The tropical west Pacific Ocean and the Philippines are often affected by tropical cyclones (TCs), with threats to human life and of severe economic damage. The performance of the Met Office global operational forecasts at predicting TC-related precipitation is examined between 2006 and 2017, the first time total TC rainfall has been analyzed in a long-term forecast dataset. All precipitation falling within 8° of a TC track point is assumed to be part of the TC rainbands. Forecasts are verified against TC tracks from the JRA-55 reanalysis and precipitation from TRMM 3B42. In composites from the forecasts, the total precipitation (TC and non-TC) is too high and the TC-related precipitation is too low, over both ocean and the Philippines. These biases exist all year-round and generally worsen with lead time, but have improved in recent years with upgrades to the forecasting system. Biases in TC-related precipitation in the Philippines are attributable mainly to TC lifetime being too short over land and ocean and (over land) possibly to individual TCs producing too little rain. There are considerable biases in predicted large-scale conditions related to TC intensification, particularly too little lower-troposphere relative humidity and too strong vertical wind shear. The shear appears to have little impact on the amount of TC precipitation, but dry biases in humidity are consistent with dry biases in TC rainfall. The forecast system accurately reproduces the impact of the MJO on TC precipitation, relative to the forecasts’ own climatology, potentially providing the opportunity for predictability out to several weeks.
and less in a suppressed MJO (e.g., Hall et al. 2001; Bessafi and Wheeler 2006; Ho et al. 2006; Kim et al. 2014).

Another large-scale phenomenon to affect TC activity is El Niño–Southern Oscillation (ENSO), which is responsible for much of the TC interannual variability. TC frequency in the west Pacific is reduced during El Niño and enhanced during La Niña events (e.g., Corporal-Lodangco et al. 2016). However, accumulated cyclone energy (ACE) is greater during El Niño than La Niña due to differences in TC lifetime and intensity (Camargo and Sobel 2005), although there is evidence that the relationship between ACE and El Niño in particular holds only during the boreal summer (Corporal-Lodangco et al. 2016).

Clearly there is a human and economic need for skillful forecasts of TCs and their impact. This study considers the ability of 11 years’ worth of global operational numerical weather prediction (NWP) forecasts from the Met Office (UKMO) to reproduce TC-related precipitation over the Philippines and the surrounding area. We believe this is the first study to analyze overall TC rainfall in a long-term forecast dataset, although Luitel et al. (2018) analyzed forecasts of case studies of landfalling TCs in the North Atlantic Ocean (see below). The NWP system evaluated here uses the Met Office Unified Model (MetUM). Liang et al. (2017) investigated TCs in two regional climate model (RCM) MetUM configurations, forced at the boundaries by ERA-Interim, with 25 km (which is the finest grid spacing analyzed here) and 12 km grid spacings. They found that while both RCMs had a realistic large-scale environment for TC activity, they had biases in the annual cycle, with TC frequency being too low in the observed peak season of June–September and too high during May and November–January. Improvements were seen with increasing resolution, with the TC track density and radial wind structure being more accurate in the 12-km RCM. Lee et al. (2018) analyzed TC activity in six models, not including the MetUM, in the Subseasonal to Seasonal (S2S) database and found a range of forecast skill, but most models had skill during the first seven days, which is also the length of the operational forecasts used in this study. Luitel et al. (2018) investigated five NWP systems, including the UKMO forecasts, and found they all have skill out to 2 days in the prediction of precipitation from 12 landfalling hurricanes over the east coast of the United States, but that no model stood out as being consistently the best.

In our study, precipitation associated with forecast TCs is verified using gridded precipitation from the Tropical Rainfall Measuring Mission (TRMM) and TCs in the Japanese 55-year Reanalysis (JRA-55; see section 2a). The methodology for compositing TC-related precipitation (see section 2b) mirrors Guo et al. (2017) and Franco-Díaz et al. (2019, manuscript submitted to Climate Dyn.). Those studies analyzed observations only, using the same TRMM product but taking TC tracks from ERA-Interim (Guo et al. 2017), and both ERA-Interim and JRA-55 (Franco-Díaz et al. 2019, manuscript submitted to Climate Dyn.) The latter, which studied TCs over Middle America (Central America, Mexico, and the southern United States), the east Pacific Ocean and the North Atlantic Ocean, found that the composite TC rainfall tends to be higher for
JRA-55 than ERA-Interim tracks, especially in the tropical east Pacific, although the TC contribution to extreme precipitation was very similar in each. Our composites of satellite-derived TC precipitation can be compared against those of Guo et al. (2017), who also analyzed the west satellite-derived TC precipitation can be compared against precipitation was very similar in each. Our composites of cal east Pacific, although the TC contribution to extreme JRA-55 than ERA-Interim tracks, especially in the tropi-

Section 2a of this paper describes the datasets used; the TC tracking and precipitation compositing techniques are detailed in section 2b. The results from observations and reanalysis are presented in section 3a; results from the forecasts are verified against them in section 3b. The annual cycle is analyzed in section 3c; section 3d considers how forecast performance has changed during the analysis period. In section 4 we consider the effect of the MJO; a discussion and conclusions are presented in section 5.

2. Data and methods

a. Datasets and forecasts

This study uses several datasets to identify TCs, and produce composites of TC-related precipitation and other large-scale meteorological conditions. The International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al. 2010) combines TC track data from numerous weather centers to produce a comprehensive database of historical tracks. Tracks are also identified in JRA-55 from the Japan Meteorological Agency (JMA; Kobayashi et al. 2015), using the technique described in section 2b. Reanalysis and forecast tracks are matched against the IBTrACS climatology to ensure that only those tracks corresponding to observed TCs are included. We use tracks in JRA-55 rather than IBTrACS since JRA-55 tracks last longer (IBTrACS tracks cyclones only during their TC phase, whereas in JRA-55 we are able to track cyclones before and after this phase, when they lack features such as the warm core that typifies a TC). Furthermore, we are able to apply an identical TC tracking method to JRA-55 and the forecasts. We choose JRA-55 over other reanalysis products as Murakami (2014) found it had the most accurate representation of TCs compared with five other reanalyses; Hodges et al. (2017) found JRA-55 had the smallest TC location errors, relative to IBTrACS, of six reanalyses studied.

Precipitation observations are taken from the TRMM 3B42 product (Huffman et al. 2007), version 7. This is a gridded product that merges satellite (using microwave instruments on satellite orbiters where available, filling in gaps with infrared-derived estimates from geostationary satellites) and gauge data. TRMM 3B42 is available as 3-hourly accumulations on a 0.25° grid between 50°N and 50°S. To examine the climatological large-scale environment, wind and relative humidity on pressure levels are taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim; Dee et al. 2011). The large-scale climatologies are very similar in ERA-Interim and JRA-55.

The Real-Time Multivariate MJO (RMM) indices of Wheeler and Hendon (2004) are used to track the amplitude and phase of the MJO. Days with RMM amplitude < 1 are excluded from MJO composites.

We analyze UKMO global operational forecasts initialized at 0000 and 1200 UTC daily, from 0000 UTC 1 January 2006 to 0000 UTC 11 July 2017. All forecasts ran for 7 days, but for 2006 and 2007 we ignore day 7 as precipitation was not archived. Of the 8419 forecasts in this period, 134 were either entirely or partially missing from the UKMO archive and were excluded from the analysis. Forecasts during the chosen period used three horizontal resolutions (Table 1); the end of the analysis period in July 2017 was the final forecast before the resolution changed again. In section 3d we bin forecasts by horizontal resolution to understand how performance has changed over time. Forecasts were not analyzed after the resolution change in 2017 as there would have been a small sample size of forecasts at the new resolution. Changes are not necessarily attributable to the resolution change since many other modifications were made over the same period (e.g., the dynamical core, physics parameterizations and data assimilation), but the infrequency of resolution changes provides a convenient way of splitting up the forecasts while retaining reasonable sample sizes.

As of the start of our analysis period, a TC initialization scheme was used in the MetUM, whereby “bogus observations,” based on TC theory, were introduced to the model near observed TCs (Heming et al. 1995). This scheme was upgraded in November 2007 to introduce these “observations” over a smaller area in the case of small TCs (Heming 2009), but switched off entirely on 17 July 2012 because by then other model upgrades

<table>
<thead>
<tr>
<th>Date range</th>
<th>Grid spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000 UTC 1 Jan 2006</td>
<td>60 km</td>
</tr>
<tr>
<td>1200 UTC 9 Mar 2010</td>
<td>40 km</td>
</tr>
<tr>
<td>1200 UTC 15 Jul 2014</td>
<td>25 km</td>
</tr>
</tbody>
</table>
caused it to degrade TC forecasts (Heming 2016). Other notable changes have included increasing vertical resolution from 50 levels with a 65-km lid to 70 levels with an 80-km lid (10 November 2009). Upgrades were made to the data assimilation over the period, including to the assimilation of surface temperature, relative humidity, and winds over land from 1 April 2008, and relative humidity from aircraft from 15 March 2016. The horizontal grid spacing of the data assimilation model increased from 200 to 140 km (9 March 2010), to 90 km (2 November 2010), then to 60 km (15 July 2014). Note that these grid spacings are stated in longitude at the equator because we are interested in the tropics only, whereas other studies typically state them in the mid-latitudes where the longitudinal and latitudinal spacings are similar. 15 July 2014 also saw the introduction of the Global Atmosphere 6.1 (GA6.1) configuration (Walters et al. 2017) with a new dynamical core, Even Newer Dynamics for General atmospheric modeling of the environment (ENDGame), and other changes such as a 25% increase in the entrainment rate for deep convection. GA6.1 and a new TC initialization scheme which assimilates central pressure estimates from international TC warning centers (introduced on 3 February 2015) both improved TC forecasts (Heming 2016).

b. TC identification

Our method for identifying TCs is described by Hodges and Klingaman (2019) and Hodges and Emerton (2015), and is summarized here. Relative vorticity data at 6-hourly intervals are processed from the JRA-55 reanalysis and the forecasts. The tracking scheme of Hodges (1994, 1995, 1999) is used.

First, in both datasets, relative vorticity is vertically averaged between 850 and 600 hPa, then spatially filtered to T63 to remove high-spatial-frequency noise. To remove the large-scale background, all variability at wavenumbers $n \leq 5$ is also removed. Local maxima exceeding $5 \times 10^{-6} \text{s}^{-1}$ in the resulting field are identified within the band $0^\circ$–$60^\circ$N. All systems satisfying this criterion are tracked and retained if they last for at least 2 days. As a measure of storm intensity, the maximum 10-m wind within a 6º radius of each track point is added to the dataset of vorticity tracks.

JRA-55 tracks are then matched against IBTrACS. If a JRA-55 track overlaps in time with an IBTrACS track for at least 10% of their track points, with a mean separation of less than 4º, it is deemed to be the same TC and retained. This matching method is explained in detail by Froude et al. (2007; see their Fig. 3). Forecast tracks are then deemed to be the same TC as a JRA-55 track if their mean separation over the first day of the forecast track (note, not necessarily the first day of the forecast itself) is less than 4º, provided the TC exists within the first three days of the forecast. For both JRA-55 and the forecasts, we retain only those TC tracks which match, so when validating forecasts against observations we use identically the same storms.

Having identified the matched TC forecast tracks, precipitation associated with each TC is composited using the technique of Guo et al. (2017). All TRMM 3B42 precipitation falling within 5º of a JRA-55 TC track point is assumed to be attributable to the TC, and the precipitation from all 6-hourly track points is composited on the TRMM grid. The composites do not include rain within 5º of the full path of the TC during its motion from one 6-h track point to the next, but the separation between track points is sufficiently small relative to the composite radius that this effect is negligible. Similarly, forecast precipitation is composited within 5º of forecast track points, on the model grid. A radius of approximately 5º is a common choice in TC studies (e.g., Dare et al. 2012; Villarini et al. 2014; Prat and Nelson 2016; Guo et al. 2017; Khouakhi et al. 2017; and many others). Bagtasa (2017) argued for a radius of 10º. However, this was to take into account not just rainfall within the TC rainbands but all the rainfall due to the change in circulation induced by the TC (e.g., strengthening of the monsoon flow). Here we choose a 5º radius because we are interested only in the precipitation directly attributable to the TC, and for consistency with Guo et al. (2017), whose methodology we reproduce.

3. TC-related precipitation

a. Observations

Figure 2a shows the year-round JRA-55 TC track density (the seasonal cycle will be presented in section 3c), calculated as the average number of tracks per year, per $2.5^\circ \times 2.5^\circ$ box. This box size is significantly larger than the typical error in track location: the distribution of track errors for JRA-55 against IBTrACS peaks at 0.5º (geodesic); almost all track points have an error $< 1^\circ$ (Hodges et al. 2017, see their Fig. 1a). In the west tropical Pacific and east Asia domain shown, the largest track density is over the South China Sea to the west and Philippine Sea to the east of Luzon ($> 8 \text{yr}^{-1}$ per box), with most of these storms making landfall over Luzon ($6 - 8 \text{yr}^{-1}$ per box). The track density over ocean to the east and west of Mindanao is almost as high around Luzon, but the rate of landfall is markedly lower, with a track density of only $2 - 4 \text{yr}^{-1}$ per box over land. The genesis density (Fig. 2b), defined using the first track point in JRA-55 for each TC, is greatest far to the east of...
the Philippines over the west Pacific, although some storms also form much closer to the Philippine coastline, both to the east and west. Unlike IBTrACS, our tracking method includes the early stages of the storms, so although all the storms we track become TCs at some stage of their lifetime, many of them will not have reached TC strength in this region of high genesis density.

In terms of maximum 10-m wind speed, TC intensity is greatest to the north of the Philippines (Fig. 2c) where TC tracks recurve northeastward, tracking along the east coast of China and southeast coast of Japan. Note that the true maximum 10-m wind intensity is likely to be higher than shown since JRA-55 has a coarse resolution (approximately 55 km) and contains errors in the pressure–wind relationship (Hodges et al. 2017). The overall TC-related rainfall is controlled more strongly by the track density than the intensity, as can be seen by comparing Figs. 2a and 3a. The latter shows TC-related precipitation is greatest over the ocean close to the northern Philippines, with TC rain tending to occur less over land than offshore. Over China, the TC-related precipitation falls to near zero within a few hundred kilometers of the coast.

By dividing by the total precipitation (Fig. 3b), we obtain the fraction of precipitation attributable to TCs (Figs. 3c,d). Over most of the ocean near Luzon, 30%–40% of precipitation is TC-related, peaking at over 40% in the South China Sea. There is strong meridional variation, with only 5%–15% of precipitation coming from TCs near Mindanao. Over the islands the contribution to rainfall is lower but with the same north–south dependence, with a 20%–30% contribution over Luzon and about 5% over Mindanao. Therefore, particularly in the north Philippines, we clearly confirm the importance of TCs for precipitation and, therefore, the socio-economic importance of skilful TC precipitation forecasts.

b. Forecast performance

The equivalent of Fig. 3 for the forecast bias is shown in Fig. 4, for four selected lead times $T$ ranging from 0 ≤ $T$ ≤ 1 days (“day 1”) to 6 ≤ $T$ ≤ 7 days (“day 7”). We first consider these composites over all TC tracks, before normalizing by track density in Fig. 5. TC-related precipitation (Figs. 4a–d) is very close to observations on day 1, with a very slight wet bias over some ocean regions, not exceeding 0.6 mm day$^{-1}$. Some coastal land

![Fig. 2. (a) JRA-55 track density (number of TC tracks per 2.5° box yr$^{-1}$); (b) JRA-55 genesis density (number of TCs with their first track point within a 2.5° box yr$^{-1}$); and (c) averaged JRA-55 maximum 10-m wind speed in TCs. All panels use year-round data.](image-url)

![Fig. 3. (a) Composite of TRMM 3B42 precipitation within a 5° radius of JRA-55 TC track points; (b) total TRMM 3B42 precipitation; (c) percentage contribution of TC-related precipitation [i.e., 100 × (a)/(b)]; (d) as in (c), but zoomed in on the Philippines. All panels use year-round data.](image-url)
regions have a dry bias, particularly Luzon and the eastern Visayan Islands, where the bias is on the order of 1 mm day$^{-1}$. TC-related precipitation steadily decreases with forecast lead time, with the dry bias over Luzon and the surrounding sea reaching 1.0–1.8 mm day$^{-1}$ by day 7.

This drift toward drier conditions is also seen in composites of total (i.e., not just TC-related) precipitation (Figs. 4e–h), although the sign of the bias is always positive in most of the domain. Day 1 has a particularly strong wet bias over ocean and much of the land of the Philippines and coastal Borneo (generally 1–3 mm day$^{-1}$). The dry bias over some coastal regions, especially Borneo, may be due to an underrepresentation of the strong diurnal cycle there (e.g., Yang and Slingo 2001; Biasutti et al. 2012).

Fig. 4. (a)–(d) Composites of forecast precipitation within a 5° radius of forecast TC track points, for four selected lead times, as a bias relative to observations (Fig. 3a); (e)–(h) As in (a)–(d), but for total precipitation (relative to Fig. 3b); (i)–(l) As in (e)–(h), but for the percentage contribution of TC-related precipitation to total precipitation (fractional difference relative to Fig. 3c expressed as a percentage; i.e., not a difference in percentage points). Regions in which Fig. 3c is below 5% are grayed out. “Day n” refers to lead times $n - 1 < T < n$.
Errors in the diurnal cycle over Maritime Continent islands frequently occur in models, including the MetUM, often due to either unrepresented or misrepresented convective physics, or the coarseness of the horizontal resolution relative to real coastal or topographical features (e.g., Love et al. 2011; Peatman et al. 2015). Following the initialization shock on day 1, the ocean wet bias reduces by a factor of $\frac{1}{2}$, but dry biases around

Fig. 5. (a) Composite of TC track density for JRA-55 (same as Fig. 2a; repeated for convenience). (b)–(d) Composites of forecast TC track density (TCs yr$^{-1}$ per 2.5° × 2.5° box), relative to (a). (e)–(h) As in (a)–(d), but for precipitation rate per TC (mm h$^{-1}$); this is calculated by dividing the precipitation composites in Figs. 3a and 4 [regridded to the 2.5° grid used in (a)–(d)] by the total composite TC track density per box, and further dividing by the TC tracking frequency (6 h). Regions in which the JRA-55 TC track density is less than 0.5 yr$^{-1}$ per box are grayed out.
the Philippines and Borneo strengthen, especially offshore.

Dividing the TC-related rainfall by the total, we again obtain the contribution of TCs to precipitation (Figs. 4i–l, shown as a percentage difference of the observed values in Fig. 3c). Although both the TC-related and total precipitation reduce with lead time, the fall in TC-related rainfall dominates as the fractional contribution also falls considerably with lead time, especially over the tropical west Pacific. Over the Philippines, although in observations there was a clear north–south variation in TC contribution, the fractional bias is broadly the same throughout the country, with a contribution approximately 30% lower than observed on day 1, and approximately 70% lower on day 7.

As explained in section 2b, TC tracks were matched between IBTrACS, JRA-55 and the forecasts, so the same storms contribute to the observed and forecast TC composites. However, each storm may have different lifetimes and intensities in each dataset, and these effects could (separately or together) account for the dry bias in TC-related precipitation.

Figures 5b–d shows composites of TC track density in the forecasts, presented as biases compared to JRA-55 (Fig. 5a). Track density is generally too low, worsening with lead time, implying forecast TC lifetime is too short. Biases in TC counts are higher over ocean than land. Meanwhile, the precipitation rate per TC is shown for observations in Fig. 5e. This is calculated by taking the sum of TC-related precipitation (as in Fig. 3a, multiplied by the number of days in the composite) and regridding to the 2.5° grid used for the track density composites. We then divide by the total number of TC tracks per grid box (as in Fig. 5a, multiplied by the number of years in the composite) to obtain the amount of rainfall in mm per TC per 6 h (since the TC tracking is 6 hourly). Dividing again by 6 h obtains the precipitation rate per TC (mm h\(^{-1}\) per TC). Over open ocean in the west tropical Pacific the rainfall rate associated with each TC is typically in the range 10–16 mm h\(^{-1}\), but is larger (>18 mm h\(^{-1}\)) over the major landfalling regions of the Philippines (especially the Visayan Islands) and the coast of Vietnam. Over these land regions the forecasts substantially underestimate the rainfall rate per TC, by a factor of \(~1/2\), at all lead times. Over the rest of the domain shown, the rainfall rate per TC is generally forecast accurately across lead times, with any biases being spatially incoherent. Therefore, the dry bias in TC rainfall in Figs. 4a–d is attributable to different causes over land and sea: over sea it is due to low storm counts (because of storms being too short-lived); over land it is primarily because individual TCs do not rain enough, with low storm counts a secondary factor.

Considering these biases, we look at large-scale atmospheric conditions related to TC intensity in the forecasts. We consider relative humidity (RH) and vertical wind shear (VWS); note that sea surface temperature is prescribed from observations in the forecasts so this is not considered here. Composites of RH at 850 hPa (RH\(_{850}\); Figs. 6a–d) show too much humidity north of 20°N (mainly the extratropics) and too little in the tropical band south of 20°N, where TC frequencies are higher (Fig. 5a). The drift in RH\(_{850}\) with lead time is very slight, but acts to worsen the bias with lead time. RH has a strong control on TCs: midlevel RH is the greatest cause of TC variability due to the MJO (Camargo et al. 2009) and one of the strongest controls on TCs following central Pacific El Niño events (Boucharel et al. 2016). The forecast track density deficit is consistent with the lack of humidity as these biases are approximately collocated. Moreover, there is a low humidity bias over the Philippines and Vietnam, where the precipitation rate per TC is also too low, but a positive humidity bias over China where there is very little bias in the precipitation rate per TC, again consistent with humidity controlling TC behavior.

We define VWS between 850 and 200 hPa (Figs. 6e–h) as

\[
VWS = \sqrt{(u_{200} - u_{850})^2 + (v_{200} - v_{850})^2},
\]

where \(u\) and \(v\) are zonal and meridional wind. VWS is calculated every 6 h and these values are averaged together (as opposed to compositing \(u\) and \(v\) first, then calculating VWS from them). There is substantially less VWS in the forecasts than in ERA-Interim over the equatorial Maritime Continent, but elsewhere the VWS is too large, especially over the tropical west Pacific. However, the largest bias tends to be farther north than the main corridor of TCs, suggesting VWS has less of an impact than RH\(_{850}\) on the track density biases in Fig. 5.

\(c.\) Seasonal cycle

As shown by Guo et al. (2017, their Fig. 2), the strong seasonal cycle in TC activity is reflected in the fractional contribution of TCs to total precipitation, with the highest contributions during July–October. As explained in section 1, we use the same methodology and precipitation observations but with TC tracks from JRA-55 rather than ERA-Interim.

Figure 7 shows the seasonal cycle of TC contribution to precipitation, splitting the year into four seasons: January–March (JFM), April–June (AMJ), July–October (JASO), and November–December (ND). These definitions were chosen by grouping months with broadly similar spatial patterns and magnitudes (not shown). Again, the highest contribution to precipitation from
TCs is during JASO, with over 50% to the northeast of the Philippines and in the South China Sea near Hainan. These results are consistent with the seasonal cycle of Guo et al. (2017; which used the same precipitation product but TC tracks from ERA-Interim), at least on the large scale. The greatest differences are just offshore, where our results tend to show a higher fractional contribution than with ERA-Interim, especially around the Vietnam coast in JFM, and west of Luzon in AMJ and ND. In each case, JRA-55 TCs account for over 50% of precipitation, and ERA-Interim TCs typically only 30%-40%.

**Fig. 6.** (a) Composite of ERA-Interim relative humidity at 850 hPa; (b)–(d) Composites of the bias in forecast relative humidity at 850 hPa, against (a). (e)–(h) As in (a)–(d), but for VWS between 850 and 200 hPa.
The forecast TC contribution to precipitation is too low in all seasons (Fig. 7). The fractional difference from observations is approximately the same, and increases with lead time at approximately the same rate, all year-round. Therefore, there is no evidence for better forecast performance during any particular time of year on the large scale. However, the bias in VWS does have a strong seasonal cycle (Fig. 8). Considering a typical Philippine TC track latitude of 10°–15°N, there is too much VWS during November–June; but over the west Pacific to the east of the Philippines there is generally too little VWS during the peak TC season of JASO, especially at later lead times. Thus, VWS biases vary markedly by season, yet the biases in TC contribution to rainfall do not, again suggesting that VWS is not a significant source of TC rainfall biases in this forecasting system. Conversely, the spatial pattern of RH biases does not vary greatly by season (not shown), consistent with RH being a more important factor than VWS.

d. Improvement in forecast performance over time

The operational forecasts used here cover a period of over 11 years (2006–17). Considerable changes were made to the system during the analysis period. Three horizontal resolutions were used (see Table 1) and a host of other physics, dynamics and assimilation changes were implemented (some of which are summarized in section 2a). To investigate the change in model performance over time, we split the forecasts into three periods based on horizontal resolution for convenience, mindful that the resolution itself is not necessarily the cause of any changes found between these periods.

Due to interannual variability (see section 1), observed TC and total precipitation differ considerably between our chosen time periods. The TC contribution to precipitation is shown for the 60-, 40-, and 25-km periods in Fig. 9, with observed contributions during those periods shown in Figs. 9a, 9e, and 9i, respectively. Over the ocean to the east and west of Luzon, a key region for TC location and rainfall (e.g., Figs. 2 and 3), the observed TC contribution to rainfall reduced over the analysis period. Over 50% of rainfall there came from TCs during the 60-km period but this reduced to 30%–40% more recently. Indeed, in the 25-km period the two “hot-spots” either side of Luzon do not stand out at all, so the major spatial features seen in composites for the entire analysis period are due mainly to TC activity within the first four years (2006–09). Although
previous studies mentioned in section 1 suggested that ENSO has a significant impact on TC interannual variability in this region (e.g., Camargo and Sobel 2005; Corporal-Lodangco et al. 2016), it does not have a clear effect on TC precipitation in these multiyear composites. The strong ENSO events were El Niño (2009/10) and La Niña (2007/08) during the 60-km period; La Niña (2010/11) during the 40-km period; and El Niño (2015/16) during the 25-km period. Over Mindanao the contribution of TCs to rainfall has remained consistently less than 10%, but over Luzon and the Visayan Islands it is greatest in the 25-km period and, unlike in the earlier periods, is not substantially smaller than it is over the surrounding seas. These changes in TC rainfall over time are all consistent with the changes in track density between the three time periods (not shown).

In all three periods the predicted percentage of TC-related rainfall reduces with lead time (Fig. 9), as seen in the overall results (Figs. 4i–l), but the sign of the bias for early lead times changes. In the 60-km period it was strongly negative everywhere, becoming even drier with lead time, especially in the east of the domain shown. However, in the 40- and 25-km periods the bias was positive over most of the South China Sea, especially at early lead times, so even though predicted percentage of TC rainfall still reduces with lead time, by later lead times the dry bias is not as severe overall as in the 60-km period. These changes in bias reflect very closely the changes in bias of the absolute amount of TC-related precipitation (not shown). However, the changes in bias of the total forecast precipitation (not shown) are less simple. In Figs. 4e–h, except for an especially wet bias on day 1, the bias varied only slightly with lead time. However, this pattern is the result of compensating biases in the three periods. In the 60-km period the wet bias dried considerably with lead time, giving a dry bias at late lead times over Luzon, the east coast of the Visayan Islands and Mindanao, and the Pacific Ocean to the west of the Philippines. In the 40-km period the wet bias remained roughly constant with lead time; in the 25-km period the tendency was toward an even wetter bias at later lead times. Hence, the model drift of total precipitation over the tropical west Pacific changed from a drying tendency to a wetting one during the 11-yr analysis period. Since the initial bias is very wet, the 60-km drying tendency restored the model to a more realistic mean state, whereas the more recent wetting tendency caused the model to drift even farther from observations with lead time.

fig. 8. As in Fig. 7, but for VWS: from ERA-Interim in (a), (e), (i), and (m) and with forecast biases presented as a difference from ERA-Interim in the remaining panels.
As in Fig. 5, we now consider the track density (not shown) and precipitation rate per TC (Fig. 10) during each period. As in Figs. 5b–d, the track density, effectively measuring TC lifetime, is generally too low and reduces with lead time, particularly during the 60- and 40-km periods. In the 25-km period the lifetime is slightly too long at very short lead times (e.g., day 2) but at longer lead times is again too short, similar to Fig. 5d. However, the change in precipitation rate per TC over time has been far more stark. Over ocean, in the 60-km period there was a strong dry bias (Figs. 10b–d) but in the 25-km period there was a strong wet bias (Figs. 10j–l). This averages out to the very small bias seen over much of the ocean in Figs. 5f–h. Over the Philippines, however, the precipitation rate per TC was too low in all periods, confirming the conclusion in section 3b that the low TC contribution to precipitation is consistent with both the short TC lifetime and the low precipitation rate within the TC rainbands. Over ocean there has been a shift from too little rain per TC to too much rain. The sign of the bias in lifetime is the same as the bias in TC precipitation contribution throughout the analysis period, suggesting the lifetime has the greatest control on the TC precipitation contribution bias.

As in sections 3b and 3c, we now consider the effect of the large-scale environment on TCs during each of the three periods. The RH850 bias in each period is shown in Fig. 11 for the same lead times as in Fig. 9. In ERA-Interim there is no substantial difference between the different periods, apart from the 40-km period having a slightly higher humidity in the southern half of the Philippines and over the sea to the southeast. In the
forecasts, north of 20°N, where the greatest change in precipitation bias was seen in Fig. 9, there is a considerable change in RH850 between periods, with a switch from a dry bias on the order 2%–3% (60 km; percentage points difference) to a wet bias on the order 4% (25 km) for day 2. South of 20°N the changes are far more slight. For VWS (not shown) there is a slight reduction in the positive bias from the 60- to the 40-km period, and little change thence to the 25-km period. Therefore, as in sections 3b and 3c, we conclude that the changes in RH850 are consistent with the changes in TC rainfall, and there is little evidence of VWS having an effect.

4. Effect of the MJO

The MJO has an impact on west Pacific TC activity (section 1). Here we quantify the effect of MJO phase on TC track density and consider whether the NWP system accurately predicts the MJO effect on TC-related precipitation for forecasts initialized during an MJO event.

Figure 12 shows the effect of each MJO phase on TC track density in JRA-55. The MJO–TC relationship has been shown in previous studies (section 1); here we wish to verify that it is reproduced well by JRA-55. One theory of MJO propagation (e.g., Matthews 2000) is of an eastward-propagating region of diabatic heating, with a Matsuno–Gill-type response (Matsuno 1966; Gill 1980), consisting in part of off-equatorial cyclonic circulation anomalies in each hemisphere, associated with a westward-propagating equatorial Rossby wave. These cyclones may persist for many days. Thus, the MJO is a possible source of west Pacific TC activity. Therefore, in Fig. 12 the track density is composited by MJO phase on the day of TC genesis. Of course, the MJO may progress to a different phase or decay as the TC develops, but here the entire track contributes...
to the composite for the MJO phase at genesis only. As is common for MJO composites, we pair successive phases to increase sample size. Common pairings are phases 2–3, 4–5, 6–7, and 8–1, because they conveniently correspond to the MJO active phase over distinct domains (the first three are the Indian Ocean, Maritime Continent, and western Pacific Ocean, respectively, and the final pairing is the transition from one MJO convective envelope to the next, with convection dying out near the date line and re-emerging over Africa). However, since there is typically a lag of a few days between a TC’s genesis and it contributing substantially to precipitation, following Klotzbach and Oliver (2015) we use the pairings 1–2, 3–4, 5–6, and 7–8.

The track density of TCs beginning in phases 5–6 is considerably higher than other phases, with high density
over Luzon, and the seas to the east and west. During these phases the MJO is active over the eastern Maritime Continent and west Pacific Ocean, consistent with TCs forming there and propagating westward as an equatorial Rossby wave. Track density is also anomalously high over Mindanao and the seas either side for TCs beginning in phases 3–4, albeit to a lesser extent and focused farther west. Again this is consistent with TCs forming in an active MJO region, since the MJO is active over the east Indian Ocean and west-to-central Maritime Continent in these phases. For these phases, however, the track density is anomalously low over most of the west Pacific domain shown. When the MJO is suppressed over the Maritime Continent, the track density is anomalously low almost everywhere (phases 1–2 and 7–8). These results are consistent with Vitart (2009), who produced similar composites using track data from the Joint Typhoon Warning Center (JTWC).

The fractional TC contributions to precipitation are shown for each pair of MJO phases in Figs. 13a–d (JRA-55 and TRMM) and Figs. 13e–h (NWP), as a fractional difference from the climatological value. NWP precipitation was calculated separately for each forecast, compositing over all lead times, then composited by the MJO phase at the initialization time. For JRA-55 the same approach was used, averaging over the same 6- or 7-day window as for each forecast. Note that the precipitation itself did not necessarily fall during the stated MJO phases as the MJO may have propagated to a different phase or decayed away since the initialization time. Note also that these composites do not correspond exactly to those in Fig. 12, which were based on MJO phase on the TC genesis date. The forecasts cannot be composited by genesis date as TCs usually already exist in the forecast initial conditions.

Despite the difference in the compositing method for Figs. 12 and 13, phases 5–6 again show widespread positive anomalies for observations. Over almost all the domain, a much higher percentage of precipitation comes from TCs during forecast windows initialized in those MJO phases than in the climatology. In the opposite phases (1–2), the percentage is much lower almost everywhere in the domain, while the remaining phases have a more mixed picture. For example, for phases 7–8 TCs provide a slightly higher fraction of rainfall than average over the sea just west of the Philippines, but slightly less immediately to the east.

The corresponding MJO anomalies for NWP are remarkably similar to the observations, on both the domain-wide and small scales. Therefore, although the Met Office NWP system underpredicts TCs rainfall (Fig. 4), due mainly to a short TC lifetime and, over land, a lack of precipitation falling within each TC (Fig. 5), the MJO phase at forecast initialization affects forecast TC precipitation in a very accurate way.

5. Discussion and conclusions

This study is motivated by the need for accurate forecasts of TCs and their impacts over and around the Philippines. We analyze the ability of the UKMO global operational NWP system to forecast the contribution of all TCs to precipitation, the first such study with a long-term dataset of forecasts. We verify forecasts against TRMM precipitation observations and TC tracks from JRA-55.

In JRA-55, climatological TC track density has a meridional gradient around the Philippines, particularly over land, with more than 8 TCs yr$^{-1}$ per 2.5° grid box over the ocean near Luzon (north Philippines) and 6–8 yr$^{-1}$ per box over land; near Mindanao (south Philippines) there
are about 8 yr\(^{-1}\) per box over ocean and about 2–4 yr\(^{-1}\) per box over land. The fractional contribution of TCs to precipitation is also greater over ocean than land, and greater to the north than the south. Over ocean, TCs contribute 30%–40% of precipitation around Luzon, but only 5%–15% around Mindanao; the corresponding figures over land are 20%–30% and 5%. These results confirm the findings of Khouakhi et al. (2017), Guo et al. (2017) and Bagtasa (2017). TC precipitation is poorly spatially correlated with mean TC maximum 10-m wind speed in JRA-55. This is consistent with Jiang et al. (2008), who found that although maximum wind speed and precipitation rate have a reasonably strong correlation close to the TC center, peaking at about 0.6 within 1.5° of the center, this falls rapidly to about 0 at a radius of approximately 5° (for TCs over ocean).

In the UKMO NWP model, TC precipitation is realistically predicted at 1-day lead time, except over Luzon and the east Visayan Islands, where there is a 1 mm day\(^{-1}\) dry bias. However, TC precipitation reduces with lead time and, owing to total forecast precipitation being too high, TCs contribute too little to total precipitation, decreasing further with lead time (Fig. 4). This bias exists year-round (Fig. 7), but NWP system upgrades during our analysis period improved the bias. During the period when the forecast model used a 60-km horizontal resolution (up to March 2010), TCs contributed too little precipitation at all lead times (Fig. 9). During this period, a TC initialization scheme was used (section 2a), which was adjusted in November 2007 for spatially small TCs. It was switched off in July 2012 as it no longer improved TC forecasts globally, although Heming (2016) reported that this may have degraded TC forecasts in the western North Pacific. This switch off occurred about halfway through the 40-km resolution period (March 2010–July 2014), during which forecsts marginally improved over the 60-km period. The bias in TC contribution to precipitation was considerably reduced again in the 25-km period (July 2014–July 2017). This period used the GA6.1 configuration; a new TC initialization scheme was introduced in February 2015. Both improved forecasts of TC location and intensity (Heming 2016); our evidence also indicates improved forecasts of TC-related precipitation.

TCs were matched between IBTrACS, JRA-55, and the forecasts, so we have the same TCs in each dataset, although they may not have the same lifetime. Therefore, TC precipitation biases may be due to TCs decaying too rapidly, or errors in the precipitation amount within each TC. The track density has only small biases at short lead times (Figs. 5b–d), implying the forecast TC lifetimes tend to be too short; this changed very little over the analysis period.

Over ocean, the mean precipitation rate per TC has fairly small (especially at short lead time) and spatially incoherent biases when averaged over the whole analysis period (Figs. 5f–h). These biases have changed substantially over time, with a very negative bias in the 60-km period but a very positive bias during the 25-km period. Biases in TC contribution of precipitation have remained negative throughout, which suggests that over ocean it is the TC lifetime bias which more strongly controls the precipitation bias. Over the Philippines there is a strong negative bias in precipitation rate per TC, which has changed little during the analysis period, so there both the lifetime bias and the bias in precipitation within the TC rainbands may be important.

The large-scale environment was investigated to understand possible reasons for the biases in lifetime and precipitation rate per TC. Composites of RH\(_{850}\) and VWS show large biases in each (Fig. 6), with a negative bias in RH\(_{850}\) over the main TC storm track, consistent with the short TC lifetime. There is a positive bias in VWS, which would also be consistent with short-lived TCs, except that it is displaced to the north of the main TC development region. Furthermore, while the RH\(_{850}\) bias persists year-round, the VWS bias is much stronger in November–March than other months (Fig. 8). The precipitation biases do not vary substantially by season (Fig. 7) which suggests the VWS bias is unimportant. Further evidence for the RH\(_{850}\) bias being more important than the VWS bias for the precipitation bias is that the RH\(_{850}\) bias improved in the 25-km period compared with earlier periods (Fig. 11), in line with the precipitation bias, whereas the VWS bias improved little.

The effect of the MJO in JRA-55 is to increase TC genesis during active phases and reduce it during suppressed phases (Fig. 12). This is consistent with many previous studies, as detailed in section 1. The TC fraction of precipitation in JRA-55 is also higher within forecast windows which begin during active MJO phases in the west Pacific, and is less during suppressed phases (Figs. 13a–d). This effect is well reproduced in the forecasts in terms of the change relative to the model’s own climatology (Figs. 13e–h). Hence, even though the model climatology contains significant biases, performance improves for forecasts initialized during an MJO event.

The techniques used in this paper are equally applicable to other models and forecasts. Short and Petch (2018) analyzed TCs in regional convection-permitting forecasts using the MetUM, including the spatial patterns of extreme rainfall and the radial distribution of precipitation within TCs. They showed that convection-permitting forecasts predict a more accurate distribution.
of precipitation within TC rainbands than the global forecasts we analyzed. Our method of compositing TC contribution to total precipitation could be used to complement metrics such as those used by Short and Petch (2018).

Overall, the Met Office global operational NWP system up to July 2017 shows considerable biases in terms of TCs’ contribution to precipitation. These biases exist year-round. Over ocean, they are mainly due to the TC lifetime being too short. Over the Philippines, where there is a clear socioeconomic need for forecasts which can predict the impact of TCs accurately, there is the additional problem of the TC rainbands producing too little rainfall, although this error is not seen consistently over ocean. This may be due to poorly resolved land–sea contrasts or orography, even at 25-km grid spacing, as orographic effects have been observed to impact TC precipitation (Tan et al. 2012; Racoma et al. 2016). However, TRMM also has errors over complex orography (Rahmawati and Lubczynski 2018) so there may also be observation errors over the Philippine islands.

Improvements in the system over time are encouraging, but biases remain considerable. Our results suggest that TC precipitation biases do not vary with the total precipitation biases in the forecasts, but TC precipitation biases could be improved by further improving RH biases. The response of the NWP system to a strong MJO in the initial conditions, relative to its own climatology, is also encouraging. Models in the S2S database generally have MJO skill out to 12–36 days (Lim et al. 2018), so it may be possible to exploit MJO predictability to make TC precipitation forecasts well beyond the normal limits of NWP.

Acknowledgments. SCP and KIH were supported by the Met Office Weather and Climate Science for Service Partnership (WCSSP) Southeast Asia, as part of the Newton Fund. NPK was supported by the U.K. Natural Environment Research Council (NE/L010976/1). This work is part of the Forecasting Air–Sea Coupled Interactions in NWP for Atmospheric Tropical Extremes (FASCINATE) project. TRMM 3B42 data are available from the NASA Goddard Space Flight Center at https://pmm.nasa.gov/trmm. JRA-55 data are available from the Japan Meteorological Agency at https://jra.kishou.go.jp/JRA-55. ERA-Interim data are available from the European Centre for Medium-Range Weather Forecasting at https://apps.ecmwf.int/datasets/data/interim-full-daily. IBTrACS data are available from the National Climatic Data Center at https://www.ncdc.noaa.gov/ibtracs. The authors are grateful to the Met Office for providing access to their operational forecast archive. We thank two anonymous reviewers, whose suggestions improved the manuscript.

REFERENCES


