Season-dependent forecast skill of the leading forced atmospheric circulation pattern over the North Pacific and North American region

XIAOJING JIA

Department of Earth Sciences, ZheJiang University, HangZhou, ZheJiang, China.

HAI LIN

Recherche en Prévision Numérique, Environment Canada, Dorval, Québec, Canada.

JUNE-YI LEE AND BIN WANG

International Pacific Research Center and Department of Meteorology, University of Hawaii, Honolulu, USA

*Corresponding author address: XiaoJing Jia, Department of Earth Sciences, ZheJiang University, HangZhou, ZheJiang, China.
E-mail: xiaojing.jia@mail.mcgill.ca
ABSTRACT

Multi-model ensemble (MME) seasonal forecasts are analyzed to evaluate numerical models’ performance in predicting the leading forced atmospheric circulation pattern over the extratropical Northern Hemisphere. Results show that the time evolution of the leading tropical Pacific sea surface temperature (SST)-coupled atmospheric pattern (MCA1), which is obtained by applying a maximum covariance analysis (MCA) between 500-hPa geopotential height (Z500) in the extratropical NH and SST in the tropical Pacific Ocean, can be predicted with a significant skill in MAM, JJA and DJF one-month ahead. However, most models perform poorly in capturing the time variation of MCA1 in SON with August 1st initial condition. Two possible reasons for the models’ low skill in SON are identified. Firstly, the models have the most pronounced errors in the mean state of SST and precipitation along the central equatorial Pacific. Due to the link between the divergent circulation forced by tropical heating and the midlatitude atmospheric circulation, errors in the mean state of tropical SST and precipitation may lead to a degradation of midlatitude forecast skill. Secondly, examination of the potential predictability of the atmosphere, estimated by the ratio of the total variance to the variance of the model forecasts due to internal dynamics, shows that the atmospheric potential predictability over the North Pacific-North American (NPNA) region is the lowest in SON comparing to the other three seasons. The low ratio in SON is due to a low variance associated with external forcing and a high variance related to atmospheric internal processes over this area.
1. Introduction

It is well known that some teleconnection patterns, such as the Pacific-North American (PNA) pattern and the North Atlantic Oscillation (NAO), can significantly influence the seasonal atmospheric condition over the Northern Hemisphere (NH) extratropics (Trenberth et al. 1998; Hoerling et al. 2001). Thompson and Wallace (1998) suggested that, comparing to the NAO, the Arctic Oscillation (AO) would more adequately represent the dominant mode of extratropical tropospheric atmospheric variability in the NH. Although there is still debate regarding the differences between the AO and NAO in the literature, many authors regard them as the same phenomenon. The PNA and NAO/AO together explain a significant part of the interannual variance of the extratropical NH atmospheric variability and are the two most important atmospheric patterns over the NH in winter time (Wallace and Guztler 1981; Barnston and Livezey 1987).

However, the PNA and NAO/AO structures may be distorted or become quite weak in other seasons than winter time and may not be the dominant atmospheric patterns that influence the weather and climate over the NH (Folland et al. 2009; Lee et al. 2011; Lee and Wang 2012). For example, using both observations and outputs from four atmospheric general circulation models (GCMs), Jia and Lin (2011) demonstrated that instead of the traditionally defined PNA and NAO/AO, the surface air temperature (SAT) over China in summer and fall are more significantly influenced by the leading forced atmospheric pattern, obtained by applying a maximum variance analysis (MCA) analysis between Z500 in the extratropical NH and SST in the tropical Pacific Ocean. It is further revealed that the relationship between the tropical Pacific SST and the extratropical atmospheric circulation
could be potentially useful to improve the forecast skill of dynamical numerical models. For instance, a statistical post-processing approach has been formulated by Lin et al. (2005a) based on the regression of the forecast models leading forced MCA patterns and the historical observations to correct the atmospheric response pattern to tropical Pacific SST anomalies. They found that this approach can significantly improve the seasonal forecasts of many variables such as the Z500 over the NH and the precipitation over Canada in winter time, and the SAT in fall over North America and China (Lin et al. 2005a, 2008; Jia et al. 2010; Jia and Lin 2011). It is therefore important for seasonal forecasting that the main features and time evolution of these tropical Pacific SST-forced large-scale atmospheric patterns in the observations are reasonably well predicted by numerical models.

Jia et al. (2010) showed that the forecast skill of numerical models in forecasting the climate over North America is quite seasonally dependent. Examination of the temporal correlation coefficients (TCC) between the observations and the MME of SAT forecasts during the period from 1969 to 2001 reveals a significant predictive skill over many regions of North America in MAM, JJA and DJF. However, only limited areas of predictive skill are found in SON over the east coast and northern Canada and parts of the Midwestern and south central United States. Examination of the percentage of significant area shows that more than 75% of North America has forecast skill significant at the 0.05 level in MAM, JJA and DJF whereas only 32% of North America has a significant forecast skill in SON. However, the reasons behind the pronounced seasonal dependence of numerical models' forecast skill remain unclear. The purpose of this study is to further investigate the forecast skill of numerical models in different seasons using more comprehensive model forecast results. Here we focus on investigating models' ability in forecasting the SST-
forced large-scale atmospheric pattern considering its importance to midlatitude seasonal
forecasting as we mentioned before. The possible causes of the seasonal dependence of
numerical models’ predictability will be explored and discussed.

The paper is organized as follows: In section 2, the data and models used in this study
are described. Section 3 presents the tropical Pacific SST forced leading atmospheric pattern
and the performance of numerical models in predicting this pattern in different seasons. It
shows that the forecast skill of MCA1 is the lowest in SON among the four seasons and this is
a common characteristic of most numerical models under examination. The possible reasons
accounting for the poor predictability in SON are investigated in section 4 and conclusions
are given in section 5.

2. Data and models

The following two hindcast datasets are used in this study. The first dataset is the en-
semble forecasts produced under the second phase of the Canadian Historical Forecasting
Project (HFP2) multimodel two-tier seasonal forecasting system conducted by Canadian Me-
teorological Centre (Kharin et al. 2009). The second dataset is from the Climate Prediction
and its Application to Society (CliPAS) project sponsored by the Asian-Pacific Economic
Cooperation (APEC) Climate Center (APCC)(Wang et al. 2009; Lee et al. 2010). It is of
interest to compare one-tier and two-tier approaches in predicting the leading atmospheric
circulation pattern over the extratropics.

One-tier seasonal forecasts use coupled GCMs (CGCMs) in which both the atmosphere
and ocean are initialized. Previous studies showed that one-tier seasonal forecasts had consid-
erable systematic errors in simulating the tropical ocean and atmosphere. As the systematic forecast errors could potentially cause biases in the global teleconnections that are associated with equatorial SSTA, the two-tier system, which uses persistent or pre-forecasted SST anomalies (SSTA) forcing has been demonstrated to have an obvious advantage over the direct use of CGCMs for the extratropics. However, recently, some studies have showed that it is important to take into account local monsoon-warm pool ocean interactions in seasonal forecasts over tropics (Wang et al. 2003, 2004, 2008; Wu and Kirtman 2005; Kumar et al. 2005). With the development of coupled climate models in the past decades it might be possible for CGCMs to improve their ability in capturing characteristics of extratropical teleconnections.

In the HFP2 project, four global atmospheric models were involved including the second and third generations of the general circulation models (GCM2 and GCM3) of the Canadian Centre for Climate Modelling and Analysis (CCCMA) (Boer et al. 1984; McFarlane et al. 1992), the reduced-resolution version of the global spectral model (SEF) (Ritchie 1991) and the Global Environmental Multiscale model (GEM) of the Recherche en Prévision Numérique (RPN) (Côté et al. 1998b,a). The first tier of the HFP2 forecasting system is the sum of the SST of the month prior to the forecast period, persisted through the forecast period, and the monthly-varying climatological SST. The SST and ice data were taken from the Seasonal Prediction Model Intercomparison Project-2 (SMIP-2) boundary data. For each model, 10 four-month forecasts were carried out starting from the first day of each month during the period from 1969 to 2001. The initial conditions were reanalyses from the National Centers for Environmental Predictions (NCEP) and the National Center for Atmospheric Research (NCAR) Reanalysis I (Kalnay et al. 1996).
For the HFP2 forecasts we make use of the ensemble mean of the seasonal forecasts spanning the 33 years from 1969 to 2001. As we are interested in the forecast skill coming from anomalies in the boundary conditions, seasonal forecasts with a one-month lead time are analyzed to minimize the influence of the initial conditions. In other words, the forecast data used for MAM, JJA, SON and DJF are from the forecasts made for February through May (FMAM), May through August (MJJA), August through November (ASON) and November through February (NDJF). The dominant external source of a skillful forecast is derived from the prescribed SST anomalies (Derome et al. 2001).

CliPAS is an international project aimed at doing multi-model intercomparison and synthesis. One of the objectives of the CliPAS project is to develop a well-validated MME prediction system and to study the forecast skill of climate variations (Wang et al. 2009; Lee et al. 2010). This project involves a large group of climate scientists from United States, South Korea, Japan, China, and Australia. Ensemble retrospective forecasts were made by one-tier and two-tier climate models from 16 different institutes. A brief summary of these institutions and model specifications have been given in Wang (2009). While each model has a different forecast length and ensemble size they all have ensemble forecasts starting from May 1 to at least September 30 for the boreal summer season (JJA) and from November 1 to at least March 31 for the boreal winter season (DJF). The two-tier predictions are from Florida State University (FSU2), Geophysical Fluid Dynamic Lab (GFDL), Institute of Atmospheric Physics/Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (IAPL), National Center for Environmental Prediction (NCT2), Seoul National University (SUT2), and University of Hawaii (UHC2 and UHT2). The one-tier predictions we use in this study are from Centre for Australia Weather and Climate Research (CAWC),
GFDL (GFT1), National Aeronautics and Space Administration (NASA), National Center for Environmental Prediction (NCEP), Seoul National University (SUT1), and University of Hawaii (UHT1). All one-tier models do not apply any flux correction. Eight of the above numerical models have forecasts for all four seasons which include CAWC, UHT1, SUT1, NCEP, GFT1, IAPL, UHC2 and NCT2. All the two-tier models except NCT2 were forced by the global SST field pre-forecasts the Seoul National University (SNU) statistical-dynamical forecast model. A detailed description of this SST forecast method is given in Kug et al. (2007). The NCT2 was forced by the forecast SST from its coupled version CFS. All models used the same initial conditions from NCEP/DOE Reanalysis II (Kanamitsu et al. 2002). For the CliPAS data, we make use of the seasonal forecasts during the common hindcast period of these numerical models which is from 1982 to 2001.

In the present study, the MME seasonal forecasts were simply made by averaging forecasts of different models. The model climatology is removed for each individual model and only anomalies are considered. The data used for the verification of the seasonal forecasts are the CRU TS 2.1 dataset, a set of monthly averaged observed SAT over the land surface from the Climate Research Unit (CRU) at the University of East Anglia, UK (Mitchell and Jones 2005).
3. Seasonal forecast skill of MCA1

a. MCA analysis

The tropical Pacific is known to be a major forcing area for some extratropical circulation patterns over the Pacific Ocean and surrounding land area on seasonal time scales. The SST anomalies of the tropical Pacific and its associated anomalous heating in the tropical atmosphere are of central importance for determining tropical-extratropical teleconnections. In this study, the atmospheric patterns that are associated with the tropical Pacific SST forcing are obtained by conducting an MCA analysis (Bretherton et al. 1992) between the Z500 north of 20°N and the tropical Pacific SST (20°N - 20°S, 120°E - 90°W) in the observations. The MCA calculation is performed on the Z500 and the simultaneous SST anomalies for the period from 1969 to 2001 for the four seasons separately. When the MCA is applied, the data on a regular lat-lon grid are weighted by the square root of the cosine of the latitude to ensure that equal areas are afforded equal weight in the analysis. Here we only concentrate on the leading MCA mode, MCA1. The MCA results are displayed in Fig. 1 where the magnitude corresponds to one standard deviation of the respective expansion coefficient. Areas of heterogeneous correlation with statistical significance passing the 0.01 level are shaded. It should be mentioned that there is some limit in physical interpretation using MCA pairs. As discussed in Newman and Sardeshmukh (1995), the MCA technique may not always be able to recover the relationship between two physically related fields.

It is found that MCA1 accounts for a dominant percentage of the total covariance between the Z500 and SST fields. The covariance between Z500 and SST explained by this MCA mode is 50%, 63% and 79% for MAM, SON and DJF, respectively, for this period.
based on a squared covariance fraction (SCF). In JJA, only 36% of the covariance between
Z500 and SST fields is explained by MCA1. As the purpose of this study is to investigate
the seasonal dependence of numerical models’ forecast skill for the leading atmospheric pat-
tern coupled with the tropical Pacific SST, we present the results of all four seasons while
keep the difference of JJA to the other three seasons in mind. The temporal correlation
coefficients (TCC) between the expansion coefficients of Z500 and SST fields during this
period for the four seasons are all statistically significant at the 0.05 level according to a
Student’s t test. Figure 1 shows that MCA1 varies from season to season with DJF having
a Tropical/Northern Hemisphere (TNH)-like pattern (Barnston et al. 1991) and JJA having
the weakest pattern among the four seasons. The corresponding SST field of MCA1 shows
a clear El Niño signal with positive SST anomalies in the eastern tropical Pacific for SON
and DJF. In MAM and JJA, the positive SST anomalies become weak in magnitude and are
distorted to some extent.

To associate the atmospheric condition over North America with MCA1, we calculate the
TCC between the CRU SAT and the atmospheric expansion coefficients of MCA1 for the
four seasons, and the results are displayed in Fig. 2. The shaded areas represent correlations
with a significance level of 0.05 according to a Student’s t-test. It shows that areas with
significant correlations account for 34%, 45%, 62% and 65% of North America for MAM,
JJA, SON and DJF, respectively. The CRU SAT is over the land surface and is seasonally
averaged from the period 1969 to 2001 on a 2.5°×2.5° grid. There are totally 484 grid
points located over North America north of 20°N. According to Livezey and Chen (1983),
if we suppose the grid data over North America are statistically independent, which they
are not, in order to reject a null hypothesis that the result presented in Fig. 2 was by
accident at the 95% level it requires that at least 7% of the area is significant (their Figure 2), which is much less than the percentages showed above. To assess the field significance of the correlation maps considering the spatial dependence a Monte Carlo approach is applied following Livezey and Chen (1983). One thousand correlation maps are calculated between the CRU SAT and randomized time series of MCA1 and the significance at each grid point is tested at the 95% level. Results show that in MAM only 2.1% of these trials have an area of significant correlation that is larger than the shaded area in Fig. 2a, indicating that the correlations shown in Fig. 2a are field significant at 0.05 level. The other three seasons even pass the 0.01 field significant test.

It is seen that the influence of MCA1 on the SAT over North America is quite season-dependent. The distribution of TCC shows a similar pattern for MAM and DJF with DJF having a larger significant TCC area than MAM. A dipole structure is seen for these two seasons with significant positive TCC appearing over northwestern North America and a negative area over southeast North America. The TCC in SON is an east-west dipole pattern with positive values over northwestern North America and negative values over the east half of North America. In JJA, significant positive correlations are seen over northern and southern North America. The links between the North American SAT and MCA1 for the four seasons are thus clear, implying that a skillful forecast of MCA1 can benefit the seasonal forecast of SAT over North America.

To see whether or not the relationship between the tropical Pacific SST and the atmosphere shown in Fig. 1 can be captured by the numerical models we examine the MCA1 in the HFP2 model atmosphere associated with the tropical Pacific SST forcing and compare it to its observed counterpart. An MCA analysis is applied between the observed tropical
Pacific SST and the MME mean Z500 over the NH and the results are depicted in Fig. 3. As we mentioned before, the forecast data used for MAM, JJA, SON and DJF are from the forecasts made for February through May (FMAM), May through August (MJJA), August through November (ASON) and November through February (NDJF). The SST forcing used in HFP2 is the sum of SSTA from the month before the start of the forecasts and the monthly-varying climatological SST. Therefore in the MCA analysis, the tropical Pacific SST used is from October, January, April and July, respectively. A comparison between Fig.1 and Fig. 3 shows that, though model prediction tends to overestimate the ENSO impact on the atmospheric variability, the MCA1 pattern of the atmospheric component in the observations is reasonably well reproduced by the HFP2 models, except in JJA, when the atmospheric pattern is quite weak in both the observations and the model forecasts. The similarity of MCA1 in the observations and in the HFP2 model forecasts are further examined by calculating the pattern correlation coefficients (PCC) of MCA1 for the four seasons (not shown). Results showed that the PCC of MCA1 are all significant at the 0.05 significant level for all the four seasons, with the highest and lowest PCC appearing in DJF and JJA, respectively.

b. Forecast skill of the time variation of MCA1

In seasonal forecasts, the signal coming from a forcing external to the atmosphere, i.e., an SST anomaly, is considered potentially predictable. The part of variability that is related to atmospheric internal dynamics is unpredictable. The MCA1 obtained above is the most pronounced atmospheric pattern that is coupled with the tropical Pacific SST and it is
reasonable to assume that MCA1 is likely the most predictable atmospheric pattern. Based on this premise, we strive to first determine to what extent the climate models and their MME mean forecast can capture the time evolution of this observed atmospheric pattern. To evaluate the performance of numerical models in predicting the time variation of MCA1 we project the one-month lead ensemble forecast of Z500 of each individual HFP2 model on to MCA1 as shown in Fig. 1. The TCC between the atmospheric expansion coefficients of MCA1 in the observations and in the ensemble forecasts are shown in Table 1. A correlation coefficient larger than 0.33 is considered to be statistically significant at a 0.05 level, and is shown in boldface in Table 1. The MME mean forecasts of the four HFP2 models as well as the averaged skill of all the individual models, which are represented as “AVE”, are also presented in Table 1.

Table 1 shows that the HFP2 ensemble forecasts capture the realistic time evolution of MCA1 in MAM, JJA and DJF where the TCC values for the three seasons exceed 0.62 for all four models, significant at the 0.01 level according to a two-tailed Student’s $t$ test. In SON, however, only SEF model has some skill in predicting the variability of MCA1. The HFP2 MME mean forecast has a TCC skill of 0.72, 0.75, and 0.70 for MAM, JJA and DJF, respectively; whereas it is only 0.32 and can not pass the significance test in SON. It also needs to point out that the total covariance between the Z500 and SST fields explained by MCA1 in JJA is much less than the other three seasons. Also noticed is that the forecast skill of the MME mean forecast is better than the averaged skill of all individual models for all four seasons, indicating that the MME method is superior in reducing forecast errors and quantifying forecast uncertainties due to model formulation, consistent with previous studies (e.g., Wang et al. 2009).
To see whether or not the results shown in Table 1 are robust, the ensemble mean forecast from the CliPAS project was also examined. The one-month lead ensemble mean forecasts from 1982 to 2001 were projected onto the atmospheric component of MCA1 as shown in Fig. 1 and the obtained time series were compared to the atmospheric expansion coefficients of MCA1 in the observations (Table 2). It can be seen that most models have a significant skill in predicting the variability of MCA1 in MAM, JJA and DJF, and the MME forecast has a TCC of 0.72, 0.74 and 0.64, respectively, significant at the 0.01 level. The lowest forecast skill also appears in SON where only two of the models (UHT1 and NCT2) can skillfully predict the time variability of MCA1 and the average forecast skill is only 0.28. Although different numerical models are examined, the results shown in Table 2 are consistent with those obtained from the HFP2 output, as expected, indicating the robustness of the results.

c. Seasonal forecast skill over the NPNA region

The TCC between the observations and the one-month lead ensemble forecasts of Z500 over the NPNA region for four seasons are presented in Fig. 4. The super MME mean Z500 forecast (Comp, hereafter) includes output from the four numerical models from HFP2 and eight numerical models from CliPAS that have one-month lead forecasts for all four seasons during the period from 1982 to 2001. As is seen that the super MME mean Z500 forecasts have statistically significant skill over a broad tropical band. In the extratropical regions, significant TCC can be seen over the North Pacific, eastern Canada and around the Gulf of Mexico in MAM and DJF, likely reflecting the atmospheric response to the ENSO signal over the eastern tropical Pacific (Derome et al. 2001). The forecast skill is relatively weaker
in MAM than that in DJF. In contrast to MAM and DJF, only a small significant correlation region is observed over northwestern and northern Canada in JJA over the extratropics while there is almost no skill over mid- high-latitude North America in SON. Also it can be seen from Fig. 4 is that the forecast skill tends to increase from JJA to DJF, which may be related to the enhanced ENSO forcing from its developing (JJA) to its mature (DJF) phase, consistent with previous studies (Kumar and Hoerling 2003; Wang et al. 2009). In the middle latitude North Pacific, the forecast skill in JJA is quite high which is, as pointed by Wang et al. (2009), Lee et al. (2011) and Lee and Wang (2012), mainly attributable to an ENSO-monsoon atmospheric teleconnection. Also noticed is that the area of significant forecast skill over the low-latitude Pacific Ocean shifts eastward from JJA to DJF, in agreement with the results of Wang et al. (2009).

The forecast skill of numerical models over the NPNA region (20°N - 80°N, 160°E - 80°W) is also evaluated by computing the PCC between the observed and Comp model forecast Z500 anomalies (Fig. 5) and then make a time-mean PCC over the entire hindcast period in order to quantify the overall MME hindcast skill (Fig. 6). Figure 5 shows that the five-year running averaged PCC is the highest in DJF and is the lowest in SON while MAM and JJA have a comparable PCC during the period. The PCC for DJF in 1997 is 0.93 which is the strongest El Niño winter during the period. In general, the forecast skill is relatively poor during early 1990s for most seasons. The time average of PCC for the Comp model forecasts shows that there is almost no skill in predicting the Z500 anomalies in SON over this region. Examination of individual models reveals that the time averaged PCC is the lowest in SON for all models except CAWC and NCT2 where JJA has an even lower forecast skill during the period under examination (Fig. 6).
The above results indicate that numerical models can reasonably well predict the variability of MCA1 in MAM, JJA and DJF, but most models have problems in forecasting the MCA1 in SON. Considering the significant influence of MCA1 on the SAT over North America as shown in Fig. 2, it is very possible that the numerical models’ ability in predicting the SAT over the NPNA region is impacted by the above results. We confirmed this by examining the TCC between the observations and the MME of SAT forecasts in HFP2 data for four seasons during the period from 1969 to 2001 (Fig. 7). Areas with a correlation score significant at the 0.05 level or better, are shaded. Significant predictive skills are found over many regions of North America in MAM, JJA and DJF. For example, skillful areas appear over central and southern North America in MAM and DJF. In JJA, skills can be found over most part of Canada and North America south of 35°N while only limited areas of predictive skill are found in SON. The predictive skill of the SAT seasonal forecast is also measured by mean squared error (MSE) (not shown). Following Smith and Livezey (1999) and Lin et al. (2008), MSE is calculated using the observed and four model averaged SAT anomalies that are normalized using their respective standard deviations. The distributions of MSE for four seasons are consistent with the correlation skill shown in Fig. 4. Overall the forecast skill of SAT for the one-month lead MME mean forecasts is the lowest over North America in SON among the four seasons.
4. Possible causes of the poor forecast skill in SON

a. The climatological mean state over the tropical Pacific region

In this section we try to investigate possible reasons why most numerical models tend to produce an erroneous forecast of the atmospheric circulation over the NPNA region in SON. The tropical Pacific is known to be an important forcing area for atmospheric variability over the Pacific Ocean and surrounding land area on a seasonal time scale, especially for boreal winter. The SST anomaly in the equatorial Pacific, for example that related to El Niño, is closely related to changes in precipitation and diabatic heating that generate anomalous vertical motion and upper-level divergence, that leads to extratropical Rossby waves and global teleconnections (Wallace and Guztler 1981; Sardeshmukh and Hoskins 1988). The 500-hPa geopotential height and SAT changes can be caused by the SST related teleconnections that account for most of the seasonal forecast skill in the NPNA region (Derome et al. 2001; Lin et al. 2005b; Jia et al. 2009). Thus it is critical for numerical models to have a realistic SST simulation over this region to have a skillful seasonal forecast. We start by examining the climatological mean state of SST in the observations and compare to that in the model forecasts. Figure 8 depicts the difference between model forecasts and the observed SST averaged over the period from 1982 to 2001. The climatology of the model forecast SST was computed using the MME of five one-tier CliPAS models which have forecasts for all four seasons (see Table 2).

It is clear that the forecast skill of the climatological SST vary by season. The forecast climatological SST errors for JJA and SON are more obvious than those in MAM and DJF. In JJA pronounced positive SST biases appear over the eastern tropical Pacific Ocean with
negative SST biases over the subtropical western Pacific. For SON, positive biases can also be found along the tropical eastern Pacific Ocean. However, if we focus on the central equatorial Pacific area, where SST has a close relationship to MCA1 as shown above (Fig.1, Fig.3), huge negative SST biases are noticed indicating that the forecast SST climatology is cooler than the observations in this region for SON. Examination of individual numerical models shows that almost all models produce a cooler SST climatology than the observations around this region in SON except UHT1 which has positive SST biases for all four seasons.

The bias of the tropical Pacific SST forecast could degrade the model’s capability in simulating the precipitation climatology over there (Lee et al. 2010). A correct representation of the divergent circulation associated with tropical heating is important for determining midlatitude atmospheric circulation patterns. Furthermore, the mean tropical SST and precipitation, through diabatic heating, have an important influence on the atmospheric mean flow that in turn affects the tropical forced Rossby wave propagation which is related to seasonal forecast skill. Here the climatological mean state precipitation fields are also examined and displayed in Fig. 9. The spatial distributions of precipitation errors are somewhat similar to each other among the four seasons. They all have negative precipitation biases around the equatorial Pacific and positive precipitation biases to the north. The magnitudes of the precipitation errors in JJA and SON are larger than those in MAM and DJF. Again, if we focus on the central equatorial Pacific, it is obvious that the negative precipitation biases in SON, which indicates a weaker tropical forcing than the observations, are more significant than other seasons. Further examination indicates that the negative precipitation biases in this region are quite consistent among the numerical models. All numerical models have pronounced negative precipitation biases around the central equatorial Pacific in SON.
To examine the current level of precipitation forecast skill we also compute the PCC between the observed and Comp model forecast precipitation anomaly fields. Figure 10 depicts the five-year running averaged PCC over the tropical Pacific region (30°S-30°N, 120°E-60°W) for four seasons. It reveals that there is clearly a seasonal dependence and interannual variability of the precipitation forecast skill. The forecast skill for SON is not always the worst during this period. However, the PCC curve of SON is generally lower than the other three seasons. The time averaged PCC of the Comp (shown as MME) and twelve individual numerical models are compared to each other in order to further quantify the forecast precipitation skill (Fig. 11). It shows that the forecast skill of Comp is higher than any individual model. Also it can be seen that the time averaged PCC for SON is the lowest among the four seasons for nine of the twelve model forecasts under examination. The time averaged pattern correlation coefficients of the twelve numerical models varies from 0.18 to 0.35 in SON while that for Comp is 0.42. It can be concluded from the above analysis that the poor forecast skill of numerical models in SON over the NPNA region is at least partly due to the fact that in SON numerical model forecasts have relatively bigger errors in the climatological SST over the tropical Pacific (Fig. 8). The cold biases of SST in the model causes weaker precipitation response around the equatorial Pacific (Fig. 9) and more pronounced erroneous precipitation pattern (Fig. 10 and Fig. 11) than the other three seasons. This precipitation biases may cause errors in the extratropical mean flow and global teleconnection patterns and can therefore degrade the prediction skill in this season.

Wang et al. (2009) examined the ability of CliPAS MME one-month lead hindcast in predicting the spatial-temporal structures of the first two leading empirical orthogonal modes of the equatorial SST anomalies for JJA and DJF. Bias of a westward shift of the SST
anomaly between the dateline and 120°E was found. In this study, we found that the cold bias of the mean state of SST and its associated erroneous precipitation response over the equatorial Pacific in the models is more significant in SON than in the other three seasons. These biases of tropical SST and precipitation may cause errors in the extratropical mean flow that influences the forced Rossby wave propagation and the global teleconnections associated with equatorial SSTA forcing, degrading seasonal climate prediction skill in the extratropical regions. As demonstrated in Lee et al. (2010), a correction of the inherent bias in the mean state is critical for improving the long-lead seasonal prediction.

In the HFP2 two-tier seasonal forecast, the SST forcing used are the sum of SSTA, obtained from observations one month preceding the forecast period, and the monthly-varying climatological SST. An examination of the precipitation response of the four HFP2 models over the tropical Pacific region reveals that the amplitude of the precipitation response is generally weaker than the observations especially over the equatorial Pacific where negative precipitation biases are observed. These precipitation biases are found larger than those shown in Fig. 9, indicating the superiority of one-tier numerical models in simulating the climatological tropical Pacific precipitation compared to the two-tier models (not shown). In the study of Wang et al. (2004), they showed that the seasonal mean SST anomalies are negatively correlated to the precipitation in the Asian-Pacific monsoon region, especially when the precipitation leads SST by one month, suggesting that SST anomalies are forced by the atmospheric circulations over this region. This relationship can be captured by the coupled models realistically. However, the two-tier models were found unable to simulate properly the local SST-precipitation correlation caused by the lack of feedback from the atmosphere.
It is known that the skill of seasonal forecasts depends on the extent to which the boundary forcing can generate strong enough signals recognizable from the chaotic internal variability of the atmosphere (Kumar and Hoerling 1995). To investigate why most numerical models have the lowest forecast skill over the NPNA region in SON among the four seasons, we examine the potential predictability of the atmosphere over this region.

Let $X_{sy}$ represent the forecast seasonal mean $Z500$, where $s = 1, 2, ..., S$ identifies a particular simulation in an ensemble of $S$ simulations and $y = 1, 2, ... Y$ identifies a particular year in the period from 1982 to 2001.

The total variance of the data can be expressed as

$$\sigma^2_T = \frac{1}{SY} \sum_{s=1}^{S} \sum_{y=1}^{Y} \hat{X}_{sy}^2,$$  \hspace{1cm} (1)

where $\hat{X}_{sy}$ is the anomaly that is defined as the deviation of the model forecast $Z500$ from its climatology.

The total variance in the model simulation can be divided into the external variance ($\sigma^2_F$) and internal variance ($\sigma^2_I$). The external variance is the part associated with the SST variability and can be obtained from the ensemble mean

$$\sigma^2_F = \frac{1}{Y} \sum_{y=1}^{Y} \hat{X}_y^2,$$  \hspace{1cm} (2)

where $\hat{X}_y$ represent the ensemble mean for year $y$ that can be obtained through $\hat{X}_y = \frac{1}{S} \sum_{s=1}^{S} \hat{X}_{sy}$.
The internal variance is due to member to member differences and can be estimated by the deviation of each member from the ensemble mean and can be expressed as

$$\sigma_I^2 = \frac{1}{SY} \sum_{s=1}^{S} \sum_{y=1}^{Y} \hat{X}_{sy}^2$$

where $\hat{X}_{sy}$ is the deviation of member from the ensemble average.

The potential predictability of the atmosphere can be estimated by calculating the variance ratio of the total variance to the internal variance (Zwiers 1996).

$$\text{ratio} = \frac{\sigma_T^2}{\sigma_I^2}$$

The variance ratio is assessed using models from the CliPAS project that have seasonal forecasts for all the four seasons and have seasonal forecasts available for individual members (CAWC, UHT1, NCEP, GFT1, UHC2 and NCT2). The spatial distributions of the ratio over the NPNA region for the four seasons are depicted in Fig. 12. Regions where the ratio is greater than 1 at a 0.05 significance level from an F-test are shaded. For MAM and DJF, high ratios can be seen over the low latitudes, the eastern North Pacific, western Canada, and southeastern North America, reminiscent of a wavetrain pattern over the NPNA region, suggesting that the potential predictability over this region is related to ENSO variability. The distribution of the variance ratio is similar for MAM and DJF with MAM having relatively higher values than DJF. In JJA, although the wavetrain signal is not as obvious as that in MAM and DJF, high variance ratios also appear over central North America. Obviously, the lowest potential predictable season is SON when areas with relative high variance ratio appear along the tropical band. The above results can, at least, partly explain
our previous results that most numerical models have difficulty in forecasting the time variability of MCA1 and have the lowest predictive skill for Z500 over the NPNA region in SON.

To further understand the seasonal dependence of the potential predictability of the atmosphere over the NPNA region, the variance due to atmospheric internal dynamics (noise) and the variance associated with the external forcing (signal) are examined and presented in Fig. 13 and Fig. 14, respectively. It can be seen from Fig. 13 that while the magnitudes of the noise level change distinctly from season to season the maximum of the variance lying over the mid-North Pacific for the whole year. The largest and weakest noise levels appear in DJF and JJA, respectively, as expected. The maximum value of noise in DJF is almost five times larger than that in JJA. The noise level in SON is seen to have a magnitude about two times of that in JJA. The variances of the model forecasts due to external forcing show that the signal level is comparable in magnitude between DJF and MAM (Fig. 14). However, the noise level in DJF is obviously larger than that in MAM making MAM more predictable than DJF as can be seen from Fig. 12. The signal levels in JJA and SON are quite similar and much weaker compared to those in MAM and DJF. The maximum value of signal during DJF is about six times larger than that in JJA and SON. The obvious seasonality of the signal is more likely the result of the dynamics of tropical-extratropical interactions since the seasonal variation of the tropical convective forcing is not obvious as discussed in Kumar and Hoerling (1998). Although JJA and SON have a similar signal level over the NPNA region, the noise level in SON is much higher than that in JJA making the potential predictability of the atmosphere in SON the lowest among the four seasons. From JJA to SON, the NH extratropical westerly jet enhances and so do transient activities
and nonlinear interactions. These atmospheric internal processes lead to strong interannual
variability which is unpredictable. The results of the potential predictability is consistent
with our previous results showing that most numerical models have problems in forecasting
the time variability of MCA1 in SON. However, as discussed in Kumar et al. (2007), it should
be kept in mind that the decomposition of external and internal components of seasonal mean
atmospheric variability may be sensitive to the GCMs employed in this study.

5. Summary and discussion

Equatorial Pacific SST forcing is of central importance for determining well-known tropical-
extratropical teleconnections and is known as the primary source of atmospheric variability
on seasonal time scale. In this study, we start by gauging numerical models’ performance
in predicting MCA1 over the NH, which is obtained by doing an MCA analysis on the
seasonally averaged Z500 over the NH and the simultaneous tropical Pacific SST for four
seasons separately. The percentage of the total covariance between the Z500 and SST fields
explained by MCA1 varies from 50% to 79% for MAM, SON and DJF while is about 36% for
JJA. It was found that MCA1 can significantly influence the SAT over North America. The
structure of the atmospheric component of MCA1 varies from season to season. Examination
of the model forecast results shows that numerical models can capture, with a high fidelity,
the time evolution of MCA1 in MAM, JJA and DJF while most models fail to predict the
variability of MCA1 in SON. The TCC and PCC results further confirm that the poor skill
for the extratropical atmospheric circulation patterns over the NPNA region in SON is quite
systematic among GCMs.
Further examination reveals two possible sources of the models low skill for SON atmospheric variability over the NPNA region. Firstly, most models have the most pronounced errors in the mean states of SST and precipitation along the central equatorial Pacific in SON. As indicated by previous studies (e.g. Lee et al. 2010), corrections of the inherent errors in the mean state are critical for improving the seasonal prediction. Secondly, the potential predictability of the atmosphere over the NPNA region is the lowest in SON in terms of the ratio of the total variance to the internal variance. In SON, high potential predictability appears only over the low latitudes. The low potential predictability in the extratropical regions in SON is attributable to low variance associated with external forcings and high variance related to atmospheric internal processes.

Many previous studies analyzed the predictability of numerical models over the tropics. Building on these studies, we examined the difference of forecast skill between one-tier and two-tier seasonal forecasting systems over extratropical regions. We divided the numerical models of HFP2 and CliPAS into one-tier and two-tier two categories and examined their abilities in predicting the time evolution of MCA1. The TCC of the MME mean seasonal forecasts of Z500 for the four seasons are also compared between the one-tier and two-tier seasonal forecasting systems in the NPNA region. Although not shown, we found that the forecast skills of one-tier models are generally better than two-tier models in predicting the time evolution of MCA1 in JJA, SON and DJF while having a comparable skill in MAM. Although not very pronounced, examination of the TCC of Z500 shows that the performances of one-tier numerical models are better than two-tier numerical models over subtropical regions and some areas over the extratropical NPNA area.
Acknowledgments.

This research was jointly funded by National Natural Sciences Fundation of China (grant 41105037) and by Fund Project of ZheJiang Provincial Department of Education (grant Y200907077). Lin is partly supported by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) and by the Natural Science and Engineering Research Council of Canada (NSERC). Lee and Wang were supported by APEC Climate Center (APCC) international research project (CliPAS) and by IPRC, which in part supported by JAMSTEC, NOAA, and NASA. We acknowledge contributions from all CliPAS team members. This is the IPRC publication 862 and SOEST publication 8582. We are grateful to the reviewers for their helpful suggestions on improving our paper.
REFERENCES

Barnston, A. and R. E. Livezey, 1987: Classification, seasonality and persistence of low-

Barnston, A., R. Livezey, and M. Halpert, 1991: Modulation of southern oscillation-northern

Boer, G., N. A. McFarlane, R. Laprise, J. D. Henderson, and J.-P. Blanchet, 1984: The cana-


Côté, J., J.-G. Desmarais, S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth,
1998a: The operational cmc-mrb global environmental multiscale (gem) model: Part ii -

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998b: The oper-
ational cmc-mrb global environmental multiscale (gem) model: Part i - design considera-

Derome, J., et al., 2001: Seasonal predictions based on two dynamical models. *Atmosphere-


Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N. Lau, and C. Ropelewski, 1998:


List of Tables

1 Correlations between the atmospheric expansion coefficients of MCA1 in the observations and in the one-month lead HFP2 model forecasts. The correlations with statistical significance passing the 0.05 level are set in boldface. 31

2 Correlations between the atmospheric expansion coefficients of MCA1 in the observations and in the one-month lead CliPAS model forecasts. The correlations with statistical significance passing the 0.05 level are set in boldface. 32
Table 1. Correlations between the atmospheric expansion coefficients of MCA1 in the observations and in the one-month lead HFP2 model forecasts. The correlations with statistical significance passing the 0.05 level are set in boldface.

<table>
<thead>
<tr>
<th></th>
<th>MAM</th>
<th>JJA</th>
<th>SON</th>
<th>DJF</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCM3</td>
<td>0.70</td>
<td>0.76</td>
<td>0.15</td>
<td>0.69</td>
</tr>
<tr>
<td>GCM2</td>
<td>0.70</td>
<td>0.73</td>
<td>0.18</td>
<td>0.64</td>
</tr>
<tr>
<td>GEM</td>
<td>0.74</td>
<td>0.65</td>
<td>0.29</td>
<td>0.66</td>
</tr>
<tr>
<td>SEF</td>
<td>0.62</td>
<td>0.68</td>
<td>0.42</td>
<td>0.63</td>
</tr>
<tr>
<td>MME</td>
<td>0.72</td>
<td>0.75</td>
<td>0.32</td>
<td>0.70</td>
</tr>
<tr>
<td>AVE</td>
<td>0.69</td>
<td>0.70</td>
<td>0.26</td>
<td>0.66</td>
</tr>
</tbody>
</table>
Table 2. Correlations between the atmospheric expansion coefficients of MCA1 in the observations and in the one-month lead CliPAS model forecasts. The correlations with statistical significance passing the 0.05 level are set in boldface.

<table>
<thead>
<tr>
<th></th>
<th>MAM</th>
<th>JJA</th>
<th>SON</th>
<th>DJF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAWC</td>
<td>0.54</td>
<td>0.67</td>
<td>0.19</td>
<td>0.66</td>
</tr>
<tr>
<td>UHT1</td>
<td>0.44</td>
<td>0.29</td>
<td>0.50</td>
<td>0.57</td>
</tr>
<tr>
<td>SUT1</td>
<td>0.67</td>
<td>0.48</td>
<td>0.33</td>
<td>0.52</td>
</tr>
<tr>
<td>NCEP</td>
<td>0.63</td>
<td>0.19</td>
<td>0.27</td>
<td>0.62</td>
</tr>
<tr>
<td>GFT1</td>
<td>0.77</td>
<td>0.86</td>
<td>0.28</td>
<td>0.67</td>
</tr>
<tr>
<td>NASA</td>
<td></td>
<td></td>
<td>0.59</td>
<td>0.41</td>
</tr>
<tr>
<td>IAPL</td>
<td>0.52</td>
<td>0.64</td>
<td>-0.01</td>
<td>0.50</td>
</tr>
<tr>
<td>UHC2</td>
<td>0.65</td>
<td>0.34</td>
<td>-0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>NCT2</td>
<td>0.59</td>
<td>0.19</td>
<td>0.59</td>
<td>0.67</td>
</tr>
<tr>
<td>FSU2</td>
<td></td>
<td>0.51</td>
<td></td>
<td>0.42</td>
</tr>
<tr>
<td>GFDL</td>
<td></td>
<td>0.47</td>
<td></td>
<td>0.61</td>
</tr>
<tr>
<td>UHT2</td>
<td></td>
<td>0.45</td>
<td></td>
<td>0.58</td>
</tr>
<tr>
<td>SUT2</td>
<td></td>
<td>0.63</td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>MME</td>
<td>0.71</td>
<td>0.74</td>
<td>0.46</td>
<td>0.64</td>
</tr>
<tr>
<td>AVE</td>
<td>0.61</td>
<td>0.48</td>
<td>0.28</td>
<td>0.53</td>
</tr>
</tbody>
</table>
List of Figures

1  Observed Z500 (top panel) and SST (bottom panel) distributions of the first SVD for (a) MAM, (b) JJA, (c) SON and (d) DJF. The contour interval is 5 m in top panels and 0.2°C in bottom panels for (a), (c) and (d) while is 3 m in top panels and 0.15°C in bottom panels for (b). The magnitude corresponds to one standard deviation of each time coefficient. Areas of heterogeneous correlation with statistical significance passing the 0.01 level are shaded. 36

2  Temporal correlation between the CRU SAT and the expansion coefficient of the atmospheric component of MCA1 for (a) MAM, (b) JJA, (c) SON and (d) DJF. The contour interval is 0.2. The shaded areas represent correlations with a significance level of 0.05 according to a Students t-test. 37

3  HFP MME mean Z500 (top panel) and SST (bottom panel) distributions of the first SVD for (a) MAM, (b) JJA, (c) SON and (d) DJF. The contour interval is 6 m for (a) and (d) and 3 m for (b) and (c) in top panels and 0.2°C in bottom panels. The magnitude corresponds to one standard deviation of each time coefficient. Areas of heterogeneous correlation with statistical significance passing the 0.01 level are shaded. 38

4  Temporal correlation coefficients of Z500 between the observed and the Comp model forecasts. Areas with statistical significance passing the 0.05 level as estimated by a Student t-test are shaded. 39
Five-year running averaged pattern correlation coefficients between observed and Comp model forecast Z500 anomalies over the NPNA region for MAM(black), JJA(red), SON(green) and DJF(blue).

The time average of pattern correlation coefficients between observed and model forecast Z500 anomalies over the Pacific North-American region (20N-80N, 160E-80W) for four seasons.

Temporal correlation coefficients of SAT between observation and HFP2 MME one-month lead seasonal forecasts for (a)MAM, (b)JJA, (c)SON and (d)DJF. The contour interval is 0.2°C.

Difference between model forecast SST and observations averaged over 1982-2001 for (a)MAM, (b)JJA, (c)SON and (d)DJF. The climatology of the model SST was computed using the MME of five one-tier CliPAS models which have forecasts for all four seasons. The contour interval is 0.3°C.

Difference between Comp model forecast precipitation and observations averaged over 1982-2001 for (a)MAM, (b)JJA, (c)SON and (d)DJF. The contour interval is 2mm.

Five-year running averaged pattern correlation coefficient between observed and Comp model forecast precipitation anomalies over the tropical Pacific region (30°S-30°N, 120°E-60°W) for four seasons.

Time average of the pattern correlation coefficient between observed and Comp model forecast precipitation anomalies over the tropical Pacific region (30°S-30°N, 120°E-60°W) for four seasons.
12 Ratio of total variance to variance of the model forecasts due to internal forcing. Regions where the ratio is significant greater than 1 at 5% significant level from an F-test are shaded.

13 The variance of the model forecasts due to atmospheric internal dynamics.

14 The variance of the model forecasts due to external forcing.
Fig. 1. Observed Z500 (top panel) and SST (bottom panel) distributions of the first SVD for (a) MAM, (b) JJA, (c) SON and (d) DJF. The contour interval is 5 m in top panels and 0.2°C in bottom panels for (a), (c) and (d) while is 3 m in top panels and 0.15°C in bottom panels for (b). The magnitude corresponds to one standard deviation of each time coefficient. Areas of heterogeneous correlation with statistical significance passing the 0.01 level are shaded.
Fig. 2. Temporal correlation between the CRU SAT and the expansion coefficient of the atmospheric component of MCA1 for (a) MAM, (b) JJA, (c) SON and (d) DJF. The contour interval is 0.2. The shaded areas represent correlations with a significance level of 0.05 according to a Students t-test.
Fig. 3. HFP MME mean Z500 (top panel) and SST (bottom panel) distributions of the first SVD for (a) MAM, (b) JJA, (c) SON and (d) DJF. The contour interval is 6 m for (a) and (d) and 3 m for (b) and (c) in top panels and 0.2 °C in bottom panels. The magnitude corresponds to one standard deviation of each time coefficient. Areas of heterogeneous correlation with statistical significance passing the 0.01 level are shaded.
Fig. 4. Temporal correlation coefficients of Z500 between the observed and the Comp model forecasts. Areas with statistical significance passing the 0.05 level as estimated by a Student t-test are shaded.
Fig. 5. Five-year running averaged pattern correlation coefficients between observed and Comp model forecast $Z_{500}$ anomalies over the NPNA region for MAM(black), JJA(red), SON(green) and DJF(blue).
Fig. 6. The time average of pattern correlation coefficients between observed and model forecast Z500 anomalies over the Pacific North-American region (20N-80N, 160E-80W) for four seasons.
Fig. 7. Temporal correlation coefficients of SAT between observation and HFP2 MME one-month lead seasonal forecasts for (a)MAM, (b)JJA, (c)SON and (d)DJF. The contour interval is 0.2°C.
Fig. 8. Difference between model forecast SST and observations averaged over 1982-2001 for (a)MAM, (b)JJA, (c)SON and (d)DJF. The climatology of the model SST was computed using the MME of five one-tier CliPAS models which have forecasts for all four seasons. The contour interval is 0.3°C.
Fig. 9. Difference between Comp model forecast precipitation and observations averaged over 1982-2001 for (a)MAM, (b)JJA, (c)SON and (d)DJF. The contour interval is 2mm.
Fig. 10. Five-year running averaged pattern correlation coefficient between observed and Comp model forecast precipitation anomalies over the tropical Pacific region (30°S-30°N, 120°E-60°W) for four seasons.
Fig. 11. Time average of the pattern correlation coefficient between observed and Comp model forecast precipitation anomalies over the tropical Pacific region (30°S-30°N, 120°E-60°W) for four seasons.
Fig. 12. Ratio of total variance to variance of the model forecasts due to internal forcing. Regions where the ratio is significant greater than 1 at 5% significant level from an F-test are shaded.
Fig. 13. The variance of the model forecasts due to atmospheric internal dynamics.
Fig. 14. The variance of the model forecasts due to external forcing.