The Impact of Analysis Correction-Based Additive Inflation on Subseasonal Tropical Prediction in the Navy Earth System Prediction Capability

Stephanie S. Rushley[a], Matthew A. Janiga[a], William Crawford[a], Carolyn A. Reynolds[a], William Komaromi[b,c], and Justin McLay[a]

[a] *Naval Research Laboratory Marine Meteorology Division, Monterey, California*
[b] *NOAA/NWS/OSTI-Modeling, Silver Spring, Maryland*
[c] *I.M. Systems Group, Silver Spring, Maryland*

ABSTRACT: Accurately simulating the Madden–Julian oscillation (MJO), which dominates intraseasonal (30–90 day) variability in the tropics, is critical to predicting tropical cyclones (TCs) and other phenomena at extended-range (2–3 week) time scales. MJO biases in intensity and propagation speed are a common problem in global coupled models. For example, the MJO in the Navy Earth System Prediction Capability (ESPC), a global coupled model, has been shown to be too strong and too fast, which has implications for the MJO–TC relationship in that model. The biases and extended-range prediction skill in the operational version of the Navy ESPC are compared to experiments applying different versions of analysis correction-based additive inflation (ACAI) to reduce model biases. ACAI is a method in which time-mean and stochastic perturbations based on analysis increments are added to the model tendencies with the goals of reducing systematic error and accounting for model uncertainty. Over the extended boreal summer (May–November), ACAI reduces the root-mean-squared error (RMSE) and improves the spread–skill relationship of the total tropical and MJO-filtered OLR and low-level zonal winds. While ACAI improves skill in the environmental fields of low-level absolute vorticity, potential intensity, and vertical wind shear, it degrades the skill in the relative humidity, which increases the positive bias in the genesis potential index (GPI) in the operational Navy ESPC. Northern Hemisphere integrated TC genesis biases are reduced (increased number of TCs) in the ACAI experiments, which is consistent with the positive GPI bias in the ACAI simulations.

KEYWORDS: Madden-Julian oscillation; Tropical cyclones; Forecast verification/skill; Forecasting techniques; Numerical weather prediction/forecasting

1. Introduction

Tropical cyclones (TCs) are one of the most destructive weather events, bringing with them risks to life and property from storm surge, flooding from extreme precipitation, and damaging winds. TCs form in regions with high sea surface temperatures (SSTs), low vertical wind shear, anomalous low-level cyclonic vorticity, and high relative humidity (RH; e.g., Gray 1968, 1979; Emanuel 1986, 1987, 1991; Raymond et al. 2007; Sessions et al. 2019). The MJO is a planetary scale, tropical wave that dominates variability in the intraseasonal (30–90 day) time scale (Madden and Julian 1971, 1972; Zhang 2005, 2013; Jiang et al. 2020). The MJO is a key source of predictability in the subseasonal (~1 week to 2 month) time scale (Zhang 2013; White et al. 2017; H. Kim et al. 2018; Stan et al. 2022; Domeisen et al. 2022) and has tight relationship with TCs activity (Maloney and Hartmann 2000a,b; Klotzbach 2010; Kossin et al. 2010; Klotzbach and Oliver 2015).

While the benefits of predicting TCs at longer lead times are undisputed, historically TC prediction has been a challenge for forecast models. Though there have been modest improvements in TC prediction skill over the last few decades, large uncertainties associated with their formation, track, and intensity ensure that extended-range prediction of TCs remains a challenge (DeMaria et al. 2014; Emanuel and Zhang 2016; Sridevi et al. 2022). These uncertainties have been linked to the representation of the mean state, convective parameterization, and radiative feedbacks in various global forecast and climate models (D. Kim et al. 2018; Wing et al. 2019). TCs in forecast models are heavily influenced by the model's representation of convection and cloud processes, which affect radiative feedbacks and surface fluxes critical to TC physics. For instance, Stan (2012) found that explicit cloud process representation in the superparameterized Community Climate System Model (CCSM) lead to an increase in the number of TCs that were more intense and longer lived. Several studies have also shown that ocean coupling reduces TC intensity in the model by reducing the surface flux feedbacks (Bender and Ginis 2000; Ogata et al. 2016; Li and Sriver 2018; Wing et al. 2019).

One of the largest challenges facing subseasonal-to-seasonal (S2S) prediction is tapping into the full potential predictability of the MJO (Stan et al. 2022). Predictability of TCs is largely influenced by the representation of the large-scale environment, which is heavily modulated on intraseasonal time scales by

Corresponding author: Stephanie Rushley, stephanie.rushley@nrlmry.navy.mil

the MJO (Aiyyer and Molinari 2008; Camargo et al. 2009; Domeisen et al. 2022; Stan et al. 2022). Previous studies have shown that there is potential prediction skill of TC activity out to weeks 3 and 4 when the MJO is well represented in models (Lee et al. 2018; Domeisen et al. 2022). This predictability is closely tied to how well models simulate the MJO, its teleconnections, and its relationship with TCs and the mean state (D. Kim et al. 2018; Domeisen et al. 2022; Stan et al. 2022).

The MJO has a tight relationship with TC genesis and tracks, and, depending on the phase of the MJO, it can either help or hinder TC activity by modifying the surrounding environment (Maloney and Hartmann 2000a,b; Aiyyer and Molinari 2008; Klotzbach 2010; Klotzbach and Oliver 2015). Locally, the MJO's active phase is associated with cyclonic vorticity and convergence off the equator, which supports TC genesis, while during the MJO's suppressed phase, TC activity is damped due to subsidence and drying (Maloney and Hartmann 2000a,b; Aiyyer and Molinari 2008; Stan et al. 2022; Domeisen et al. 2022). The MJO can influence TCs remotely through global teleconnections that alter the environmental vertical wind shear, moisture, and low level vorticity (Klotzbach 2010; Klotzbach and Oliver 2015; Lee et al. 2020), which can be at least as large as those driven by El Niño–Southern Oscillation (ENSO; Maloney and Hartmann 2000a).

In the Atlantic basin, TC days are twice as high and the number of major hurricane days are 3 times as high when the enhanced MJO convection is in the Indian Ocean compared to when the enhanced MJO convection is over the central Pacific Ocean (Klotzbach 2010). In S2S models, TC activity is amplified in the main development region of the Atlantic Ocean when forecasts are initialized when the MJO enhanced convection is in the central Pacific due to a reduction in the vertical wind shear over the Atlantic. TC activity is reduced when forecasts are initialized when the MJO's enhanced convection is over the Maritime Continent due to increased vertical wind shear over the Atlantic (Stan et al. 2022). Similarly, when the MJO-scale winds are easterly over the eastern Pacific there is anomalous anticyclonic vorticity, suppressing TC activity. When the MJO-scale winds over the eastern Pacific are westerly the environment is characterized by cyclonic vorticity and low vertical wind shear, which leads to a greater preference for TC genesis in the Gulf of Mexico and Caribbean basins (Maloney and Hartmann 2000a,b).

Improving the representation of the MJO in forecast and climate models has been a priority of the modeling community, resulting in improved skill over the last decade due to ocean-atmosphere coupling and improved representation of the moisture-precipitation relationship (H. Kim et al. 2014; Vitart 2014; Kim et al. 2016; Lim et al. 2018). Improvements in the predictability of the MJO can have a large impact on extended-range TC prediction. When forecasts are initialized during MJO events, predictions of TC activity are skillful out to 3–4 weeks (Lee et al. 2018; Domeisen et al. 2022). Recent studies have found that while the MJO is potentially predictable out to 6–7 weeks, current operational models only show MJO predictive skill out to 2–5 weeks (H. Kim et al. 2014, 2018). To take full advantage of the potential prediction skill of the MJO, reducing model errors that lead to errors in the

MJO is vital. Biases in the MJO have been attributed to biases in the mean state and uncertainties in the parameterization and representation of subgrid-scale processes, such as those in the convection and cumulus parameterization schemes (Shutts 2005; Kim et al. 2009; D. Kim et al. 2014; Kim 2017; Gonzalez and Jiang 2017; Lim et al. 2018; Kim et al. 2019; Vannitsem et al. 2021; Rushley et al. 2022). Forecast errors can arise from systematic errors which are due to errors in parameterizations and model numerical calculations, errors in the boundary conditions and missing processes, and random errors attributable to uncertainty in unresolved subgrid-scale physics, parameterization, and noise (Saha 1992; Buizza et al. 1999; Danforth and Kalnay 2008; Klocke and Rodwell 2014; Piccolo et al. 2018; Bhargava et al. 2018; Chang et al. 2019; Crawford et al. 2020; Reynolds et al. 2022). Forecast skill can be increased through improvements to the data assimilation and modeling systems (including bias reduction), and, in an ensemble system, through improvements to the representation of initial condition and model uncertainty. Many methods have been proposed to account for and correct model biases, including both offline postprocessing and in-line bias correction.

Different in-line corrections methods have been successful in reducing systematic error growth. Saha (1992) reduced the systematic errors in the medium range U.S. National Centers for Environmental Prediction (NCEP) model by implementing artificial sources and sinks of heat, momentum, and mass. Other studies have used data assimilation (DA) and numerical weather prediction (NWP) models to estimate bias using analysis increments to identify and correct for model drift, which when implemented in-line acts to reduce systematic errors in both weather and climate models (Dee and Da Silva 1998; Rodwell and Palmer 2007; Mapes and Bacmeister 2012; Klocke and Rodwell 2014; Ulate et al. 2015; Cavallo et al. 2016; Bhargava et al. 2018; Piccolo et al. 2018; Crawford et al. 2020; Crawford 2021). Corrections based on tendency errors have been shown to improve model performance, reliability, and error reduction on subseasonal time scales (Chang et al. 2019).

The impacts of stochastic forcing on biases cannot be ignored when considering bias correction. Several studies have implemented stochastic perturbations into forecast models in an attempt to account for model uncertainties with great success, resulting in a reduction of the model bias, as well as an improvement in the ensemble spread (Buizza et al. 1999; Weisheimer et al. 2014; Ollinaho et al. 2017; Berner et al. 2017; Li et al. 2019). Additionally, stochastic perturbations can help improve simulations of regime changes and responses to external forcing (Berner et al. 2017). Stochastic models are designed to account for random errors associated with physical parameterizations by perturbing different variables based on a known probability distribution function (PDF; Buizza et al. 1999; Batté and Déqué 2016; Berner et al. 2017; Ollinaho et al. 2017; Piccolo et al. 2018). One such method is stochastic kinetic energy backscatter (SKEB), which works to inject energy back into the system to account for uncertainties that arise from unresolved subgrid scale processes (Shutts 2005; Weisheimer et al. 2014; Berner et al. 2017; Li et al. 2019). In addition to increasing ensemble spread, SKEB has been found to reduce model bias in

key forecast variables such as 500-hPa geopotential, 2-m temperature, upper- and lower-level winds, RH, cloud cover, and tropical convection (Weisheimer et al. 2014; Batté and Déqué 2016; Berner et al. 2017; Li et al. 2019). Piccolo et al. (2018) applied random analysis increments from a DA model that constrained the stochastic perturbations to a set of known observations, which improved the reliability of the model. Li et al. (2019) applied stochastic perturbation methods to the Global Ensemble Forecast System (GEFS) and showed improvements in the MJO skill and the environmental variables important for the MJO–TC relationship including tropical RH, zonal winds, and temperature. Weisheimer et al. (2014) and Berner et al. (2017) also found that improved skill in terms of MJO amplitude and frequency results in improved prediction skill in the subseasonal time scale, suggesting promising progress in improving the predictability gap in this time scale.

Both systematic and stochastic forcing can play an important role in reducing forecast bias (Crawford et al. 2020) and previous studies have shown significant improvements by implementing one or the other. In this study, we examine the effects of analysis correction-based additive inflation (ACAI), which applies perturbations to the tendency terms with the goal of reducing the systematic errors and improving ensemble performance (Bowler et al. 2017; Crawford et al. 2020). ACAI estimates model error using an archive of analysis corrections, which is applied to the full state vector of the model tendency of surface pressure, temperature, humidity, and wind speed (Crawford et al. 2020; Reynolds et al. 2022). Crawford et al. (2020) implemented ACAI into the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014), the atmospheric component of the Navy Earth System Prediction Capability (ESPC), and found that both the systematic and random component of ACAI play a significant role in the adjustment of the bias. The systematic component resulted in larger bias corrections for the upper and lower level winds while the stochastic component resulted in larger bias corrections for the temperature field. Some degradation in skill occurred at longer lead times, particularly when the bias tended to change sign or trend toward zero at longer lead times, resulting in an overcorrection by ACAI (Crawford et al. 2020). In the fully coupled version of the Navy ESPC, ACAI had a beneficial effect on integrated vapor transport forecast bias (Reynolds et al. 2022). While ACAI reduced biases in both the forecasted winds and moisture fields, the reduction in integrated vapor transport bias was due primarily to the reduction in the wind speed bias. As in Crawford et al. (2020), Reynolds et al. (2022) found that ACAI was less effective or even had a negative impact on forecast biases in regions where the bias changed with forecast lead time, and in regions where biases were dependent on intraseasonal or seasonal variability, such as the South Asian monsoon region.

In this paper, we examine the effects of different configurations of ACAI on the forecast skill in the Navy ESPC. The effects of ACAI on forecast skill of TCs is examined using the genesis potential index (GPI), and the environmental variables associated with TC genesis including RH, vertical wind shear, low-level absolute vorticity, and potential intensity (PI). The impact of ACAI on MJO forecast skill is also explored. In addition to documenting the skill in the ACAI versions of the Navy ESPC, we discuss the role that the physical parameters directly impacted by ACAI (surface pressure, temperature, moisture, and horizontal wind speed) may play in changes to forecast skill.

The remainder of this paper is organized as follows: data and methods are described in section 2, model skill is examined in section 3, the physical processes in the model that contribute to the biases are discussed in section 4, and last, section 5 presents the conclusions of this study.

2. Data and methods

a. Observations

Observations of moisture (q), sea surface temperature (SST), geopotential height (Φ), horizontal winds (u, v), and temperature (T) are taken from the European Centre for Medium-Range Forecasts (ECMWF) interim reanalysis (ERA-Interim) four times daily from 2007 to 2018 (Dee et al. 2011) at a $2.5° \times 2.5°$ horizontal resolution. Observations of outgoing longwave radiation (OLR) are obtained over 2007–18 from the National Oceanic and Atmospheric Administration (NOAA) Climate Data Record (Lee and NOAA CDR Program 2011).

TC track and intensity are obtained using the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al. 2010; Schreck et al. 2014; Knapp et al. 2018). The IBTrACS data are gridded onto a $2.5° \times 2.5°$ grid using the Joint Typhoon Warning Center (JTWC) and National Hurricane Center (NHC) tracks (Schreck et al. 2014).

b. Navy ESPC

The Navy ESPC is a global, fully coupled atmosphere–ocean–ice forecast system developed at the U.S. Naval Research Laboratory (NRL). The components of the Navy ESPC are as follows: the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014) atmosphere, the Community Ice Code (CICE; Hunke and Lipscomb 2008), and the Hybrid Coordinate Ocean Model (HYCOM; Bleck 2002). See Barton et al. (2021) for additional details on the Navy ESPC configuration and coupling methods.

In this study, we examine the effects of ACAI on the performance of the Navy ESPC. ACAI is an in-line bias correction and stochastic forcing method that applies perturbations to the tendency terms of surface pressure, temperature, humidity, and meridional and zonal wind speeds to correct for model errors (Crawford et al. 2020; Reynolds et al. 2022). The ACAI perturbations are comprised of a mean component aimed at reducing systematic error and a stochastic component aimed at improved ensemble spread–skill. The mean component of the ACAI perturbation is calculated using a multimonth average analysis correction and the stochastic component is from a random analysis correction selected from the same multimonth time period used to compute the average correction (Crawford et al. 2020; Reynolds et al. 2022). The ACAI perturbations are applied to all prognostic model variables at each model grid point and at each model time-step (Crawford et al. 2020). The effect of ACAI is compared to the current operational version of the Navy ESPC ensemble (Ops), which has no implementation of

TABLE 1. List of experiments used in this study.

Ensemble	Simulation	Ensemble size	Perturbation method	HYCOM grid	CICE version
Ops	Operational	16	None	1/12°	4
ACAI_static	ACAI	7	2011 static archive	1/4°	5
ACAI_60tr	ACAI	7	60-day running archive	1/4°	5
ACAI_static_hr	ACAI	7	2011 static archive	1/12°	5

ACAI. The Ops simulation has 16 ensemble members and is run on a 1/12° HYCOM grid and uses CICE version 4.

The Ops simulation is compared to three versions of ACAI. The differences in the ACAI simulations are the resolution of the ocean model and the method of how the ACAI tendency perturbation is calculated. The ACAI_static simulation uses perturbations computed from a static archive of analysis increments in a fixed year, in this case 2011. The mean component of the ACAI_static perturbations is derived from a 3-month average centered on the initialization month of the forecast and the stochastic component is drawn from the same 3-month period. ACAI_static is run on a 1/4° HYCOM grid (Crawford 2021). The ACAI_60tr is also run on a 1/4° HYCOM grid; however, it calculates the ACAI perturbations from a 60-day trailing window from a running archive of analysis increments (Crawford 2021). The mean component of the ACAI_60tr perturbations is derived from a 60-day average prior to the initialization date of the forecast and the stochastic component is drawn from the same 60-day period. Last, the ACAI_static_hr simulation calculates the tendency perturbation from the 2011 fixed year archive of analysis increments, but is run on a higher resolution 1/12° HYCOM grid (Crawford 2021). All three versions of ACAI are run using CICE version 5 and have 7 ensemble members (Reynolds et al. 2022). The simulations used in this study are summarized in Table 1.

All simulations (Ops and ACAI) in this study are run for 45 days and initialized once a week from 1 February 2017 to 31 January 2018. Here, we examine the extended boreal summer months of May–November (MJJASON), selecting forecasts based on the initialization date to avoid artificially weighting earlier lead times in forecasts that are initialized later in November (Rushley et al. 2022). The atmospheric component of the forecasts is run on a T359 grid (~37-km equatorial resolution) (Barton et al. 2021), and has been interpolated onto a 2.5° × 2.5° grid and averaged to daily values.

c. Diagnostics

1) SKILL METRICS

Ideally, the spread of the ensemble (measured in terms of ensemble standard deviation) should be equal to the root-mean-square error (RMSE) of the ensemble mean (i.e., the spread–skill relationship). Following previous studies on forecast model performance (e.g., Crawford et al. 2020; Reynolds et al. 2022, etc.), we examine the model spread–skill by examining the RMSE, ensemble spread, and absolute value of the bias on the forecast anomaly fields. The observed climatology is calculated over the years 2007–18 and smoothed using a 2-week moving average, then removed from both the observations and the forecast to obtain the anomalies prior to calculating the RMSE,

spread, and bias. An increase in RMSE indicates a degradation of the prediction skill in the model, while a reduction shows an improvement in the prediction skill. The absolute values of the bias can be interpreted similarly. Ideally the ensemble spread will match the RMSE of the ensemble mean. An ensemble is under dispersive if the spread is lower than the RMSE and over dispersive if the spread is higher than the RMSE. In the calculation of the RMSE, ensemble spread, and bias, the analysis is restricted to the global tropics between 30°N and 30°S, over ocean grids only. The RMSE is calculated by first calculating the ensemble mean of the forecast, then taking the square root of the mean squared error between the ensemble mean of the forecast and observations. The RMSE is then averaged over the latitude and longitude. For the bias, the absolute value of the bias is taken for each ensemble member, then average over the ensembles, all forecasts, and finally latitude and longitude. The ensemble spread is calculated as the standard deviation of the ensemble members, then averaged over all forecasts and latitude and longitude. Due to the larger ensemble in the Ops simulation (16 members), the ensemble spread for the Ops simulation is calculated using 1000 bootstraps without resampling to obtain 7 ensemble members to match the number of ensembles in the ACAI version (Reynolds et al. 2022), the average of which is used as the representative spread in the operational model.

The improvements in the bias from ACAI in several mean state variables are examined by comparing the bias in the Ops simulation to the bias in the ACAI simulations. To better examine the impacts of ACAI on the bias in the Navy ESPC, the differences in the bias are calculated relative to the Ops simulation.

In addition to the changes to the RMSE, bias, and spread, the relative effect of these improvement on the model skill are evaluated using percent changes [%Change$_{ACAI}$ = 100 × (Metric$_{ACAI}$ − Metric$_{Ops}$)/Metric$_{Ops}$], where the Metric is RMSE, spread, or bias. The statistical significance of the biases and the changes of ACAI biases are examined using a bootstrap with random resampling of ensemble members and forecasts. The ensemble members are assumed to be independent of each other, as the initialization of each ensemble member is independent and uses independent EDA cycles (Barton et al. 2021). The bootstrap has 1000 simulations. The changes in the bias maps are considered statistically significant at the 99.5% confidence interval, to account for the correlation of adjacent grids and false rejections of the null hypothesis due to a large sample size (Wilks 2016).

2) MJO FILTER

The MJO is isolated in the OLR and u_{850} fields using wavenumber-frequency filtering following Wheeler and
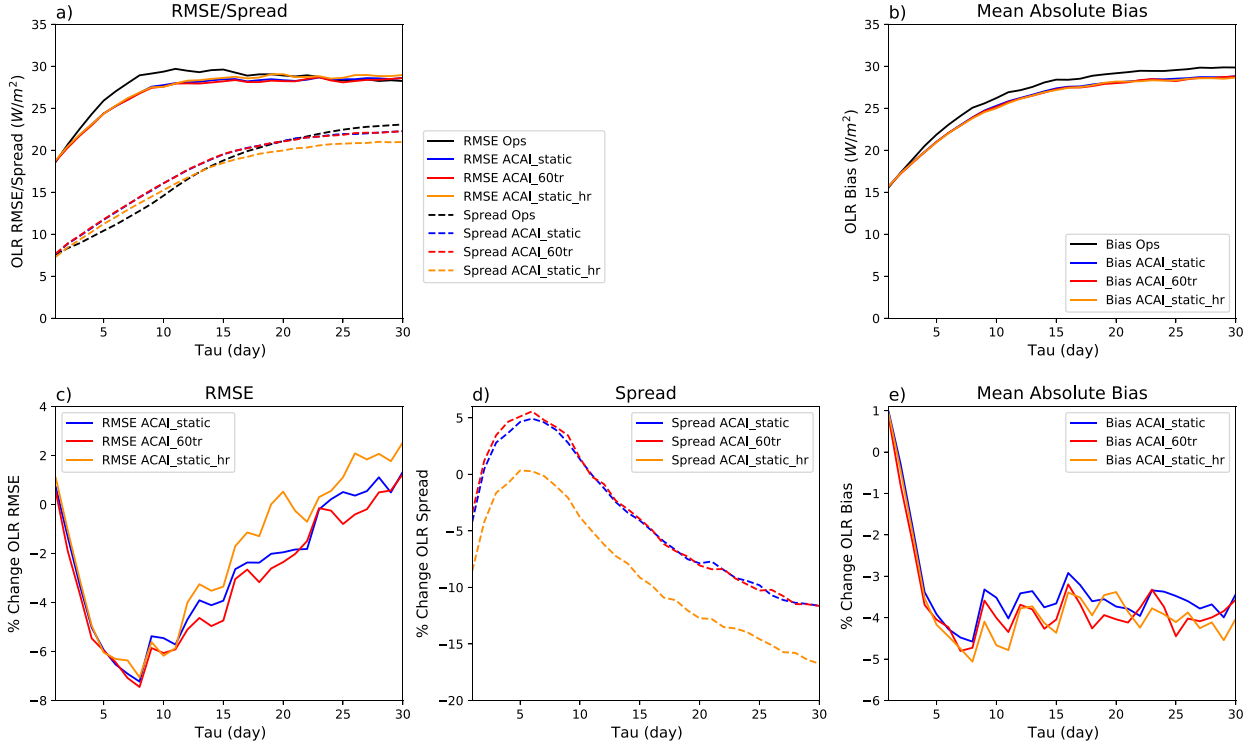
OLR Anomalies MJJASON

FIG. 1. (a) Ensemble mean RMSE (solid lines) and ensemble spread (dashed lines) and (b) mean absolute bias of OLR for the tropical band (30°S–30°N) during MJJASON. Black lines indicate the operational model (Ops), colored lines are the different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold). The percent changes to (c) RMSE, (d) spread, and (e) bias relative to OPS. The colored lines are the percent changes in versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold).

Kiladis (1999), filtering for the MJO band, which is defined here as periods between 30 and 90 days and eastward wavenumbers 1–6. Similar to the examination of the MJO in previous versions of the Navy ESPC (Janiga et al. 2018; Rushley et al. 2022), we use the padded filtering method to extend the 45-day forecast with observations (Janiga et al. 2018). Two years of observations preceding the forecast initialization date are concatenated with the 45-day forecast and a matrix of zeros is added to the end of the forecast period to create a 4-year long dataset for each forecast. To maintain consistency with the comparison to observations, the observations are similarly padded (Janiga et al. 2018). The observed climatology is removed from each forecast and corresponding observations, then filtered for the MJO band. Due to spectral leakage due to the zero's matrix (Janiga et al. 2018), only the first 30 days in the forecast are examined. We use this method of MJO filtering in favor of the real-time multivariate MJO (RMM) index (Wheeler and Hendon 2004), because this method allows for a clear separation of the impact of ACAI on the MJO's circulation and convection and has been a successful forecast metric in examining the MJO in previous studies (Janiga et al. 2018; Rushley et al. 2022).

3) TROPICAL CYCLONE ENVIRONMENT

The genesis potential index (GPI; Emanuel and Nolan 2004; Camargo et al. 2007, 2009) is used to evaluate the skill in the environmental variables important for TC genesis in the ACAI runs. The GPI combines the environmental variables that are important to TC genesis (Emanuel and Nolan 2004; Camargo et al. 2007):

$$\text{GPI} = |10^5 \eta|^{3/2} \left(\frac{\mathcal{H}}{50}\right)^3 \left(\frac{V_{\text{pot}}}{70}\right)^3 (1 + 0.1 V_{\text{shear}})^{-2}, \qquad (1)$$

where η is the absolute vorticity at 850 hPa, $\mathcal{H}$ is the RH at 700 hPa, V_{pot} is the PI, and V_{shear} is the vertical shear between 850 and 200 hPa (Emanuel and Nolan 2004; Camargo et al. 2007, 2009). The PI is the theoretical maximum wind speed that a TC can reach in a given environment if it were to act as an idealized Carnot cycle (Emanuel et al. 1988). The PI is calculated using the pyPI Python package (Gilford 2021), which adapts the algorithm from Bister and Emanuel (2002). The algorithm developed by Gilford (2021) assumes that 1) work done against friction by outflow is negligible, 2) an anticyclone is fully developed when the TC reaches maximum intensity, and 3) that cyclostrophic balance can be used to estimate
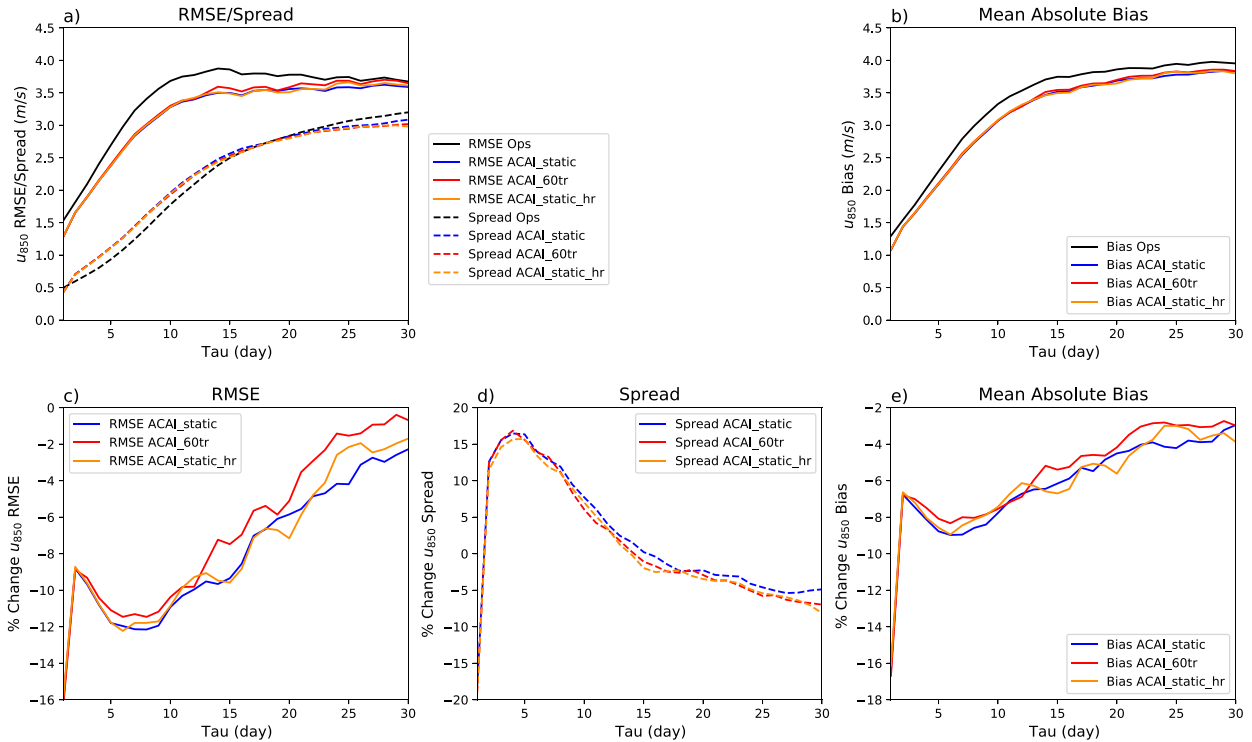
u_{850} Anomalies MJJASON

FIG. 2. (a) Ensemble mean RMSE (solid lines) and ensemble spread (dashed lines) and (b) mean absolute bias of u_{850} for the tropical band (30°S–30°N) during MJJASON. Black lines indicate the operational model (Ops), colored lines are the different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold). The percent changes to (c) RMSE, (d) spread, and (e) bias relative to OPS. The colored lines are the percent changes in different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold).

gradient wind at the radius of maximum winds. Under these assumptions, Gilford (2021) estimates the PI as

$$V_{pot} = (V_{max})^2 = \frac{C_k}{C_D}\frac{(T_s - T_0)}{T_0}(h_o^* - h^*), \qquad (2)$$

where $(V_{max})^2$ is the PI; T_s is the SST; T_0 is the outflow temperature; C_k and C_D are the enthalpy and momentum surface exchange coefficients, respectively, the ratio of which is considered constant here; and h_o^* and h^* are the saturation of moist static energy at the sea surface and of the air above the boundary layer, respectively (Gilford 2021). The GPI is calculated daily at each grid point.

4) TROPICAL CYCLONE ACTIVITY

TC tracks in the Navy ESPC are obtained using the TempestExtremes algorithm (Ullrich and Zarzycki 2017; Ullrich et al. 2021; Zarzycki and Ullrich 2017). For the identification of TCs in the Navy ESPC the high resolution (1/2° grid, 3 hourly) output of geopotential heights at 500 and 300 hPa, 10-m horizontal wind speed, and mean sea level pressure (mslp) are used. TCs are identified in TempestExtremes by first identifying points that have a sea level pressure minima and upper level warm core. Second, candidate cyclones are checked for a closed contour

based on two criteria: 1) an increase in sea level pressure of at least 200 Pa from the center over a distance of 4° in great-circle distance and 2) the geopotential height differences between 300 and 500 hPa decreases by 6 m over a great-circle distance of 6.5°. Third, candidates are eliminated if the wind speed does not exceed 10 m s^{-1} within 15 h before or after the identified storm, or if they do not occur in one of the defined TC basins (see Fig. 1 in the online supplemental material). Tracks are then stitched together if there is less than 6° distance between candidates, with a minimum lifetime of 48 h and a maximum of 12 h between two points, and the first and last point are separated by at least 20°.

The skill in the Navy ESPC with regards to TC activity is examined using accumulated TC genesis and TC days over the Northern Hemisphere. TCs are only considered if they form and track within defined basins in the Northern Hemisphere: the Northern Indian Ocean (5°–25°N, 47.5°–100°E), northwest Pacific Ocean (5°–40°N, 102.5°E–180°), east-central Pacific (5°–25°N, 182.5°–280°E), and Atlantic Ocean (5°–40°N, 265°–352.5°E), which include only ocean grid points (supplemental Fig. 1). TC genesis is considered as the location and time when a track identified from TempestExtremes first occurs and when the wind speed exceeds 33 kt (1 kt ≈ 0.51 m s^{-1}). TCs present at the forecast initialization date are omitted to prevent counting a pre-existing storm as a genesis event. The accumulated TC

Total OLR Bias MJJASON

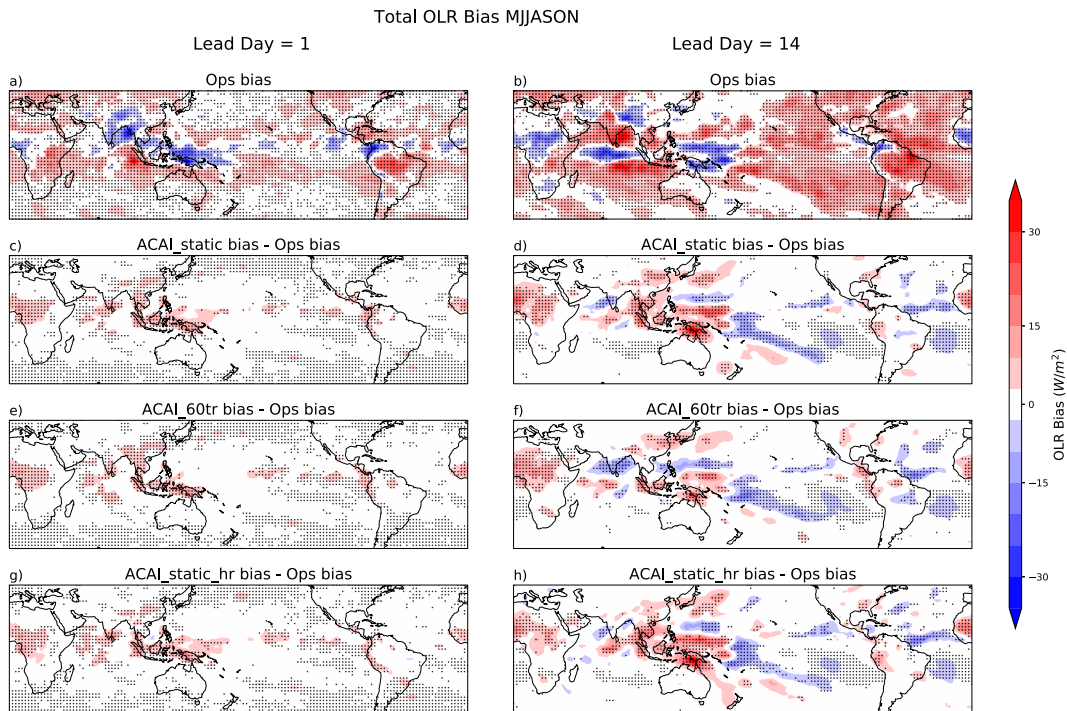Lead Day = 1   Lead Day = 14

FIG. 3. (a),(b) Total OLR anomaly bias (shading) during MJJASON for the Ops at lead days 1 and 14, respectively. (c),(d) Difference in the bias (relative to the Ops simulation) in ACAI_static at leads days 1 and 14, respectively. (e),(f) As in (c) and (d), but for ACAI_60tr. (g),(h) As in (c) and (d), but for ACAI_static_hr. Hatching indicates statistically significant differences at the 99.5% level.

genesis is calculated by first averaging the genesis events over all ensemble members, then taking the sum of the number of TC genesis events in the Northern Hemisphere TC basins. Next, the TC genesis is averaged over all the forecasts and cumulatively summed over lead day. TC days is calculated as all days along the track of TCs when the wind speed is greater than 33 kt. Similar to the accumulated TC genesis, an accumulated TC days is calculated by first taking the ensemble mean, then summing over the latitudes and longitudes in the TC basins in the Northern Hemisphere. The TC days are then averaged over all forecasts and cumulatively summed over lead day. The accumulated TC genesis and TC days is calculated for ease of interpretability.

3. Impact of ACAI on forecast skill

The addition of ACAI into the Navy ESPC improves the skill in tropical band OLR and u_{850} (Figs. 1 and 2, respectively). The RMSE in the ACAI simulations is reduced for OLR (~7%) and u_{850} (~12%) at shorter lead times (between lead times 5–10), converging with the Ops RMSE around lead day 20 for OLR and lead day 25 for u_{850}. ACAI also increases ensemble spread for u_{850} by ~15% between lead days 5–10, reducing to 0% around lead day 15 (Fig. 2d) and at shorter lead times for OLR (~5% between lead times 5–10), decreasing the spread beyond lead day 12. ACAI_static_hr shows the smallest increase in the spread compared to the other versions of ACAI and quickly becomes less dispersive than Ops. Inclusion of

ACAI reduces the bias magnitude for both OLR (~4%) and u_{850} (~5%) at all lead times. The low-level zonal winds are directly impacted by ACAI and it is therefore expected to see that there is an overall improvement in this field. More encouragingly, ACAI indirectly improves OLR, suggesting that reductions in the bias of fields directly impacted by ACAI result in improvements to the representation of convection in the model.

The pattern of the mean OLR bias in the Ops simulation (Ops − observations; Figs. 3a,b) shows, for lead day 1, a negative bias over the Bay of Bengal, South China Sea, and north of the Maritime Continent, and over portions of the tropical eastern Pacific and Colombia and Venezuela, and positive bias throughout the rest of the tropics, with the largest bias occurring in the Indian Ocean west of Indonesia and over the Amazon rain forest. By lead day 14, the positive OLR bias dominates most of the tropics, with negative biases in OLR growing in the Indian Ocean, west Pacific, and Africa. The ACAI biases relative to observations (not shown) look very similar to the Ops biases, such that it is difficult to visually identify the impact that ACAI has on the biases from these plots. Therefore, the biases in the ACAI simulations are examined relative to the Ops simulations (ACAI − Ops; Figs. 3c–h) to show the impact of ACAI on the model bias. The biases in all the ACAI simulations relative to the Ops simulation have similar spatial structures. At lead day 1 positive values are primarily in the regions where the bias in the Ops simulation are negative, indicating an improvement in the negative bias from ACAI. At lead day 14,
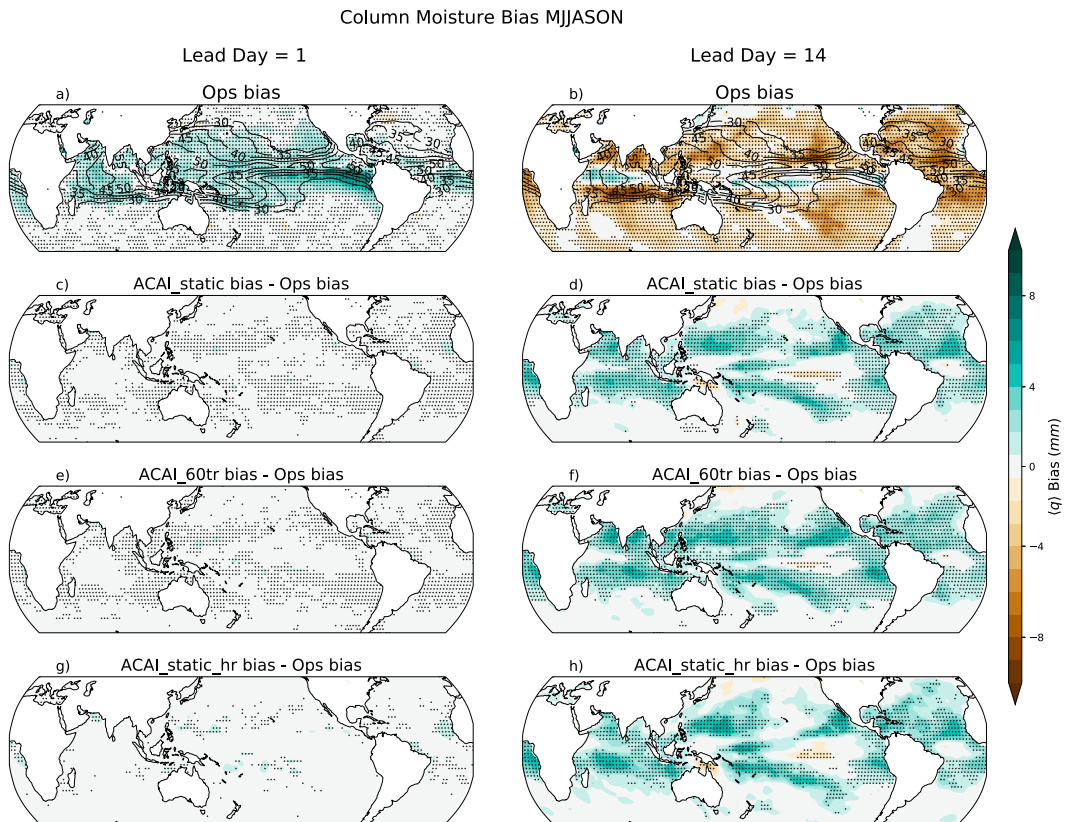
FIG. 4. (a),(b) Column integrated moisture bias (shading) during MJJASON for the Ops simulation at lead days 1 and 14. Contours in (a) and (b) are the observed column integrated moisture from ERA-interim. (c),(d) Differences in the bias (relative to the Ops simulation) in ACAI_static at lead day 1 and 14, respectively. (e),(f) As in (c) and (d), but for ACAI_60tr. (g),(h) As in (c) and (d), but for ACAI_static_hr. Hatching indicates statistically significant differences at the 99.5% level.

the magnitude of the ACAI biases relative to the Ops simulation have grown, and continue to show an opposite sign to the Ops bias. There are large positive differences over Africa and the Indo-Pacific warm pool that are opposite in sign to the Ops bias, indicating that ACAI is correctly accounting for the biases in these regions. At lead day 14 there is a reduction of the high OLR bias over the South Pacific convergence zone (SPCZ) in the ACAI simulations and some reduction in the high bias over the Atlantic Ocean. The differences between Ops and observations are statistically significant, at lead day 14 the largest changes in the biases from ACAI are statistically significant.

ACAI also reduces the mean state moisture biases (Fig. 4). The mass-weighted vertical integration of moisture is calculated as: $\langle q \rangle = -(1/g)\int_{P_{\text{sfc}}}^{P_T} q \, dP$, where g is the acceleration due to gravity, P is the pressure, and P_{sfc} and P_T are, respectively, the pressures at the surface and top of the troposphere, here defined as 100 hPa. In the Ops simulation (Figs. 4a,b), the initial moist bias at lead day 1 is followed by a dry bias at longer lead times, consistent with previous moisture biases seen in the Navy ESPC (Rushley et al. 2022). The ACAI runs show little difference in the bias from the Ops bias at lead day 1. At lead day 14, the difference in the bias in ACAI and Ops shows a statistically significant increase in the moisture over a broad

region off the equator, which contributes to reducing the large dry bias in the Ops simulation off the equator (Fig. 4b). Thus, while the dry bias still exists in the ACAI versions of the Navy ESPC (not shown), the magnitude of the bias is reduced as compared to the Ops simulation.

Although ACAI does not impact the SSTs directly (i.e., the ACAI perturbations are only applied to the atmospheric component of the ESPC), the ACAI simulations show some structural differences in the SST biases compared to Ops (Fig. 5). While the differences in the bias in the ACAI and Ops simulations are small at lead day 1, by lead day 14 some notable differences in the bias appear. Even small changes in the ACAI simulation at lead day 14 are statistically significant at the 99.5% level. In the Ops simulations at lead day 14 there is a warm bias throughout most of the tropical oceans, with the largest biases occurring in the Indian Ocean, Maritime Continent, Atlantic Ocean, and ENSO region. ACAI reduces the SST biases in the equatorial eastern Pacific in the ENSO region; however, the ACAI simulations have a cold bias in the southeastern Pacific Ocean, south of the main ENSO region, indicative of an overcorrection of the bias in this region, which leads to a net cold bias in ACAI. The ACAI simulations also contribute to lowering the positive SST bias in the
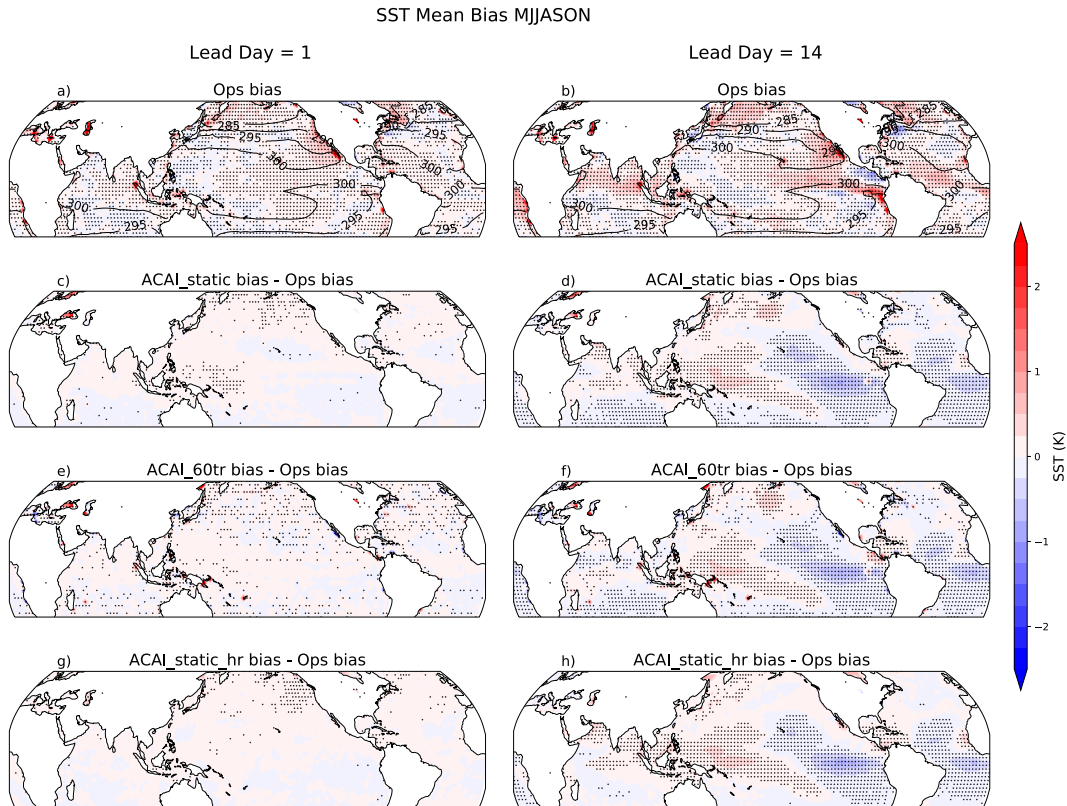
SST Mean Bias MJJASON

FIG. 5. (a),(b) Mean SST bias (shading) during MJJASON for the Ops at lead days 1 and 14, respectively. The observed pattern of SSTs is shown in the contours in (a) and (b). (c),(d) Differences in the bias (relative to the Ops simulation) in ACAI_static at lead days 1 and 14, respectively. (e),(f) As in (c) and (d), but for ACAI_60tr. (g),(h) As in (c) and (d), but for ACAI_static_hr. Hatching indicates statistically significant differences at the 99.5% level.

Ops simulation in the Atlantic Ocean at lead day 14. In the Indian Ocean, Indo-Pacific warm pool, and west Pacific ACAI is warmer relative to Ops, indicating that it is contributing to the warm biases in these regions.

There is a clear reduction in the RMSE of the MJO-filtered OLR and u_{850} in the ACAI simulations (Fig. 6), the bias is similarly reduced (not shown). The RMSE is reduced more at the beginning of the forecast (~10% reduction) for both MJO-filtered OLR and u_{850} and has a more moderate decrease in the RMSE by the end of the forecast (~5% reduction). There is a small increase in the ensemble spread for the MJO filtered u_{850} for all the ACAI runs over the Ops simulation, with ACAI_static showing the largest increase in the spread (~7.5% at early lead times, decreasing to 5% by day 12). The impact of ACAI on MJO-filtered OLR spread is less definitive, with an increase of less than 4% for ACAI_static and ACAI_60tr that goes to 0% midway through the forecast period, and a decrease in ACAI_static_hr of 2% at the beginning of the forecast that evolves into a decreases of 8% by the end of the forecast period. This improvement in the MJO skill is encouraging for the potential predictability in the S2S time scale in the ACAI simulations. The relatively large value of RMSE at lead day 1 and the decrease in the RMSE beyond lead day 25 are artifacts of the FilPad filtering method,

which includes some of the observations at shorter lead times and the zeros padding at longer lead times. While this structure of the RMSE development with lead time is an artifact of the FilPad filtering method (Janiga et al. 2018), it does not affect the magnitude of the reduction of the RMSE in the ACAI simulations. It is found that the phase speed of the MJO in the different versions of ACAI does not change (supplemental Fig. S2). While the decrease in the RMSE and biases in the MJO are encouraging, it would require a longer dataset to fully understand how ACAI affects the processes responsible for the maintenance of the MJO, such as the vertical moisture advection, as was determined by Rushley et al. (2022) to be a key driver of the amplitude bias in the version of the Navy ESPC used for the Subseasonal Experiment (SubX).

4. Tropical cyclone environment

Despite the improvements in the MJO skill, total OLR, and u_{850} fields, there is a degradation in the skill in the GPI in the ACAI runs in terms of ensemble mean RSME and bias magnitude (Fig. 7). GPI combines the contributions from the low-level absolute vorticity, vertical wind shear, PI, and RH, all of which incorporate variables directly perturbed by the ACAI method through the horizontal wind speed, surface pressure,
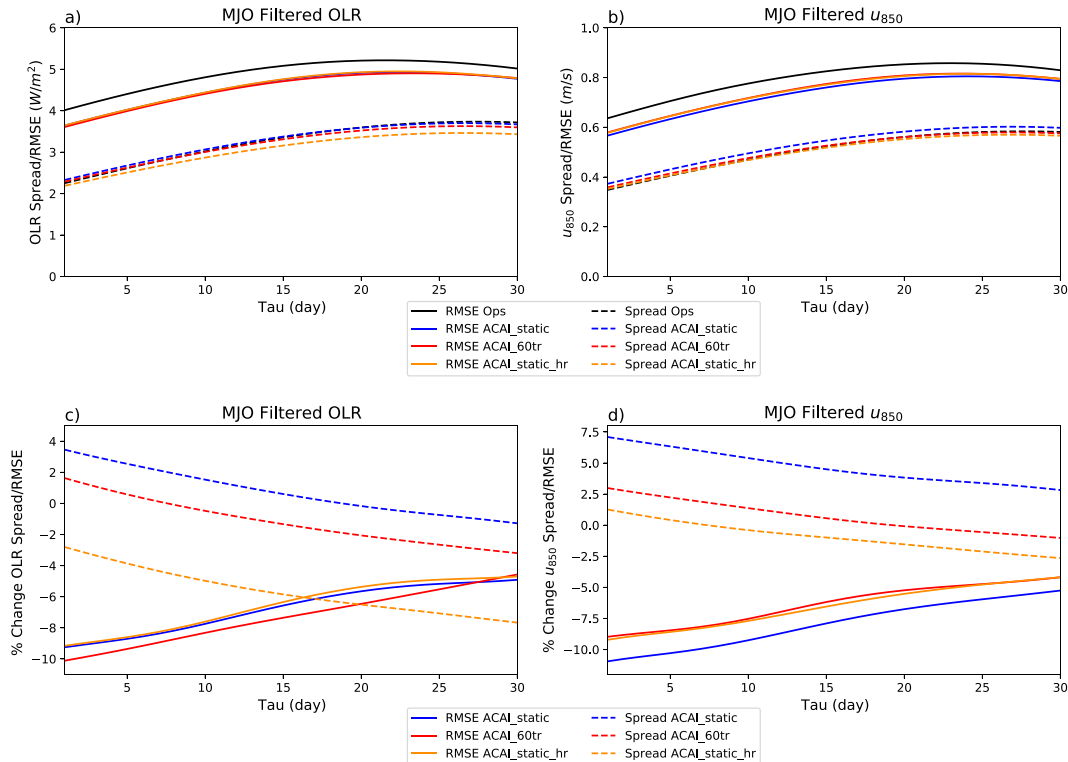
FIG. 6. Ensemble mean RMSE (solid lines) and ensemble spread (dashed lines) of MJO-filtered (a) OLR and (b) u_{850} for the tropical band (30°S–30°N) during MJJASON. Black lines indicate the operational model (Ops), colored lines are the different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold). Percent changes relative to Ops are shown for RMSE (solid lines) and spread (dashed lines) for (c) MJO-filtered OLR and (d) MJO filtered u_{850}, colored lines are the percent changes from different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold).

temperature, and/or moisture. Despite this, there is a large increase in the RMSE and bias at all lead times in the GPI. The RMSE and bias increase in the ACAI versions by 5%–15% and 10%, respectively. There is an improvement in the ensemble spread in the GPI of about 20% between lead days 5–10 that decreases as the lead time increases.

It is of interest to look at the spatial distribution of the GPI, as the absolute value of the bias and RMSE does not inform us as to the direction of the bias, i.e., if the GPI is increased or decreased in the ACAI runs. In the Ops simulation, the dominant GPI biases are positive and statistically significant in the Indo-Pacific warm pool, eastern Pacific, and Atlantic Oceans (Figs. 8a,b). These positive GPI biases are accentuated in the ACAI run (Figs. 8c–h), indicating that the ACAI method produces an environment that is even more favorable for TCs than the Ops simulation. In many regions at lead day 14, these increases in the biases due to ACAI are statistically significant, with the most significant changes occurring in the western Pacific, tropical eastern Pacific, and Atlantic, and Indian Oceans.

As previously mentioned, GPI combines the environmental variables of absolute vorticity, vertical wind shear, RH, and PI, all of which are composed of variables directly perturbed by of ACAI. To understand the relative importance of these

variables in the GPI errors, each term is examined individually. Figure 9 shows the RMSE, ensemble spread and bias for the four terms in the GPI. There is a modest improvement in the ensemble mean RMSE, spread and bias of the low-level absolute vorticity at shorter lead times, but the ACAI simulations perform similarly to Ops at long lead times (Figs. 9a,b). On average there is a 5% reduction in the RMSE, a 2% reduction in the bias, and a 5%–20% increase in the ensemble spread in the absolute vorticity (See supplemental Fig. 3). There is a reduction in the bias at all lead times in the vertical wind shear, with a reduction in the RMSE out to lead day 25, and some increase in the ensemble spread at shorter lead times (Figs. 9c,d). On average there is an 8% reduction in the RMSE, a 5% reduction in the bias, and a 5%–30% increase in the ensemble spread in the vertical wind shear (see supplemental Fig. 3). There is degradation in the skill of the RH, associated with an increase in the RMSE and bias, particularly at longer lead times, with some improvement in the ensemble spread at shorter lead times (Figs. 9e,f). The RMSE of the RH increases approximately 6%–8% for the ACAI simulations, with the ACAI_static_hr performing the best of the three ACAI simulations, the bias is similarly increased at the beginning of the forecast period (7%–9%). The ensemble spread increases by ~20% at the beginning of the forecast

GPI Anomalies MJJASON

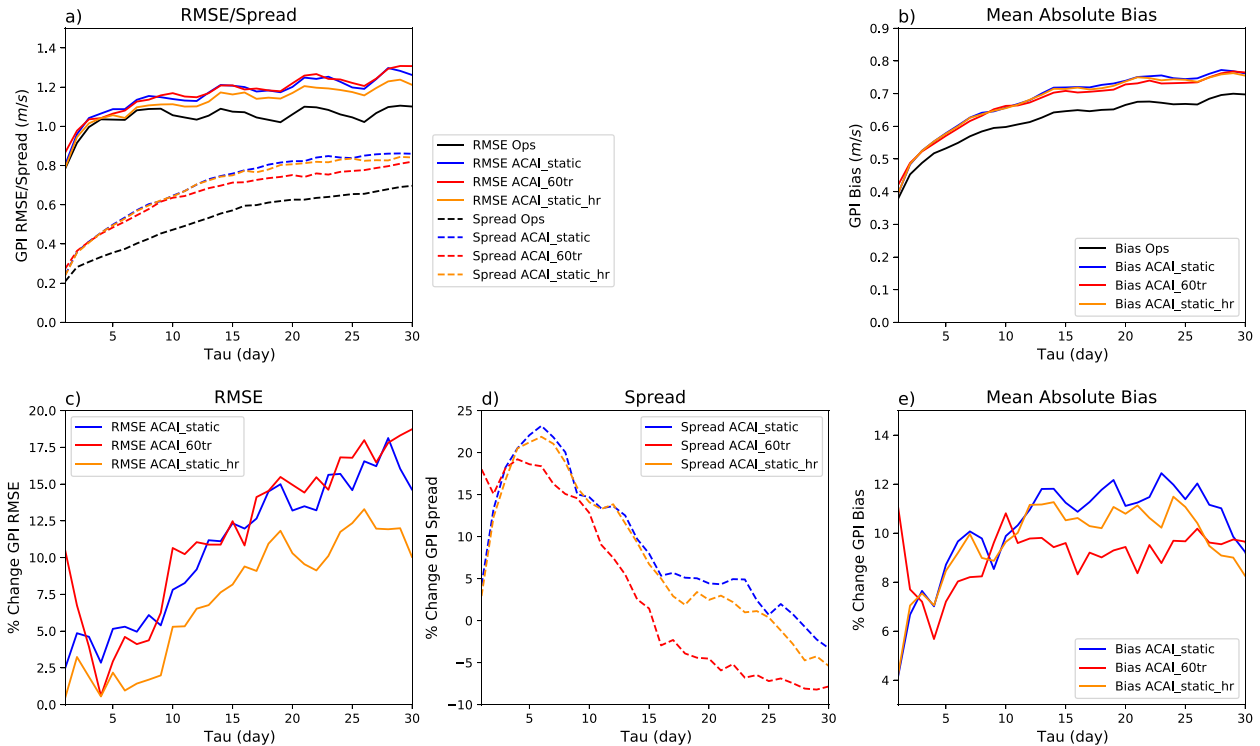FIG. 7. (a) Ensemble mean RMSE (solid lines) and ensemble spread (dashed lines) and (b) mean absolute bias of GPI for the tropical band (30°S–30°N) during MJJASON. Black lines indicate the operational model (Ops), colored lines are the different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold). The percent changes to (c) RMSE, (d) spread, and (e) bias relative to OPS. The colored lines are the percent differences in the different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold).

period, but decreases to less than 5% near the end of the forecast period (Figs. 10a–c). Last, the PI shows small changes in the ensemble spread, but a clear decrease in the RMSE and bias, especially at longer lead times (Figs. 9g,h). The RMSE and bias in the PI decrease up to 8% by the end of the forecast period with the ACAI_static showing the largest reduction in the RMSE and bias (Figs. 10d,f). The ensemble spread in the PI also increases, maximizing around lead day 5–10. ACAI_static shows the smallest increase in the spread compared to the other versions of ACAI (Fig. 10e). All variables show an increase in the ensemble spread that maximizes between lead day 5–10.

Despite the large increases in the bias and RMSE in the GPI, most of the individual components that contribute to the GPI calculation (with the exception of RH) are improved by the addition of ACAI. To test if there were changes to the contributions of each variable in the GPI in ACAI, the GPI was recalculated following the component method of substituting three of the forecast variables in the GPI calculation with the three corresponding reanalysis fields (Camargo et al. 2007; Reynolds et al. 2022). The contributions to the RMSE, bias, and spread is (in order from largest to smallest) the absolute vorticity, RH, vertical shear, and PI; this relationship does not change in the ACAI runs,

suggesting that there is no change in the overall contributions to the GPI from its components (not shown). The increase in the RMSE and bias in the GPI is likely tied to the increase in the bias and RMSE in the RH. Therefore, the spatial distribution of RH is closely examined. (See supplemental Figs. 4 and 5 for additional discussion on other variables).

The increase in the bias and RMSE of the RH is driven by a positive difference in the bias in the 700-hPa RH throughout the tropics in the ACAI simulations (Fig. 11). Similar to the column moisture, there is a large moist bias in 700-hPa RH at lead day 1 in the Ops simulation, then some drying out of the environment by lead day 14. Changes in the Ops simulation are statistically significant at the 99.5% level. Changes from ACAI are small at lead day 1, but by lead day 14 ACAI adjusts the RH into a moister regime (Figs. 11c–h). These changes in the bias due to ACAI are statistically significant. This partly accounts for the dry bias in the Ops simulations at 700 hPa, but the ACAI simulations over correct into a too moist regime, similar to the overcorrections seen in previous studies of ACAI (Crawford et al. 2020; Reynolds et al. 2022). As the ACAI simulation reduces the dry bias in the column moisture in the Navy ESPC (Fig. 4), some difference in the vertical profile of the biases in moisture or temperature lead

Genesis Potential Index Bias MJJASON

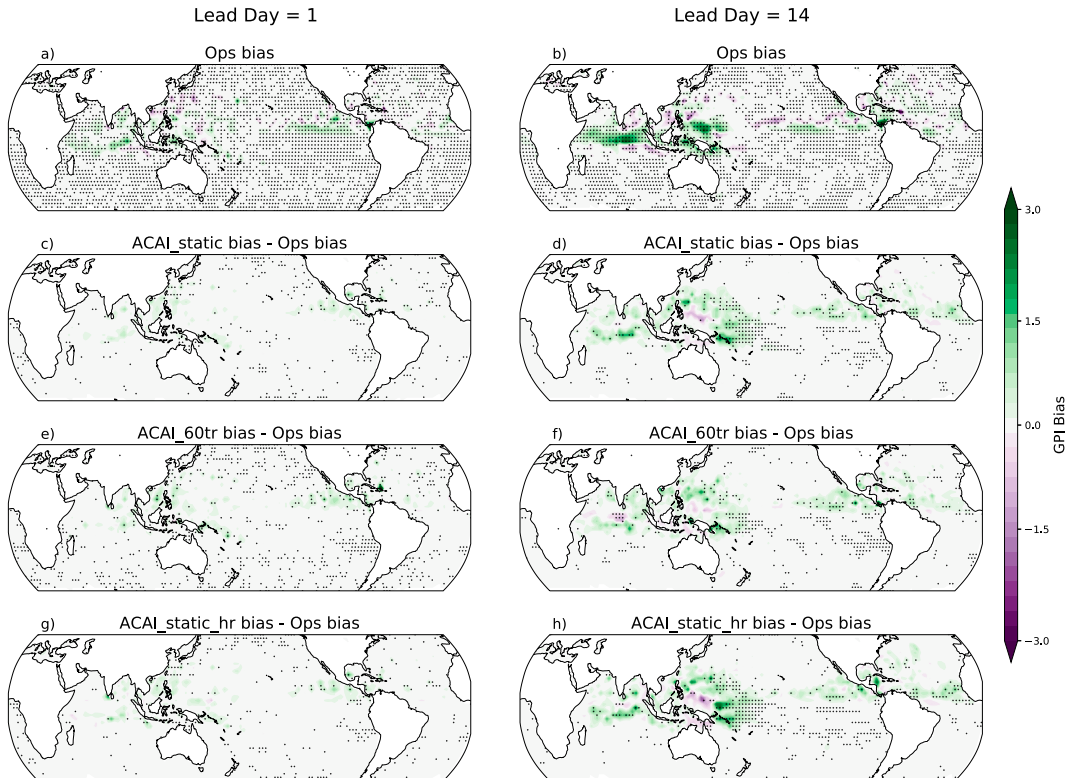Lead Day = 1   Lead Day = 14

FIG. 8. (a),(b) The bias in the GPI for the Ops simulation at lead day 1 and 14, respectively. (c),(d) Differences in the bias (relative to the Ops simulation) in ACAI_static, at lead day 1 and 14, respectively. (e),(f) As in (c) and (d), but for ACAI_60tr. (g),(h) As in (c) and (d), but for ACAI_static_hr. Hatching indicates statistically significant differences at the 99.5% level.

to an overall too high RH in the ACAI runs. The high RH bias in the ACAI simulations contributes to a more favorable (moister) environment for TC genesis. This is consistent with Li et al. (2022), who also found that positive RH biases were the dominant factor in GPI biases in the Community Earth System Model version 2 (CESM2).

As both the moisture and temperature are affected by ACAI, it is of interest to look at each component in regard to the RH bias. The spatial bias maps of 700-hPa specific humidity and temperature are shown in supplemental Figs. 6 and 7. The pattern of the specific humidity bias well matches that of the RH bias, while there is a widespread cold bias throughout the tropics at 700 hPa with no consistent pattern with the RH biases, therefore it likely dominated by the moisture biases. The vertical profiles of the moisture and temperature biases in the Northern Hemisphere (Fig. 12) show that there is not a consistent bias throughout the depth of the troposphere for either of these variables. Early in the forecast, the moisture bias is very similar in the ACAI simulations and Ops, with a large amplitude moist bias in the boundary layer, a large dry bias around 900 hPa, and a moist bias around 800–700 hPa. The temperature bias at lead day 1 shows similar biases in Ops and ACAI_static and ACAI_60tr, ACAI_static_hr is in

general warmer than the other simulations, potentially related to the high resolution ocean in this simulation. At lead day 14, the biases in the moisture and temperature have increased for all simulations at most levels. There is a persistent moist bias below 900 hPa and a very dry bias at 900 hPa. This dry bias is consistent with previous versions of the Navy ESPC (Rushley et al. 2022). The ACAI simulations correct the dry bias that develops in lead day 14 in the Ops simulations between 800 and 700 hPa. The Ops simulation is largely dry above 900 hPa at lead day 14. This moist bias at 700 hPa has a strong impact on the relative humidity which is calculated at that level in the GPI. At lead day 14, a cold bias appears in all simulations above about 800 hPa, largest in Ops, with ACAI improving the bias in the mid-to-upper atmosphere. ACAI_static_hr is warmer than the other simulations at almost all levels, which improves the bias in the mid-to-upper atmosphere, but increases the bias below 800 hPa. The other ACAI simulations are consistent with the bias in the Ops simulation in the lower atmosphere.

At 700 hPa there is a local maximum in the cold bias in all simulations. This cold bias, paired with the positive moisture bias at 700 hPa at lead day 1, and near-zero moisture bias at lead day 14, results in an increase in RH in the ACAI
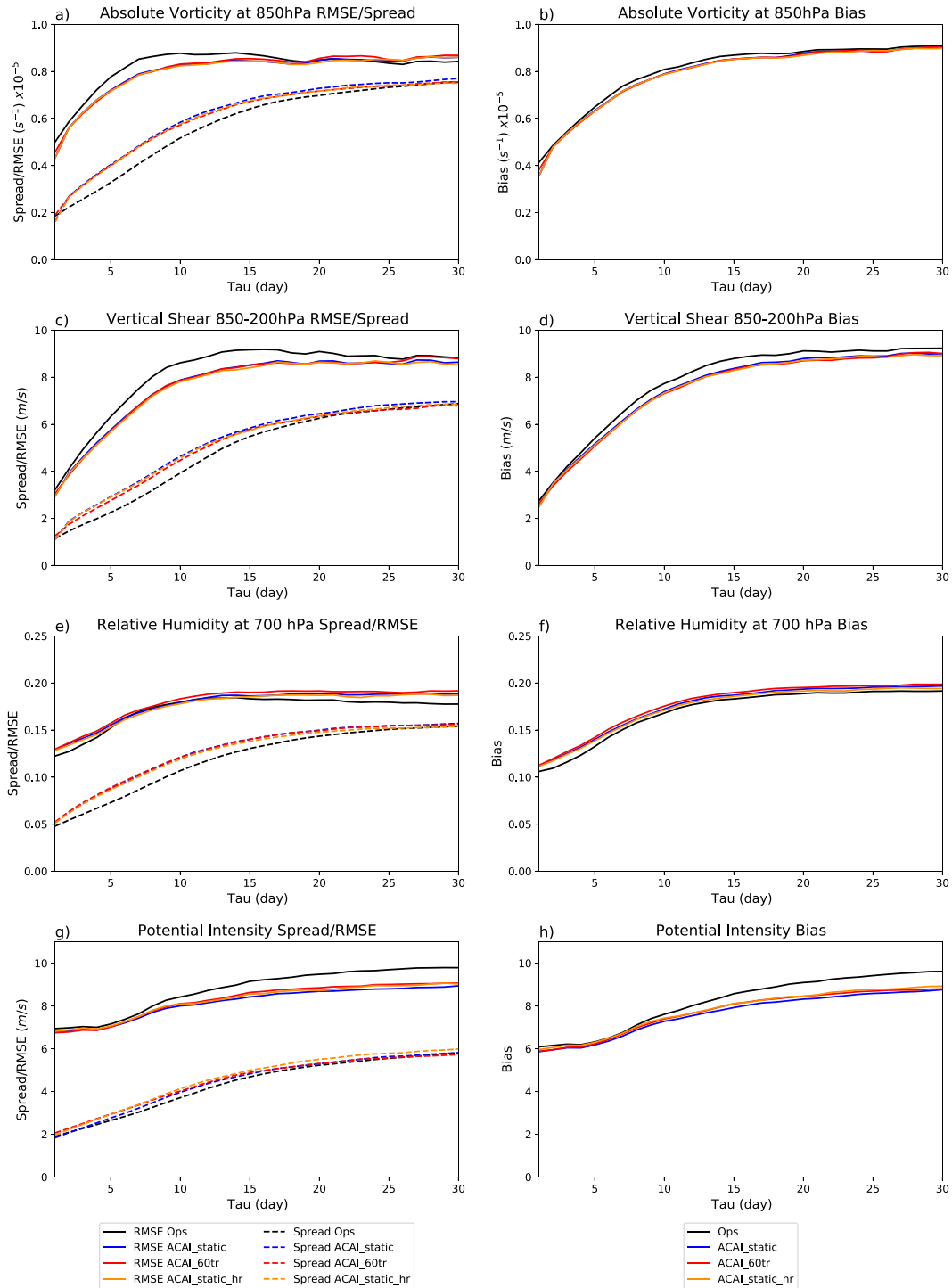
FIG. 9. Ensemble mean RMSE (solid lines) and ensemble spread (dashed lines) for (a) absolute vorticity at 850 hPa scaled by 10^{-5}, (c) vertical shear from 850 to 200 hPa, (e) relative humidity at 700 hPa, and (g) potential intensity over the tropical band (30°S–30°N) during MJJASON. The mean absolute bias of (b) absolute vorticity at 850 hPa scaled by 10^{-5}, (d) vertical shear from 850 to 200 hPa, (f) relative humidity at 700 hPa, and (h) potential intensity over the tropical band (30°S–30°N) during MJJASON. On all panels black lines indicate the operational model (Ops), and colored lines are the different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold).
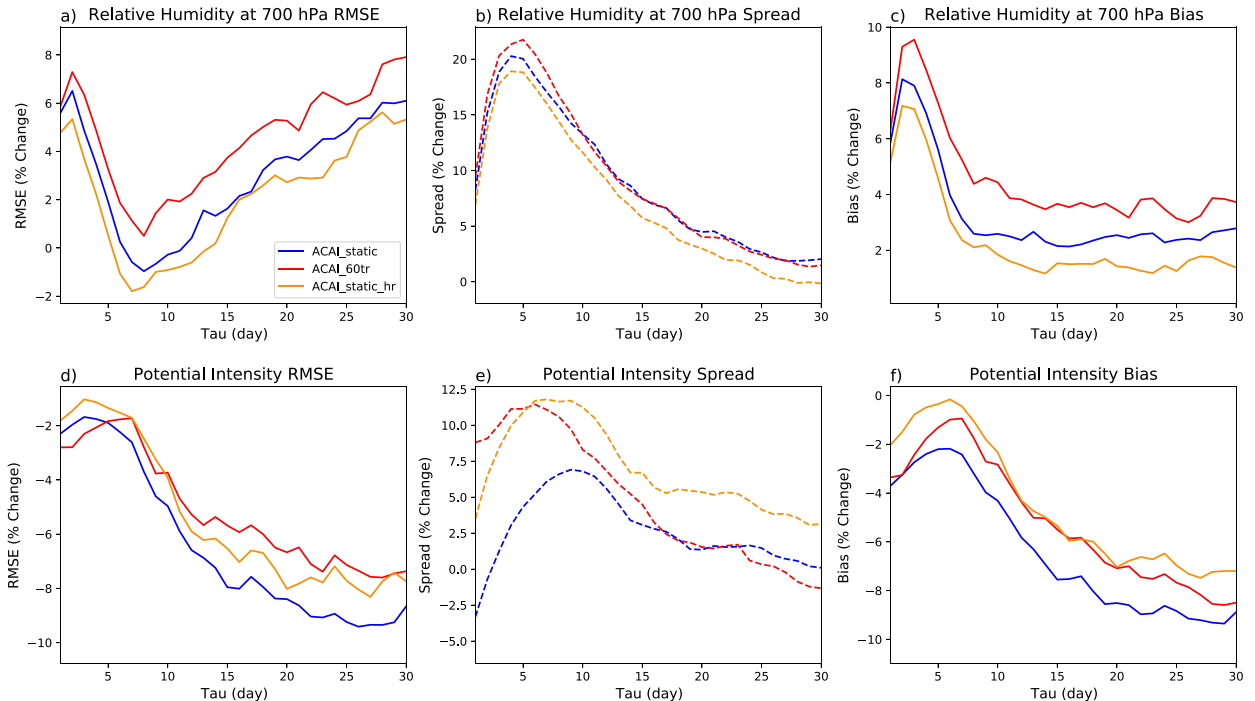
FIG. 10. Percent changes relative to Ops in the RMSE for (a) relative humidity at 700 hPa and (d) potential intensity. Percent changes in ensemble spread relative to the Ops simulation (dashed lines) for (b) relative humidity at 700 hPa and (e) potential intensity. Percent changes in mean absolute bias relative to the Ops simulation for (c) relative humidity at 700 hPa and (f) potential intensity. Percent changes are calculated over the tropical band (30°S–30°N) during MJJASON. Colored lines are the percent changes from Ops for different versions of ACAI: ACAI_static (blue), ACAI_60tr (red), and ACAI_static_hr (gold).

simulations. While outside the scope of this study, it would be of interest to examine the role that the high resolution ocean plays in the ACAI_static_hr column temperature and moisture biases, despite the fact that the SST fields look very similar in all the ACAI simulations (Fig. 5).

The increase in the GPI bias suggests that the environment in the Navy ESPC is too favorable for TC formation. Therefore, the TC activity in the Northern Hemisphere in the ACAI simulations is examined. The accumulated TC genesis in the simulations increases at a similar but slightly slower rate as is seen in IBTrACS (Fig. 13a). While there is some improvement in the total number of genesis events from the ACAI runs, particularly for ACAI-static, all simulations slightly underestimate the total number of TC genesis events. In contrast, the accumulated TC days are overestimated in the simulations as compared to IBTrACS (Fig. 12b). This in tandem with the underestimation of TC genesis events in the Navy ESPC indicates that TCs in the Navy ESPC last longer than in IBTrACS. A detailed analysis of TC intensity and structure in the Navy ESPC and ACAI would likely require a higher a higher resolution to resolve the inner core and a larger sample size of TCs; therefore, we focus on genesis rather than intensity of TCs. To make a fair comparison between the Navy ESPC and IBTrACS, the observed data are selected to match each forecast. In this method, duplicate TC counts (e.g., for the same genesis event that occurs in subsequent forecasts) are treated the same way in the forecasts and

observations, therefore, we do not expect that this would change our conclusions in Fig. 12.

Despite the overestimation of the GPI, the TC genesis remains too low in the Navy ESPC ACAI versions. It is worth noting that the increase in GPI in the ACAI simulations also corresponds to an increase in TC genesis events (Fig. 13). However, this may be evidence that there is some compensating factor in the Navy ESPC that is damping TC formation, while the GPI is overestimated. This may be related to a location disconnect between the regions where GPI is overestimated and where TCs tend to form in the Navy ESPC, for instance TC genesis events in the northwest Pacific in the Navy ESPC tend to form too far east of the Philippines (See supplemental Fig. 8), which is coincidentally a region of negative GPI bias in the ACAI simulations (Fig. 8). The GPI and TC genesis spatial patterns are generally consistent, with a correlation coefficient of 0.47 for observations and 0.6 for the Navy ESPC simulations. However, there remain many regions where the GPI may not well explain the simulated TC genesis (or lack thereof), such as the tropical Atlantic. The disconnect between the decrease in the number of TC genesis events and the increase in the GPI highlights a potential limitation in using GPI to predict TC genesis.

5. Summary and conclusions

In this study, the effect of ACAI on the MJO and TCs in the Navy ESPC model have been explored. ACAI is an in-line bias

RH 700hpa Mean Bias MJJASON

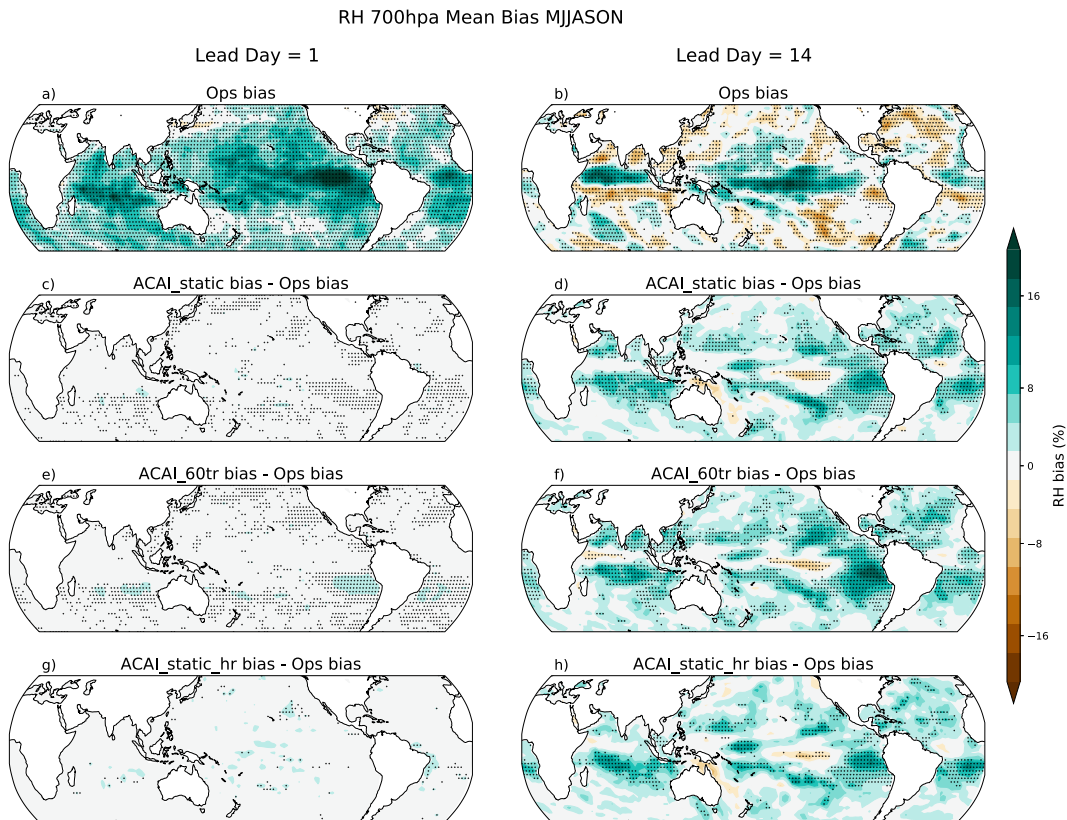Lead Day = 1   Lead Day = 14

FIG. 11. (a),(b) Mean RH bias (shading) at 700 hPa during MJJASON for the Ops simulation at lead day 1 and 14, respectively. (c),(d) Differences in the bias (relative to the Ops simulation) in ACAI_static, at lead day 1 and 14, respectively. (e),(f) As in (c) and (d), but for ACAI_60tr. (g),(h) As in (c) and (d), but for ACAI_static_hr. Hatching indicates statistically significant differences at the 99.5% level.

correction and stochastic forcing method that is designed to reduce systematic errors and account for model uncertainty. The ACAI tendency perturbations are applied to the tendency terms of moisture, horizontal winds, temperature, and surface pressure at every model grid point and forecast time step. Three versions of ACAI are compared to the Operational version (Ops) of the Navy ESPC (Table 1). The differences are that the ACAI_static and ACAI_static_hr simulations uses a static archive of analysis increments from 2011 to get the tendency perturbation, while ACAI_60tr uses a 60-day trailing window from a running archive of analysis increments to get the tendency perturbations. Additionally, the ACAI_static and ACAI_60tr are run with a lower resolution 1/4° HYCOM grid, while ACAI_static_hr is run using a high resolution 1/12° HYCOM grid.

An examination of ensemble forecast performance shows that the inclusion of ACAI reduces the ensemble mean RMSE and bias in the tropical band OLR and u_{850} fields as well as increases the ensemble spread, therefore reducing ensemble underdispersion. This is encouraging, as while the zonal winds are directly impacted by ACAI, the OLR is not, showing evidence that ACAI is beneficial for the representation of convection in the Navy ESPC in most regions. The addition of ACAI improves the mean state moisture by reducing the dry bias seen in the Ops simulation over many regions, although it increases a moist biases in

some areas of the tropics as well. ACAI increases the magnitude of a cold bias over the southeastern Pacific, which may be related to changes in the near-surface winds. Notably, the ACAI simulations show an improvement in the skill of the MJO scale OLR and u_{850}.

Despite these improvements in OLR and winds, and the encouraging improvements in the MJO from ACAI, the ACAI simulations show some degradation in the environment in which TCs form. We examine the GPI, which combines the influence of low-level absolute vorticity at 850 hPa, vertical wind shear from 200 to 850 hPa, PI, and RH at 700 hPa. It is found that there is an increase in the RMSE and bias in the GPI in the ACAI simulations, as well as an improvement in the ensemble spread. This is due to a positive GPI bias in the Navy ESPC that is enhanced through the addition of ACAI, indicating that the Navy ESPC Ops simulation has an overly favorable environment for TC formation, made worse by the addition of ACAI.

The performance of the components of the GPI were examined to attempt to identify which terms are contributing to the increase in the bias and RMSE of the GPI. There is an improvement in the model skill in the PI, absolute vorticity, and vertical shear in the ACAI models. The 700-hPa RH term, however, showed an increase in the bias in the ACAI versions. The positive GPI bias is therefore likely linked to the increase in the RH

Mean Vertical Moisture Biases
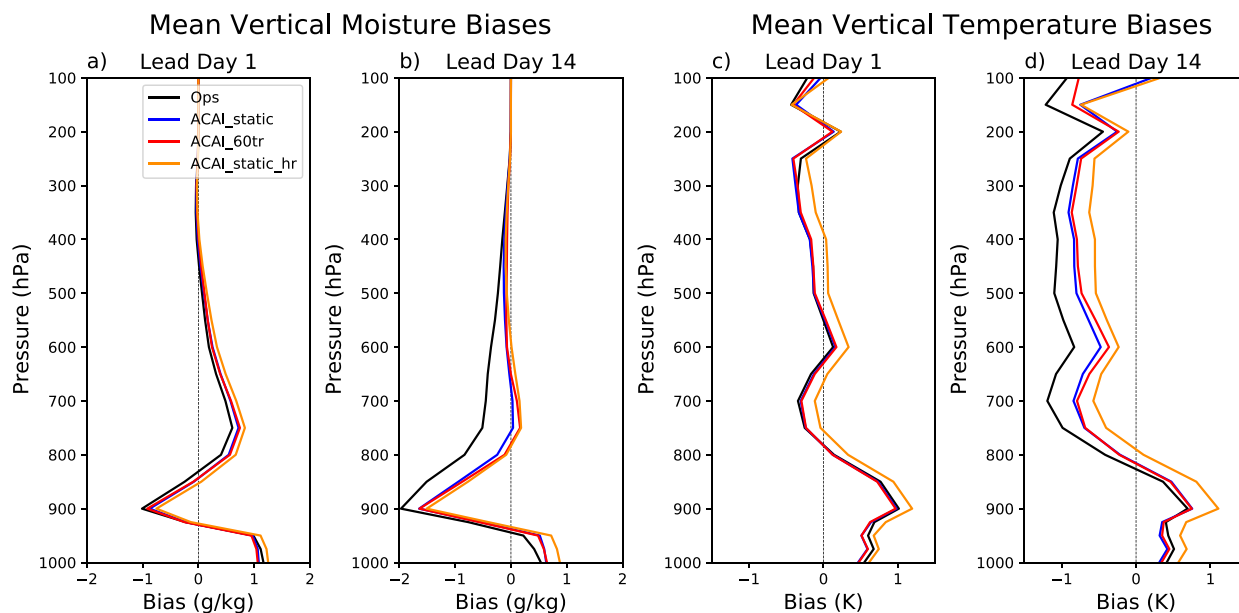## Mean Vertical Temperature Biases

FIG. 12. Vertical profiles of the moisture bias at (a) lead day 1 and (b) lead day 14. Vertical profiles of the temperature bias at (c) lead day 1 and (d) lead day 14. The vertical profiles are averaged over the Northern Hemisphere TC basins for the Ops (black line), ACAI_static (blue line), ACAI_60tr (red line), and ACAI_static_hr (gold line).

bias at 700 hPa, with this moister middle troposphere supporting TC formation. While ACAI does well at reducing the dry bias in the Navy ESPC in the column, the RH at 700 hPa shows a large positive bias, upward of 15% in the deep tropics. The pattern of the 700-hPa moisture bias in the ACAI versions matches well the pattern of the RH bias. There is likely some contribution from the cold bias in 700-hPa temperature, but the spatial pattern of specific humidity biases matches well the pattern of the biases seen in RH, suggesting that the dominating source of the RH biases is the specific humidity. There may also be some compensating errors that ACAI corrects that influence

the GPI. For instance, in the Ops simulation the western Pacific is too moist, which would increase the GPI; however, there is too-strong vertical shear in this simulation, which would reduce the GPI (supplemental Figs. 5 and 6). The ACAI simulations reduce the vertical shear bias, which would also support an increase in the GPI.

Last, we examined the TC activity in the Navy ESPC simulations using TempestExtremes (Ullrich and Zarzycki 2017; Ullrich et al. 2021) to obtain the TC genesis and TC days in the Navy ESPC. Compared to IBTrACS, the Ops and ACAI simulations reproduce the observed accumulated TC genesis
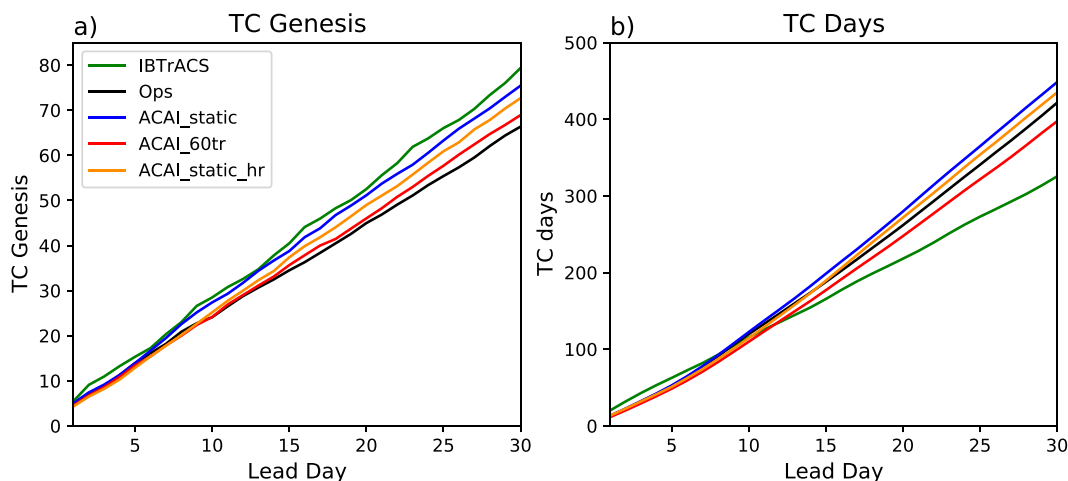

FIG. 13. Accumulated TC genesis (a) and TC days (b) over the Northern Hemisphere TC basins for observed TCs from IBTrACS (green lines), the Ops (black lines), ACAI_static (blue lines), ACAI_60tr (red lines), and ACAI_static_hr (gold lines).

and TC days fairly well. While the total number of TC genesis events in the Navy ESPC is lower than observed, the ACAI simulation shows a small increase (improvement) in the number of genesis events. This is consistent with the increase in the GPI seen in the ACAI simulations, such that the environment in the Navy ESPC supports TC formation, encouraging TC genesis. There is an inconsistency disconnect between the overestimation of GPI and the underestimation of TC genesis events in Ops and ACAI. This may be due to a location disconnect between TC formation and where GPI is too large in the Navy ESPC. Accumulated TC days is overestimated in the Navy ESPC in both the Ops and ACAI versions. The discrepancy between the accumulated TC days and TC genesis suggest some behavioral differences in the TCs produced in the Navy ESPC, such that TCs may be longer lived in the Navy ESPC. A detailed analysis into the structure of TCs in the Navy ESPC ACAI version is warranted, but would likely require a higher resolution to resolve the inner core as well as a larger sample size of TCs to form robust conclusions, and is thus beyond the scope of this study.

For several forecast metrics considered here, there is a clear separation between the ACAI simulations and Ops (e.g., Figs. 1, 2, and 6). However, for most of these metrics, there is not much difference in the skill between the different versions of ACAI, indicating little sensitivity of the forecast improvements to the specifics of the ACAI implementation or ocean resolution. This does not indicate that ocean resolution is unimportant in the modeling of the Earth system. In the context of the ACAI simulations, this lack of impact from ocean resolution is likely due to the fact that the ACAI method does not apply any perturbations directly to the ocean, although there is some indirect influence on the ocean through surface fluxes and due to the fact that the ocean model is coupled to the atmosphere. Additionally, Reynolds et al. (2022) note that the perturbation method using the static archive of analysis increments in 2011 may be suboptimal due to annual variability, such as ENSO, being in different phases in the sampled year versus the simulation year (2017). The results shown here indicate that, for 2017 at least, this is not a large problem, as the improvements in the ACAI_static and ACAI_static_hr simulations are similar (sometimes better) than the improvements in the simulations which use the 60-day trailing analysis increments (ACAI_60tr).

Overall, there is a promising improvement in the bias, RMSE, and spread in many fields relevant to subseasonal tropical prediction with the addition of ACAI into the Navy ESPC. Similar to previous works, the sign of the bias and the spatial pattern of the bias shows that the improvements from ACAI are not universal and that there are areas of degradation of the forecast, particularly at longer lead times (Crawford et al. 2020; Reynolds et al. 2022). It is of interest to examine which of the ACAI tendency terms contributes to these changes. However, as ACAI is run in-line, the changes in the tendency terms interact and evolve to impact a wide range of different aspects of the model state, making it difficult to identify the impact that one particular tendency term has on the simulations. Regardless, with the addition of ACAI, there is an improvement in the absolute vorticity, vertical shear, and potential intensity.

Furthermore, ACAI is shown to improve the MJO's prediction skill in the Navy ESPC, which warrants more detailed examination of the role that ACAI plays in improving the mean state.

REFERENCES

Aiyyer, A., and J. Molinari, 2008: MJO and tropical cyclogenesis in the Gulf of Mexico and eastern Pacific: Case study and idealized numerical modeling. *J. Atmos. Sci.*, **65**, 2691–2704, https://doi.org/10.1175/2007JAS2348.1.

Barton, N., and Coauthors, 2021: The Navy's Earth system prediction capability: A new global coupled atmosphere-ocean-sea ice prediction system designed for daily to subseasonal forecasting. *Earth Space Sci.*, **8**, e2020EA001199, https://doi.org/10.1029/2020EA001199.

Batté, L., and M. Déqué, 2016: Randomly correcting model errors in the ARPEGE-climate v6.1 component of the CNRM-CM: Applications for seasonal forecasts. *Geosci. Model Dev.*, **9**, 2055–2076, https://doi.org/10.5194/gmd-9-2055-2016.

Bender, M. A., and I. Ginis, 2000: Real-case simulations of hurricane–ocean interaction using a high-resolution coupled model: Effects on hurricane intensity. *Mon. Wea. Rev.*, **128**, 917–946, https://doi.org/10.1175/1520-0493(2000)128<0917:RCSOHO>2.0.CO;2.

Berner, J., and Coauthors, 2017: Stochastic parameterization: Toward a new view of weather and climate models. *Bull. Amer. Meteor. Soc.*, **98**, 565–588, https://doi.org/10.1175/BAMS-D-15-00268.1.

Bhargava, K., E. Kalnay, J. A. Carton, and F. Yang, 2018: Estimation of systematic errors in the GFS using analysis increments. *J. Geophys. Res. Atmos.*, **123**, 1626–1637, https://doi.org/10.1002/2017JD027423.

Bister, M., and K. A. Emanuel, 2002: Low frequency variability of tropical cyclone potential intensity 1. Interannual to interdecadal variability. *J. Geophys. Res.*, **107**, 4801, https://doi.org/10.1029/2001JD000776.

Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, **4**, 55–88, https://doi.org/10.1016/S1463-5003(01)00012-9.

Bowler, N. E., and Coauthors, 2017: Inflation and localization tests in the development of an ensemble of 4D-ensemble variational assimilations. *Quart. J. Roy. Meteor. Soc.*, **143**, 1281–1302, https://doi.org/10.1002/qj.3004.

Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, https://doi.org/10.1002/qj.49712556006.

Camargo, S. J., A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007: Tropical cyclone genesis potential index in climate models. *Tellus*, **59A**, 428–443, https://doi.org/10.1111/j.1600-0870.2007.00238.x.

——, M. C. Wheeler, and A. H. Sobel, 2009: Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *J. Atmos. Sci.*, **66**, 3061–3074, https://doi.org/10.1175/2009JAS3101.1.

Cavallo, S. M., J. Berner, and C. Snyder, 2016: Diagnosing model errors from time-averaged tendencies in the Weather Research and Forecasting (WRF) Model. *Mon. Wea. Rev.*, **144**, 759–779, https://doi.org/10.1175/MWR-D-15-0120.1.

Chang, Y., S. D. Schubert, R. D. Koster, A. M. Molod, and H. Wang, 2019: Tendency bias correction in coupled and uncoupled global climate models with a focus on impacts over North America. *J. Climate*, **32**, 639–661, https://doi.org/10.1175/JCLI-D-18-0598.1.

Crawford, W., 2021: Impact of model error techniques on the forecast skill of the Navy ESPC ensemble. NRL Memo. Rep. NRL/7530/MR—2021-1, 28 pp., https://apps.dtic.mil/sti/citations/AD1130118.

——, S. Frolov, J. McLay, C. A. Reynolds, N. Barton, B. Ruston, and C. H. Bishop, 2020: Using analysis corrections to address model error in atmospheric forecasts. *Mon. Wea. Rev.*, **148**, 3729–3745, https://doi.org/10.1175/MWR-D-20-0008.1.

Danforth, C. M., and E. Kalnay, 2008: Impact of online empirical model correction on nonlinear error growth. *Geophys. Res. Lett.*, **35**, L24805, https://doi.org/10.1029/2008GL036239.

Dee, D. P., and A. M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295, https://doi.org/10.1002/qj.49712454512.

——, and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

DeMaria, M., C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, https://doi.org/10.1175/BAMS-D-12-00240.1.

Domeisen, D. I. V., and Coauthors, 2022: Advances in the subseasonal prediction of extreme events: Relevant case studies across the globe. *Bull. Amer. Meteor. Soc.*, **103**, E1473–E1501, https://doi.org/10.1175/BAMS-D-20-0221.1.

Emanuel, K. A., 1986: An air–sea interaction theory for tropical cyclones. Part I: Steady-state maintenance. *J. Atmos. Sci.*, **43**, 585–605, https://doi.org/10.1175/1520-0469(1986)043<0585:AASITF>2.0.CO;2.

——, 1987: The dependence of hurricane intensity on climate. *Nature*, **326**, 483–485, https://doi.org/10.1038/326483a0.

——, 1988: The maximum intensity of hurricanes. *J. Atmos. Sci.*, **45**, 1143–1155, https://doi.org/10.1175/1520-0469(1988)045<1143:TMIOH>2.0.CO;2.

——, 1991: The theory of hurricanes. *Annu. Rev. Fluid Mech.*, **23**, 179–196, https://doi.org/10.1146/annurev.fl.23.010191.001143.

——, and D. S. Nolan, 2004: Tropical cyclone activity and the global climate system. *26th Conf. on Hurricanes and Tropical Meteorology*, Miami, FL, Amer. Meteor. Soc., 10A.2, https://ams.confex.com/ams/pdfpapers/75463.pdf.

——, and F. Zhang, 2016: On the predictability and error sources of tropical cyclone intensity forecasts. *J. Atmos. Sci.*, **73**, 3739–3747, https://doi.org/10.1175/JAS-D-16-0100.1.

Gilford, D. M., 2021: pyPI (v1.3): Tropical cyclone potential intensity calculations in python. *Geosci. Model Dev.*, **14**, 2351–2369, https://doi.org/10.5194/gmd-14-2351-2021.

Gonzalez, A. O., and X. Jiang, 2017: Winter mean lower tropospheric moisture over the Maritime Continent as a climate model diagnostic metric for the propagation of the Madden-Julian Oscillation. *Geophys. Res. Lett.*, **44**, 2588–2596, https://doi.org/10.1002/2016GL072430.

Gray, W. M., 1968: Global view of the origin of tropical disturbances and storms. *Mon. Wea. Rev.*, **96**, 669–700, https://doi.org/10.1175/1520-0493(1968)096<0669:GVOTOO>2.0.CO;2.

——, 1979: Hurricanes: Their formations, structure and likely role in the tropical circulation. *Supplement to Meteorology over the Tropical Oceans*, D. B. Shaw, Ed., James Glaisher House, 155–218.

Hogan, T. F., and Coauthors, 2014: The Navy global environmental model. *Oceanography*, **27** (3), 116–125, https://doi.org/10.5670/oceanog.2014.73.

Hunke, E. C., and W. Lipscomb, 2008: CICE: The Los Alamos Sea ice model, documentation and software user's manual, version 4.1. Tech. Rep., LA-CC-06-012, 76 pp., https://csdms.colorado.edu/w/images/CICE_documentation_and_software_user's_manual.pdf.

Janiga, M. A., C. J. Shreck III, J. A. Ridout, M. Flatau, N. P. Barton, E. J. Metzger, and C. A. Reynolds, 2018: Subseasonal forecasts of convectively coupled equatorial waves and the MJO: Activity and predictive skill. *Mon. Wea. Rev.*, **146**, 2337–2360, https://doi.org/10.1175/MWR-D-17-0261.1.

Jiang, X., A. F. Adames, D. Kim, E. D. Maloney, H. Lin, H. Kim, C. Zhang, C. A. DeMott, and N. P. Klingaman, 2020: Fifty years of research on the Madden-Julian Oscillation: Recent progress, challenges, and perspectives. *J. Geophys. Res. Atmos.*, **125**, e2019JD030911, https://doi.org/10.1029/2019JD030911.

Kim, D., and Coauthors, 2009: Applications of MJO simulation diagnostics to climate models. *J. Climate*, **22**, 6413–6436, https://doi.org/10.1175/2009JCLI3063.1.

——, J.-S. Kug, and A. H. Sobel, 2014: Propagating versus nonpropagating Madden–Julian oscillation events. *J. Climate*, **27**, 111–125, https://doi.org/10.1175/JCLI-D-13-00084.1.

——, and Coauthors, 2018: Process-oriented diagnosis of tropical cyclones in high-resolution GCMS. *J. Climate*, **31**, 1685–1702, https://doi.org/10.1175/JCLI-D-17-0269.1.

Kim, H., 2017: The impact of the mean moisture bias on the key physics of MJO propagation in the ECMWF reforecast. *J. Geophys. Res. Atmos.*, **122**, 7772–7784, https://doi.org/10.1002/2017JD027005.

——, P. J. Webster, V. E. Toma, and D. Kim, 2014: Predictability and prediction skill of the MJO in two operational forecasting systems. *J. Climate*, **27**, 5364–5378, https://doi.org/10.1175/JCLI-D-13-00480.1.

——, D. Kim, F. Vitart, V. E. Toma, J.-S. Kug, and P. J. Webster, 2016: MJO propagation across the maritime continent in the ECMWF ensemble prediction system. *J. Climate*, **29**, 3973–3988, https://doi.org/10.1175/JCLI-D-15-0862.1.

——, F. Vitart, and D. E. Waliser, 2018: Prediction of the Madden–Julian oscillation: A review. *J. Climate*, **31**, 9425–9443, https://doi.org/10.1175/JCLI-D-18-0210.1.

——, M. A. Janiga, and K. Pegion, 2019: MJO propagation processes and mean state biases in the SubX and S2S reforecasts. *J. Geophys. Res. Atmos.*, **124**, 9314–9331, https://doi.org/10.1029/2019JD031139.

Klocke, D., and M. J. Rodwell, 2014: A comparison of two numerical weather prediction methods for diagnosing fast-physics errors in climate models. *Quart. J. Roy. Meteor. Soc.*, **140**, 517–524, https://doi.org/10.1002/qj.2172.

Klotzbach, P. J., 2010: On the Madden–Julian oscillation–Atlantic hurricane relationship. *J. Climate*, **23**, 282–293, https://doi.org/10.1175/2009JCLI2978.1.

——, and E. C. J. Oliver, 2015: Variations in global tropical cyclone activity and the Madden–Julian Oscillation since the midtwentieth century. *Geophys. Res. Lett.*, **42**, 4199–4207, https://doi.org/10.1002/2015GL063966.

Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best track data. *Bull. Amer. Meteor. Soc.*, **91**, 363–376, https://doi.org/10.1175/2009BAMS2755.1.

——, H. J. Diamond, J. P. Kossin, M. C. Kruk, and C. J. Schreck III, 2018: International Best Track Archive for Climate Stewardship (IBTrACS) project, version 4. NOAA National Centers for Environmental Information, accessed 2 July 2021, https://doi.org/10.25921/82ty-9e16.

Kossin, J. P., S. J. Camargo, and M. Sitkowski, 2010: Climate modulation of North Atlantic hurricane tracks. *J. Climate*, **23**, 3057–3076, https://doi.org/10.1175/2010JCLI3497.1.

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

——, ——, F. Vitart, A. H. Sobel, J. Camp, S. Wang, M. K. Tippett, and Q. Yang, 2020: Subseasonal predictions of tropical cyclone occurrence and ACE in the S2S dataset. *Wea. Forecasting*, **35**, 921–938, https://doi.org/10.1175/WAF-D-19-0217.1.

Lee, H.-T., and NOAA CDR Program, 2011: NOAA Climate Data Record (CDR) of Daily Outgoing Longwave Radiation (OLR), version 1.2. NOAA National Climatic Data Center, accessed 1 July 2017, https://doi.org/10.7289/V5SJ1HH2.

Li, H., and R. L. Sriver, 2018: Tropical cyclone activity in the high-resolution community earth system model and the impact of ocean coupling. *J. Adv. Model. Earth Syst.*, **10**, 165–186, https://doi.org/10.1002/2017MS001199.

——, J. H. Richter, C.-Y. Lee, and H. Kim, 2022: Subseasonal tropical cyclone prediction and modulations by MJO and ENSO in CESM2. *J. Geophys. Res. Atmos.*, **127**, e2022JD036986, https://doi.org/10.1029/2022JD036986.

Li, W. Y., and Coauthors, 2019: Evaluating the MJO prediction skill from different configurations of NCEP GEFS extended forecasts. *Climate Dyn.*, **52**, 4923–4936, https://doi.org/10.1007/s00382-018-4423-9.

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

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

——, and ——, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. *J. Atmos. Sci.*, **29**, 1109–1123, https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2.

Maloney, E. D., and D. L. Hartmann, 2000a: Modulation of hurricanes activity in the Gulf of Mexico by the Madden-Julian Oscillation. *Science*, **287**, 2002–2004, https://doi.org/10.1126/science.287.5460.2002.

——, and ——, 2000b: Modulation of eastern North Pacific hurricanes by the Madden–Julian oscillation. *J. Climate*, **13**, 1451–1460, https://doi.org/10.1175/1520-0442(2000)013<1451:MOENPH>2.0.CO;2.

Mapes, B. E., and J. T. Bacmeister, 2012: Diagnosis of tropical biases and the MJO from patterns in the MERRA analysis tendency fields. *J. Climate*, **25**, 6202–6214, https://doi.org/10.1175/JCLI-D-11-00424.1.

Ogata, T., R. Mizuta, Y. Adachi, H. Murakami, and T. Ose, 2016: Atmosphere-ocean coupling effect on intense tropical cyclone distribution and its future change with 60 km-AOGCM. *Sci. Rep.*, **6**, 29800, https://doi.org/10.1038/srep29800.

Ollinaho, P., and Coauthors, 2017: Towards process-level representation of model uncertainties stochastically perturbed parameterizations in the ECMWF ensemble. *Quart. J. Roy. Meteor. Soc.*, **143**, 408–422, https://doi.org/10.1002/qj.2931.

Piccolo, C., M. J. P. Cullen, W. J. Tennant, and A. T. Semple, 2018: Comparison of different representations of model error in ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **145**, 15–27, https://doi.org/10.1002/qj.3348.

Raymond, D. J., S. L. Sessions, and Ž. Fuchs, 2007: A theory for the spinup of tropical depressions. *Quart. J. Roy. Meteor. Soc.*, **133**, 1743–1754, https://doi.org/10.1002/qj.125.

Reynolds, C. A., and Coauthors, 2022: Analysis of integrated vapor transport bias. *Mon. Wea. Rev.*, **150**, 1097–1113, https://doi.org/10.1175/MWR-D-21-0198.1.

Rodwell, M. J., and T. N. Palmer, 2007: Using numerical weather prediction to assess climate models. *Quart. J. Roy. Meteor. Soc.*, **133**, 129–146, https://doi.org/10.1002/qj.23.

Rushley, S. S., M. A. Janiga, J. A. Ridout, and C. A. Reynolds, 2022: The impact of mean state moisture biases on MJO skill in the Navy ESPC. *Mon. Wea. Rev.*, **150**, 1725–1745, https://doi.org/10.1175/MWR-D-21-0225.1.

Saha, S., 1992: Response of the NMC MRF Model to systematic-error correction within integration. *Mon. Wea. Rev.*, **120**, 345–360, https://doi.org/10.1175/1520-0493(1992)120<0345:ROTNMM>2.0.CO;2.

Schreck, C. J., III, K. R. Knapp, and J. P. Kossin, 2014: The impact of best track discrepancies on global tropical cyclone climatologies using IBTrACS. *Mon. Wea. Rev.*, **142**, 3881–3899, https://doi.org/10.1175/MWR-D-14-00021.1.

Sessions, S. L., S. Sentić, and D. J. Raymond, 2019: Balanced dynamics and moisture quasi-equilibrium in DYNAMO convection. *J. Atmos. Sci.*, **76**, 2781–2799, https://doi.org/10.1175/JAS-D-18-0173.1.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.*, **131**, 3079–3102, https://doi.org/10.1256/qj.04.106.

Sridevi, C., and Coauthors, 2022: Tropical cyclone track and intensity prediction skill of GFS model over NIO during 2019 and 2020. *Trop. Cyclone Res. Rev.*, **11**, 36–49, https://doi.org/10.1016/j.tcrr.2022.04.002.

Stan, C., 2012: Is cumulus convection the concertmaster of tropical cyclone activity in the Atlantic? *Geophys. Res. Lett.*, **39**, L19716, https://doi.org/10.1029/2012GL053449.

——, and Coauthors, 2022: Advances in the prediction of MJO-teleconnections in the S2S forecast system. *Bull. Amer. Meteor. Soc.*, **103**, E1426–E1447, https://doi.org/10.1175/BAMS-D-21-0130.1.

Ulate, M., C. Zhang, and J. Dudhia, 2015: Role of water vapor and convection-circulation decoupling in MJO simulations by a tropical channel model. *J. Adv. Model. Earth Syst.*, **7**, 692–711, https://doi.org/10.1002/2014MS000393.

Ullrich, P. A., and C. M. Zarzycki, 2017: TempestExtremes: A framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geosci. Model Dev.*, **10**, 1069–1090, https://doi.org/10.5194/gmd-10-1069-2017.

——, ——, E. E. McClenny, M. C. Pinheiro, A. M. Stansfield, and K. A. Reed, 2021: TempestExtremes v2.1: A community framework for feature detection, tracking and analysis in large datasets. *Geosci. Model Dev.*, **14**, 5023–5048, https://doi.org/10.5194/gmd-14-5023-2021.

Vannitsem, S., and Coauthors, 2021: Statistical postprocessing for weather forecasts: Review, challenges, and avenues in a big data word. *Bull. Amer. Meteor. Soc.*, **102**, E681–E699, https://doi.org/10.1175/BAMS-D-19-0308.1.

Vitart, F., 2014: Evolution of ECMWF sub-seasonal forecast skill scores. *Quart. J. Roy. Meteor. Soc.*, **140**, 1889–1899, https://doi.org/10.1002/qj.2256.

Weisheimer, A., S. Corti, T. Palmer, and F. Vitart, 2014: Addressing model error through atmospheric stochastic physical parameterizations: Impact on the coupled ECMWF seasonal forecasting system. *Philos. Trans. Roy. Soc.*, **A372**, 20130290, https://doi.org/10.1098/rsta.2013.0290.

Wheeler, M. C., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wave-number-frequency domain. *J. Atmos. Sci.*, **56**, 374–399, https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2.

——, and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, https://doi.org/10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.

White, C. J., and Coauthors, 2017: Potential applications of sub-seasonal-to-seasonal (S2S) predictions. *Meteor. Appl.*, **24**, 315–325, https://doi.org/10.1002/met.1654.

Wilks, D. S., 2016: "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. *Bull. Amer. Meteor. Soc.*, **97**, 2263–2273, https://doi.org/10.1175/BAMS-D-15-00267.1.

Wing, A. A., and Coauthors, 2019: Moist static energy budget analysis of tropical cyclone intensification in high-resolution climate models. *J. Climate*, **32**, 6071–6095, https://doi.org/10.1175/JCLI-D-18-0599.1.

Zarzycki, C. M., and P. A. Ullrich, 2017: Assessing sensitivities in algorithmic detection of tropical cyclones in climate data. *Geophys. Res. Lett.*, **44**, 1141–1149, https://doi.org/10.1002/2016GL071606.

Zhang, C., 2005: Madden-Julian Oscillation. *Rev. Geophys.*, **43**, RG2003, https://doi.org/10.1029/2004RG000158.

——, 2013: Madden–Julian oscillation: Bridging weather and climate. *Bull. Amer. Meteor. Soc.*, **94**, 1849–1870, https://doi.org/10.1175/BAMS-D-12-00026.1.