RESEARCH ARTICLE

The effect of atmosphere–ocean coupling on the prediction of 2016 western North Pacific tropical cyclones

Xiangbo Feng1 | Nicholas P. Klingaman1 | Kevin I. Hodges2

1National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK
2Department of Meteorology, University of Reading, Reading, UK

Correspondence
Xiangbo Feng, National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK.
Email: xiangbo.feng@reading.ac.uk

Funding information
Natural Environment Research Council, NE/L010976/1; WCSSP Southeast Asia

Abstract
We examine the merit of atmosphere–ocean coupled models for tropical cyclone (TC) predictions in the western North Pacific (WNP), where accurate TC predictions remain challenging. The UK Met Office operational atmospheric global numerical weather prediction (NWP) model is compared with two trial coupled configurations, in which the operational atmospheric model is coupled to a one-dimensional mixed-layer ocean model and a three-dimensional dynamical ocean model. Reforecasts for the 2016 TC season show that the coupled models outperform the NWP model for TC location predictions, with a systematic improvement of 50–100 km over the seven-day forecasts, but the coupled models amplify the underestimation of TC intensity in the NWP model. Nearly identical TC predictions (for both location and intensity) are found in the two coupled models, indicating the dominance of thermodynamic processes at the air–sea interface for TC predictions on these timescales. The improved prediction of the TC position in the coupled models is associated with an enhanced Western North Pacific Subtropical High (WNPSH), which introduces an anticyclonic steering flow anomaly that shifts TC tracks further west in the southern part of the region and further east in the northern part. Based on sensitivity experiments, we show that these improvements in the coupled models are due mainly to colder initial sea-surface temperatures (SSTs). Air–sea feedbacks do not change the WNPSH or TC tracks noticeably. Apart from the effect of the initial SSTs, tropical ocean warming due to air–sea interaction in the coupled forecasts can also reduce the predicted TC intensity, presumably due to a stronger regional Hadley circulation with increased subtropical subsidence.

KEYWORDS
air–sea coupling, tropical cyclone, weather forecast, western North Pacific

1 INTRODUCTION

Tropical cyclones (TCs) are among the most hazardous weather phenomena on the planet. The western North Pacific (WNP) is the most active area for TC activity. The tremendous damage caused by TCs arises from their strong winds, heavy rainfall, associated storm surge, and large waves, with other indirect effects such as saltwater intrusions, coastline erosion, and ecosystem recessions (e.g., Feng et al., 2012; Feng and Tsimplis, 2014; Guo et al., 2017). In the research...
community and operational weather forecasting centres, considerable efforts have been made to deliver operationally reliable predictions of TCs, using statistical, dynamical, or hybrid approaches, at forecast timescales of synoptic to seasonal (e.g., Goerss, 2000; Chan et al., 2001; LaRow et al., 2010; Vecchi et al., 2011; Camp et al., 2015; Yamaguchi et al., 2015; Mogensen et al., 2017).

Global numerical weather prediction (NWP) models, the weather-timescale general circulation models (GCMs), are widely used as operational tools to predict TC activity. TC track prediction in NWP models has been substantially improved over the past decades, due to improved simulation of the large-scale environment through better and higher-resolution models and better model initializations based on more advanced data assimilation schemes (Heming, 2016). However, large errors remain in TC predictions by global NWP models (Hodges and Emerton, 2015; Heming, 2016; Yamaguchi et al., 2017; Hodges and Klingaman, 2019, submitted, personal communication). For example, in the west Pacific, the state-of-the-art UK Met Office and European Centre for Medium-Range Weather Forecasts (ECMWF) global NWP systems have mean location errors that grow from 50 to 400 km during the five-day forecast (Yamaguchi et al., 2017). Global NWP models consistently underestimate TC intensity, mainly due to coarse resolution (e.g., Short and Petch, 2018; Hodges and Klingaman, 2019, submitted, personal communication). However, there has been only limited improvement in the ability of global models to predict TC intensity over the past decades (DeMaria et al., 2014; Yamaguchi et al., 2017). Refining model resolution and resolving TC-related physical processes explicitly are among the main efforts to advance TC prediction skill further in global NWP models. For the latter, implementing two-way air–sea coupling in NWP models is seen as important (Takaya et al., 2010; Mogensen et al., 2017), as air–sea coupled models can capture the complicated flux exchanges at the air–sea interface more realistically.

Atmosphere–ocean coupled GCMs were first used for TC simulations in the 1970s (e.g., Manabe et al., 1970; Elsberry et al., 1976; Chang and Anthes, 1979). Air–sea coupling is one of the key factors contributing to the prediction skill of subseasonal to seasonal forecast systems, related to improved skill in capturing the sea-surface temperature (SST) variations that are important for predicting intraseasonal climate variability, such as the Madden–Julian oscillation (MJO: Vitart et al., 2017). Following the same concept, coupled models have been investigated for operational synoptic weather forecasts (usually <10 days: Bender and Ginis, 2000; Yablonsky and Ginis, 2009; Wu et al., 2016; Mogensen et al., 2017; Pianeze et al., 2018). In June 2018, ECMWF launched its first operational atmosphere–ocean coupled NWP model for global weather forecasts (Buizza et al., 2018).

Most studies show that ocean coupling reduces TC intensity and intensification, especially for slow-moving TCs that cause strong SST cooling. Model simulations also show that the effects of air–sea coupling on TC predictions are larger in higher horizontal resolution models (e.g., Bender et al., 1993; Vitart and Stockdale, 2001). Mogensen et al. (2017) further confirmed that, in the ECMWF coupled NWP model, the coupling effect is maximized wherever the ocean mixed layer is shallow with strong ocean stratification: for example, in the mid-latitudes. These studies mostly focused on TC case studies or used higher-resolution regional models. Meanwhile, there are only limited studies on the effect of ocean coupling on TC track predictions. With increased computing capability, many weather service centres may be able to run coupled NWP models for weather forecasts. Thus, it is necessary to assess the performance of coupled models rigorously for TC forecasts, relative to their counterpart atmosphere-only NWP models.

In coupled models, air–sea coupling can alter TC predictions through different effects. A passing TC induces a cold wake in the upper ocean, due to the upwelling of cooler subsurface water and increased evaporation, forced by the strong surface wind (Chen et al., 2007). TCs can also cool the ocean surface through increased cloud cover with reduced downward shortwave radiation, similar to other atmospheric convective events such as equatorial waves and the MJO (e.g., Feng et al., 2018). These negative feedbacks change the local thermodynamic structure of the TCs through reduced upward surface heat fluxes, to discourage rapid TC intensification. This damping effect is strongest in regions where the ocean mixed layer is shallow (Mogensen et al., 2017). On the other hand, air–sea interactions in remote regions can change the large-scale atmospheric environment in the TC development region through teleconnections (e.g., Wu et al., 2010; Takaya et al., 2017). This effect, which has been demonstrated in longer forecasts, could also alter shorter-scale TC predictions for both location and intensity. To identify the relative roles of local and remote air–sea interactions for TC predictions, sensitivity experiments using global NWP models are required.

This article investigates the added value of air–sea coupling in the UK Met Office operational global NWP model for predicting TCs, with a focus on the WNP basin. The WNP basin is chosen because it is an area where NWP models have large TC prediction errors (Hodges and Emerton, 2015; Heming, 2016; Yamaguchi et al., 2017). The UK Met Office operational global atmosphere-only NWP model is compared with two newly developed coupled models for seven-day forecasts of the 2016 WNP TC season. To minimize the impacts of differences in model resolution and physics, all configurations use the same atmospheric model version and resolution. We investigate whether the coupled models outperform the atmosphere-only NWP model for TC predictions in the WNP basin, and the causes of differences between the coupled and
uncoupled TC predictions, to indicate where model development is required to improve WNP TC predictions.

The article is structured as follows. In section 2, the UK Met Office operational global atmosphere-only NWP model and the counterpart coupled models used for reforecasts of the 2016 TC season are described, along with other data and analysis methods used. In section 3, TC position errors in the NWP model are diagnosed and linked with errors in the large-scale environment. These errors are compared with prediction errors in the coupled models. In section 4, the TC intensity predictions in the NWP and coupled models are characterized and compared. Possible reasons for changes in TC predictions (for both position and intensity) between models are also proposed and analysed, from the perspective of air–sea interaction. In section 5, sensitivity experiments for a short period (15 days) are run and analysed, to distinguish the effects of air–sea coupling on TC predictions further. Finally, conclusions and discussion are provided in section 6.

2 | DATA AND METHODOLOGY

2.1 | Forecast models

We compare forecasts and reforecasts of the 2016 WNP TC season from three configurations of the Met Office atmospheric model. All configurations use the Global Atmosphere 6.1 (GA6.1) scientific configuration of the Met Office atmospheric model (Walters et al., 2017), at N768 horizontal resolution (approximately 0.23° × 0.16°). The first configuration is the operational NWP model (hereafter “NWP”), in which GA6.1 is forced with observed SSTs and sea ice from the Operational Sea-surface Temperature and sea Ice Analysis (OSTIA: Donlon et al., 2012). The initial SST and ice are persisted over the seven-day forecast.

The second configuration is a trial coupled NWP configuration, in which GA6.1 is coupled to the Nucleus for European Modelling of the Oceans (NEMO) dynamical ocean model every hour (hereafter “NEMO coupled”). NEMO uses a tripolar horizontal grid at approximately 0.25° resolution, referred to as ORCA025, with 75 vertical levels and with a top level approximately 1 m thick. This configuration mimics the Global Coupled 2.0 configuration of the Met Office coupled model (Williams et al., 2015) used for the seasonal forecast. GA6.1 and NEMO are also coupled every hour to the Los Alamos sea-ice model (CICE), which predicts sea-ice extent and depth. All three models are coupled via the Ocean Atmosphere Sea Ice Soil (OASIS: Craig et al., 2017) coupler. GA6.1 passes net surface heat, moisture, and momentum fluxes to NEMO, which returns SST and surface currents to GA6.1.

In the third configuration, GA6.1 is coupled every hour via OASIS to the Many Column configuration of the K Profile Parameterisation (MC-KPP) one-dimensional ocean mixed-layer model. Hereafter, we refer to this model as “KPP coupled.” MC-KPP is based on the KPP mixing scheme of Large et al. (1994), with enhancements to run as a matrix of columns and couple to an atmospheric model. Each GA6.1 column is coupled to one KPP column, such that the effective MC-KPP horizontal resolution is the same as the GA6.1 resolution (approximately 0.23° × 0.16°). MC-KPP is configured with 100 points in a 1,000 m column depth, with 70 points in the top 300 m and a top layer of approximately 1 m. This configuration mimics the Global Ocean Mixed Layer configuration of the Met Office model (Hirons et al., 2015). GA6.1 and MC-KPP exchange the same fields as GA6.1 and NEMO, except that for MC-KPP the surface currents are set to zero as the model lacks ocean dynamics. For this study, MC-KPP was extended to include a representation of vertical Ekman transports, which are an important component of TC-induced SST cooling (e.g., Vincent et al., 2012). Following Lu et al. (2017), the Ekman vertical velocity (w_e) is

\[
    w_e = \frac{1}{\rho_w f} \nabla \times \tilde{\tau},
\]

where \( \rho_w \) is the density of seawater (kg/m³), \( w \) is the Coriolis parameter (s⁻¹) and \( \tau \) is the surface wind stress (N/m²). We assume that the Ekman vertical velocity reaches its maximum at the Ekman depth \( d_{ek} \):

\[
    d_{ek} = 0.4 \sqrt{\tau \rho_w f^2},
\]

and decays sinusoidally to the surface and to a subjectively chosen depth to represent the pycnocline (300 m).

Otherwise, MC-KPP simulates only vertical mixing; there is no horizontal advection. It is important to note that the KPP coupled model does not include any flux correction, as we expect drift from the lack of ocean dynamics to be small over the forecast period. As MC-KPP lacks a sea ice module, the initial sea ice is persisted over the forecast period as in the NWP model. Due to computational limitations, MC-KPP is coupled to GA6.1 only across 25°S–35°N, at all longitudes. Outside the coupled region, the initial SST is persisted as in the NWP model. At the boundary of the coupled and uncoupled regions, the coupled and persisted SSTs are blended across a region of approximately 7° (45 grid points), with linearly decreasing weight given to the coupled SST further away from the coupled region.

The NEMO and KPP coupled models differ in (a) their treatment of ocean dynamics, (b) their vertical mixing scheme (NEMO uses the Turbulent Kinetic Energy mixing scheme), (c) the coupling region, (d) their vertical resolutions, and (e) the treatment of sea ice.

Forecasts or reforecasts of the 2016 WNP TC season, with six-hourly output, were obtained or performed, respectively, with each model configuration running for the period May
FIGURE 1  (a) TC tracks at six-hourly intervals from IBTrACS, observed during May 15–September 30, 2016, coloured by intensity categorized on the Saffir–Simpson Hurricane Wind Scale. (b) TC prediction sample sizes binned into $2^\circ \times 2^\circ$ grid boxes, using all TC track predictions in the NWP forecasts, after matching with the observed TC tracks in (a).

15–September 30, 2016, inclusive. For the NWP model, these are the operational seven-day high-resolution deterministic forecasts. For the NEMO and KPP coupled models, these are 15-day deterministic reforecasts. The NWP forecasts and KPP coupled reforecasts are initialized twice a day (0000 and 1200 UTC), while the NEMO coupled reforecasts are initialized only once a day (0000 UTC). Here, we analyse only the common forecast set: the first seven days of each 0000 UTC forecast. Atmospheric initial conditions for all forecasts come from the Met Office operational analyses. Oceanic initial conditions for the NEMO and KPP coupled reforecasts come from the analysis of the operational Met Office Forecast Ocean Assimilation Model (FOAM) system (Martin et al., 2007; Blockley et al., 2014; Waters et al., 2015), which is interpolated to the ocean grid of each model using a conservative remapping scheme. The main difference between the FOAM and OSTIA products is that the former is produced by assimilating various ocean observations (e.g., in situ and satellite SST, satellite altimeter sea-level anomalies, and ocean temperature and salinity profiles) into a three-dimensional ocean model, while the latter is produced from only satellite SST observations, without a background ocean model.

2.2 Observation and reanalysis data

The International Best Track Archive for Climate Stewardship (IBTrACS: Knapp et al., 2010) dataset is used to validate TC predictions. IBTrACS is an analysed global TC record created from TC observations from many TC operational agencies, with certain limitations. For instance, there are noticeable discrepancies in TC frequency and intensity in the WNP basin between agencies, due to applying different TC observational algorithms (Ren et al., 2011; Barcikowska et al., 2012). Barcikowska et al. (2012) showed that, after 1998, the annual mean difference in TC position between data providers is below 30 km, which is close to the spatial resolution of our models. In IBTrACS, TC intensity (i.e., measured by maximum 10-m wind speed or minimum mean sea-level pressure, MSLP) observations typically contain missing values, especially for maximum 10-m wind speed, as intensity is recorded only when it exceeds certain criteria. This article mainly investigates the difference in TC predictions (both position and intensity) between models, which reduces the effect of observational uncertainty. During May15–September 30, 2016, IBTrACS contains 24 TCs in the WNP basin ($0^\circ$–$60^\circ$N, $100^\circ$–$180^\circ$E), with a range of intensities (Figure 1a).

ERA5 reanalysis (the successor to ERA-Interim; Hersbach and Dee, 2016) is used to verify the predicted atmospheric fields. ERA5 is the fifth generation of ECMWF atmospheric reanalyses of the global climate. Innovations in ERA5, relative to ERA-Interim (Dee et al., 2011), include higher resolution and improvements in the numerical model, data assimilation method, observations, and external forcing. SSTs in ERA5 are derived from Hadley Centre Ice and Sea Surface Temperature (HadISST2) data prior to 2007 and from OSTIA thereafter (Hirahara et al., 2016). ERA5 is a 10-member ensemble of reanalyses together with a single higher resolution reanalysis. We use the high-resolution reanalysis (31 km globally).
2.3 | Methods

2.3.1 | TC tracking and matching

TCs are identified and tracked in the seven-day forecasts at six-hourly intervals using the same methods as used in Hodges and Emerton (2015); a brief summary follows. The basic tracking uses the scheme described in Hodges (1994, 1995, 1999). First, the vertical average of the relative vorticity between 850 and 600 hPa is obtained. This is then spatially filtered using spherical harmonics to T63 resolution; the large-scale background with total wave numbers $n \leq 5$ is removed. Vorticity maxima (in the Northern Hemisphere) are determined on the T63 grid and then used as starting points to obtain the off-grid locations using B-spline interpolation and maximization methods (Hodges, 1995). In the first instance, all positive vorticity centres that exceed a threshold of $0.5 \times 10^{-5} \text{s}^{-1}$ in the range $0^\circ–60^\circ\text{N}$ are tracked. This approach produces the most coherent tracks, and the full lifecycles of the systems including their pre-TC and post-TC stages. The tracking is performed over the full length of the seven-day forecasts.

After tracking, other variables are added to the tracks, including the maximum 10-m wind speed and minimum MSLP. This is done by searching for the maximum 10-m winds within a $6^\circ$ geodesic radius around the TC track and for the true minimum MSLP within a $5^\circ$ radius using the B-splines and minimization method. In this article, the precipitation averaged over an area within $5^\circ$ geodesic radius around the TC track is taken as the TC-related precipitation.

To identify the TCs from amongst all tracked features, a matching method (Hodges and Emerton, 2015) is used to match the forecast tracks against the verification tracks (IBTrACS). A forecast track is matched to a verification track if the mean spatial separation of its first four points (1 day) is $\leq 4^\circ$ and it is the track with the smallest separation for the four points. Only forecast tracks that have their first point within the first three days of the forecast are considered, to exclude matches by chance. Our tracking and matching approaches are similar to the Met Office TC verification techniques (Heming, 2017), but ours use the smoothed vertical average of vorticity. Our methods also provide a larger sample size of TC tracks, which is critical for our composite analysis, because we track the TC in any forecasts before the observed storm forms and after the observed storm decays.

2.3.2 | TC verification and comparison

On completion of the TC tracking and matching, there are 154 TC tracks commonly identified amongst the coupled and uncoupled forecasts in the WNP basin ($0^\circ–60^\circ\text{N}, 100^\circ–180^\circ\text{E}$). The number of forecast tracks (154) is much greater than the number of observed tracks (24), as the same observed storm often appears in several forecasts. The TC position and intensity, at six-hourly intervals, of these tracks are then verified against IBTrACS to obtain the TC prediction errors. The errors are then binned and averaged into $2^\circ \times 2^\circ$ grid boxes, which are based on the TC positions produced by the NWP model. As well as computing prediction errors, the TC predictions (position and intensity) are also compared between the NWP and coupled models. The sample size of the TC tracks is highest in the Philippine Sea and to the south of Japan (>15 tracks; Figure 1b).

2.3.3 | WNPSH index

The Western North Pacific Subtropical High (WNPSH) is a strong modulator of TC tracks in the WNP basin (Wang et al., 2013; Camp et al., 2019, and references therein). We analyse the TC position errors with associated errors in the WNPSH. Similar to Wang et al. (2013), we take the average of 500-hPa GPH over the region $20^\circ–40^\circ\text{N}$ and $120^\circ–140^\circ\text{E}$ as an index to measure the westward extent of the WNPSH.

3 | CHANGES IN TC TRACK, AND THEIR ASSOCIATIONS WITH AIR–SEA COUPLING

3.1 | TC position errors

3.1.1 | NWP model

Compared with observations, the NWP model predicts TCs with an average location error ranging from $0.5^\circ–5^\circ$ for the six-hourly positions in the seven-day forecasts for May–September 2016 (Figure 2a). The regional mean TC location error is about $2.7^\circ$, which is consistent with the seven-day average TC prediction error in the Met Office global NWP model during 2008–2017 for this region (Hodges and Klingaman, 2019, submitted, personal communication). The seven-day mean error is relatively smaller in the south (below $20^\circ\text{N}$) and larger in the north (above $20^\circ\text{N}$). The larger error in the north is mainly related to recurving TCs (Figure 1a), which have a faster translation speed after they recurve to the north or when they approach land, which the NWP model struggles to capture.

The NWP TCs generally track $0.5^\circ–2^\circ$ too far east in the southern part of the basin, whilst they track too far west, with the same magnitude bias, in the northern part of the basin, relative to observations (Figure 3a). This means that the predicted TCs move more slowly than expected in the southern part, and that they recurve too late to the north. In the meantime, the predicted TCs are generally $1^\circ–3^\circ$ further south than observed (Figure 3b), again indicating that in the south, translation speeds are too slow for both straight-moving TCs and recurving TCs, while in the north, recurving TCs recurve too late.
In the South China Sea, where some TCs form and travel northwest (Figure 1a), the NWP model predicts tracks too far southeast (Figure 3a and b), with the same biases as for northwest-moving TCs that originate in the open ocean, but with larger amplitudes. East of Taiwan, the NWP model has a 1°–2° northward bias, which differs from the overall southward bias in neighbouring areas (i.e., the South China Sea and southeast of Japan). In general, the NWP forecast has a southeast bias in the southern part of the WNP basin and a southwest bias in the northern part, with respect to observations.

3.1.2 Coupled models

In the seven-day forecasts of the two coupled models, the TC location errors are systematically reduced (Figure 2b and c), by about 0.5°–1.0° (50–100 km) on average. The largest reduction occurs in the Luzon Strait, where the sample size of TC predictions is largest (for sample size, please see Figure 1b). These improvements in TC track predictions are nearly identical between the KPP and NEMO coupled models. This suggests that the vast majority of the improvement over the NWP model is related to thermodynamic air–sea feedbacks; ocean dynamics, which are not represented in MC-KPP, beyond the idealized vertical Ekman transport (see section 2.1), have little impact on the seven-day TC predictions. The TC position error near the Japanese coasts increases in the coupled models (Figure 2b and c), but results are less robust here due to the limited sample size of TC predictions (Figure 1b).

The improved TC track predictions in the coupled models are dominated by a longitudinal shift of TC tracks (Figure 3c–e). The tracks of the coupled models are further west in the south and further east in the north, relative to the NWP tracks, both of which reduce the NWP biases (Figure 3a). In contrast, TC latitudes are better predicted only east of Taiwan and in the area around the Philippines (Figure 3d–f), where the NWP-predicted TCs are biased northward (Figure 3b). In the rest of the region, there is no consistent change in TC latitudes. The NEMO and KPP
FIGURE 3  (a) Longitudinal and (b) latitudinal errors of TC predictions in the NWP model against IBTrACS. (c) Longitudinal and (d) latitudinal changes of the TC predictions in the KPP coupled model over the NWP model. (e) Longitudinal and (f) latitudinal changes of the TC predictions in the NEMO coupled model over the NWP model. Forecast period is May 15–September 30, 2016. TC errors are first calculated from all seven-day forecasts and then binned into $2^\circ \times 2^\circ$ grid boxes as in Figure 1b. Only the boxes with TC prediction sample size $\geq 4$ in Figure 1b are shown.

coupled models predict TC tracks with similar longitudinal and latitudinal changes, relative to the NWP model, although the NEMO coupled model produces slightly larger amplitudes of change in the northern part of the region (Figure 3e–f).

We now examine the association between TC position errors and WNPSH errors. For the mean of all forecasts with TCs identified, the NWP model predicts the 5,880-m contour of 500-hPa GPH further east than in ERA5, indicating a weak WNPSH in the model (Figure 2). In contrast, the coupled models predict the western edge of the WNPSH better. To understand the effect of the stronger WNPSH in the coupled models on TC predictions, we composite those TC forecasts in which the seven-day average WNPSH index in the KPP coupled model is $\geq 3$ m larger than in the NWP model. There are 45 such forecasts (out of 139), with 67 common TC tracks. In these 45 forecasts, TC track predictions in the
FIGURE 4  (a) Difference of TC position errors between the KPP coupled model and the NWP model, averaged only for the seven-day TC forecasts in which the KPP-predicted WNPSH is $> 3$ m higher than the NWP-predicted WNPSH, during May 15–September 30, 2016. (b) Differences of the TC longitude and (c) latitude between the KPP coupled model and the NWP model, corresponding to (a). TC errors are binned into $2^\circ \times 2^\circ$ grid boxes as in Figure 1b. Only the values where the TC prediction sample size $\geq 4$ in Figure 1b are plotted.

coupled model are improved nearly everywhere in the southwestern part of the region, with the largest error reduction of $1.0^\circ$–$1.5^\circ$ around Taiwan (Figure 4a). Such an improvement is, again, clearly associated with a southwest shift of TC tracks in the KPP coupled model (Figure 4b and c). Similar results are found when comparing the NEMO coupled and NWP models. Thus, there is a close association between the prediction skill of the TC position and the simulation of the WNPSH index.

Because the KPP and NEMO coupled models give nearly identical results for TC predictions, for the rest of this article we focus on comparing the KPP coupled and NWP models.

3.2 Association of TC track changes with environmental field changes

In forecasts with TCs, the WNPSH index in the NWP model is consistently too weak relative to ERA5 (Figure 5a). The weakened WNPSH is accompanied by a cyclonic wind anomaly centred around $20^\circ$–$30^\circ$N and $120^\circ$–$140^\circ$E at 500 hPa. This anomaly is associated with a westerly anomaly in the south and an easterly anomaly in the north. The results are similar for winds at other pressure levels. The cyclonic anomaly tends to steer TCs in the southern part of the region to the east and, consequently, delay them recurving to the north (Figure 3a and b). In contrast, the KPP coupled model produces an enhanced WNPSH relative to the NWP model (Figure 5b). The largest enhancements ($>3$ m) appear around $20^\circ$–$30^\circ$N and $120^\circ$–$140^\circ$E as well. The enhanced WNPSH is associated with an anticyclonic wind anomaly, which favours TCs in the southern part of the region to move southwest and those in the northern part of the region to move east (Figure 3c and d).

For the 45 forecasts in which the seven-day average WNPSH in the KPP coupled model is $>3$ m larger than in the NWP model, the anticyclonic circulation anomaly produced by the coupled model (Figure 5d) is stronger than that in all forecasts with TCs (Figure 5b). These changes compensate for the deficiencies of a suppressed WNPSH and associated cyclonic circulation bias in the NWP model (Figure 5c), leading to greater improvements of TC position predictions (Figure 4). For the forecasts in which there are no TCs identified (mostly May–June), the NWP model still has similar WNPSH and circulation biases (Figure 5e). In these forecasts, the easterly wind anomaly produced by the KPP coupled model relative to the NWP model remains in the central part.
of the region (10°–20°N and 120°–140°N), but with reduced amplitude (Figure 5f). Thus, the coupled model predicts an enhanced WNPSH in most forecasts, not only those with TCs, with the enhanced WNPSH appearing to contribute to improved TC track predictions.

There are two additional points worth noting. The first is that the anticyclonic wind anomaly between the KPP and NWP models (Figure 5b and d) is asymmetric, with relatively larger easterly anomalies in the south and smaller westerly anomalies in the north, corresponding to the greater improvements in the prediction of TC longitudes in the southern part of the region (Figures 3c and 4b). This suggests that the easterly wind anomalies may not be related solely to the enhanced WNPSH. Other relevant factors will be explored later.

The second point is that the changes produced by the coupled model reduce, but do not remove, the biases of TC tracks and environmental fields in the NWP model. This suggests that other factors, not related to coupling, in the NWP model also result in TC track errors. The reasons for the WNPSH bias in the NWP model are likely complex. It may be caused by deficiencies in the parametrization of unresolved physical processes, such as convection, or biases in boundary conditions, such as the SSTs. The investigation is not

**Figure 5**  (a) Difference of SST (background), wind (arrow) and GPH (green and magenta contours for suppressed and enhanced GPH, respectively, in metres) at 500 hPa in the NWP model against ERA5, averaged over all the seven-day forecasts with TC detected in the WNP basin (0°–60°N, 100°–180°E), during May 15–September 30, 2016. (b) Same as (a), but for the difference of the KPP coupled model over the NWP model. (c, d) Same as (a, b), but only for the seven-day TC forecasts in which the KPP-predicted WNPSH index is >3 m higher than the NWP-predicted WNPSH. (e, f) Same as (a, b), but only for the seven-day forecasts without a TC detected in the WNP basin.
3.3 Association of WNPSH changes with SST changes

As seen in section 3.2, the improved TC track predictions in the coupled models are closely associated with the enhanced WNPSH. In this subsection, we discuss the potential causes of the WNPSH changes between the KPP coupled and NWP models (KPP forecast minus NWP forecast), in the context of air–sea coupling. We propose two hypotheses to explain the changed WNPSH: the local SST change (ΔSST\textsubscript{local}) and the remote SST change (ΔSST\textsubscript{remote}).

There are two reasons why the coupled SSTs differ from the NWP SSTs: the difference in SST initial conditions (ΔSST\textsuperscript{init}) and air–sea coupling (ΔSST\textsuperscript{AO}). Assuming these effects of SST on WNPSH (denoted as f) are independent, the change in WNPSH can be expressed as a linear combination of these terms (Equation 3):

\[
\Delta \text{WNPSH} = f(\Delta \text{SST}\textsubscript{local}) + f(\Delta \text{SST}\textsubscript{remote}) = f(\Delta \text{SST}\textsuperscript{init}\textsubscript{local}) + f(\Delta \text{SST}\textsuperscript{AO}\textsubscript{local}) + f(\Delta \text{SST}\textsuperscript{init}\textsubscript{remote}) + f(\Delta \text{SST}\textsuperscript{AO}\textsubscript{remote}).
\]  

The initial SST (SST\textsuperscript{init}) differs, as the NWP model is initialized from OSTIA and the coupled models are initialized from FOAM (Section 2.1). ΔSST\textsuperscript{AO} is calculated by subtracting initial SST from the predicted SST in the KPP coupled model. “Local” and “remote” areas are the areas of WNPSH (20°–40°N and 120°–140°E) and the tropical West Pacific (0°–15°N and 100°–160°E), respectively. Sensitivity experiments are described and analysed in section 5 to test these hypotheses.

3.3.1 Local SST cooling

Figure 5b shows that, in the composite of the seven-day forecasts with TCs presented, relative to the NWP model, the enhanced WNPSH in the KPP model is associated with about 0.1−0.4 °C colder SSTs around the area of the WNPSH. In the seven-day TC forecasts in which the WNPSH is most strongly enhanced, this SST cold anomaly is even larger (Figure 5d). In the composite of forecasts without TCs, the SST in KPP tends to be warmer than the SST in NWP in the WNP basin, and the WNPSH enhancement reduces strongly (Figure 5f).

We analysed further the time series of the WNPSH difference (ΔWNPSH, Figure 6c) and the local SST difference (ΔSST\textsubscript{local}) between the KPP coupled and NWP models (Figure 6a). The two time series of seven-day forecast averaged fields, which show strong intraseasonal variability, are negatively correlated with \( r = -0.48 \), which is statistically significant at the 95% confidence level. A cold SST anomaly can enhance the WNPSH locally, through suppressing local convection due to reduced instability. This is confirmed by the largely reduced precipitation over this region (see Figure S1).

We next investigate the two sources of ΔSST\textsubscript{local} between the two models: ΔSST\textsuperscript{init}\textsubscript{local} and ΔSST\textsuperscript{AO}\textsubscript{local}. The associations between ΔWNPSH and the two effects at different lead times are examined (Figure 6). ΔWNPSH has more variability from August–September and less variability from May–July, while ΔSST\textsuperscript{init}\textsubscript{local} shows consistently strong intraseasonal variability (Figure 6a and c). ΔWNPSH and ΔSST\textsuperscript{init}\textsubscript{local} are not dependent on their mean states with low frequency (Figure 6b and d), indicating the dominance of high-frequency variability in model difference. The seven-day average ΔWNPSH (Figure 6c) has a similar but slightly stronger correlation with ΔSST\textsuperscript{init}\textsubscript{local} \( (r = -0.51) \) than with the seven-day average ΔSST\textsubscript{local} \( (r = -0.48; \) Figure 6a), indicating the strong effect of the local initial SST. We analyse the lagged correlation between ΔSST\textsuperscript{init}\textsubscript{local} and ΔWNPSH further at different lead times (Figure 7a). The negative correlation grows with lead time, with values peaking at \( r = -0.55 \) when the initial SST difference leads the WNPSH by about 5–6 days, indicating the timescales for the local initial SST forcing the WNPSH in the coupled model. Thus, the SST initial conditions in the KPP model are statistically associated with the WNPSH changes, with the largest effect at 5–6 days of lead time.

On the other hand, ΔWNPSH is also linked to the seven-day forecast SST, which is induced by air–sea interactions (ΔSST\textsuperscript{AO}local). The lagged correlation between the time series of ΔSST\textsuperscript{AO}\textsubscript{local} and ΔWNPSH (Figure 7b) shows that in the KPP coupled model the local SST change, due to air–sea coupling, tends to affect the WNPSH immediately. This coupled effect is maximized in the second and third days of the forecast, with \( r = -0.30 \sim -0.20 \). However, the relationship between the WNPSH and air–sea coupled SST changes is much weaker than the relationship between the WNPSH and the initial SST (Figure 7a). When ΔSST\textsuperscript{AO}\textsubscript{local} is superimposed on the effect of ΔSST\textsuperscript{init}\textsubscript{local}, the total SST shows a robust negative correlation with the WNPSH \( (r = -0.50 \sim -0.30) \) at nearly all lead times. In all, the statistical correlation between local SST difference and WNPSH difference suggests that cooler SSTs are related to a stronger WNPSH in the coupled model.

3.3.2 Remote SST warming

Figure 5b and d show that the KPP coupled SSTs are warmer than the NWP SSTs over the Maritime Continent and east of the Philippines. Warm tropical SSTs may encourage stronger tropical deep convection and hence more subsiding air in the subtropics through the local Hadley circulation. Thus, this SST difference over the Maritime Continent and east of the Philippines may favour an enhanced WNPSH in the north Pacific. Alternatively, warmer tropical SSTs may be a result...
of reduced convection, and hence a weaker Hadley circulation. We average SSTs over $0^\circ$–$15^\circ$N and $100^\circ$–$160^\circ$E to diagnose the link between warmer tropical SSTs ($\Delta SST_{remote}$) and the enhanced WNPSH ($\Delta WNPSH$) in the KPP coupled model, relative to the NWP model. Figure 8a shows that, across the season, the initial tropical SST in the coupled model is generally warmer than in the NWP model (i.e., $\Delta SST_{init,remote} > 0$); tropical SSTs in the coupled model nearly always continue to warm over the seven-day forecast, due to air–sea interactions (i.e., $\Delta SST_{AO,remote} > 0$). In the meantime, the WNPSH nearly always strengthens over the forecast (i.e., $\Delta WNPSH > 0$, Figure 6c). This composite analysis suggests a positive relationship between $\Delta WNPSH$ and $\Delta SST_{remote}$. Note that $\Delta SST_{AO,remote}$ causes a tropical SST warm bias, compared with observed SST (NWP SST, OSTIA; Figure 8b).

Across the 139 forecasts, the seven-day $\Delta SST_{remote}$ has a large intraseasonal variability, following the SST variability itself (Figure 8a and b). $\Delta SST_{remote}$ (Figure 8a) and $\Delta WNPSH$ (Figure 6c) are significantly negatively correlated ($r = -0.27$) at the 95% confidence level. The correlation between $\Delta SST_{init,remote}$ and $\Delta WNPSH$ is weaker ($r = -0.19$), but still significant at the 95% confidence level. We also calculate the lagged correlation of $\Delta SST_{remote}$ against $\Delta WNPSH$ (Figure 9a). $\Delta SST_{init,remote}$ has a positive but statistically insignificant (95% confidence level) correlation with $\Delta WNPSH$ in the first three days of the forecast. At longer lead times, the correlation becomes negative and significant. Figure 9b shows the lagged correlation of $\Delta SST_{AO,remote}$ with $\Delta WNPSH$. $\Delta SST_{AO,remote}$ is significantly negatively correlated with $\Delta WNPSH$, a relationship that strengthens at longer lead times for $\Delta SST_{AO,remote}$. Combining the effects of $\Delta SST_{init,remote}$ and $\Delta SST_{AO,remote}$ shows that $\Delta SST_{remote}$ has a weak correlation with $\Delta WNPSH$ in the first four days of the forecast (from $\Delta SST_{init,remote}$), but a strong negative correlation in the last three days of the forecast (from $\Delta SST_{AO,remote}$; Figure 9a).

Unlike the positive relationship in the composite analysis, the negative correlation between $\Delta SST_{remote}$ and $\Delta WNPSH$ suggests that the warmer tropical SSTs in the KPP coupled model might be associated with a weaker WNPSH (i.e., warm SSTs result from reduced tropical convection, which is associated with a weaker Hadley circulation and a weaker WNPSH). Thus, the relationship between $\Delta SST_{remote}$ and $\Delta WNPSH$ is inconclusive from these analysis methods. This could be
due to the competing effects of $\Delta S_{\text{local}}$ (Figure 7), which outweigh the impact of $\Delta S_{\text{remote}}$ (Figure 9).

Overall, the changed WNPSH in the KPP coupled model could potentially be related to SST changes in two regions: the local subtropics and the remote Tropics. In each region, there are two effects that can cause the coupled SSTs to depart from the NWP SSTS: the difference in SST initial conditions and the air–sea coupling. To identify further which areas and which effects dominate in the WNPSH change, sensitivity experiments with the KPP and NWP models are carried out in section 5.

4 | CHANGES IN TC INTENSITY

The NWP model generally underestimates TC intensity in terms of maximum 10-m wind speeds and minimum MSLP (Figure 10a and b). A companion article showed that this overall underestimation has persisted in all versions of the
UK Met Office NWP model from 2008 to 2017 (Hodges and Klingaman, 2019, submitted, personal communication). An important reason for this underestimation is that the spatial resolution of the model is not high enough to resolve the TC structure or the physical processes for rapid TC intensification. In Figure 10a, the TC sample size is limited because the maximum 10-m wind speed is only recorded in observations when it exceeds 17 m/s (34 knots). Minimum MSLP is too strong in the NWP model in the northeastern part of the region, where TCs transition to extratropical cyclones.

In the KPP coupled model, the TC intensity becomes even weaker north of 20°N (Figure 10c and d). The maximum 10-m wind speed is about 1–5 m/s slower than in the NWP model, corresponding to a minimum MSLP of 2–10 hPa higher, in the northern part of the region. TCs tend to be stronger in the southern section, that is, east of the Philippines. We find that this pattern of intensity change matches the pattern of SST difference between two models over the seven days of forecast (Figure 5b). Generally speaking, TCs are intensified in the areas where the coupled SSTs are warmer (e.g., the deep Tropics) and weakened in the areas where the coupled SSTs are colder (e.g., the northern and central parts of the region).

There are several effects expected to damp the TC intensification. As mentioned in section 1, in coupled models, negative atmosphere–ocean feedbacks via cooler SSTs can damp convection and vorticity and discourage TC development. In the meantime, tropical SST changes due to atmospheric feedbacks may also affect TC intensity via altering the Hadley circulation, as discussed in the last section for TC tracks. Apart from air–sea coupling, different SST initial conditions, primarily the consistently colder initial SST in the coupled model, may contribute to the TC intensity difference, as well via reducing TC convection. In the seven-day forecast, the TC-related precipitation, an indicator of TC convection, in the coupled model is 0.5–1.5 mm/day less than that in the NWP model (see Figure S2). In our analysis so far, however, it remains unclear whether the SST initialization or the air–sea feedbacks dominate the TC intensity reduction in the KPP coupled model, or whether the two effects are equally important. To investigate this, sensitivity experiments are performed and analysed in the next section.

5 | SENSITIVITY EXPERIMENTS

To distinguish the sensitivity of TC predictions to air–sea interactions and initial SST conditions, we performed two additional sets of KPP coupled reforecasts and two additional sets of NWP reforecasts (Table 1). Due to computational limitations, we performed reforecasts for only the period September 1–15, 2016, when the KPP coupled model has the largest improvement for TC position predictions. The SST difference between the control KPP coupled forecast and control NWP forecast is shown in Figure 11a, which consists of effects from air–sea interactions during the forecast (Figure 11b) and the initial SST difference (Figure 11c). During the period, there are two TCs: Typhoons Meranti and Malakas. During this
To test the relative roles of local and remote air–sea coupling on WNP TCs, one KPP coupled reforecast set denies air–sea coupling in the Pacific subtropical band of 17°–35°N ("KPP_noSubTrop"); the other denies air–sea coupling in the Pacific tropical band of 0°–17°N ("KPP_noTrop"). The differences between the KPP_noSubTrop set and the control KPP set show the effects of local air–sea interactions on TC predictions, while the differences between the KPP_noTrop set and the control KPP set show the effects of remote (tropical) air–sea interactions. Similarly, the differences between each KPP set and the control NWP set show the effects of mainly remote or local air–sea interactions, respectively, plus the effect of differences in SST initialization.

Two further NWP reforecast sets are also performed, which differ from the standard NWP set only in their SSTs. One set (“NWP_ODA”) uses the FOAM initial SSTs (as in the standard KPP coupled model); the other (“NWP_blend”) uses a blended SST condition, in which the FOAM SSTs in the WNP basin (0°–40°N and 100°–180°E) are blended with the OSTIA SSTs elsewhere. At the boundaries of the WNP, the

| Air–sea coupling     | Initial SST                  |
|----------------------|------------------------------|
| KPP                  | 25°S–35°N Ocean DA SST       |
| KPP_noTrop           | 25°S–0°N and 17°–35°N Ocean DA SST |
| KPP_noSubTrop        | 25°S–17°N Ocean DA SST       |
| NWP                  | None OSTIA                  |
| NWP_ODA              | None Ocean DA SST           |
| NWP_blend            | None Ocean DA SST in 0°–40°N and 100°–180°E; OSTIA elsewhere |

TABLE 1  SST and air–sea coupling configuration in sensitivity experiments of the KPP coupled and atmosphere-only NWP models, over the period September 1–15, 2016

period, changes in the SST climatology over the seven-day forecast are negligible.
two products are blended across a distance of approximately 7° (45 grid points), using the same technique as in the KPP model (section 2.1). The differences of these two NWP sets from the standard NWP set will show the effects of the change in SST initial conditions globally (Figure 11c) or in the WNP basin alone (Figure 11d), respectively. The comparison between the NWP_ODA and standard KPP sets indicates the effect of air–sea coupling across the entire 25°S–35°N coupled region (Figure 11b).

These new reforecast sets all predict TC tracks that are nearly identical to the standard KPP coupled model for Typhoons Meranti and Malakas (Figure 12), indicating that TC track predictions are more sensitive to the initial SST conditions, mostly in the WNP region, than to air–sea coupling in any region. The better prediction of TC tracks in these sets is associated with a consistent westward extension of the 5,880-m contour of GPH at 500 hPa, that is, an enhanced WNPSH (Figure 13). The difference in the WNPSH between the new sets and the standard KPP coupled model set is negligible. Therefore, the colder SST initial conditions from the ocean data assimilation system FOAM in the WNP basin, relative to OSTIA, are responsible for the local atmospheric circulation change and increased subsidence. The anticyclonic wind anomaly, associated with the enhanced WNPSH (Figure 13), steers TCs further west in the southern part of the basin (Figure 12).

The WNPSH also interacts with the TC intensity. Because TCs travel along the western edge of the WNPSH following the steering flow, in the forecast a weakened TC is usually associated with a locally increased GPH and a westward extension of the WNPSH. Comparing the new sets and the standard KPP coupled model set against the standard NWP model set, the 500-hPa GPH around the western edge of the WNPSH is largely increased (Figure 13), meaning that TC

FIGURE 11 (a) SST difference between the standard KPP coupled model run and the standard NWP model run. (b) SST difference between the standard KPP coupled model run and the NWP_ODA model run. (c) SST difference between the NWP_ODA model run and the standard NWP model run. (d) SST difference between the NWP_blend model run and the standard NWP model run. Forecast period is September 1–15, 2016. Experiment setups are detailed in Table 1

FIGURE 12 Typhoon (a) Meranti and (b) Malakas positions predicted in the sensitivity experiments, for the forecasts initialized on September 9 and 10, 2016, along with the IBTrACS track (black). Experiment setups are detailed in Table 1
FIGURE 13  Difference of 500-hPa GPH predicted in the (a) standard KPP coupled, (b) KPP_noSubTrop, (c) KPP_noTrop, (d) NWP_ODA, and (e) NWP_blend model runs against the standard NWP model run. Forecast period is September 1–15, 2016. The position of the 5,880-m contour of 500-hPa GPH in each experiment is also provided. Experiment setups are detailed in Table 1

strength is greatly reduced at maximum intensity (Figure 14). Thus, apart from an increased subsidence due to the colder SST initial condition, the weakened TC intensity also contributes to the enhanced WNPSH.

Figure 14a shows that in the standard KPP coupled model the TC intensity, in terms of maximum 10-m wind speed, is 2–5 m/s slower than in the NWP model for the seven-day forecast. The weakened TC intensity is due largely to the colder initial SSTs, particularly in the WNP basin (Figure 14a–c). In the meantime, air–sea coupling also plays an important role in TC intensity changes. Comparing the NWP_ODA and new KPP coupled sets against the standard KPP set shows the effects of air–sea coupling on TC intensity predictions (Figure 14d–f). Air–sea coupling reduces TC intensity, reflected in a 1–4 m/s maximum 10-m wind speed reduction (Figure 14d). This TC intensity reduction is less related to the air–sea coupling in the region of the TC tracks (17°–35°N, Figure 14e), but instead is more related to the coupling in the Tropics (0°–17°N, Figure 14f). The SST warming tendency in the Maritime Continent and east of the Philippines (0°–15°N, 100°–160°E, Figure 11b), due to air–sea coupling, encourages stronger deep convection over this region, as indicated by more precipitation (see Figure S3). We hypothesize that the stronger deep convection drives a stronger local Hadley cell circulation, which is associated with increased subtropical subsidence and increased GPH (see Figure S4) that suppresses TC intensification.

6 | CONCLUSIONS AND DISCUSSION

The performance of the atmosphere–ocean coupled and atmosphere-only global NWP models for TC predictions in the WNP basin has been rigorously assessed. In
the 2016 TC season, the UK Met Office atmosphere-only global NWP model forecasts TC positions with a consistent eastward bias in the south of the basin and a westward bias in the north. This is associated with a strong cyclonic steering wind bias caused by a suppressed WNPSH. The atmosphere–ocean-mixed-layer (KPP) coupled model and the atmosphere–3D-ocean (NEMO) coupled model show significantly reduced TC position errors, by $0.5^\circ$–$1.0^\circ$ (50–100 km) on average over the seven-day forecasts of the TC season. The improvement is more robust in the southwest WNP. The two coupled models show nearly identical results for TC predictions. This improvement is related to an enhanced WNPSH in the coupled models, resulting in an anticyclonic steering wind anomaly favouring TCs moving faster in the south of the region and recurving earlier to the north. The coupled models tend to worsen TC intensity predictions, which are underestimated in the NWP model, further in general, even though they slightly improve TC intensity predictions in the southern sector, where the coupled SSTs are warmer.

Possible reasons for the TC prediction differences between the coupled and uncoupled models were investigated. The different SST initial conditions and air–sea coupling seem to contribute to the prediction differences. After further sensitivity experiments for the period September 1–15, 2016, when the TC prediction differences are largest, we confirmed that in the coupled models the enhanced WNPSH, associated with
increased subsidence and reduced TC intensity, is mostly a result of colder initial SSTS, with the effect of coupling being negligible. The ocean component of the coupled models is initialized from the UK Met Office FOAM analysis (Martin et al., 2007; Blockley et al., 2014; Waters et al., 2015), while the NWP SST is prescribed by the observation-based OSTIA product (Donlon et al., 2012). In the 2016 TC season, the former provides a generally colder initial SST in the WNP basin. Therefore, it is important to quantify the effects of any differences in SST initial conditions when evaluating the effect of air–sea coupling on weather predictions.

Sensitivity experiments also showed that changes in TC intensity predictions are sensitive to SST initialization and air–sea coupling. The colder initial SSTS suppress local convection and reduce available potential energy for TCs to develop. In the meantime, the warm SST tendencies around the Maritime Continent and east of the Philippines induced by the air–sea interaction encourage a stronger deep convection and a stronger local Hadley cell circulation. The increased subsidence in the subtropics due to this stronger local Hadley cell circulation consequently reduces TC intensity. Interestingly, in our experiments, we also found that the reduced TC intensity is not clearly related to local SST changes imposed by negative atmospheric feedbacks. This is inconsistent with some other case studies (Mogensen et al., 2017), which showed that local air–sea feedbacks are largely responsible for TC intensity decreases in coupled models. The inconsistency is presumably caused by the SST warming over the forecast in this region (Figure 11b), which reduces the effects of negative local air–sea feedbacks on SST.

The SST warming tendency in the western Pacific during the forecast might also be related to the colder SST initial conditions, indicating that these two effects in coupled models might not be independent. When the coupled models start with a relatively colder SST, the atmosphere will react by reducing convection and cloud, which leads to an increase in surface shortwave radiation and SST warming over the forecast. This two-way air–sea interaction, acting through the air–sea feedback on SST, is induced by the imbalance between atmosphere and SST initial conditions. This effect becomes less distinguishable if it is superimposed with other longer-time air–sea coupling effects, such as the seasonal cycle or El Niño/La Niña events. Coupled data assimilation, which provides dynamically consistent SST and atmospheric initial conditions (Feng et al., 2018), might benefit TC predictions by reducing such atmosphere–ocean imbalances in coupled forecasts, which, at the moment, are initialized separately in the ocean and atmosphere.

There are a few further points worth noting. Firstly, we also analysed the growth of TC prediction errors with lead time, which gave similar results to those of Hodges and Klingaman (2019 submitted, personal communication). However, the error growths in the NWP and coupled models are not statistically different—despite substantial changes in the mean error over the seven-day forecast—probably due to the strong spatial variations of coupling effects on TC predictions (as seen in Figure 2) and limited prediction samples. Secondly, we performed our sensitivity experiments for only a short period (15 days) of the 2016 TC season. Thus, the competing roles of SST initialization and local and remote air–sea coupling for TC predictions may differ if the full 2016 TC season is considered. To address this question fairly, further studies should either repeat the atmosphere-only NWP forecasts for the entire 2016 TC season with persisted SSTS from the FOAM SST analysis (like the NWP_ODA set) or repeat the coupled forecasts but adjust the initial SSTS (and also mixed-layer temperature) to be consistent with the NWP initial SSTS. We were unable to perform a longer set of experiments, due to limited computational resources. Thirdly, the relative performance of the coupled and atmosphere-only NWP models for TC track predictions may vary in other TC seasons. In some other years, for example in strong El Niño and La Niña years, when the large-scale atmospheric circulation and air–sea interaction differ strongly from those in 2016, the merit of coupled models for TC predictions might change. Future work is required to address these issues.

ACKNOWLEDGEMENTS

This work and two of its authors (Xiangbo Feng and Kevin Hodges) were supported by the Weather and Climate Science for Services Partnership for Southeast Asia (WCSSP–Southeast Asia) through the Newton Fund. Nicholas Klingaman was funded by the UK 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. ERA5 is generated by ECMWF and distributed by the Climate Data Store (CDS), under the Copernicus Climate Change Service (C3S). ITrACS data were retrieved from the National Climatic Data Center. The authors thank the UK Met Office for making operational NWP forecasts and the trial NEMO coupled reforecasts available. The authors also thank two anonymous reviewers for their constructive comments and suggestions.

ORCID

Xiangbo Feng https://orcid.org/0000-0003-4143-107X

REFERENCES

Barcikowska, M., Feser, F. and Von Storch, H. (2012) Usability of best track data in climate statistics in the western North Pacific. Monthly Weather Review, 140(9), 2818–2830.
Bender, M.A. and Ginis, I. (2000) Real-case simulations of hurricane–ocean interaction using a high-resolution coupled model: effects on hurricane intensity. Monthly Weather Review, 128(4), 917–946.

Bender, M.A., Ginis, I. and Kurita, Y. (1993) Numerical simulations of tropical cyclone–ocean interaction with a high-resolution coupled model. Journal of Geophysical Research – Atmospheres, 98(D12), 23245–23263.

Blockley, E.W., Martin, M.J., McLaren, A., Ryan, A.G., Waters, J., Lea, D.J., Mirouze, I., Peterson, K.A., Sellar, A. and Storkey, D. (2014) Recent development of the Met Office operational ocean forecasting system: an overview and assessment of the new Global FOAM forecasts. Geoscientific Model Development, 7(6), 2613–2638.

Buizza, R., Balsamo, G. and Haiden, T. (2018) IFS upgrade brings more seamless coupled forecasts. ECMWF Newsletter, 156, 18–22. https://doi.org/10.21957/729r3bdx6.

Camp, J., Roberts, M., MacLachlan, C., Wallace, E., Hermanson, L., Brookshaw, A., Arribas, A. and Scaife, A.A. (2015) Seasonal forecasting of tropical storms using the Met Office GloSea5 seasonal forecast system. Quarterly Journal of the Royal Meteorological Society, 141(691), 2206–2219.

Chan, J.C.L., Gray, W.M., Shi, J.E. and Lam, C.M. (2001) Improvements in the seasonal forecasting of tropical cyclone activity over the western North Pacific. Weather and Forecasting, 16, 491–498.

Chang, S.W. and Anthes, R.A. (1979) The mutual response of the tropical cyclone and the ocean. Journal of Physical Oceanography, 9(1), 128–135.

Chen, S.S., Price, J.F., Zhao, W., Donelan, M.A. and Walsh, E.J. (2007) The CBLAST-Hurricane program and the next-generation fully coupled atmosphere–ocean–mixed-layer models for hurricane research and prediction. Bulletin of the American Meteorological Society, 88(3), 311–318.

Craig, A., Valcke, S. and Coquart, L. (2017) Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0. Geoscientific Model Development, 10(9), 3297–3308.

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

DeMaria, M., Sampson, C.R., Knaff, J.A. and Musgrave, K.D. (2014) Is tropical cyclone intensity guidance improving? Bulletin of the American Meteorological Society, 95(3), 387–398.

Donlon, C.J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E. and Wimmer, W. (2012) The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sensing of Environment, 116, 140–158.

Elswberry, R.L., Fraim, T.S. and Trapnell, R.N. (1976) A mixed layer model of the oceanic thermal response to hurricanes. Journal of Geophysical Research, 81(6), 1153–1162.

Feng, X., Haines, K., Liu, C., de Boisseson, E. and Polo, I. (2018) Improved SST–precipitation intraseasonal relationships in the ECMWF coupled climate reanalysis. Geophysical Research Letters, 45(8), 3664–3672.

Feng, X. and Tsimpis, M.N. (2014) Sea level extremes at the coasts of China. Journal of Geophysical Research, Oceans, 119(3), 1593–1608.

Feng, X., Zheng, J. and Yan, Y. (2012) Wave spectra assimilation in typhoon wave modeling for the East China Sea. Coastal Engineering, 69, 29–41.

Goess, J.S. (2000) Tropical cyclone track forecasts using an ensemble of dynamical models. Monthly Weather Review, 128(4), 1187–1193.

Guo, L., Klingaman, N.P., Vidale, P.L., Turner, A.G., Demory, M.E. and Cobb, A. (2017) Contribution of tropical cyclones to atmospheric moisture transport and rainfall over East Asia. Journal of Climate, 30(10), 3853–3865.

Heming, J.T. (2016) Met Office Unified Model tropical cyclone performance following major changes to the initialization scheme and a model upgrade. Weather and Forecasting, 31(5), 1433–1449.

Heming, J.T. (2017) Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models. Meteorological Applications, 24(1), 1–8.

Hersbach, H. and Dee, D.P. (2016) ERA-5 is in production. ECMWF Newsletter, 147, 7.

Hirahara, S., Alonso-Balmaseda, M., de Boisseson, E. and Hersbach, H. (2016) Sea surface temperature and sea ice concentration for ERA5. ERA Report Series no. 26. Reading: ECMWF.

Hirons, L.C., Klingaman, N.P. and Woolnough, S.J. (2015) MetUM-GOML: a near-globally coupled atmosphere–ocean-mixed-layer model. Geoscientific Model Development, 8, 363–379.

Hodges, K.I. (1994) A general method for tracking analysis and its application to meteorological data. Monthly Weather Review, 122(11), 2573–2586.

Hodges, K.I. (1995) Feature tracking on the unit sphere. Monthly Weather Review, 123(12), 3458–3465.

Hodges, K.I. (1999) Adaptive constraints for feature tracking. Monthly Weather Review, 127(6), 1362–1373.

Hodges, K.I. and Emerdon, R. (2015) The prediction of Northern Hemisphere tropical cyclone extended life cycles by the ECMWF ensemble and deterministic prediction systems. Part I: tropical cyclone stage. Monthly Weather Review, 143(12), 5091–5114.

Knapp, K.R., Kruk, M.C., Levinson, D.H., Diamond, H.J. and Neumann, C.J. (2010) The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data. Bulletin of the American Meteorological Society, 91(3), 363–376.

Large, W.G., McWilliams, J.C. and Doney, S.C. (1994) Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. Reviews of Geophysics, 32(4), 363–403.

LaRow, T.E., Stefanova, L., Shin, D.W. and Cocke, S. (2010) Seasonal Atlantic tropical cyclone hindcasting/forecasting using two sea surface temperature datasets. Geophysical Research Letters, 37(2), 1–5.

Lu, W., Yan, X.H., Han, L. and Jiang, Y. (2017) One-dimensional ocean model with three types of vertical velocities: a case study in the South China Sea. Ocean Dynamics, 67(2), 253–262.
Manabe, S., Holloway, J.L., Jr. and Stone, H.M. (1970) Tropical circulation in a time-integration of a global model of the atmosphere. *Journal of the Atmospheric Sciences,* 27(4), 580–613.

Martin, M.J., Hines, A. and Bell, M.J. (2007) Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. *Quarterly Journal of the Royal Meteorological Society,* 133(625), 981–995.

Mogensen, K.S., Magnusson, L. and Bidlot, J.R. (2017) Tropical cyclone sensitivity to ocean coupling in the ECMWF coupled model. *Journal of Geophysical Research, Oceans,* 122(5), 4392–4412.

Pianezze, J., Barthe, C., Bielli, S., Tulet, P., Jullien, S., Cambon, G., Bouquet, O., Claey, M. and Cordier, E. (2018) A new coupled ocean–waves–atmosphere model designed for tropical storm studies: example of tropical cyclone Bejisa (2013–2014) in the South-West Indian Ocean. *Journal of Advances in Modeling Earth Systems,* 10(3), 801–825.

Ren, F., Liang, J., Wu, G., Dong, W. and Yang, X. (2011) Reliability analysis of climate change of tropical cyclone activity over the western North Pacific. *Journal of Climate,* 24(22), 5887–5898.

Short, C.J. and Petch, J. (2018) How well can the Met Office Unified Model forecast tropical cyclones in the western North Pacific? *Weather and Forecasting,* 33(1), 185–201.

Takaya, Y., Kubo, Y., Maeda, S. and Hirahara, S. (2017) Prediction and attribution of quasi-annual tropical cyclone activity in the early summer of 2016: case study of lingering effects by preceding strong El Niño events. *Atmospheric Science Letters,* 18(8), 330–335.

Takaya, Y., Vitart, F., Balsamo, G., Balmsada, M., Leutbecher, M. and Molteni, F. (2010) Implementation of an ocean mixed layer model in IFS. Technical report 622. ECMWF: Reading, UK.

Vecchi, G.A., Zhao, M., Wang, H., Villarini, G., Rosati, A., Kumar, A., Held, I.M. and Gudgel, R. (2011) Statistical-dynamical predictions of seasonal North Atlantic hurricane activity. *Monthly Weather Review,* 139(4), 1070–1082.

Vincent, E.M., Lengaigne, M., Madec, G., Viallard, J., Samson, G., Jourdain, N.C., Menkes, C.E. and Jullien, S. (2012) Processes setting the characteristics of sea surface cooling induced by tropical cyclones. *Journal of Geophysical Research, Oceans,* 117(C2), C0220.

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

Vitart, F. and Stockdale, T.N. (2001) Seasonal forecasting of tropical storms using coupled GCM integrations. *Monthly Weather Review,* 129(10), 2521–2537.

Walters, D., Brooks, M., Boule, I., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P.L., Wilkinson, J., Willett, M., Woolnough, S. and Xavier, P. (2017) The Met Office unified model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geoscientific Model Development,* 10(4), 1487–1520.

Wang, B., Xiang, B. and Lee, J.Y. (2013) Subtropical high predictability establishes a promising way for monsoon and tropical storm predictions. *Proceedings of the National Academy of Sciences,* 110(8), 2718–2722.

Waters, J., Lea, D.J., Martin, M.J., Mirozue, I., Weaver, A. and While, J. (2015) Implementing a variational data assimilation system in an operational 1/4 degree global model ocean. *Quarterly Journal of the Royal Meteorological Society,* 141(687), 333–349.

Williams, K.D., Harris, C.M., Boda-Salcedo, A., Camp, J., Comer, R.E., Copsey, D., Ferleday, D., Graham, T., Hill, R., Hinton, T., Hyder, P., Ineson, S., Masato, G., Milton, S.F., Roberts, M.J., Rowell, D.P., Sanchez, C., Shelly, A., Sinha, B., Walters, D.N., West, A., Woolings, T. and Xavier, P.K. (2015) The Met Office global coupled model 2.0 (GC2) configuration. *Geoscientific Model Development,* 8(5), 1509–1524.

Wu, B., Li, T. and Zhou, T. (2010) Relative contributions of the Indian Ocean and local SST anomalies to the maintenance of the western North Pacific anomalous anticyclone during the El Niño decaying summer. *Journal of Climate,* 23(11), 2974–2986.

Wu, C.C., Tu, W.T., Pun, I.F., Lin, I.I. and Peng, M.S. (2016) Tropical cyclone–ocean interaction in Typhoon Megi (2010)—a synergy study based on ITOP observations and atmosphere–ocean coupled model simulations. *Journal of Geophysical Research–Atmospheres,* 121(1), 153–167.

Yablonsky, R.M. and Ginis, I. (2009) Limitation of one-dimensional ocean models for coupled hurricane–ocean model forecasts. *Monthly Weather Review,* 137(12), 4410–4419.

Yamaguchi, M., Ishida, J., Sato, H. and Nakagawa, M. (2017) WGNE intercomparison of tropical cyclone forecasts by operational NWP models: a quarter century and beyond. *Bulletin of the American Meteorological Society,* 98(11), 2337–2349.

Yamaguchi, M., Vitart, F., Lang, S.T., Magnusson, L., Elsberry, R.L., Elliott, G., Kyouda, M. and Nakazawa, T. (2015) Global distribution of the skill of tropical cyclone activity forecasts on short-to medium-range time scales. *Weather and Forecasting,* 30(6), 1695–1709.

**SUPPORTING INFORMATION**

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