An Assessment of the Skill of GEOS-5 Seasonal forecasts

Yoo-Geun Ham\textsuperscript{1,2}, Siegfried Schubert\textsuperscript{1}, and Michele M. Rienecker\textsuperscript{1}

\textsuperscript{1}Global Modeling and Assimilation Office, GSFC/NASA, Greenbelt, Maryland, USA,

\textsuperscript{2}Goddard Earth Sciences Technology and Research Studies and Investigations, Morgan state university

September, 2013

Submitted to Climate Dynamics

\underline{Corresponding author address:} Dr. Yoo-Geun Ham, NASA Goddard Space Flight Center, Code 610.1, Greenbelt, Maryland, USA.

E-mail: yoo-geun.ham@nasa.gov

Tel.: +1-301-614-6545

Fax: +1-301-614-6246
Abstract

The seasonal forecast skill of the NASA Global Modeling and Assimilation Office (GMAO) coupled global climate model (CGCM) is evaluated based on an ensemble of 9-month lead forecasts for the period 1993 to 2010. The results from the current version (V2) of the CGCM consisting of the GEOS-5 AGM coupled to the MOM4 ocean model are compared with those from an earlier version (V1) in which the AGCM (the NSIPP model) was coupled to the Poseidon Ocean Model. It was found that the correlation skill of the Sea Surface Temperature (SST) forecasts is generally better in V2, especially over the sub-tropical and tropical central and eastern Pacific, Atlantic, and Indian Ocean. Furthermore, the improvement in skill in V2 mainly comes from better forecasts of the developing phase of ENSO from boreal spring to summer. The skill of ENSO forecasts initiated during the boreal winter season, however, shows no improvement in terms of correlation skill, and is in fact slightly worse in terms of root mean square error (RMSE).

The degradation of skill is found to be due to an excessive ENSO amplitude. For V1, the ENSO amplitude is too strong in forecasts starting in boreal spring and summer, which causes large RMSE in the forecast. For V2, the ENSO
amplitude is slightly stronger than that in observations and V1 for forecasts starting in boreal winter season. An analysis of the terms in the SST tendency equation, shows that this is mainly due to an excessive zonal advective feedback. In addition, V2 forecasts that are initiated during boreal winter season, exhibit a slower phase transition of El Nino, which is consistent with larger amplitude of ENSO after the ENSO peak season. It is found that this is due to weak discharge of equatorial Warm Water Volume (WWV). In both observations and V1, the discharge of equatorial WWV leads the equatorial geostrophic easterly current so as to damp the El Nino starting in January. This process is delayed by about 2 months in V2 due to the slower phase transition of the equatorial zonal current from westerly to easterly.
1. Introduction

During the last few decades, dynamical forecast systems using coupled atmosphere-ocean GCMs (CGCMs) have produced significant improvements in seasonal forecast skill, particularly during times when El Nino Southern Oscillation (ENSO) is active. This progress is reflected in a number of operational or semi-operational dynamical seasonal forecast systems in various institutes around the world (e.g., Bacmeister et al. 2000; Wang et al. 2002; Palmer et al. 2004; Luo et al. 2005; Saha et al. 2006; Molteni et al. 2011; Ham et al. 2012a). In general, most current seasonal forecast systems exhibit successful prediction of ENSO at lead time ranging from two to four seasons, which beats the prediction skill using persistence or other more sophisticated statistical models. The improvements in skill appear to be the result of a combination of factors including more realistic model formulations (Luo et al. 2005; Yuan et al. 2011), improved initialization procedures (Rosati et al. 1997; Zhang et al. 2007; Keppenne et al 2008; Balmaseda and Anderson 2009; Yin et al. 2011), and the use of optimal initial perturbations for ensemble forecasting (Yang et al. 2006, 2008; Kug et al. 2011; Ham et al. 2009, 2012c).

In addition to the success of individual forecast systems, Multi-Model Ensemble (MME) prediction systems can lead to additional improvements in forecast skill, by reducing the uncertainties due to the model errors in
individual forecast systems (Yoo and Kang, 2005; Wang et al. 2009). The positive impact of MME approaches is now well documented as a result of several international projects including the European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) (Palmer et al. 2004), the Asia–Pacific economic cooperation climate center (APCC)/climate prediction and its application to society (CliPAS) project (Wang et al. 2009), and the U.S. National Multi-Model Ensemble (NMME \(^1\), http://www.cpc.ncep.noaa.gov/products/NMME/) project.

The current version of the NASA GMAO seasonal forecast system (CGCMv2: Vernieres et al. 2013; Ham et al. 2012c,d) consists of the GEOS-5 AGCM (Molod et al. 2012) coupled to the MOM4 ocean model (Griffies et al. 2005). This is a major update of an earlier forecast system (CGCMv1) consisting of the NASA Seasonal-to-Interannual Prediction Project (NSIPP) AGCM (Bacmeister et al. 2000) coupled to the Poseidon ocean model (Schopf and Loughe 1995). Further details of the two forecast systems are provided in the next section.

The aim of this study is to evaluate and compare the forecast skill of the two versions of the NASA GMAO seasonal forecast systems. An important aspect

---

\(^1\) The US National Multi-Model Ensemble (NMME) is an experimental multi-model seasonal forecasting system consisting of coupled models from NOAA/NCEP, NOAA/GFDL, IRI, NCAR, NASA, and Canada's CMC.
of the study is that we address the physical mechanisms responsible for any
differences in the forecast skill between the two versions.

Section 2 describes the two GMAO seasonal forecast systems, the hindcast
experiments, and the observations used in this study. In section 3, we compare
the seasonal forecast skill of the two forecast systems. The physical mechanisms
responsible for the differences in forecast skill between the two systems are
discussed in section 4. The summary and discussion are given in section 5.

2. The GMAO seasonal forecast systems, and validating observations

a. The V1 forecast system

The CGCMv1 consists of the NASA Seasonal-to-Interannual Prediction
Project atmospheric general circulation model (NSIPP AGCM) described in
Bacmeister et al. (2000), the Poseidon ocean GCM (Schopf and Loughe 1995),
and the Mosaic land surface model (Koster and Suarez 1992). The convective
parameterization is the Relaxed Arakawa-Schubert (RAS) scheme (Moorthi and
Suarez, 1992). The resolution of the NSIPP AGCM is 2.5° longitude by 2°
latitude with 34 sigma vertical levels. The Poseidon OGCM is a reduced-gravity
isopycnal model, with a resolution of 5/8° longitude by 1/3° latitude, and 27
vertical layers. The atmosphere and ocean is coupled daily without any flux
corrections.

The oceanic initial conditions are from an analysis that uses Optimal Interpolation to assimilate ocean temperature observations, while the atmospheric states are taken from AMIP-style simulations with the NSIPP AGCM – runs forced with observed sea surface temperatures (SSTs). The forecasts start at the 1st day of the month from 1993 to 2010. The six ensemble members used in this study have initial conditions that consist of the assimilated fields and 5 other states constructed by perturbing either the oceanic or atmospheric states.

b. The V2 forecast system

The NASA/GMAO GEOS-5 AOGCM consists of the GEOS-5 AGCM and the Modular Ocean Model version 4 (MOM4) (Griffies et al., 2005). The atmospheric component of the GEOS-5 model used here has 72 vertical levels and 2° latitude by 2.5° longitude grid spacing. The dynamic core is based on a finite volume method (Lin 2004). The convective parameterization is again the Relaxed Arakawa-Schubert (RAS) scheme (Moorthi and Suarez, 1992), though with substantial improvements (Molod et al. 2012). More details about the GEOS-5 atmospheric model are provided in Rienecker et al. (2007) and Molod et al. (2012). The ocean model uses a B-grid finite difference treatment of the
primitive equations of motion, Boussinesq and hydrostatic approximations in spherical coordinates, and covers the global oceans with realistic coastlines and bathymetry. The resolution is 50 vertical levels and a 1°x1° horizontal grid telescoping to 1/3° meridional spacing near the equator. The vertical grid spacing is a constant 10 m over the top 225 m. Air–sea fluxes are exchanged at every time step.

The initial conditions are obtained from an ocean assimilation employing the GEOS-5 AOGCM and a multi-variate Ensemble Optimal Interpolation (EnOI) analysis scheme that ingests various temperature and salinity observations, while the atmosphere is constrained by MERRA (Rienecker et al. 2011). We utilize initial condition from assimilated states and 5 more ensemble members from the perturbed initial conditions using the breeding and random perturbation (i.e. total 6 ensemble members). The forecast period is 1981-2010, but we only utilize the data from 1993 for the consistency with V1 forecast period.

c. Observational data

The observed SST data are from the National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperatures version 3 from NOAA/OAR/ESRL (ERSST V3b, http://www.esrl.noaa.gov/psd/).
The monthly-mean zonal wind stress data are from ERA Interim (Dee et al. 2011). The ocean temperature and current data are obtained from the NCEP Global Ocean Data Assimilation system (GODAS, Behringer and Xue, 2004). For precipitation, we use the Global Precipitation Climatology Project (GPCP) monthly-mean precipitation data from WMO/WCRP/GEWEX (Adler et al. 2003). The analyzed periods for all observations are from 1993 to 2010.

3. Seasonal forecast skill

We first compare the bias (i.e. the time mean ensemble-mean forecast minus the time mean observed) in 3-month averaged Sea Surface Temperature (SST) fields for 2 to 4 month lead times, for forecasts initiated at the beginning of March, June, September, and December (Figure 1). In V1, the general spatial pattern of the SST bias over the Pacific is similar regardless of the month the forecasts were initiated. The cold bias is mainly over the western and equatorial far-eastern Pacific, while there is positive SST bias of about 1.5°C over the southern far-eastern Pacific. The cold SST bias over the western Pacific is relatively strong over the northern hemisphere in the forecast initiated in March, and June, and along the South Pacific convergence zone (SPCZ) in the forecast initiated in September, and December. In V2, the cold bias is mainly over the
central Pacific, and the magnitude of the SST bias is generally smaller than that in V1. For example, in the case of the forecast initiated in March, the equatorial SST bias is almost zero in V2, while for V1 the magnitude of the cold SST bias reaches -1.5°C (-3°C) over the western (eastern) Pacific.

Figure 2 shows the correlation between the observed and ensemble-mean 3-month averaged Sea Surface Temperature (SST) anomaly from 2 to 4 months lead time initiated at the beginning of March, June, September, and December. It is clear that the anomaly correlation for V2 is systematically higher than that or V1 in almost all regions and seasons. For example, the V2 forecasts initiated in March, have anomaly correlations over the equatorial and sub-tropical Pacific, Indian Ocean, and Atlantic Ocean that are systematically higher than those in V1. The same is true for the forecasts starting in June. The anomaly correlation over the eastern Pacific is more than 0.8 in V2, while it is about 0.7 in V1. Such an improvement is also evident in forecasts starting in September, and December over the off-equatorial and subtropical Pacific, Atlantic, and Indian Ocean, while there is little if any improvement over the equatorial eastern Pacific.

The above results indicate a systematic improvement in SST forecast skill in V2 compared with V1. There are, however, some exceptions, most notably the equatorial western Pacific where V2 exhibits negative correlations for forecasts
initiated in June and September. This is due to fact that the simulated ENSO in V2 extends too far to the west (not shown), resulting in simulated anomalies that are of opposite sign compared to the observations over the western Pacific during ENSO.

Figure 3 shows the ENSO forecast skill in more detail, in terms of the Nino3.4 index (averaged SST anomaly over 170-120°W, 5°S-5°N) for each start month and forecast lead time. The correlation has a clear seasonal dependency for both sets of forecasts. In V1, the forecasts starting in boreal winter or early spring exhibit relatively high skill during the early forecast lead time, however, the forecasts show relatively lower skill at long leads for the other seasons. For example, the forecasts starting in October exhibit anomaly correlations over 0.9 up to 5 month lead time, however, they show an abrupt decrease in skill after that. This is likely related to the so-called spring barrier in ENSO prediction, which is characterized by a drop in skill when the forecasts enter the boreal spring period (Chen et al. 2004). On the other hand, forecasts starting in April, exhibit correlations above 0.7 up to 9 month lead time.

The V2 forecast skill has a general behavior that is similar to that of V1, but there is clear improvement. For example, the drop in skill for forecasts starting in the boreal spring season (i.e. from February to April) does not occur in V2. In addition, there is also improvement in skill for forecasts starting in June, and
November (Fig. 2c). Even though the V2 forecasts show a slight decrease in skill after 8 months lead time for January and August, the general improvement in skill as measured by correlation is quite obvious.

The improvement in V2 forecasts is less obvious when the skill is evaluated in terms of Root Mean Square Error (RMSE). Figure 4 shows the V1 and V2 SST RMSE and the Mean Squared Skill Score (MSSS; Murphy, 1988). Here we define the MSSS as

$$\text{MSSS} = \frac{\text{MSE}_{V1} - \text{MSE}_{V2}}{\text{MSE}_{V1}},$$  \hspace{1cm} (1)$$

where $\text{MSE}_{V1}$ and $\text{MSE}_{V2}$ denote the mean-squared-error (MSE) of V1 and V2, respectively. The upper limit of this value is 1 when the seasonal forecast in V2 is perfect, and a positive value indicates that the skill in V2 is better than that of the V1 in terms of MSE.

The RMSE in V1 is relatively small for forecasts initiated in August through October. In V2, the RMSE for forecasts starting in the boreal spring season is systematically smaller than that in V1. However, for forecasts starting in boreal winter the RMSE is larger in V2. As a result, the MSSS is positive (i.e. improvement in V2) for forecasts starting in boreal spring and summer, but is negative for forecasts starting in boreal fall and winter. In the case of the boreal spring and summer seasons, the improvement in V2 as measured by RMSE is
consistent with that based on the anomaly correlation, however, the decrease in
skill for forecasts starting in boreal winter is only shown in the RMSE-based
validation.

To summarize, the ENSO forecast skill in the two GMAO systems can be
characterized as follows. 1) For forecasts starting in boreal spring and summer,
V2 consistently shows better skill than V1 in terms of both the anomaly
correlation and RMSE. 2) For forecasts starting in boreal fall and winter, V2 skill
is similar to that of V1 in terms of correlation, but lower in terms of RMSE. In
next section, we examine the causes of these differences in forecast skill.

One can wonder what causes the above apparent inconsistency between the
correlation and RMSE measures of skill in forecast initiated at boreal winter
season. It might be related to the differences in the simulated ENSO amplitude.
The relationship between the RMSE and the correlation is as follows (Barnston,

$$\text{RMSE}_{t,o} = \left[ S_f^2 + S_o^2 - 2S_f S_o r_{t,o} + b^2 \right]^{1/2} \quad (2)$$

where $S_f$, $S_o$, $r_{t,o}$, $b$ is a standard deviation of forecast, observation,
correlation coefficient between forecast and observation, and bias, respectively.
Even though the correlation skill is same, the different amplitude in forecasted
anomaly can cause the differences in RMSE. For example, if the simulated
ENSO amplitude is larger, it can cause the larger RMSE even though the
correlation skill is in similar degree. This implies that there is a systematic error
in simulated ENSO amplitude in V2 initiated at boreal winter season. We will
also examine this point in a next section.

4. A comparison of ENSO-related feedbacks

In order to understand the cause of the differences in forecast skill between
the two systems, we first examine in Figure 5 the forecast and observed time
series of the Nino3.4 index. Here we selected the forecasts starting in March,
and September, which show the largest forecast-lead-month-averaged positive
and negative MSSS values, respectively. Even though the general variation of
the Nino3.4 index is well predicted for the March forecasts, the growth of the
Nino3.4 index in V1 is often excessive. For example, during 1995, 2000, 2003,
2006, 2007, 2008, 2009, and 2010, the evolution of the Nino3.4 index in V1 is
faster than that in V2 or the observations. On the other hand, for forecasts
starting in September, the V2 Nino3.4 index is slightly stronger than that of V1
or the observations, though this is not as pronounced as for the March forecasts.
2009/10 show an oversho
stronger in V1 than in V2 (i.e. 1994/95, 2000/01 cases).

Figure 6 shows the standard deviation (STD) of MJJA SST anomalies for forecasts starting in March and September. The observed STD of the SST anomalies is about 1°C, with maximum values over the far-eastern Pacific. For the V1 forecasts starting in March, the STD of the SST anomalies is about 2.5°C over the equatorial eastern Pacific, which is more than twice the observed. For the V2 forecasts, the STD of the SST anomalies is more realistic, though the highest values extend too far to the west. In any event, the key result here is that the magnitude of the variability is excessive in V1 compared to both V2 and the observations. During the boreal winter season (when ENSO tends to peak), the STD of SST anomalies is larger than that of the other seasons (the upper panel of Figure 6). The STD of the observed SST anomalies is more than 1°C over the equatorial central-eastern Pacific. Both forecast systems simulate a too strong ENSO magnitude, with V2 having slightly larger values than V1 except over the far-eastern Pacific. For example, the simulated Nino3.4 magnitude in V1 is 1.03, while it is 1.39 in V2.

In short, it appears that to first order, deficiencies in the simulation of the amplitude of ENSO account for the differences in the forecast skill. In order to understand the causes of the excessive ENSO variability in the forecasts, we consider the Bjerknes (BJ) index (Jin et al. 2006)
\[
\frac{dT}{dt} = 2I_{Bj} T + F[h] \quad (3a),
\]

where

\[
2I_{Bj} = - \left( a_1 \frac{(\Delta u)_E}{L_x} + a_2 \frac{(\Delta v)_E}{L_y} \right) - \alpha_s + \mu_a \beta_u \left( \frac{dT}{dx} \right)_E + \mu_a \beta_h \left( \frac{\partial \theta}{\partial y} \right)_E a_h + \mu_a \beta_w \left( \frac{dT}{dz} \right)_E \quad (3b)
\]

and

\[
F = - \left( \frac{dT}{dx} \right)_E \beta_{uh} + \left( \frac{H(\theta)\theta}{H_m} \right) \quad (3c).
\]

\( \langle A \rangle_E, [A], H(A) \) denote a volume averaged quantity over the eastern region (the Nino3.4 region in this study), zonal mean of the equatorial Pacific basin (i.e. 120°E-90°W, 5°S-5°N in this study), and a step function whose value is 1 (0) when \( A \) is bigger (smaller) than zero, respectively. Here \( u, v, w \), and \( T \) denote the zonal, meridional, vertical velocity, and temperature anomalies, respectively. \( L_x \) and \( L_y \) represent the longitudinal and latitudinal length of the Nino3.4 region. \( a_1 \) and \( a_2 \) are estimated using the scale of the anomalous SSTs as in Kim and Jin (2011). The other coefficients in equation 3b, c are computed using a linear regression from the approximated balance equations, namely,

\[
\langle Q \rangle_E = -\alpha_s \langle T \rangle_E , \quad \langle H(\theta)T_{sub} \rangle_E = a_h \langle h \rangle_E , \quad \langle h \rangle_E - \langle h \rangle_w = \beta_h [\tau_x] , \quad \langle H(\theta)w \rangle_E = -\beta_w [\tau_x] , \quad \langle u \rangle_E = \beta_u [\tau_x] + \beta_{uh} \langle h \rangle_w , \quad \text{and} \quad [\tau_x] = \mu_u \langle T \rangle_E . \quad \text{Note that, for decomposing} \ \beta_u \ \text{and} \ \beta_{uh}, \ \text{we used multiple linear regression. Also} \ Q, \ T_{sub}, \ h,
\( \tau_x \) denote the net heat flux from atmosphere to ocean, the subsurface
temperature (50m in this study), the thermocline depth (20°C isotherm depth in
this study), and the zonal wind stress anomalies, respectively. The definition of
the western box (e.g. \( A_W \)) is 120-180°E, 5°S-5°N.

The BJ index consists of five terms (i.e. right hand side of equation 3b); the
damping due to the mean currents (first), net heat flux (second), growth due to
the zonal advective feedback (third), thermocline feedback (fourth), and Ekman
pumping feedback (fifth). This index is a proxy for the strength of various air-
sea coupled feedback processes, and also a parameter to help determine
whether ENSO is unstable or a damped mode. Note that the zonal advective
feedback in the BJ index only takes into account the impact of the wind-driven
zonal current anomalies (\( \beta_u \)). We will discuss the role of zonal advection due to
the anomalous geostrophic current later.

Figure 7 shows the magnitude of each feedback process and the BJ index
during MJJA season for forecasts starting in March. For the observations, the
advection due to the mean currents and net heat flux exhibits negative values,
which mean that those are ENSO damping mechanisms. The magnitude of the
zonal advective feedback is about 0.6 °C/year, and that of the thermocline
feedback is slightly smaller than that. The Ekman pumping feedback is weak
positive. The general features are similar to those in Kim and Jin (2011), which
first applied the BJ index to climate models, though there are some differences. The biggest difference is that the strength of the thermocline feedback is stronger than that of the zonal advective feedback in Kim and Jin (2011), which is opposite to what we found here. This might be due to the fact that the eastern Pacific box is shifted 30° to the west in this study to focus on the Nino3.4 region (Kang et al., 2001). Other possible reasons for the differences in the results are the difference in seasons (i.e. all seasons in Kim and Jin (2011), and MJJA in this study), and the use of different reanalysis products.

In the case of V2, the strength of each of the feedback processes is generally well simulated. The zonal advective feedback has the largest magnitude among all feedback processes, and thermal damping exhibits the minimum value, which is consistent with observations. The value of the BJ index (weak negative) is also quite similar between V2 and observations. However, the magnitude of the ENSO damping due to advection by mean current, the zonal advective feedback, and thermocline feedback is slightly stronger in V2 than for the observations. The differences are, however, less than 0.5°C/year for all feedback processes.

On the other hand, the BJ index in the V1 forecasts is too strong, which is consistent with stronger STD of SST shown in Fig. 6. The BJ index is more than 1°C/year, which is significantly larger than for V2 or observations. This is due to
the excessive zonal advection feedback, thermocline feedback, and Ekman pumping feedback. The weak thermal damping also plays some role. Among them, zonal advective feedback exhibits the biggest difference with respect to the observations, suggesting that may be the main factor in producing the stronger ENSO magnitude in V1.

We next investigate the reasons for the stronger zonal advective feedback in V1. The zonal advection feedback is determined by three factors; air-sea coupling strength $\mu_a$, the sensitivity of ocean current to the surface wind forcing $\beta_u$, and mean zonal temperature gradient $\left(\frac{dT}{dx}\right)_E$. Once a positive El Nino SST anomaly is produced, it induces anomalous equatorial westerlies, which in turn generate westerly currents. The current advects the mean temperature over the western Pacific to the eastern Pacific to induce additional positive El Nino SST anomalies. Checking the above three terms, we found that the sensitivity of the ocean current to the surface wind forcing (i.e. $\beta_u$) is too strong in the forecasts.

Figure 8 shows the linear regression coefficients of surface-layer zonal current (surface to 50m) onto the equatorial Pacific-mean zonal wind stress anomalies during MJJA season. As mentioned earlier, the contribution of the geostrophic current is excluded in this calculation. For the observations, the response of the surface-layer zonal current is between 5 and 10 m/s/N/m² over
the equatorial Pacific, comparable to the results of Jin et al. (2006). In contrast, the responses in both forecasts are substantially stronger than the observed. This is especially true for V1, for which the response of the ocean zonal current to the zonal wind stress is about 40 m/s/N/m², while for V2 it is about 20-30 m/s/N/m². Over the Nino3.4 region, the observational value is 5.8 m/s/N/m², while for V1, and V2 it is 22.5 m/s/N/m², and 15.1 m/s/N/m², respectively. This shows clearly that the excessive zonal advective feedback in the forecasts is due to the unrealistic response of the surface-layer oceanic current to the zonal wind stress forcing.

Figure 9 shows the strength of each of the air-sea coupled processes and the BJ index during NDJF for forecasts starting in September. The BJ index based on the observations shows a negative value (i.e. -0.46). The BJ index for V1 is slightly larger (-0.23). On the other hand, V2 has a positive value of the index, which is consistent with the excessive ENSO variability in the forecasts. The zonal advective feedback appears to be the main contributor to the excessive BJ index in V2. While also stronger than observed in V1, zonal advective feedback is not as excessive as that in V2. This is opposite to the results for the March forecast where V1 shows excessive zonal advective feedback strength.

We next assess the causes of the change in the strength of the zonal advective feedback between two sets of forecast initiated at March and
September. We found that the sensitivity of ocean current to the surface wind forcing $\beta_u$, and mean zonal temperature gradient $\frac{dT}{dx}$ are similar between two sets of forecasts. However, there is a large difference in the air-sea coupling strength ($\mu_a$). In the observations, the air-sea coupling strength (defined as the linear regression of zonal wind stress anomalies averaged over the equatorial Pacific (120-90°W, 5°S-5°N) onto the Nino3.4 SST anomalies during NDJF season) is 0.0047 N/m²/°C. That in V2 is quite similar to the observed (0.0045 N/m²/°C), however, in V1 the air-sea coupling strength is only half that (0.0020 N/m²/°C), indicating that the response of equatorial wind stress anomalies to El Nino SST forcing is significantly weaker in V1. As a result, for the V1 September forecasts, the weak air-sea coupling strength mitigates the excessive zonal advective feedback due to the strong sensitivity of the oceanic current ($\beta_u$). On the other hand, the excessive sensitivity of the oceanic current in the V2 forecasts is directly impacted by the too strong zonal advective feedback.

5. The impact of the recharge/discharge process

In a previous section, we utilized the BJ index to investigate ENSO characteristic in forecasts and observation. This, however, does not take into account the impact of the geostrophic current associated with the
recharge/discharge of equatorial heat content, which is particularly important for determining the ENSO period (Jin et al. 2006). As such, the zonal-mean heat content and related geostrophic zonal current are important factors for determining the growth and demise of ENSO, and are likely to be particularly important during the ENSO peak season.

To examine how the geostrophic current influences the phase transition or demise of the ENSO, Figure 10 shows the composite of the thermocline depth anomaly and surface-layer zonal currents in observations, V1, and V2 during the El Nino events. We selected 5 El Nino cases (i.e. 1994/95, 1997/98, 2002/03, 2006/07, and 2009/10) for the composite analysis. Note that we focus on the forecasts starting in September since we did not find any systematic differences between the forecasts starting earlier in the year (not shown). For the observations, there is maximum over the equatorial eastern Pacific with negative values over the western Pacific during November. In addition, off-equator around 5°N, the negative thermocline depth anomaly is clearly evident. This induces an anomalous negative meridional gradient in the height, which is associated with an anomalous westerly current through the geostrophic balance. There is also a weak easterly current over the equatorial central Pacific, which becomes stronger beginning in December, while the off-equatorial westerly current disappears. The change in sign of the equatorial zonal current anomaly
from westerly to easterly in December acts as a brake on the evolution of the El
Nino event. The equatorial easterly current is directly linked to the spatial
pattern of the thermocline depth anomaly (a minimum over the equator) which
results from the discharge of equatorial heat content during the El Nino due to
divergent meridional currents (Jin, 1997).

In V1, this discharge process is slower than observed, however, the
thermocline depth anomaly does exhibit a minimum at the equator starting in
January. In addition, the positive thermocline depth anomaly is shifted from the
equator in November to off the equator at February, indicating that the
discharge of equatorial heat content is simulated to some extent in V1. On the
other hand, this discharge process is not effective in V2. The positive
thermocline depth anomaly still has a maximum over the equator, and the off-
equatorial negative anomaly continues into February. This implies that the
discharge process in V2 is much slower than what occurs in V1 or the
observations. As a result, the westerly current in V2 is maintained up to January,
and the magnitude of the easterly current is systematically weaker than that in
either V1 or observation. This results in excessive warm temperature advection,
which is not evident in V1 or the observations, and El Nino SST anomalies
continue to grow even after it starts to decay in the observations and V1. This
appears to be one of the mechanisms contributing to the overshoot of ENSO in
V2 as shown in Fig. 5.

The discharge process is mainly driven by the equatorial divergent meridional flow over the thermocline layer (Jin, 1997), so we would expect that the V2 equatorial meridional current would be weak. Figure 11 shows the El Nino composite of the meridional current averaged from surface to 200m over the central Pacific (160°E-80°W). In V1, the magnitude of divergent flow is less than 1 cm/s, while the observational value is over 1.2 cm/s. In addition, there is northward current south of the equator, which is related to the mass recharge at the equator. In V2, the magnitude of the divergent flow is similar to that in V1 poleward of 5°N, but it is systematically weaker between 0-5°N. This is consistent with Fig. 10 in that the discharge of the equatorial heat content is not effective in V2, and this is responsible for the persistent positive equatorial thermocline depth anomaly and westerly current.

In order to understand why the equatorial meridional current is systematically weaker in V2 we examine the following Sverdrup balance equation:

\[ \rho v_s H = -\frac{\Omega}{\beta} \frac{\partial \tau_s}{\partial y}, \quad (4) \]

where \( \rho \), \( v_s \), and \( H \) denote the water density, meridional current for Sverdrup transport, and the mean thermocline depth, respectively. Once there are
westerlies at the equator, the meridional gradient of zonal wind stress is negative north of equator, which leads to equatorial divergent flows. To check whether this relationship holds in the observations and forecasts, we show in Figure 12 the composite of the zonal wind stress over the central Pacific (160°E-80°W). For the observations, the peak of the westerly wind forcing is at the equator, and so the meridional gradient of the zonal wind stress is negative north of the equator. On the other hand, the positive meridional gradient of zonal wind stress south of the equator is quite weak, consistent with the hemispheric asymmetric feature in divergent flow (Kug et al. 2003).

In V1, the peak of westerly wind stress is at 2°S, and so the northward mass transport occurs north of 2°S, which is consistent with Fig. 11b that the northward flow is south of the equator, consistent with the Sverdrup. In V2, the peak in the westerly wind stress occurs north of the equator (i.e. 2°N), so that the sign of the meridional gradient of the zonal wind stress is positive between 0-2°N, which is opposite to that in V1 or observation. The associated weak divergent flow at the equator, suggests that the unrealistic atmospheric response during El Nino might be responsible for the slower ENSO demise in V2.

Figure 13 shows the El Nino composite of precipitation and zonal wind stress anomaly in the observations and forecasts. The observed precipitation
anomaly during El Nino events shows maximum values over the central Pacific. There are two peaks in the precipitation anomaly; one south of the equator between 5-0°S, and the other between 0-5°N. The northern part of the precipitation anomaly is elongated over the region where the climatological precipitation is relatively large (i.e. Intertropical Convergence Zone, ITCZ). In V1, the overall pattern is well simulated, however, the precipitation anomaly is shifted to the south. On the other hand, the anomalous precipitation in V2 is confined to the northern hemisphere, lacking the maximum south of the equator, though that to the north along the ITCZ is as strong as observed.

The spatial pattern of the zonal wind stress anomaly is dynamically linked to the precipitation anomaly (the lower panel of Figure 13). For the observations, the peak in the zonal wind stress is over the central Pacific south of the equator where the precipitation anomaly is a maximum, (Harrison and Vecchi, 1999). As the anomalous precipitation extends to the east along the ITCZ, the north-eastern edge of the zonal wind stress extends slightly to the east. In V1, the zonal wind stress anomaly and precipitation anomaly are both confined to be south of the equator. In V2, the center of zonal wind stress is slightly shifted to the north compared to the observed and V1, even though it still lies south of the equator. Another aspect of the V2 stress is that the off-equatorial westerly stress along the ITCZ is stretched too far to the east. This additional westerly stress
along the ITCZ contributes to the unrealistic peak of zonal wind stress north of
the equator. This is consistent with the precipitation anomaly composite which
is confined to be north of the equator.

The unrealistic extension of the zonal wind stress along the ITCZ, suggests a
possible bias in the precipitation over this region. In particular, a wet
climatology provides favorable conditions for stronger anomalous convection
(Watanabe et al. 2011; Kim et al. 2011; Ham and Kug, 2012), leading to enhanced
surface winds. Figure 14 shows the climatological precipitation fields for the
observations and the model simulations. V2 clearly has a wet bias along the
ITCZ and that is strongest over the central Pacific between 180-120°W – the
region where the ENSO-related convection is elongated and the westerly wind
stress is unrealistic. This suggests that the wet bias over the central Pacific
ITCZ might provide the background for inducing the excessive precipitation
anomalies and the related anomalous westerlies along the ITCZ.

6. Summary and discussions

In this study, the forecast skills of two versions of GMAO seasonal
prediction system are evaluated. The more recent version of the forecast system
with the GEOS-5 model (V2) exhibits overall better skill than the previous
version based on the NSIPP model (V1). The correlation skill in V2 is systematically higher in most regions except for the equatorial western Pacific. This improvement is robust especially for forecasts starting early in the year. On the other hand, when the forecast skill is measured by RMSE, the skill of the Nino3.4 index is slightly worse in V2 for forecasts initiated during the boreal winter season. We found that the forecast skill degrades relatively fast because the simulated ENSO is generally too strong. In particular, for forecasts starting early in the year, V1 tends to simulate too strong ENSO amplitudes (so V1 skill is less than that of V2), while for forecasts starting in boreal winter there is some overshooting of the ENSO anomaly in V2, and the skill in V2 is slightly worse than that in V1.

By utilizing the BJ index from Jin et al. (2006), we examined the reasons for why the ENSO magnitude in V1 (V2) is excessive in the forecast starting early (late) in the year. For the forecasts starting in March, it was shown that the positive ENSO feedbacks (the zonal advective feedback, thermocline feedback, and Ekman feedback) are excessive in V1. This together with weak thermal damping explains the excessive ENSO magnitude. Among the above factors, the magnitude of the zonal advective feedback exhibits the largest differences with the observations. This difference occurs primarily because the response of
the wind-driven current to the wind forcing is too sensitive in V1, since the
other processes are generally well simulated. This basic sensitivity is also
excessive in V2, however, less so compared with V1.

For the forecast starting late in the year, the ENSO amplitude is stronger
than the observed in both versions, but more excessive in V2. This is again due
to a strong oceanic sensitivity to the wind forcing. However, in V1 the air-sea
coupling strength is only half of the observed, which mitigates the excessive
oceanic sensitivity. Therefore, the overall strength of the zonal advective
feedback in V1 is similar degree with the observed. On the other hand, the
excessive sensitivity of the oceanic current in the V2 forecasts is directly
impacted by the too strong zonal advective feedback.

In addition to the wind-driven zonal current, which is taken into account in
the BJ index, it is found that geostrophic currents are also important in
determining the ENSO characteristics in the forecasts during its peak phase. For
the V2 forecasts that are initiated during boreal winter, there is a too slow phase
transition of El Nino due to the weak discharge of equatorial WWV. This delays
the transition of equatorial current from westerly to easterly, and contributes to
the overshooting of the ENSO in V2. This weak discharge of the equatorial heat
content in V2 is related to an erroneous spatial pattern of the anomalous zonal wind stress. That is, the ENSO-related zonal wind stress in V2 shows unrealistic peak along the ITCZ, while that in V1 and the observation show maximum at the south of the equator. This unrealistic peak in ENSO-related zonal wind stress is linked to the confinement of ENSO-related precipitation anomaly at the north of the equator, which is possibly due to the excessive climatological precipitation in the ITCZ.

This study focused on the physical mechanisms which control the ENSO prediction skill in two substantially different forecast systems. The hope is that this can provide some guidance to model development aimed at improving forecast skill on seasonal and longer time scales. On the other hand, it must be acknowledged that a large portion of the forecast skill is also controlled by the quality of the initial conditions (Rosati et al. 1997; Behringer et al. 1998; Tang and Kleeman 2002; Rogel et al. 2005). Since the initial conditions between two GMAO seasonal forecast systems is different by using different observations and perturbing method even though the initialization method is similar, we believe that differences documented here are also affected by the differences in initial conditions.
References


Barnston AG (1992) Correspondence among the correlation, RMSE, and Heidke forecast verification measures; Refinement of the Heidke Score. Weather and Forecasting, 7: 699-709.


Figure 1. The climatological bias in 3-month averaged SST from 2 to 4 month lead initiated at March, June, September, and December 1st in V1 (left panels), and V2 (right panels).
Figure 2. Same as Fig. 1, but for correlation skill of SST anomaly between the observed and the predicted.
Figure 3. The correlation skill of Nino3.4 index with respect to the start month (x-axis) and forecast lead month (y-axis) in (a) V1, and (b) V2. The panel (c) shows the difference of the correlation skill (i.e. V2-V1).
Figure 4. Same as Fig. 3, but for RMSE. The panel (c) shows the MSSS score.
Figure 5. Time series of predicted ensemble-mean Nino3.4 index in V1 (blue), V2 (red) in the forecast starting at (a) March, and (b) September 1st. The observed Nino3.4 index is shown as a grey line.
Figure 6. The left panel shows the standard deviation (STD) of SST anomaly in (a) observation, (b) V1, and (c) V2 forecast for MJJA season. We utilize the forecast starting at March 1\textsuperscript{st}. The right panel is the STD of SST anomaly during NDJF season using the forecast starting at September 1\textsuperscript{st}.
Figure 7. The strength of the damping due to the mean currents (1, denoted as MA), net heat flux (2, denoted as TD), and growth due to the zonal advective feedback (3, denoted as ZA), thermocline feedback (4, denoted as TH), and Ekman pumping feedback (5, denoted EK), and BJ index (6, denoted as BJ) during MJJA in observation (grey bar), V1 (blue line), and V2 (red line). We utilize the forecast starting at March 1st.
Figure 8. The linear regression of oceanic zonal current at the surface-layer (0-50m average) to the equatorial Pacific-mean (i.e. 120°E-90°W, 5°S-5°N) zonal wind stress anomaly in (a) observation, (b) V1, and (c) V2 during MJJA season. We utilize the forecast starting at March 1st.
Figure 9. The strength of the damping due to the mean currents (1, denoted as MA), net heat flux (2, denoted as TD), and growth due to the zonal advective feedback (3, denoted as ZA), thermocline feedback (4, denoted as TH), and Ekman pumping feedback (5, denoted EK), and BJ index (6, denoted as BJ) during NDJF in observation (grey bar), V1 (blue line), and V2 (red line). We utilize the forecast starting at September 1st.
Figure 10. The El Nino composite of the thermocline depth anomaly (shading) and surface-layer zonal currents (contour) in observations (left panel), V1 (mid-panel), and V2 (right panel) from November to subsequent February. We utilize the forecast starting at September 1st.
Figure 11. Time-latitude section of the meridional current averaged from surface to 200m over the central Pacific (160°E-80°W) in (a) observation, (b) V1, and (c) V2 during the El Nino event. We utilize the forecast starting at September 1st.
Figure 12. The El Nino composite of the zonal wind stress over the central Pacific (160°E-80°W) in observation (black), V1 (blue), and V2 (red) during OND season.
Figure 13. The El Nino composite of precipitation (upper panels) and zonal wind stress anomaly (lower panels) in observation and forecasts during OND season.
Figure 14. The climatological precipitation fields in (a) observation, (b) V1, and (c) V2 during OND season.