JGR Atmospheres

RESEARCH ARTICLE
10.1029/2024JD042580

Toward Dynamical Annual to Decadal Climate Prediction Using the IAP-CAS Model

Yao Tang[1,2], Qing Bao[1], Xiaofei Wu[3], Tao Zhu[1], Bian He[1], Yimin Liu[1], Guoxiong Wu[1], Siyuan Zhou[1], Yangke Liu[1,2], and Ankang Qu[4,5]

[1]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [2]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China, [3]School of Atmospheric Sciences/Plateau Atmosphere and Environment Key Laboratory of Sichuan Province, Chengdu University of Information Technology, Chengdu, China, [4]Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [5]University of Chinese Academy of Sciences, Beijing, China

Abstract Annual to decadal (A2D) climate prediction provides key insights for public policy and individual decision-making over the next 1–10 years, but most current dynamical models exhibit limited skill at the A2D scale. To address this challenge, the IAP-CAS A2D dynamical ensemble climate prediction system has been developed by expanding the existing operational sub-seasonal to seasonal (S2S) prediction system approved by the WMO/WWRP S2S panel. Using a full-field atmosphere-ocean initialization experiment which covers the period from 1981 to 2015, several key findings are revealed: First, the model demonstrates significant positive skill for regional surface temperature predictions globally, except for the North Atlantic, likely due to the initial shock. Despite this, the model effectively captures the global mean surface temperature warming trend. Second, the model exhibits relatively high predictability for the Pacific Decadal Oscillation (PDO), with correlation skill up to 3 years, comparable to the sixth Coupled Model Intercomparison Project Decadal Climate Prediction Project multi-model ensemble mean. The spread-error ratios close to 1 in the PDO predictions indicate high reliability. Additionally, the model shows significant skill in predicting the El Niño-Southern Oscillation (ENSO) for up to 1 year, comparable to leading seasonal dynamical prediction models. Further analysis reveals an established teleconnection between ENSO and the North Pacific atmosphere in the IAP-CAS model, likely underpinning the PDO predictive skill at forecast year 1. This study also assesses the effect of initialization by comparing initialized hindcast data with uninitialized historical simulations.

Plain Language Summary Predictive climate information for the next 1–10 years is increasingly important for policy and individual planning, especially due to accelerated climate change. However, predictions on this timescale are still developing and face more uncertainties compared to shorter-term forecasts. We developed a climate model called the IAP-CAS annual to decadal model to enhance these predictions by building on our existing sub-seasonal to seasonal forecasting system. This model produces a data set of 35 predictions, each covering 129 months. Its core consists of equations that describe the movement of Earth's fluids and runs on supercomputers. To generate the data set, we input reconstructed past weather conditions and climate change factors. Our evaluation shows the model effectively predicts global surface temperatures and the warming trend. It can also predict the Pacific Decadal Oscillation (PDO) up to 3 years in advance and the El Niño-Southern Oscillation (ENSO) up to 1 year ahead. Additionally, the model realistically reproduces atmosphere-ocean interactions during El Niño and La Niña events, likely underpinning its skill for PDO in the first forecast year.

Key Points:

- An annual to decadal climate prediction model is developed. A detailed description of the unique initialization and advanced skill is given
- The model is highly skilled at predicting global surface temperature, and the Pacific Decadal Oscillation (PDO) for up to 3 years
- The model predicts ENSO events up to 1 year and shows a physical link to the North Pacific, possibly enhancing PDO prediction in year one

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Q. Bao,
baoqing@mail.iap.ac.cn

Citation:

Tang, Y., Bao, Q., Wu, X., Zhu, T., He, B., Liu, Y., et al. (2025). Toward dynamical annual to decadal climate prediction using the IAP-CAS model. *Journal of Geophysical Research: Atmospheres*, *130*, e2024JD042580. https://doi.org/10.1029/2024JD042580

Received 27 SEP 2024
Accepted 21 FEB 2025

Author Contributions:

Conceptualization: Yao Tang, Qing Bao
Data curation: Xiaofei Wu
Formal analysis: Yao Tang
Funding acquisition: Qing Bao
Investigation: Yao Tang, Qing Bao, Xiaofei Wu, Tao Zhu
Methodology: Yao Tang, Qing Bao, Xiaofei Wu
Project administration: Qing Bao
Resources: Qing Bao
Software: Yao Tang
Supervision: Qing Bao
Visualization: Yao Tang
Writing – original draft: Yao Tang
Writing – review & editing: Yao Tang, Qing Bao, Xiaofei Wu, Tao Zhu, Bian He, Yimin Liu, Guoxiong Wu, Siyuan Zhou, Yangke Liu, Ankang Qu

1. Introduction

Annual to decadal (A2D) predictions, which focus on providing predictive information over the next 1–10 years, have garnered increased attention in recent years. These predictions are crucial for enhancing our understanding of the climate system's response to external forcing and natural variability (Boer et al., 2016). In the context of accelerating global warming and more frequent extreme weather events (Gulev et al., 2021; Seneviratne et al., 2021), reliance on historical information alone is insufficient to support future planning (Findell et al., 2023). Skillful and reliable A2D predictions are increasingly important for risk assessment, uncertainty management, and policy development (Kushnir et al., 2019; Malakar et al., 2024).

Given the importance of A2D predictions, several international research initiatives have intensified their focus on this area. The earliest organized effort originated from the fifth phase of the Coupled Model Intercomparison Project (CMIP5), which mobilized numerous climate modeling groups to conduct initialized experiments and explore near-term climate predictability (Taylor et al., 2012). This was followed by the Decadal Climate Prediction Project (DCPP) as part of CMIP6, which established a more structured and demanding framework, with hindcasts starting annually rather than every 5 years as in CMIP5 (Boer et al., 2016). More recently, the World Climate Research Programme's (WCRP) Lighthouse Activity on Explaining and Predicting Earth System Change (EPESC) has aimed to develop an integrated capability for understanding, explaining, and predicting Earth system changes on A2D timescales (Findell et al., 2023).

Previous research on A2D predictions mainly focused on two aspects: sources of predictability and prediction skill. Climate variations in A2D can arise from internal variations of the climate system and external forcing including anthropogenic and volcanic aerosols, greenhouse gases, and changes in solar radiation (Boer et al., 2016). The oceans play an important role in generating large-scale and long-duration internal variability, which constitutes the main source of predictability on the A2D timescale, owing to their thermal and dynamical inertia, as well as atmosphere-sea surface temperature (SST) interactions and teleconnections (Cassou et al., 2018). The El Niño-Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) are the two dominant SST variabilities in the Pacific Ocean (Barnston et al., 2019; Choi & Son, 2022; Tippett et al., 2012; Weisheimer et al., 2022). The Atlantic Multidecadal Oscillation (AMO) (Trenberth & Shea, 2006) is the primary SST variability in the North Atlantic and plays a significant role in modulating climate on A2D scales. These internal variabilities directly impact key meteorological variables such as temperature and precipitation, which provide crucial information on future temporal and spatial scales. Since 2020, the World Meteorological Organization's (WMO) Global Annual to Decadal Climate Update has provided real-time predictions for the next five years annually, covering surface temperature, precipitation, sea level pressure, and major climate variabilities on a global scale (Hermanson et al., 2022). However, current A2D prediction capabilities remain limited, and significant challenges persist (Kushnir et al., 2019; Meehl et al., 2021). For example, the North Pacific has been identified as the ocean with the lowest prediction skill in several A2D prediction systems (Doblas-Reyes et al., 2013; Guemas et al., 2012), exhibiting its inherently low predictability. Correspondingly, the limited predictive skill of PDO is also commonly observed in A2D models, with the improvement of initialization being confined to the first few years (Choi & Son, 2022; Z. Liu & Di Lorenzo, 2018). Additionally, a recent study suggests that accurately predicting A2D climate variability, such as the PDO, is becoming more difficult in the context of global warming (S. Li et al., 2020).

To address the limitations in A2D prediction, an A2D model has been developed by the Institute of Atmospheric Physics, Chinese Academy of Sciences (IAP-CAS) (hereafter referred to as the IAP-CAS model). The IAP-CAS model is a real-time operational ensemble prediction system that has recently conducted a series of hindcast experiments to evaluate its skill on an A2D timescale. Unlike other models, the IAP-CAS model, which is based on a mature subseasonal-to-seasonal (S2S) operational prediction system coordinated by WMO WWRP S2S panel (https://confluence.ecmwf.int/display/S2S/Models). The IAP-CAS model can be carried out consistently in the modern climate and commenced annually to ensure timeliness. This study focuses on evaluating the performance of the IAP-CAS model in predicting key surface variables and simulating major climate variabilities on the A2D timescale. By assessing the model's strengths and limitations, we aim to provide insights for improving A2D predictions.

The article is structured as follows: Section 2 describes the validation data sets, the composition of the IAP-CAS model, the experimental design, and the methodology for skill evaluation. Section 3 presents the skill of the IAP-CAS model in predicting key meteorological variables and internal climate variability, the model's biases, as well as the simulated air-sea interactions on the results. Section 4 concludes the study.

2. Data, Model and Methodology

2.1. Verification Data Sets

In this study, we extensively utilize observational data commonly employed in A2D prediction assessments to evaluate various aspects of predictive skill. To verify surface temperature predictions, we use SST data from the HadISST 1.1 data set (Rayner et al., 2003) and land surface air temperature data from the HadCRUT.5.0.1.0 data set (Morice et al., 2021). A merged data set from the IAP-CAS model, which includes 2 m air temperature (T2m)

over land and SST over the ocean, is used to align with the observational data. For precipitation, we utilize the GPCC Full Data Monthly Product Version 2022 (U. Schneider et al., 2022). Observed sea level pressure (SLP) data are obtained from the NCEP-NCAR Reanalysis 1 data set, provided by NOAA PSL, Boulder, Colorado, USA, and available at https://psl.noaa.gov (Kalnay et al., 1996). Additionally, data from nine models in the CMIP6 DCPP hindcast data sets are accessed via the Earth System Grid Federation (ESGF) (https://aims2.llnl.gov/search/cmip6/) for evaluation purposes to demonstrate predictive skill (Table S1 in Supporting Information S1). All model data are interpolated to the corresponding observation grids, with missing grid points in the observations also designated as missing in the models.

2.2. The Model

The IAP-CAS A2D model is an extension of the existing S2S operational prediction system (Y. Liu et al., 2024) endorsed by the WMO/WWRP S2S panel. The climate model utilized by this system is the Flexible Global Ocean-Atmosphere-Land System, Finite Volume version 2 (FGOALS-f2) (Bao et al., 2019). FGOALS-f2 shares the same atmospheric component as FGOALS-f3, which has participated in some intercomparison projects under CMIP6 including the Global Monsoons Model Intercomparison Project (GMMIP) Tier-1 and Tier-3 (B. He et al., 2020), the Atmospheric Model Intercomparison Project (AMIP) (B. He et al., 2019) and the High-Resolution Model Intercomparison Project (HighResMIP) Tier 2 (An et al., 2022).

FGOALS-f2 is a fully coupled climate system that integrates four components: atmosphere, ocean, land, and sea ice, all connected through a coupler, as presented in Figure 1a. Each component features a horizontal resolution of approximately 1°. The atmospheric component, the Finite-volume Atmospheric Model in IAP/LASG (FAMIL) version 2 (Zhou et al., 2015), operates on a C96 grid with 32 hybrid sigma-pressure levels, extending vertically up to 1 hPa. FAMIL2 utilizes the finite-volume dynamical core (FV3) developed by the Geophysical Fluid Dynamics Laboratory (GFDL) (Lin, 2004; Putman & Lin, 2007), on a cubed-sphere grid to address numerical challenges at the poles. FV3's Lagrangian vertical coordinate enhances computational efficiency and accuracy (https://www.gfdl.noaa.gov/fv3/). The oceanic component uses the Parallel Ocean Program, version 2 (POP2) (Kerbyson & Jones, 2005), which includes 60 vertical layers. The land component is represented by the Community Land Model, version 4 (CLM4.0) (Lawrence et al., 2011; Oleson et al., 2010). The sea ice component utilizes the Los Alamos Sea Ice Model, version 4.0 (CICE4) (Hunke & Lipscomb, 2010). Data exchange, interpolation, and parallel coupling among these components are managed by the coupler version 7 in the Community Earth System Model (CESM) (Craig et al., 2012). The IAP-CAS model is configured on supercomputing platforms to achieve high-performance computing capabilities, both x86 and ARM architectures are supported.

The IAP-CAS model spin-up began on 1 January 1976, and was integrated up to the forecast start time. Throughout this period, the model was constrained by initial conditions from reanalysis data, a process also referred as reanalysis nudging in the IAP-CAS framework. This nudging process is illustrated in Figure 1b. Oceanic and atmospheric variables are initialized using a full-field initialization strategy, where raw values from reanalysis data sets are directly applied as initial conditions. For the atmospheric component, initialization involves temperature, zonal and meridional winds on isobaric surfaces, surface pressure and specific humidity from the CRA40 reanalysis (Z. Liu et al., 2023). The oceanic component initializes SST data from the OISST data set (Reynolds et al., 2007). The Newtonian nudging technique (Jeuken et al., 1996) with time-varying weighting factor was introduced to reduce assimilation errors. Atmospheric variables were nudged with a 6-hr relaxation timescale, whereas oceanic variables were nudged using a 1-day relaxation timescale. Detailed documentation of the time-varying weighting factor is available in Zeng et al. (2023) and Y. Liu et al. (2024). The external forcing conditions incorporated in the IAP-CAS model are consistent with those used in CMIP6 historical simulations (Eyring et al., 2016).

To generate the IAP-CAS A2D hindcast (reforecast) data set, predictions are conducted annually from 1981 to 2015, with each integration lasting 129 months and involving 8 ensemble members. The forecast start date is March 20, aligned with the S2S operational prediction system, which is widely used in the national flood season outlook in China. Ensembles are generated using time-lagged perturbations (Figure 1c) as described by Hoffman and Kalnay (1983). This approach combines historical observations and dynamic models, ensuring robust performance in prediction accuracy and stability. The workflow of the IAP-CAS A2D model is summarized in Figure 1d. It includes nudging with reanalysis and external forcing data to maintain proximity to observations,

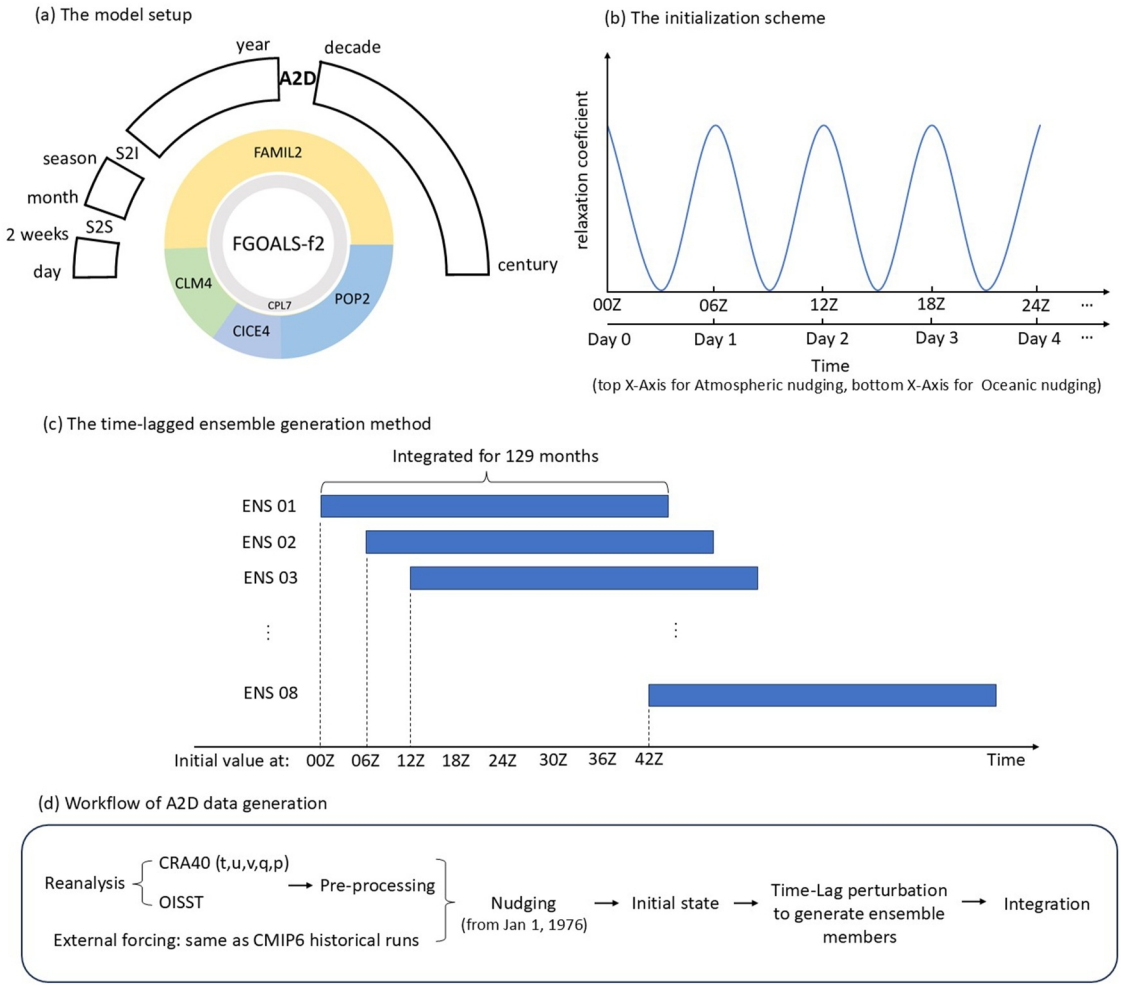

Figure 1. Schematic of the IAP-CAS A2D ensemble prediction system. (a) The setup of FGOALS-f2 model. (b) The relaxation coefficient over time in reanalysis nudging. (c) The time-lag perturbation is used to generate 8 ensemble members. (d) The workflow of A2D data generation.

initializing the model with input states at the forecast start time, generating ensembles using a time-lagged approach, and performing integrations for 129 months for each prediction.

2.3. Skill Evaluation Method

In full-field initialization, models unavoidably drift from observed initial states toward their preferred climate states throughout the prediction period. Although methods like trend-based and initial condition-based drift corrections can be applied, previous studies reveal their improvements are limited compared to mean drift correction (Choudhury et al., 2017). To remove the drift, model anomalies are calculated based on forecast-time-dependent climatologies to produce bias-corrected predictions (Boer et al., 2016). For a specific forecast time τ, the model climatology $\overline{H}_\tau$ is the average of all model predictions $H_{i\tau}$ at that forecast time (where i refers to the start date, corresponding to 35 predictions initialized from 1981 to 2015). For example, the model climatology of forecast year 1 is the average of the first year taken from 35 predictions, forming a continuous time series spanning from 1981 to 2015. Similarly, the model climatology of forecast year 2 is the average of all second-year predictions, spanning from 1982 to 2016. The model's anomaly H' is then constructed by subtracting the forecast-time-dependent climatology $\overline{H}_\tau$ from each of the model predictions $H_{i\tau}$. For observations, the climatology is calculated as the mean over the 1981–2010 period, and anomalies are derived by subtracting this climatology, with the time periods aligned to the model anomalies for consistency. To ensure analysis consistency, all variables and data sets used in this study are selected from data available prior to 2020, as the GPCC data set extends only up to 2020. Therefore, for the IAP-CAS model, all the 35 hindcast experiments initialized from 1981 to 2015 are

used for the first six forecast years, 34 experiments initialized from 1981 to 2014 are used for the forecast year 6, decreasing incrementally until only the 30 experiments initialized from 1981 to 2010 are used for forecast year 10. As the initialization starts in March, a full year for the IAP-CAS model is defined from April to March of the following year. For historical simulations, the same approach as observations is used: the climatology is based on 1981–2010, and anomalies are calculated with the time periods aligned to the hindcast anomalies. However, since the historical simulations only extend to 2014, the sample size is somewhat reduced.

To construct the CMIP6 DCPP multi-model ensemble (DCPP-MME), each ensemble member of every model is assigned the same weight. Some models provide more than 10 ensemble members, but only the first 10 are selected. All models are interpolated to a resolution of $1° \times 1°$. The DCPP hindcast experiments begin from 1960, but only those beginning from 1981 to 2014 are selected. The initialization dates of each model vary, including October, November, and January. To ensure consistency, the selection uniformly starts in January. Therefore, for DCPP-MME, a full year refers to the period from January to December, which is considered the forecast months 1–12 for each forecast year. Although the calendar months do not align with those of the IAP-CAS hindcasts, the two data sets are not directly compared in this study. Instead, their prediction skill is assessed separately by comparing each data set to observations over the corresponding time periods, ensuring that the evaluation reflects their respective initialization timings.

The climate indices are calculated as follows. The global mean surface temperature (GMST) is the area-weighted surface temperature anomaly, with T2m for land and SST for ocean. For ENSO, the Nino 3.4 index is computed from the area-weighted SST anomaly (SSTA) over 5°N–5°S, 170°–120°W. For the PDO index, the principal component of the first empirical orthogonal function (EOF1) of annual SSTA over 20°–70°N in the North Pacific is computed. Before EOF analysis, the global mean SSTA over 60°N–60°S is removed from each grid. The spatial pattern of EOF1 in the model may be distorted as the forecast year increases. To avoid the difficulty in recognizing the model's spatial pattern, the model's PDO index is estimated by projecting SSTA onto the EOF1 of observations (Choi & Son, 2022; Kataoka et al., 2020; Lee et al., 2019).

To evaluate prediction skill, the anomaly correlation coefficient (ACC) and mean squared skill score (MSSS) are used. These two indices are also used in the Global Annual to Decadal Climate Update released annually (e.g., World Meteorological Organization (WMO), 2024) and in many studies (e.g., Goddard et al., 2013; Nicolì et al., 2023; Yeager et al., 2023). Anomaly correlation coefficient assesses the prediction of the phase, while MSSS is sensitive to the amplitude of variability. For ensemble predictions, reliability is also a critical metric for assessing the quality. The perfect reliability indicates that the predicted probabilities align with the observed relative frequencies (Wilks, 2011). In our study, the spread-error ratio (SER) is used to evaluate forecast reliability (Ho et al., 2013). These metrics are calculated as follows:

$$ACC(H,O)_\tau = \frac{\sum_{i=1}^{N}(H_{i\tau}-\overline{H}_\tau)(O_{i\tau}-\overline{O}_\tau)}{\sqrt{\sum_{i=1}^{N}(H_{i\tau}-\overline{H}_\tau)^2}\sqrt{\sum_{i=1}^{N}(O_{i\tau}-\overline{O}_\tau)^2}} \quad (1)$$

$$MSSS(H,O)_\tau = 1 - \frac{\sum_{i=1}^{N}\left[(H_{i\tau}-\overline{H}_\tau)-(O_{i\tau}-\overline{O}_\tau)\right]^2}{\sum_{i=1}^{N}(O_{i\tau}-\overline{O}_\tau)^2} \quad (2)$$

where $H_{i\tau}$ ($O_{i\tau}$) is the hindcasts (observations) of forecast year τ at a certain start date i, and $\overline{H}_\tau$ ($\overline{O}_\tau$) is the average of N start dates, representing the climatology. Equations 1 and 2 are also applied for skill assessment of historical simulations and CMIP6-DCPP MME, where the hindcast data are replaced with historical simulations or CMIP6-DCPP MME respectively. It is important to note that the MSSS in this context focuses on evaluating skill relative to reality, with the reference data being the observations. When comparing predictive skill across different experiments (i.e., hindcasts vs. historical simulations or hindcasts vs. CMIP6-DCPP MME), the ACC difference and the MSSS(H,P), as shown in the formula below, are computed accordingly:

$$MSSS(H,P)_\tau = 1 - \frac{\sum_{i=1}^{N}\left[(H_{i\tau}-\overline{H}_\tau)-(O_{i\tau}-\overline{O}_\tau)\right]^2}{\sum_{i=1}^{N}\left[(P_{i\tau}-\overline{P}_\tau)-(O_{i\tau}-\overline{O}_\tau)\right]^2} \quad (3)$$

where $P_{i\tau}$ represents the historical simulations or CMIP6-DCPP MME of forecast year τ at a certain start date i, and $\overline{P}_{\tau}$ is the climatology derived from the historical simulations or CMIP6-DCPP MME.

$$SER_{\tau} = \sqrt{\frac{m+1}{m}} \frac{\sigma_{\tau}}{RMSE_{\tau}} \tag{4}$$

where m is the number of ensemble members, σ_{τ} is the time-averaged intraensemble standard deviation of forecast year τ, and $RMSE_{\tau}$ is the root mean squared error of the ensemble mean of forecast year τ. An SER less than 1 suggests that the ensemble is under-dispersive and the predictions are overconfident, while an SER greater than 1 indicates that the ensemble is over-dispersive and the predictions lack confidence.

Due to the limited sample size in decadal predictions, we employ the nonparametric block bootstrap approach for conducting statistical significance tests (Goddard et al., 2013). A block size of 5 years is chosen to account for autocorrelation. To estimate confidence intervals for the test statistic, we perform 1000 resamplings to construct the density function, setting the statistical significance level at 90%.

2.4. SST Variability in Uninitialized Experiments of the IAP-CAS Model

In addition to serving as a reference for initialized experiments, uninitialized historical simulations reflect the fundamental performance of the model. Before evaluating the hindcast skill of the IAP-CAS model, we first present its inherent capability to simulate key SST variability indices relevant to A2D predictions, specifically the PDO and ENSO, using uninitialized historical simulations. These simulations cover the period from 1850 to 2014 and include three ensemble members.

Figure S1 in Supporting Information S1 presents the PDO and ENSO simulations from the historical run of the IAP-CAS model. It can be seen that the spatial pattern of the PDO simulated by the IAP-CAS model is basically consistent with the observation, except that it is stronger in the northeast Pacific region, and the response of equatorial SST is weaker. The PDO amplitude of the model is close to the observations. The power spectrum analysis (Figures S1c and S1f in Supporting Information S1) shows that the PDO time series simulated by the model has significant spectral peaks at 3 and above 10 years, while the observations are mainly concentrated above 10 years. As for ENSO, the simulated Nino 3.4 time series by the model has obvious interannual oscillations except a long positive stage lasting in the later periods. The spectral peaks of simulated Nino 3.4 are concentrated around 2–5 years, which is consistent with observations.

3. Results

3.1. Predictive Skill of Global-Mean Surface Temperature

Figure 2 shows the predictive skill of the initialized hindcasts of the IAP-CAS model in capturing GMST. The GMST skill of the uninitialized historical simulations is shown in Figure S2 in Supporting Information S1 to illustrate the benefit of initialization. For forecast year 1, the hindcasts successfully capture the general rapid warming trend and the interannual variability, including the cooling due to the eruption of Mount Pinatubo in 1991. The historical simulations underestimate GMST in the early phase of year 1 and overestimate them in the later phase. As a result, despite achieving the same correlation coefficient (ACC = 0.88) as the hindcast experiments, their MSSS is considerably lower. For forecast years 1–5 and 6–10, the hindcasts capture the overall warming trend, achieving high skill scores with an ACC of 0.96 (MSSS of 0.92) and 0.93 (MSSS of 0.80), respectively. However, they tend to underestimate the magnitude of GMST warming during these periods, and the ensemble spread fails to adequately encompass the observations. The improved skill compared to year 1 can be attributed to multi-year averaging, which filters out interannual variability. Similarly, the historical simulations also show enhanced skill for years 1–5 and 6–10, except for the MSSS skill for years 6–10, which may be influenced by the shorter study period. However, the MSSS skill of historical simulations still differs markedly from that of the hindcasts, highlighting the significant advantage of initialization in enhancing the model's skill in GMST prediction.

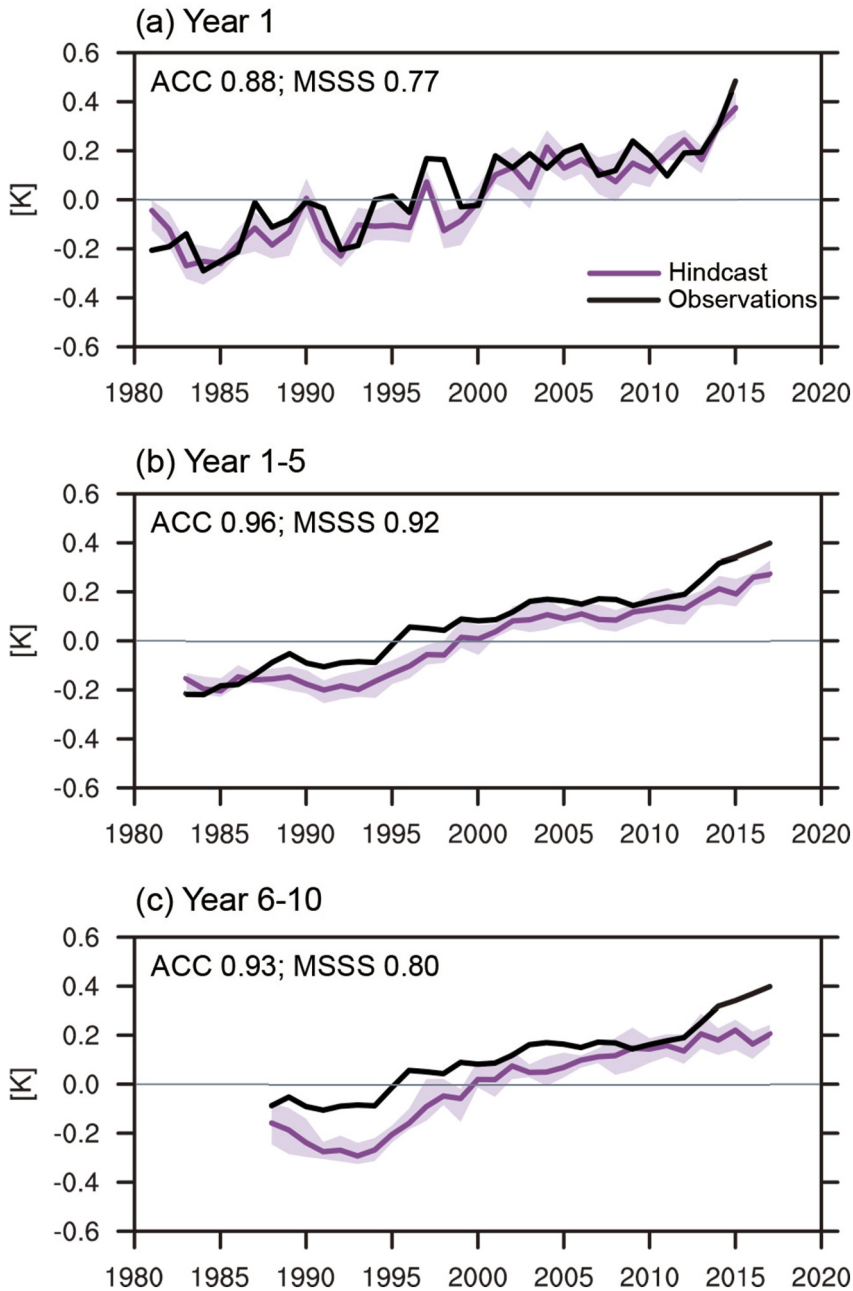

Figure 2. The skill of GMST predictions from the hindcasts of the IAP-CAS model for (a) forecast year 1, (b) years 1–5 and (c) years 6–10. The black line represents the observations, while the purple line represents the IAP-CAS model predictions. Shading indicates the range of maximum and minimum values of the ensemble members. For 5-year averages, values are plotted at the midpoint year. Note that a complete year is defined as the period from April to the following March. Numbers in the upper left corner represent the ACC(H,O) and MSSS(H,O) skill of the hindcasts relative to the observational time series.

3.2. Regional Predictions of Major Meteorological Variables

This section presents the regional predictions of major meteorological variables including surface temperature, precipitation and SLP on a global scale, providing information for diagnosing spatial skill. The ACC and MSSS are still used as skill metrics. Note that due to some missing values in the observation data sets, only non-missing grid points throughout the entire calculated period are retained to ensure an adequate sample size over time.

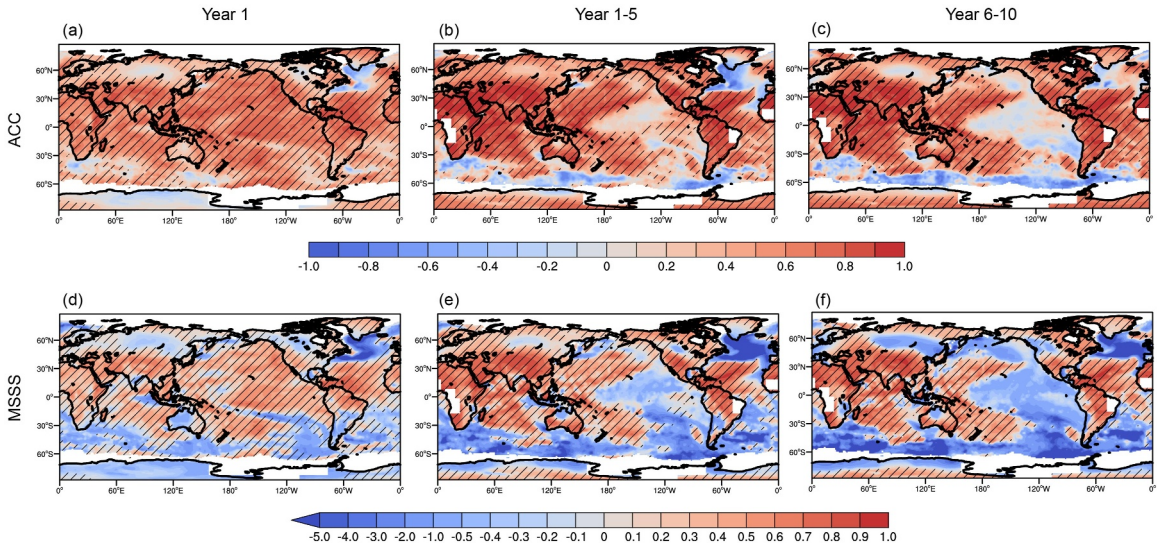

Figure 3. (a–c) The ACC(H,O) and (d–f) MSSS(H,O) skill maps of surface temperature. The first column displays skill scores for forecast year 1, the second column for years 1–5, and the third column for years 6–10. Shaded regions represent skill scores that are significant at the 90% confidence level. Note that the MSSS color scale is not symmetric around 0.

The hindcast skill of surface temperature predictions for forecast year 1 is significant across most parts of the globe (Figure 3a). In nearly all of the Indian Ocean, the Pacific Ocean, and most of the Atlantic Ocean, as well as the majority of land areas, the ACC is positive and passes the significance test. However, there are distinct areas with low skill: the North Atlantic region south of Greenland and west of Europe; Antarctica and the nearby ocean; and the high-latitude interior of Eurasia. For MSSS skill at forecast year 1 (Figure 3d), the Indian Ocean, most of the Pacific Ocean, the low latitudes of the Atlantic Ocean, and land areas including low and mid-latitude Eurasia, most parts of Africa, the central Americas, and Greenland perform well. The skill of the hindcasts in predicting surface temperature is compared with the historical simulations in Figure S3 in Supporting Information S1. In forecast year 1 (Figures S3a and S3d in Supporting Information S1), the hindcasts show a significant improvement over the historical simulations, particularly SST. Notable enhancements are observed in the Pacific, Indian Ocean, and Southern Ocean regions, indicating the positive impact of initialization. However, the large negative ACC difference and MSSS values in the North Atlantic indicate that initialization has caused substantial adverse effects on SST predictions in this region.

Regions with positive ACC skill at forecast year 1 of hindcasts are further enhanced in forecast years 1–5 (Figure 3b), except for parts of the central and eastern Pacific, and the Southern Ocean. Notably, the North Pacific, especially off the west coast of North America, shows high skill in forecast years 1–5. There is also significant ACC skill off the west coast of equatorial South America. Compared to forecast year 1, the low-skill area in the North Atlantic extends to the adjacent seas of Europe and Northwest Africa. However, the skill for surface temperature predictions in Antarctica is higher than at forecast year 1. The surface temperatures in Antarctica are characterized by large interannual fluctuations related to sea ice coverage in this region (Comiso, 2000), which are filtered out by multi-year averaging in the IAP-CAS model, thereby reducing prediction difficulty. The MSSS in regions with positive skill at forecast year 1 is also enhanced in forecast years 1–5 (Figure 3e), except for the tropical ENSO region. Compared to the historical simulations, the improvements in surface temperature prediction skill by the hindcasts during years 1–5 are relatively limited (Figures S3b and S3e in Supporting Information S1), appearing only sporadically in regions such as the northeastern and equatorial eastern Pacific, the Southern Ocean, and parts of the Indian Ocean. This suggests that the positive impacts of initialization are not consistently sustained over such long periods. Furthermore, positive skill areas are more prevalent over the ocean than over land, indicating that the ocean retains the beneficial effects of initialization for a longer duration. In the North Atlantic, the extent of negative skill expands compared to year 1, likely reflecting the lingering effects of the initial negative impacts from the first year of initialization. The negative SST skill in the central subpolar North Atlantic may propagate through air-sea interactions and internal oceanic processes, further degrading SST predictions over adjacent regions.

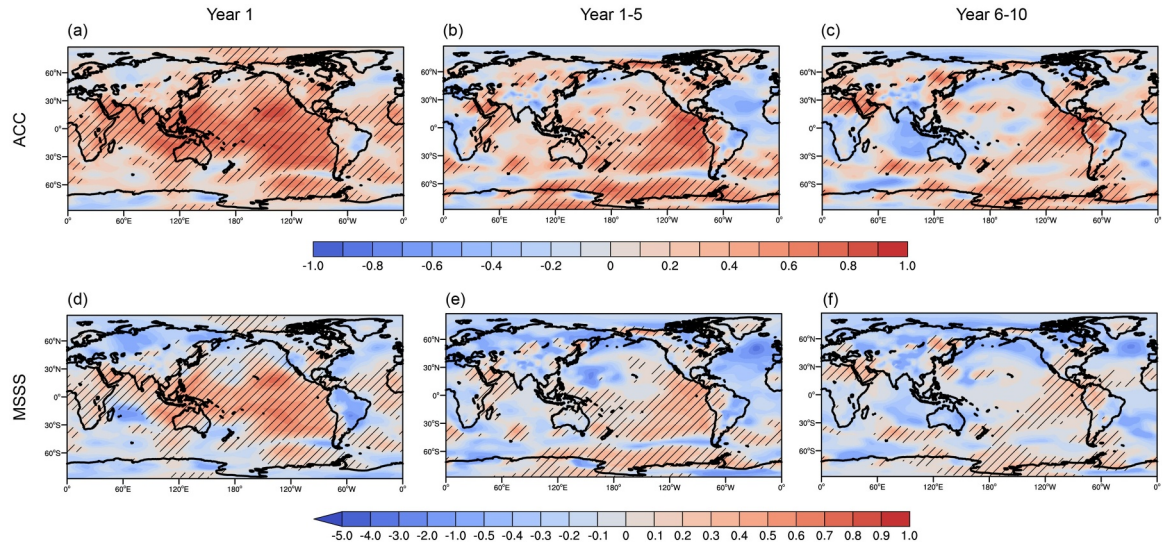

Figure 4. As Figure 3 but for annual mean sea-level pressure.

The pattern for forecast years 6–10 (Figures 3c and 3f) is generally similar to that for forecast years 1–5. However, the low skill area in the interior of Eurasia expands, while the low skill area in the North Atlantic decreases. The skill in the tropical ENSO region and the eastern North Pacific is almost non-existent in the last 5 years. The improvements by the hindcasts during years 6–10 (Figures S3c and S3f in Supporting Information S1) are also comparable to those observed in years 1–5, although the extent of regions with enhanced skill has diminished.

A trend analysis of the hindcasts (Figure S4 in Supporting Information S1) is provided as a supplement to the skill assessment. Although accurate trend prediction is not a sufficient condition for achieving high skill, most regions with high skill show relatively accurate trend predictions. However, the IAP-CAS model does not capture the observed trend of the North Atlantic SSTA, that is, the warming off the northeastern coast of North America and cooling in the southern Greenland sea, known as the North Atlantic warming hole (Keil et al., 2020; Rahmstorf et al., 2015). The SST trend in this region can influence the location of the North Atlantic jet, leading to the northward shift of the intertropical convergence zone (ITCZ) over the Indian Ocean, and impacting sensible weather in Europe (Gervais et al., 2020; Karnauskas et al., 2021). Instead, the IAP-CAS model seems to amplify the North Atlantic warming hole. Additionally, the IAP-CAS model does not capture the cooling pattern of temperatures in the interior of Eurasia as observed.

For SLP (Figure 4), the forecast year 1 has hindcast skill in most parts of the globe, especially over the oceans. The SLP skill over the mid-to low-latitude oceans is relatively better and more spatially coherent compared to other regions. Notably, the SLP skill in the tropical East Pacific remained skillful throughout the entire decade. The western North Pacific is also included in the high-skill region, where the tropical cyclones generated account for about one-third of the global total, significantly impacting the East Asian climate (Basconcillo & Moon, 2022; Wu et al., 2005). The predictive skill in this region is also evident in the IAP-CAS model's seasonal predictions, as demonstrated by skillful prediction of tropical cyclones in the western Pacific (J. Li et al., 2021). In the North Atlantic, the North Atlantic Oscillation (NAO) skill remains low even with atmospheric initialization (figure not shown), which may be related to the relatively low SST skill in this region (Figures 3a–3c). Previous study (Karnauskas et al., 2021) showed that the SST trends in the North Atlantic (as discussed in the previous trend analysis) can influence the phase shifts of NAO. Conversely, in the North Pacific at similar latitudes, the SLP shows skill at forecast year 1. The skill of SLP in forecast years 1–5 performs better than forecast years 6–10, suggesting the importance of initialization.

Prediction skill of precipitation of the IAP-CAS model is sporadic compared to SST and SLP (Figure 5). For forecast year 1, skillful regions include ENSO-related precipitation areas such as the Maritime Continent, Australia, and northern South America (Dai & Wigley, 2000), as well as parts of Asia and Africa, western North America, southern South America, and parts of Greenland. For forecast years 1–5 and 6–10, the significant skill is mainly located in Northeast Asia, which is consistent with other studies (e.g., Hu et al., 2023; Kataoka et al., 2020;

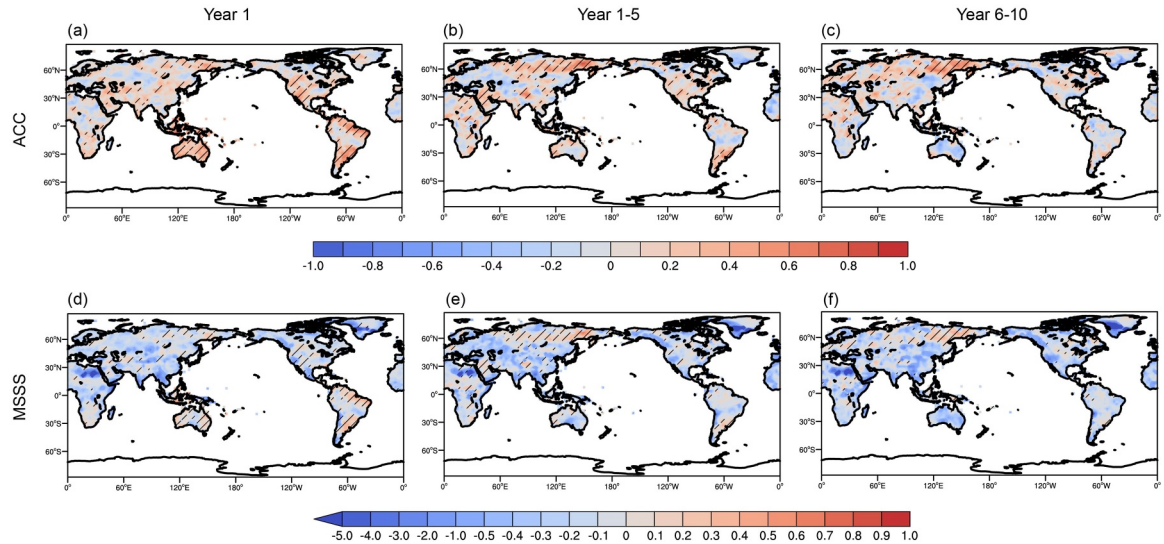

Figure 5. As Figure 3 but for annual mean precipitation.

Nicolì et al., 2023). In the study by Xin et al. (2024), it was shown that benefiting from the reproduction of circulation anomalies over the North Pacific and Northeast Asia, winter precipitation in Northeast Asia can be predicted by most CMIP6 DCPP models and the multi-model mean at the A2D timescale.

3.3. Mean Bias Assessment

In full-field initialization, biases are typically small in the initial stages but gradually increase as the forecast progresses, in contrast to anomaly initialization, where biases remain more consistent over the forecast period (Bellucci et al., 2015). Although statistical bias correction (as mentioned in Section 2.3) is applied during skill assessment, biases cannot be completely eliminated. To enhance the understanding of the IAP-CAS A2D model's systematic errors and the benefits of initialization, we analyze the mean bias in this section. Based on the International CLIVAR Project Office (ICPO, 2011) and previous studies (e.g., Bellucci et al., 2015; Nicolì et al., 2023), the mean bias is defined as the deviation of the model's climatology at different forecast years from the observational climatology over the entire study period. The evolution of mean biases with forecast year is presented in Figure 6. It is evident that cold SST biases are widespread across the forecast years in all of the involved oceans, including most regions of the Atlantic, the South Pacific and western North Pacific, and the central and eastern Indian Ocean, while a warm bias is observed in the oceans near Antarctica, the equatorial Pacific, and the eastern North Pacific. As the forecast period progresses, the IAP-CAS model gradually approaches its preferred climatology. It can be seen that, in the later years of the forecast, the hindcast biases (Figures 6b and 6c) closely resemble those of the historical simulations (Figure S5 in Supporting Information S1). Besides, the biases at forecast years 5 and 10 are comparable, indicating that the model's drift stabilizes regionally at least by year 5.

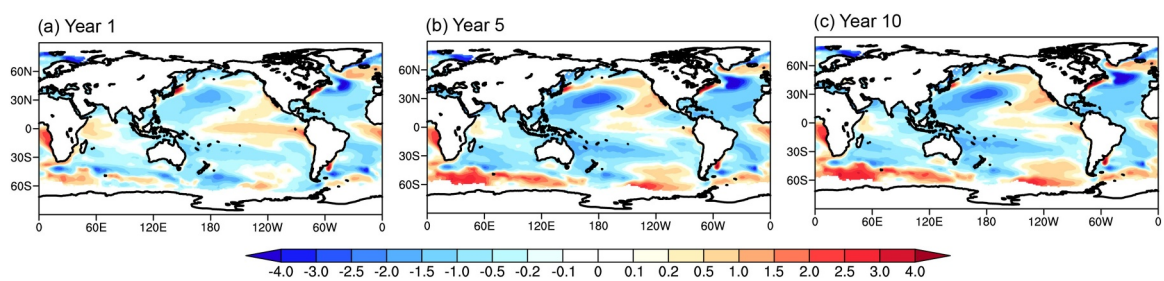

Figure 6. Mean bias of predicted SST in the IAP-CAS model (unit: K). The spatial distribution of SST bias for (a) forecast year 1, (b) year 5 and (c) year 10.

The pronounced cold mean bias is evident in the historical simulations in the central subpolar North Atlantic, further highlighting the systematic nature of the IAP-CAS model's biases in this region. However, the cold bias is also observed at the first forecast year of the hindcasts (Figure 6a), indicating that this region also suffered from initialization shock. Previous studies (Polkova et al., 2023; Robson et al., 2012) showed that the Atlantic Meridional Overturning Circulation (AMOC) and its northward heat transport have a substantial impact on the North Atlantic subpolar gyre, and initialization errors and drift in the AMOC can affect models' predictive skill for this region. We therefore speculate that the bias near the central subpolar North Atlantic in the IAP-CAS model may be linked to suboptimal AMOC simulations and their initial biases; however, a more detailed investigation is beyond the scope of this study and will be addressed in future work. On the other hand, a widely accepted view in the literature is that Arctic sea ice can induce variability in the AMOC and its northward heat transport as a forcing factor (e.g., Mauritzen & Häkkinen, 1997; Sévellec et al., 2017). Recent studies have indicated that incorporating sea ice initialization into models can enhance predictive skill of the North Atlantic region (Navarro et al., 2022; Passos et al., 2023). Since sea ice initialization is not included in the IAP-CAS A2D model, its potential inclusion could be a valuable direction for future research. Although noticeable biases persist in the North Atlantic of the hindcasts, the biases are still moderately reduced compared to the historical simulations, particularly during the first forecast year (Figure 6a) and to a lesser extent in the second year (not shown). This reduction is even more evident in other ocean regions.

3.4. Prediction of Climate Indices in the Pacific Ocean

A robust evaluation of model performance should encompass the simulation of climate modes and teleconnection patterns, as these are the primary sources of predictability (Merryfield et al., 2020). Based on the previous section, we find that SST in the Pacific and Indian Oceans has the highest predictive skill for the IAP-CAS model, while the Atlantic Ocean shows negative skill due to some initialization problems. Our preliminary diagnosis reveals that the hindcast skill for climate variability of the Indian Ocean, primarily the Indian Ocean Dipole (IOD), is limited. Therefore, we primarily focus on the Pacific region to examine the two main sources of predictability: PDO and ENSO.

3.4.1. PDO

The PDO (Mantua et al., 1997) is the main SST variability mode in the North Pacific and is closely linked to global climate and ecosystems. Figures 7a and 7b show the hindcast skill of the IAP-CAS model predicting the PDO index as a function of start forecast year and forecast period averaged. On the interannual scale (the first row in Figure 7a), the skill is significant for up to 3 years. The first forecast year has the highest skill with an ACC of 0.78 (MSSS of 0.57), while forecast year 3 has an ACC skill of 0.31 (MSSS fails significance test), suggesting that the IAP-CAS model not only captures the PDO's phase but also the amplitude at a 2-year lead. Multi-year averaging (rows 2–10 in Figure 7a) promotes positive ACC skill, with approximately one-third of the forecast years exceeding 0.5. The forecast years 1–2, 2–3, 1–3, 1–4, 2–10, and 1–10 exhibit remarkable skill. This is reasonable for the fact that the multi-year average filters out the internal variability, making the effect of external forcing more pronounced and thus improving skill (García-Serrano & Doblas-Reyes, 2012). For MSSS, positive skill only lasts until the forecast years 2–3, 2–4, 1–4, 1–5, and 1–6 in the forecast period averaged from 2 to 6 years respectively. We also noticed an improvement in skill for both ACC and MSSS during the final years, which currently lacks an explanation and may be attributable to the limited number of samples, as the years progress the number of samples decreases. This phenomenon of skill increase has also been noted in other models (e.g., Kim et al., 2012; Shaffrey et al., 2017).

Given the inherent challenges in predicting variability over the North Pacific, the skill of the IAP-CAS model in predicting the PDO aligns well with the range of existing A2D models. The DCPP MME represents the average performance of current A2D models. Figure S6 in Supporting Information S1 compares the PDO prediction skill between the IAP-CAS model and DCPP MME. The DCPP-MME demonstrates predictive skill for the PDO up to 2 years (Figures S6a and S6b in Supporting Information S1), which is consistent with previous studies (e.g., Choi & Son, 2022; Y. He et al., 2023). Positive skill differences (Figures S6c and S6d in Supporting Information S1) across most forecast periods indicate the good performance of the IAP-CAS model in PDO prediction. Compared to ACC, the difference in MSSS between the IAP-CAS model and DCPP-MME is more pronounced. This suggests that IAP-CAS not only performs well in predicting the phase of the PDO but also has an advantage in capturing its magnitude.

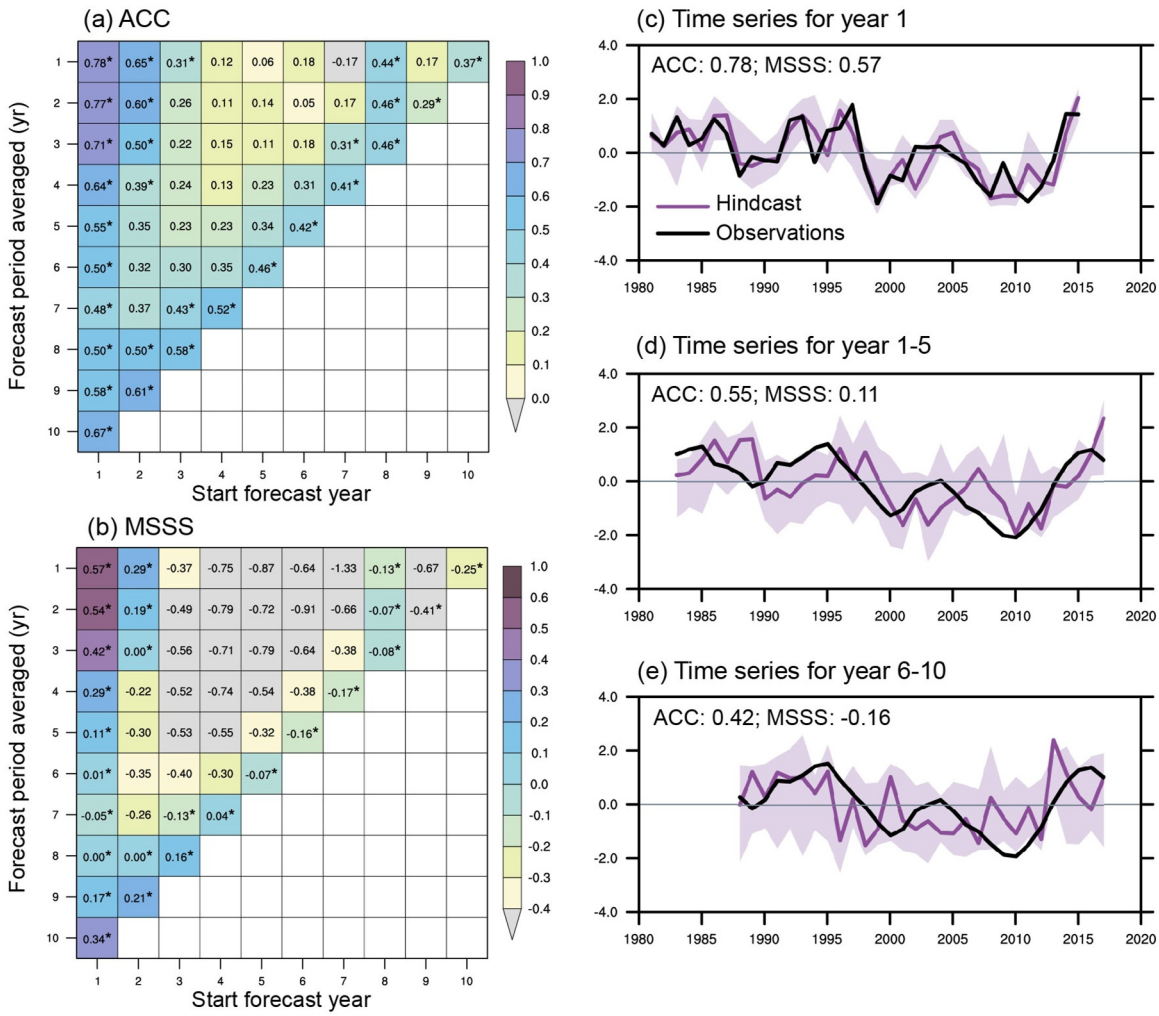

Figure 7. (a) ACC(H,O) of predicted PDO index as a function of the start forecast year (horizontal axis) and forecast period averaged (vertical axis). The numbers in each grid box represent skill score and the marker * denotes values significant at the 90% confidence level. (b) Is the same as (a) but for MSSS(H,O). Note that the range of label bar is different with (a). (c) The PDO time series for forecast year 1 of the hindcast (in purple line) and observations (in black line). The light purple shading denotes the ensemble spread (range of minimum and maximum among ensemble members). (d)–(e) is the same as (c) but for forecast years 1–5 and 6–10 respectively. For the 5-year forecast period averaged, the values are centered at the middle year.

The hindcast and observed PDO time series are presented in Figures 7c–7e. The observed PDO is in a positive phase during the early part of the study period, shifts to a negative phase from the late 1990s, and transitions to a positive phase around 2013. For forecast year 1, the IAP-CAS model aligns well with observations, capturing most of the interannual variability and accurately identifying the phase shift in the late 1990s. The main error arises from the late 1990s onwards, the so-called hiatus period, during which the PDO is increasingly affected by external forcing (Hua et al., 2018). The error may be associated with the limitations of the model's response to external forcing. For forecast years 1–5, the model generally captures the basic phase but with significant error. For forecast years 6–10, phase changes are not captured. Additionally, for all forecast years, the range of the PDO index predicted by the IAP-CAS model is comparable to observations. Even though the ensemble mean predictions for forecast years 6–10 are inaccurate, the ensemble spread is sufficiently large to cover the majority of the observations. As a supplement, Figure S7 in Supporting Information S1 presents the PDO time series from the historical simulations, which exhibit poor predictive skill. While the historical simulations successfully capture the spatial pattern and amplitude of the PDO (Section 2.4), the absence of initialization leads to a phase largely opposite to that of the observations.

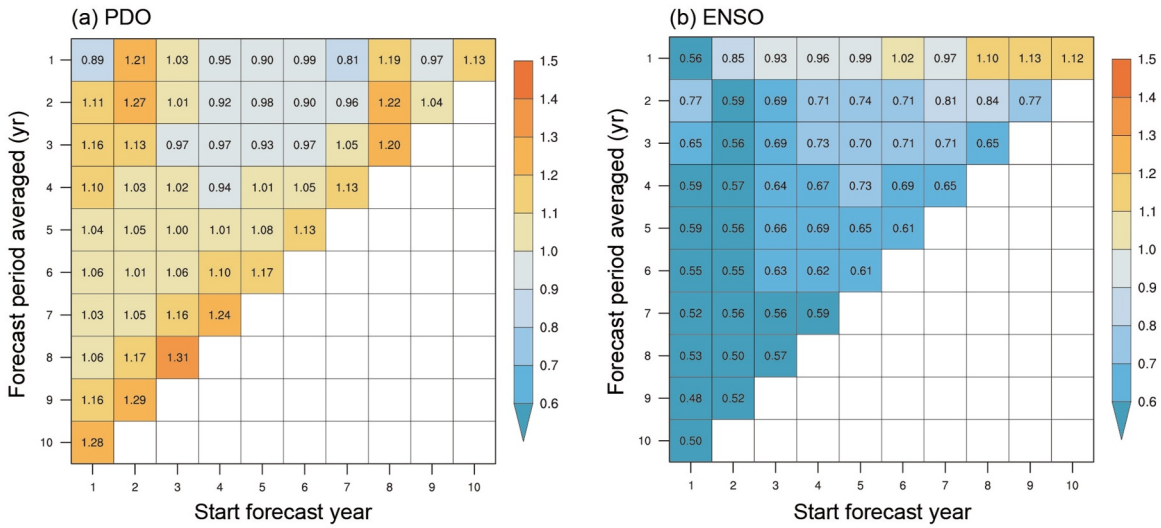

Figure 8. The spread–error ratio of predicted (a) PDO and (b) ENSO in the IAP-CAS model.

The SER is useful for assessing the effectiveness of the ensemble perturbation method. Figure 8a illustrates the SER of the IAP-CAS model predicting the PDO index. The SERs for most forecast years are close to 1, indicating generally robust forecast reliability. However, on the interannual scale (see the first row of Figure 8a), predictions tend to be overconfident (ensemble underdispersion). The disparity between ensemble spread and MSE reaches its peak at forecast year 2 and then stabilizes between forecast years 3 and 6. For multi-year averages (see rows 2–10 of Figure 8a), the ensemble is slightly overdispersed, with SERs greater than 1 for most forecast years. SERs for the 5-year average forecast period are relatively close to 1, indicating high reliability.

3.4.2. ENSO

El Niño-Southern Oscillation is a major player in the interannual cycle occurring in the tropical Pacific and has impact on the wide region of the globe through teleconnections (Yeh et al., 2018). Figure 9 shows that the skills for Nino3.4 of the IAP-CAS model are significant at forecast years 1, 1–2 and 1–3, with the ACC for forecast year 1 being 0.75 and the MSSS being 0.53. The skill shows a slight improvement in the latter forecast time, similar to the pattern with the predicted PDO, but the reason for this remains unclear. Monthly analysis of ENSO reveals that the correlation skill exceeds 0.5 at a forecast time of 13 months (Figures S8a and S8b in Supporting

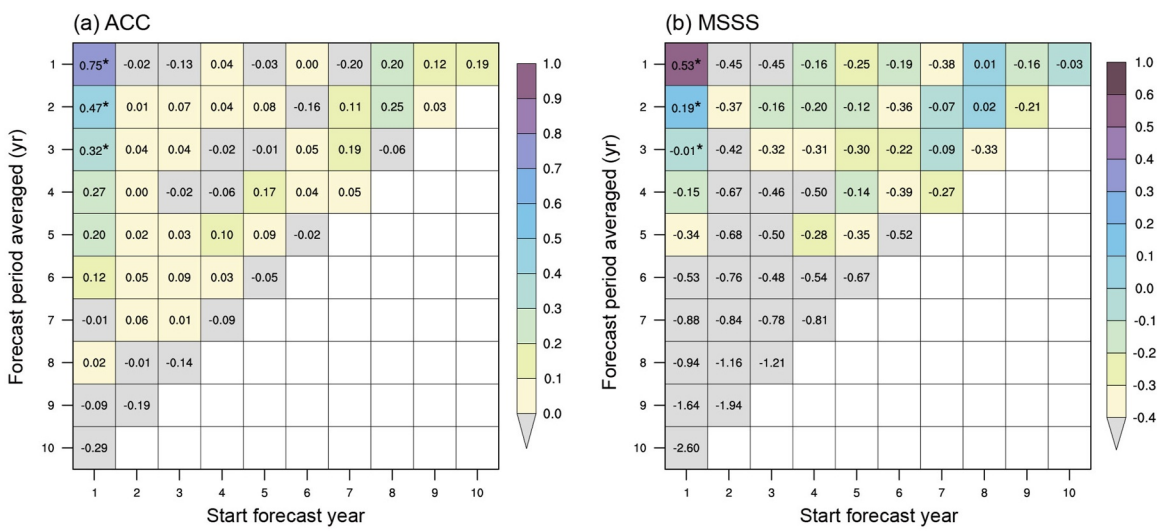

Figure 9. (a) ACC(H,O) and (b) MSSS(H,O) skill for predicted Nino 3.4 index in the IAP-CAS model.

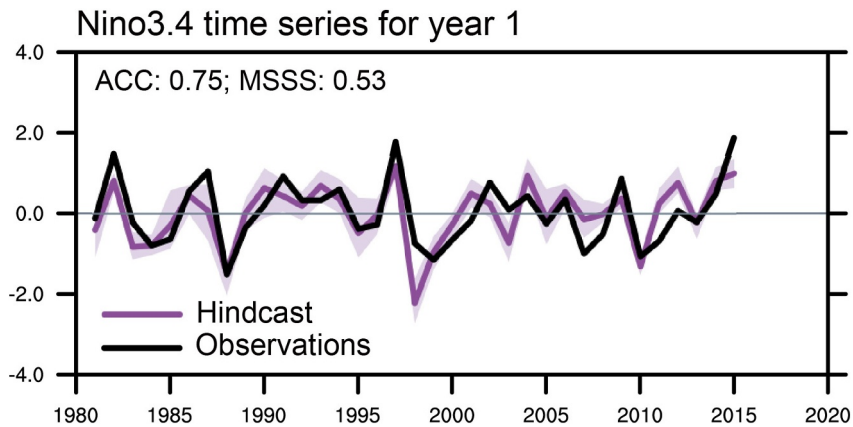

Figure 10. The time series for Nino 3.4 index of the IAP-CAS model hindcast (purple) and observations (black) at forecast year 1.

Information S1). Skill demonstrates significant levels in the first 12 months, showing all ACC greater than 0.6 and MSSS greater than 0.2. However, skill declines rapidly when entering the spring of forecast year 2 due to the spring barrier (Duan & Wei, 2013; Hou et al., 2019), which is also difficult to overcome in the CMIP5 and CMIP6 multi-model ensembles (e.g., Choi & Son, 2022). In contrast, as expected, the historical simulations exhibit poor skill for the Nino3.4 index (Figure S9 in Supporting Information S1). The monthly skill scores (Figures S9b and S9c in Supporting Information S1) for years 1 and 2 are shown, with all values remaining low, highlighting the difficulty of the IAP-CAS model in accurately predicting ENSO without initialization. In addition, the time series of the historical simulations (Figure S9a in Supporting Information S1) deviates significantly from the observations.

Figure 10 shows the predicted Nino3.4 time series of hindcasts for forecast year 1. The IAP-CAS model captures the overall observed phase changes except for years like 2003 and 2011, but it tends to overestimate ENSO events such as the 1998 La Niña event and the 2004 El Niño event, while underestimate events such as the 1982 and 2015 El Niño events. Interestingly, the IAP-CAS model predicts ENSO at forecast year 1 with a larger error during the hiatus phase compared to the pre-hiatus phase, which is also observed at the forecast year 1 prediction of the PDO. El Niño-Southern Oscillation forecast skill is closely linked to its amplitude. Lou et al. (2023) demonstrated a similar evolution of ENSO prediction skill, which may be associated with the observed high ENSO variance in the late 20th century and its subsequent decline in the early 21st century, as shown in Figure 6d of their study. Additionally, the ensemble spread of the IAP-CAS model's ENSO predictions is relatively small, indicating lower uncertainty in the forecasts. This low forecast uncertainty is also seen in the results of Lou et al. (2023) and Weisheimer et al. (2022), where the recent decades exhibit lower prediction uncertainty compared to the period prior to about 1960. This may be attributed to more robust observations during that time, which likely facilitated better representation in model initialization.

For ENSO prediction, single-model ensembles often exhibit an overly narrow spread or are overconfident (Larson & Kirtman, 2015). The results in Figure 8b show that the ENSO ensemble of the IAP-CAS model is under-dispersed in most of the years. The SER is low at forecast year 1 (at 0.56), indicating that the spread is relatively small while the error is large. The SER gradually approaches 1 as forecast time increases, achieving a balance between MSE and spread at forecast year 5, after which the spread exceeds the MSE. From the monthly analysis (Figure S8c in Supporting Information S1), the SER exhibits a similar trend, being lowest in the first month of the forecast period and increasing over time. When it reaches the spring and summer of the second year, the SER is closest to 1, then it slowly declines, but remains higher than at forecast year 1. Similar changes have been observed in other models (e.g., Bilbao et al., 2021).

3.5. The PDO Skill Benefits From Ocean-Atmosphere Interaction

As previously mentioned, the PDO prediction skill of the IAP-CAS model demonstrates comparable performance over the DCPP-MME, with ACC skill reaching 0.78 and MSSS skill of 0.57 at the first year. The PDO is recognized as a combination of local air-sea interactions and remote tropical forcing processes (Fang &

Yang, 2016; Newman et al., 2003, 2016). Constraining air-sea interactions in the North Pacific during initialization has been shown to significantly improve the A2D predictive skill for the PDO index (Y. He et al., 2023). On the interannual timescale, the forcing of ENSO and Aleutian Low on North Pacific SST response are approximately equal (N. Schneider & Cornuelle, 2005). The relationship between ENSO and PDO in CMIP models is associated with the ENSO amplitude and the response of the North Pacific SLP in the Aleutian Low region to ENSO (Nidheesh et al., 2017).

To understand the air-sea interactions in the IAP-CAS model, we conducted a composite analysis of global SST and SLP differences during El Niño and La Niña events (Figure 11), following the methodology of Alexander et al. (2002). The four strongest ENSO events during the study period are selected: 1982, 1987, 1997, and 2015 for El Niño, and 1988, 1998, 2007, and 2010 for La Niña. These events were identified based on NOAA's Climate Prediction Center (CPC)'s Oceanic Niño Index (ONI) (NOAA's Climate Prediction Center, 2024), which determines whether the 3-month running mean SST anomalies in the Niño 3.4 region exceed ±0.5°C. Considering all ENSO events during the study period yields similar results (figure not shown). Given the 13-month predictive skill of ENSO in the IAP-CAS model, we analyze only the first forecast year to ensure result robustness.

In observations (Figure 11a), ENSO events develop from summer, peak in winter, and persist into the following spring. Warm (cold) water gradually accumulates in the tropical eastern Pacific, while a pressure difference forms between the eastern and western hemispheres (Alexander et al., 2002). In response, an SLP anomaly develops over the extratropical Pacific, guiding the formation of cold water in the western ocean. For example, the SLP difference over the North Pacific begins to form in summer, and by winter, the Aleutian Low anomaly over the northeastern Pacific can exceed 15 mb. This anomaly strengthens the anomalous north-westerly winds over the central-western North Pacific and the anomalous southerly winds along the North American west coast, reinforcing the warm-cold water structure. El Niño-Southern Oscillation can also induce tropical North Atlantic SST warming through multiple teleconnection mechanisms (e.g., Chang et al., 2006; García-Serrano et al., 2017). In the tropical North Atlantic, SST anomalies develop from summer and peak in the following spring. Similarly, in the Indian Ocean, the weakened Walker circulation causes warm water to accumulate in the western Indian Ocean while cooling continues in the eastern Indian Ocean during El Niño events, forming a dipole SST distribution that triggers a positive IOD event (Saji et al., 1999).

The IAP-CAS model effectively simulates teleconnections between ENSO and the North Pacific (Figure 11b). Overall, the simulation of equatorial SST and SLP changes over the North Pacific aligns well with observations, though the intensity is somewhat diminished. Benefiting from initialization, during the first 2 months (April and May), it accurately captures the structure of equatorial ENSO. Although the atmospheric initialization information dissipates more quickly, the north-south SLP distribution over the North Pacific is well preserved. As ENSO progresses, the hemispheric SLP differences are captured by the model with the 0 mb contour shifted more westward than observation. In autumn and winter, the observed patterns are also captured, including the northeastward movement of the negative SLP anomaly over the North Pacific, although the structures are simpler and the SST intensity is weaker. Besides, the equatorial SST decays more significantly in spring compared to observations. In the Indian Ocean, the response in the IAP-CAS model is less pronounced, with weaker cold SST and less positive SLP anomalies on the eastern side, and later appearance of warm water in the western side during the IOD events. This indicates a weaker Walker circulation response in the model. Additionally, the IAP-CAS model does not reproduce the observed positive SST anomaly near 30°N in the tropical North Atlantic during summer and autumn, possibly due to the adverse impact of initialization shocks near the North Atlantic warming hole (as described in Section 3.2).

As a comparison, an analysis of SST and SLP differences for the four ENSO events simulated by the DCPP models at forecast year 1 was performed (Figure S10 in Supporting Information S1). The teleconnection between ENSO and the North Pacific is weaker in the DCPP MME. The seasonal evolution of equatorial SST aligns with observations, with development beginning in summer and peaking in autumn and winter. However, the amplitude and spatial extent are significantly smaller, likely due to the multi-model averaging that filters single-model signals. The cold-warm SST structure in the North Pacific is consistent with observations, though their intensities are reduced. Despite this, a weak negative SLP center is evident over the North Pacific in December, suggesting that the atmospheric bridge between ENSO and the North Pacific persists. However, its reduced intensity likely limits its influence on North Pacific SST and the maintenance of the SST structure.

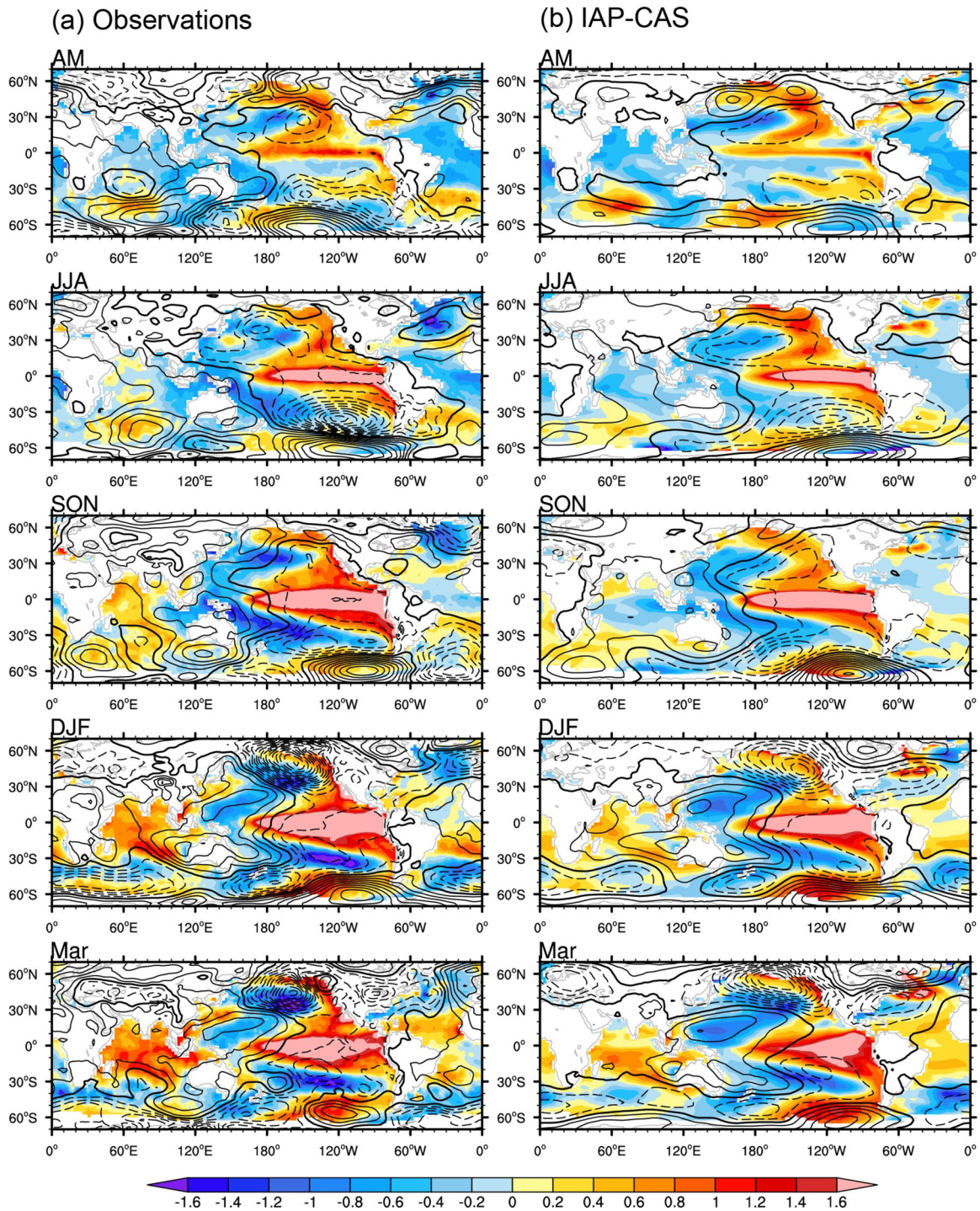

Figure 11. Differences of SST (shading, interval 0.2 K) and SLP (contours, from −15 to 10 mb, interval 1 mb; solid lines for positive values, dashed lines for negative values, thick solid line for zero value) during El Niño and La Niña events at forecast year 1 based on observations (a) and the IAP-CAS model (b) for April-May (AM), June-August (JJA), September-November (SON), December-February (DJF), and March. Months are corresponded to forecast months 1–12. SST observational data are from HadISST, and SLP data are from NCEP-NCAR Reanalysis 1.

4. Summary and Conclusions

In this study, we evaluate the predictive skill of the IAP-CAS A2D model for surface meteorological variables and major climate indices using a multi-decade hindcast data set (1981–2015) with 8-member ensemble initiated annually. This system represents an extension of the operational S2S prediction approved by the WMO/WWRP S2S panel. The climate model employed is the fully coupled atmosphere-ocean-land-ice FGOALS-f2 model, which adopts a full-field atmosphere-ocean initialization strategy. The effect of initialization is assessed by comparing initialized hindcast data with uninitialized historical simulations. Consistent with most decadal prediction systems, the IAP-CAS model demonstrates notable predictive skill for surface temperature, as indicated by positive ACC values over much of the globe, including the Indian Ocean, the Pacific Ocean, the mid-to-low latitudes of the Atlantic Ocean, and most land areas. Notably, the SST predictions for the eastern North Pacific and equatorial eastern Pacific averaged over the 1–5-year forecast period exhibit significantly positive ACC skill, which is rarely observed in other models, underscoring the distinct predictive capability of the IAP-CAS model. Comparison with historical simulations reveals that initialization significantly improves SST prediction skill during the first forecast year, particularly over the Pacific, Indian Ocean, and Southern Ocean regions. The GMST predictions capture the observed global warming trend and major interannual variations, including responses to key volcanic eruption events. Initialization notably enhances the amplitude of GMST predictions; however, it appears to overcorrect the model's response to external forcings, leading to an underestimation of global warming in multi-year average GMST predictions, particularly beyond the 5-year forecast horizon. Sea level pressure predictions during the first forecast year demonstrate robust hindcast skill particularly over the oceans. In contrast, the skill of regional precipitation predictions shows greater spatial variability, with notable accuracy confined to ENSO-related regions during the first forecast year and Northeast Asia in 5-year means.

A critical issue identified in this study is the limited SST prediction skill in the subpolar North Atlantic, characterized by persistently negative hindcast skill throughout the forecast period, which is lower than that of historical simulations. This limitation may arise from initialization shocks that amplify within the first 1–5 years of the forecast. Previous studies have reported potential shocks in ocean transport and heat advection within the SPG region triggered by full-field initialization, resulting in poor predictive skill in SST and ocean heat content, implying that it may not be an optimal initialization strategy for this region (Bilbao et al., 2021; Kröger et al., 2018). To further examine the effects of full-field initialization and associated model biases, we evaluated the SST mean bias. Our analysis reveals a systematic cold bias of the IAP-CAS model in the central subpolar North Atlantic, which is rapidly established during the first forecast year, suggesting a potential influence from initialization shocks. This cold bias is also observed in other A2D prediction models (e.g., Bellucci et al., 2015; Nicolì et al., 2023). However, the distinction between biases arising from initialization shocks and those resulting from inherent model drift has not been explicitly addressed in this study. Nevertheless, our findings underscore the critical need to address biases in the North Atlantic and mitigate initialization shocks in A2D predictions. Given the significant role of the North Atlantic in global climate variability and teleconnections, enhancing A2D prediction skill in this region remains an urgent priority for advancing predictive capabilities.

The prediction skill for the two primary modes of SST variability in the Pacific, ENSO and PDO, is evaluated. El Niño-Southern Oscillation predictions show notable skill at the first forecast year and are comparable to the performance of leading seasonal dynamical prediction systems. For PDO predictions, the IAP-CAS model demonstrates significant annual-mean PDO correlation skill for up to 3 years. The historical simulations exhibit low skill in capturing these two climate variabilities, which is expected due to the absence of initialization information. The DCPP MME is analyzed as a benchmark to assess the IAP-CAS model's prediction skill for the PDO. The results show that the IAP-CAS model's skill is comparable to the DCPP MME across most forecast years. The IAP-CAS model demonstrates better performance in predicting the amplitude of PDO variability, as indicated by pronounced positive MSSS difference compared with the DCPP-MME.

Simulations of teleconnections arising from internal variabilities and air-sea interactions are crucial for A2D predictions (Cassou et al., 2018; Meehl et al., 2009), yet less extensively studied in previous studies. For instance, ENSO's influence on other regions is largely mediated by atmospheric responses (Trenberth et al., 1998; Yeh et al., 2018). The extratropical atmospheric responses, in turn, impact local climate variability, including the ocean-atmospheric interactions between Aleutian Low and PDO variations. Composite analysis of the first forecast year in this study demonstrates that the IAP-CAS model realistically simulates the seasonal evolution of equatorial SSTA during ENSO events and captures the structure of SLP anomalies of the Aleutian Low over the

North Pacific. This consistent representation implies an established teleconnection between ENSO and the North Pacific atmosphere in the IAP-CAS model, which likely underpins the predictive skill of PDO. While the weaker SST and SLP intensity in the DCPP MME implies reduced ENSO teleconnections and limited influence on North Pacific SST, it may consequently fail to adequately support the prediction of PDO. Our results underscore the importance of simulating teleconnections and air-sea interactions to improve predictive A2D skill, potentially offering valuable insights for refining the A2D prediction framework in future studies.

This study provides guidance for users of the IAP-CAS A2D prediction data set, serving as a key reference for its application. It is important to note, however, that the limited sample size of historical simulations used in this study may have partially influenced the evaluation of A2D initialization effectiveness. Notable SST prediction biases in the North Atlantic are identified in this study, which highlight a potential direction for future model improvements. The model's strong skill in predicting near-surface temperature and the PDO system suggests significant practical value. Meanwhile, the increasing share of renewable energy sources worldwide is driving a growing demand for climate services (Bloomfield et al., 2021; Troccoli et al., 2018; Yang et al., 2022). We aim to extend the capabilities of the IAP-CAS A2D model to deliver reliable predictions for renewable energy applications and to evaluate its predictive skill in this context.

Data Availability Statement

The hindcast data of the IAP-CAS A2D model used in this study can be accessed via Zenodo at https://zenodo.org/records/13834418. The CMIP6 DCPP (Boer et al., 2016) model data used in this study is accessible via the ESGF node (https://aims2.llnl.gov/search/cmip6/). Temperature observation data (Morice et al., 2021; Rayner et al., 2003) are available at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html and https://climatedataguide.ucar.edu/climate-data/global-land-ocean-surface-temperature-data-hadcrut5. Precipitation observation data (U. Schneider et al., 2022) can be obtained at https://opendata.dwd.de/climate_environment/GPCC/html/download_gate.html. Sea level pressure observation data (Kalnay et al., 1996) are available at https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html. The analysis was performed using NCL version 6.6.2 (Brown et al., 2012), and CDO version 1.9.8 (CDO Core Team, 2019; Schulzweida, 2019) and NCO version 4.9.1 (Zender, 2008). Some of the color schemes used in the figures were based on Tol (2021).

Acknowledgments
This study was supported by the National Natural Science Foundation of China (Grant 42475155 and 42175161), and the Earth System Numerical Simulation Facility project of the National Key Scientific and Technological Infrastructure. We would like to acknowledge the climate model groups that contributed to the CMIP6 DCPP project.

References

Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N.-C., & Scott, J. D. (2002). The atmospheric bridge: The influence of ENSO teleconnections on Air–Sea interaction over the global oceans. *Journal of Climate*, *15*(16), 2205–2231. https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2

An, B., Yu, Y., Bao, Q., He, B., Li, J., Luan, Y., et al. (2022). CAS FGOALS-f3-H dataset for the high-resolution model intercomparison project(HighResMIP) tier 2. *Advances in Atmospheric Sciences*, *39*(11), 1873–1884. https://doi.org/10.1007/s00376-022-2030-5

Bao, Q., Wu, X., Li, J., Wang, L., He, B., Wang, X., et al. (2019). Outlook for El Nino and the Indian Ocean Dipole in autumn-winter 2018-2019 (in Chinese). *Chinese Science Bulletin*, *64*(1), 73–78. https://doi.org/10.1360/n972018-00913

Barnston, A. G., Tippett, M. K., Ranganathan, M., & L'Heureux, M. L. (2019). Deterministic skill of ENSO predictions from the North American multimodel ensemble. *Climate Dynamics*, *53*(12), 7215–7234. https://doi.org/10.1007/s00382-017-3603-3

Basconcillo, J., & Moon, I.-J. (2022). Increasing activity of tropical cyclones in East Asia during the mature boreal autumn linked to long-term climate variability. *Npj Climate and Atmospheric Science*, *5*(1), 1–11. https://doi.org/10.1038/s41612-021-00222-6

Bellucci, A., Haarsma, R., Gualdi, S., Athanasiadis, P. J., Caian, M., Cassou, C., et al. (2015). An assessment of a multi-model ensemble of decadal climate predictions. *Climate Dynamics*, *44*(9–10), 2787–2806. https://doi.org/10.1007/s00382-014-2164-y

Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P.-A., et al. (2021). Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth. *Earth System Dynamics*, *12*(1), 173–196. https://doi.org/10.5194/esd-12-173-2021

Bloomfield, H. C., Gonzalez, P. L. M., Lundquist, J. K., Stoop, L. P., Browell, J., Dargaville, R., et al. (2021). The importance of weather and climate to energy systems: A workshop on next generation challenges in energy–climate modeling. *Bulletin of the American Meteorological Society*, *102*(1), E159–E167. https://doi.org/10.1175/BAMS-D-20-0256.1

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., et al. (2016). The decadal climate prediction project (DCPP) contribution to CMIP6. *Geoscientific Model Development*, *9*(10), 3751–3777. https://doi.org/10.5194/gmd-9-3751-2016

Brown, D., Brownrigg, R., Haley, M., & Huang, W. (2012). NCAR Command Language (NCL) [Software]. *UCAR/NCAR - Computational and Information Systems Laboratory (CISL)*. https://doi.org/10.5065/D6WD3XH5

Cassou, C., Kushnir, Y., Hawkins, E., Pirani, A., Kucharski, F., Kang, I.-S., & Caltabiano, N. (2018). Decadal climate variability and predictability: Challenges and opportunities. *Bulletin of the American Meteorological Society*, *99*(3), 479–490. https://doi.org/10.1175/BAMS-D-16-0286.1

CDO Core Team. (2019). Climate data operators (version 1.9.8) [Software]. Retrieved from https://code.mpimet.mpg.de/projects/cdo

Chang, P., Fang, Y., Saravanan, R., Ji, L., & Seidel, H. (2006). The cause of the fragile relationship between the Pacific El Niño and the Atlantic Niño. *Nature*, *443*(7109), 324–328. https://doi.org/10.1038/nature05053

Choi, J., & Son, S.-W. (2022). Seasonal-to-decadal prediction of El Niño–Southern Oscillation and Pacific decadal oscillation. *Npj Climate and Atmospheric Science*, *5*(1), 29. https://doi.org/10.1038/s41612-022-00251-9

Choudhury, D., Sen Gupta, A., Sharma, A., Mehrotra, R., & Sivakumar, B. (2017). An assessment of drift correction alternatives for CMIP5 decadal predictions. *Journal of Geophysical Research: Atmospheres*, *122*(19), 10282–10296. https://doi.org/10.1002/2017JD026900

Comiso, J. C. (2000). Variability and trends in Antarctic surface temperatures from in situ and satellite infrared measurements. *Journal of Climate*, *13*(10), 1674–1696. https://doi.org/10.1175/1520-0442(2000)013<1674:vatias>2.0.co;2

Craig, A. P., Vertenstein, M., & Jacob, R. (2012). A new flexible coupler for earth system modeling developed for CCSM4 and CESM1. *The International Journal of High Performance Computing Applications*, *26*(1), 31–42. https://doi.org/10.1177/1094342011428141

Dai, A., & Wigley, T. M. L. (2000). Global patterns of ENSO-induced precipitation. *Geophysical Research Letters*, *27*(9), 1283–1286. https://doi.org/10.1029/1999GL011140

Doblas-Reyes, F. J., Andreu-Burillo, I., Chikamoto, Y., García-Serrano, J., Guemas, V., Kimoto, M., et al. (2013). Initialized near-term regional climate change prediction. *Nature Communications*, *4*(1), 1. https://doi.org/10.1038/ncomms2704

Duan, W., & Wei, C. (2013). The 'spring predictability barrier' for ENSO predictions and its possible mechanism: Results from a fully coupled model. *International Journal of Climatology*, *33*(5), 1280–1292. https://doi.org/10.1002/joc.3513

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fang, J., & Yang, X.-Q. (2016). Structure and dynamics of decadal anomalies in the wintertime midlatitude North Pacific ocean–atmosphere system. *Climate Dynamics*, *47*(5), 1989–2007. https://doi.org/10.1007/s00382-015-2946-x

Findell, K. L., Sutton, R., Caltabiano, N., Brookshaw, A., Heimbach, P., Kimoto, M., et al. (2023). Explaining and predicting earth system change: A World climate research Programme call to action. *Bulletin of the American Meteorological Society*, *104*(1), E325–E339. https://doi.org/10.1175/BAMS-D-21-0280.1

García-Serrano, J., Cassou, C., Douville, H., Giannini, A., & Doblas-Reyes, F. J. (2017). Revisiting the ENSO teleconnection to the tropical North Atlantic. *Journal of Climate*, *30*(17), 6945–6957. https://doi.org/10.1175/JCLI-D-16-0641.1

García-Serrano, J., & Doblas-Reyes, F. J. (2012). On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Climate Dynamics*, *39*(7–8), 2025–2040. https://doi.org/10.1007/s00382-012-1413-1

Gervais, M., Shaman, J., & Kushnir, Y. (2020). Impact of the North Atlantic warming hole on sensible weather. *Journal of Climate*, *33*(10), 4255–4271. https://doi.org/10.1175/JCLI-D-19-0636.1

Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., et al. (2013). A verification framework for interannual-to-decadal predictions experiments. *Climate Dynamics*, *40*(1–2), 245–272. https://doi.org/10.1007/s00382-012-1481-2

Guemas, V., Doblas-Reyes, F. J., Lienert, F., Soufflet, Y., & Du, H. (2012). Identifying the causes of the poor decadal climate prediction skill over the North Pacific. *Journal of Geophysical Research*, *117*(D20). https://doi.org/10.1029/2012JD018004

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., et al. (2021). Changing state of the climate system. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 287–422). Cambridge University Press.

He, B., Bao, Q., Wang, X., Zhou, L., Wu, X., Liu, Y., et al. (2019). CAS FGOALS-f3-L model datasets for CMIP6 historical atmospheric model intercomparison project simulation. *Advances in Atmospheric Sciences*, *36*(8), 771–778. https://doi.org/10.1007/s00376-019-9027-8

He, B., Liu, Y., Wu, G., Bao, Q., Zhou, T., Wu, X., et al. (2020). CAS FGOALS-f3-L model datasets for CMIP6 GMMIP Tier-1 and Tier-3 experiments. *Advances in Atmospheric Sciences*, *37*(1), 18–28. https://doi.org/10.1007/s00376-019-9085-y

He, Y., Wang, B., Fang, J., Yu, Y., Li, L., Liu, J., et al. (2023). Constraint of Air–Sea interaction significant to skillful predictions of North Pacific climate variations. *Journal of Climate*, *36*(17), 5941–5962. https://doi.org/10.1175/JCLI-D-22-0635.1

Hermanson, L., Smith, D., Seabrook, M., Bilbao, R., Doblas-Reyes, F., Tourigny, E., et al. (2022). WMO global annual to decadal climate update: A prediction for 2021–25. *Bulletin of the American Meteorological Society*, *103*(4), E1117–E1129. https://doi.org/10.1175/BAMS-D-20-0311.1

Ho, C. K., Hawkins, E., Shaffrey, L., Bröcker, J., Hermanson, L., Murphy, J. M., et al. (2013). Examining reliability of seasonal to decadal sea surface temperature forecasts: The role of ensemble dispersion. *Geophysical Research Letters*, *40*(21), 5770–5775. https://doi.org/10.1002/2013GL057630

Hoffman, R. N., & Kalnay, E. (1983). Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, *35A*(2), 100–118. https://doi.org/10.1111/j.1600-0870.1983.tb00189.x

Hou, M., Duan, W., & Zhi, X. (2019). Season-dependent predictability barrier for two types of El Niño revealed by an approach to data analysis for predictability. *Climate Dynamics*, *53*(9), 5561–5581. https://doi.org/10.1007/s00382-019-04888-w

Hu, S., Wu, B., Wang, Y., Zhou, T., Yu, Y., He, B., et al. (2023). CAS FGOALS-f3-L model datasets for CMIP6 DCPP experiment. *Advances in Atmospheric Sciences*, *40*(10), 1911–1922. https://doi.org/10.1007/s00376-023-2122-x

Hua, W., Dai, A., & Qin, M. (2018). Contributions of internal variability and external forcing to the recent Pacific decadal variations. *Geophysical Research Letters*, *45*(14), 7084–7092. https://doi.org/10.1029/2018GL079033

Hunke, E. C., & Lipscomb, W. H. (2010). CICE: The Los Alamos Sea Ice model documentation and software user's manual version 4.

ICPO (International CLIVAR Project Office). (2011). Decadal and bias correction for decadal climate predictions *(150; CLIVAR Publication Series)*. International CLIVAR Project Office.

Jeuken, A. B. M., Siegmund, P. C., Heijboer, L. C., Feichter, J., & Bengtsson, L. (1996). On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *Journal of Geophysical Research*, *101*(D12), 16939–16950. https://doi.org/10.1029/96JD01218

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–472. https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Karnauskas, K. B., Zhang, L., & Amaya, D. J. (2021). The atmospheric response to North Atlantic SST trends, 1870–2019. *Geophysical Research Letters*, *48*(2), e2020GL090677. https://doi.org/10.1029/2020GL090677

Kataoka, T., Tatebe, H., Koyama, H., Mochizuki, T., Ogochi, K., Naoe, H., et al. (2020). Seasonal to decadal predictions with MIROC6: Description and basic evaluation. *Journal of Advances in Modeling Earth Systems*, *12*(12), e2019MS002035. https://doi.org/10.1029/2019MS002035

Keil, P., Mauritsen, T., Jungclaus, J., Hedemann, C., Olonscheck, D., & Ghosh, R. (2020). Multiple drivers of the North Atlantic warming hole. *Nature Climate Change*, *10*(7), 667–671. https://doi.org/10.1038/s41558-020-0819-8

Kerbyson, D. J., & Jones, P. W. (2005). A performance model of the Parallel Ocean Program. *The International Journal of High Performance Computing Applications*, *19*(3), 261–276. https://doi.org/10.1177/1094342005056114

Kim, H.-M., Webster, P. J., & Curry, J. A. (2012). Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophysical Research Letters*, *39*(10). https://doi.org/10.1029/2012GL051644

Kröger, J., Pohlmann, H., Sienz, F., Marotzke, J., Baehr, J., Köhl, A., et al. (2018). Full-field initialized decadal predictions with the MPI earth system model: An initial shock in the North Atlantic. *Climate Dynamics*, *51*(7), 2593–2608. https://doi.org/10.1007/s00382-017-4030-1

Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., et al. (2019). Towards operational predictions of the near-term climate. *Nature Climate Change*, *9*(2), 2–101. https://doi.org/10.1038/s41558-018-0359-7

Larson, S. M., & Kirtman, B. P. (2015). An alternate approach to ensemble ENSO forecast spread: Application to the 2014 forecast. *Geophysical Research Letters*, *42*(21), 9411–9415. https://doi.org/10.1002/2015GL066173

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., et al. (2011). Parameterization improvements and functional and structural advances in version 4 of the community land model. *Journal of Advances in Modeling Earth Systems*, *3*(1). https://doi.org/10.1029/2011MS00045

Lee, J., Sperber, K. R., Gleckler, P. J., Bonfils, C. J. W., & Taylor, K. E. (2019). Quantifying the agreement between observed and simulated extratropical modes of interannual variability. *Climate Dynamics*, *52*(7), 4057–4089. https://doi.org/10.1007/s00382-018-4355-4

Li, J., Bao, Q., Liu, Y., Wu, G., Wang, L., He, B., et al. (2021). Dynamical seasonal prediction of tropical cyclone activity using the FGOALS-f2 ensemble prediction system. *Weather and Forecasting*, *36*(5), 1759–1778. https://doi.org/10.1175/WAF-D-20-0189.1

Li, S., Wu, L., Yang, Y., Geng, T., Cai, W., Gan, B., et al. (2020). The Pacific Decadal Oscillation less predictable under greenhouse warming. *Nature Climate Change*, *10*(1), 1–34. https://doi.org/10.1038/s41558-019-0663-x

Lin, S.-J. (2004). A "vertically Lagrangian" finite-volume dynamical core for global models. *Monthly Weather Review*, *132*(10), 2293–2307. https://doi.org/10.1175/1520-0493(2004)132<2293:AVLFDC>2.0.CO;2

Liu, Y., Bao, Q., He, B., Wu, X., Yang, J., Liu, Y., et al. (2024). Dynamical Madden–Julian Oscillation forecasts using an ensemble subseasonal-to-seasonal forecast system of the IAP-CAS model. *Geoscientific Model Development*, *17*(16), 6249–6275. https://doi.org/10.5194/gmd-17-6249-2024

Liu, Z., & Di Lorenzo, E. (2018). Mechanisms and predictability of Pacific decadal variability. *Current Climate Change Reports*, *4*(2), 128–144. https://doi.org/10.1007/s40641-018-0090-5

Liu, Z., Jiang, L., Shi, C., Zhang, T., Zhou, Z., Liao, J., et al. (2023). CRA-40/Atmosphere—The first-generation Chinese atmospheric reanalysis (1979–2018): System description and performance evaluation. *Journal of Meteorological Research*, *37*(1), 1–19. https://doi.org/10.1007/s13351-023-2086-x

Lou, J., Newman, M., & Hoell, A. (2023). Multi-decadal variation of ENSO forecast skill since the late 1800s. *Npj Climate and Atmospheric Science*, *6*(1), 1–14. https://doi.org/10.1038/s41612-023-00417-z

Malakar, Y., Snow, S., Fleming, A., Fielke, S., Jakku, E., Tozer, C., & Darbyshire, R. (2024). Multi-decadal climate services help farmers assess and manage future risks. *Nature Climate Change*, *14*(6), 586–591. https://doi.org/10.1038/s41558-024-02021-2

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., & Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on Salmon production. *Bulletin of the American Meteorological Society*, *78*(6), 1069–1080. https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2

Mauritzen, C., & Häkkinen, S. (1997). Influence of sea ice on the thermohaline circulation in the Arctic-North Atlantic Ocean. *Geophysical Research Letters*, *24*(24), 3257–3260. https://doi.org/10.1029/97GL03192

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., et al. (2009). DECADAL PREDICTION can it Be skillful? *Bulletin of the American Meteorological Society*, *90*(10), 1467–1485. https://doi.org/10.1175/2009BAMS2778.1

Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F., et al. (2021). Initialized Earth System prediction from subseasonal to decadal timescales. *Nature Reviews Earth and Environment*, *2*(5), 340–357. https://doi.org/10.1038/s43017-021-00155-x

Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A. S., et al. (2020). Current and emerging developments in subseasonal to decadal prediction. *Bulletin of the American Meteorological Society*, *101*(6), E869–E896. https://doi.org/10.1175/BAMS-D-19-0037.1

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., et al. (2021). An updated assessment of near-surface temperature change from 1850: The HadCRUT5 data set. *Journal of Geophysical Research: Atmospheres*, *126*(3), e2019JD032361. https://doi.org/10.1029/2019JD032361

Navarro, J. C. A., García-Serrano, J., Lapin, V., & Ortega, P. (2022). Added value of assimilating springtime Arctic sea ice concentration in summer-fall climate predictions. *Environmental Research Letters*, *17*(6), 064008. https://doi.org/10.1088/1748-9326/ac6c9b

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., et al. (2016). The Pacific decadal oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427. https://doi.org/10.1175/JCLI-D-15-0508.1

Newman, M., Compo, G. P., & Alexander, M. A. (2003). ENSO-forced variability of the Pacific decadal oscillation. *Journal of Climate*, *16*(23), 3853–3857. https://doi.org/10.1175/1520-0442(2003)016<3853:EVOTPD>2.0.CO;2

Nicolì, D., Bellucci, A., Ruggieri, P., Athanasiadis, P. J., Materia, S., Peano, D., et al. (2023). The Euro-Mediterranean center on climate change (CMCC) decadal prediction system. *Geoscientific Model Development*, *16*(1), 179–197. https://doi.org/10.5194/gmd-16-179-2023

Nidheesh, A. G., Lengaigne, M., Vialard, J., Izumo, T., Unnikrishnan, A. S., & Cassou, C. (2017). Influence of ENSO on the Pacific decadal oscillation in CMIP models. *Climate Dynamics*, *49*(9), 3309–3326. https://doi.org/10.1007/s00382-016-3514-8

NOAA's Climate Prediction Center. (2024). Historical El Nino/La Nina episodes (1950-present). Retrieved from https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php

Oleson, K., Lawrence, D., B, G., Flanner, M., Kluzek, E., Lawrence, P., et al. (2010). Technical description of version 4.0 of the community land model (CLM).

Passos, L., Langehaug, H. R., Årthun, M., Eldevik, T., Bethke, I., & Kimmritz, M. (2023). Impact of initialization methods on the predictive skill in NorCPM: An Arctic–Atlantic case study. *Climate Dynamics*, *60*(7), 2061–2080. https://doi.org/10.1007/s00382-022-06437-4

Polkova, I., Swingedouw, D., Hermanson, L., Köhl, A., Stammer, D., Smith, D., et al. (2023). Initialization shock in the ocean circulation reduces skill in decadal predictions of the North Atlantic subpolar gyre. *Frontiers in Climate*, *5*. https://doi.org/10.3389/fclim.2023.1273770

Putman, W. M., & Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, *227*(1), 55–78. https://doi.org/10.1016/j.jcp.2007.07.022

Rahmstorf, S., Box, J. E., Feulner, G., Mann, M. E., Robinson, A., Rutherford, S., & Schaffernicht, E. J. (2015). Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nature Climate Change*, *5*(5), 475–480. https://doi.org/10.1038/nclimate2554

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14), 2002JD002670. https://doi.org/10.1029/2002JD002670

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*(22), 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Robson, J., Sutton, R., Lohmann, K., Smith, D., & Palmer, M. D. (2012). Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s. *Journal of Climate*, *25*(12), 4116–4134. https://doi.org/10.1175/JCLI-D-11-00443.1

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., & Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature*, *401*(6751), 360–363. https://doi.org/10.1038/43854

Schneider, N., & Cornuelle, B. D. (2005). The forcing of the Pacific decadal oscillation. *Journal of Climate*, *18*(21), 4355–4373. https://doi.org/10.1175/JCLI3527.1

Schneider, U., Hänsel, S., Finger, P., Rustemeier, E., & Ziese, M. (2022). GPCC full data monthly Product version 2022 at 2.5°: Monthly land-surface precipitation from rain-gauges built on GTS-based and historical data. [Dataset]. https://doi.org/10.5676/DWD_GPCC/FD_M_V2022_250

Schulzweida, U. (2019). CDO user guide. *Zenodo*. https://doi.org/10.5281/ZENODO.4246855

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., et al. (2021). Weather and climate extreme events in a changing climate. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (pp. 1513–1766). Cambridge University Press.

Sévellec, F., Fedorov, A. V., & Liu, W. (2017). Arctic sea-Ice decline weakens the Atlantic meridional overturning circulation. *Nature Climate Change*, *7*(8), 604–610. https://doi.org/10.1038/nclimate3353

Shaffrey, L. C., Hodson, D., Robson, J., Stevens, D. P., Hawkins, E., Polo, I., et al. (2017). Decadal predictions with the HiGEM high resolution global coupled climate model: Description and basic evaluation. *Climate Dynamics*, *48*(1), 297–311. https://doi.org/10.1007/s00382-016-3075-x

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Tippett, M. K., Barnston, A. G., & Li, S. (2012). Performance of recent multimodel ENSO forecasts. *Journal of Applied Meteorology and Climatology*, *51*(3), 637–654. https://doi.org/10.1175/JAMC-D-11-093.1

Tol, P. (2021). Paul Tol's notes. Retrieved from https://personal.sron.nl/~pault/

Trenberth, K. E., Branstator, G. W., Karoly, D., Kumar, A., Lau, N.-C., & Ropelewski, C. (1998). Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *Journal of Geophysical Research*, *103*(C7), 14291–14324. https://doi.org/10.1029/97JC01444

Trenberth, K. E., & Shea, D. J. (2006). Atlantic hurricanes and natural variability in 2005. *Geophysical Research Letters*, *33*(12). https://doi.org/10.1029/2006GL026894

Troccoli, A., Goodess, C., Jones, P., Penny, L., Dorling, S., Harpham, C., et al. (2018). Creating a proof-of-concept climate service to assess future renewable energy mixes in Europe: An overview of the C3S ECEM project. *Advances in Science and Research*, *15*, 191–205. https://doi.org/10.5194/asr-15-191-2018

Weisheimer, A., Balmaseda, M. A., Stockdale, T. N., Mayer, M., Sharmila, S., Hendon, H., & Alves, O. (2022). Variability of ENSO forecast skill in 2-year global reforecasts over the 20th century. *Geophysical Research Letters*, *49*(10), e2022GL097885. https://doi.org/10.1029/2022GL097885

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed., Vol. 100). Academic Press.

World Meteorological Organization (WMO). (2024). *WMO global annual to decadal climate update target years: 2024 and 2024-2028*. WMO.

Wu, L., Wang, B., & Geng, S. (2005). Growing typhoon influence on east Asia. *Geophysical Research Letters*, *32*(18). https://doi.org/10.1029/2005GL022937

Xin, X., Wu, T., Zheng, M., Fang, Y., Lu, Y., & Zhang, J. (2024). Decadal prediction of Northeast Asian winter precipitation with CMIP6 models. *Climate Dynamics*, *62*(5), 3245–3259. https://doi.org/10.1007/s00382-023-07063-4

Yang, D., Wang, W., Gueymard, C. A., Hong, T., Kleissl, J., Huang, J., et al. (2022). A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality. *Renewable and Sustainable Energy Reviews*, *161*, 112348. https://doi.org/10.1016/j.rser.2022.112348

Yeager, S. G., Chang, P., Danabasoglu, G., Rosenbloom, N., Zhang, Q., Castruccio, F. S., et al. (2023). Reduced Southern Ocean warming enhances global skill and signal-to-noise in an eddy-resolving decadal prediction system. *Npj Climate and Atmospheric Science*, *6*(1), 107. https://doi.org/10.1038/s41612-023-00434-y

Yeh, S.-W., Cai, W., Min, S.-K., McPhaden, M. J., Dommenget, D., Dewitte, B., et al. (2018). ENSO atmospheric teleconnections and their response to greenhouse gas forcing. *Reviews of Geophysics*, *56*(1), 185–206. https://doi.org/10.1002/2017RG000568

Zender, C. S. (2008). Analysis of self-describing gridded geoscience data with netCDF Operators (NCO). *Environmental Modelling and Software*, *23*(10), 1338–1342. https://doi.org/10.1016/j.envsoft.2008.03.004

Zeng, L., Bao, Q., Wu, X., He, B., Yang, J., Wang, T., et al. (2023). Impacts of humidity initialization on MJO prediction: A study in an operational sub-seasonal to seasonal system. *Atmospheric Research*, *294*, 106946. https://doi.org/10.1016/j.atmosres.2023.106946

Zhou, L., Bao, Q., Liu, Y., Wu, G., Wang, W.-C., Wang, X., et al. (2015). Global energy and water balance: Characteristics from finite-volume atmospheric model of the IAP/LASG (FAMIL1). *Journal of Advances in Modeling Earth Systems*, *7*(1), 1–20. https://doi.org/10.1002/2014MS000349

References From the Supporting Information

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., et al. (2021). NorCPM1 and its contribution to CMIP6 DCPP. *Geoscientific Model Development*, *14*(11), 7073–7116. https://doi.org/10.5194/gmd-14-7073-2021

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002010. https://doi.org/10.1029/2019MS002010

Sospedra-Alfonso, R., Merryfield, W. J., Boer, G. J., Kharin, V. V., Lee, W.-S., Seiler, C., & Christian, J. R. (2021). Decadal climate predictions with the Canadian earth system model version 5 (CanESM5). *Geoscientific Model Development*, *14*(11), 6863–6891. https://doi.org/10.5194/gmd-14-6863-2021

Yeager, S. G., Danabasoglu, G., Rosenbloom, N. A., Strand, W., Bates, S. C., Meehl, G. A., et al. (2018). Predicting near-term changes in the earth system: A large ensemble of initialized decadal prediction simulations using the community earth system model. *Bulletin of the American Meteorological Society*, *99*(9), 1867–1886. https://doi.org/10.1175/BAMS-D-17-0098.1