The First Operational Version of Taiwan Central Weather Bureau's One-Tier Global Atmosphere–Ocean Coupled Forecast System for Seasonal Prediction

HANN-MING HENRY JUANG,[b,d] TZU-YU WU,[d] PANG-YEN BRIAN LIU,[a] HSIN-YI LIN,[d] CHING-TENG LEE,[a] MIEN-TZE KUEH,[f] JIA-FONG FAN,[d] JEN-HER RIVER CHEN,[c] MONG-MING LU,[e] AND PAY-LIAM LIN[d]

[a] *Meteorology Research and Development Center, Central Weather Bureau, Taipei, Taiwan*
[b] *Environmental Modeling Center, NOAA/NWS/NCEP, College Park, Maryland*
[c] *Meteorological Information Center, Central Weather Bureau, Taipei, Taiwan*
[d] *Department of Atmospheric Sciences, National Central University, Taoyuan, Taiwan*
[e] *Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan*
[f] *Research Center for Environmental Changes, Academia Sinica, Taipei, Taiwan*

(Manuscript received 10 June 2022, in final form 25 January 2024, accepted 19 February 2024)

ABSTRACT: The first version of the Taiwan Central Weather Bureau one-tier (TCWB1T) fully coupled global atmospheric and oceanic modeling forecast system had been developed and implemented as a routine operation for seasonal prediction at Central Weather Bureau (CWB) in 2017, with a minor revision in 2020. Based on NCEP CFSv1, the global atmospheric model in NCEP CFSv1 was replaced by CWB's atmospheric global spectral model (GSM) and coupled with the GFDL MOM3. Several parameters have been tested and tuned in the CWB atmospheric GSM, achieving an optimal configuration with better sea surface temperature (SST) predictions for integration more than one year. Using NCEP CFSR as the initial condition, TCWB1T conducted hindcasts from 1982 to 2011 and forecasts from 2012 to 2019 to analyze its performance. The results of these hindcasts and forecasts show that the TCWB1T can make useful predictions as verified against the observations of OISST, ERSST, CFSR, and GPCP based on the methods of EOF, RMSE, anomaly correlation, ranked probability skill score (RPSS), reliability diagram (RD), and relative operating characteristics (ROCs). TCWB1T also has the same level of skill scores as NCEP CFSv2 and/or the ECMWF fifth-generation seasonal forecast system (SEAS5), based on EOF, anomaly pattern correlation, climatological bias, RMSE, temporal correlation, and anomaly correlation percentage of forecast skill. TCWB1T shows forecast skill that is better in winter than in summer. Overall, it indicates that TCWB1T can be used for seasonal ENSO predictions.

KEYWORDS: Climate models; Coupled models; Ensembles; Model comparison; Model evaluation/performance; Numerical weather prediction/forecasting

1. Introduction

Numerical weather and climate predictions are initial and boundary-value problems. In the case of long-range climate predictions, boundary values such as external climate forcings are important and necessary to provide along the temporal integration. Depending on the applications and resources, the boundary-value and external climate forcings in space and time can be provided by statistic values from datasets or integrated values from other components of Earth system models. The better the boundary values provided, the more accurate the prediction will be. Thus, the exchange of boundary values along spatial and temporal integration among different components of Earth system models as a coupling system has been adapted by current Earth modeling systems, which has been recognized as the best boundary providing technique, which is referred to as a fully coupling with two-way interaction; we called it one-tier.

Due to the robust and full coupling by one-tier systems that results in feedback among all model dynamics and physics, most seasonal and climate forecast modeling systems evolve from an uncoupled system to a one-tier fully coupled system, at least between the atmosphere and ocean. Examples of this include the European Centre for Medium-Range Weather Forecasts (ECMWF) model, which evolved from a research model directly into fully atmosphere–ocean coupled systems (Roberts et al. 2018) for operational use, and the National Centers for Environmental Prediction (NCEP) model, which also evolved from research and quasi-operational models (Ji et al. 1994) into an operational seasonal forecast model (SFM; Kanamitsu et al. 2002) and then into the Climate Forecast System, version 1 (CFSv1; Saha et al. 2006), and further into the Climate Forecast System, version 2 (CFSv2; Saha et al. 2014). This evolution can be seen by the fact that most climate forecast models operate without flux correction, as noted in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4; Randall et al. 2007), and there are many coupled models from all over the world listed in IPCC reports. This indicates that each institute would like to have a coupling forecast system for their own use and purposes. Therefore, it is reasonable for the Central Weather Bureau (CWB) in Taiwan to have their own coupling forecast system for seasonal predictions.

CWB has established numerical weather predictions since the early 1980s, starting with a first-generation global atmospheric gridpoint model, initially operational in 1989 (Liou et al. 1989),

Corresponding author: Hann-Ming Henry Juang, henry.juang@g.ncu.edu.tw

DOI: 10.1175/WAF-D-22-0128.1

TABLE 1. Elements of TCWB1T.

TCWB1T CFSv1	Contents
Atmosphere model	Taiwan Central Weather Bureau GSM (T119L40)
Ocean model	GFDL MOM3
Initial conditions	CFSR (Saha et al. 2010) OISST (Reynolds et al. 2002)
Integration	Each member starts at 0000 UTC, 10-month integration, no flux correction, freely coupled once per 24 h. And it is one member per day
Ensemble members (30-day daily lagged ensemble)	Hindcast: 900 members (30 members per month for 30 years from 1982 to 2011 as climatology) Forecast: 30 members per month

and a second-generation global atmospheric spectral model (GSM) with a resolution of T79L18 (79 wave triangular truncation with 18 vertical layers), operational in 1994 (Liou et al. 1997); the resolution improved to T119L30 in 2000. After 2000, CWB started to implement a global data assimilation Gridpoint Statistical Interpolation analysis system (GSI) three-dimensional variational data assimilation (3DVAR; Kleist et al. 2009) together with CWB GSM. They called their global atmospheric modeling with the data assimilation system the global forecast system (GFS) with a resolution of T319L40 routine operation in 2011 and T511L60 in 2015 with reduced Gaussian grid (Juang 2004) and two-dimensional message passing interface (MPI) programming (Juang and Kanamitsu 2001; Juang et al. 2007) and then Tco639L72 with noniteration dimensional-split semi-Lagrangian (NDSL) dynamics (Juang 2007, 2008). Nevertheless, for seasonal and climate forecasts, CWB started developing an uncoupled multimodel ensemble system in 2002 and made it operational in 2009. The multimodel ensemble climate forecast system includes the T42L18 European Centre Hamburg general circulation model 5 (ECHAM5; Roeckner et al. 2003) and CWB GSM using CWB optimized global sea surface temperature (OPGSST; Weng et al. 2005) and NCEP CFSv1 sea surface temperature (SST).

It is clear that CWB has a well-developed global atmospheric model for routine operational weather forecasts, seasonal, and climate predictions. However, CWB has less capability for oceanic modeling. Moving to a one-tier atmospheric and oceanic fully coupled modeling system may not be easy. The strategy to start with NCEP CFSv1 and only replace the NCEP global atmospheric model with the CWB global atmospheric model was an acceptable method to start from. However, it required tuning model physics to a fully coupled system to perform a similar phenomenon as the observed Earth system through the reforecast period and then the forecast period. This project started in 2010, and the final coupling system is called the Taiwan Central Weather Bureau one-tier (TCWB1T) seasonal and climate forecast system. TCWB1T became a routine operational system in 2017 (Wu et al. 2019), with a minor operational revision in 2020.

This article documents the successful development of the first version of TCWB1T as a fully coupled system without flux correction for operational seasonal prediction. We describe coupling strategy and methods with the required dataset in section 2, briefly illustrate our procedure for the fully coupled model for seasonal prediction in section 3, evaluate hindcast results in section 4, verify forecasts in section 5, and finally wrap up with our conclusions and discussion in section 6.

2. One-tier fully coupled atmospheric and oceanic modeling

a. *Methodology and model components*

As mentioned, CWB had been developing its own atmospheric general circulation GSM for years. CWB GSM was familiar to the dynamics and physics scientists in the bureau, and thus, it was quite reasonable to use CWB GSM for the atmospheric component of the atmosphere–ocean coupling system. Since the National Oceanic and Atmospheric Administration (NOAA) NCEP CFS system was a successful coupling system and available for external institutes to use, we adapted the CFSv1 package system, including all scripts and

TABLE 2. Resolution and physical schemes of TCWB1T.

Atmospheric model resolution	T119 (about 1° × 1°), 40 vertical levels
Radiation scheme	Fu and Liou (1993)
Boundary layer parameterization	MRF PBL A first-order nonlocal scheme (Hong and Pan 1996; Troen and Mahrt 1986)
Land surface model	Noah 4-layer soil model (Ek et al. 2003)
Ice model	Three-layer sea ice model (Winton 2000)
Cumulus parameterization	New SAS (Han and Pan 2011)
Ocean model resolution (MOM3) 74°S to 64°N	Zonal resolution: 1° domain wide Meridional resolution: 1/3° from 10°S to 10°N, increasing gradually from 1/3° to 1° meridional resolution poleward to 30°S and 30°N, and 1° for the remaining 40 vertical layers (see Table 3) with 10-m thickness from surface to 240 m, with 27 layers in the top 400 m to resolve the mixed layer and then gradually increasing layer thickness of the remaining 13 layers to bottom depth at 4500 m with a bottom layer thickness of about 511 m

TABLE 3. MOM3 model layers in meters from sea surface to sea bottom.

Vertical coordinate (m)	5	15	25	35	45
	55	65	75	85	95
	105	115	125	135	145
	155	165	175	185	195
	205	215	225	238.4779	262.2945
	303.0287	366.7978	459.091	584.6193	747.187
	949.5881	1193.53	1479.588	1807.187	2174.619
	2579.091	3016.798	3483.029	3972.294	4478.478

coupling system, and replaced the atmospheric component, NCEP GSM in NCEP CFSv1, with the CWB operational GSM (T319L40 operation in 2011, based on Liou et al. 1997). With the assumption that the ocean model was already a well-tuned component in the NCEP CFSv1 coupling system, only the replaced atmospheric component, CWB GSM, had to be tuned to harmonically couple with the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model, version 3 (MOM3; Pacanowski and Griffies 1999).

Table 1 shows the elements of TCWB1T, including the atmosphere model CWB GSM with T119 and 40 vertical layers, and the ocean model of GFDL MOM3. Table 2 continuously shows the major model physics in CWB GSM, including a radiation scheme (Fu and Liou 1993), a first-order nonlocal boundary-layer parameterization (Hong and Pan 1996; Troen and Mahrt 1986), Noah (NOAA/NCEP–Oregon State University–Air Force Research Laboratory–NOAA/Office of Hydrology land surface model) land surface model with the four-layer soil model (Ek et al. 2003), and cumulus parameterization of simplified Arakawa–Schubert scheme (Han and Pan 2011). We are using climatologic values for ice at initial and daily update with a three-layer sea ice physics (Winton 2000). There is no ice prediction, so there is no ice drift issue. And GFDL MOM3 covers from 74°S to 64°N with zonal resolution of 1° and meridional resolution of 1/3° from 10°S to 10°N, then increasing grid spacing from 1/3° to 1° poleward to 30°N and 30°S, and 40 vertical layers, whose thickness was 10 m from the ocean surface to 240 m deep, with total 27 levels in top 400 m to resolve the mixed layer, and the bottom depth is 4500 m with the bottom layer thickness of about 511 m (Saha et al. 2006). The details of the layer depth can be found in Table 3.

b. *Data, integration, and ensemble members*

To establish a climate coupling system, we must make sure that we have enough consistent initial conditions from the past several decades for reforecast to the future for forecast, and thus, we must either have our own reanalysis for the past and future or obtain available data from other operational centers. The CFSv2 was the operational system in NCEP with reforecast and forecast initial conditions through its analysis. It was the best candidate for CWB to use for the CWB coupling system. Table 1 shows the initial condition from Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) based on CFSv2; the CFSR initial data contain atmosphere data (upper air data and surface data) which have high resolution and needed to interpolate to T119 with 40 layers for CWB GSM. CFSR also has ocean data of GFDL MOM4 (Griffies et al. 2005) and needed to interpolate to MOM3 domain and resolution. The Optimum

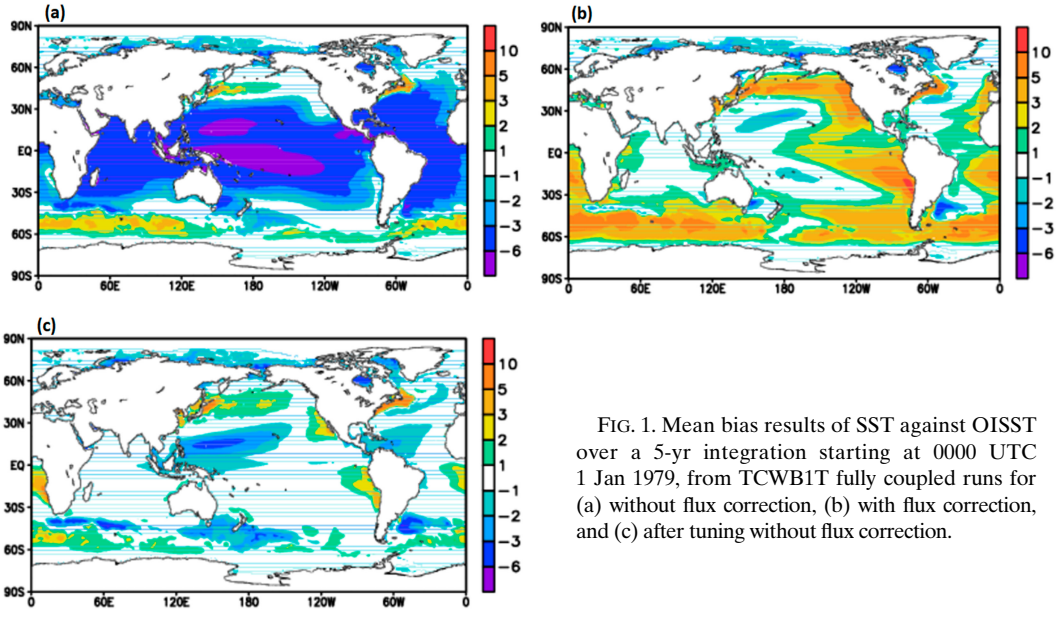

FIG. 1. Mean bias results of SST against OISST over a 5-yr integration starting at 0000 UTC 1 Jan 1979, from TCWB1T fully coupled runs for (a) without flux correction, (b) with flux correction, and (c) after tuning without flux correction.

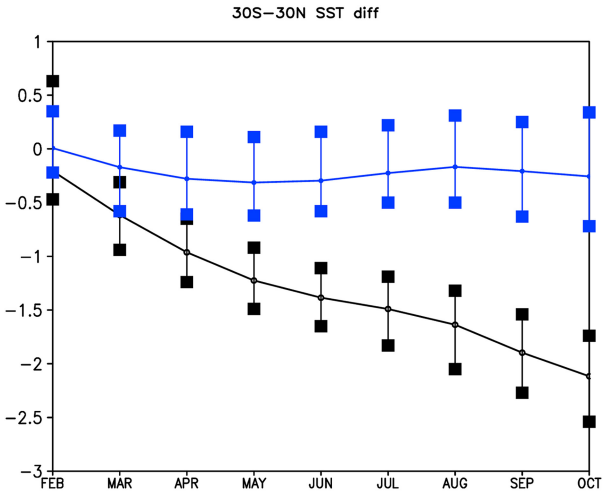

FIG. 2. SST mean bias with respect to OISST within 30°S–30°N global band of 9-month ensemble forecasts starting from January with six members (1, 6, 11, 16, 21, and 26 Jan) and 29 years (1980–2008). The black curve is the result of TCWB1T before model physics tuning, and the blue curve is after model physics tuning. The vertical lines with end marks along the black and blue curves indicate the ranges of maximal and minimal values of SST mean bias among 174 members.

Interpolation Sea Surface Temperature (OISST; Reynolds et al. 2002) dataset is used for the atmosphere SST initial condition.

For reforecast and forecast, TCWB1T conducted 10 months of integration with a coupling frequency of once every 24 h. The fully coupled system integrates in a sequential way: First, it integrates the atmospheric component (CWB GSM) with OISST as the initial boundary condition. After 1 day, the atmospheric component passes daily mean fluxes to MOM3 and then integrates MOM3 for 1 day to update ocean fields. Then, MOM3 passes SST to CWB GSM, which completes the first fully couple cycle. Then it returns to integrate CWB GSM with forecast MOM3 SST for the next couple cycle, and this process repeats until the end of integration.

Each TCWB1T integration is called a member of the ensemble forecast, which starts at 0000 UTC of any given date as an initial condition. For simplicity and other concerns (discussed later in section 2c), all members are conducted with a 10-month integration. For ensemble forecast, past 29 days of any given date and the given date form a 30-member daily lagged ensemble forecast. For consistency between forecast and hindcast, each hindcast uses the same daily lagged ensemble as forecast. For hindcast years from 1982 to 2011 (30 years), each climatological hindcast ensemble contains 900 members: 30 members per year for 30 years. Each forecast ensemble contains 30 members.

c. *Operational strategy*

On the 20th day of any given month, the operational center in CWB issues the upcoming TCWB1T monthly and seasonal forecasts. It uses the past 30 days, counting backward from the 14th of the current month until around the 15th of the previous month, as the 30 members of daily lagged ensemble forecast, it then uses 5 days (from the 15th to 19th of the current month) to do postprocessing and product generation, and then releases monthly and seasonal forecasts on the 20th of the current month. For example, the 20 April monthly and seasonal forecasts release, daily lagged initial dates are 16 March–14 April, and postprocessing and product generation dates are 15–19 April. Thus, the monthly mean result of 1–31 May is lead 0 forecast; that of 1–30 June is lead 1 forecast, etc.

We follow the original design for ensemble initial conditions as used in NCEP CFSv1, except that we used only initial conditions of 0000 UTC. The degradation due to longer time lagged members as compared to shorter time lagged members has been realized. However, we found that more distant different dates have better ensemble spread than closer different dates, and they also have the same conclusion that multiple cycles have

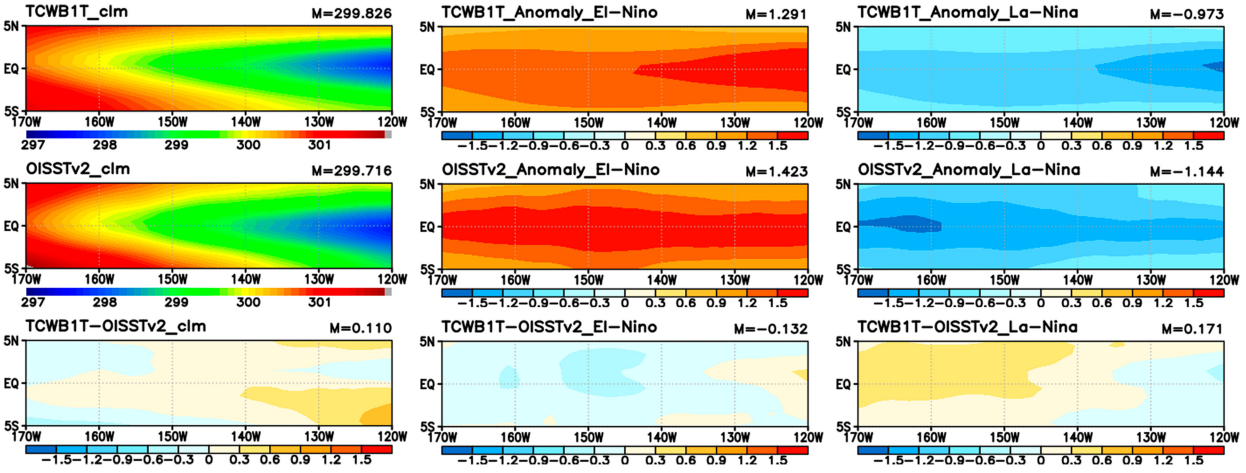

FIG. 3. The DJF SST anomaly in Niño-3.4 area. (top) The result from TCWB1T with initial conditions of August, (middle) the result from OISSTv2, and (bottom) the difference of the top and middle rows (TCWB1T minus OISSTv2); (left) climate average, (center) mean anomaly of all El Niño cases, and (right) mean anomaly of all La Niña cases, in hindcast years from 1982 to 2011.

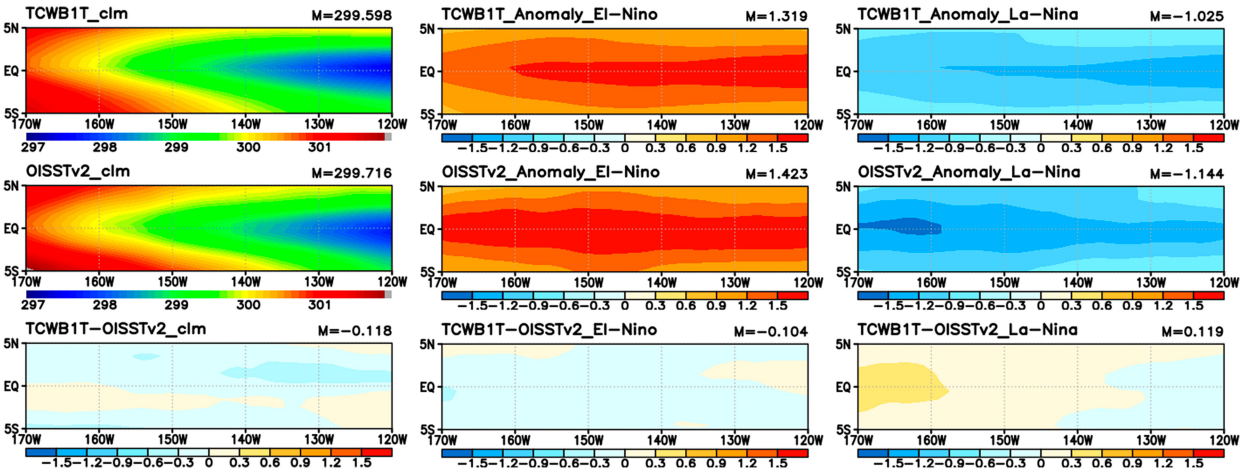

FIG. 4. As in Fig. 3, but with initial conditions of November.

less ensemble spread (figure not shown here). Note that our findings are similar to those in Vitart and Takaya (2021). Thus, using a long time lag up to 30 days as ensemble members, as did NCEP CFSv1, may not be an issue.

For consistency and for anomaly forecast in monthly and seasonal predictions, we must pick the reforecast members with 30-member initial conditions at the same range of dates used in the operational forecast as illustrated in the previous paragraph. Note that, due to the lagged initial conditions, the 10-month integration will leave the months after lead 8 with incomplete members. Since the official seasonal forecasts in CWB issue only two seasons, it is practical for simplicity for all members to use 10-month integration.

3. Physical parameterization tuning for one-tier fully coupled system

In this section, we will illustrate our procedure to tune the fully coupled model for seasonal forecasts based on the prediction of SST, especially on the tropical SST from 30°S to 30°N. Figure 1 shows the SST bias of TCWB1T fully coupled results averaged over 5-yr integration with respect to OISST, (i) without flux correction, (ii) with flux correction before model physics tuning for operational seasonal forecast, and (iii) after tuning without flux correction. It implies there is a cold bias for the current operational atmospheric model while it is coupled with MOM3. Based on the statistical bias, we can make a bias correction on atmospheric fluxes (Anderson et al. 2003) to be used in MOM3; however, the flux correction made a better result but overcorrected to a warm bias. As mentioned in the previous section, we would like to have a fully coupled model, and the flux correction cannot be helpful to atmospheric model physics parameterization. It will be better to have no flux correction so that the atmospheric response to the ocean can be fully coupled.

We used physical parameters to test the sensitivities of TCWB1T behavior on SST and investigate the model results of all fields. We found that RHc is the most sensitive and effective to alter tropical SST. RHc is defined as a given relative humidity threshold value in the model for phase change from vapor to water, which is used in the large-scale precipitation model physics, usually less than 100%. When RHc is larger, the water vapor will not easily condensate, so that more water vapor will be kept in the atmosphere, which could potentially increase

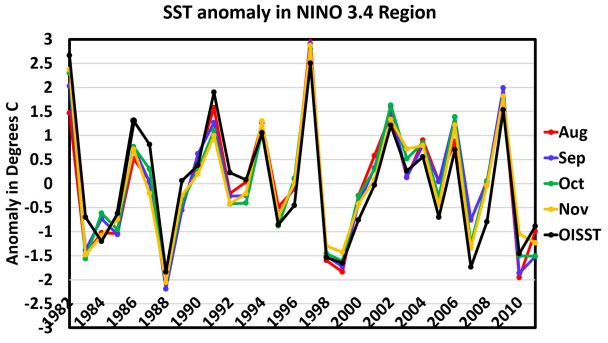

FIG. 5. DJF SST anomaly forecast over Niño-3.4 by TCWB1T with initial conditions from August to November and from 1982 to 2011, as compared to OISSTv2. The horizontal axis is for hindcast years, and the vertical axis is for SST anomaly (°C). The legend at the right-hand side indicates the curve colors to the different initial conditions and the black color is for OISSTv2.

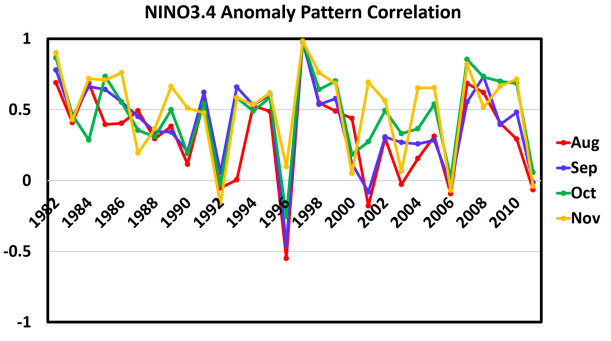

FIG. 6. As in Fig. 5, but for anomaly pattern correlation.

TABLE 4. The Niño-3.4 area-averaged SST anomaly pattern correlation for DJF by TCWB1T with the initial conditions from August to November of years from 1982 to 2011.

	August	September	October	November
Area mean	0.322 888	0.393 491	0.456 913	0.502 938

the greenhouse effect to warm up SST. We are not using it related to convection but large-scale precipitation parameterization.

We have the final optimal model physics package with the new simplified Arakawa–Schubert cumulus scheme (NSAS; Han and Pan 2011) to catch up with the operational GSM for weather forecast, land ice/snow with an albedo of 0.65 instead of 0.55 to reduce surface heating, the ocean sea ice temperature changes to avoid rapid melting, and RHc of 0.95 instead of 0.85 to keep more vapor in the air. Figure 2 shows the comparison of statistical climatological mean SST bias with respect to OISST for 9-month integrations. Again, it is within the 30°S–30°N global band mean SST from six members of each January of all 29-yr means and their variation with maximal and minimal values of the SST mean bias among 174 members. The black curve indicates the results before the final tuning, and the blue curve shows the results after the final tuning. It is clearly shown that the final tuning results have remarkably close to zero mean bias, but before tuning, it had a cold bias close to 2°C colder after 9-month integrations. Note that we did not conduct flux correction results for Fig. 2 because we will not use flux correction and we did save resources. The final tuned model physics package will be used in TCWB1T for hindcast in the next section and prediction in section 5 for operational purposes.

Based on the 30°S–30°N global band mean SST bias with respect to OISST, we have model physics tuned for minimal SST bias. This tuned model physics with some related modifications is used in TCWB1T for hindcasts and forecasts in sections 4 and 5, respectively. The related modifications include (i) changing machine accuracy from 1.0×10^{308} to 1.0×10^{307} to fit on the current operational supercomputer; (ii) adjusting all negative moisture by borrowing moisture from nearby lower layers. There is no moisture conservation issue because we borrow existing moisture, so moisture is conserved, but it results in slightly more moisture in the upper layers; (iii) changing the first near ground surface pressure by the mean of the first level and ground surface pressures; and (iv) changing four-point interpolation to optimum interpolation with effective radius of 60 km for interpolating the ocean initial condition from MOM4 grids to MOM3 grids.

4. 1982–2011 hindcast results

Based on the model configuration in the previous section for TCWB1T, we conducted hindcast from 1982 to 2011. To evaluate the hindcast results, we emphasize the climatological phenomena of El Niño–Southern Oscillation (ENSO) first. The datasets used for verification are OISSTv2 (in 0.25° resolution which is higher resolution than original version of OISST; Reynolds et al. 2007, 2009), the Extended Reconstructed Sea Surface Temperature (ERSST; Smith and Reynolds 2004; Smith et al. 2008; Banzon et al. 2010), ECMWF interim reanalysis (ERA-Interim; Berrisford et al. 2011), CFS Reanalysis and Reforecast (CFSRR; Saha et al. 2014) project, and Global Precipitation Climatology Project (GPCP; Huffman et al. 1997; Adler et al. 2003). The predicted variables used to compute their skill scores are SST, 2-m temperature (T2M), and precipitation. Since we emphasized on ENSO, the domain to be used for validation is mainly Niño-3.4; however, we also include domains of the global band from 40°S to 40°N, from 60°S to 60°N, and entire globe for related scores to have more coverage for hindcast verification.

a. *Results over Niño-3.4*

Since we used CFSR for initial condition and verification, we found that CFSR has one discrepancy with SST jump in

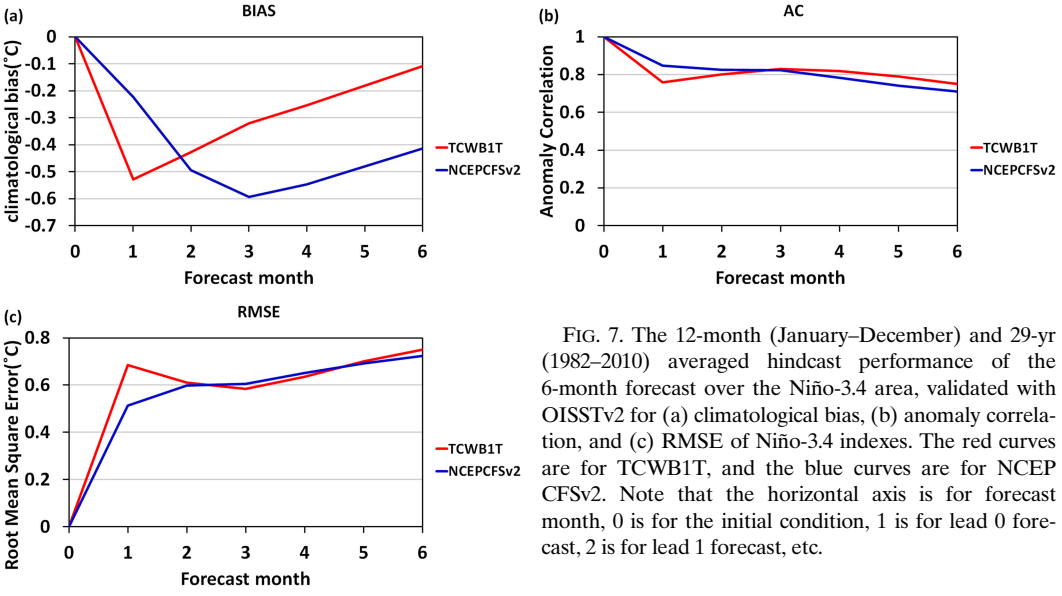

FIG. 7. The 12-month (January–December) and 29-yr (1982–2010) averaged hindcast performance of the 6-month forecast over the Niño-3.4 area, validated with OISSTv2 for (a) climatological bias, (b) anomaly correlation, and (c) RMSE of Niño-3.4 indexes. The red curves are for TCWB1T, and the blue curves are for NCEP CFSv2. Note that the horizontal axis is for forecast month, 0 is for the initial condition, 1 is for lead 0 forecast, 2 is for lead 1 forecast, etc.

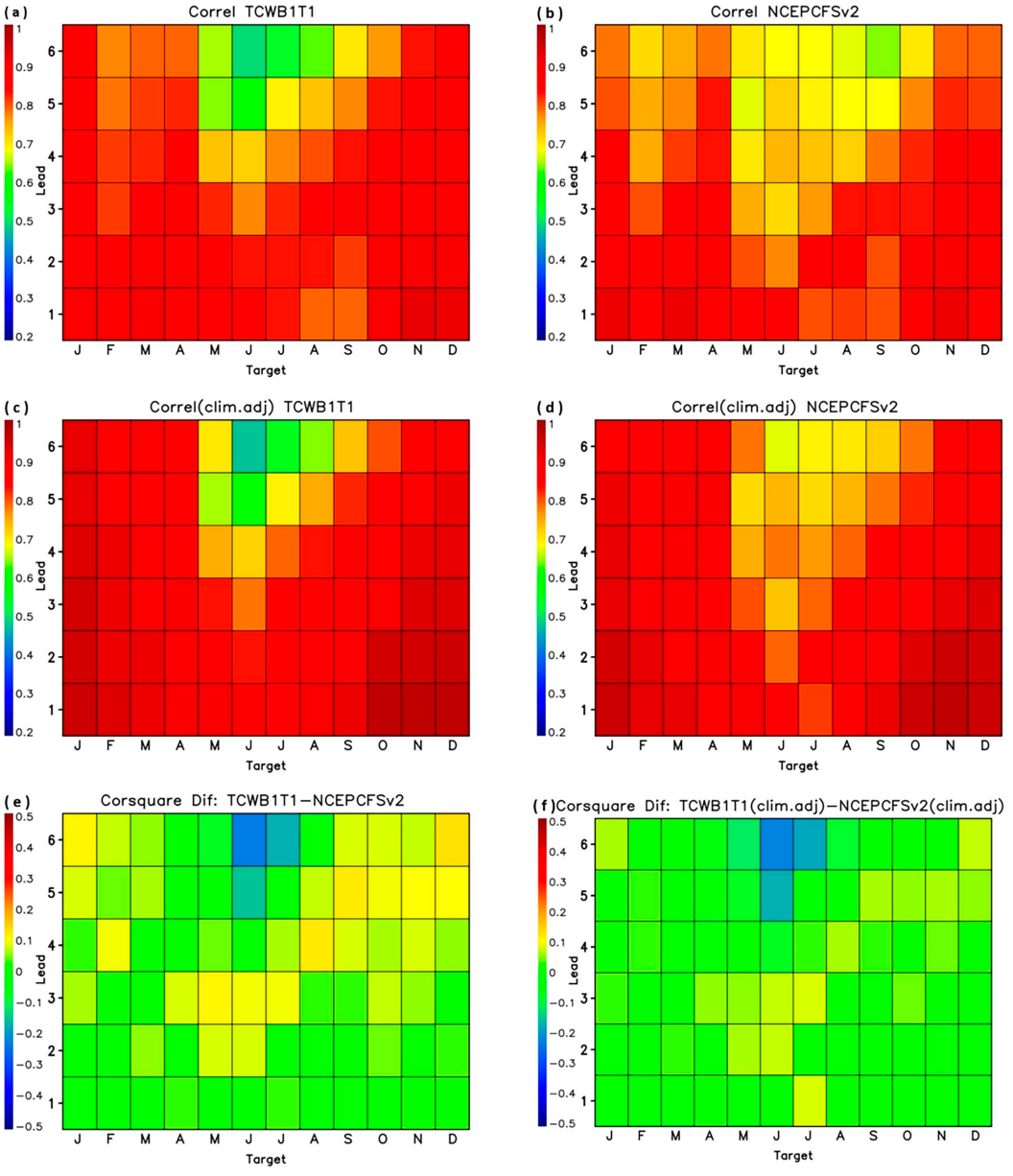

FIG. 8. Temporal correlation between (a) TCWB1T and (b) NCEP CFSv2 predictions of Niño-3.4 SST and verifying with NCEP CFSR ground temperature over the 1982–2010 period. Temporal correlation for (c) TCWB1T and (d) NCEP CFSv2 following the elimination of discontinuities in the predictions of each model by using split climatology. (e),(f) Differences of squared correlation between TCWB1T and NCEP CFSv2 without the split climatology and with the split climatology for each model, respectively.

October 1998, due to the start of using Advanced TIROS Operational Vertical Sounder (ATOVS) satellite observation data, as pointed out and illustrated with three possible reasons in Xue et al. (2011). This discrepancy is also mentioned in Saha et al. (2014). This SST jump results in two kinds of climatology, colder before October 1998 and warmer after October 1998, which is referred to as the split climatology. We made corrections to the hindcast results in this subsection due to this split climatology for Niño-3.4 area.

The elements of ENSO can be characterized by El Niño and La Niña. During 1982–2011, there are 10 years of El Niño and 11 years of La Niña. The El Niño years are 1982, 1986, 1987, 1991, 1994, 1997, 2002, 2004, 2006, and 2009, and the La Niña years are 1983, 1984, 1988, 1995, 1999, 2000, 2005, 2007, 2008, 2010, and 2011. The first row of Figs. 3 and 4 shows the TCWB1T December–February (DJF) hindcast for climatology of all cases, the mean anomaly of all El Niño cases, and the mean anomaly of all La Niña cases, with initial conditions

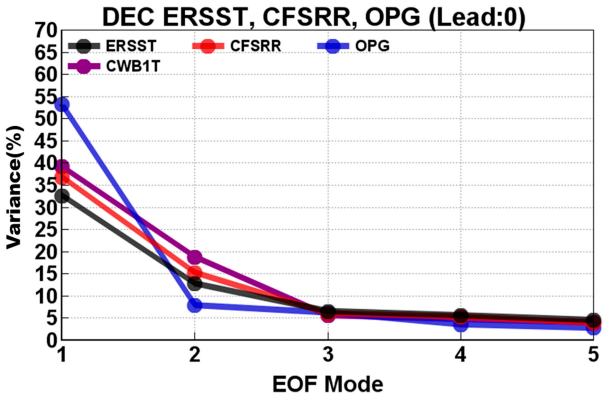

FIG. 9. Fraction of variance explained by EOF modes 1–5 for the SST anomaly. The black line is ERSST, the purple line is TCWB1T, the red line is CFSRR, and the blue line is CWB OPGSST, with initial conditions of November to forecast December for all years from 1982 to 2011.

from August in Fig. 3 and November in Fig. 4. The second row of Figs. 3 and 4 shows the OISSTv2 DJF climatology, the mean anomaly of all El Niño cases, and the mean anomaly of all La Niña cases. The third row is the differences between TCWB1T and OISSTv2. From the third row, we can see that the differences are small with magnitudes of area mean values from 0.11 to 0.17 for August and from 0.10 to 0.12 for November. It indicates that TCWB1T has reasonable hindcast performance in terms of climatology and mean anomaly of ENSO over Niño-3.4 as compared to OISSTv2. Furthermore, Figs. 3 and 4 show that results from El Niño are better than those from La Niña, but both have not extended enough westward.

Figure 5 shows the DJF forecasts of regional SST anomaly over Niño-3.4 for years from 1982 to 2011 by TCWB1T as compared to OISSTv2 with the different initial conditions from August to November of each year. It indicates that TCWB1T has anomaly forecast on phase with OISSTv2 even up to 4 months ahead for all ENSO years. Figure 6 shows the DJF forecasts of regional SST anomaly pattern correlation with respect to OISST over Niño-3.4 for years from 1982 to 2011 by TCWB1T for different initial conditions from August to November of each year. Even though Fig. 5 shows about 87% of anomaly forecasts on phase with observation from all different initial conditions, to investigate whether SST is well predicted, we should check SST anomaly pattern correlation score in Fig. 6. The SST anomaly pattern correlation score must be larger than 0.5 for better forecast of SST anomaly distributions. In 30 years, there are half of the total number of hindcast years with three initial conditions having SST anomaly pattern correlation larger than 0.5. Within these 15 years, there are 6 El Niño years, 6 La Niña years, and 3 normal years. And there are 4 years that have negative anomaly pattern correlation scores of all different initial months: 1992, 1996, 2006, and 2011. Within these 4 years, 1992 and 1996 have an out of phase to the OISSTv2 SST anomaly, and 2006 and 2011 have an overprediction of the SST anomaly. The remaining 11 years have a positive anomaly pattern correlation, but it is less than 0.5. Table 4 shows the 30-yr mean of the SST anomaly pattern correlation for different initial conditions; it indicates that the closer to DJF the initial condition is, the better the result is, and the November initial condition has mean score over 0.5.

The analyzed results so far are mostly concentrated on certain initial conditions for one season forecast (DJF). We will use different verification methods to investigate 6-month forecasts for all seasons and compare them to NECP CFSv2. First, we start from the model result of Niño-3.4 climatological bias, anomaly correlation, root-mean-square error (RMSE), and temporal correlation.

Hindcasts of TCWB1T and NCEP CFSv2 have some differences: the TCWB1T hindcast period is from 1982 to 2011, every day at 0000 UTC is for the initial condition, and 30 initial

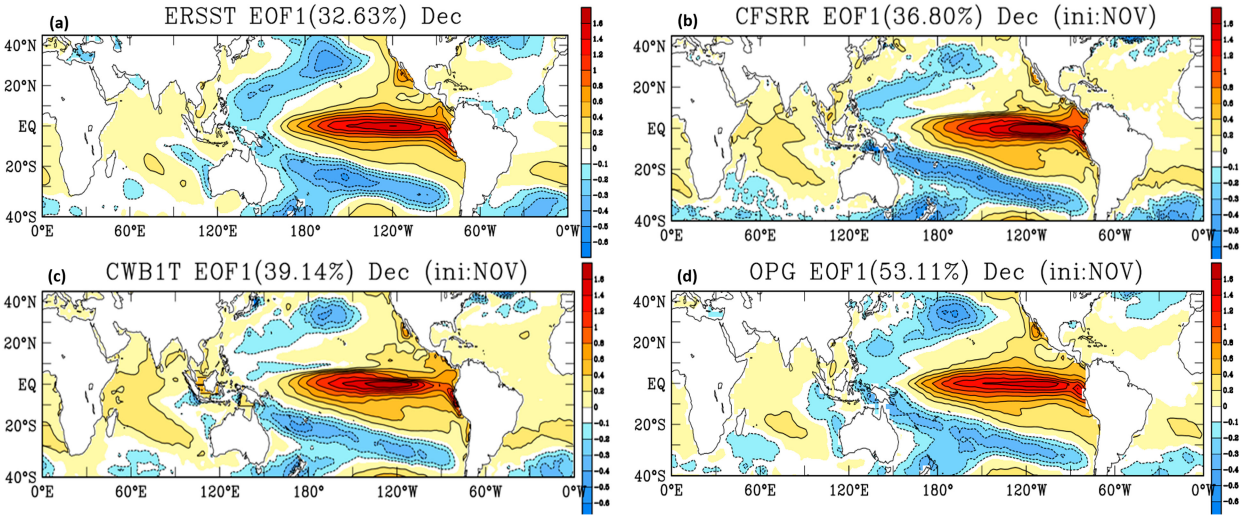

FIG. 10. Spatial patterns of the EOF first mode of SST anomaly calculated for the period from 1982 to 2011 over the global band of 40°S–40°N for (a) ERSST, (b) CFSRR, (c) TCWB1T, and (d) OPGSST.

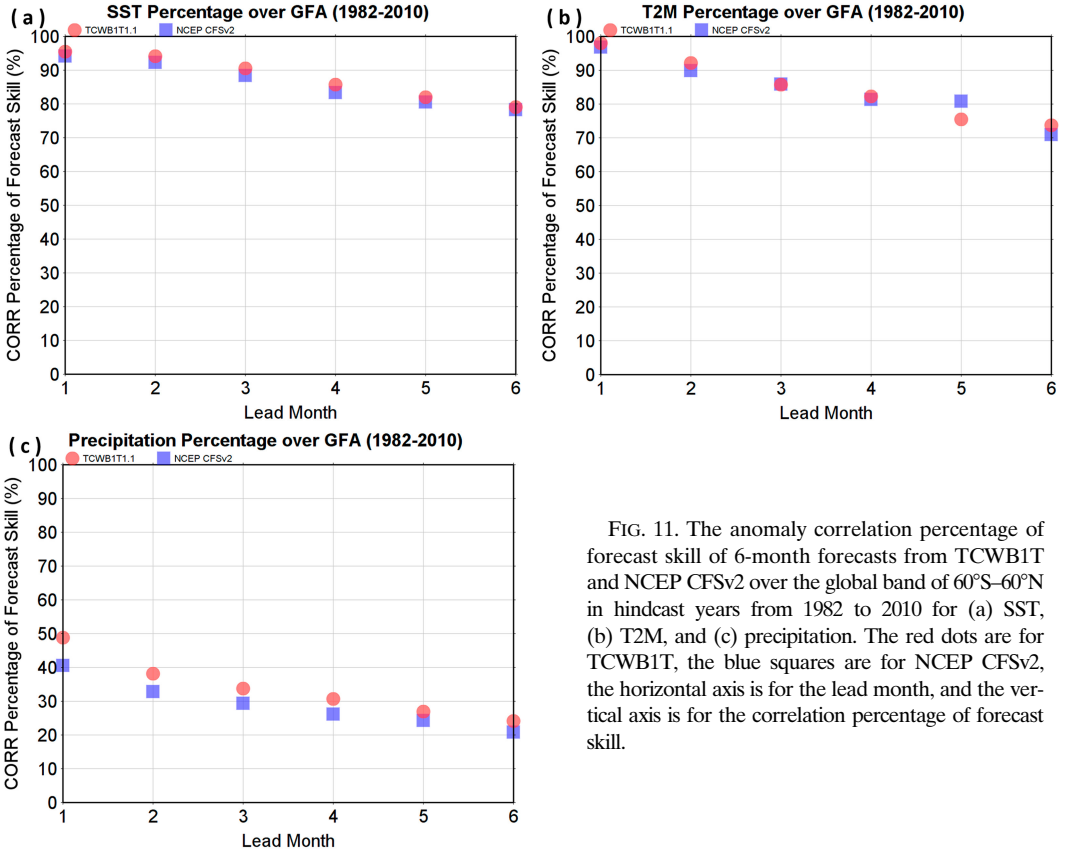

FIG. 11. The anomaly correlation percentage of forecast skill of 6-month forecasts from TCWB1T and NCEP CFSv2 over the global band of 60°S–60°N in hindcast years from 1982 to 2010 for (a) SST, (b) T2M, and (c) precipitation. The red dots are for TCWB1T, the blue squares are for NCEP CFSv2, the horizontal axis is for the lead month, and the vertical axis is for the correlation percentage of forecast skill.

conditions (30 members) are in 1 month. The NCEP CFSv2 hindcast period is from 1981 to 2010, four cycles (0000, 0600, 1200, and 1800 UTCs) of dates of 1, 6, 11, 16, 21, and 26, and a total of 24 members are in 1 month. For consistency, we select the period from 1982 to 2010 for both model hindcasts for comparison. Also, the results of Niño-3.4 here from TCWB1T and NCEP CFSv2 have been corrected based on the split climatology (1982–98, 1999–2010).

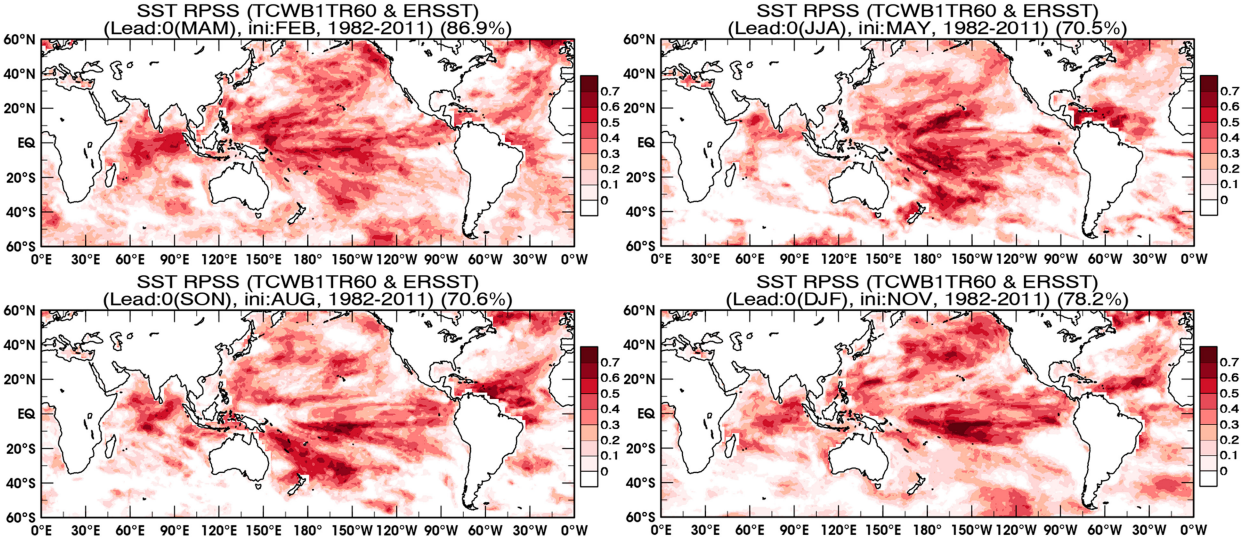

FIG. 12. The lead 0 first seasonal forecasts of SST ranked probability skill scores from TCWB1T with respect to ERSST over 60°S–60°N, with initial conditions of February, May, August, and November as the sequence from left to right and from top to bottom, respectively. We used three categories to calculate RPSS.

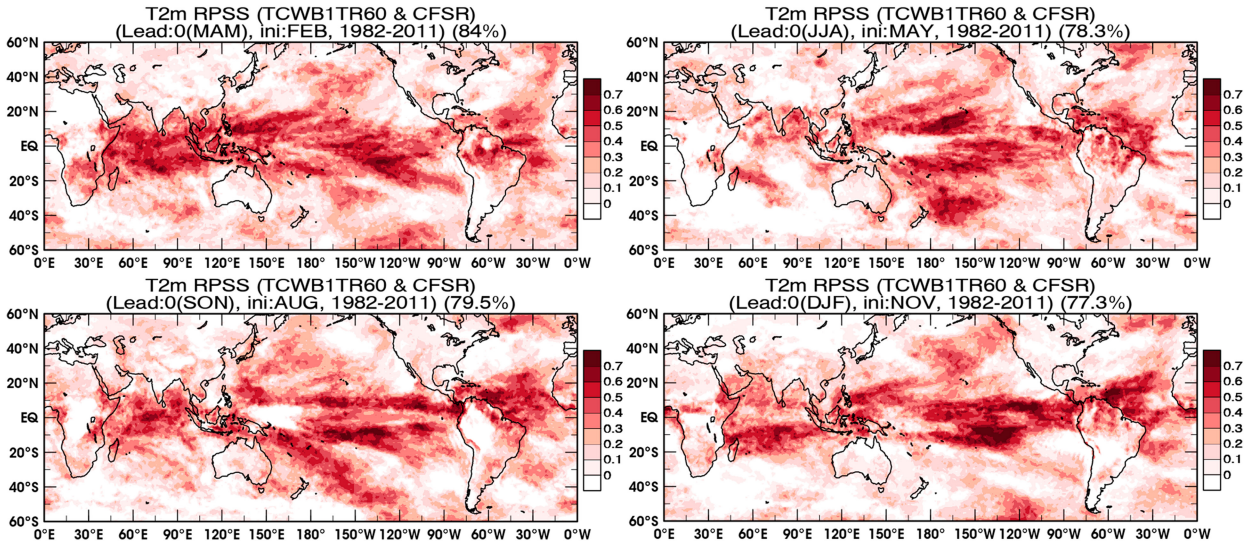

FIG. 13. As in Fig. 12, but for T2M.

Figure 7 shows hindcast results that cover 1982–2010 with forecast months from January to December. The forecast length is 6 months; red curves are for TCWB1T, and blue curves are for NCEP CFSv2. Based on the investigation method of Johnson et al. (2019), hindcast results were validated with OISSTv2 in terms of bias, anomaly correlation, and RMSE of Niño-3.4 indexes.

From Fig. 7a, we can see that the TCWB1T bias of Niño-3.4 indexes with respect to OISSTv2 rapidly increases negative magnitudes to about −0.53 in the 1-month forecast and then decreases its negative magnitudes with time to −0.11 in the 6-month forecast. However, the NCEP CFSv2 bias of Niño-3.4 indexes increases negative magnitudes to −0.22 in 1 month and −0.59 in 3 months and then decreases negative magnitudes slowly to about −0.41 in the 6-month forecast. We can compare Fig. 7a to Fig. 2a in Johnson et al. (2019), which shows slowly increasing negative bias to about −0.4 in the 4-month forecast and then slowly decreases with time to about −0.25 in the 12-month forecast in ECMWF's fifth generation seasonal forecast system (SEAS5). SEAS5 and NCEP CFSv2 had similar character, which had a maximum negative bias at the third or fourth month forecast and gradually reduced the negative bias. However, TCWB1T has a maximum bias at the first month and quickly decreases the magnitude of negative bias. We do not know how to explain this behavior; we can only say that it is the character of TCWB1T. The result of the climatology bias in Fig. 7a can explain the anomaly correlation of the Niño-3.4 index in Fig. 7b. TCWB1T has lower anomaly

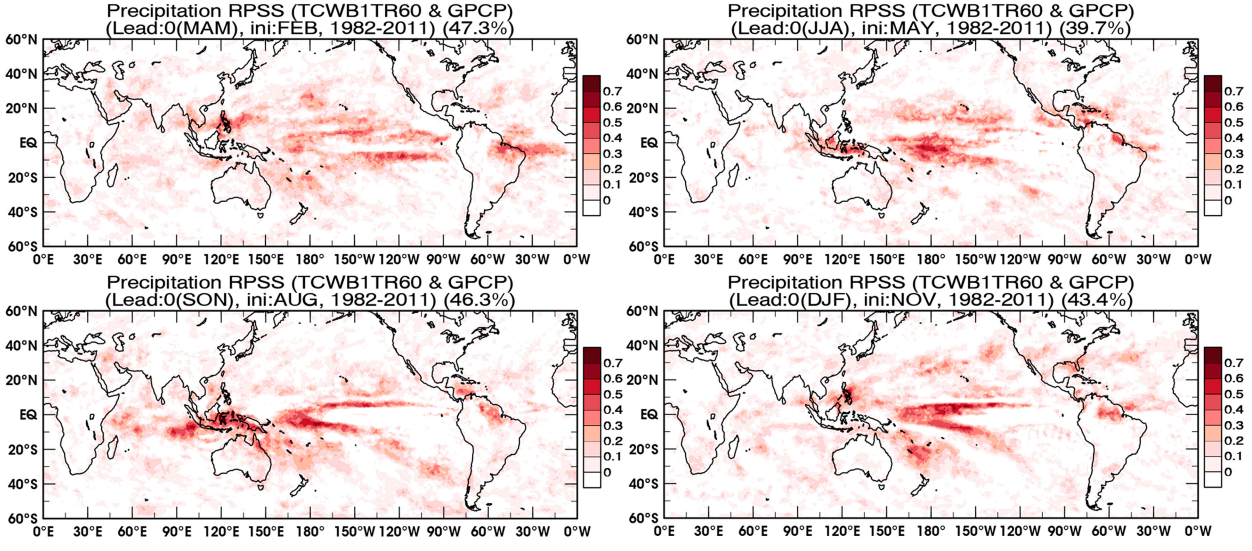

FIG. 14. As in Fig. 12, but for precipitation.

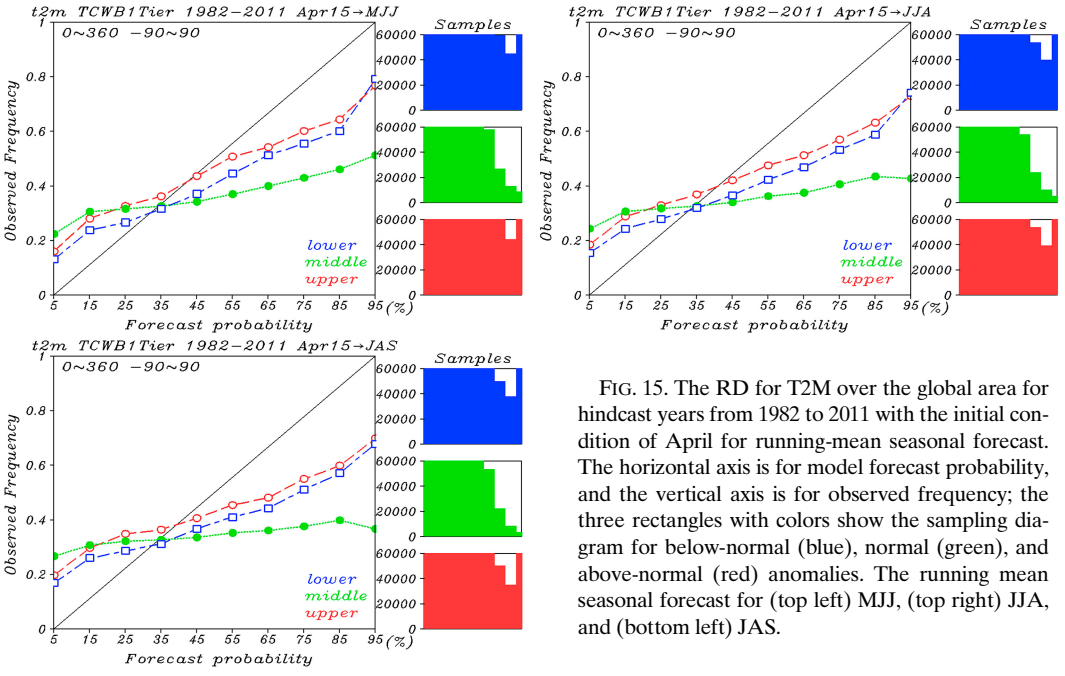

FIG. 15. The RD for T2M over the global area for hindcast years from 1982 to 2011 with the initial condition of April for running-mean seasonal forecast. The horizontal axis is for model forecast probability, and the vertical axis is for observed frequency; the three rectangles with colors show the sampling diagram for below-normal (blue), normal (green), and above-normal (red) anomalies. The running mean seasonal forecast for (top left) MJJ, (top right) JJA, and (bottom left) JAS.

correlation than NCEP CFSv2 at the first month because TCWB1T has the largest bias at the first month, and the negative bias is larger than NCEP CFSv2. And after two months, the climatology bias of TCWB1T is of small magnitude, which implied that the anomaly correlation of TCWB1T is better than that of NCEP CFSv2. In Fig. 7c, the RMSE of TCWB1T increases rapidly in the first month, which is larger than NCEP CFSv2, and after the second month, the RMSE decreases to the same magnitude as NCEP CFSv2. In general, TCWB1T is worse than NCEP CFSv2 at the first month, and in the second month, TCWB1T is similar to NCEP CFSv2 in terms of RMSE and anomaly correlation, but it is better in terms of bias.

Based on the method in Barnston and Tippett (2013) for intercomparison between NCEP CFSv1 and NCEP CFSv2 on SST temporal correlation over Niño-3.4, we validated with NCEP CFSR ground temperature. We do the same calculations to compare between TCWB1T and NCEP CFSv2. Figures 8a and 8b show the SST temporal correlation over Niño-3.4 without split climatology correction for TCWB1T and

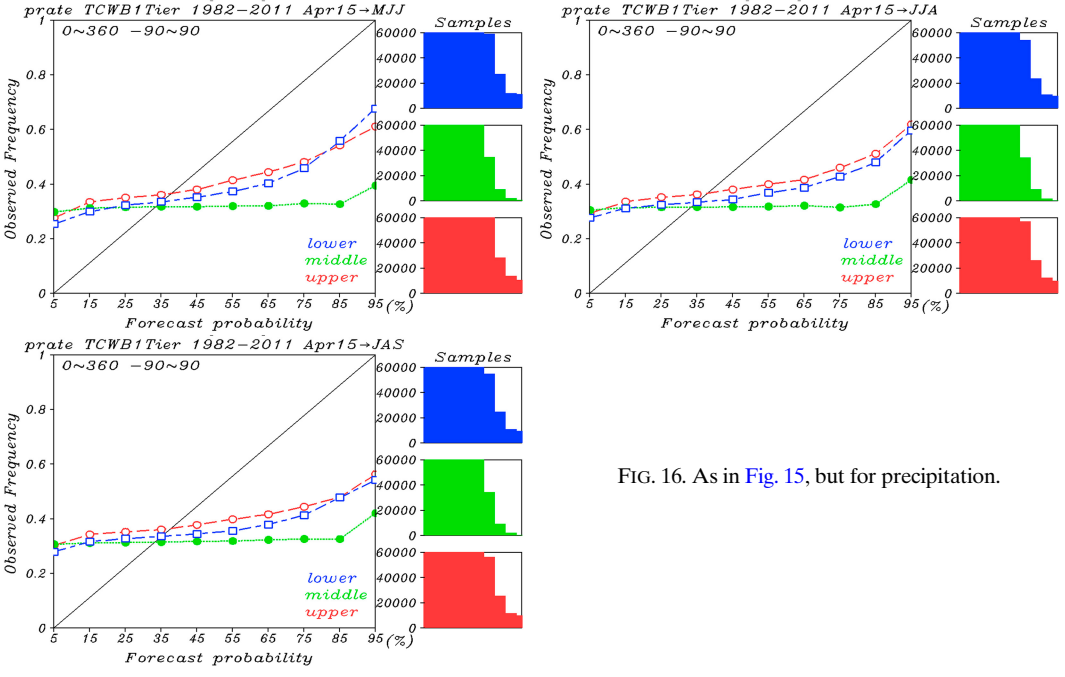

FIG. 16. As in Fig. 15, but for precipitation.

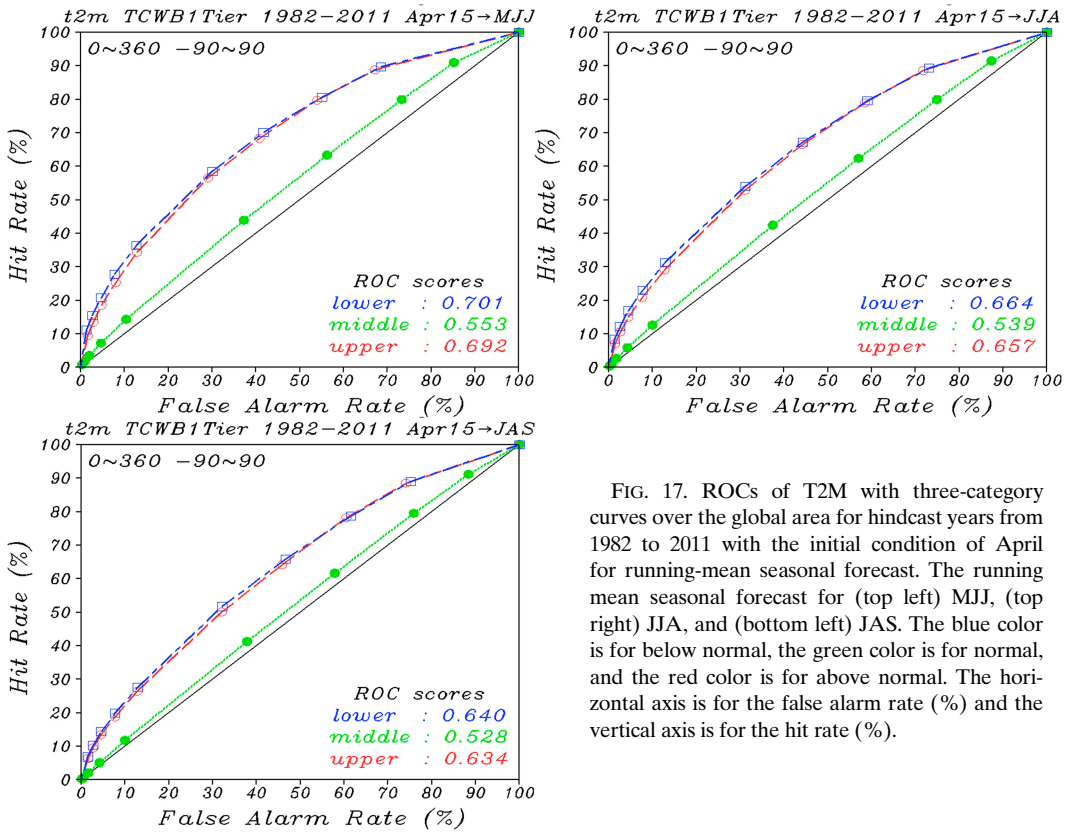

FIG. 17. ROCs of T2M with three-category curves over the global area for hindcast years from 1982 to 2011 with the initial condition of April for running-mean seasonal forecast. The running mean seasonal forecast for (top left) MJJ, (top right) JJA, and (bottom left) JAS. The blue color is for below normal, the green color is for normal, and the red color is for above normal. The horizontal axis is for the false alarm rate (%) and the vertical axis is for the hit rate (%).

NCEP CFSv2, respectively, and their difference is shown in Fig. 8e. In Fig. 8a, TCWB1T can maintain a correlation above 0.7 over lead 6 months from Northern Hemisphere autumn to winter, but Northern Hemisphere summer keeps the correlation under lead 4 months. In Fig. 8b, NCEP CFSv2 also shows similar behavior, but the correlation is extended above 0.7 more than lead 4 months in Northern Hemisphere summer. We also can see that TCWB1T shows a better score over the beginning forecast of Northern Hemisphere summer and lead 3–4 months of Northern Hemisphere autumn and winter than those of NCEP CFSv2, but TCWB1T shows a worse score on lead 5–6 months of Northern Hemisphere summer in Fig. 8e. This may be because TCWB1T1 is still using GFDL MOM3 as the oceanic model with 1° horizontal resolution, which is much coarser than GFDL MOM4 with 1/4° in NCEP CFSv2. Increasing the horizontal resolution of the ocean model can represent more details characteristic of the currents over Niño-3.4 (Stockdale et al. 2018), which is the main reason why NCEP CFSv2 has better skill in Northern Hemisphere summer. Figs. 8c, 8d, and 8f have same configuration with Figs. 8a, 8b, and 8e, but with split climatology for the calculation of SST temporal correlation over Niño-3.4, which shows that the correlation has some improvement in both TCWB1T and NCEP CFSv2 after using split climatology in Figs. 8c and 8d, especially for NCEP CFSv2, so we can see that the difference of correlation between TCWB1T and NCEP CFSv2 decreased in Fig. 8f as compared to Fig. 8e.

b. *Results over the area of 180°–180° and 40°S–40°N*

In addition to investigating the scores over the Niño-3.4 area, we would like to expand the investigated area to cover the global band from 40°S to 40°N on the empirical orthogonal functions (EOFs) of the variance of SST anomalies. Figure 9 shows the fraction of variance explained by EOF modes from 1 through 5 for SST anomaly of TCWB1T as compared to ERSST, CFSRR, and OPGSST for the lead 0 prediction. The variances of mode 1 from model results of ERSST, CFSRR, TCWB1T, and OPGSST are 32.63%, 36.8%, 39.14%, and 53.11%, respectively. The result of OPGSST is different from that of the others, OPG is a statistical model, and it happened to be the best only in the first mode in this case. For all other modes, the variances of TCWB1T and CFSRR are close to observation ERSST, which indicates that CFSRR and TCWB1T are capable of handling annual variation of ENSO.

Messié and Chavez (2011) mentioned that the leading mode in the SST anomalies explains better than any other modes. The EOF1 displays the well-known ENSO pattern, with most of the variance in the central/eastern equatorial Pacific. Figure 10 shows EOF spatial patterns of the first mode of the SST anomaly for ERSST, CFSRR, TCWB1T, and OPGSST. TCWB1T and CFSRR had narrower warm anomalies than ERSST over equatorial east Pacific Ocean. TCWB1T has a weaker cold anomaly than ERSST over the Northern Hemisphere central Pacific Ocean, and TCWB1T has a cold anomaly west of the

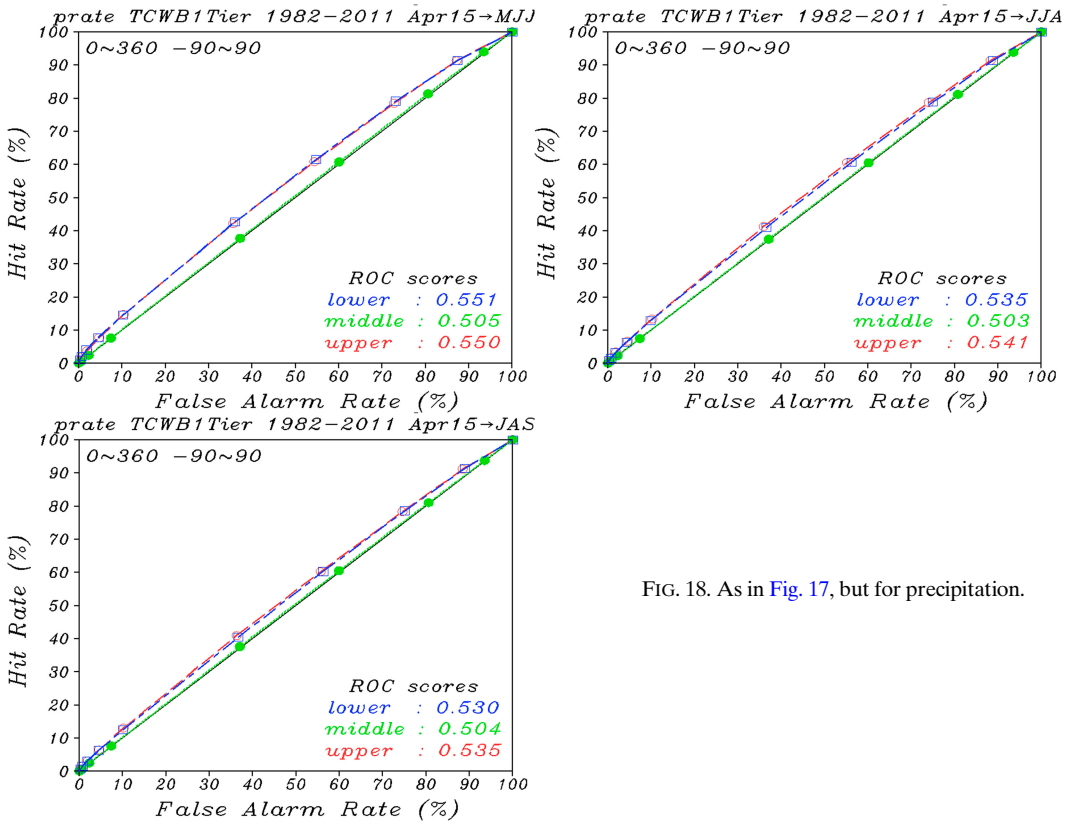

FIG. 18. As in Fig. 17, but for precipitation.

warm anomaly over the tropical Pacific Ocean. CFSRR is closer to ERSST than TCWB1T on the spatial distribution of cold anomaly over the Pacific Ocean. Even though OPGSST had much larger difference of variance of the first EOF mode to ERSST than others, it was closest to ERSST on spatial distribution of the first EOF mode than others.

c. *Results over the area of 180°–180° and 60°S–60°N*

We continue our comparison of TCWB1T and NCEP CFSv2 hindcasts but expand the area covering the global band from 60°S to 60°N. First, we emphasized the correlation percentage of forecast skill by counting the number of grid points for anomaly correlation with the correlation coefficient at the 99% significance level. Figure 11 shows the correlation percentage of forecast skill with red dots from TCWB1T and blue dots from NCEP CFSv2. The forecast scores of SST, T2M, and precipitation of TCWB1T are slightly better than those of NCEP CFSv2, except lead 3 and 5 months in T2M. The overall results of TCWB1T are comparable with NCEP CFSv2.

Ranked probability skill score (RPSS; Müller et al. 2005; Weigel et al. 2007) measures the improvement of the multicategory probabilistic forecast relative to a reference forecast (usually the long-term or sample climatology). Taking the climatological frequency into account, the values of RPSS are between negative infinity and 1; when RPSS is equal to 1, it means it is a perfect forecast. In Figures 12–14, we used three categories to calculate RPSS and show the first season forecast of SST with respect to ERSST, T2M with respect to CFSR, and precipitation with respect to GPCP from TCWB1T with initial conditions of February, May, August, and November. Wherever the shaded color is deeper while the value of RPSS is approaching 1, the forecast is better. There are higher RPSS scores of SST over the Pacific, Indian, and tropical Atlantic Oceans. There are higher RPSS scores over 0.6 for T2M along the tropical area. Nevertheless, the higher RPSS scores for rainfall are concentrated over the tropical Pacific Ocean. This indicates that SST and T2M correlation scores are significantly better than precipitation. Overall, February is the preferable initial condition to forecast MAM for all three variables as compared to other initial conditions to forecast other three seasons. The worst initial condition for SST is May, for T2M it is November, and for precipitation it is May.

d. *Results over the global area*

Due to the analyzed package in CWB, we show the forecast reliability over the global area and using hindcast data for the period from 1982 to 2011. And we discuss forecast reliability based on the reliability diagrams (RDs; Hartmann et al. 2002) and the relative operating characteristics (ROCs; Mason 1982; Jolliffe and Stephenson 2008). Since all results of RD and ROC from different initial months showed similar features for T2M and precipitation, we use April as the initial condition forecasting the running seasonal means of May–July (MJJ), June–August (JJA), and July–September (JAS) for T2M and precipitation as an example.

A reliability diagram is used to analyze the relationship between observed frequency (OF) and forecast probability

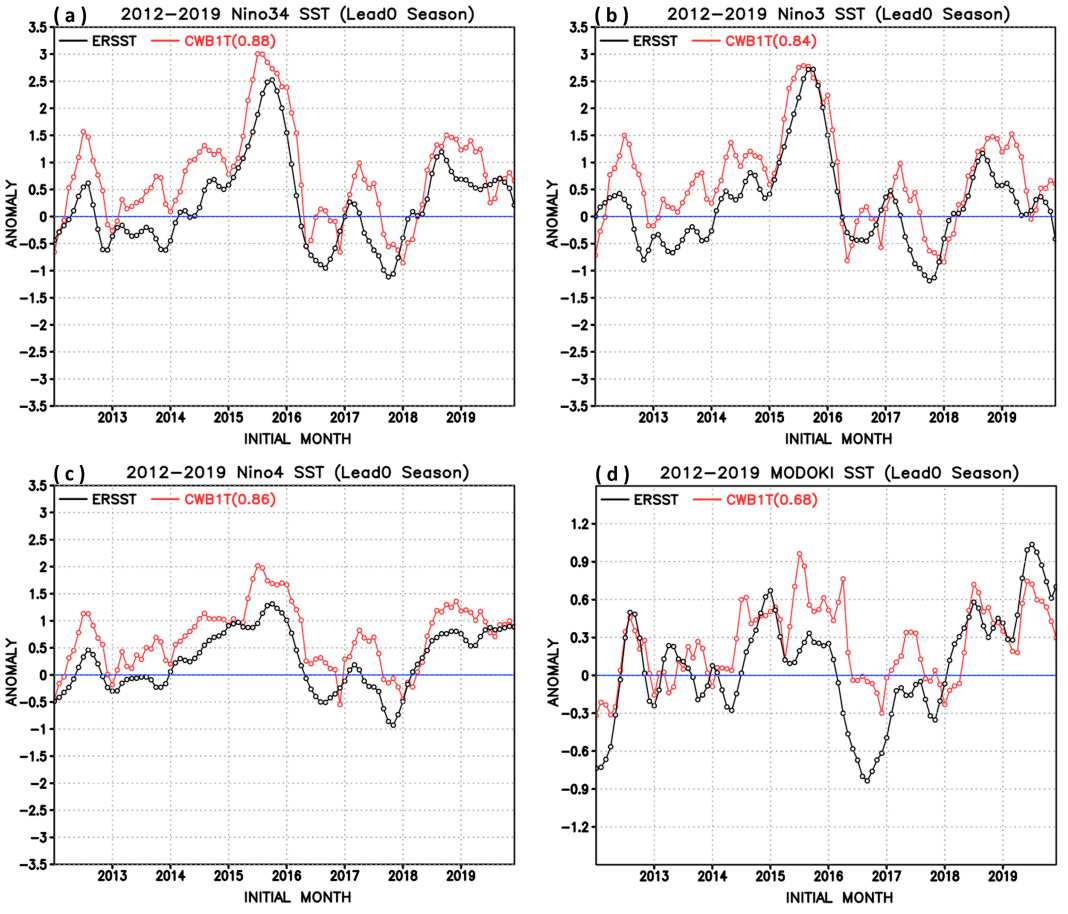

FIG. 19. The lead 0 monthly forecast results of SST anomaly over (a) Niño-3.4, (b) Niño-3, (c) Niño-4, and (d) Modoki areas from 2012 to 2019 by TCWB1T (red curves) as compared to ERSST (black curves).

(FP). When FP is equal to OF, the result is perfect reliability. Figure 15 shows the RD of T2M, the black line from the top-right corner to the bottom-left corner is the perfect reliability, and the red, green, and blue curves are for anomalies that are above normal, normal, and below normal, respectively. Overall, when the value of FP is larger than 0.45, these three curves show the overforecast probability (or overconfidence), and the anomaly above normal (red curve) shows the closest value to perfect reliability (black line) compared with the normal and below-normal curves. When the value of FP is smaller than 0.35, these three curves show underforecast probability (or underconfidence) and the anomaly below normal (blue curve) is closest to perfect reliability (black line) compared with the normal and above-normal curves. The slope of three curves in the first running season is larger (means better reliability) than those of the second and third running seasons.

Figure 16 shows the same features as Fig. 15 for precipitation, except these three curves are packed together and have much less slope than T2M in Fig. 15. From the sampling for these three categories, there are much smaller values for precipitation in Fig. 16 than those for T2M in Fig. 15. This may be why precipitation has less reliability than T2M. In summary, Figs. 15 and 16 show that the model has limited resolution or less dispersion since the probability of occurrence does not change much with forecast probability, and they are overconfident in general, especially for precipitation, which is approaching uniform overconfidence.

Figures 17 and 18 show ROC diagrams for T2M and precipitation, respectively, with three categories of below-normal, normal, and above-normal anomalies using 10 probability thresholds for running-mean seasonal forecasts of MJJ, JJA, and JAS with April as the initial condition from TCWB1T hindcast years from 1982 to 2011. The horizontal axis is for the false alarm rates and the vertical axis is for the hit rates, so the black line from the lower-left corner to the upper-right corner indicates no reliable forecast with the same values of hit and false alarm rates. When the curves are away from the black line and closer to the upper-left corner, the forecast reliability scores are higher. The lower-right corner of each diagram shows the ROC values of three categories: the higher the value, the better the reliability. It shows that the normal group has scores below 0.56, which is no score, and the other two categories have scores above 0.64. For T2M in Fig. 17, the first running-mean season is about 0.7, which is considered a skillful discrimination (Buizza et al. 1999) and better than the second running-mean season, which is about 0.66. For the

precipitation in Fig. 18, no matter the category, there is no score larger than 0.56, so precipitation is considered to have nearly no skill for these three running-mean seasons.

5. 2012–19 forecast results

With the reasonable hindcast results as mentioned in the previous section, we used the same frozen version of TCWB1T as in the hindcast for forecasts starting the years after 2011 and based on the past hindcast climatological mean of years from 1982 to 2011 for the anomaly prediction. We have 2012–19 results shown here to verify the forecast skill of TCWB1T for seasonal forecasts.

Figure 19 shows SST anomaly for (i) Niño-3.4, (ii) Niño-3, (iii) Niño-4, and (iv) Modoki for the forecast period from 2012 to 2019 for lead 0 monthly results from TCWB1T (red curves) as compared to ERSST (black curves). Modoki is referred to as the El Niño Modoki index (EMI; Ashok et al. 2007), which is defined as

$$\mathrm{EMI} = [\mathrm{SSTA}]_A - 0.5 \times [\mathrm{SSTA}]_B - 0.5 \times [\mathrm{SSTA}]_C, \quad (1)$$

where the brackets represent the area-averaged sea surface temperature anomaly (SSTA) over region A (165°E–140°W, 10°S–10°N), region B (110°–70°W, 15°S–5°N), and region C (125°–145°E, 10°S–20°N). The results in Fig. 19 show that the evolutions of anomaly phases from TCWB1T are very similar to ERSST for these four different areas. And there is evidence that El Niño events are stronger and La Niña events are weaker as compared to ERSST. Also, in Fig. 19d, the EMI score is worse than others, possibly because the forecast results of TCWB1T have a large error over the east Pacific area (region B) that contributes to worse EMI results.

We selected a strong El Niño case from 2012 to 2019 as an example to investigate the TCWB1T forecast result. Figure 20 shows the DJF anomaly over the entire globe from (i) TCWB1T with initial conditions in November 2015 as compared to (ii) ERSST and (iii) their difference. The major positive anomaly areas include the western coast of America, the tropical area of the central Pacific Ocean, and Atlantic and Indian Oceans. TCWB1T has a similar negative anomaly to ERSST over 40°–60°N of northern Atlantic and western Pacific Oceans. It indicates that the results from TCWB1T catch all major anomaly areas. Figure 20c shows the difference between TCWB1T and ERSST; TCWB1T has a stronger El Niño compared to ERSST, which is the same conclusion as in Fig. 19.

After we examined SST results of TCWB1T during the forecast period, we compared TCWB1T forecast results with the results from NCEP CFSv2, within 2012 and 2019. In the forecast period, TCWB1T has 30 members per month (0000 UTC daily). NCEP CFSv2 data are still from the North American Multi-Model Ensemble (NMME) website, which has 32 members in a month (the last day of the previous month and 7 days of this month with four cycles of 0000, 0600, 1200, and 1800 UTC). We used only the first 24 members because the 25th–32nd members sometimes are not available. As the same as the previous paragraph, we compared both anomaly scores of SST, T2M, and precipitation in Fig. 21. TCWB1T may show higher

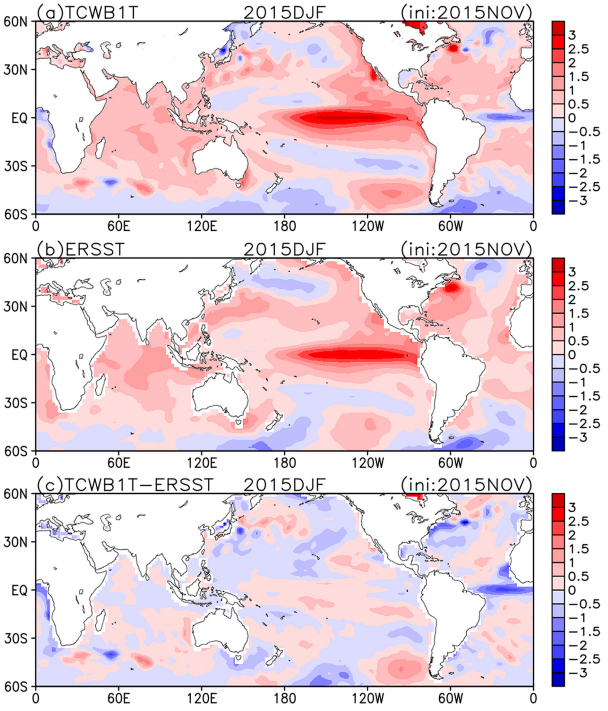

FIG. 20. The lead 0 DJF forecast results of SST anomaly from (a) TCWB1T with initial conditions from November 2015 as compared to (b) ERSST and (c) the difference between TCWB1T and ERSST.

skill of SST and T2M than NCEP CFSv2 because TCWB1T has larger spread in initial conditions than those of NCEP CFSv2. Comparing Fig. 11 with Fig. 21, we can see that forecast SST, T2M, and precipitation show lower scores in forecast than those in hindcast for both TCWB1T and NCEP CFSv2. A possible reason why both models have low scores in forecast years is because both models have the lack of skill for modeling the continuous ENSO events, which are shown in Fig. 19 during the forecast years.

We can use the first month result in Fig. 21 as an example to show the 2D diagram in Fig. 22 for comparison between TCWB1T and NCEP CFSv2 on anomaly correlation with respect to observations or reanalysis. It is clearly shown that the shaded areas of TCWB1T are larger in coverage and darker in color than those of NCEP CFSv2 for SST and T2M, which also applies to other forecast months, and there is not much difference in precipitation.

6. Discussion and conclusions

It can be concluded that building the TCWB1T coupled model by adopting the NCEP CFSv1 coupled model and substituting the atmospheric model with CWB GSM is a feasible method. We let the model stably predict the SST from the fully coupled system first and then reduced bias with adjustment of physical parameterization. In the preliminary results before adjustment, the model forecasts of SST showed cold bias increasing

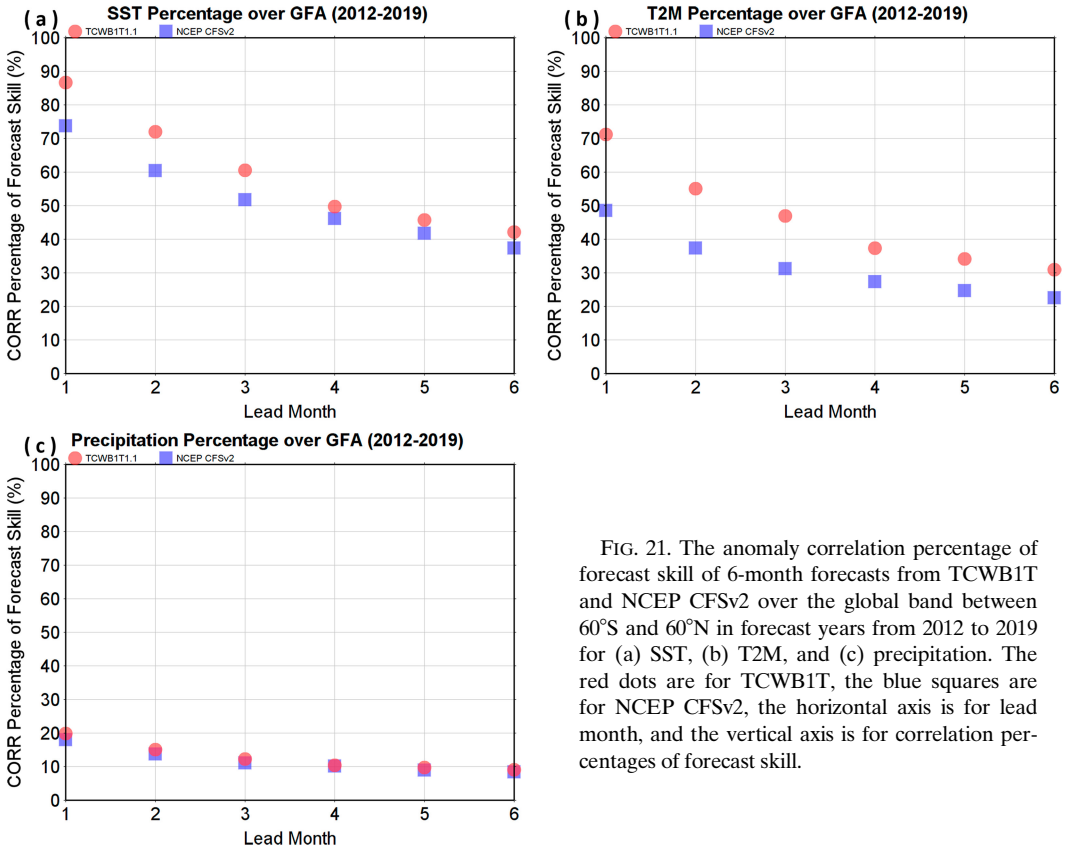

FIG. 21. The anomaly correlation percentage of forecast skill of 6-month forecasts from TCWB1T and NCEP CFSv2 over the global band between 60°S and 60°N in forecast years from 2012 to 2019 for (a) SST, (b) T2M, and (c) precipitation. The red dots are for TCWB1T, the blue squares are for NCEP CFSv2, the horizontal axis is for lead month, and the vertical axis is for correlation percentages of forecast skill.

with time in the tropical area. The adjustment and evaluation processes found that there was too little water vapor in the air of the model. We found that increasing the saturated relative humidity threshold in the model-physics parameterization can successfully keep the water vapor content in the air and reduce the cold bias of SST via the greenhouse effect. This parameterization adjustment happened to be very successful; however, it may not be the only solution to reduce the bias of cooling SST.

After adjusting the SST bias, the other meteorological fields are also relatively reasonable among many tests as compared to analysis data, such as sea level pressure field, 500-hPa geopotential field, and 850-hPa and surface wind field. The results of the ENSO forecast tests are also quite reasonable. Thus, reducing SST bias could be the main key to improving ENSO forecasts for TCWB1T.

When designing the reforecast experiment for TCWB1T using 0000 and 1200 UTC initial conditions, the same initial dates as used in NCEP CFSv1, 6 days per month, the monthly mean anomaly results of 0000 and 1200 UTC of the same six initial dates are similar, but the monthly mean anomaly results for the 0000 UTC of different 12 initial dates are quite different. This implies that the spread of different cycles is less than the spread of different dates. Thus, we chose to increase the number of initial-condition dates from 6 to 30 days as a lagged ensemble, but only run 0000 UTC for TCWB1T. In line with the Taiwan CWB operational suite to provide operational forecast information at the end of each month, the model needs to provide forecast data on the 20th of each month.

For the Niño-3.4 index in forecasting DJF, it is shown that half the hindcast years have an SST anomaly pattern correlation larger than 0.5, and only 4 years have negative anomaly pattern correlation, so TCWB1T has skill on forecasting DJF for the SST anomaly. TCWB1T has similar skills as compared to NCEP CFSv2 based on anomaly correlation, RMSE, and climatology bias. In general, TCWB1T is worse than NCEP CFSv2 at the first month, but in the second month later, TCWB1T has similar scores to NCEP CFSv2 in terms of RMSE and anomaly correlation, except it is better in terms of bias. In comparison with NCEP CFSv2, based on SST anomaly pattern correlation in different months with different lead forecast, it is better in the beginning for the Northern Hemisphere summer but worse after lead 4-month forecast. A possible reason may be because TCWB1T1 is using coarser resolution (MOM3) than those (MOM4) in NCEP CFSv2. However, TCWB1T has better score after the lead 3-month forecast than NCEP CFSv2 for Northern Hemisphere winter. For the correlation percentage of forecast skill comparison, TCWB1T has a similar score as NCEP CFSv2 for the hindcast period. Surprisingly, TCWB1T has better scores than those of NCEP CFSv2 in T2M and SST during the forecast period. It is possible that both models have lower scores during the forecast period than the hindcast period because they lack skills for prediction on continuous ENSO, which should be

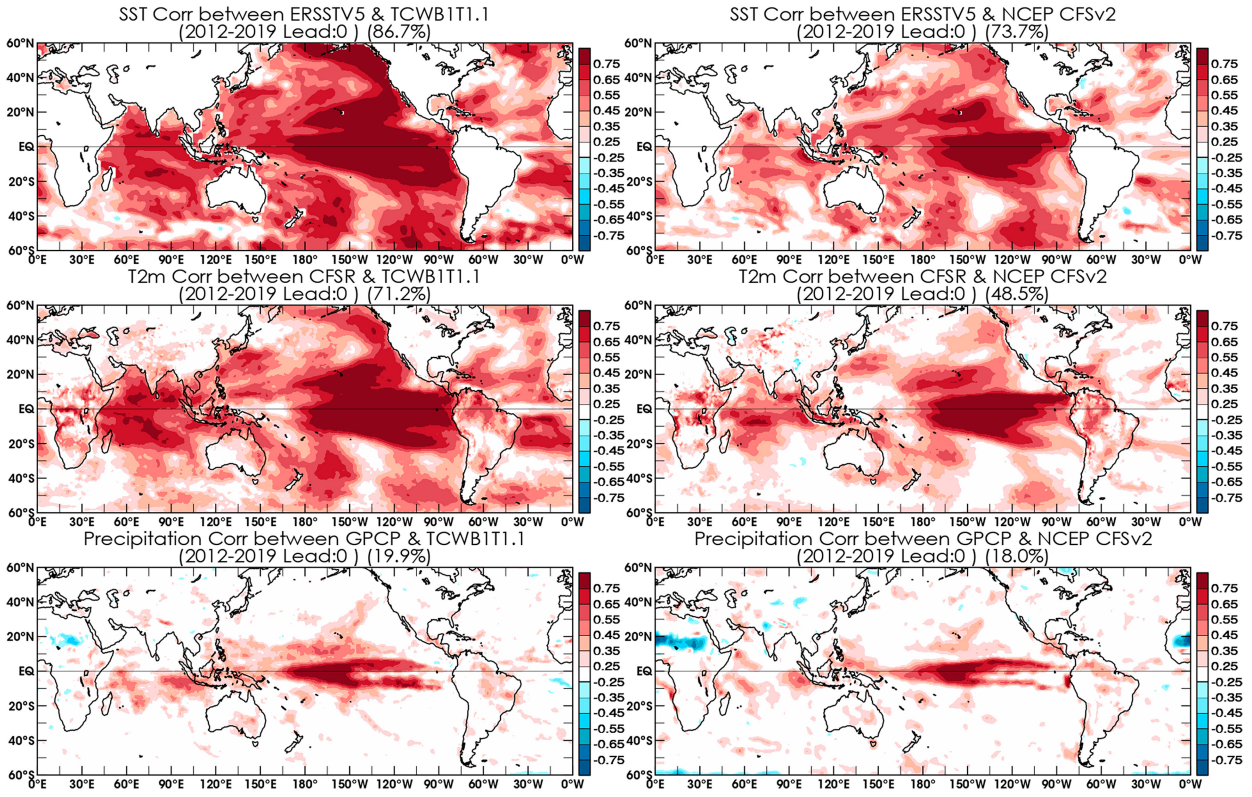

FIG. 22. The first month forecast anomaly correlations of (top) SST, (middle) T2M, and (bottom) precipitation for (left) TCWB1T and (right) NCEP CFSv2 with respect to observations, averaged from 2012 to 2019. Only correlations with an absolute value greater than 0.25 are shaded. The percentiles of shaded areas with respect to all grid points between 60°S and 60°N are indicated.

investigated in the future. TCWB1T has similar variances of EOF to ERSST and CFSRR for the lead 0 prediction for December, which indicates that TCWB1T is comparable to CFSRR to handle annual variation of ENSO.

There are higher RPSS scores for T2M along the tropical area. Nevertheless, it has a higher RPSS score for rainfall concentrated over the tropical Pacific Ocean. Overall, February is the preferable initial condition for SST, T2M, and precipitation. For the RD of T2M over the global area, overall, the forecasts have limited resolution or are less dispersive, and they are overconfident in general. For RD of precipitation over the global area, it approaches uniform overconfidence. For ROC of T2M, the first running-mean season is considered skillful discrimination and is better than the other two running-mean seasons. For ROC of precipitation, there is nearly no skill for all running-mean seasons. Through all these analyzed methods, we can conclude that TCWB1T can be used for a seasonal forecast, especially for the first season, for SST and T2M on probabilistic forecasts.

We considered ENSO, west North Pacific, and East Asia forecasts with relation to the Taiwan area. The forecast SST anomalies over Niño-3.4, Niño-3, Niño-4, and Modoki have a relative correlation coefficient within 0.68 to 0.88 during 2012–19 as compared to observed ERSST. TCWB1T is inclined to have weak La Niña but better El Niño forecasts. It can be shown in the case of the November 2015 DJF SST anomaly that TCWB1T caught all major anomaly areas. A possible reason that TCWB1T shows a higher anomaly correlation percentage of forecast skill of SST and T2M than NCEP CFSv2 is that TCWB1T has a larger spread in initial conditions than those of NCEP CFSv2. However, further investigation of this possible reason is necessary in the future.

The next generation of TCWB1T has been developed into a higher-resolution modeling system in both atmospheric and oceanic components than the current operational one-tier climate forecast system. The atmospheric model has been tested by using the current operational GFS with resolution as TL359L60, which has horizontal resolution increasing from 1° to 0.5° and vertical resolution increasing from 42 to 60 layers. In addition to increased spatial resolution in all dimensions, the GFS TL359L60 has improved dynamics and physics. Instead of Eulerian advection, the current operational GFS uses NDSL advection, which has advantages of (i) avoiding first-guess and iteration errors for advection, (ii) time saving with a large time step, and (iii) no spectral noise for tracer advection and also has an option to have reduced linear grid or reduced cubic octahedral grid. The original development of this project started more than 10 years ago, and we initially followed NCEP CFSv1 for coupling frequency without altering the ocean model and the associated coupling. We will consider testing a higher frequency coupling for the next version. The oceanic model has been tested with GFDL MOM5, and MOM6 will be tested as well.

Furthermore, as mentioned, the major objective of this TCWB1T is emphasized on ENSO seasonal prediction for operational use. Nevertheless, in the future, in addition to the seasonal forecasts, the development of TCWB1T will include monthly and 2–4-week subseasonal predictions, emphasizing the Madden–Julian oscillation (MJO), boreal summer intraseasonal oscillation (BSISO), monsoon index, and mei-yu forecasts.

Acknowledgments. During the project years, Dr. Juang has been salary-employed by NCEP in the United States and nonsalary appointed as adjunct professor by National Central University (NCU) and overseas consultant by CWB in Taiwan. We also want to recognize the support from CWB and NCU to Dr. Juang for this work. We thank Ying-Ju Chen for reviewing and invaluable suggestions and thank Benjamin Juang for proofreading. We would also very much like to thank the journal reviewers and editor for their constructive suggestions. This project had been funded by the Taiwan National Science Council through Taiwan CWB P5 plan contracted to Taiwan NCU (contracts 992136N, 1002127B, 1012128D, 1022145A, 1032048D, and 1042035A).

Data availability statement. Data are available through collaboration with the Taiwan Central Weather Bureau.

REFERENCES

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, **4**, 1147–1167, https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Anderson, D., and Coauthors, 2003: Comparison of the ECMWF seasonal forecast systems 1 and 2, including the relative performance for the 1997/8 El Nino. ECMWF Tech. Memo. 404, 95 pp., https://doi.org/10.21957/bnb7k5yjf.

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007, https://doi.org/10.1029/2006JC003798.

Banzon, V. F., R. W. Reynolds, and T. M. Smith, 2010: The role of satellite data in extended reconstruction of sea surface temperatures. *Proc. "Oceans from Space" Venice 2010*, Venice, Italy, European Commission, 27–28, https://doi.org/10.2788/8394.

Barnston, A. G., and M. K. Tippett, 2013: Predictions of Nino3.4 SST in CFSv1 and CFSv2: A diagnostic comparison. *Climate Dyn.*, **41**, 1615–1633, https://doi.org/10.1007/s00382-013-1845-2.

Berrisford, P., and Coauthors, 2011: The ERA-Interim archive version 2.0. ERA Rep. Series 1, 27 pp., https://www.ecmwf.int/sites/default/files/elibrary/2011/8174-era-interim-archive-version-20.pdf.

Buizza, R., A. Hollingsworth, F. Lalaurette, and A. Ghelli, 1999: Probabilistic predictions of precipitation using the ECMWF ensemble prediction system. *Wea. Forecasting*, **14**, 168–189, https://doi.org/10.1175/1520-0434(1999)014<0168:PPOPUT>2.0.CO;2.

Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, https://doi.org/10.1029/2002JD003296.

Fu, Q., and K. N. Liou, 1993: Parameterization of the radiative properties of cirrus clouds. *J. Atmos. Sci.*, **50**, 2008–2025, https://doi.org/10.1175/1520-0469(1993)050<2008:POTRPO>2.0.CO;2.

Griffies, S. M., and Coauthors, 2005: Formulation of an ocean model for global climate simulations. *Ocean Sci.*, **1**, 45–79, https://doi.org/10.5194/os-1-45-2005.

Han, J., and H.-L. Pan, 2011: Revision of convection and vertical diffusion schemes in the NCEP global forecast system. *Wea. Forecasting*, **26**, 520–533, https://doi.org/10.1175/WAF-D-10-05038.1.

Hartmann, H. C., T. C. Pagano, S. Sorooshiam, and R. Bales, 2002: Confidence builders: Evaluating seasonal climate forecasts from user perspectives. *Bull. Amer. Meteor. Soc.*, **83**, 683–698, https://doi.org/10.1175/1520-0477(2002)083<0683:CBESCF>2.3.CO;2.

Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339, https://doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20, https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

Ji, M., A. Kumar, and A. Leetmaa, 1994: An experimental coupled forecast system at the National Meteorological Center. Some early results. *Tellus*, **46A**, 398–419, https://doi.org/10.1034/j.1600-0870.1994.t01-3-00006.x.

Johnson, S. J., and Coauthors, 2019: SEAS5: The new ECMWF seasonal forecast system. *Geosci. Model Dev.*, **12**, 1087–1117, https://doi.org/10.5194/gmd-12-1087-2019.

Jolliffe, I. T., and D. B. Stephenson, 2008: Proper scores for probability forecasts can never be equitable. *Mon. Wea. Rev.*, **136**, 1505–1510, https://doi.org/10.1175/2007MWR2194.1.

Juang, H.-M. H., 2004: A reduced spectral transform for the NCEP seasonal forecast global spectral atmospheric model. *Mon. Wea. Rev.*, **132**, 1019–1035, https://doi.org/10.1175/1520-0493(2004)132<1019:ARSTFT>2.0.CO;2.

——, 2007: Semi-Lagrangian advection without iteration. *Proc. Conf. on Weather Analysis and Forecasting*, Taoyan, Taiwan, Central Weather Bureau.

——, 2008: Mass conserving and positive-definite semi-Lagrangian advection in NCEP GFS: Decomposition of massively parallel computing without halo. *Proc. 13th Workshop on Use of High Performance Computing in Meteorology*, Reading, United Kingdom, European Centre for Medium-Range Weather Forecasts, 50, https://www.ecmwf.int/sites/default/files/elibrary/2008/15346-mass-conserving-and-positive-definite-semi-lagrangian-advection-ncep-gfs-decomposition.pdf.

——, and M. Kanamitsu, 2001: The computational performance of the NCEP seasonal forecast model on Fujitsu VPP5000 at ECMWF. *Developments in Teracomputing: Proc. Ninth ECMWF Workshop on the Use of High-Performance Computing in Meteorology*, Reading, United Kingdom, World Scientific, 338–347, https://doi.org/10.1142/9789812799685_0029.

——, W.-K. Tao, X. Zeng, C.-L. Shie, S. Lang, and J. Simpson, 2007: Parallelization of the NASA Goddard cumulus ensemble model for massively parallel computing. *Terr. Atmos. Ocean. Sci.*, **18**, 593–622, https://doi.org/10.3319/TAO.2007.18.3.593(A).

Kanamitsu, M., and Coauthors, 2002: NCEP dynamical seasonal forecast system 2000. *Bull. Amer. Meteor. Soc.*, **83**, 1019–1038,

https://doi.org/10.1175/1520-0477(2002)083<1019:NDSFS>2.3.CO;2.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Liou, C.-S., C.-T. Terng, W.-S. Kau, T. Rosmond, C.-S. Chen, J.-H. Chen, and C.-Y. Tsai, 1989: Global weather forecast system at Central Weather Bureau. *Pap. Meteor. Res.*, **12**, 205–228.

——, and Coauthors, 1997: The second-generation global forecast system at the Central Weather Bureau in Taiwan. *Wea. Forecasting*, **12**, 653–663, https://doi.org/10.1175/1520-0434(1997)012<0653:TSGGFS>2.0.CO;2.

Mason, I. B., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.

Messié, M., and F. Chavez, 2011: Global modes of sea surface temperature variability in relation to regional climate indices. *J. Climate*, **24**, 4314–4331, https://doi.org/10.1175/2011JCLI3941.1.

Müller, W. A., C. Appenzeller, F. J. Doblas-Reyes, and M. A. Liniger, 2005: A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. *J. Climate*, **18**, 1513–1523, https://doi.org/10.1175/JCLI3361.1.

Pacanowski, R. C., and S. M. Griffies, 1999: The MOM 3 manual. Geophysical Fluid Dynamics Laboratory Tech. Rep., 708 pp., https://mom-ocean.github.io/assets/pdfs/MOM3_manual.pdf.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 589–662.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15**, 1609–1625, https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2.

——, T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, https://doi.org/10.1175/2007JCLI1824.1.

——, N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2009: What's new in version 2. OISST web page. NOAA/NCDC, www.ncdc.noaa.gov/sst/papers/oisst_daily_v02r00_version2-features.pdf.

Roberts, C. D., R. Senan, F. Molteni, S. Boussetta, M. Mayer, and S. P. E. Keeley, 2018: Climate model configurations of the ECMWF integrated forecasting system (ECMWF-IFS cycle 43r1) for HighResMIP. *Geosci. Model Dev.*, **11**, 3681–3712, https://doi.org/10.5194/gmd-11-3681-2018.

Roeckner, E., and Coauthors, 2003: The atmospheric general circulation model ECHAM 5. Part I: Model description. Max-Planck-Institut für Meteorologie Rep. 349, 127 pp., https://pure.mpg.de/rest/items/item_995269_4/component/file_995268/content.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517, https://doi.org/10.1175/JCLI3812.1.

——, and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1.

——, and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1.

Smith, T. M., and R. W. Reynolds, 2004: Improved extended reconstruction of SST (1854–1997). *J. Climate*, **17**, 2466–2477, https://doi.org/10.1175/1520-0442(2004)017<2466:IEROS>2.0.CO;2.

——, ——, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land–ocean surface temperature analysis (1880–2006). *J. Climate*, **21**, 2283–2296, https://doi.org/10.1175/2007JCLI2100.1.

Stockdale, T., and Coauthors, 2018: SEAS5 and the future evolution of the long-range forecast system. ECMWF Tech. Memo. 835, 83 pp., https://www.ecmwf.int/sites/default/files/elibrary/2018/18750-seas5-and-future-evolution-long-range-forecast-system.pdf.

Troen, I. B., and L. Mahrt, 1986: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Bound.-Layer Meteor.*, **37**, 129–148, https://doi.org/10.1007/BF00122760.

Vitart, F., and Y. Takaya, 2021: Lagged ensembles in sub-seasonal predictions. *Quart. J. Roy. Meteor. Soc.*, **147**, 3227–3242, https://doi.org/10.1002/qj.4125.

Weigel, A. P., M. A. Liniger, and C. Appenzeller, 2007: The discrete Brier and ranked probability skill scores. *Mon. Wea. Rev.*, **135**, 118–124, https://doi.org/10.1175/MWR3280.1.

Weng, S.-P., Y.-C. Tung, and W.-H. Huang, 2005: Predictions of global sea surface temperature anomalies: Introduction of CWB/OPGSST1.1 forecast system. *2005 Conf. on Weather Analysis and Forecasting*, Taipei, Taiwan, 341–345, https://photino.cwa.gov.tw/rdcweb/lib/cd/cd01conf/dissertation/2005-1/072.pdf.

Winton, M., 2000: A reformulated three-layer sea ice model. *J. Atmos. Oceanic Technol.*, **17**, 525–531, https://doi.org/10.1175/1520-0426(2000)017<0525:ARTLSI>2.0.CO;2.

Wu, T.-Y., H.-M. H. Juang, Y.-L. Chen, P.-Y. Liu, S.-I. Lin, J.-H. Chen, and M.-M. Lu, 2019: CWB CFS 1-Tier hindcast analysis and forecast verification. Climate Prediction S&T Digest: NWS Science & Technology Infusion Climate Bulletin Supplement, *43rd Climate Diagnostics and Prediction Workshop 2018*, Santa Barbara, CA, NOAA, 172–174, https://doi.org/10.25923/ae2c-v522.

Xue, Y., B. Huang, Z.-Z. Hu, A. Kumar, C. Wen, D. Behringer, and S. Nadiga, 2011: An assessment of oceanic variability in the NCEP Climate Forecast System Reanalysis. *Climate Dyn.*, **37**, 2511–2539, https://doi.org/10.1007/s00382-010-0954-4.