge Press, 547 pp.