



A study of predictability of coupled ocean–atmosphere system using attractor radius and global attractor radius

Haoran Zhao¹ · Shaoqing Zhang^{1,2,3,4} · Jianping Li^{1,4} · Youwei Ma¹

Received: 4 June 2020 / Accepted: 10 November 2020 / Published online: 2 January 2021
© The Author(s) 2021

Abstract

While ocean–atmosphere coupled models play an increasingly important role in weather-climate simulation and prediction, the predictability theory based on an atmosphere-only model has significant limitations in interpreting prediction results and guiding predictability studies. Here we use a conceptual ocean–atmosphere coupled model that describes the typical interactions of a synoptic-scale atmosphere with a seasonally-interannually varying upper ocean as well as a deep ocean that varies on decadal timescales to systematically study the predictability of a coupled system. Moving from an atmosphere-only system to an ocean–atmosphere coupled system, the initial-value predictability problem becomes a joint initial-value and boundary-value problem. Although the coupling process increases the uncertainties of the boundary, ocean signals with longer timescales are added to the atmosphere system, thus increasing its predictability. We then investigate the predictability characteristics of the National Centers for Environmental Prediction coupled Climate Forecast System (CFS) and the uncoupled Global Ensemble Forecast System (GEFS). In the coupled CFS system, the practical predictability limit of the lower troposphere is significantly longer than in the uncoupled GEFS due to the contribution of low-frequency boundary signals from air-sea interactions. While further deep and thorough examination is necessary for understanding ocean predictability in the climate system, a preliminary discussion for the predictability of the upper and deep oceans within a coupled ocean–atmosphere framework is also presented in this study.

Keywords Coupled models · The predictability theory · CFS · GEFS · Air-sea interactions

1 Introduction

Thompson (1957) and Lorenz (1963) first proposed the predictability limit of deterministic systems, and since then the predictability of the atmosphere has been extensively studied

using theoretical, numerical, and statistical models (Chou 1989; Simmons and Hollingsworth 2002). In particular, the study of error growth and propagation has become an important part of predictability theory (Dalcher and Kalnay 1987; Farrell 1990; Palmer 2006; Li and Wang 2008). The ocean and atmosphere, as nonlinear dynamical systems, have complex deterministic features that are sensitive to initial values and boundary conditions. It has long been recognized that the upper limit of weather predictability for the synoptic and larger scales is about 2 weeks (Lorenz 1965).

In previous studies, Lorenz (1969) sought to calculate the predictability limit by using the doubling time of an initial infinitesimally small error in the model, but more recent research has shown that such a method has a dependence on the numerical model used (Dalcher and Kalnay 1987; Li et al. 2000). The Lyapunov exponent is a physical quantity to measure the long-term average exponential divergence or convergence of initial nearby orbits in phase space, which can estimate the mean growth rate of the initial infinitesimal error (Oseledec 1968). The nonlinear local Lyapunov

✉ Shaoqing Zhang
szhang@ouc.edu.cn

✉ Jianping Li
ljp@ouc.edu.cn

¹ Key Laboratory of Physical Oceanography, Institute for Advanced Ocean Studies, Frontiers Science Center for Deep Ocean Multispheres and Earth System (FDOMES), Ocean University of China, Qingdao, China

² Pilot National Laboratory for Marine Science and Technology (QNLN), Qingdao, China

³ International Laboratory for High-Resolution Earth System Model and Prediction (iHESP), Qingdao, China

⁴ The College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao, China

exponent (NLLE; Ding and Li 2007; Li et al. 2008; Li and Wang 2008; Li and Ding 2011) was introduced to quantitatively measure the predictability limits of nonlinear systems. Using reanalysis data, local dynamical analogs (LDA) are applied to study the predictability limits of the geopotential height, vector wind field, and sea surface temperature. (SST; Li and Ding 2011, 2013, 2015). However, the NLLE method cannot overcome the uncertainties associated with numerical prediction models. According to the definition of atmospheric attractors, Li et al. (2018) proposed two invariant statistical quantities, the attractor radius (AR) and the global attractor radius (GAR), to define the geometric characteristics, average behavior of a chaotic system and its error growth. Moreover, GAR measures the average distance between two randomly selected states on an attractor, while AR quantifies the average distance of all states on the attractor from the average state. Both the AR and GAR are intrinsic properties of a chaotic system and independent of the forecast model and model errors, and thus provide more accurate, objective threshold to assess the global predictability limits of forecast models. Feng et al. (2019) used the AR and GAR to study the predictability limits of deterministic and ensemble mean forecasts, focusing on quantifying the relationship between forecast errors, especially in individual prediction cases.

The ocean and atmosphere are coupled dynamically and thermodynamically by the exchange of fluxes, including heat, freshwater and momentum fluxes, at the air-sea interface (Webster and Lukas 1992; Saenko et al. 2002). An atmospheric general circulation model (AGCM) simulates the atmosphere using a specified SST, while a coupled general circulation model (CGCM) couples an AGCM and an oceanic general circulation model (OGCM). Some pioneering works (Nese and Dutton 1993; Nese et al. 1996) recognized that the average predictability is significantly improved when an ocean circulation is coupled to the atmosphere. There is large literature discussing predictability of the atmosphere induced by boundary conditions, especially in the tropics (Charney and Shukla 1981; Shukla 1998; Palmer 1994; Reichler and Roads 2003; Bach et al. 2019). Boundary conditions can provide long-term predictability in the tropics, because baroclinic instability is less significant there (Charney and Shukla 1981). Shukla (1998) found that the flow and rainfall of the tropical atmosphere are more deeply influenced by SST than the extratropical atmosphere. Bach et al. (2019) used SST and low-level atmospheric variables to calculate the detailed spatial structure of ocean-to-atmosphere predictability. The ocean improves the atmosphere predictability most significantly in the tropics.

Moving from an AGCM to a CGCM, the initial-value predictability problem becomes a joint initial-value and boundary-value problem, so evaluating the effect of boundary conditions on the predictability limit of the atmosphere

remains a crucial issue. This study attempts to examine quantitatively the practical predictability limits and decorrelation limits of the atmosphere in a CGCM system using the AR and GAR, compared with an AGCM.

The remainder of this paper is organized as follows. Section 2 describes the methodology, including simple conceptual coupled models and AR methods, as well as the CGCM and AGCM reanalysis and prediction data used in this study. Section 3 first systematically examines the predictability of the atmospheric and oceanic variables in simple conceptual coupled models using AR methods, and then derives the theory of the practical predictability limit of the ocean-atmosphere coupled system. Section 4 analyzes the global practical predictability limits of both the coupled Climate Forecast System (CFS) and the uncoupled Global Ensemble Forecast System (GEFS), illustrating the influence of low-frequency signals in a coupled system on the predictability limit. Finally, a summary and discussion are given in Sect. 5.

2 Model, methodology, and data

2.1 Conceptual coupled model

Lorenz (1963) proposed a simple model with only three variables to represent the chaotic characteristics of the atmosphere. The uncoupled Lorenz63 model cannot reflect the coupling processes between the atmosphere and the ocean. Zhang et al. (2011) added a slowly changing variable w , which is coupled with the Lorenz 3-variable (x , y , and z are the high frequency variables of atmosphere) chaotic model to simulate the interaction of the fast system with the slow upper ocean. We call this the 4-variable coupled model (4VCM):

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + (1 + c_1 w)\kappa x - y \\ \dot{z} &= xy - bz \\ O_m \dot{w} &= c_2 y - O_d w + S(t)\end{aligned}\quad (1)$$

At the same time, Zhang (2011) derived a conceptual pycnocline prediction model, which added the deep ocean pycnocline (η) on the 4VCM. We call this a 5-variable coupled model (5VCM):

$$\begin{aligned}\dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= -xz + (1 + c_1 w)\kappa x - y \\ \dot{z} &= xy - bz \\ O_m \dot{w} &= c_2 y + c_3 \eta + c_4 w \eta - O_d w + S_m + S_s \cos(2\pi t/S_{pd}) \\ \Gamma \dot{\eta} &= c_5 w + c_6 w \eta - O_d \eta\end{aligned}\quad (2)$$

where setting the external forcing as $S_m + S_s \cos(2\pi t/S_{pd})$ simulates the constant and seasonal forcing for the “climate”

system. The ‘‘ocean’’ can get the energy from the external forcing when the linear damping $-O_d w$ dissipate energy continuously. The parameters S_{pd} , S_m and S_s define the period of seasonal cycle, the magnitudes of the annual mean and seasonal cycle of the forcing respectively Zhang (2011). described the values of the model parameters in detail. Using a fourth-order Runge–Kutta time stepping scheme with $\Delta t = 0.01$ dimensionless time units (TU), where 1 TU is the time taken for the model to pass through an attractive lobe, the model is specified with parameter values: $(\sigma, \kappa, b, c_1, c_2, O_m, O_d, S_m, S_s, S_{pd}, \Gamma, c_3, c_4, c_5, c_6) = (9.95, 28, 8/3, 10^{-1}, 1, 10, 1, 10, 1, 10, 100, 10^{-2}, 10^{-2}, 1, 10^{-3})$. The model is spun up for an initial 10^4 TU starting from $(x, y, z, w, \eta) = (0, 1, 0, 0, 0)$.

2.2 Attractor radius and global attractor radius

Li et al. (2018) proposed an attractor theory to measure the predictability of nonlinear dynamical systems. Three statistical quantities are used to describe the attractor: Attractor radius (AR), Global attractor radius (GAR) and Global average distance (GAD). First, consider x to be the state column vector on a compact attractor, the average root-mean-square distance between any point in the attractor and the center of the attractor is defined as the attractor radius (AR, R_E) of a compact attractor Θ_1 :

$$R_E = \sqrt{E(\|x - E(x)\|^2)}, \quad x \in \Theta_1, \tag{3}$$

where E is the expectation of the time series and $\|\cdot\|$ is the L^2 -norm of a vector.

Second, the global attractor radius (GAR, R_G) represents the average root-mean-square distance between any two points in the attractor:

$$R_G = \sqrt{E(R_L^2)} = \sqrt{E(\|x - y\|^2)}, \quad x, y \in \Theta_1 \tag{4}$$

where x and y are two randomly selected state vectors from Θ_1 . There is a constant proportional relationship between R_G and R_E of a compact attractor Θ_1 : $R_G = \sqrt{2}R_E$. It is also consistent with the conclusions in Kalnay (2002). Third, the global average distance (GAD) between two attractors can be calculated from the AR:

$$R_G(\Theta_1, \Theta_2) = \sqrt{E(\|x - y\|^2)} = \sqrt{R_E^2(\Theta_1) + R_E^2(\Theta_2) + d^2(x_E, y_E)}, \quad x \in \Theta_1, y \in \Theta_2 \tag{5}$$

where x_E (y_E) is referred to as the center of the compact attractor Θ_1 (Θ_2). And $d(x_E, y_E)$ is defined as the distance between the mean states of Θ_1 and Θ_2 .

While the detailed and exhaustive mathematical derivations can be referred to the aforementioned literature (Li et al. 2018),

thus we mainly comment on the computational implementation. For a real n -dimensional nonlinear dynamical system,

$$dx(t)/dt = F(x(t)) \tag{6}$$

where x and F are n -dimensional nonlinear column vectors. Since the forecast model error can't be ignored, the dynamical system is redefined as $dx(t)/dt = \tilde{F}(x(t))$. Let Θ and Θ_M correspond to the attractor of the real system and its approximate forecast model. Typically, they are different and thus their predictabilities are also different. $x(t)$ and $y(t)$ are fiducial orbits on Θ and Θ_M , respectively. For an imperfect model, there are both initial error and model error in forecasting. And the initial value $y_0 = x_0 + \delta_0$, where δ_0 is a small initial error. Then, we can measure the RMSE between $x(t)$ and $y(t)$,

$$e_M(x_0, \delta_0, t) = \|\delta_M(t)\| = \|y(t) - x(t)\|, \quad x \in \Theta, y \in \Theta_M \tag{7}$$

For global predictability, the global ensemble average of error growth can be defined as.

$$\bar{e}_M(\delta_0, t) = \sqrt{\langle e_M^2(x_0, \delta_0, t) \rangle_{N \rightarrow \infty}} = \sqrt{\int_{\Theta} e_M^2(x(t_0), \delta(t_0), t) dx} \tag{8}$$

where $\langle \rangle_{N \rightarrow \infty}$ denotes the ensemble average of all samples of sufficiently large number N . For the chaotic system, as the integral step increases, $\bar{e}_M(\delta_0, t)$ will converge to the error saturation.

According to the definition of GAD in Eq. (5), it's the expectation of the RMS distance between two different attractors. Hence, the global ensemble RMSE between the real attractor Θ and the model attractor Θ_M tends to the GAD with time. However, the GAD cannot directly be used as an objective threshold to determine the predictability limit, because the forecast model has associated model errors. Following the definition of GAD, to avoid the effect of the model errors, let $R_e = \min(R_E(\Theta_M), R_E(\Theta))$, and $R_g = \min(R_G(\Theta_M), R_G(\Theta))$, then the global practical predictability limits (T_{pr}^M) and decorrelation limits (T_{de}^M) of the forecast model are determined by.

$$\bar{e}_M(\delta_0, t)|_{t < T_{pr}^M} < R_e \text{ and } \bar{e}_M(\delta_0, t)|_{t = T_{pr}^M} = R_e \tag{9}$$

$$\bar{e}_M(\delta_0, t)|_{t < T_{de}^M} < R_g \text{ and } \bar{e}_M(\delta_0, t)|_{t = T_{de}^M} = R_g \tag{10}$$

where T_{pr}^M is the time at which the error reaches the saturation value R_e , and T_{de}^M is the time at which the error reaches the saturation value R_g .

The global ensemble average of error growth (Eq. (8)) tends to a saturation value, which depend on model error. In most studies, the predictability limit has been defined as the time when the forecast error exceeds 95% of the saturation value. This means that the saturation value of forecast error depends on the used numerical model, and thus is not an objective threshold to quantify and compare the predictability limits between different models. In contrast, the AR and GAR is calculated using the ‘perfect’ model as an objective threshold, which only depend on the real attractor described by the ‘perfect’ model and are not affected by model drift. Thus, the AR and GAR are more suitable metrics to measure the predictability limit of a forecast model.

When the forecasting error has surpassed the AR but has not yet reached the GAR, the single model has lost its forecasting skill compared with the climate mean states. Leith (1974) suggested that the error variance of a single deterministic forecast is twice that of a climate prediction in the long-range forecast. The anomaly field is used to calculate a final best estimate of the state of the atmosphere using a linear regression of dynamic forecasting. Hence, the range of prediction error between AR and GAR is still related to predictability. Meanwhile, Feng et al. (2019) pointed out that the forecast error asymptotes to a saturation value of the GAR in a single forecast, while the error saturation level will be the AR in an ensemble forecast. This result is consistent with the conclusions in Leith (1974) and Kalnay (2002). In this work, we only use the single forecast data to study the deterministic forecast.

2.3 Data for the atmosphere-only and coupled model forecasts

In addition to the three simple models, the 45-day reforecasts of the Climate Forecast System (CFS) and the 35-day reforecasts of the Global Ensemble Forecast System (GEFS) are used to study the predictability limit of earth system models. This makes it possible to extend the results from simple conceptual models to operational forecasting models. Details of the CFS and GEFS models are given in Table 1.

The GEFS and CFS data are the medium- and long-range meteorological forecasts produced by NCEP, respectively.

Table 1 Some details about CFS and GEFS

	CFS	GEFS
Model	Coupled	Uncoupled
Frequency	6 h	Once a week
Period	1999.01–2010.12	1999.01–2016.12
Variable	Geopotential Height in 850-hPa, 500-hPa, 200-hPa	
Forecast days	45 days	35 days

CFS is produced by a fully coupled model that represents the interaction between the ocean and atmosphere, while the GEFS is produced by an uncoupled model. We use a common period (1999–2010) of the CFS and GEFS reforecast datasets (note that only the control forecast of the GEFS system is used). The prediction results are verified by using the 00Z initial conditions of NCEP Climate Forecast system reanalysis (CFSR) from Jan 1999 to Dec 2010. When we calculated the AR and GAR using the CFSR, the annual cycle of the time series has been removed.

3 Results of the conceptual coupled model

3.1 Uncertainties in coupled vs. uncoupled systems

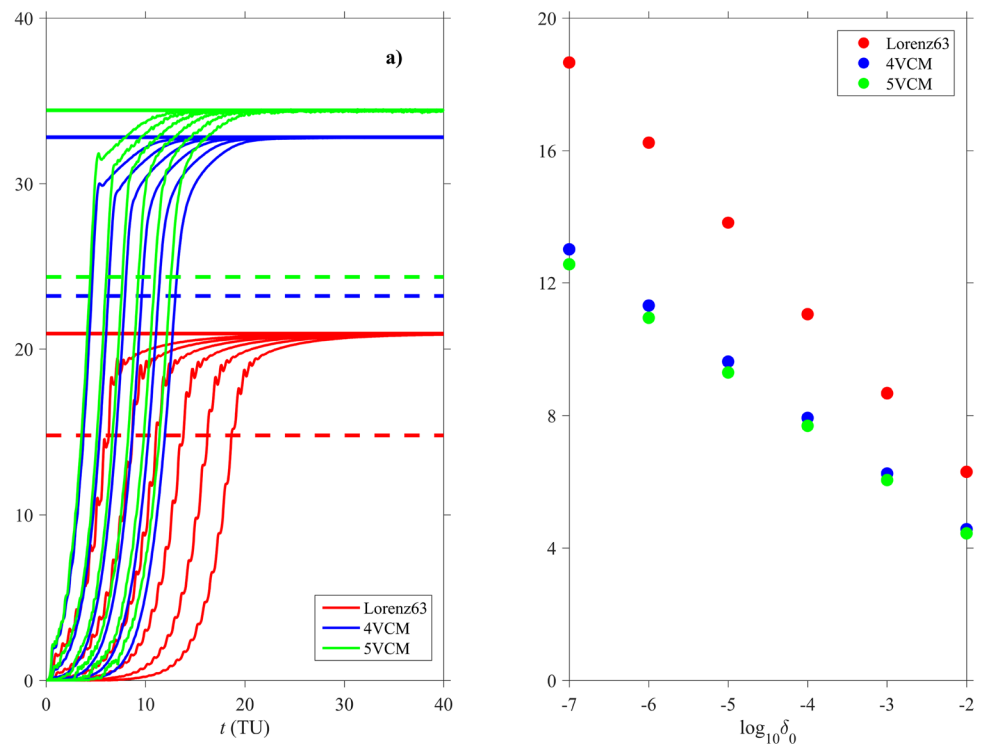
AR being the expected RMSE if the mean state of the attractor is chosen as a predictor, while GAR is the expected RMSE if a random point is chosen. Compared with the standard deviation (SD), AR has the same form to measure the variability of variables. AR (GAR) is objective threshold to determine the practical predictability limits and decorrelation limits of a forecast model. Table 2 lists the AR and GAR for the variables x , y , z and the vector (x, y, z) in the Lorenz63, 4VCM, and 5VCM perfect models. The results reveal that the 5VCM has the largest values of AR and GAR for three atmospheric variables in these models. As more coupling processes are added, the natural variability of the chaotic system gradually increase.

The predictability limit is affected by the magnitude of initial errors (Ding and Li 2007). Figure 1a gives the forecast RMS error statistics of the vector for three atmospheric variables for these three simple conceptual models. Based on the perfect models of Lorenz63, 4VCM, and 5VCM, a set of initial errors with magnitudes 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , and 10^{-7} is generated with white noise superimposed on the atmospheric variables (x, y, z) of the unperturbed run to produce a set of initial conditions (total of 2×10^7). Forecasts for 50 TUs are then carried out. Figure 1b shows the global practical predictability limit (T_{pr}) of these three perfect conceptual models. As the magnitude of the initial errors decreases, the practical predictability limit increases. For a given magnitude of initial errors, it is clear that the practical predictability limit of the 5VCM is shorter than that

Table 2 The AR (GAR) of the variables x , y , z and the vector (x, y, z) in different simple conceptual models

Model	x	y	z	$x-y-z$
Lorenz63	7.92(11.21)	9.01(12.74)	8.62(12.19)	14.78(20.90)
4VCM	11.63(16.45)	14.48(20.47)	13.90(19.65)	23.19(32.80)
5VCM	12.07(17.07)	15.26(21.58)	14.61(20.67)	24.33(34.41)

Fig. 1 a Time series of global ensemble mean RMSE of vector (x, y, z) in different simple models and the magnitude of initial errors. The magnitudes of initial perturbation are 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , and 10^{-7} , from left to right with same color. The solid line and the dashed line represent the corresponding GAR and AR of each model. **b** Global practical predictability limit T_{pr} of those three models with respect to initial error. The red, blue and green colors correspond to the Lorenz63, 4VCM and 5VCM models

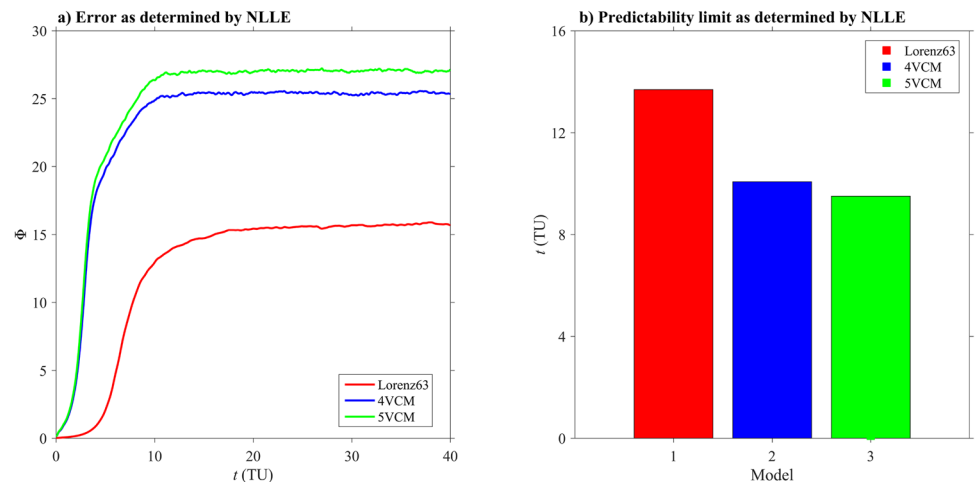


of the 4VCM, and the Lorenz63 model has the maximum predictability limit. In the 4VCM and 5VCM, the initial error superimposed on the atmospheric variable is gradually introduced into the ocean through a forcing term from the ‘atmosphere’ (c_2y). In return, the error growth of the ocean component accelerates the error amplification of the atmosphere component. Finally, the practical predictability limit decreases with the increase of variables.

According to the AR theory, the global practical predictability limit T_{pr} can be regarded as the objective criteria to measure the time when model forecasts have skill. Meehl et al. (2014) point out that predictability is the ability that a system signal can be predicted rather than the ability of

current human being to predict some feature or quantity. Hence, we also use the NLE (Ding and Li 2007; Li and Ding 2011) method to calculate the intrinsic predictability limits of those three perfect conceptual models based on their chaotic error growth. The intrinsic predictability limit is defined as the timescale on which the error reaches 98% of its saturation level (Li and Ding 2011). Figure 2 shows the intrinsic predictability limit of vector (x, y, z) using the NLE method, which is consistent with the T_{pr} in Fig. 1. The 5VCM has the maximum error saturation, while the Lorenz63 has the minimum error saturation. The predictability limits for the Lorenz63, 4VCM, and 5VCM models of the vector (x, y, z) are 13.69, 10.07, and 9.50, respectively, while

Fig. 2 a Time series of the error growth of vector (x, y, z) in different simple models as estimated by NLE. **b** Predictability limit of those three models of Lorenz63, 4VCM and 5VCM. The red, blue and green colors correspond to these three models



the average initial error of those three models are 10^{-2} , 10^{-1} and 10^{-1} .

It is of great significance to objectively measure the forecasting skills of operational forecast systems and provide guidance for their development. As stated by Li et al. (2018), the AR method is more suitable than the NLLE to quantitatively measure the time when model forecasts have skill. And compared with the asymptotic value (AV), the AR method has the advantage that it can objectively use a criterion to measure the practical predictability limit instead of using forecast models that have a strong dependence on model errors. Hence, we will only use the AR method to measure the time when model forecasts have skill in the following research.

3.2 Predictability of coupled vs. uncoupled systems

When an uncoupled model is developed as a coupled model, the initial-value predictability problem becomes a joint initial-value and boundary-value problem. Here we assume that the 5VCM includes ‘perfect’ coupling processes between the ‘atmosphere’ and ‘ocean’ of the ‘real world.’ Let Θ and Θ_M correspond to the attractor of the perfect 5VCM and its imperfect model with missing one or more coupling processes. Then, we’ll measure how well the imperfect model can predict the 5VCM.

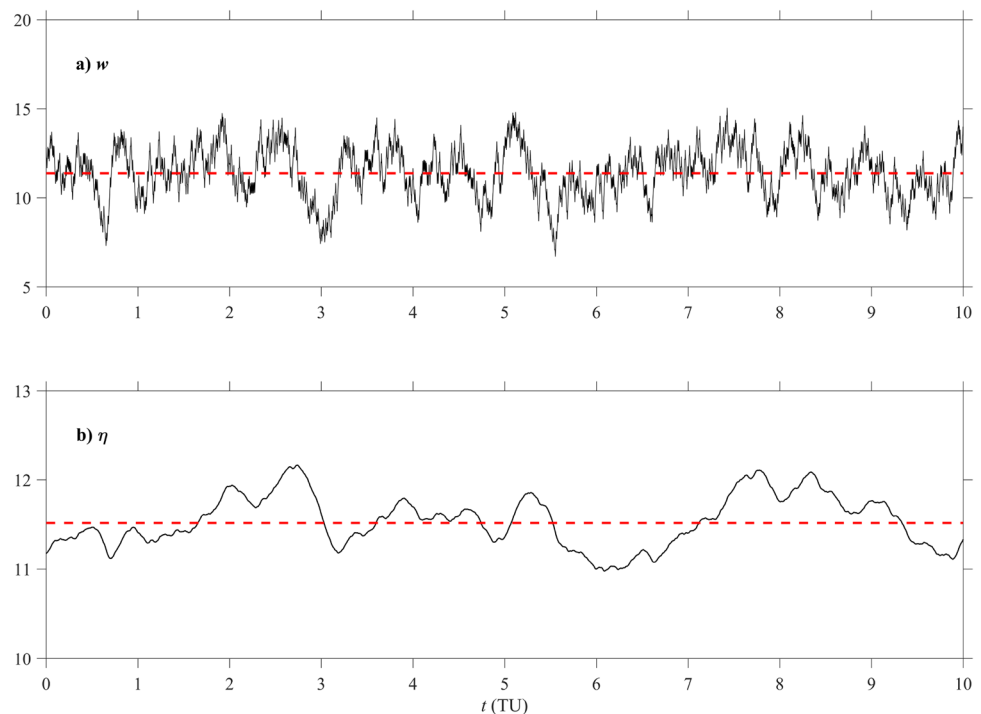
To compare the effects of coupling processes on the predictability limit, several sets of experiments are designed, as listed in Table 3. By changing the equations of both w

Table 3 The control equations of simple models from Case1 to Case7

Model	Equations for low frequency variables	Equations for high frequency variables
Case1	$w = 11.3854$ $\eta = 0$	$\dot{x} = -\sigma x + \sigma y$ $\dot{y} = -xz + (1 + c_1 w)\kappa x - y$ $\dot{z} = xy - bz$
Case2	$\dot{w} = (c_2 y + c_3 \eta + c_4 w \eta + o_d w + S_m + S_s \cos(2\pi t/S_{pd}))$ $\eta = 0$	
Case3	$\dot{w} = (c_2 y + c_3 \eta + c_4 w \eta + o_d w + S_m + S_s \cos(2\pi t/S_{pd}))/O_m$ $\eta = 0$	
Case4	$\dot{w} = (c_2 y + c_3 \eta + c_4 w \eta + o_d w + S_m + S_s \cos(2\pi t/S_{pd}))/O_m$ $\eta = (c_5 w + c_6 w \eta - O_d \eta)\Gamma$	
Case5	$\dot{w} = (c_2 y + c_3 \eta + c_4 w \eta + o_d w + S_m + S_s \cos(2\pi t/S_{pd}))/O_m$ $\eta = 0$	
Case6	$\dot{w} = (c_2 y + c_3 \eta + c_4 w \eta + o_d w + S_m + S_s \cos(2\pi t/S_{pd}))/O_m$ $\eta = c_5 w + c_6 w \eta - O_d \eta/\Gamma$	
Case7	$\dot{w} = (c_2 y + c_3 \eta + c_4 w \eta + o_d w + S_m + S_s \cos(2\pi t/S_{pd}))/O_m$ $\dot{\eta} = c_5 w + c_6 w \eta - O_d \eta/\Gamma$	$\dot{x} = -\sigma x + \sigma y$ $\dot{y} = -xz + (1 + c_1 w)\kappa x - y$ $\dot{z} = xy - bz$

and η in 5VCM, the first four experiments are designed to study the impact of the coupling mechanism on the predictability limit of the vector (x, y, z) . The interaction between the slab ocean and the deep ocean pycnocline is ignored

Fig. 3 Time series of (a) the upper-ocean variable w (black line) and (b) the deep ocean pycnocline depth anomaly η (black line) in the 5VCM. The red dashed line represents the corresponding mean value of each variable



in experiment Case1, where w is regarded as a mean value of the time series ($w = 11.3854$), calculated by the ‘perfect model’ (Fig. 3a). This experiment represents an atmosphere forced by constant bottom boundary conditions in an uncoupled model. In Case2, we retain the interaction between the high-frequency variables and the slab ocean, but neglect the impact of the deep ocean pycnocline. Based on Case2 ($\eta = 0$), the η value is replaced by the climatological mean 11.5196 to form Case3 (Fig. 3b). Case4 is a series of perturbation experiments, with initial errors of magnitudes from 10^{-2} to 10^{-7} superimposed on the initial conditions of the control run (or ‘truth’ run) to calculate the growth of forecast RMSEs, as in Sect. 3.1.

Figure 4a shows the forecast RMS error statistics of the vector for three atmospheric variables between perturbed runs and the control run. The global practical predictability limit (T_{pr}) of the forecast model is shown in Fig. 4b. The first three experiments represent the influence of boundary conditions on the predictability limit. The global practical predictability limits for three atmospheric variables in Case1, Case2, and Case3 are 0.84, 3.73, and 6.43, respectively. As the boundary conditions become more and more accurate, the global practical predictability limit of the nonlinear dynamical system increases significantly. The Case4 experiments reflect the impact of initial errors with different magnitude. From Case4-1 to 4-6, the global practical predictability limits for three atmospheric variables are 4.44, 6.05, 7.68, 9.29, 10.94, and 12.55, respectively. When the magnitude of the initial error is smaller than 10^{-3} , the

predictability limit of the initial-value is longer than that of the boundary-value.

The last three experiments (Case5, Case6, and Case7) in Table 3 are designed to study the predictability of the slab ocean w . In Case5 and Case6, the equations of high-frequency variables are different from the ‘real model’ since we set $c_1 w k x = 0$, meaning that the atmosphere is not affected by the bottom boundary condition. Consequently, the ‘ocean’ components in these two experiments look like an ocean-only model which has forcing from the atmosphere but without feedback to the atmosphere. In Case5, we ignore the influence of the deep ocean pycnocline by setting $\eta = 0$, while the interaction between the slab ocean and the deep ocean pycnocline is retained in Case6. Case7 is designed as the coupled model to study the initial-value predictability for w , just as in Case4.

Figure 5a shows the forecast RMS error statistics of w for the imperfect model relative to the ‘real model’, and the global practical predictability limit of the forecast model is shown in Fig. 5b. The global practical predictability limit of the variable w is 1.92 in Case5 and 1.94 in Case6. Compared with the Case7, the predictability limit of the initial-value problem is longer than that of the boundary-value problem when the upper boundary forcing is closed.

From an atmosphere-only system to an ocean–atmosphere coupled system, the initial-value predictability problem becomes a joint initial-value and boundary-value. Both boundary conditions and initial conditions can affect the predictability of the system. Improving the accuracy of the

Fig. 4 **a** Time series of global ensemble mean RMSE of vector (x, y, z) between the 5VCM ‘real model’ and imperfect model. The yellow, green, red, and blue lines correspond to the Case1, Case2, Case3, and Case4. Especially the magnitudes of initial perturbation are 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , and 10^{-7} , from left to right with blue color. The solid line and the dashed line represent the corresponding GAR and AR of the 5VCM ‘real model’. **b** Global practical predictability limit T_{pr} of the vector (x, y, z) with respect to model errors or initial errors. The Case1, Case2 and Case3 represent the influence of model errors, while the Case4 reflect the impact of initial errors. The magnitudes of initial perturbation are 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , and 10^{-7} from Case4-1 to Case4-6

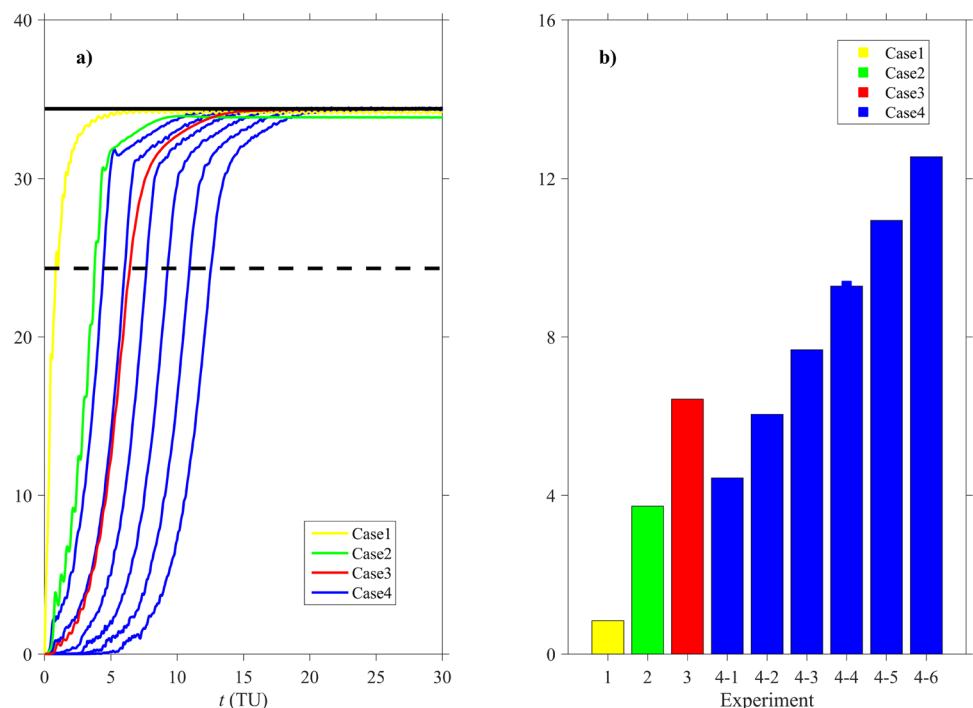
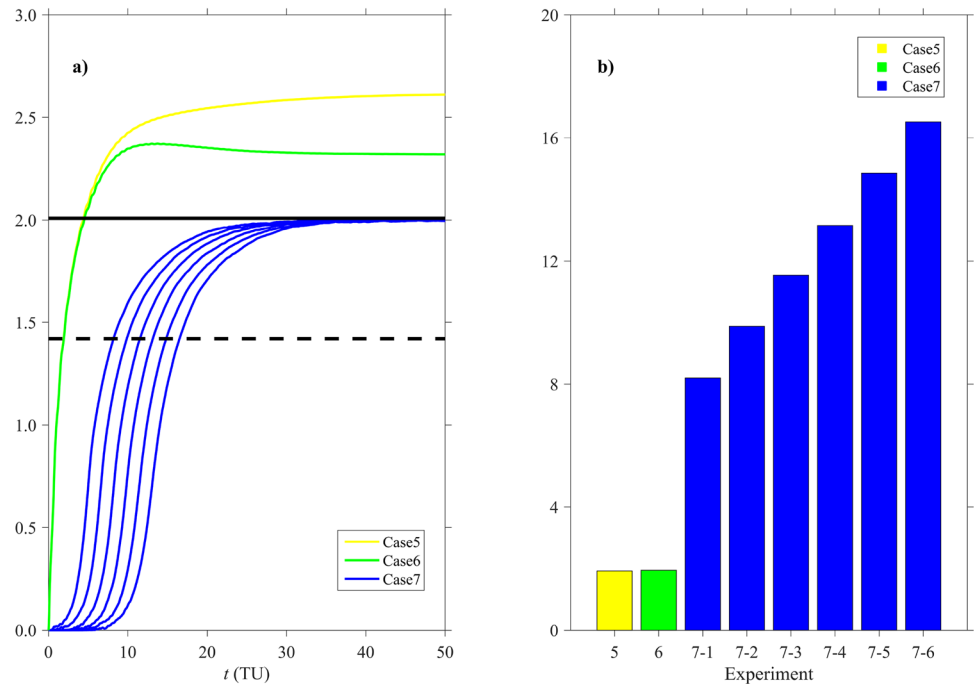


Fig. 5 Same as in Fig. 4, but for the variable w in the Case5, Case6, and Case7. The yellow, green, and blue lines correspond to the Case5, Case6, and Case7. The magnitudes of initial perturbation are 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , and 10^{-7} from Case7-1 to Case7-6



boundary conditions and reducing the initial errors of the models are key to improving forecasting skills in the future.

4 Results of CGCM predictions system

4.1 Practical predictability limit of CFS vs. GEFS

From the results of simple conceptual models, we learned that the predictability limit of coupled models is longer than that of uncoupled models. However, we still need to examine operational forecasting systems to further validate conclusions and understand the mechanisms. We first compare the CFS and GEFS data to verify the conclusions and use CFSR data to verify the forecasts and calculate the AR. According to the AR theory, first, we assume that the "true" state is described by the NCEP CFSR. Then the prediction results are verified by using CFSR to get the global mean RMSE. In the following, the AR and GAR are calculated using the CFSR. And the AR is an objective threshold to determine the practical limits of different models, where AR is independent of the model errors and initial errors.

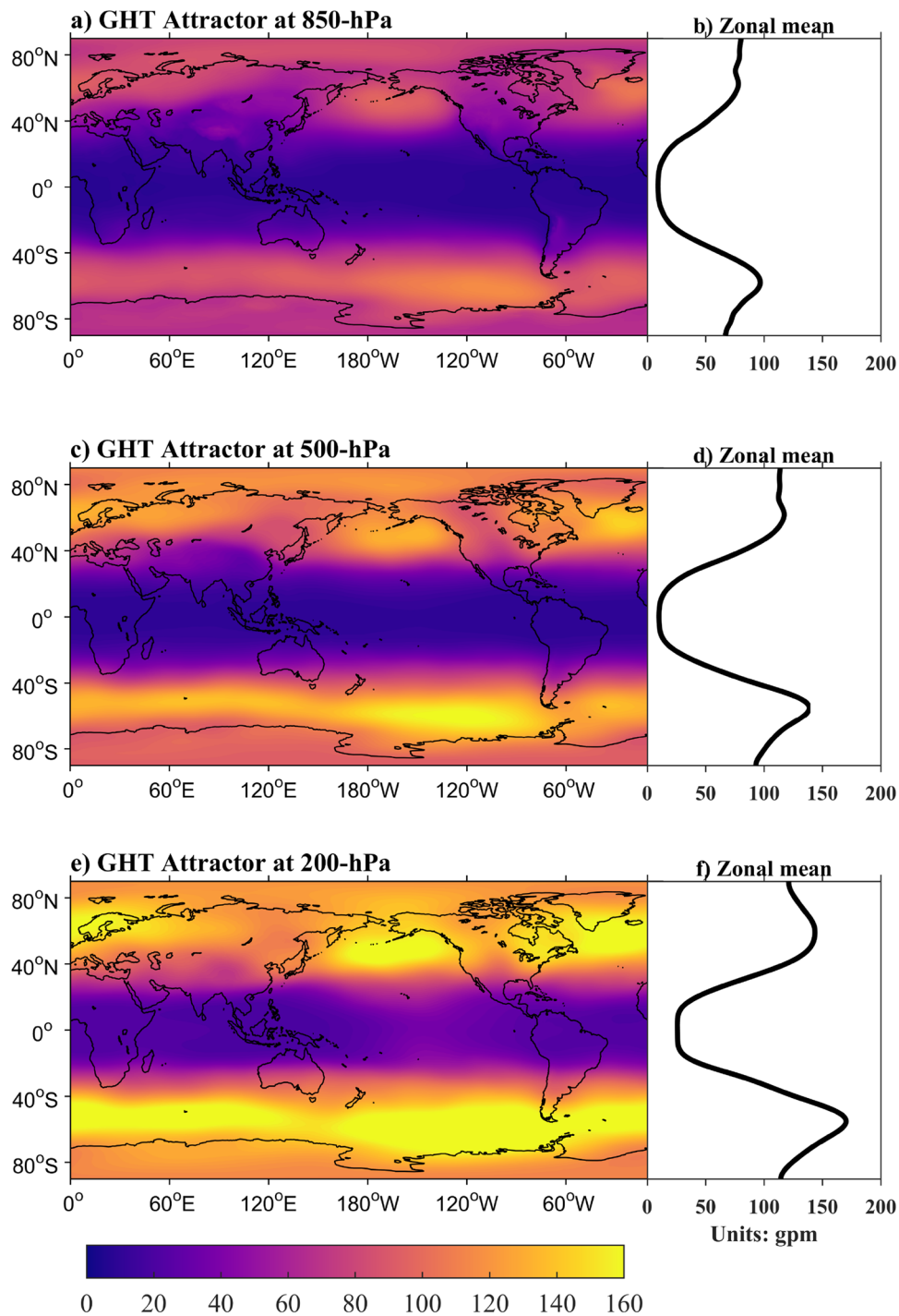
Figure 6 illustrates the spatial distribution of the AR of the geopotential height (GHT) at 850-hPa, 500-hPa, and 200-hPa from Jan 1999 to Dec 2010. The AR of CFSR appears to have a regional characteristic on various pressure levels, with a maximum over mid-latitudes in the Southern Hemisphere (SH), and the lowest values over the tropics. Meanwhile, the spatial distribution of the AR is consistent with the trough and ridge positions in the Northern Hemisphere (NH). The centers of the maximum AR are observed

at the positions of the East Asian trough, the North American trough, and the European trough. Baroclinic instability is generally negligible in the tropics, where barotropic and convective instabilities and their interactions are dominant. In the midlatitudes, baroclinicity is the dominant instability responsible for the growth of small errors at the synoptic (weather) scales. The existence of storm tracks over the midlatitudes explains the higher variability in these regions.

Theoretically, the model attractor radius should be calculated using the model long free forecasts. But we cannot get sufficiently long free forecast time-series of the model. Considering that the reforecast frequency of GEFS is once a week, we use the first 7 days of each forecast case to form a continuous time series as the model long free forecasts to calculate the model attractor radius. And the model attractor radius for CFS is calculated similarly. The model attractor radius of CFS and GEFS are greater than the real AR in most areas on these three levels (not shown).

Figure 7 shows the spatial distribution of the mean practical predictability limit of CFS and GEFS at three pressure layers. In Fig. 7a, the practical predictability limits of the CFS at 850 hPa appear to have a regional characteristic; there are many maximum or minimum centers, with a maximum of 11–13 days over the Western Pacific Warm Pool area and a lowest limit of 2–3 days in the Andes, the Tibetan Plateau, and the Sahara. However, the practical predictability limits of the GEFS at 850 hPa have a significant difference (Fig. 7b), with 9–11 days over the North Pacific and of 3–7 days over the tropics. The zonal mean variation of practical predictability limits with latitude shows that the

Fig. 6 Spatial distribution of (left) the AR of the GHT from Jan 1999 to Dec 2010 and (right) their zonal mean profiles. Upper, middle, and lower panels are for 850-hPa, 500-hPa, and 200-hPa, respectively



practical predictability limits of CFS over the tropics are significantly longer than those of GEFS (Fig. 7c).

As shown in Fig. 7d, the practical predictability limits of 500 hPa GHT in CFS appear to have a zonal distribution. The practical predictability limits of 500 hPa GHT in CFS appear to be highest over the NH midlatitudes. However, relatively low predictability of GHT is observed over the tropics between 10°S and 10°N. In contrast, the practical predictability limits of the GEFS have a significant difference

(Fig. 7e), with 10–12 days over the tropics, especially in the central equatorial Pacific. Over the NH mid–high latitudes, the practical predictability limits can reach 9–11 days. Predictability limits of geopotential height on 200 hPa are markedly longer than those on 500 hPa over the middle-east equatorial Pacific Ocean.

Figure 8 shows the difference in the practical predictability limit of GHT between CFS and GEFS. CFS has longer practical predictability limit than GEFS at 850 hPa

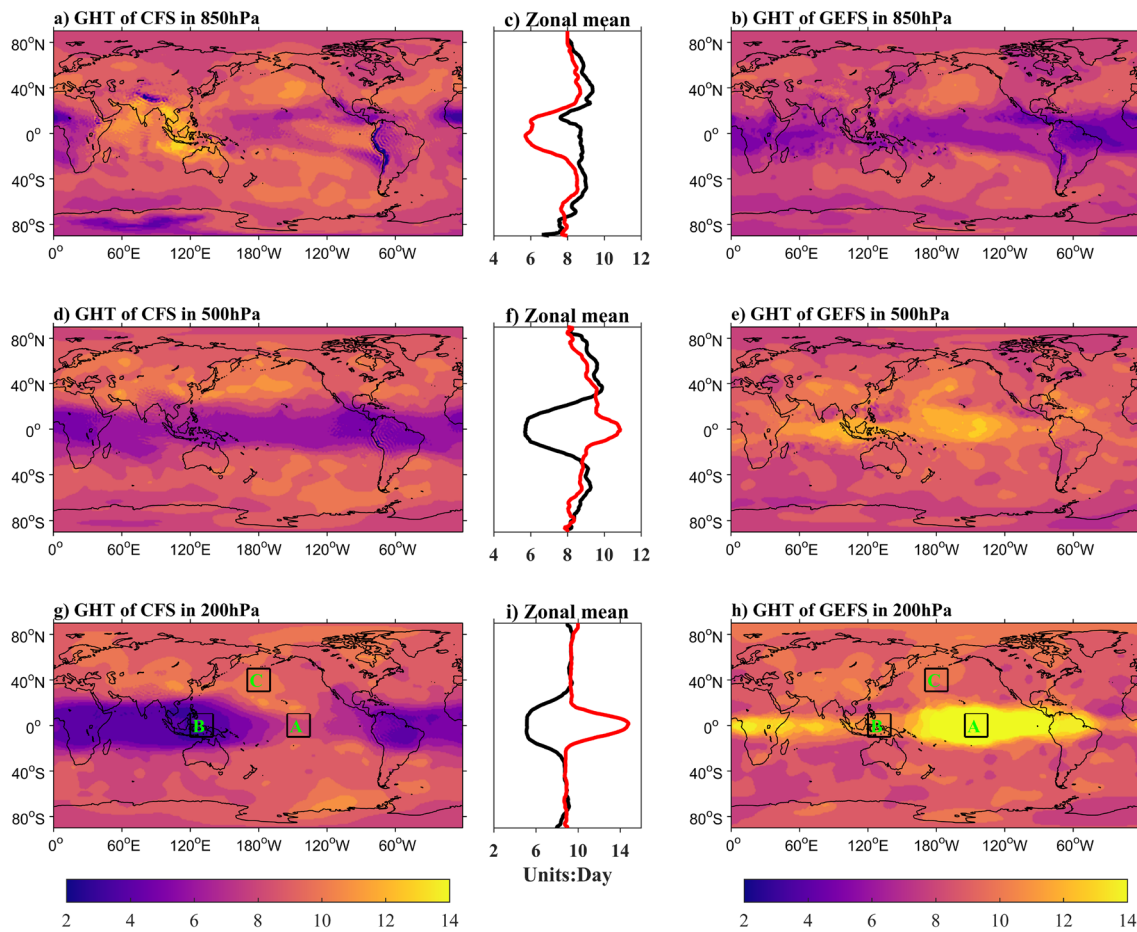


Fig. 7 Spatial distribution of the practical predictability limit of GHT for CFS (left panel) and GEFS (right panel). The black and red lines in the middle panels of zonal mean profile are practical predictability

limit of CFS and GEFS, respectively. Upper, middle, and lower panels are for 850-hPa, 500-hPa, and 200-hPa, respectively

over the tropics (Fig. 8a), while GEFS is more predictable than CFS at 500 hPa and 200 hPa (Fig. 8b, c). From 500 to 200 hPa, the difference between GEFS and CFS increased significantly in the tropics, especially over the middle-east equatorial Pacific Ocean. However, coupling process has a limited effect on the practical predictability limit in mid-high latitudes.

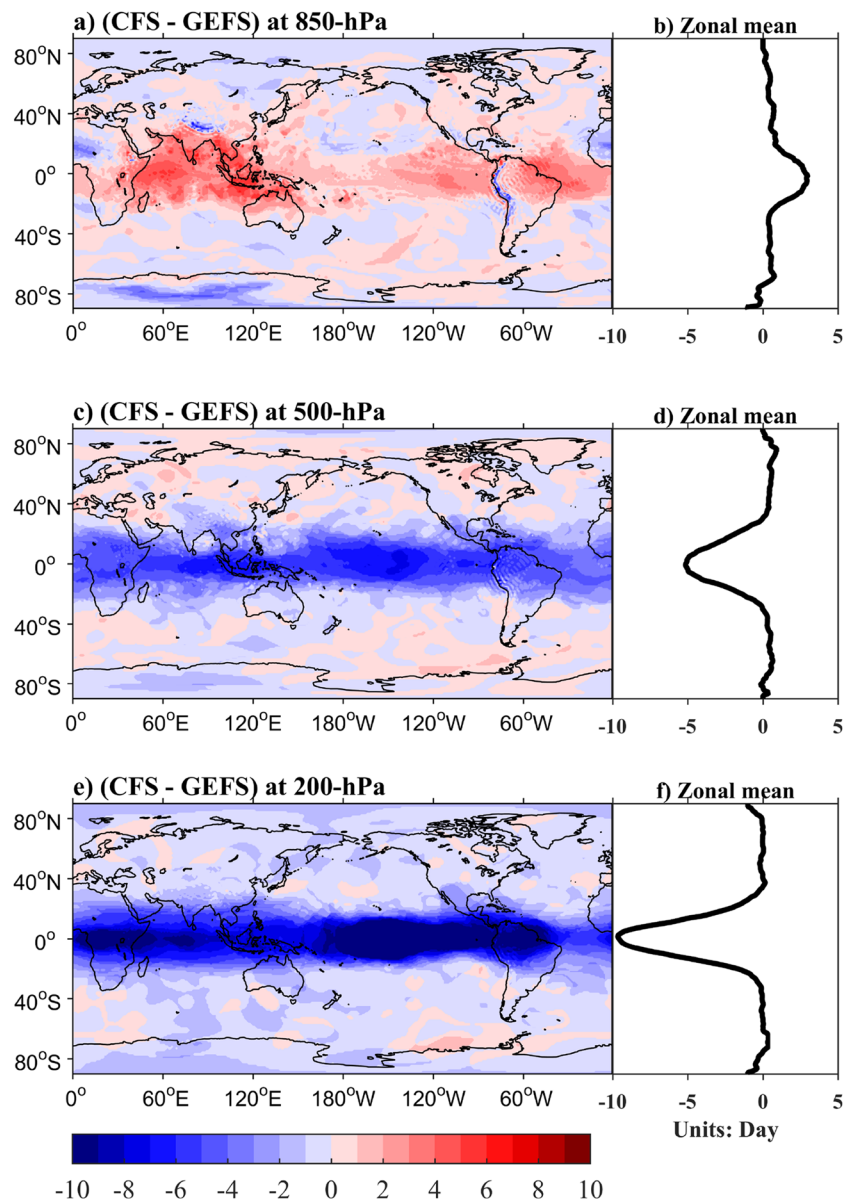
The spatial distribution of practical predictability limits on 200 hPa varies with the season (Fig. 9). In the boreal spring, the practical predictability limit of both GEFS and CFS in the Northern Hemisphere is significantly longer than in the southern hemisphere. The situation in the boreal autumn is the reverse of that in the boreal spring. For those four seasons, there is little difference in the practical predictability limit of the zonal mean between CFS and GEFS over the mid-high latitudes. But GEFS is more predictable in the tropics than CFS, especially in boreal spring. In the boreal spring, the 200-hPa practical predictability limits over the middle-east equatorial Pacific Ocean of GEFS can be more than 22 days. And the seasonal difference of the practical

predictability limits between CFS and GEFS are shown in Fig. 10. The practical predictability limits of GEFS have similar distributions to those of CFS, but with much higher values over the tropics, especially over the middle-east equatorial Pacific Ocean. It is worth noting that the difference in the practical predictability between two models over the tropics varies with the seasons: increasing in the winter and reaching its peak in spring, then decreasing and achieving its minimum in autumn finally.

As shown in Figs. 7 and 9, the CFS has longer practical predictability limit than GEFS at 850 hPa over the tropics, while it is opposite at higher levels. The tropical oceans are crucial for large-scale ocean–atmosphere interactions, and have a profound impact on global climate variability. As the bottom boundary conditions of the atmosphere, the low-frequency signals generated by the ocean–atmosphere interaction can extend the predictability limits.

Previous studies have shown that the ocean has an important effect on the predictability of the lower atmosphere (Bach et al. 2019; Kumar et al. 2013; Peña et al.

Fig. 8 Difference (CFS–GEFS) in the practical predictability limit of GHT (left panels) between CFS and GEFS and their zonal mean profiles (right panels). Upper, middle, and lower panels are for 850-hPa, 500-hPa, and 200-hPa, respectively



2003; Wu and Kirtman 2007). Bach et al. (2019) found that the ocean improves predictions of the low-level atmosphere most significantly in the tropics, which is consistent with our results on 850 hPa. As Pegion et al. (2019) pointed out, the uncoupled atmosphere model could have higher predictability than a coupled atmosphere model with the ocean if the atmosphere model itself has a larger model error. They compared the 2-m temperature and precipitation with five coupled models and two uncoupled models using sub-seasonal prediction experiments. Their results show that the prediction skill of 2-m temperature of GEFS is better than that of all other models. Thus, we believe that there are two reasons for the difference in the predictability between CFS and GEFS over the tropics:

coupling processes and atmosphere model. We will discuss these two points in details next.

First, the CFS is a coupled model. A complete reforecast of CFS have been made with the T126L64 GFS with half-hourly coupling to the ocean (MOM4 at $1/4^\circ$ equatorial, $1/2^\circ$ global) (<https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/climate-forecast-system-version2-cfsv2>). But GEFS is an uncoupled model. In the Environmental Modeling Center (EMC)-GEFS forecast system, SSTs are specified by relaxing the SST analysis to a combination of climatological SST and bias-corrected SST from operational NCEP-CFSv2 forecasts. The longer the lead time, the more weight given to the bias-corrected NCEP-CFSv2 forecast SST (Pegion et al. 2019).

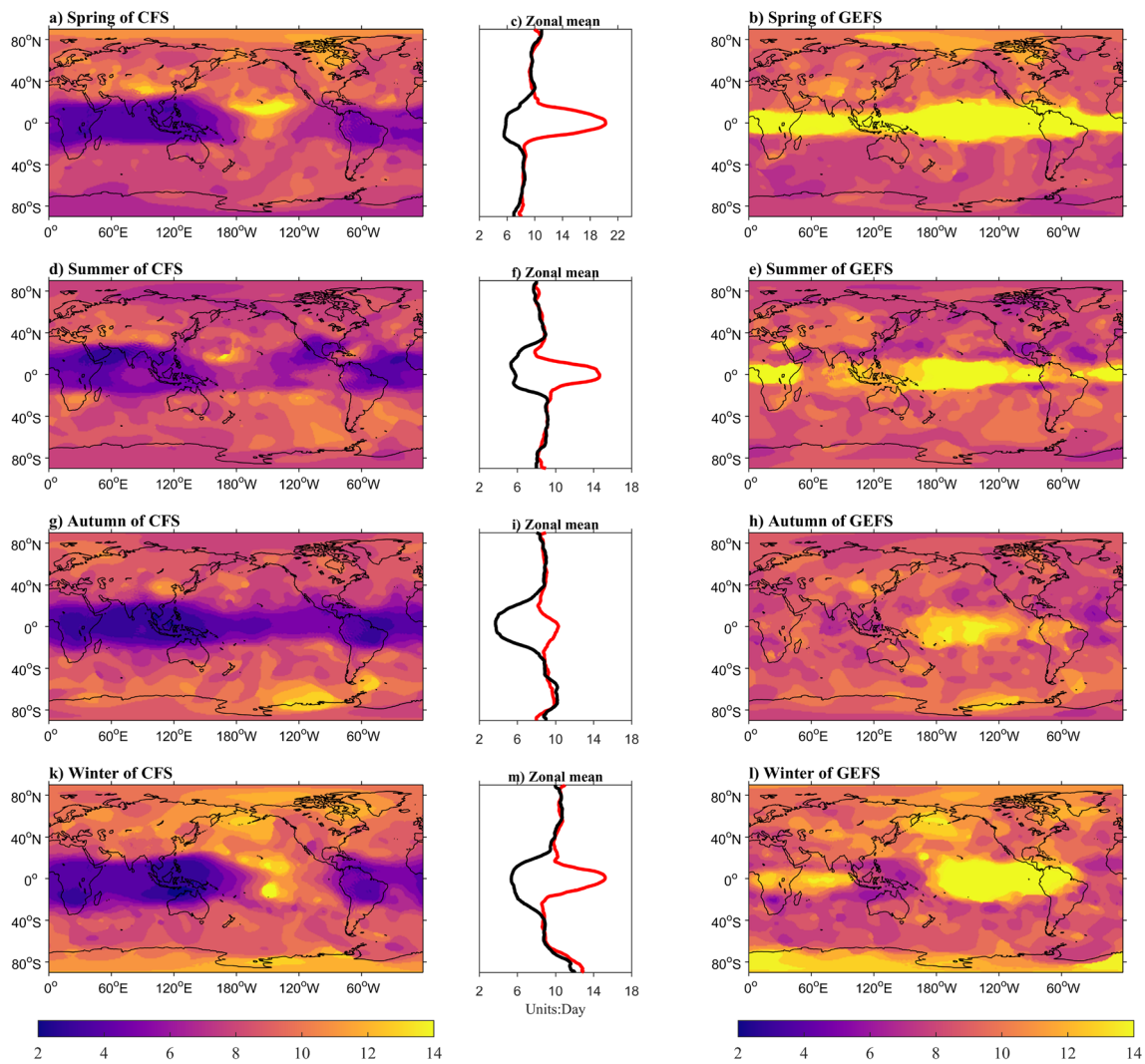


Fig. 9 Spatial distribution of the seasonal mean practical predictability limit of GHT at 200-hPa for CFS (left panel) and GEFS (right panel). The black and red lines in the middle panels of zonal mean

profile are seasonal mean practical predictability limit of CFS and GEFS, respectively. First, second, third, and lower panels are for spring, summer, autumn and winter, respectively

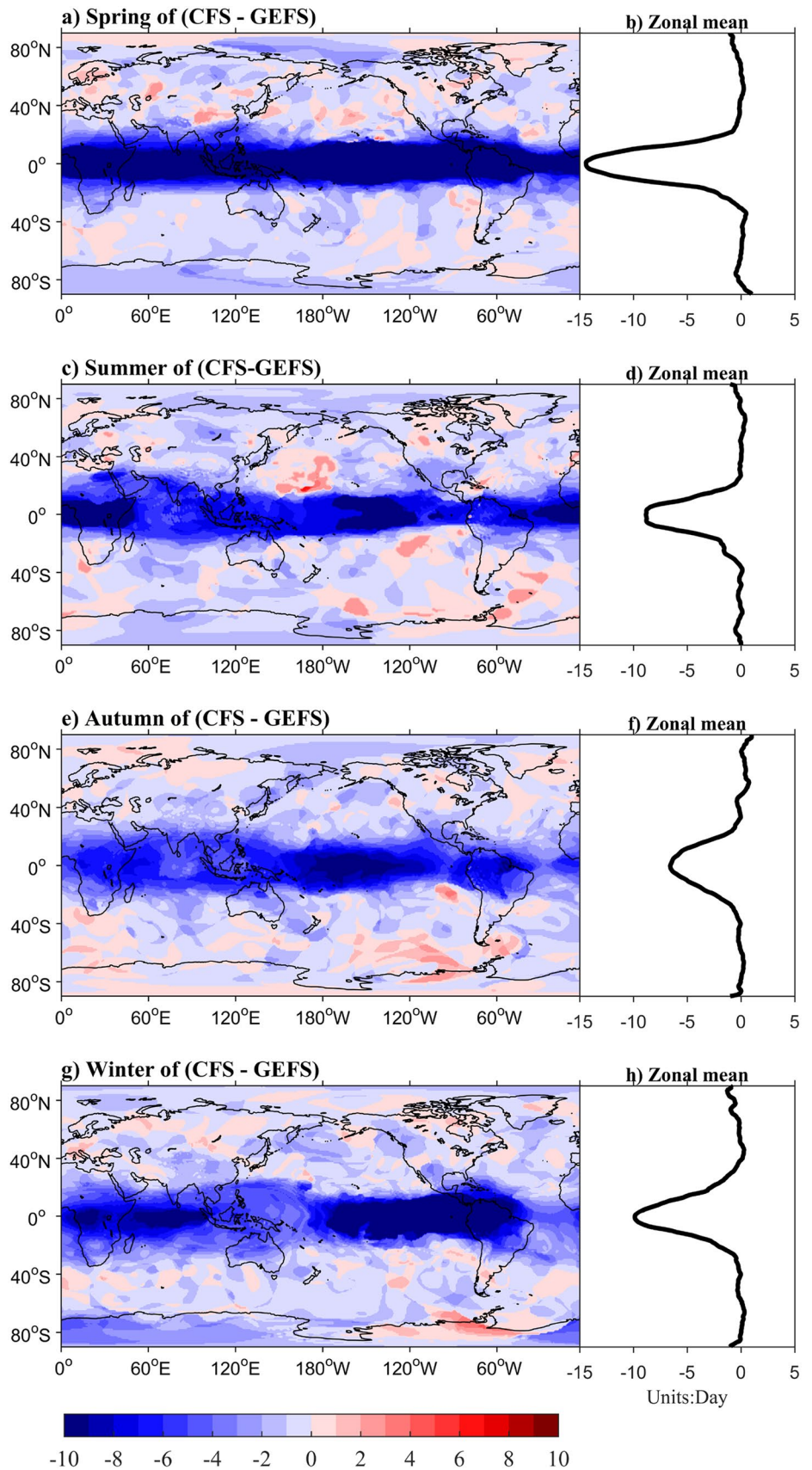
Second, the GEFS atmosphere model has better performance than the CFS atmosphere model. The 35 days reforecasts of GEFS used the new operational Global Ensemble Forecast System version (GEFS; version11). In contrast to the atmospheric component of the CFS, there are several advantages in version11 (Zhou et al. 2017), including: (1) improved initial perturbations using an ensemble Kalman filter (EnKF) data assimilation system; (2) increased horizontal resolution to TL574(34 km)L64 in the first 192 h for weather, to better represent the interaction around multiple scales; (3) Eulerian dynamics is replaced with two time-level semi-implicit semi-Lagrangian dynamic Zhou et al. (2017). found that the RMSE of 500-hPa geopotential height in the new GEFS version (v11) is significantly smaller in the Northern Hemisphere than the old version.

Although the new GEFS version (v11) outperformed the T126L64 GFS in the geopotential height, the coupling process in the lower atmosphere results in a higher predictability limit of CFS than GEFS, especially in equatorial regions. However, in the middle and upper atmosphere, the coupling process has a smaller impact on the atmosphere, while the atmosphere model errors play an important role in the forecasting. As a result, at 200 hPa and 500 hPa, the practical predictability limit in GEFS over the tropics is longer than that of CFS.

4.2 Influence of low-frequency signals on predictability limit

The AR and GAR are introduced to quantitatively estimate the predictability limit of strange attractors. According to the attractor theory, error growth curves will converge to

Fig. 10 Difference (CFS–GEFS) in the practical predictability limit of GHT (left panels) between CFS and GEFS and their zonal mean profiles (right panels). First, second, third, and lower panels are for spring, summer, autumn and winter, respectively



the GAR. Over the 200-hPa central-east equatorial Pacific of the GEFS, GAR is obviously larger than the error saturation, which leads to the long predictability limits in this region. What causes the contradiction between this result and attractor theory? Compared with the atmosphere, the ocean varies slowly. The ocean can suppress the fast-varying signals and enhance the low-frequency signals in the climate system through air–sea interaction, thus lengthening the predictability limit. The ocean influence on atmospheric components may cause the changes in the properties of the attractors. We will use 5VCM to explore the mechanism.

For atmospheric components x , y , and z in the 5VCM, we add a low-frequency term in the x equation to simulate the effects of the periodic signals, so the conceptual pycnocline prediction model becomes

$$\begin{aligned} \dot{x} &= -\sigma x + \sigma y + T_l(t) \\ \dot{y} &= -xz + (1 + c_1 w)\kappa x - y \\ \dot{z} &= xy - bz \\ O_m \dot{w} &= c_2 y + c_3 \eta + c_4 w \eta - O_d w + S(t) \\ \Gamma \dot{\eta} &= c_5 w + c_6 w \eta - O_d \eta \end{aligned} \tag{10}$$

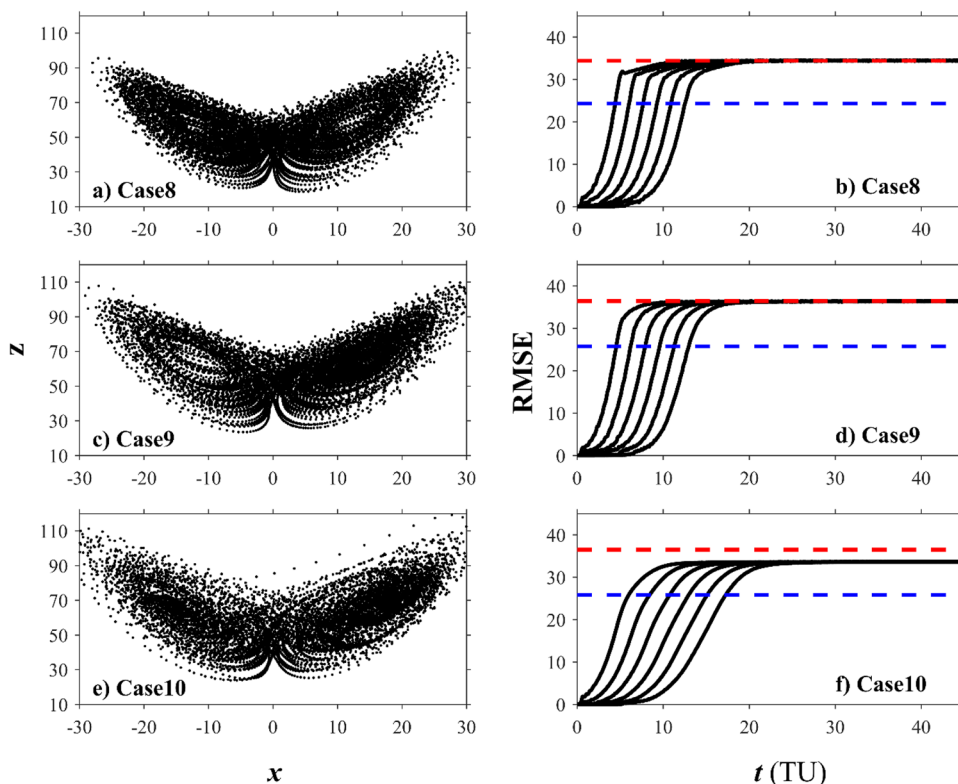
We simply set $T_l(t) = T_0 + A \sin(2\pi t/T_{pd})$ to simulate a low-frequency forcing, where T_{pd} and T_0 are set as 10 and 20, respectively. To compare the effects of signal magnitudes, we run three experiments: the magnitude of A is set to 10 in Case9 and 70 in Case10, while Case8 is defined as

the control experiment, in which atmospheric components are not affected by the low-frequency signal.

For these three experiments, projections on the x - z plane in the phase space of the 5VCM are presented in Fig. 11. In Case8 the atmospheric components are not affected by the low-frequency signals and the attractors are distributed symmetrically on the x - z plane. By adding a low-frequency forcing to the variable x , the spatial distribution of attractor described by Case10 is significantly different from that of Case8 (Fig. 9a, e). Figures 11b,d,f give the forecast RMS error statistics of the vector for three atmospheric variables in these three experiments. The GAR in Case10 is larger than the error saturation, unlike in Case8 and Case9. This phenomenon is consistent with the situation in the GEFS over the middle-east equatorial Pacific. Compared with the results of the standard 5VCM, the low-frequency forcing of the chaotic system has an important effect on the properties of the atmospheric attractors. The addition of an external periodic forcing causes long-term trajectories to be closer than any two random points on the attractor (i.e., for the GAR to be higher than the error saturation), since the forcing can cause the trajectories to experience similar ‘weather’.

The distribution of the practical predictability limits of GHT at 200 hPa shows that both the CFS and GEFS have the longest predictability limit over the middle-east equatorial Pacific compared with other regions in the tropics. This phenomenon varies seasonally, and the practical predictability limits in spring show larger spatial variability

Fig. 11 Projections (a) Case8, (c) Case9, and (e) Case10 on the x - z plane in the phase space of the 5VCM. 1×10^5 points are plotted. Time series of global ensemble mean RMSE of vector (x , y , z) in different experiments (b) Case8, (d) Case9, and (f) Case10 in the 5VCM with respect to the magnitude of initial error. The magnitudes of initial perturbation are 10^{-2} , 10^{-3} , 10^{-4} , 10^{-5} , 10^{-6} , and 10^{-7} , from left to right. Blue and red dashed lines are the AR and the GAR



than in autumn. Compared with other tropical regions, the middle-east equatorial Pacific Ocean has very long predictability limits, but these long predictability limits are not seen over the midlatitudes. We use the Ensemble Empirical Mode Decomposition (EEMD) to investigate this phenomenon further. We analyze the effect of significant low-frequency signals for three regions: A (155°W–135°W, 10°S–10°N), B (120°E–140°E, 10°S–10°N), and C (170°E–170°W, 30°N–50°N), and find the significant low-frequency signals lengthen the predictability limits. Those three regions have been drawn in Fig. 7g, h.

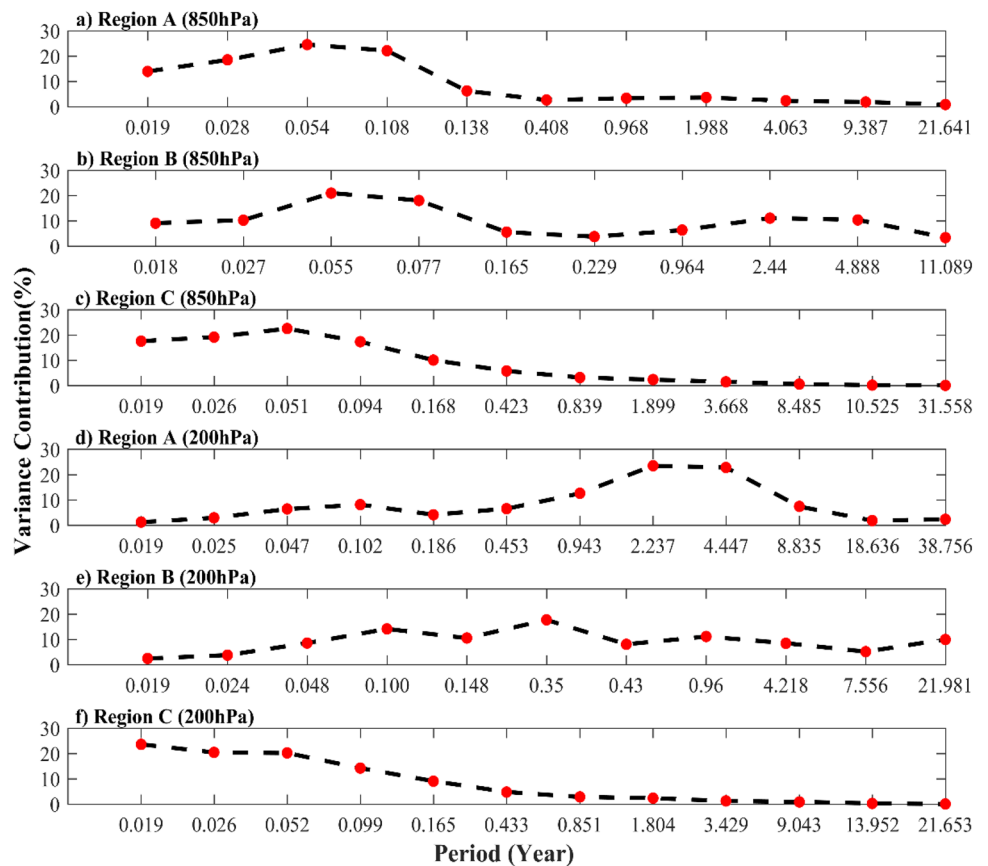
The Ensemble Empirical Mode Decomposition (EEMD; Wu and Huang 2008) is a method of separating data into different components by their scales. Each component is defined as an intrinsic mode function (IMF). Compared with the Empirical Mode Decomposition (EMD), it overcome the scale mixing problem through a new noise-assisted data analysis method. It can be seen in Fig. 7 that the AR and GAR are calculated by CFSR from Jan 1999 to Dec 2010. However, in order to calculate the significant periods, we use both the CFSR (Jan 1979 to May 2011) and CFSv2 (Apr 2011 to Dec 2017).

Figure 12 shows the significant period calculated by the EEMD method. At the 200 hPa level, the low-frequency signals in regions A and B play a dominant role compared

with region C, especially those in region A. The quasi-biennial oscillation (QBO) and quasi-4-year period contribute a greater variance than the seasonal oscillation. In contrast, the quasi-2-week period is the most significant period in region C. However, in the lower troposphere the main periods in the three regions are all intra-seasonal oscillations (ISO), where the largest variance contribution is greater than 20%.

To determine the influence of low-frequency signals on the predictability limits, a high-pass Butterworth filter of order 1 is used to filter signals longer than two years. A Butterworth filter can extract the significant period of a timeseries and obtain its main modes. Murakami (1979) first used the band-pass Butterworth filter to study the Madden–Julian oscillation (MJO), which is the largest element of intraseasonal (30–90 day) variability in the tropical atmosphere. Since then, it has been widely used in meteorology. Here, we define AR_2 as the attractor radius obtained by filtered data, while AR_1 is the attractor radius calculated with the original data, shown in Fig. 13. There is a very small difference between AR_1 and AR_2 for region C at 200 hPa and 850 hPa, while AR_2 is significantly smaller than AR_1 over regions A and B at 200 hPa. The results indicate that the low frequency signals in the atmosphere have an important influence on the attractor radius. When we use AR_2 as a criterion to measure the global practical predictability limits,

Fig. 12 Periods through 95% significance level, obtained using EEMD method. The periods of mean GHT from Jan 1979 to Dec 2017 for CFSR averaged region A (155°W–135°W, 10°S–10°N) in (a) 850-hPa and (d) 200-hPa, region B (120°E–140°E, 10°S–10°N) in (b) 850-hPa and (e) 200-hPa, and region C (170°E–170°W, 30°N–50°N) in (c) 850-hPa and (f) 200-hPa



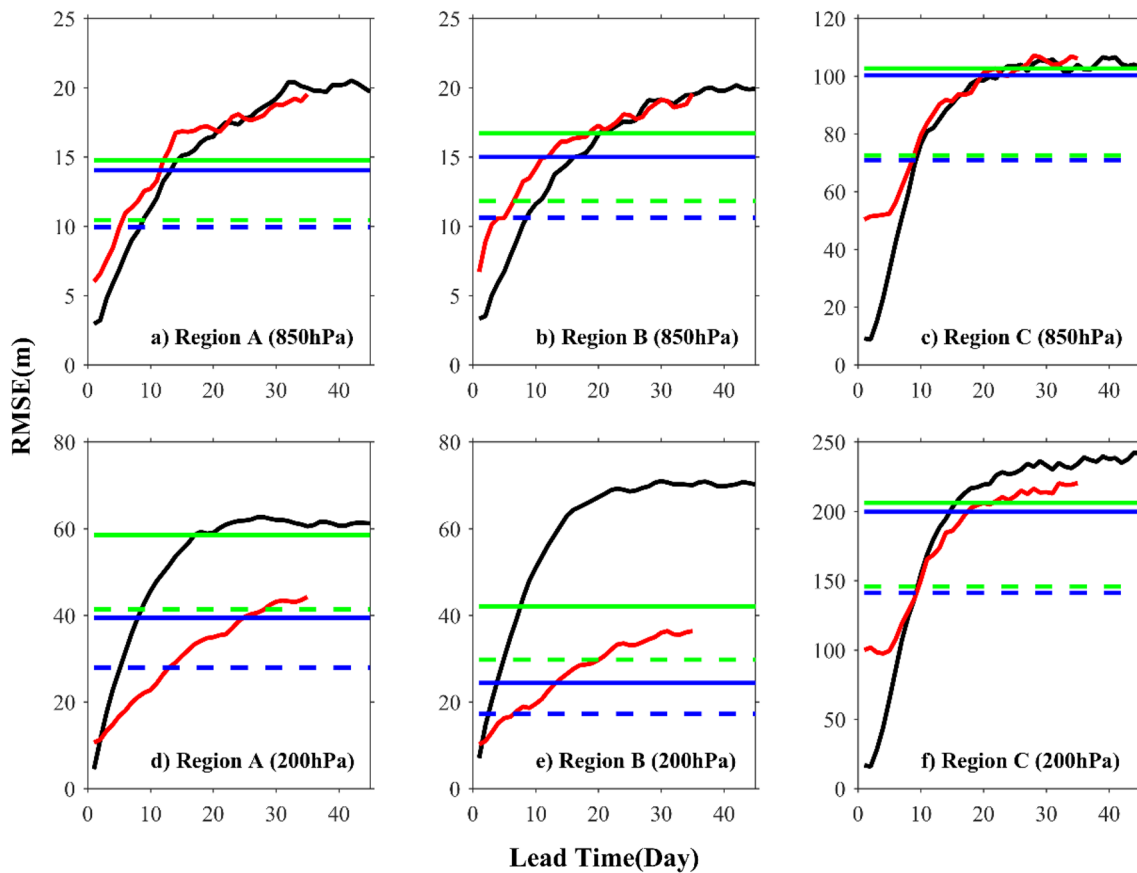


Fig. 13 Time series of mean RMSE from Jan 1999 to Dec 2010 of GHT at 850-hPa and 200-hPa for CFS and GEFS operational forecast data averaged the Region A (155°W–135°W, 10°S–10°N) (left panel), the Region B (120°E–140°E, 10°S–10°N) (middle panel), and the Region C (170°E–170°W, 30°N–50°N) (right panel). Upper and lower panels are for GHT at 850-hPa and 200-hPa, respectively.

Black and red solid lines respectively represent the RMSE of the CFS and GEFS. Green dashed lines are AR_1 calculated by CFSR, while blue lines are AR_2 where the signals over 2 years are filtered by high pass Butterworth filter of order 1. Green and blue solid lines are GAR_1 and GAR_2

both CFS and GEFS have significantly shorter predictability limits over regions A and B in 200 hPa. Comparing regions of A, B, and C, we find that the low-frequency signals have a greater impact on the global practical predictability limits at 200 hPa over the tropics than those over the mid-latitude regions, especially for GEFS.

Some previous studies have shown that the variability and predictability of the 200-hPa seasonal mean GHT can be divided into external and internal variability (Kumar et al. 2003). The seasonality of internal variability depends on the seasonality of the atmospheric mean state alone, while the seasonality in the external variability depends further on the seasonality of the SST forcing, especially the tropical SST variability associated with the El Niño–Southern Oscillation (ENSO) (Hoerling and Kumar 2002). In addition, Baldwin et al. (2003) reported that persistent circulation anomalies in the lowermost stratosphere could affect the troposphere through changes in waves in the upper troposphere. Abnormal signals in the stratosphere can travel down to the

troposphere and influence its weather and climate. Because the general circulation of the stratosphere has a long time-scale, it may enhance the forecast skill in the troposphere.

5 Summary and discussion

We compared the difference in the global practical predictability limits of coupled and uncoupled models, using the attractor radius (AR), the global attractor radius (GAR), and the nonlinear local Lyapunov exponent (NLLE). We first used three simple models to systematically study the characteristics of predictability of a coupled system. Starting from the Lorenz butterfly model (Lorenz63), a slab ‘ocean’ variable and a ‘pycnocline’ variable are combined with the Lorenz63 through coupling processes (the 4VCM and 5VCM models). The coupling processes include the air-sea interaction and the coupling between the upper and deep oceans. Compared with the Lorenz63 model, the 4-variable coupled

model (4VCM) and 5-variable coupled model (5VCM) have larger natural variability due to the low-frequency boundary signals from air–sea interactions. The 5VCM has the maximum AR and GAR. From an atmosphere-only system to an ocean–atmosphere coupled system, the initial-value predictability problem becomes a joint initial-value and boundary-value problem. The results show that both boundary condition errors and initial condition errors play important roles in growing forecast errors.

We performed quantitative analysis of the spatiotemporal distributions of the practical predictability limits of geopotential height (GHT), comparing the coupled Climate Forecast System (CFS) and the uncoupled Global Ensemble Forecast System (GEFS) at 850 hPa, 500 hPa, and 200 hPa. Retrospective forecasts of the CFS and GEFS were used to evaluate the practical predictability limit. In the low-level atmosphere, the ocean improves prediction of the atmosphere most significantly in the tropics. The practical predictability limits of the CFS at 850 hPa appear to have a regional characteristic, with many maxima and minima, with a maximum of 11–13 days over the Western Pacific Warm Pool and 3–5 days over the Antarctic. In contrast, the practical predictability limits of the GEFS at 850 hPa have a significant difference, with a maximum of 9–11 days over the North Pacific and 3–7 days over the tropics. Note that the predictability limits of coupled models at 850-hPa geopotential height are higher than those of uncoupled models, especially in the tropics.

CFS has longer practical predictability limit than GEFS at 850 hPa over the tropics, while GEFS is more predictable than CFS at 500 hPa and 200 hPa. From 500 to 200 hPa, the difference between GEFS and CFS increased significantly in the tropics, especially over the middle-east equatorial Pacific Ocean. Moreover, it is worth noting that the difference in the practical predictability at 200 hPa between two models over the tropics varies with the seasons: increasing in the winter and reaching its peak in spring, then decreasing and achieving its minimum in autumn finally. The tropical oceans are crucial areas where large-scale air–sea interactions occur that have profound impacts on global climate variability. Using air–sea interactions as the atmospheric bottom boundary conditions could extend the predictability limits for geopotential height in lower troposphere. In addition, based on the 5VCM, we find that low-frequency forcing of atmospheric components has an important effect on the properties of atmospheric attractors, which may enhance the forecast skill in the upper troposphere.

We believe that there are two reasons for the difference in the predictability between CFS and GEFS over the tropics: coupling processes and atmosphere model. The coupling process in the lower atmosphere results in a higher predictability limit of CFS than GEFS, especially in equatorial regions. However, in the middle and upper atmosphere, the

coupling process has a smaller impact on the atmosphere, while the atmosphere model errors play an important role in the forecasting. As a result, at 200 hPa and 500 hPa, the predictability limit in GEFS over the tropics is longer than that of CFS.

We have only taken the first step in predictability studies for the coupled system. Compared with an uncoupled system, we find that a coupled system has longer practical predictability limits in the lower troposphere. However, further studies are required to clarify the physical processes at the air–sea interface on different temporal and space scales that extend the predictability limits in the coupled system. In addition to its application to the atmosphere within a coupled system, the concept of predictability limits can be extended to the ocean, land, and sea ice. In this paper, the retrospective forecasts of geopotential height in both the CFS and GEFS were used to evaluate predictability of atmospheric systems. In the future, more coupled models need to be employed to further explore the predictability limits of other components of coupled systems.

Acknowledgements This research was supported by the National Key R&D Program of China (2017YFC1404100, 2017YFC1404104) and the National Natural Science Foundation of China (Grant No. 41775100, 41830964), as well as Shandong Province’s “Taishan” Scientist Project and Qingdao “Creative and Leadership” Scientist Program. This research was also supported by the Center for High Performance Computing and System Simulation, Pilot National Laboratory for Marine Science and Technology (Qingdao).

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

References

- Bach E, Motesharrei S, Kalnay E, Ruiz-Barradas A (2019) Local atmosphere-ocean predictability: dynamical origins, lead times, and seasonality. *J Clim* 32(21):7507–7519
- Baldwin MP, Stephenson DB, Thompson DWJ, Dunkerton TJ, Charlton AJ, O’Neill A (2003) Stratospheric memory and skill of extended-range weather forecasts. *Science* 301(5633):636–640
- Chou JF (1989) Predictability of the atmosphere. *Adv Atmos Sci* 6:335–346
- Charney JG, Shukla J (1981) Predictability of monsoons. In: Light-hill J, Pearce RP (eds) *Monsoon Dynamics*. Cambridge University Press, 99–110

- Dalcher A, Kalnay E (1987) Error growth and predictability in operational ECMWF forecasts. *Tellus* 39A:474–491
- Ding RQ, Li JP (2007) Nonlinear finite-time Lyapunov exponent and predictability. *Phys Lett* 364A:396–400
- Farrell BF (1990) Small error dynamics and the predictability of atmospheric flows. *J Atmos Sci* 47:2409–2416
- Feng J, Li JP, Zhang J, Liu DQ, Ding RQ (2019) The relationship between deterministic and ensemble mean forecast errors revealed by global and local attractor radii. *Adv Atmos Sci* 36:271–278
- Hoerling MP, Kumar A (2002) Atmospheric response patterns associated with tropical forcing. *J Clim* 15:2184–2203
- Kalnay E (2002) Atmospheric modeling, data assimilation and predictability. Cambridge University Press, Cambridge, New York, p 230
- Kumar A, Schubert SD, Suarez MS (2003) Variability and predictability of 200-mb seasonal mean heights during summer and winter. *J Geophys Res* 108:4169
- Kumar A, Chen M, Wang W (2013) Understanding prediction skill of seasonal mean precipitation over the tropics. *J Clim* 26:5674–5681
- Leith CE (1974) Theoretical skill of Monte Carlo forecasts. *Mon Weather Rev* 102(6):409–418
- Li JP, Ding RQ (2011) Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogues. *Mon Weather Rev* 139:3265–3283
- Li JP, Ding RQ (2013) Temporal-spatial distribution of the predictability limit of monthly sea surface temperature in the global oceans. *Int J Climatol* 33:1936–1947
- Li JP, Ding RQ (2015) Seasonal and interannual weather prediction. In: North G, Pyle J, Zhang F (eds) *Encyclopedia of atmospheric sciences*, vol 6. 2nd edn. Academic Press and Elsevier, Amsterdam, Boston, pp 303–312
- Li JP, Ding RQ, Ha KJ (2008) Nonlinear local Lyapunov exponent and quantification of local predictability. *Chin Phys Lett* 25:1919–1922
- Li JP, Feng J, Ding RQ (2018) Attractor radius and global attractor radius and their application to the quantification of predictability limits. *Clim Dyn* 12:1–16
- Li JP, Wang S (2008) Some mathematical and numerical issues in geophysical fluid dynamics and climate dynamics. *Commun Comput Phys* 3(4):759–793
- Li JP, Zeng QC, Chou JF (2000) Computational uncertainty principle in nonlinear ordinary differential equations. I Numerical results *Science in China (Series E)* 43(5):449–460
- Lorenz EN (1963) Deterministic nonperiodic flow. *J Atmos Sci* 20:130–141
- Lorenz EN (1965) A study of the predictability of a 28-variable atmospheric model. *Tellus* 17:321–333
- Lorenz EN (1969) Atmospheric predictability as revealed by naturally occurring analogues. *J Atmos Sci* 26:636–646
- Meehl GA et al (2014) Decadal climate prediction: an update from the trenches. *Bull Am Meteor Soc* 95(2):243–267
- Murakami M (1979) Large-scale aspects of deep convective activity over the GATE area. *Mon Weather Rev* 107:994–1013
- Nese JM, Dutton JA (1993) Quantifying predictability variations in a low-order ocean-atmosphere model: a dynamical systems approach. *J Clim* 6(2):185–204
- Nese JM, Miller AJ, Dutton JA (1996) The nature of predictability enhancement in a low-order ocean-atmosphere model. *J Clim* 9(9):2167–2172
- Oseledec VI (1968) A multiplicative ergodic theorem: lyapunov characteristic numbers for dynamical systems. *Trans Moscow Math Soc* 19:197–231
- Palmer TN (1994) Chaos and predictability in forecasting the monsoons. *Proc Indian Natl Sci Acad* 60(1A):57–66
- Palmer TN (2006) Predictability of weather and climate: From theory to practice. In: Palmer T, Hagedorn R (eds) *Predictability of Weather and Climate*, Cambridge University Press, 1–10
- Pegion K et al (2019) The subseasonal experiment (SubX): a multi-model subseasonal prediction experiment. *Bull Am Meteor Soc* 100(10):2043–2060
- Peña M, Kalnay E, Cai M (2003) Statistics of locally coupled ocean and atmosphere intraseasonal anomalies in reanalysis and AMIP data. *Nonlinear Processes Geophys* 10:245–251
- Reichler TJ, Roads JO (2003) The role of boundary and initial conditions for dynamical seasonal predictability. *Nonlinear Process Geophys* 10(3):211–232
- Saenko OA, Gregory JM, Weaver AJ, Michael E (2002) Distinguishing the influence of heat, freshwater, and momentum fluxes on ocean circulation and climate. *J Clim* 15(24):3686–3697
- Shukla J (1998) Predictability in the Midst of Chaos: a scientific basis for climate forecasting. *Science* 282(5389):728–731
- Simmons AJ, Hollingsworth A (2002) Some aspects of the improvement in skill of numerical weather prediction. *Quart J Roy Meteor Soc* 128:647–677
- Thompson PD (1957) Uncertainty of initial state as a factor in the predictability of large-scale atmospheric flow pattern. *Tellus* 9:275–295
- Webster PJ, Lukas R (1992) TOGA COARE: the coupled ocean-atmosphere response experiment. *Bull Amer Meteor Soc* 73:1377–1416
- Wu Z, Huang NZ (2008) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adapt Data Anal* 1:1–41
- Wu R, Kirtman BP (2007) Regimes of seasonal air-sea interaction and implications for performance of forced simulations. *Climate Dyn* 29:393–410
- Zhang SQ (2011) A study of impacts of coupled model initial shocks and state-parameter optimization with observations on climate predictions using a simple pycnocline prediction model. *J Clim* 24:6210–6226
- Zhang SQ, Liu Z, Rosati A, Delworth T (2011) A study of enhance parameter correction with coupled data assimilation for climate estimation and prediction using a simple coupled model. *Tellus A* 64:10963
- Zhou X, Zhu Y, Hou D, Luo Y, Peng J, Wobus R (2017) Performance of the New NCEP global ensemble forecast system in a parallel experiment. *Wea Forecasting* 32:1989–2004

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.