



Decadal prediction of Northeast Asian winter precipitation with CMIP6 models

Xiaoge Xin^{1,2,3} · Tongwen Wu^{1,2,3} · Mengzhe Zheng^{1,4} · Yongjie Fang^{1,2,3} · Yixiong Lu^{1,2,3} · Jie Zhang^{1,2,3}

Received: 28 March 2023 / Accepted: 9 December 2023
© The Author(s) 2023

Abstract

This study evaluates the decadal prediction skill of 13 forecast systems in predicting winter precipitation over Eurasia, contributing to the Decadal Climate Prediction Project of the Coupled Model Intercomparison Project Phase 6. Northeast Asia stands out as a region with improved decadal prediction skill for forecast years 2–5 due to the initialization. Observations show anticyclonic and cyclonic wind anomalies over the North Pacific and Northeast Asia, respectively, with southwesterly flow to the east of Northeast Asia. Ten forecast systems reproduce such circulation anomalies favoring abundant winter precipitation in Northeast Asia. The significant positive (negative) correlations between the detrended Northeast Asian precipitation (NEAP) and AMV (PDO-like) time series are reproduced by seven (nine) forecast systems. However, most forecast systems underestimate the correlation between the NEAP and the AMV, and have relatively low skill in predicting the PDO. Further improvements in these aspects will help to improve the decadal prediction skill of winter precipitation over Northeast Asia. The multi-model ensemble (MME) is able to reproduce both links of NEAP with AMV and PDO-like variability. The MME demonstrates significant skill and outperforms the individual forecast systems in predicting the NEAP for all 4-year averaged periods in the range of 1–8 years, demonstrating the benefits of using the ensemble mean of multiple models.

Keywords Decadal prediction · CMIP6 · DCPD · PDO · AMV

1 Introduction

Decadal climate predictions include both the effects of external forcings and internal climate variability, with the goal of predicting climate evolution up to 10 years ahead. The main difference between decadal predictions and historical simulations is that the former incorporates information on the current state of the climate system throughout the model initialization process to align the modeled

internal variability with the observed, while the latter aims only to reproduce the effects of external forcings such as anthropogenic greenhouse gases (Smith et al. 2007). In the fifth phase of the Coupled Model Intercomparison Project (CMIP5, Taylor et al. 2012), decadal predictions performed by climate models have a skillful prediction in sea surface temperature (SST) over the North Atlantic, western Pacific, eastern tropical Pacific, and the Indian Oceans (Wu and Zhou 2012; Doblas-Reyes et al. 2013). The Atlantic multi-decadal variability (AMV) is a prominent signal that could be skillfully predicted by decadal predictions (Meehl et al. 2014). The skillful decadal predictions over land are mainly distributed in Europe, America, and Africa, which are attributed to the predictable signal of the AMV and the associated atmospheric circulation (Corti et al. 2012; Meehl et al. 2014). Enhanced skill is also found for summer surface air temperature in East Asia, which is related to the skill in forecasting western Pacific SST (Xin et al. 2018).

It is more difficult to improve the skill of precipitation forecasts than of surface air temperature for decadal climate predictions (Goddard et al. 2012; Doblas-Reyes et al. 2013).

✉ Xiaoge Xin
xinxg@cma.gov.cn

¹ Earth System Modeling and Prediction Centre, China Meteorological Administration, Beijing 100081, China

² Key Laboratory of Earth System Modeling and Prediction China Meteorological Administration, Beijing 100081, China

³ State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China

⁴ Chinese Academy of Meteorological Sciences, China Meteorological Administration, Beijing 100081, China

Predicting near-term climate was identified by the World Climate Research Programme (WCRP) as one of the grand challenges facing international climate research. In phase 6 of CMIP (CMIP6, Eyring et al. 2016), the Decadal Climate Prediction Project (DCPP) is to enhance the prediction skill in more detailed processes (Boer et al. 2016). The differences between the experimental protocols of CMIP6 and CMIP5 include more frequent starting dates for the hindcast and larger ensembles in order to provide robust estimates of decadal predictive skill.

Recent studies have found improvements in the decadal prediction of the North Atlantic subpolar gyre SST in CMIP6 relative to CMIP5 (Borchert et al. 2021). With larger ensembles, the contribution of the initialized component of the forecast to the overall skill increases, as does the percentage of global area with predictable variance due to initialization (Sospedra-Alfonso and Boer 2020). The predictions of CMIP6 generally show high skill in predicting surface air temperature over most land regions and the AMV index, while the skill is still limited for precipitation (Delgado-Torres et al. 2022, 2023). Skill in precipitation forecasting was shown to be significant in tropical monsoon regions by decadal predictions (Dunstone et al. 2020). In a large ensemble of CMIP6 models, skillful prediction of rainfall on the interior Tibetan Plateau was found to result from the Silk Road pattern driven by the SST over the subpolar gyre region in the North Atlantic (Hu and Zhou 2021). Significant skill of precipitation is found in regions of northern Europe, western and central Africa, and northern Asia (Smith et al. 2019; Delgado-Torres et al. 2022). Although the decadal prediction skill of precipitation in northern Europe and Sahel was found to be related to the AMV (Gaetani and Mohino 2013; Sheen et al. 2017; Yeager et al. 2018; Simpson et al. 2019), few studies have investigated the mechanism for the decadal prediction skill of precipitation in Northern Asia.

The precipitation in Asia exhibits distinct seasonal features, with dominant influence from different climate variabilities in different seasons. This study explores the performance of 13 forecast systems participating in the CMIP6 DCPP in predicting winter precipitation over Eurasia. The role of initialization is identified by comparing initialized (Init) decadal hindcasts with non-initialized (NoInit) historical simulations. For the regions where the decadal prediction is skillful, the mechanisms in relation to the atmospheric circulation and the SST will be further investigated. The paper is organized as follows. Section 2 introduces the models, experiments, and methods used in this study. Section 3 presents the main results. Conclusions are given in Sect. 4.

2 Models, experiments, and methods

2.1 Models and experiments

The decadal climate predictions from 13 forecast systems contributing to the CMIP6 were used in this study. To investigate the role of initialization in the skill of decadal climate predictions, historical simulations of CMIP6 performed with the corresponding models were used for comparison. For each model, decadal hindcasts and historical simulations have the same external forcing. Therefore, the main difference between the decadal hindcast and the historical simulation of the same model is whether initialization is used.

Table 1 lists the spatial resolutions of 11 models and their ensemble sizes for decadal hindcasts and historical simulations. The horizontal resolution of these models ranges from 70 km (EC-Earth3) to 280 km (CanESM5). The ensemble size of the decadal hindcast varies from 8 to 40. The EC-Earth3 have two sets of decadal predictions using the methods of full-field initialization (EC-Earth3-i1) with 10 members and anomaly initialization (EC-Earth3-i2) with 5 members. NorCPM1 also has two sets of decadal predictions with different anomaly assimilation methods. NorCPM1-i1 uses a 1980–2010 reference climatology for computing anomalies and the data assimilation only updates the physical ocean state, while NorCPM1-i2 uses a 1950–2010 reference climatology and additionally updates the sea ice state via strongly coupled assimilation of ocean observations (Bethke et al. 2021). In total, there are 172 members for the decadal hindcasts, and 153 members for the historical simulations. Previous studies have highlighted the need for large ensembles to obtain robust estimates of the skill (Yeager et al. 2018; Smith et al. 2019; Dunstone et al. 2020). In this study, the average of the individual ensemble means of the different forecasts systems (climate models) is used as the multi-model ensemble (MME).

2.2 Methods

Model initialization can lead to initial shocks in the decadal predictions (García-Serrano and Doblas-Reyes 2012). Predictions also suffer from drift from the predicted state towards the model's mean climate, which is usually different from the observed mean climate. Therefore, the anomalies of all decadal predictions are used in the analysis. The anomalies are calculated relative to the climatology of the ensemble mean and lead times of each model as used in previous studies, taking into account the different climatology of each model (Kim et al. 2012). The climatology is calculated from the full period of each forecast year.

Table 1 Descriptions of the climate models and experiments used in this study

Model	Resolutions (horizontal grids)	Experiments	Member size	Model and experiment references
BCC-CSM2-MR	Atm: 320 × 160 Ocn: 360 × 232	Historical Hindcast	3 8	Wu et al. (2019)
CanESM5	Atm: 128 × 64 Ocn: 360 × 291	Historical Hindcast	20 20	Swart et al. (2019)
CESM1-1-CAM5-CMIP5	Atm: 320 × 160 Ocn: 360 × 232	Historical Hindcast	40 40	Yeager et al. (2018)
CMCC-CM2-SR5	Atm: 288 × 192 Ocn: 362 × 292	Historical Hindcast	11 20	Nicoli et al. (2023)
EC-Earth3	Atm: 512 × 256 Ocn: 362 × 292	Historical Hindcast (i1)	9 10	Döscher et al. (2022) Bilbao et al. 2021
FGOALS-f3-L	Atm: 288 × 180 Ocn: 360 × 218	Hindcast (i2)	5	
		Historical hindcast	6 9	He et al. 2019 Hu et al. 2023
Had-GEM3-GC31-MM	Atm: 432 × 324 Ocn: 1440 × 1205	Historical Hindcast	4 10	Williams et al. (2017)
IPSL-CM6A-LR	Atm: 144 × 143 Ocn: 362 × 332	Historical	10	Boucher et al. (2020)
		Hindcast	10	
MIROC6	Atm: 256 × 128 Ocn: 360 × 256	Historical	10	Tatebe et al. 2019
		Hindcast	10	Kataoka et al. 2020
MPI-ESM1-2-HR	Atm: 384 × 192 Ocn: 802 × 404	Historical	10	Müller et al. (2018)
		Hindcast	10	
NorCPM1	Atm: 144 × 96 Ocn: 320 × 384	Historical	30	Bethke et al. (2021)
		Hindcast (i1)	10	
		Hindcast (i2)	10	

The hindcasts of these models are initialized in different calendar months of each year starting from 1960 and perform ten-year forecasts. In order to have full calendar years in the analysis, the first months were discarded. The historical simulations of these models were carried out from 1850 to 2014. Therefore, we choose the common time period of 1961–2014. This study focuses on the 4-year average forecast

period of 2–5 years. The hindcasts for the years 1962–1965 started in 1961 are calculated and denoted as 1962. And so on, the average for the years 2011–2014 of the prediction started in 2010 is shown as 2011. A 4-year running mean is applied to the observations and historical simulations to validate the decadal predictions. In the comparisons of the decadal prediction and historical simulations (Sect. 3.1), we use 50-year hindcast denoted as 1962, 1963, ..., 2011. In the analysis of the decadal hindcasts alone (Sect. 3.2–3.4), we use 53 forecast years' over the period 1962–2014 period (start dates: 1961–2013). Since the observational data end in 2020, for the forecast years 1–4, 2–5, 3–6, and 5–8, we use 53 years of forecasts starting from 1961 to 2013 to compare with the observations. For the forecast years 6–9 and 7–10, we use 52 and 51 forecast years with the start dates over the periods 1961–2012 and 1961–2011, respectively.

The evaluation of the prediction skill is mainly in terms of the anomaly correlation coefficient (ACC). The ACC measures the linear relationship between two time series calculated at each grid point between the simulated and observed anomalies. The ACC ranges from –1 to 1. A value of 1 indicates a perfect forecast, while values close to zero or negative indicate a forecast with no skill. The statistical significance of the correlation is evaluated using the two-tailed Student's *t*-test. The effective degrees of freedom were used to take into account the autocorrelation of the time series (Bretherton et al. 1999). Winter is defined as the average of the months from November to the following March.

The AMV index is defined as the mean SST anomaly in the North Atlantic (80°W–0, 0–60°N) relative to the global mean SST (60°S–60°N). The PDO patterns are explored in both observation and predictions by performing an empirical orthogonal function (EOF) analysis on detrended winter SST anomalies over the North Pacific (20°–70°N, 110°–260°E). To highlight the role of initialization, both the linear trend of atmospheric circulation and precipitation are detrended in the analysis of atmospheric circulation associated with precipitation.

2.3 Observational data

The observed data used to validate the prediction skill are as follows. The land precipitation data are from the Climatic Research Unit Time-Series Version 4.06 (CRU-TS4.06) dataset with a horizontal resolution of 0.5° (Harris, 2020). The SST data are from the Extended Reconstructed Sea Surface Temperature (ERSST) Version 5 dataset with a resolution of 2° (Huang et al. 2017). The atmospheric circulation data were obtained from the Japanese 55-year reanalysis (JRA-55) dataset, which has a horizontal resolution of 1.25° (Kobayashi et al. 2015; Harada et al. 2016). The outputs of all models are interpolated to the corresponding

observations using the bilinear interpolation method for the validation of SST and atmospheric variables. For the validation of precipitation, the observational and model data are all interpolated to the 1.25° grid, which is consistent with the resolution of JRA-55.

3 Results

3.1 Skill in predicting winter precipitation over Eurasia

The ACC of the climate models in predicting winter precipitation over Eurasia is compared between decadal hindcasts and historical simulations for the forecast years 2–5 during 1962–2011. As shown in Figs. 1 and 2, the ACC of the decadal hindcast by each model has larger significant area than the historical simulation over the middle and high latitudes (40° – 65°N) of Eurasia. EC-Earth3-i1 (NorCPM1-i1) and EC-Earth3-i2 (NorCPM1-i2) show some differences in the distribution of the skillful prediction over northern Europe, but they have similar distribution of the skillful prediction over northern Asia (Fig. 2). Results of the decadal predictions over Eurasia are similar to the 1–5 year decadal prediction for annual mean precipitation shown in a previous study (Fig. S13 in Delgado-Torres et al. 2022). However, the historical simulations of these models generally have lower skill (Figs. 1 and 2). Both the historical simulations and the decadal hindcasts are intended to capture the forced response to changing external forcing, but only the decadal hindcast carries the initial state of the internal variability modes from the initialization. This suggests that the decadal hindcasts benefit from the initialization in predicting winter precipitation in the mid-latitudes of Eurasia. It is noted that in some areas including the Middle East, South and East Asia, the ACC is negative in most forecast systems.

The MME of the decadal predictions has significant ACC for winter precipitation over Europe and Northeast Asia (Fig. 3a). The skillful prediction of precipitation in northern Europe and Asia was also found in the CMIP5 MME (Smith et al. 2019). Similar to the individual models, the Init MME has larger area with significant ACC than the NoInit MME (Fig. 3b). After removing the linear trend, the Init MME still has significant ACC in Europe and Northeast Asia, while the NoInit MME has almost no significant skill in these regions (Fig. 3c and d). This indicates that the variability of the precipitation could be better predicted by the Init MME due to the initialization of the model. The added value due to the initialization in Asian precipitation is also found in (Delgado-Torres et al. 2022), denoted as residual ACC with the method of Smith et al. (2019).

We choose Northeast Asia (85° – 137°E , 45° – 58°N) (outlined in Fig. 3c) with significant ACC to investigate the variability of the precipitation. Time series of winter precipitation averaged over Northeast Asia are calculated in the decadal hindcasts and the historical simulations by the MME and individual forecast systems for the forecast years 2–5 (Fig. 4). The decadal hindcasts and historical simulations of all models can reproduce the increasing trend of Northeast Asian precipitation (NEAP) during 1962–2011, indicating a common impact of the external forcing. The temporal evolution of the NEAP in the Init MME is in better agreement with the observations than the NoInit MME with a higher correlation coefficient (Fig. 4a and b). The correlation coefficients of all models in the decadal prediction are significant above the 90% confidence level, while none of the models has a significant correlation coefficient in the historical simulations.

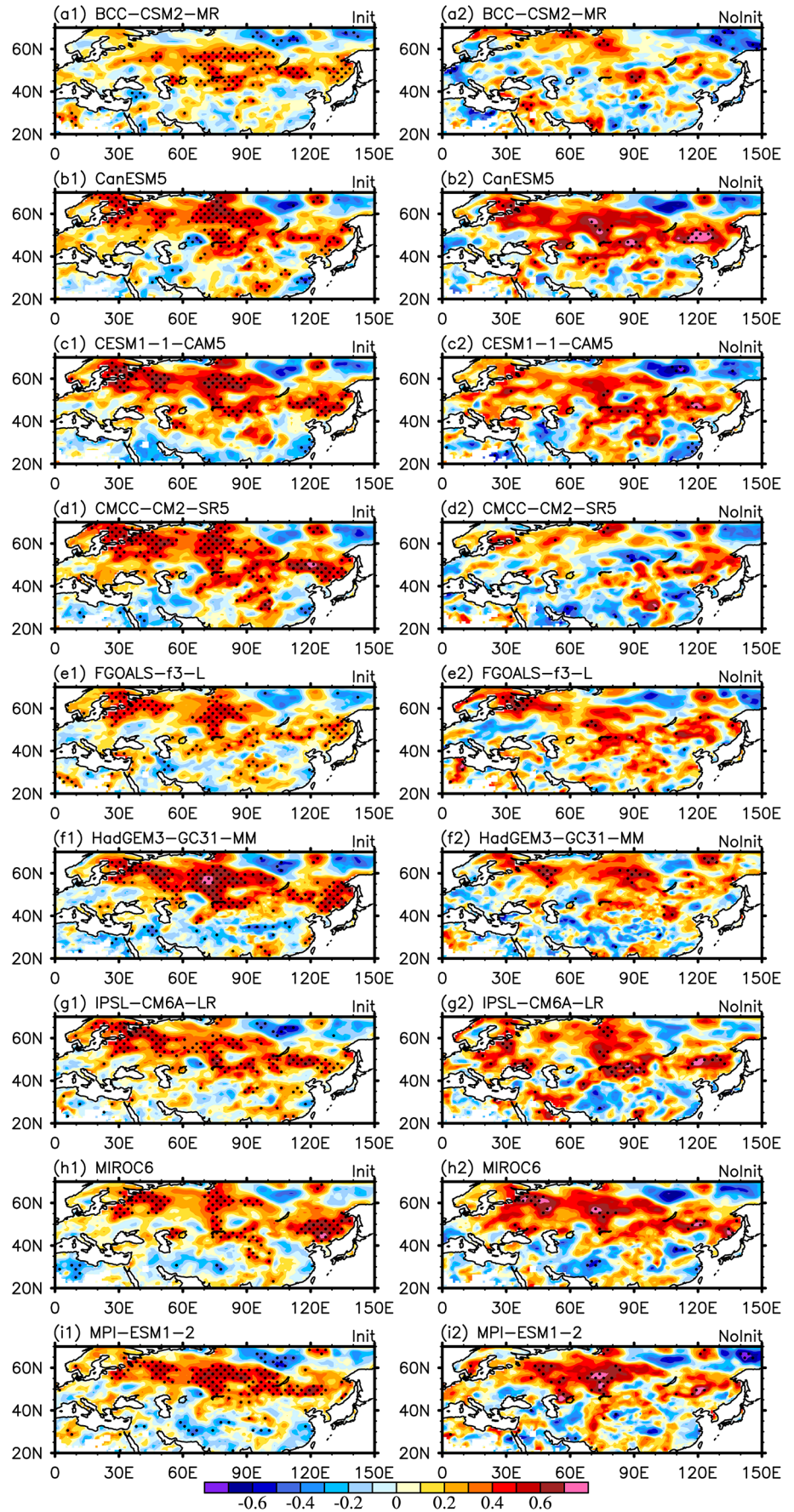
The linear trend of the time series is removed in the decadal hindcasts and historical simulations and shown in Fig. 4c and d. The correlation coefficients of the detrended NEAP in the decadal hindcasts are still above the 90% confidence level for all models except for MIROC6 (Fig. 4c). Although some models of the historical simulation have high correlation coefficients, but they do not exceed the 90% confidence level due to the small effective sample size (Fig. 4d).

The correlation coefficient of the Init MME for the detrended NEAP is 0.62, which is much higher than any of the individual models. Smith et al. (2019) found that a large ensemble is needed to reveal significant skill for precipitation over land and atmospheric circulation. This was attributed to the low signal-to-noise ratio of climate models (Scaife and Smith 2018). Recent studies have shown the benefit of using a larger ensemble for the high-latitude blocking and North Atlantic Oscillation (NAO) predictions (Athanasiadis et al. 2020; Smith et al. 2020; Donat et al. 2023). The skill of the precipitation forecasts in this study also demonstrates the benefit of using a large number of predictions.

3.2 Atmospheric circulation associated with NEAP

We further investigate the atmospheric circulations associated with the winter NEAP predicted by the decadal hindcasts. Here, the NEAP precipitation and atmosphere circulation are both detrended to reveal the climate variability. Figure 5 shows the regression of detrended winter horizontal winds at 700 hPa on the NEAP during 1962–2014. The observation shows an anomalous cyclonic circulation over Northeast Asia, and an anomalous anticyclonic circulation over the North Pacific. The southwesterly flow between the cyclonic and anticyclonic circulations tends

Fig. 1 ACC of winter precipitation predicted by decadal hindcasts (left) and historical simulations (right) of each model for forecast years 2–5 relative to observation. The dotted indicates the region where the ACC is above the 90% confidence level. Init denotes decadal hindcast, and NoInit denotes historical simulation



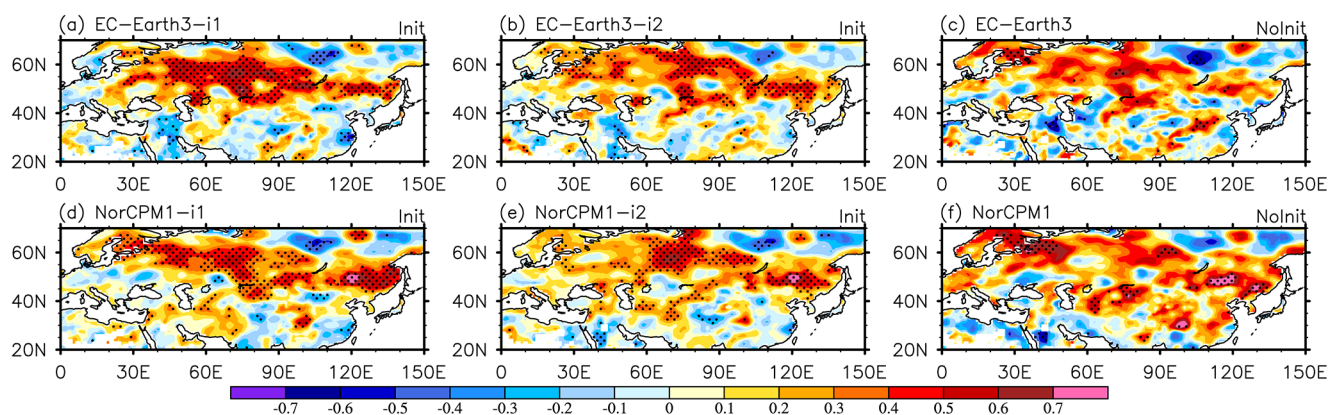


Fig. 2 ACC of winter precipitation predicted by decadal hindcasts of EC-Earth3 with initialization method i1 (a) and i2 (b), decadal hindcasts of NorCPM1 with initialization method i1 (d) and i2 (e), and historical simulation of EC-Earth3 (c) and NorCPM1 (f) relative to

observation. The dotted indicates the region where the ACC is above the 90% confidence level. Init denotes decadal hindcast, and NoInit denotes historical simulation

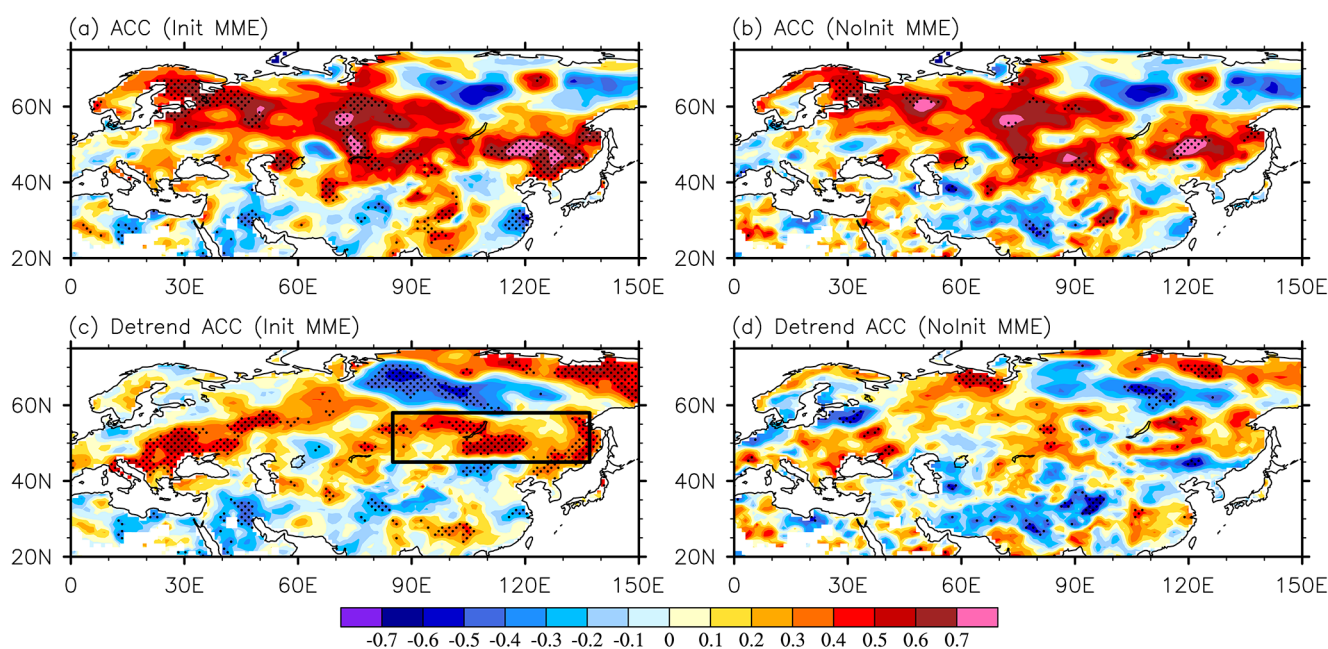


Fig. 3 ACC of winter precipitation (a-b) and detrended winter precipitation (c-d) predicted by the MME of decadal hindcasts (left) and historical simulations (right) for forecast years 2–5 relative to observation. The dotted indicates the region where the ACC is above the 90% confidence level

to bring moisture northward to Northeast Asia resulting in more precipitation (Fig. 5a). Such circulation anomalies are reproduced by ten forecast systems including BCC-CSM2-MR, CMCC-CM2-SR5, EC-Earth3-i1, EC-Earth3-i2, FGOALS-f3-L, HadGEM3-GC31-MM, IPSL-CM6A-LR, MPI-ESM1-2-HR, NorCPM1-i1 and NorCPM1-i2. The anticyclonic anomalies over the North Pacific predicted by most models indicate the possible influence from the North Pacific on NEAP in the decadal predictions. Although CanESM5 and CESM1-1-CAM5 do not capture the circulation anomalies in the North Pacific, they predict the anomalous westerlies in Northeast Asia. MIROC6 is unable to reproduce the cyclonic anomaly over Northeast Asia,

corresponding to the low skill of this model in predicting the detrended NEAP (Fig. 4c).

Note that some models and the MME predict cyclonic anomalies in the subtropical North Atlantic, including BCC-CSM3-MR, CanESM5, CESM1-1-CAM5, CMCC-CM2-SR5, EC-Earth3-i1, EC-Earth3-i2, FGOALS-f3-L, and HadGEM3-GC31-MM, NorCPM1-i2, suggesting possible links between NEAP and North Atlantic SST in the forecast systems.

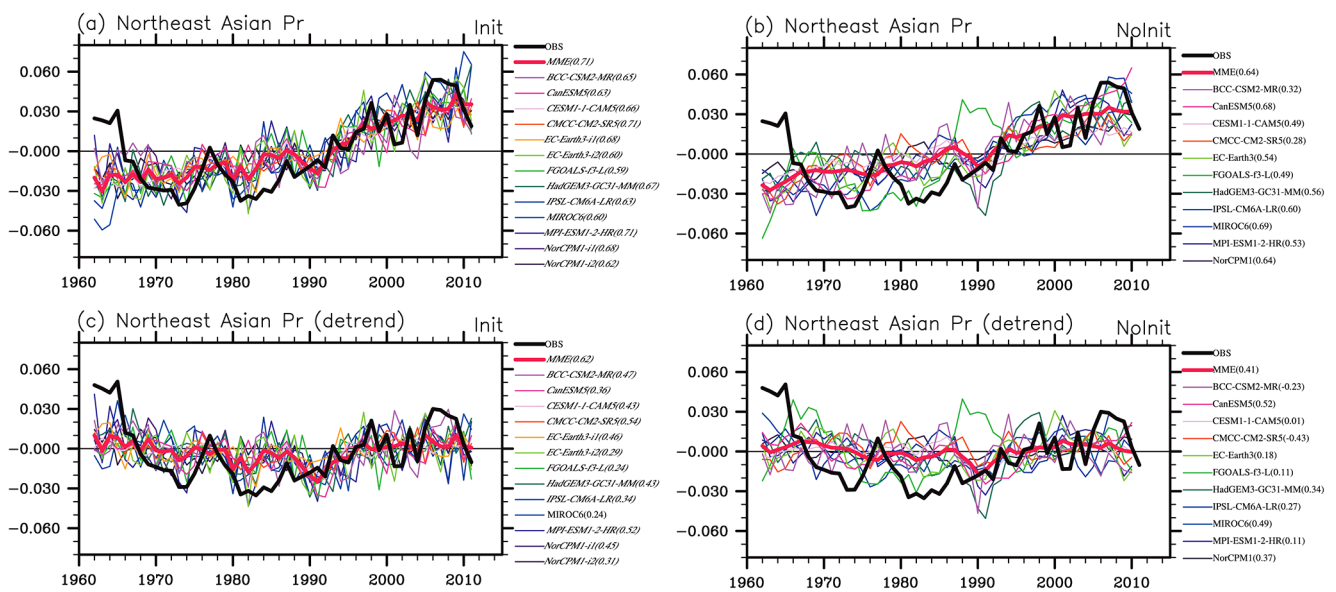


Fig. 4 Time series of winter precipitation during 1962–2014 averaged over Northeast Asia (85° – 137° E, 45° – 58° N) in the decadal forecasts (a) and historical simulations (b) by the MME and individual forecast systems for the forecast years 2–5. The detrended time series for the

decadal forecasts (c) and historical simulations (d) of each model and the MME. The number in parentheses indicates correlation coefficient between the hindcast (simulation) and the observations. Italic identifies correlation exceeding the 90% confidence level

3.3 Possible influences from the PDO and AMV

3.3.1 Prediction skill of the SST

Before investigating the possible influences of the ocean on Northeast Asian winter precipitation, we examine the skill of the forecast systems in predicting winter SST for forecast years 2–5. As shown in Fig. 6, all forecast systems have skillful SST predictions in the tropical Atlantic, while only six forecast systems, including CESM1-1-CAM5, CMCC-CM2-SR5, HadGEM3-GC31-MM, MIROC6, NorCPM1-i1, and NorCPM1-i2, have significant ACC in the subpolar North Atlantic. CESM1-1-CAM5, CMCC-CM2-SR5, IPSL-CAM6A-LR, MPI-ESM1-2-HR, and NorCPM1-i2 have relatively lower skill than other models in the central North Pacific. The MME outperforms individual models, with significant ACC in the tropical and subpolar regions of the North Atlantic, western and central North Pacific.

3.3.2 Prediction skill of the PDO and AMV

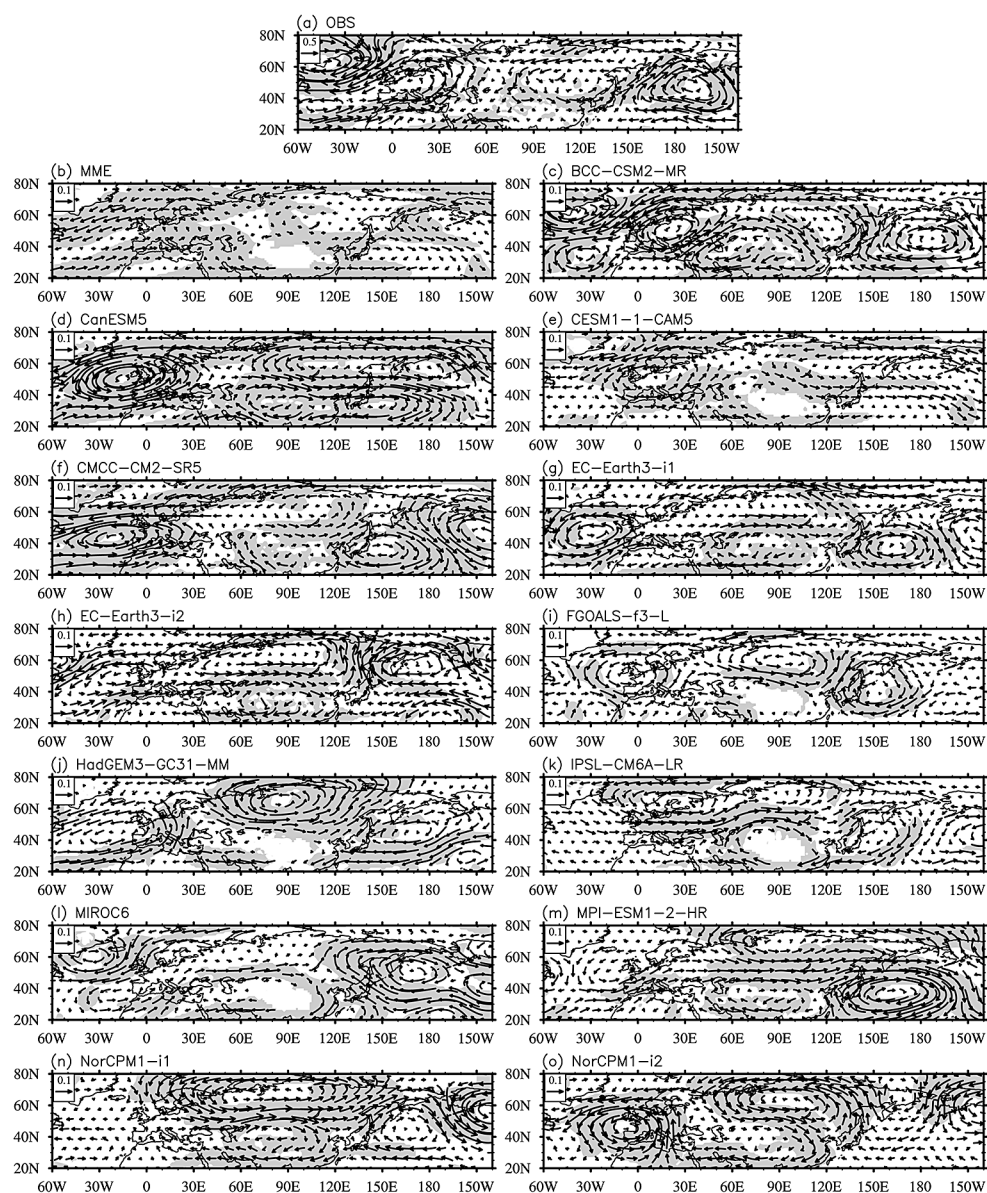
The AMV and the Pacific Decadal Oscillation (PDO) are the prominent modes for the decadal variability of the SST in the North Atlantic and the North Pacific, respectively. Previous studies have found the influence of the AMV on Siberian warm season precipitation (Sun et al. 2015), and the impact of the AMV and PDO on Asian climate (Si and Ding 2016; Huang et al. 2019). On the other hand, the AMV is considered to be one of the most predictable aspects of the decadal climate (Yeager and Robson 2017). Improved

prediction skill through initialization has also been found for the PDO, although it is less skillful than the AMV (Kim et al. 2012; Doblas-Reyes et al. 2013). Here, we further explore the prediction skill of the winter AMV and PDO predicted by the forecast systems and their relationship with the NEAP.

As described in Boer and Sospedra-Alfonso (2019), the assessment of the overall skill of the PDO is mainly approached in several ways. One approach is to expand both the forecasts and the observation in terms of the EOFs of the latter, and evaluate the skill of the associated PDO indices. Another approach is to extend the observation and the forecasts in terms of their own EOFs and compare the PDO indices. We use the second approach by performing EOFs on detrended winter SST anomalies over the North Pacific.

The observation shows a PDO pattern in the leading EOF with negative SST anomalies over the central North Pacific and positive anomalies along the eastern Pacific coast and the subtropical Pacific (Fig. 7a). The leading EOF mode explains 42% of the total variance in the observation. As shown in Fig. 7, eight forecast systems show basin-wide SST anomalies over the central North Pacific in the leading EOF with an approximate amount of the explained variance (33–49%). Five other forecast systems produce the basin-wide SST anomalies in the second EOF. However, the oppositely signed anomalies in the eastern North Pacific are missing in some models and the MME. Such biases were also present in CMIP5 models (Kim et al. 2014). Only BCC-CSM2-MR, CESM1-1-CAM5, FGOALS-fs-L, IPSL-CM6A-LR, NorCPM1-i1 and NorCPM1-i2 can partially reproduce the

Fig. 5 Regression of detrended winter horizontal winds at 700 hPa on the NEAP index in the observation (a), and hindcast by the MME (b) and each model (c–j) for the forecast years 2–5. The shaded indicates the region where the zonal wind or meridional wind is above the 90% confidence level



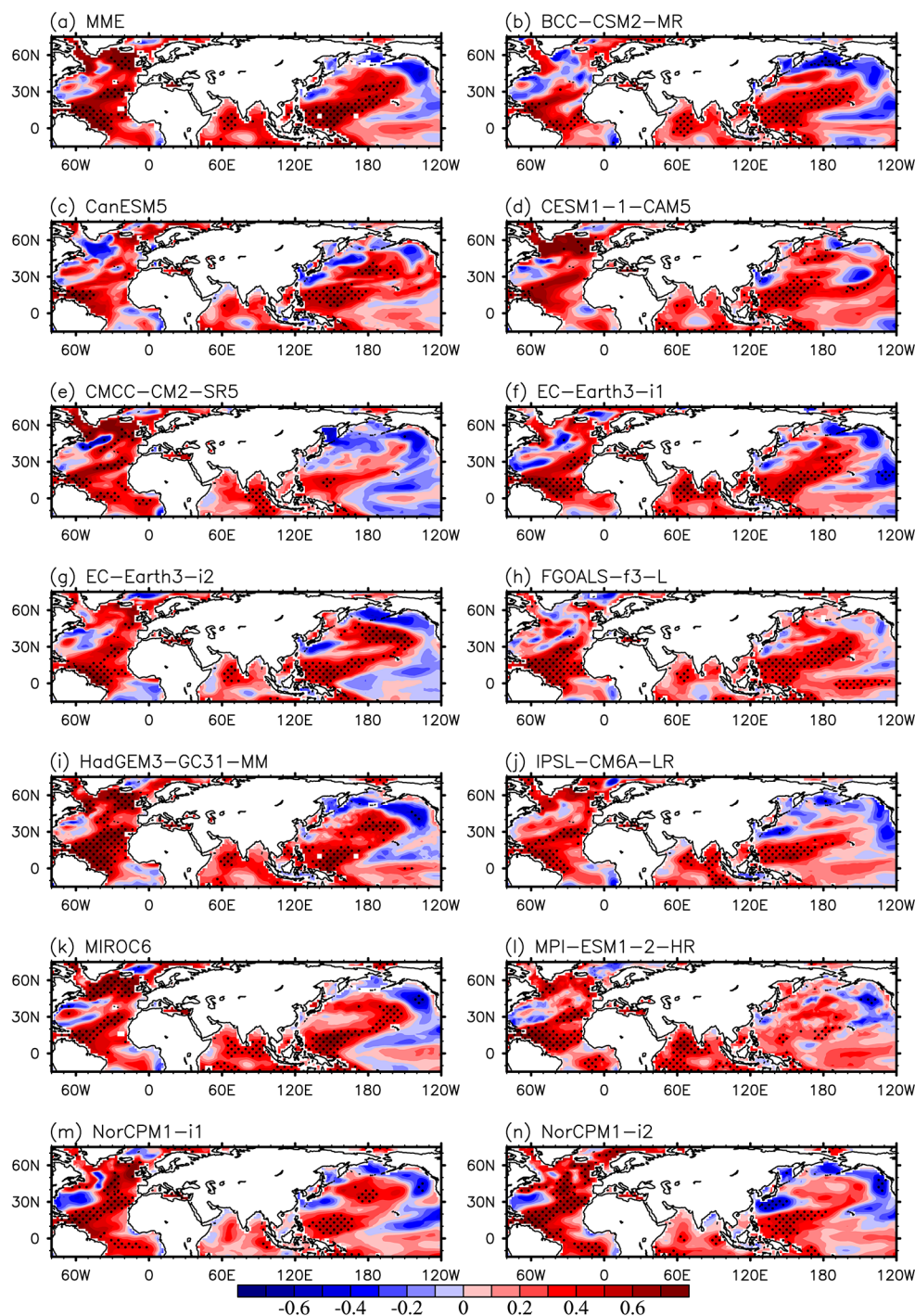
opposite variation of the SST in the eastern Pacific coast and subtropical Pacific.

Although the predicted EOF mode in the North Pacific is systematically different from the observation in some models, the corresponding principal component time series shows a decadal variability similar to the observed one (Fig. 8a). Therefore, we refer to the EOF pattern in the decadal predictions shown in Fig. 7 as the PDO-like pattern. The PDO-like index predicted in the forecast system is defined as the normalized principal component time series corresponding to the EOF mode shown in Fig. 7. As shown in Fig. 8a, most models reproduce the positive phase of the PDO in the 1980 and 1990 s, and the negative phase of the PDO in the 1960 and 2000 s. The PDO-like index of some models, including CanESM5, EC-Earth3-i1, EC-Earth3-i2, FGOALS-f3-L, HadGEM3-GC31-MM and MIROC6, have

correlation coefficients above 0.45 with the observation. However, they do not exceed the 90% confidence level due to the small effective sample size when considering autocorrelation. FGOALS-f3-L and MIROC6 are significant at the 85% level, while CanESM5, EC-Earth3-i1, EC-Earth3-i1, MPI-ESM1-2-HR and the MME are significant at the 75% level. The ACC of these forecast systems could be improved if the decadal predictions have more hindcast start dates in the future.

The prediction skill of the AMV is much higher than the PDO predicted by these models and their ensemble. All forecast systems except for CanESM5, EC-Earth3-i1, EC-Earth3-i2 and NorCPM1-i1 can predict the time evolution of the AMV time series well with correlation coefficients above the 90% confidence level (Fig. 8b). CESM1-1-CAM5 has the highest correlation coefficient (0.82) among

Fig. 6 ACC of detrended winter SST hindcast by the MME (a) and each model (b–k) for the forecast years 2–5 with respect to the observation. The dotted identifies the region where the ACC is above the 90% confidence level



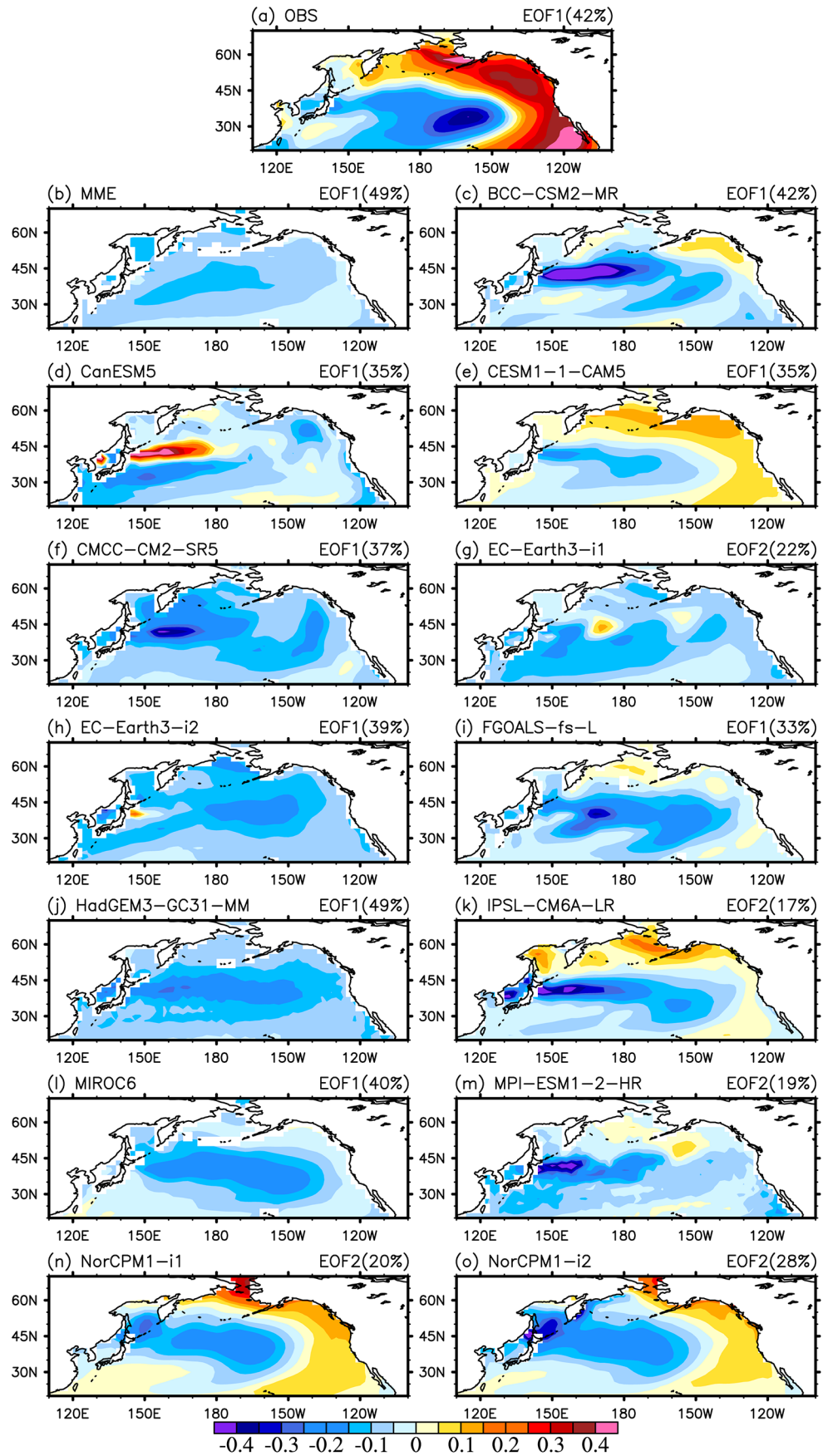
the forecast systems, which may be attributed to its largest ensemble size (40).

3.3.3 Connections of PDO and AMV with NEAP

Normalized time series of detrended NEAP and AMV, and PDO-like indices are shown in Fig. 9. In the observations, the NEAP shows an in-phase relationship with the AMV and an out-of-phase relationship with the PDO,

with the correlation coefficients of 0.80 and -0.41 , respectively. The MME reproduces the relationships of AMV and PDO-like variability with the NEAP, but it underestimates the positive correlation between NEAP and AMV (0.55), and overestimates the negative correlation between the NEAP and PDO-like index (-0.74). Seven forecast systems (BCC-CSM2-MR, CMCC-CM2-SR5, EC-Earth3-i1, EC-Earth3-i2, FGOALS-f3-L, NorCPM1-i1, NorCPM1-i2) can reproduce the significant positive correlation between the

Fig. 7 The EOF mode of detrended SST anomalies over the North Pacific in the observation (a), and decadal hindcasts by the MME (b) and individual forecast systems (c–j) for the forecast years 2–5. The label on the right corner of each figure indicates which EOF mode is used and the percentage of variance explained



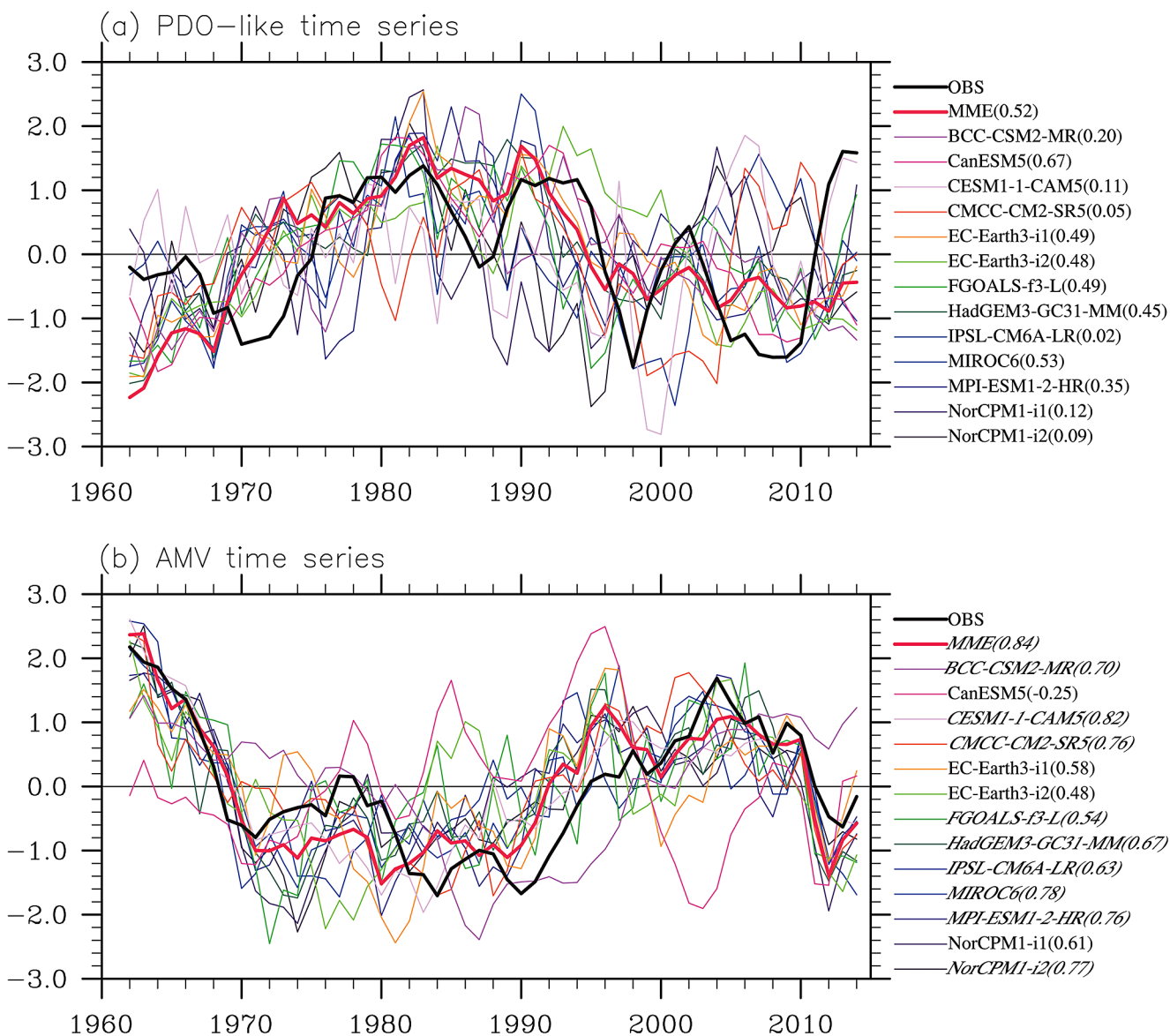


Fig. 8 Time series of PDO-like index (a), and AMV (b) in the observation, the MME and each forecast system for forecast years 2–5. The number in parentheses indicates the correlation coefficient between

the decadal hindcast and the observation. Italic indicates correlation exceeding the 90% confidence level

NEAP and the AMV, but with relatively lower correlation coefficients than the observation. Among the 7 models, 4 models (BCC-CSM2-MR, CMCC-CM2-SR5, FGOALS-fs-L, NorCPM1-i2) have skillful prediction in the AMV (Fig. 8), corresponding to their skillful prediction of the NEAP.

Nine models reproduce the significant negative correlation between the NEAP and the PDO-like index, including BCC-CSM2-MR, CanESM5, CMCC-CM2-SR5, EC-Earth3-i1, EC-Earth3-i2, HadGEM3-GC31-MM, IPSL-CM6A-LR, MIROC6, and MPI-ESM1-2-HR. This implies that the decadal variability in the North Pacific play an important role in the precipitation over Northeast Asia in

the decadal predictions of these forecast systems. Given the relatively low skill of the PDO predicted by most forecast systems, further improvement in this respect is needed in the future. The close relationships between the AMV, PDO and the NEAP in the observations and decadal predictions explains why the NEAP time series exhibits obvious decadal variations.

A previous study also reported the North Atlantic-Eurasian teleconnection in summer (Li and Ruan 2018). Skillful decadal prediction was found in summer surface air temperature over Northeast Asia, which was attributed to the circumglobal teleconnection pattern associated with the AMV (Monerie et al. 2018). However, there is almost no

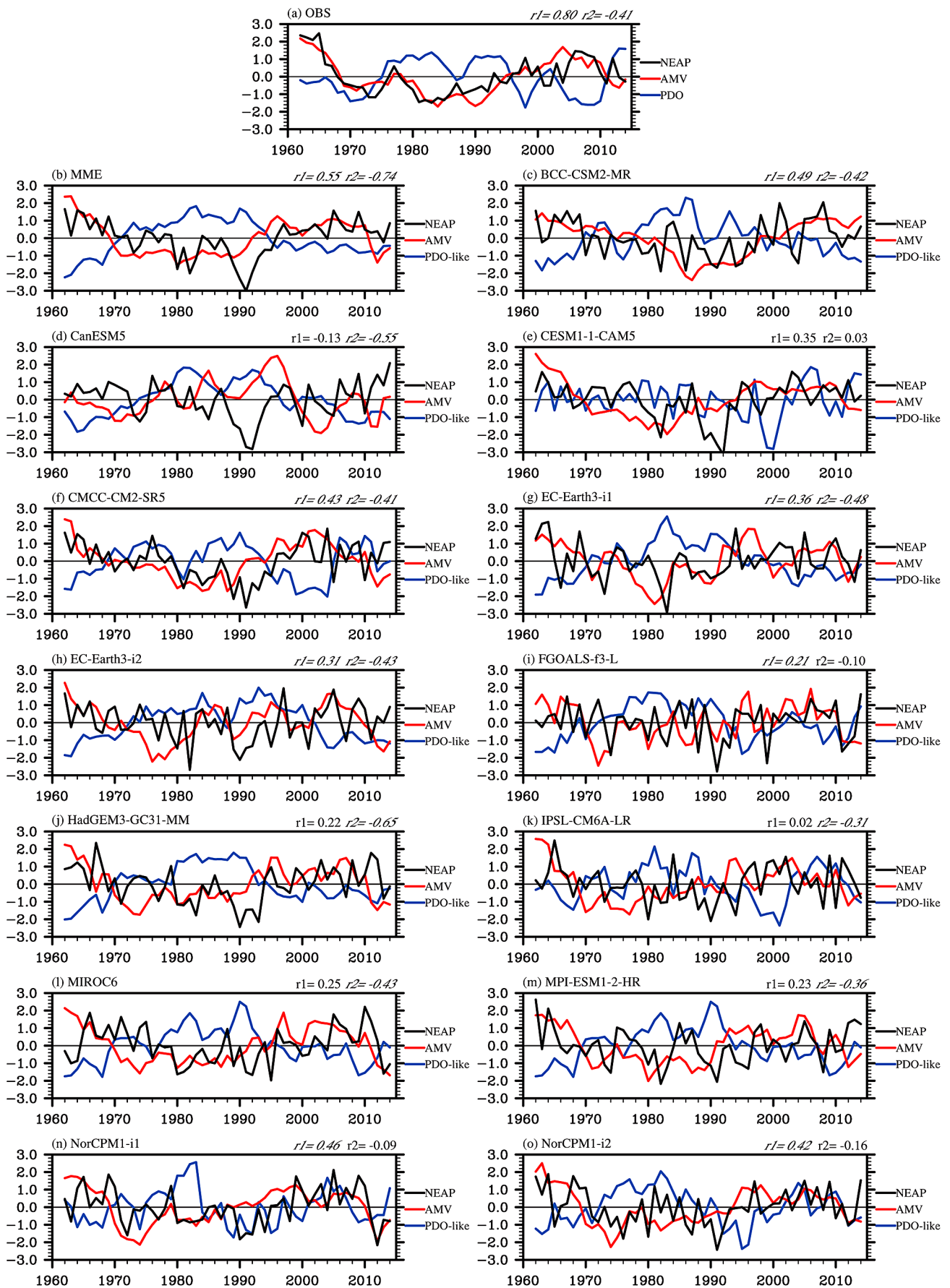


Fig. 9 Normalized NEAP, AMV and PDO-like time series during 1962–2014 in the observation (a), and decadal hindcasts by the MME and individual forecast systems (c–j) for the forecast years 2–5. In each figure, r_1 denotes the correlation coefficient between NEAP and AMV, and r_2 denotes the correlation coefficient between NEAP and PDO-like index. *Italic* indicates correlation exceeding the 90% confidence level

significant ACC in summer precipitation over Northeast Asia in the decadal predictions of all forecast systems (figure not shown).

3.4 ACC of NEAP within one decade

The ACC of the detrended NEAP in the decadal hindcasts along the forecast time for the 4-year averages is further examined and shown in Fig. 10. Ten forecast systems have significant ACC of NEAP for forecast years 3–6. Seven forecast systems have significant ACC of NEAP for forecast years 4–7 and 5–8. Four forecast systems including CMCC-CM2-SR5, MIROC6, MPI-ESM1-2-HR and NorCPM1-i2 have significant ACC of NEAP for forecast years 6–9. Only two forecast systems (CMCC-CM2-SR5, NorCPM1-i2) have significant ACC for forecast years 7–10. The ACCs of the MME are significant within 5–8 forecast years.

As can be seen in Fig. 10, the MME has higher ACC than the individual models in within the 5–8 forecast years. Therefore, the MME is still a useful strategy to improve the decadal prediction skill of winter precipitation in Northeast Asia. This is consistent with the study by Delgado-Torres et al. (2022), who suggest that the MME is a reasonable choice

for not having to select the best system for each particular variable and forecast period.

4 Conclusions

Decadal hindcasts from 13 forecast systems participating in the CMIP6 DCPD were compared with their historical simulations to predict winter precipitation over Eurasia for forecast years 2–5. Northeast Asia emerges as an area of improved decadal prediction skill, even after removing the linear trend. While none of the models have statistically significant correlation coefficients in historical simulations for the NEAP, all forecast systems except MIROC6 have significant correlation coefficients above the 90% confidence level. The atmospheric circulation associated with the detrended NEAP shows a cyclone over Northeast Asia and an anticyclone over the North Pacific, favoring abundant precipitation over Northeast Asia. Such circulation anomalies are reproduced by 10 forecast systems.

Six forecast systems (BCC-CSM2-MR, CESM1-1-CAM5, FGOALS-fs-L, IPSL-CM6A-LR, NorCPM1-i1 and NorCPM1-i2) are able to partially reproduce the opposite variation of the SST in the eastern Pacific coast and the subtropical Pacific as observed. The corresponding time series of PDO-like pattern in most forecast systems show decadal variability similar to the observation, but their ACCs are not above the 90% confidence level. The prediction skill of the AMV is much higher than that of the PDO predicted by these models, with 9 forecast systems having significant ACCs.

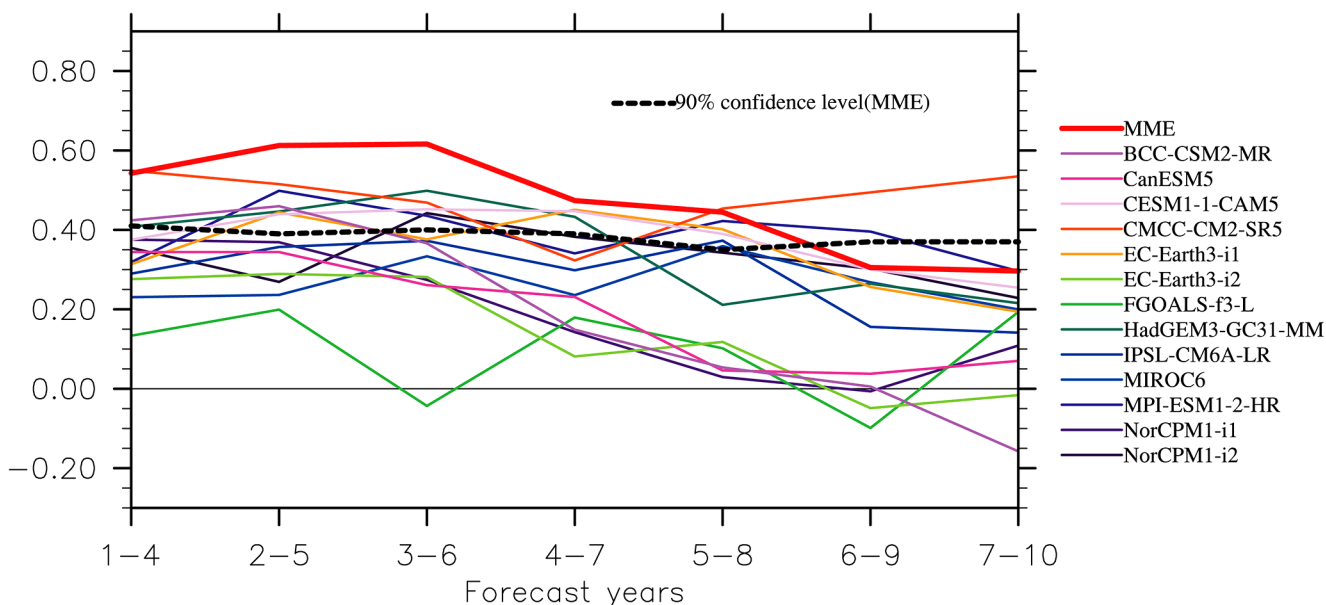


Fig. 10 ACC of winter detrended NEAP in the decadal hindcast by the MME and each forecast system along the forecast time for 4-year averages. The dashed line indicates the 90% confidence level of the MME

The detrended NEAP shows an in-phase relationship with the detrended AMV and an out-of-phase relationship with the PDO in the observations. The MME underestimates the positive correlation between NEAP and AMV, and overestimates the negative correlation between NEAP and PDO-like variability. Seven (nine) forecast systems are able to reproduce the significant positive correlation between the NEAP and the AMV (PDO-like variability). Due to the connection with the decadal variability in the North Atlantic or North Pacific, the time series of the NEAP also show apparent decadal variations in the decadal predictions. With higher ACCs than individual forecast systems, the MME is able to predict the NEAP for all 4-year averaged periods in the range of 1–8 years, demonstrating the advantages of using the multi-model ensemble. Since the majority of the forecast systems underestimate the correlation between the NEAP and the AMV, and have relatively low skill in predicting the PDO, further improvement of the forecast systems in these aspects helps to achieve higher skill in winter precipitation over Northeast Asia.

Funding This research is supported by the National Natural Science Foundation of China grant numbers 42230608, 42175165 and 42077427.

Data Availability CMIP6 outputs were downloaded from the ESGF node <https://esgf-node.llnl.gov/search/cmip6/>. The precipitation data of CRU version 4.05 is available at <http://badc.nerc.ac.uk/data/cru/>. ERSST version 5 data were obtained from <https://psl.noaa.gov/data/gridded/data.noaa.ersst.v5.html>. The atmospheric circulation data of JRA-55 were obtained from https://jra.kishou.go.jp/JRA-55/index_en.html.

Declarations

Competing interests The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Athanasiadis PJ, Yeager S, Kwon YO et al (2020) Decadal predictability of North Atlantic blocking and the NAO. *npj Clim Atmos Sci* 3:20. <https://doi.org/10.1038/s41612-020-0120-6>

- Bethke I, Wang Y, Counillon F et al (2021) NorCPM1 and its contribution to CMIP6 DCP. *Geosci Model Dev* 14:7073–7116. <https://doi.org/10.5194/gmd-14-7073-2021>
- Bilbao R, Wild S, Ortega P et al (2021) Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth. *Earth Syst Dynam* 12:173–196. <https://doi.org/10.5194/esd-12-173-2021>
- Boer GJ, Sospedra-Alfonso R (2019) Assessing the skill of the Pacific Decadal Oscillation (PDO) in a decadal prediction experiment. *Clim Dyn* 53(9–10):5763–5775. <https://doi.org/10.1007/s00382-019-04896-w>
- Boer GJ, Smith DM, Cassou C et al (2016) The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geosci. Model Dev* 9:3751–3777. <https://doi.org/10.5194/gmd-9-3751-2016>
- Borchert LF, Menary MB, Swingedouw D et al (2021) Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6. *Geophys Res Lett* 48. <https://doi.org/10.1029/2020GL091307>
- Boucher O, Servonnat J, Albright AL et al (2020) Presentation and evaluation of the IPSL-CM6A-LR climate model. *J Adv Model Earth Syst* 12. <https://doi.org/10.1029/2019MS002010>. e2019MS002010
- Bretherton CS, Widmann M, Dymnikov VP et al (1999) The effective number of spatial degrees of freedom of a time-varying field. *J Clim* 12:1990–2009. [https://doi.org/10.1175/1520-0442\(1999\)012%3C1990:TENOSD%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012%3C1990:TENOSD%3E2.0.CO;2)
- Corti S, Weisheimer A, Palmer TN et al (2012) Reliability of decadal predictions. *Geophys Res Lett* 39:L21712. <https://doi.org/10.1029/2012GL053354>
- Delgado-Torres C, Donat MG, Gonzalez-Reviriego N et al (2022) Multi-model forecast quality assessment of CMIP6 decadal predictions. *J Clim* 35:4363–4382. <https://doi.org/10.1175/JCLI-D-21-0811.1>
- Delgado-Torres C, Donat MG, Soret A et al (2023) Multi-annual predictions of the frequency and intensity of daily temperature and precipitation extremes. *Environ Res Lett* 18(3):034031. <https://doi.org/10.1088/1748-9326/acbbe1>
- Doblas-Reyes FJ, Andreu-Burillo I, Chikamoto Y et al (2013) Initialized near-term regional climate change prediction. *Nat Commun* 4:1715. <https://doi.org/10.1038/ncomms2704>
- Donat MG, Delgado-Torres C, De Luca P et al (2023) How credibly do CMIP6 simulations capture historical mean and extreme precipitation changes? *Geophys Res Lett* 50(14). <https://doi.org/10.1029/2022GL102466>. e2022GL102466
- Döscher R, Acosta M, Alessandri A et al (2022) The EC-Earth3 Earth system model for the coupled model intercomparison project 6. *Geosci Model Dev* 15:2973–3020. <https://doi.org/10.5194/gmd-15-2973-2022>
- Dunstone N, Smith D, Yeager S et al (2020) Skilful interannual climate prediction from two large initialised model ensembles. *Environ Res Lett* 15:094083. <https://doi.org/10.1088/1748-9326/ab9f7d>
- Eyring V, Bony S, Meehl GA et al (2016) Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci Model Dev* 9:1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Gaetani M, Mohino E (2013) Decadal prediction of the Sahelian precipitation in CMIP5 simulations. *J Clim* 26:7708–7719. <https://doi.org/10.1175/JCLI-D-12-00635.1>
- García-Serrano J, Doblas-Reyes FJ (2012) On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Clim Dyn* 39:2025–2040. <https://doi.org/10.1007/s00382-012-1413-1>
- Goddard L, Kumar A, Solomon A et al (2012) A verification framework for interannual-to-decadal predictions experiments. *Clim Dyn* 40:245–272. <https://doi.org/10.1007/s00382-012-1481-2>

- Harada Y, Kamahori H, Kobayashi C et al (2016) The JRA-55 reanalysis: representation of atmospheric circulation and climate variability. *J Meteor Soc Japan* 94:269–302. <https://doi.org/10.2151/jmsj.2016-015>
- Harris I, Osborn TJ, Jones P et al (2020) Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci Data* 7:109. <https://doi.org/10.1038/s41597-020-0453-3>
- He B, Bao Q, Wang X et al (2019) CAS FGOALS-f3-L model datasets for CMIP6 historical atmospheric model intercomparison project simulation. *Adv Atmos Sci* 36:771–778. <https://doi.org/10.1007/s00376-019-9027-8>
- Hu S, Zhou TJ (2021) Skillful prediction of summer rainfall in the Tibetan Plateau on multiyear time scales. *Sci Adv* 7:eabf9395. <https://www.science.org/doi/https://doi.org/10.1126/sciadv.abf9395>
- Hu S, Wu B, Wang Y et al (2023) CAS FGOALS-f3-L model datasets for CMIP6 DCPP experiment. *Adv Atmos Sci*. <https://doi.org/10.1007/s00376-023-2122-x>
- Huang B, Thorne PW, Banzon VF et al (2017) Extended reconstructed sea surface temperature, Version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J Clim* 30:8179–8205. <https://doi.org/10.1175/JCLI-D-16-0836.1>
- Huang D, Dai A, Yang B, Yan P, Zhu J, Zhang Y (2019) Contributions of different combinations of the IPO and AMV to recent changes in winter East Asian jets. *J Clim* 32:1607–1626. <https://doi.org/10.1175/JCLI-D-18-0218.1>
- Kataoka T, Tatebe H, Koyama H et al (2020) Seasonal to decadal predictions with MIROC6: description and basic evaluation. *J Adv Model Earth Syst* 12. <https://doi.org/10.1029/2019MS002035>. e2019MS002035
- Kim HM, Webster PJ, Curry JA (2012) Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys Res Lett* 39:L10701. <https://doi.org/10.1029/2012GL051644>
- Kim H, Ham Y, Scaife AA (2014) Improvement of initialized decadal predictions over the North Pacific Ocean by systematic anomaly pattern correction. *J Climate* 27:5148–5162. <https://doi.org/10.1175/JCLI-D-13-00519.1>
- Kobayashi S, Ota Y, Harada Y et al (2015) The JRA-55 reanalysis: general specifications and basic characteristics. *J Meteor Soc Jpn* 93:5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Li JP, Ruan CQ (2018) The North Atlantic–Eurasian teleconnection in summer and its effects on Eurasian climates. *Environ Res Lett* 13:024007. <https://doi.org/10.1088/1748-9326/aa9d33>
- Meehl GA, Goddard L, Boer G et al (2014) Decadal climate prediction: an update from the trenches. *Bull Amer Meteorol Soc* 95:243–267. <https://doi.org/10.1175/BAMS-D-12-00241.1>
- Monerie P-A, Robson J, Dong B, Dunstone N et al (2018) A role of the Atlantic Ocean in predicting summer surface air temperature over North East Asia? *Clim Dyn* 51:473–491. <https://doi.org/10.1007/s00382-017-3935-z>
- Müller WA, Jungclaus JH, Mauritsen T et al (2018) A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *J Adv Model Earth Syst* 10:1384–1413. <https://doi.org/10.1029/2017MS001217>
- Nicoli D, Bellucci A, Ruggieri P et al (2023) The Euro-Mediterranean Center on Climate Change (CMCC) decadal prediction system. *Geosci Model Dev* 16:179–197. <https://doi.org/10.5194/gmd-16-179-2023>
- Scaife AA, Smith D (2018) A signal-to-noise paradox in climate science. *npj Clim Atmos Sci* 1:28. <https://doi.org/10.1038/s41612-018-0038-4>
- Sheen KL, Smith DM, Dunstone NJ et al (2017) Skillful prediction of Sahel summer rainfall on inter-annual and multi-year timescales. *Nat Commun* 8:14966. <https://doi.org/10.1038/ncomms14966>
- Si D, Ding Y (2016) Oceanic forcings of the interdecadal variability in east Asian summer rainfall. *J Clim* 29:7633–7649. <https://doi.org/10.1175/JCLI-D-15-0792.1>
- Simpson IR, Yeager SG, McKinnon KA et al (2019) Decadal predictability of late winter precipitation in western Europe through an ocean–jet stream connection. *Nat Geosci* 12:613–619. <https://doi.org/10.1038/s41561-019-0391-x>
- Smith DM, Cusack S, Colman AW et al (2007) Improved surface temperature prediction for the coming decade from a global climate model. *Science* 317(5839):796–799. <https://www.science.org/doi/https://doi.org/10.1126/science.1139540>
- Smith DM, Eade R, Scaife AA et al (2019) Robust skill of decadal climate predictions. *npj Clim Atmos Sci* 2:13. <https://doi.org/10.1038/s41612-019-0071-y>
- Smith DM, Scaife AA, Eade R et al (2020) North Atlantic climate far more predictable than models imply. *Nature* 583:796–800. <https://doi.org/10.1038/s41586-020-2525-0>
- Sospedra-Alfonso R, Boer GJ (2020) Assessing the impact of initialization on decadal prediction skill. *Geophys Res Lett* 47. <https://doi.org/10.1029/2019GL086361>. e2019GL086361
- Sun C, Li JP, Zhao S (2015) Remote influence of Atlantic multidecadal variability on Siberian warm season precipitation. *Sci Rep* 5:16853. <https://doi.org/10.1038/srep16853>
- Swart NC, Cole JNS, Kharin VV et al (2019) The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci Model Dev* 12:4823–4873. <https://doi.org/10.5194/gmd-12-4823-2019>
- Tatebe H, Ogura T, Nitta T et al (2019) Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geosci Model Dev* 12: 2727–2765, <https://doi.org/10.5194/gmd-12-2727-2019>, 2019
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Am Meteorol Soc* 93:485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Williams KD, Copsey D, Blockley EW et al (2017) The Met office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *J Adv Model Earth Syst* 10:357–380. <https://doi.org/10.1002/2017MS001115>
- Wu B, Zhou TJ (2012) Prediction of decadal variability of sea surface temperature by a coupled global climate model FGOALS_gl developed in LASG/IAP. *Chin Sci Bull* 57:2453–2459. <https://doi.org/10.1007/s11434-012-5134-y>
- Wu T, Lu Y, Fang Y et al (2019) The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6. *Geosci Model Dev* 12:1573–1600. <https://doi.org/10.5194/gmd-12-1573-2019>
- Xin X, Gao F, Wei M et al (2018) Decadal prediction skill of BCC-CSM1.1 climate model in East Asia. *Int J Climatol* 38:584–592. <https://doi.org/10.1002/joc.5195>
- Yeager SG, Robson JI (2017) Recent progress in understanding and predicting Atlantic decadal climate variability. *Curr Clim Change Rep* 3:112–127. <https://doi.org/10.1007/s40641-017-0064-z>
- Yeager SG, Danabasoglu G, Rosenbloom NA (2018) Predicting near-term changes in the earth system: a large ensemble of initialized decadal prediction simulations using the community Earth system model. *Bull Am Meteorol Soc* 99:1867–1886. <https://doi.org/10.1175/BAMS-D-17-0098.1>